



(12) 发明专利申请

(10) 申请公布号 CN 115956125 A

(43) 申请公布日 2023.04.11

(21) 申请号 202180038779.4

(74) 专利代理机构 北京市中咨律师事务所
11247

(22) 申请日 2021.03.31

专利代理师 胡志君 黄革生

(30) 优先权数据

63/003,281 2020.03.31 US

63/124,397 2020.12.11 US

(51) Int.Cl.

C12N 15/86 (2006.01)

(85) PCT国际申请进入国家阶段日

2022.11.28

(86) PCT国际申请的申请数据

PCT/IB2021/052710 2021.03.31

(87) PCT国际申请的公布数据

W02021/198963 EN 2021.10.07

(71) 申请人 医学噬菌体生物制药公司

地址 加拿大安大略省

(72) 发明人 R·斯拉夫切夫 N·纳菲西

权利要求书8页 说明书51页
序列表45页 附图13页

(54) 发明名称

用于产生病毒样颗粒的载体及其用途

(57) 摘要

本公开提供了用于产生病毒样颗粒 (VLP) 的表达载体和无细菌序列载体, 如微链 DNA (msDNA), 以及其组合物和方法。在一些方面, 所述方法包括用所述载体、组合物和 VLP 治疗受试者的病毒感染。

1. 表达载体,其包含

表达盒,所述表达盒包含编码重组蛋白的核酸序列,所述重组蛋白包含与免疫原性氨基酸序列融合的来自病毒的保守氨基酸序列;

位于所述表达盒每一侧翼的用于第一重组酶的靶序列;和

整合在用于所述第一重组酶的靶序列的非结合区内的用于一个或多个额外重组酶的一个或多个额外靶序列,

其中从所述表达盒细胞内表达的蛋白质能够形成病毒样颗粒 (VLP)。

2. 根据权利要求1所述的表达载体,其中所述免疫原性氨基酸序列与所述保守氨基酸序列来自同一病毒。

3. 根据权利要求1或2所述的表达载体,其中所述保守氨基酸序列来自病毒糖蛋白。

4. 根据权利要求3所述的表达载体,其中所述免疫原性氨基酸序列来自相同的病毒糖蛋白。

5. 根据权利要求1-4中任一项所述的表达载体,其中所述表达盒还包含编码病毒包膜蛋白的核酸序列和/或编码病毒基质蛋白的核酸序列。

6. 根据权利要求5所述的表达载体,其中所述病毒包膜蛋白和/或病毒基质蛋白来自与保守氨基酸序列相同的病毒。

7. 根据权利要求1-6中任一项所述的表达载体,其中所述保守氨基酸序列、免疫原性氨基酸序列、病毒包膜蛋白和/或病毒基质蛋白是共有序列。

8. 根据权利要求1-7中任一项所述的表达载体,其中所述重组蛋白能够刺激针对病毒的免疫应答,包括产生中和抗体。

9. 根据权利要求1-8中任一项所述的表达载体,其中所述重组蛋白能够刺激Th1细胞介导的针对病毒的免疫应答。

10. 根据权利要求8或9所述的表达载体,其中所述免疫应答对相关病毒或毒株是交叉反应的。

11. 根据权利要求1-10中任一项所述的表达载体,其中所述重组蛋白排除了病毒中刺激包括产生非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。

12. 根据权利要求1-11中任一项所述的表达载体,其中所述表达盒包含单个可读框,所述单个可读框包含每个编码蛋白质的核酸序列之间的编码自切割肽的核酸序列。

13. 根据权利要求1-12中任一项所述的表达载体,其中所述病毒是冠状病毒、流感病毒、人类免疫缺陷病毒、人类乳头瘤病毒、肝炎病毒或溶瘤病毒。

14. 根据权利要求13所述的表达载体,其中所述病毒是冠状病毒。

15. 根据权利要求14所述的表达载体,其中所述冠状病毒是COVID-19。

16. 根据权利要求14或15所述的表达载体,其中所述表达盒包含编码冠状病毒膜 (M) 蛋白、冠状病毒包膜 (E) 蛋白和重组蛋白的核酸序列,所述重组蛋白包含来自冠状病毒刺突 (S) 蛋白的保守氨基酸序列和免疫原性氨基酸序列。

17. 根据权利要求16所述的表达载体,其中所述保守氨基酸序列来自S蛋白S2'切割位点和内部融合肽 (IFP)。

18. 根据权利要求16或17所述的表达载体,其中所述保守氨基酸序列包含SEQ IDNO: 12。

19. 根据权利要求16-18中任一项所述的表达载体,其中所述免疫原性氨基酸序列来自S蛋白受体结合结构域(RBD)。

20. 根据权利要求16-19中任一项所述的表达载体,其中所述免疫原性氨基酸序列与SEQ ID NO:11至少约90%同一。

21. 根据权利要求16-20中任一项所述的表达载体,其中所述重组蛋白进一步包含来自S蛋白的跨膜(TM)结构域序列。

22. 根据权利要求16-21中任一项所述的表达载体,其中所述重组蛋白排除了S蛋白中刺激包含非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。

23. 根据权利要求16-21中任一项所述的表达载体,其中所述重组蛋白的氨基酸序列与SEQ ID NO:55至少约90%同一。

24. 根据权利要求16-23中任一项所述的表达载体,其中所述表达盒包含翻译为与SEQ ID NO:57至少约90%同一的氨基酸序列的单个可读框。

25. 根据权利要求16-24中任一项所述的表达载体,其中所述重组蛋白能够刺激针对COVID-19的免疫应答。

26. 根据权利要求16-25中任一项所述的表达载体,其中所述重组蛋白能够刺激Th1细胞介导的针对COVID-19的免疫应答。

27. 根据权利要求25或26所述的表达载体,其中所述免疫应答对其他冠状病毒是交叉反应的。

28. 根据权利要求27所述的表达载体,其中所述免疫应答对其他严重急性呼吸综合征冠状病毒和/或人类β冠状病毒是交叉反应的。

29. 根据权利要求1-28中任一项所述的表达载体,其中用于第一重组酶的靶序列和用于一个或多个额外重组酶的一个或多个额外靶序列选自PY54 pal位点、N15 telRL位点、loxP位点、ΦK02 telRL位点、FRT位点、phiC31 attP位点和λattP位点。

30. 根据权利要求29所述的表达载体,其中所述表达载体包含每个所述靶序列。

31. 根据权利要求30所述的表达载体,其中所述表达载体包含Tel重组酶pal位点和整合在pal位点内的telRL、loxP和FRT重组酶靶结合序列。

32. 根据权利要求1-31中任一项所述的表达载体,其中所述表达载体用于产生无细菌序列载体。

33. 根据权利要求32所述的表达载体,其中所述无细菌序列载体具有环状共价闭合末端。

34. 根据权利要求32所述的表达载体,其中所述无细菌序列载体具有线性共价闭合末端。

35. 根据权利要求1-34中任一项所述的表达载体,所述表达载体还包含至少一个位于第一重组酶靶序列每一侧的增强子序列。

36. 根据权利要求35所述的表达载体,其中所述至少一个增强子序列是至少两个增强子序列。

37. 根据权利要求35或36所述的表达载体,其中所述至少一个增强子序列是SV40增强子序列。

38. 一种载体生产系统,其包含经设计为在可诱导启动子的控制下编码至少第一重组

酶的重组细胞,其中所述细胞包含权利要求1-37中任一项所述的表达载体。

39. 根据权利要求38所述的载体生产系统,其中所述诱导型启动子是热调节的、化学调节的、IPTG调节的、葡萄糖调节的、阿拉伯糖诱导的、T7聚合酶调节的、冷休克诱导的、pH诱导的或它们的组合。

40. 根据权利要求38或39所述的载体生产系统,其中所述第一重组酶选自telN和tel,并且所述表达载体并入有所述至少第一重组酶的靶序列。

41. 根据权利要求38-40中任一项所述的载体生产系统,其中所述重组细胞已被进一步设计为编码核酸酶基因组编辑系统,并且其中所述表达载体还包含含有用于核酸酶基因组编码系统的切割位点的主链序列。

42. 根据权利要求41所述的载体生产系统,其中所述核酸酶基因组编辑系统是包含Cas核酸酶和gRNA的CRISPR核酸酶系统,并且所述表达载体包含主链序列内gRNA的靶序列。

43. 一种产生无细菌序列载体的方法,包括在表达第一重组酶合适的条件下孵育权利要求38-42中任一项所述的载体生产系统。

44. 一种产生无细菌序列载体的方法,包括在表达所述第一重组酶和所述核酸酶基因组编辑系统合适的条件下孵育权利要求41或42所述的载体生产系统。

45. 根据权利要求43或44所述的方法,其进一步包括收获所述无细菌序列载体。

46. 通过权利要求43-45中任一项所述的方法产生的无细菌序列载体。

47. 一种无细菌序列载体,其包含表达盒,所述表达盒包含编码重组蛋白的核酸序列,所述重组蛋白包含与免疫原性氨基酸序列融合的来自病毒的保守氨基酸序列,其中从所述表达盒胞内表达的蛋白质能够形成VLP。

48. 根据权利要求47所述的无细菌序列载体,其中所述免疫原性氨基酸序列与所述保守氨基酸序列来自同一病毒。

49. 根据权利要求47或48所述的无细菌序列载体,其中所述保守氨基酸序列来自病毒糖蛋白。

50. 根据权利要求49所述的无细菌序列载体,其中所述免疫原性氨基酸序列来自相同的病毒糖蛋白。

51. 根据权利要求47-50中任一项所述的无细菌序列载体,其中所述表达盒还包含编码病毒包膜蛋白的核酸序列和/或编码病毒基质蛋白的核酸序列。

52. 根据权利要求51所述的无细菌序列载体,其中所述病毒包膜蛋白和/或病毒基质蛋白来自与保守氨基酸序列相同的病毒。

53. 根据权利要求47-52中任一项所述的无细菌序列载体,其中所述保守氨基酸序列、免疫原性氨基酸序列、病毒包膜蛋白和/或病毒基质蛋白是共有序列。

54. 根据权利要求47-53中任一项所述的无细菌序列载体,其中所述重组蛋白能够刺激针对病毒的包含中和抗体的免疫应答。

55. 根据权利要求47-54中任一项所述的无细菌序列载体,其中所述重组蛋白能够刺激Th1细胞介导的针对病毒的免疫应答。

56. 根据权利要求54或55的无细菌序列载体,其中所述免疫应答对相关病毒或毒株是交叉反应的。

57. T根据权利要求47-56中任一项所述的无细菌序列载体,其中所述重组蛋白排除了

病毒中刺激包含非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。

58. 根据权利要求47-57中任一项所述的无细菌序列载体,其中所述表达盒包含单个可读框,所述单个可读框包含每个编码蛋白质的核酸序列之间的编码自切割肽的核酸序列。

59. 根据权利要求47-58中任一项所述的无细菌序列载体,其中所述病毒是冠状病毒、流感病毒、人类免疫缺陷病毒、人类乳头瘤病毒、肝炎病毒或溶瘤病毒。

60. 根据权利要求59所述的无细菌序列载体,其中所述病毒是冠状病毒。

61. 根据权利要求60所述的无细菌序列载体,其中所述冠状病毒是COVID-19。

62. 根据权利要求60或61所述的无细菌序列载体,其中所述表达盒包含编码冠状病毒M蛋白、冠状病毒E蛋白和重组蛋白的核酸序列,所述重组蛋白包含来自冠状病毒S蛋白的保守氨基酸序列和免疫原性氨基酸序列。

63. 根据权利要求62所述的无细菌序列载体,其中所述保守氨基酸序列来自S蛋白S2'切割位点和IFP。

64. 根据权利要求62或63所述的无细菌序列载体,其中所述保守氨基酸序列包含SEQ ID NO:12。

65. 根据权利要求62-64中任一项所述的无细菌序列载体,其中所述免疫原性氨基酸序列来自S蛋白RBD。

66. 根据权利要求62-65中任一项所述的无细菌序列载体,其中所述免疫原性氨基酸序列与SEQ ID NO:11至少约90%同一。

67. 根据权利要求62-66中任一项所述的无细菌序列载体,其中所述重组蛋白进一步包含来自S蛋白的TM结构域序列。

68. 根据权利要求62-67中任一项所述的无细菌序列载体,其中所述重组蛋白排除了S蛋白中刺激包含非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。

69. 根据权利要求62-68中任一项所述的无细菌序列载体,其中所述重组蛋白的氨基酸序列与SEQ ID NO:55至少约90%同一。

70. 根据权利要求62-69中任一项所述的无细菌序列载体,其中所述表达盒包含翻译为与SEQ ID NO:57至少约90%同一的氨基酸序列的单个可读框。

71. 根据权利要求62-70中任一项所述的无细菌序列载体,其中所述重组蛋白能够刺激针对COVID-19的免疫应答。

72. 根据权利要求62-71中任一项所述的无细菌序列载体,其中所述重组蛋白能够刺激Th1细胞介导的针对COVID-19的免疫应答。

73. 根据权利要求71或72所述的无细菌序列载体,其中所述免疫应答对其他冠状病毒是交叉反应的。

74. 根据权利要求73所述的无细菌序列载体,其中所述免疫应答对其他严重急性呼吸综合征冠状病毒和/或人类β冠状病毒是交叉反应的。

75. 根据权利要求47-74中任一项所述的无细菌序列载体,其还包含至少一个位于所述表达盒每一侧的至少一个增强子序列。

76. 根据权利要求75所述的无细菌序列载体,其中所述至少一个增强子序列是至少两个增强子序列。

77. 根据权利要求75或76所述的无细菌序列载体,其中所述至少一个增强子序列是

SV40增强子序列。

78. 根据权利要求47-77中任一项所述的无细菌序列载体,其包含环状共价闭合末端。

79. 根据权利要求47-77中任一项所述的无细菌序列载体,其包含线性共价闭合末端。

80. 一种多核苷酸,其编码与SEQ ID NO:57至少约90%同一的氨基酸序列。

81. 重组细胞,其包含权利要求1-37中任一项所述的表达载体或权利要求46-79中任一项所述的无细菌序列载体。

82. 一种生产VLP的方法,包括在从表达载体或无细菌序列载体生产VLP的合适条件下培养权利要求81所述的重组细胞。

83. 根据权利要求82所述的方法,其进一步包括分离所述VLP。

84. 权利要求83的方法,其中所述分离是通过亲和纯化。

85. 根据权利要求82-84中任一项所述的方法,其中所述VLP由权利要求14-37中任一项所述的表达载体或权利要求60-79中任一项所述的无细菌序列载体产生。

86. 根据权利要求85所述的方法,其中所述亲和纯化包括血管紧张素转换酶2 (ACE2) 受体肽或抗S蛋白单克隆抗体。

87. 根据权利要求86所述的方法,其中所述ACE2受体肽包含与SEQ ID NO:70至少约90%同一的氨基酸序列。

88. 根据权利要求86或87所述的方法,其中所述ACE2受体肽在所述肽的C端或N端包生物素受体肽 (BAP) 标签。

89. 根据权利要求88所述的方法,其中所述BAP标签包含与SEQ ID NO:71至少约90%同一的氨基酸序列。

90. 根据权利要求86-89中任一项所述的方法,其中所述ACE2受体肽或抗S蛋白单克隆抗体是生物素化的并固定在链霉亲和素包被的珠上。

91. 根据权利要求84-90中任一项所述的方法,其中所述亲和纯化包括微流体法和/或层析法。

92. 通过权利要求82-91中任一项所述的方法产生的VLP。

93. 一种VLP,其包含重组蛋白,所述重组蛋白包含融合至免疫原性氨基酸序列的来自病毒的保守氨基酸序列。

94. 根据权利要求93所述的VLP,其中所述免疫原性氨基酸序列和保守氨基酸序列来自同一病毒。

95. 根据权利要求93或94所述的VLP,其中所述保守氨基酸序列来自病毒糖蛋白。

96. 根据权利要求93-95中任一项所述的VLP,其中所述免疫原性氨基酸序列来自相同的病毒糖蛋白。

97. 根据权利要求93-96中任一项所述的VLP,其还包含病毒包膜蛋白和/或病毒基质蛋白。

98. 根据权利要求97所述的VLP,其中所述病毒包膜蛋白和/或病毒基质蛋白来自与保守氨基酸序列相同的病毒。

99. 根据权利要求93-98中任一项所述的VLP,其中所述保守氨基酸序列、免疫原性氨基酸序列、病毒包膜蛋白和/或病毒基质蛋白是共有序列。

100. 根据权利要求93-99中任一项所述的VLP,其中所述重组蛋白能够刺激针对病毒的

包含中和抗体的免疫应答。

101. 根据权利要求93-100中任一项所述的VLP, 其中所述重组蛋白能够刺激针对病毒的Th1细胞介导的免疫应答。

102. 根据权利要求100或101所述的VLP, 其中所述免疫应答对相关病毒或毒株是交叉反应的。

103. 根据权利要求93-102中任一项所述的VLP, 其中所述重组蛋白排除了病毒中刺激包含非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。

104. 根据权利要求93-103中任一项所述的VLP, 其中所述病毒是冠状病毒、流感病毒、人类免疫缺陷病毒、人类乳头瘤病毒、肝炎病毒或溶瘤病毒。

105. 根据权利要求104所述的VLP, 其中所述病毒是冠状病毒。

106. 根据权利要求105所述的VLP, 其中所述冠状病毒是COVID-19。

107. 根据权利要求105或106所述的VLP, 其包含冠状病毒膜(M)蛋白、冠状病毒包膜(E)蛋白和重组蛋白, 所述重组蛋白包含来自冠状病毒刺突(S)蛋白的保守氨基酸序列和免疫原性氨基酸序列。

108. 根据权利要求107所述的VLP, 其中所述保守氨基酸序列来自S蛋白S2'切割位点和内部融合肽(IFP)。

109. 根据权利要求107或108所述的VLP, 其中所述保守氨基酸序列包含SEQ ID NO:12。

110. 根据权利要求107-109中任一项所述的VLP, 其中所述免疫原性氨基酸序列来自S蛋白受体结合结构域(RBD)。

111. 根据权利要求107-110中任一项所述的VLP, 其中所述免疫原性氨基酸序列与SEQ ID NO:11至少约90%同一。

112. 根据权利要求107-111中任一项所述的VLP, 其中所述重组蛋白进一步包含来自S蛋白的跨膜(TM)结构域序列。

113. 根据权利要求107-112中任一项所述的VLP, 其中所述重组蛋白排除了S蛋白中刺激包含非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。

114. 根据权利要求107-113中任一项所述的VLP, 其中所述重组蛋白的氨基酸序列与SEQ ID NO:55至少约90%同一。

115. 一种VLP, 其包含与SEQ ID NO:55至少约90%同一的重组蛋白、与SEQ ID NO:1至少约90%同一的M蛋白和与SEQ ID NO:3至少约90%同一的E蛋白。

116. 根据权利要求107-115中任一项所述的VLP, 其中所述重组蛋白能够刺激针对COVID-19的免疫应答。

117. 根据权利要求107-116中任一项所述的VLP, 其中所述重组蛋白能够刺激Th1细胞介导的针对COVID-19的免疫应答。

118. 根据权利要求116或117所述的VLP, 其中所述免疫应答对其他冠状病毒是交叉反应的。

119. 根据权利要求118的VLP, 其中所述免疫应答对其他严重急性呼吸综合征冠状病毒和/或人类β冠状病毒是交叉反应的。

120. 一种组合物, 其包含权利要求1-37中任一项所述的表达载体、权利要求46-79中任一项所述的无细菌序列载体或权利要求92-119中任一项所述的病毒样颗粒。

121. 根据权利要求120所述的组合物,其进一步包含递送剂。
122. 根据权利要求121所述的组合物,其中所述递送剂是纳米粒子。
123. 根据权利要求121或122所述的组合物,其中所述递送剂包含靶向性配体。
124. 根据权利要求123所述的组合物,其中所述靶向性配体包含S蛋白肽。
125. 根据权利要求124所述的组合物,其中所述S蛋白肽包含与任一SEQ ID NO:76-99至少约90%同一的氨基酸序列。
126. 一种治疗受试者病毒感染的方法,所述方法包括向受试者施用权利要求1-37中任一项所述的表达载体、权利要求46-79中任一项所述的无细菌序列载体、权利要求92-119中任一项所述的VLP或权利要求120-125中任一项所述的组合物,其中所述表达载体或无细菌序列载体的细胞内表达产生VLP。
127. 根据权利要求126所述的方法,其中所述施用是通过肠胃外或非肠胃外施用。
128. 根据权利要求127所述的方法,其中所述施用是通过口服、肺、鼻内、静脉内、表皮、经皮、皮下、肌肉或腹膜内施用或吸入施用。
129. 根据权利要求126-128中任一项所述的方法,其中所述VLP刺激受试者的免疫应答,包括针对病毒感染的中和抗体。
130. 根据权利要求126-129中任一项所述的方法,其中所述VLP刺激受试者对病毒感染的Th1细胞介导的免疫应答。
131. 根据权利要求129或130所述的方法,其中所述免疫应答对相关病毒或毒株是交叉反应的。
132. 根据权利要求126-131中任一项所述的方法,其中所述VLP不刺激受试者中包含非中和抗体的免疫应答和/或不刺激受试者的Th2细胞介导的免疫应答。
133. 根据权利要求126-132中任一项所述的方法,其中所述VLP与感染病毒交叉竞争对病毒受体的结合。
134. 根据权利要求133所述的方法,其中所述VLP与相关病毒或毒株交叉竞争对病毒受体的结合。
135. 根据权利要求126-134中任一项所述的方法,其中所述病毒感染是冠状病毒、流感病毒、人类免疫缺陷病毒、人类乳头瘤病毒、肝炎病毒或溶瘤病毒感染。
136. 根据权利要求135所述的方法,其中所述病毒感染是冠状病毒感染。
137. 根据权利要求136所述的方法,其中所述病毒感染是COVID-19。
138. 根据权利要求137所述的方法,其中所述VLP刺激受试者的免疫应答,包括产生针对COVID-19的中和抗体。
139. 根据权利要求137或138所述的方法,其中所述VLP刺激受试者针对COVID-19的Th1细胞介导的免疫应答。
140. 根据权利要求138或139所述的方法,其中所述免疫应答对其他冠状病毒是交叉反应的。
141. 根据权利要求140所述的方法,其中所述免疫应答对其他严重急性呼吸综合征冠状病毒和/或人类 β 冠状病毒是交叉反应的。
142. 根据权利要求137-141中任一项所述的方法,其中所述VLP不刺激受试者中包含非中和抗体的免疫应答和/或不刺激受试者的Th2细胞介导的免疫应答。

143. 根据权利要求137-142中任一项所述的方法,其中所述施用是通过吸入施用。

144. 根据权利要求137-143中任一项所述的方法,其中所述VLP与COVID-19交叉竞争结合ACE2受体、神经毡蛋白-1和/或其他受体。

145. 根据权利要求144所述的方法,其中所述VLP与其他冠状病毒交叉竞争结合ACE2受体、神经毡蛋白-1和/或其他受体。

146. 根据权利要求145所述的方法,其中所述VLP与其他严重急性呼吸综合征冠状病毒和/或人类 β 冠状病毒交叉竞争结合ACE2受体、神经毡蛋白-1和/或其他受体。

用于产生病毒样颗粒的载体及其用途

[0001] 与相关申请的交叉引用

[0002] 本PCT申请要求2020年3月31日提交的美国临时申请号63/003,281和2020年12月11日提交的美国临时申请号63/124,397的优先权利益,这些申请通过引用全部并入本文作为参考。

技术领域

[0003] 本公开提供了用于产生病毒样颗粒 (VLP) 的载体和用其治疗受试者的方法。

背景技术

[0004] 尽管疫苗技术取得了许多进步,但病毒感染仍然是一种普遍存在的健康问题,通常受到有限的控制。例如,无论是在全球传播还是在经济影响方面,COVID-19冠状病毒大流行都不同于一个多世纪以来全世界所看到的任何疫情。它导致许多发达国家反复停业,死亡人数和新感染不断增加。

[0005] COVID-19引起呼吸道感染,严重时还会引起急性呼吸窘迫综合征。与其他冠状病毒如SARS-CoV相比,COVID-19在疾病进程早期的症状发生前的/无症状的(Pre/asymptomatic)空气传播和高病毒滴度显著增加了其传染性,这使得疫苗的开发对于大流行的管理至关重要。

[0006] VLP代表模拟病毒物理化学性质和结构而不引起病毒生长的有效疫苗候选物(Cimica,V.,&Galarza,J.M.,Clin.Immunol.183:99-108(2017))。因此,它们赋予强烈的体液应答,但通常对“全病毒”产生有限的细胞介导的应答,因为它们仍然是外源性抗原。此外,它们的生产、纯化和储存都很昂贵。

[0007] 现有的疫苗在不同病毒株之间往往表现出有限的交叉保护,但由于病毒在进化压力下继续变异其基因组,这一事实使得更加复杂。

[0008] 需要改进的VLP和治疗病毒感染的方法。

[0009] 发明概述

[0010] 本公开涉及一种表达载体,其包含表达盒,所述表达盒包含编码重组蛋白的核酸序列,所述重组蛋白包含与免疫原性氨基酸序列融合的来自病毒的保守氨基酸序列;包含位于所述表达盒每一侧翼的用于第一重组酶的靶序列;和整合在用于所述第一重组酶的靶序列的非结合区内的用于一个或多个额外重组酶的一个或多个额外靶序列,其中从所述表达盒细胞内表达的蛋白质能够形成病毒样颗粒(VLP)。

[0011] 在一些方面,免疫原性氨基酸序列与保守氨基酸序列来自同一病毒。在一些方面,保守氨基酸序列来自病毒糖蛋白。在一些方面,免疫原性氨基酸序列来自相同的病毒糖蛋白。

[0012] 在一些方面,表达盒还包含编码病毒包膜蛋白的核酸序列和/或编码病毒基质蛋白的核酸序列。在一些方面,病毒包膜蛋白和/或病毒基质蛋白来自与保守氨基酸序列相同的病毒。

- [0013] 在一些方面,保守氨基酸序列、免疫原性氨基酸序列、病毒包膜蛋白和/或病毒基质蛋白是共有序列。
- [0014] 在一些方面,重组蛋白能够刺激针对病毒的免疫应答,包括产生中和抗体。
- [0015] 在一些方面,重组蛋白能够刺激Th1细胞介导的针对病毒的免疫应答。
- [0016] 在一些方面,免疫应答对相关病毒或毒株具有交叉反应。
- [0017] 在一些方面,重组蛋白排除了病毒中刺激包括产生非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。
- [0018] 在一些方面,表达盒包含单个可读框,其包含每个编码蛋白质的核酸序列之间的编码自切割肽的核酸序列。
- [0019] 在一些方面,病毒是冠状病毒、流感病毒、人类免疫缺陷病毒、人类乳头瘤病毒、肝炎病毒或溶瘤病毒。
- [0020] 在一些方面,病毒是冠状病毒。在一些方面,冠状病毒是COVID-19。
- [0021] 在一些方面,表达盒包含编码冠状病毒膜(M)蛋白、冠状病毒包膜(E)蛋白和重组蛋白的核酸序列,所述重组蛋白包含来自冠状病毒刺突(S)蛋白的保守氨基酸序列和免疫原性氨基酸序列。在一些方面,保守氨基酸序列来自S蛋白S2'切割位点和内部融合肽(IFP)。
- [0022] 在一些方面,保守氨基酸序列包含SEQ ID NO:12。
- [0023] 在一些方面,免疫原性氨基酸序列来自S蛋白受体结合结构域(RBD)。
- [0024] 在一些方面,所述免疫原性氨基酸序列与SEQ ID NO:11至少约90%同一。
- [0025] 在一些方面,重组蛋白进一步包含来自S蛋白的跨膜(TM)结构域序列。
- [0026] 在一些方面,重组蛋白排除了S蛋白中刺激包含非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。
- [0027] 在一些方面,重组蛋白的氨基酸序列与SEQ ID NO:55至少约90%同一。
- [0028] 在一些方面,表达盒包含翻译为与SEQ ID NO:57至少约90%同一的氨基酸序列的单个可读框。
- [0029] 在一些方面,重组蛋白能够刺激针对COVID-19的免疫应答。
- [0030] 在一些方面,免疫应答对其他冠状病毒具有交叉反应。在一些方面,免疫应答对其他严重急性呼吸综合征冠状病毒和/或人类 β 冠状病毒具有交叉反应。
- [0031] 在一些方面,第一重组酶的靶序列和一个或多个额外重组酶的一个或多个额外靶序列选自PY54 pal位点、N15 telRL位点、loxP位点、 ϕ K02 telRL位点、FRT位点、phiC31attP位点和 λ attP位点。在一些方面,表达载体包含每个所述靶序列。在一些方面,表达载体包含Tel重组酶pal位点和整合在pal位点内的telRL、loxP和FRT重组酶靶结合序列。
- [0032] 在一些方面,表达载体用于生产无细菌序列的载体。在一些方面,无细菌序列载体具有环状共价闭合末端。在一些方面,无细菌序列载体具有线性共价闭合末端。
- [0033] 在一些方面,表达载体还包含至少一个位于第一重组酶靶序列每一侧的增强子序列。在一些方面,所述至少一个增强子序列是至少两个增强子序列。在一些方面,所述至少一个增强子序列是SV40增强子序列。
- [0034] 本公开涉及一种载体生产系统,其包含重组细胞,所述重组细胞被设计为在可诱导启动子的控制下编码至少第一重组酶,其中所述细胞包含任一上述表达载体。在一些方

面,诱导型启动子是热调节的、化学调节的、IPTG调节的、葡萄糖调节的、阿拉伯糖诱导的、T7聚合酶调节的、冷休克诱导的、pH诱导的或其组合。在一些方面,所述第一重组酶选自telN和tel,并且所述表达载体并入有所述至少第一重组酶的靶序列。在一些方面,重组细胞已被进一步设计为编码核酸酶基因组编辑系统,并且其中表达载体还包含含有核酸酶基因组编码系统的切割位点的主链序列。在一些方面,核酸酶基因组编辑系统是包含Cas核酸酶和gRNA的CRISPR核酸酶系统,并且表达载体包含主链序列内gRNA的靶序列。

[0035] 本公开涉及一种产生无细菌序列载体的方法,包括在合适的条件下孵育任一上述载体生产系统,以表达第一重组酶。

[0036] 本公开涉及一种产生无细菌序列载体的方法,包括在合适的条件下孵育任一上述载体生产系统,所述载体生产系统包含设计为编码核酸酶基因组编辑系统的重组细胞,以表达第一重组酶和核酸酶基因组编辑系统。

[0037] 在一些方面,任一上述产生无细菌序列载体的方法还包括收获无细菌序列载体。

[0038] 本公开涉及通过任一上述产生无细菌序列载体的方法产生的无细菌序列载体。

[0039] 本公开涉及一种无细菌序列载体,其包含表达盒,所述表达盒包含编码重组蛋白的核酸序列,所述重组蛋白包含与免疫原性氨基酸序列融合的来自病毒的保守氨基酸序列,其中从所述表达盒胞内表达的蛋白质能够形成VLP。

[0040] 在一些方面,免疫原性氨基酸序列与保守氨基酸序列来自同一病毒。在一些方面,保守氨基酸序列来自病毒糖蛋白。在一些方面,免疫原性氨基酸序列来自相同的病毒糖蛋白。

[0041] 在一些方面,表达盒还包含编码病毒包膜蛋白的核酸序列和/或编码病毒基质蛋白的核酸序列。在一些方面,病毒包膜蛋白和/或病毒基质蛋白来自与保守氨基酸序列相同的病毒。

[0042] 在一些方面,保守氨基酸序列、免疫原性氨基酸序列、病毒包膜蛋白和/或病毒基质蛋白是共有序列。

[0043] 在一些方面,重组蛋白能够刺激针对病毒的包含中和抗体的免疫应答。

[0044] 在一些方面,重组蛋白能够刺激Th1细胞介导的针对病毒的免疫应答。

[0045] 在一些方面,免疫应答对相关病毒或毒株具有交叉反应。

[0046] 在一些方面,重组蛋白排除了病毒中刺激包含非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。

[0047] 在一些方面,表达盒包含单个可读框,其包含编码蛋白质的每个核酸序列之间的自切割肽的编码核酸序列。

[0048] 在一些方面,病毒是冠状病毒、流感病毒、人类免疫缺陷病毒、人类乳头瘤病毒、肝炎病毒或溶瘤病毒。

[0049] 在一些方面,病毒是冠状病毒。在一些方面,冠状病毒是COVID-19。

[0050] 在一些方面,表达盒包含编码冠状病毒M蛋白、冠状病毒E蛋白和重组蛋白的核酸序列,所述重组蛋白包含来自冠状病毒S蛋白的保守氨基酸序列和免疫原性氨基酸序列。

[0051] 在一些方面,保守氨基酸序列来自S蛋白S2'切割位点和IFP。

[0052] 在一些方面,所述保守氨基酸序列包含SEQ ID NO:12。

[0053] 在一些方面,免疫原性氨基酸序列来自S蛋白RBD。

- [0054] 在一些方面,所述免疫原性氨基酸序列与SEQ ID NO:11至少约90%同一。
- [0055] 在一些方面,重组蛋白进一步包含来自S蛋白的TM结构域序列。
- [0056] 在一些方面,重组蛋白排除了S蛋白中刺激包含非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。
- [0057] 在一些方面,重组蛋白的氨基酸序列是SEQ ID NO:55。
- [0058] 在一些方面,表达盒包含翻译为与SEQ ID NO:57至少约90%同一的氨基酸序列的单个可读框。
- [0059] 在一些方面,重组蛋白能够刺激针对COVID-19的免疫应答。
- [0060] 在一些方面,重组蛋白能够刺激Th1细胞介导的针对COVID-19的免疫应答。
- [0061] 在一些方面,免疫应答对其他冠状病毒具有交叉反应。在一些方面,免疫应答对其他严重急性呼吸综合征冠状病毒和/或人类 β 冠状病毒具有交叉反应。
- [0062] 在一些方面,无细菌序列载体还包含位于表达盒每侧的至少一个增强子序列。在一些方面,所述至少一个增强子序列是至少两个增强子序列。在一些方面,所述至少一个增强子序列是SV40增强子。
- [0063] 在一些方面,无细菌序列载体包含环状共价闭合末端。
- [0064] 在一些方面,无细菌序列载体包含线性共价闭合末端。
- [0065] 本公开涉及编码与SEQ ID NO:57至少约90%同一的氨基酸序列的多核苷酸。
- [0066] 本公开涉及包含任一上述表达载体或任一上述无细菌序列载体的重组细胞。
- [0067] 在一些方面,本公开涉及一种生产VLP的方法,包括在合适的条件下培养所述重组细胞,以从表达载体或无细菌序列载体生产VLP。
- [0068] 在一些方面,生产VLP的方法还包括分离该VLP。在一些方面,分离是通过亲和纯化。在一些方面,VLP由任何上述表达载体或任何上述无细菌序列载体产生,其中病毒是冠状病毒。在一些方面,亲和纯化包含血管紧张素转换酶2 (ACE2) 受体肽或抗S蛋白单克隆抗体。在一些方面,ACE2受体肽包含与SEQ ID NO:70至少约90%同一的氨基酸序列。在一些方面,ACE2受体肽在肽的C端或N端包含生物素受体肽 (BAP) 标签。在一些方面,所述BAP标签包含与SEQ ID NO:71至少约90%同一的氨基酸序列。在一些方面,ACE2受体肽或抗S蛋白单克隆抗体是生物素化的并固定在链霉亲和素包被的珠上。在一些方面,亲和纯化包括微流体法和/或层析法。在一些方面,本公开涉及通过任何产生VLP的方法产生的VLP。
- [0069] 本公开涉及一种VLP,其包含融合至免疫原性氨基酸序列的来自病毒的保守氨基酸序列的重组蛋白。在一些方面,免疫原性氨基酸序列和保守氨基酸序列来自同一病毒。在一些方面,保守氨基酸序列来自病毒糖蛋白。在一些方面,免疫原性氨基酸序列来自相同的病毒糖蛋白。
- [0070] 在一些方面,VLP还包含病毒包膜蛋白和/或病毒基质蛋白。在一些方面,病毒包膜蛋白和/或病毒基质蛋白来自与保守氨基酸序列相同的病毒。
- [0071] 在一些方面,保守氨基酸序列、免疫原性氨基酸序列、病毒包膜蛋白和/或病毒基质蛋白是共有序列。
- [0072] 在一些方面,重组蛋白能够刺激针对病毒的包含中和抗体的免疫应答。
- [0073] 在一些方面,重组蛋白能够刺激针对病毒的Th1细胞介导的免疫应答。
- [0074] 在一些方面,免疫应答对相关病毒或毒株具有交叉反应。

[0075] 在一些方面,重组蛋白排除了病毒中刺激包含非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。

[0076] 在一些方面,病毒是冠状病毒、流感病毒、人类免疫缺陷病毒、人类乳头瘤病毒、肝炎病毒或溶瘤病毒。在一些方面,病毒是冠状病毒。

[0077] 在一些方面,冠状病毒是COVID-19。

[0078] 在一些方面,VLP包含冠状病毒膜(M)蛋白、冠状病毒包膜(E)蛋白和重组蛋白,所述重组蛋白包含来自冠状病毒刺突(S)蛋白的保守氨基酸序列和免疫原性氨基酸序列。

[0079] 在一些方面,保守氨基酸序列来自S蛋白S2'切割位点和内部融合肽(IFP)。

[0080] 在一些方面,所述保守氨基酸序列包含SEQ ID NO:12。

[0081] 在一些方面,免疫原性氨基酸序列来自S蛋白受体结合结构域(RBD)。

[0082] 在一些方面,所述免疫原性氨基酸序列与SEQ ID NO:11至少约90%同一。

[0083] 在一些方面,重组蛋白进一步包含来自S蛋白的跨膜(TM)结构域序列。

[0084] 在一些方面,重组蛋白排除了S蛋白中刺激包含非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。

[0085] 在一些方面,重组蛋白的氨基酸序列与SEQ ID NO:55至少约90%同一。

[0086] 本公开涉及一种VLP,其包含与SEQ ID NO:55至少约90%同一的重组蛋白、与SEQ ID NO:1至少约90%同一的M蛋白和与SEQ ID NO:3至少约90%同一的E蛋白。

[0087] 在一些方面,重组蛋白能够刺激针对COVID-19的免疫应答。

[0088] 在一些方面,重组蛋白能够刺激Th1细胞介导的针对COVID-19的免疫应答。

[0089] 在一些方面,免疫应答对其他冠状病毒具有交叉反应。

[0090] 在一些方面,免疫应答对其他严重急性呼吸综合征冠状病毒和/或人类β冠状病毒具有交叉反应。

[0091] 本公开涉及包含任何上述表达载体、任何上述无细菌序列载体或任何上述病毒样颗粒的组合物。在一些方面,该组合物还包含递送剂。在一些方面,递送剂是纳米粒子。在一些方面,递送剂包含靶向性配体。在一些方面,靶向性配体包含S蛋白肽。在一些方面,S蛋白肽包含与任一SEQ ID NO:76-99至少约90%同一的氨基酸序列。

[0092] 本公开涉及治疗受试者病毒感染的方法,包括向受试者施用任何上述表达载体、任何上述无细菌序列载体、任何上述VLP或任何上述组合物,其中表达载体或无细菌序列载体的胞内表达产生VLP。

[0093] 在一些方面,施用是通过肠胃外或非肠胃外施用。在一些方面,施用是通过口服、肺、鼻内、静脉内、表皮、经皮、皮下、肌肉或腹膜内施用或吸入施用。

[0094] 在一些方面,VLP刺激受试者的免疫应答,包括针对病毒感染的中和抗体。

[0095] 在一些方面,VLP刺激受试者对病毒感染的Th1细胞介导的免疫应答。

[0096] 在一些方面,免疫应答对相关病毒或毒株具有交叉反应。

[0097] 在一些方面,VLP不刺激受试者中包含非中和抗体的免疫应答和/或不刺激受试者的Th2细胞介导的免疫应答。

[0098] 在一些方面,VLP与感染病毒交叉竞争对病毒受体的结合。

[0099] 在一些方面,VLP与相关病毒或毒株交叉竞争对病毒受体的结合。

[0100] 在一些方面,病毒感染是冠状病毒、流感病毒、人类免疫缺陷病毒、人类乳头瘤病

毒、肝炎病毒或溶瘤病毒感染。

[0101] 在一些方面,病毒感染是冠状病毒。在一些方面,病毒感染是COVID-19。

[0102] 在一些方面,VLP刺激受试者的免疫应答,包括针对COVID-19的中和抗体。

[0103] 在一些方面,VLP刺激受试者针对COVID-19的Th1细胞介导的免疫应答。

[0104] 在一些方面,免疫应答对其他冠状病毒具有交叉反应。

[0105] 在一些方面,免疫应答对其他严重急性呼吸综合征冠状病毒和/或人类β冠状病毒具有交叉反应。

[0106] 在一些方面,VLP不刺激受试者中包括非中和抗体的免疫应答和/或不刺激受试者的Th2细胞介导的免疫应答。

[0107] 在一些方面,施用是通过吸入施用。

[0108] 在一些方面,VLP与COVID-19交叉竞争结合ACE2受体、神经毡蛋白-1(neuropilin-1)和/或其他受体。

[0109] 在一些方面,VLP与其他冠状病毒交叉竞争结合ACE2受体、神经毡蛋白-1和/或其他受体。

[0110] 在一些方面,VLP与其他严重急性呼吸综合征冠状病毒和/或人类β冠状病毒交叉竞争结合ACE2受体、神经毡蛋白-1和/或其他受体。

附图说明

[0111] 图1显示了用于产生冠状病毒VLP的示例性表达盒的示意图,其含有猴病毒40增强子(SV40E);巨细胞病毒启动子(P_{CMV});编码冠状病毒包膜(E)蛋白的序列;编码冠状病毒膜(M)蛋白的序列;编码重组蛋白的序列,所述重组蛋白包含来自冠状病毒刺突(S)蛋白的受体结合结构域(RBD)、第二亚基切割结构域和内部融合肽(S2' IFP)以及跨膜(TM)结构域的序列(本文中称为重组刺突(S)蛋白,RBD::S2' IFP::TM);编码来自猪捷申病毒(Teschovirus)-1的2A自切割肽(P2A)的序列,以分隔表达盒的多个蛋白质编码序列;和聚腺苷酸化(pA)信号。

[0112] 图2显示了包含如图1所述表达盒的示例性表达载体(pGL2-SS-CMV-VLP-BGH-SS)的载体图,其中pA信号来自牛生长激素。

[0113] 图3显示了来自图2的表达载体的基因和蛋白质的体外表达。(A)显示了一个柱状图,描述了如图1所示的E蛋白、M蛋白和重组S蛋白(RBD::S2' IFP::TM)编码基因自含有图2表达载体的细胞(VLP)以及不含该表达载体的对照细胞(CTL)的相对表达,***=p<0.001和****=p<0.0001。(B)显示了一个代表性的蛋白质印迹,描述了使用结合RBD的抗体(抗刺突蛋白(RBD))检测重组S蛋白的表达。使用β-肌动蛋白抗体检测β-肌动蛋白作为上样对照。对照=来自不含表达载体的细胞的蛋白质。VLP=来自含有图2的表达载体的细胞的蛋白质。(C)显示了如(B)所述的来自蛋白质印迹的重组S蛋白表达的相对平均强度(n=3)。

[0114] 图4显示了如本文所述的由图2的表达载体编码的示例性msDNA-VLP(msDN.A VLP Cov 19-BGH poly)。

[0115] 图5显示了C57小鼠在第0天和第14天(加强)肌肉注射图2的表达载体后,在第0天、第7天、第14天、第21天、第28天、第35天、第42天和第49天C57小鼠血清中结合COVID-19刺突蛋白S1亚基的抗体(刺突AB)的浓度(ng/mL)。(A)和(B)分别显示抗体浓度的线图和柱状图。

[0116] 图6显示了代表性COVID-19基因组的序列保守性分析。(A)显示了柱状图,其中水平柱表示根据参考基因组(NC_045512.2),在y轴上列出的每个COVID-19基因在x轴上的基因组位置。(B)显示了柱状图,其中柱高度对应于3928个代表性COVID-19基因组在每个基因组位置不同于参考基因组的百分数。

[0117] 图7和图8显示了直方图,其中柱高度对应于在每个基因组位置与参考基因组不同的分析基因组的百分数,分析基因组为:(7) 3928个代表性COVID-19基因组,120个严重急性呼吸综合征冠状病毒(SARS-CoV)基因组,和257个中东呼吸综合征冠状病毒(MERS-CoV)基因组,(8A)变异株B.1.1.7的233个COVID-19基因组,(8B)变异株B.1.351的104个COVID-19基因组,(8C)变异株P.1的39个COVID-19基因组,和(8D)变异株B.1.427/429的62个COVID-19基因组。

[0118] 图9显示了如本文所述的用于在真核细胞中产生VLP的示例性真核表达载体(pFastBacTM Dual-VLP),其包含如图1所述的E、M和重组S蛋白。

具体实施方式

[0119] 本公开提供了用于生产病毒样颗粒(VLP)的表达载体和无细菌序列载体(例如,微链DNA(ministring DNA(msDNA))、载体生产系统和VLP及其组合物和方法。本发明的一些方面涉及治疗受试者的病毒感染(例如,人类受试者中的冠状病毒感染,例如COVID-19)。

[0120] 本文引用的所有出版物,包括但不限于本文引用的全部期刊文章、书籍、手册、专利申请和专利,均通过引用的方式并入本文作为参考,其程度与每一份单独出版物均通过引用明确表示的相同。

[0121] I. 术语

[0122] 为了更容易理解本公开,首先定义了某些术语。除非本申请另有明确规定,否则本申请中使用的下列术语应具有以下含义。附加定义在整个申请中阐述。

[0123] 需要注意的是,术语“一个”或“一种”实体是指该实体中的一个或多个;例如,“核苷酸序列”被理解为代表一个或多个核苷酸序列。因此,术语“一个”或“一种”、“一个或多个”和“至少一个”在本文中可互换使用。

[0124] 本文中使用的术语“和/或”应视为对两个指定特征或组件中的每一个的具体披露,包括或不包括另一个。因此,本文中“A和/或B”等短语中使用的术语“和/或”意在包括“A和B”、“A或B”、(仅)“A”和(仅)“B”。同样,“A、B和/或C”等短语中使用的术语“和/或”意在涵盖以下各个方面:A、B和C;A、B或C;A或C;A或B;B或C;A和C;A和B;B和C;(仅)A;(仅)B;和(仅)C。

[0125] 应当理解,在本文中使用“包含”的语言描述方面时,也提供了用“由...组成”和/或“基本上由...组成”描述的其他类似方面。

[0126] 术语“约”或“基本上包含”是指在本领域普通技术人员确定的特定值或组成的可接受误差范围内的值或组成,这将部分取决于如何测量或确定值或组成(即测量系统的限制)。例如,根据本领域的实践,“约”或“基本上包含”可以表示在1个或多于1个标准偏差内。或者,“约”或“基本上包含”可以指高达10%的范围。此外,特别是关于生物系统或过程,这些术语可以表示高达一个数量级或高达5倍的数值。当在本申请和权利要求中提供特定值或组成时,除非另有说明,否则“约”或“基本上包含”的含义应认为是在该特定值或组成的

可接受误差范围内。

[0127] 如本文所述,除非另有说明,否则任何浓度范围、百分数范围、比率范围或整数范围应理解为包括所述范围内的任何整数的值,并且根据需要包括其分数(例如整数的十分之一和百分之一)。数字范围包括定义范围的数字。

[0128] 除非另有定义,否则本文使用的所有技术和科学术语具有与本公开相关领域的普通技术人员通常理解的相同含义。例如,生物医学和分子生物学简明词典 (Concise Dictionary of Biomedicine and Molecular Biology), Juo, Pei-Show, 2nd ed., 2002, CRC Press; 细胞和分子生物学词典 (The Dictionary of Cell and Molecular Biology), 5th ed., 2013, Academic Press; 和牛津生物化学和分子生物学词典 (Oxford Dictionary Of Biochemistry And Molecular Biology), 2006, Oxford University Press 为技术人员提供了本公开中使用的许多术语的通用词典。

[0129] 单位、前缀和符号以国际单位制 (SI) 认可的形式表示。

[0130] 除非另有说明,核苷酸序列以5'至3'方向从左向右书写。氨基酸序列以氨基到羧基的方向从左到右书写。

[0131] 本文提供的标题不是对本公开的各个方面的限制,可以通过参考本说明书作为一个整体来理解。因此,下面定义的术语通过参考整个说明书更充分地定义。

[0132] “氨基酸”是具有以下结构的分子:中心碳原子(α -碳原子)连接到氢原子、羧基团(其碳原子在本文中称为“羧基碳原子”)、氨基(其氮原子在本文中称为“氨基氮原子”)和侧链基团R。当氨基酸引入肽、多肽或蛋白质时,在将一个氨基酸连接到另一个氨基酸的脱水反应中一个氨基酸失去一个或多个其氨基酸羧基原子。因此,当引入到蛋白质中时,氨基酸被称为“氨基酸残基”

[0133] “蛋白质”或“多肽”是指通过肽键连接的两个或多个单个氨基酸(无论是否天然存在)的任何聚合物,当与一个氨基酸(或氨基酸残基)的 α 碳键合的羧基的羧基碳原子与相邻氨基酸的非 α 碳键的氨基的氨基氮原子共价键合时产生“蛋白质”或“多肽”。术语“蛋白质”理解为包括其含义内的术语“多肽”和“肽”(在本文中有时可以互换使用)。此外,包含多个多肽亚基的蛋白质也应理解为包含在本文所用的“蛋白质”的含义内。类似地,蛋白质和多肽的片段也在本公开的范围,在本文中称为“蛋白质”。在本公开的一个方面,多肽包含两个或多个亲本肽片段的嵌合体。术语“多肽”还意指并包含多肽翻译后修饰(“PTM”)的产物,包括但不限于二硫键形成、糖基化、氨甲酰化、脂化、乙酰化、磷酸化、酰胺化、已知保护/阻断基团的衍生化、蛋白水解切割、通过非天然存在的氨基酸进行修饰,或任何其他操作或修饰,例如与标记成分的缀合。多肽可以来源于天然生物源或通过重组技术生产,不必从指定的核酸序列翻译。它可以以任何方式产生,包括通过化学合成。“分离的”多肽或其片段、变体或衍生物是指不在其自然环境中的多肽。不需要特定的纯化水平。例如,分离的多肽可以是仅从其天然或自然环境中移出。出于本公开的目的,认为宿主细胞中表达的重组产生的多肽和蛋白质是分离的,认为是通过任何合适的技术分离、分级或部分或基本纯化的天然或重组多肽。

[0134] 本文所用的“结构域”可与术语“肽节段”互换使用,并指较大多肽或蛋白质的一部分或片段。一个结构域本身不需要具有功能活性,尽管在某些情况下,一个结构域可以具有自己的生物活性。

[0135] 当提及两个或多个结构域时,“融合的”、“有效连接的”和“有效关联的”在本文中可互换地使用,以广义地指本文所公开的重组多肽形成过程中两个或更多结构域的任何化学或物理耦合。在一个实施方案中,本文公开的重组多肽是包含来自两种或多种不同多肽的多个结构域的嵌合多肽。

[0136] 如本文所公开的包含两个或多个结构域和/或蛋白质的重组多肽(即重组蛋白质)可以由单个编码序列编码,该编码序列包含编码每个结构域和或蛋白质的多核苷酸序列。除非另有说明,否则编码每个结构域和/或蛋白质的多核苷酸序列处于“符合读框的”,使得包含多核苷酸序列的单个mRNA的翻译产生包含每个结构域和/或蛋白质的单个多肽。通常,本文所述的重组多肽中的结构域和/或蛋白质将直接相互融合或通过肽接头分离。编码肽接头的各种多核苷酸序列是本领域已知的,并且包括例如自切割肽。

[0137] 本文所用的“多核苷酸”或“核酸”是指核苷酸的聚合形式。在某些情况下,多核苷酸包含与编码序列不直接邻接或与衍生多核苷酸的生物体的天然基因组中的编码序列直接邻接(5'端或3'端)的序列。因此,该术语包括例如重组DNA,其被并入载体、自主复制的质粒或病毒、或原核生物或真核生物的基因组DNA中,或作为独立于其他序列的单独分子(例如,cDNA)存在。本公开的核苷酸可以是核糖核苷酸、脱氧核糖核苷酸或两者之一核苷酸的修饰形式。本文所用的多核苷酸尤其是指单链和双链DNA、单链和双链区域的混合物的DNA、单链和双链RNA、单链和双链区的混合的RNA、包含DNA和RNA的杂合分子,所述DNA和RNA可以是单链的、或更一般地是双链的或是单链和双链区的混合物。术语多核苷酸包含基因组DNA或RNA(取决于生物体,即病毒的RNA基因组),以及基因组DNA编码的mRNA和cDNA。在某些实施方案中,多核苷酸包含常规磷酸二酯键或非常规键(例如酰胺键,如在肽核酸(PNA)中发现的酰胺键)。“分离”的核酸或多核苷酸是指已从天然环境中取出的核酸分子,例如DNA或RNA。例如,出于本公开的目的,包含在载体中含有的编码重组多肽的多核苷酸的核酸分子被认为是“分离的”。分离的多核苷酸的进一步实例包括维持在异源宿主细胞中或在溶液从其他多核苷酸(部分或基本上)纯化的重组多核苷酸。分离的RNA分子包括本公开的多核苷酸的体内或体外RNA转录物。根据本公开的分离的多核苷酸或核酸进一步包括合成产生的多核苷酸和核酸(例如核酸分子)。

[0138] 如本文所用,“编码区”或“编码序列”是由可翻译成氨基酸的密码子组成的多核苷酸的一部分。尽管“终止密码子”(TAG、TGA或TAA)通常不翻译成氨基酸,但它可以被认为是编码区的一部分,但任何侧翼序列,例如启动子、核糖体结合位点、转录终止子、内含子等,都不是编码区的组成部分。编码区的边界通常由编码所得多肽的氨基末端的5'末端的起始密码子和编码所得多肽的羧基末端的3'末端的翻译终止密码子确定。

[0139] 如本文所用,术语“表达控制区”是指与编码区可操作地关联以指导或控制由编码区编码的产物表达的转录控制元件,包括例如启动子、增强子、操纵子、阻遏子、核糖体结合位点、翻译前导序列、内含子、聚腺苷酸化识别序列、RNA加工位点、效应子结合位点、茎环结构和转录终止信号。例如,如果启动子功能的诱导导致包含编码产物的编码区的mRNA的转录、如果启动子和编码区之间的连接的性质不干扰启动子引导由编码区编码的产物表达的能力或不干扰DNA模板转录的能力,则编码区和启动子是“有效关联的”(即“有效连接的”)。表达控制区包括位于编码区上游(5'非编码序列)、内部或下游(3'非编码顺序)的核苷酸序列,其影响相关编码区的转录、RNA加工、稳定性或翻译。如果打算在真核细胞中表达编码

区,则聚腺苷酸化信号和转录终止序列通常位于编码序列的3'处。

[0140] 如本文所用,术语“宿主细胞”和“细胞”可以互换使用,并且可以指任何类型的细胞或细胞群,例如原代细胞、培养中的细胞或来自细胞系的细胞,其含有或能够含有核酸分子(例如重组核酸分子)。宿主细胞可以是原核细胞,或者,宿主细胞可以为真核细胞,例如真菌细胞,如酵母细胞,以及各种动物细胞,如昆虫细胞或哺乳动物细胞。

[0141] 本文所用的“培养”是指在允许细胞生长或分裂的体外条件下孵育细胞或将细胞维持在存活状态。本文所用的“培养细胞”是指在体外繁殖的细胞。

[0142] “受试者”包括任何人或非人动物。术语“非人动物”包括但不限于脊椎动物,如哺乳动物、鸟类、宠物、农场动物、非人灵长类动物、绵羊、奶牛、山羊、猪、鸡、狗、猫和啮齿动物,如小鼠、大鼠和豚鼠。在优选方面,受试者是人类。术语“受试者”和“患者”在本文中互换使用。

[0143] “施用”是指使用本领域技术人员已知的各种方法和递送系统中的任何一种将治疗剂物理引入受试者。

[0144] 本文所用的受试者的术语“治疗”是指对受试者进行的任何类型的干预或过程,或向受试者施用活性剂,目的是逆转、缓解、改善、抑制或减缓或防止症状、并发症、疾病的进展、发展、严重程度或复发、或者是逆转、缓解、改善、抑制或减缓或防止与疾病相关的生化指标、或提高总体生存率。治疗可以是对患有疾病的受试者或没有疾病的受试者(例如,用于预防,如接种疫苗)。

[0145] 术语“有效剂量”或“有效量”定义为足以达到或至少部分达到预期效果的量。药物或治疗剂的“治疗有效量”或“治疗有效剂量”是指当单独使用或与另一种治疗剂联合使用时,通过降低疾病症状的严重程度、增加疾病症状无症状期的频率和持续时间、促进疾病消退的任何量的药物、总体存活率的增加(从诊断之日或疾病开始治疗之日起被诊断为该疾病的患者仍然存活的时间长度)或预防由于疾病折磨而导致的损伤或残疾。药物的治疗有效量或剂量包括“预防有效量”或“预防有效剂量”,这是指当单独或与另一种治疗剂联合施用于有疾病发展或疾病复发风险的受试者时,抑制疾病发展或复发的任何量的药物。治疗剂促进疾病消退或抑制疾病发展或复发的能力可以使用本领域技术人员已知的多种方法来评估,例如在临床试验期间在人类受试者中,在预测人类疗效的动物模型系统中,或通过体外试验中测定该物质的活性来评估。

[0146] 本公开的各个方面在以下子部分中进一步详细描述。

[0147] II. 用于产生VLP的载体

[0148] 美国专利号9,290,778和9,862,954;Nafissi和Slavcev, *Microbial Cell Factories* 11:154(2012);和Nafissi等人, *Nucleic Acids* 3(6):e165(2014)描述了无细菌序列载体及其生产,所述文献通过引用全文并入本文作为参考。这些无细菌序列的载体是由表达载体(例如质粒)产生的,该表达载体含有包含重组酶靶序列的特化“超级序列”(“Super Sequence(SS)”)位点。SS位点位于包含目的核酸的表达盒侧。当表达载体存在于表达适当重组酶的重组细胞中时,含有表达盒的无细菌序列载体与表达载体的主链DNA分离。为了产生环状共价闭合(CCC)无细菌序列载体,使用这样的生产系统,其中例如重组细胞表达Cre或Flp重组酶,且表达载体包含重组酶的相应靶序列。为了产生线性共价闭合(LCC)无细菌序列载体,本文也称为微链DNA(msDNA),使用这样的生产系统,其中例如重组

细胞表达TelN或Tel重组酶,且表达载体包含重组酶的相应靶序列。然后可以从细胞中纯化无细菌序列载体并直接用作递送载体。参见美国专利号9,290,778和9,862,954,Nafissi和Slavcev,以及Nafissi等人。

[0149] 具有LCC末端的msDNA载体是无扭转的,在大肠杆菌中生产时不会受到回旋酶针对的不利超螺旋。示例性的msDNA载体携带带有真核启动子、目的基因(GOI)、内含子和polyA序列以及核易位增强序列的表达盒(Nafissi和Slavcev,以及Nafisi等人)。此外,由于其双链LCC拓扑结构,msDNA整合到细胞染色体中会导致染色体断裂,从而将细胞从群体中剔除。因此,msDNA消除了任何插入突变的风险,保护了施用msDNA的患者免受潜在的遗传毒性和癌症(Nafissi等人)。

[0150] 在一些方面,本文公开的用于产生VLP的无细菌序列载体包含根据本领域已知的任何其他方法产生的CCC或LCC载体。

[0151] A. 表达载体、表达盒和用于产生无细菌序列载体的载体生产系统和VLP

[0152] 本文提供了一种表达载体,其包含表达盒,所述表达盒包含编码重组蛋白的核酸序列,所述重组蛋白包含与免疫原性氨基酸序列融合的病毒的保守氨基酸序列,其中从表达盒细胞内表达的蛋白质能够形成VLP。

[0153] 本文提供了一种表达载体,其包含表达盒,所述表达盒包含编码重组蛋白的核酸序列,所述重组蛋白包含与免疫原性氨基酸序列融合的来自病毒的保守氨基酸序列;包含位于所述表达盒每一侧翼的用于第一重组酶的靶序列;和整合在用于所述第一重组酶的靶序列的非结合区内的用于一个或多个额外重组酶的一个或多个额外靶序列,其中从所述表达盒细胞内表达的蛋白质能够形成VLP

[0154] 保守和免疫原性氨基酸序列包含本领域已知的序列以及通过已知技术确定的序列。例如,基于基因组的反向疫苗学可应用于比较基因组学分析,这是一个生物学研究领域,可用于比较不同致病株之间的基因组序列(参见例如,Sieb等人,Clin.Microbiol.Infect.

[0155] 18(Suppl.5):109-116(2012))。也可以使用其他测序、结构和计算方法(参见例如,Liljeroos等人,J.Immunol.Res.2015:156241;Sette和Rappuoli,Immunity 33(4):530-541(2010))。

[0156] 在一些方面,免疫原性氨基酸序列与保守氨基酸序列来自同一病毒。在一些方面,保守氨基酸序列来自病毒糖蛋白。在一些方面,免疫原性氨基酸序列来自相同的病毒糖蛋白。

[0157] 在一些方面,表达盒还包含编码病毒包膜蛋白的核酸序列和/或编码病毒基质蛋白的核酸序列。在一些方面,病毒包膜蛋白和/或病毒基质蛋白来自与保守氨基酸序列相同的病毒。

[0158] 在一些方面,保守氨基酸序列、免疫原性氨基酸序列、病毒包膜蛋白和/或病毒基质蛋白是共有序列。

[0159] 在一些方面,重组蛋白能够刺激针对病毒的免疫应答,包括产生中和抗体。例如,保守位点通常被广泛中和抗体识别,并且易受抗体失活的影响(参见例如Nabel,N.Engl.J.Med.368(6):551-560(2013))。

[0160] 在一些方面,重组蛋白能够刺激Th1细胞介导的针对病毒的免疫应答。细胞介导的

免疫是细胞毒性T细胞识别抗原感染细胞以诱导细胞裂解的过程。

[0161] 在一些方面,免疫应答对相关病毒或毒株具有交叉反应。例如,不同病毒血清型/毒株之间的保守序列可用于提供对多种血清型/毒株的保护,包括作为通用疫苗。

[0162] 在一些方面,重组蛋白排除了病毒中刺激包括产生非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。

[0163] 在一些方面,表达盒包含单个可读框,其包含每个编码蛋白质的核酸序列之间的编码自切割肽的核酸序列,使得表达盒的翻译产物在细胞内切割成两个或多个蛋白质。在一些方面,自切割肽是2A自切割肽。在一些方面,2A自切割肽是来自猪捷申病毒-1的P2A。在一些方面,2A自切割肽是来自明脉扁刺蛾病毒2A(thosea asigna virus 2A)的T2A。

[0164] 在一些方面,表达盒包含编码自切割肽的核酸序列,所述自切割肽位于编码病毒基质蛋白和病毒包膜蛋白的核酸序列之间,位于编码病毒基质蛋白和重组蛋白的核酸序列之间,和/或位于编码病毒包膜蛋白和重组蛋白的核酸序列之间。在一些方面,表达盒包含编码病毒基质蛋白、自裂解肽、病毒包膜蛋白、自切割肽和重组蛋白的5'至3'核酸序列。在一些方面,表达盒包含编码病毒包膜蛋白、自裂解肽、病毒基质蛋白、自切割肽和重组蛋白的5'至3'核酸序列。

[0165] 在一些方面,表达盒还包含编码基因表达标志物的核酸序列。在一些方面,基因表达标志物是荧光报告基因,例如绿色荧光蛋白(GFP)、红色荧光蛋白(RFP)、黄色荧光蛋白(YFP)或近红外荧光蛋白(iRFP);生物发光报告基因,如荧光素酶;选择性抗生素标志物;或LacZ。在一些方面,表达盒包含编码自切割肽的核酸序列,该自切割肽位于编码用于基因表达的标志物的核酸序列和编码蛋白质的任何其他核酸序列之间。

[0166] 表达盒可以包含本领域技术人员已知的任何可操作地连接到蛋白质编码核酸序列的表达控制区。在一些方面,表达控制区是启动子、增强子、操纵子、阻遏子、核糖体结合位点、翻译前导序列、内含子、聚腺苷酸化识别序列、RNA加工位点、效应子结合位点、茎环结构、转录终止信号或其组合。

[0167] 在一些方面,用于第一重组酶的靶序列和用于一个或多个额外重组酶的一个或多个额外靶序列选自PY54 pal位点、N15 telRL位点、loxP位点、 ϕ K02 telRL位点、FRT位点、phiC31 attP位点和 λ attP位点。在一些方面,表达载体包含每个所述靶序列。在一些方面,表达载体包含Tel重组酶pal位点和整合在pal位点内的telRL、loxP和FRT重组酶靶结合序列。

[0168] 在一些方面,表达载体用于生产无细菌序列的载体。在一些方面,无细菌序列载体具有环状共价闭合末端。在一些方面,无细菌序列载体具有线性共价闭合末端。

[0169] 在一些方面,表达载体还包含至少一个位于第一重组酶靶序列每一侧的增强子序列。在一些方面,所述至少一个增强子序列是至少两个增强子序列。在一些方面,所述至少一个增强子序列是SV40增强子序列。

[0170] 本文公开的保守氨基酸序列、免疫原性氨基酸序列和/或病毒蛋白的来源可以是与人类或动物感染相关的任何病毒。

[0171] 在一些方面,病毒是冠状病毒、流感病毒、人类免疫缺陷病毒、人类乳头瘤病毒、肝炎病毒或溶瘤病毒。

[0172] 在一些方面,流感病毒是甲型流感病毒。在一些方面,甲型流感病毒是H1N1、H5N1

或H3N2。

[0173] 在一些方面,流感病毒是乙型流感病毒。

[0174] 在一些方面,冠状病毒是人类冠状病毒,例如但不限于HCoV-229E,HCoV-NL63,HCoV-OC43,HCoV-HKU1,SARS-CoV-1,SARS-CoV-2(即,COVID-19)和/或MERS-CoV。

[0175] 在一些方面,冠状病毒是COVID-19(即,WHCV或其变体,例如但不限于英国变体B.1.1.7、南非变体B.1.351、巴西变体P.1或加利福尼亚变体B.1.427/429)。

[0176] 本文提供了一种载体生产系统,其包含重组细胞,所述重组细胞被设计为在可诱导启动子的控制下编码至少第一重组酶,其中所述细胞包含本文公开的表达载体,所述表达载体包含所述至少一种重组酶的靶。在一些方面,诱导型启动子是热调节的、化学调节的、IPTG调节的、葡萄糖调节的、阿拉伯糖诱导的、T7聚合酶调节的、冷休克诱导的、pH诱导的或其组合。在一些方面,所述至少第一重组酶选自telN和tel,并且所述表达载体包含所述至少第一重组酶的靶序列。在一些方面,所述至少第一重组酶选自Cre或Flp,并且所述表达载体包含所述至少第一重组酶的靶序列。在一些方面,重组细胞已被进一步设计为编码核酸酶基因组编辑系统,并且表达载体还包含含有核酸酶基因组编码系统的切割位点的主链序列。在一些方面,核酸酶基因组编辑系统是包含Cas核酸酶和gRNA的CRISPR核酸酶系统,并且表达载体包含主链序列内gRNA的靶序列。

[0177] 本文提供了一种产生无细菌序列载体的方法,包括在合适的条件下孵育本文所述的载体生产系统,以表达至少第一重组酶或第一重组酶和核酸酶基因组编辑系统。在一些方面,该方法还包括收获无细菌序列载体。本公开还涉及通过该方法产生的无细菌序列载体。

[0178] A.1包含冠状病毒序列的表达盒

[0179] 冠状病毒包括冠状病毒科(Coronaviridae)的任何病毒,包括冠状病毒亚科,包括Alphacoronavirus,Betacoronavirus,Gammacoronavirus和Deltacoronavirus属。例如,参见Fung和Liu(2019)。冠状病毒包括人类冠状病毒(HCoV),例如HCoV-229E,HCoV-NL63,HCoV-OC43,HCoV-HKU1,严重急性呼吸综合征冠状病毒(SARS-CoV,如,SARS-CoV-1和SARS-CoV-2(即,COVID-19)),中东呼吸综合征冠状病毒(MERS-CoV),人畜共患冠状病毒(例如,SARS-CoV和MERS-CoV),蝙蝠冠状病毒(BtCoVs),禽冠状病毒,鼠冠状病毒和bulbol冠状病毒(BuCoV)。

[0180] 冠状病毒基因组是正义、非节段的、单链RNA,约为27至32千碱基(参见例如,Fung和Liu,Annu.Rev.Microbiol.73:529-557(2019))。例如,与基因组分别为27.9kb和30.1kb的SARS-CoV和MERS-CoV相比,COVID-19(也称为WHCV、SARS-CoV-2和2019-nCoV)的完整基因组的大小为29.9kb(Zhou等人,Nature 579:270-273(2020))。已发现COVID-19基因组与蝙蝠CoV RaTG13基因组96.2%同一,后者是在蝙蝠中发现的一种SARS-CoV-2类型,可能是通过未知中间宿主传播给人类的病毒来源。

[0181] 冠状病毒具有膜(M)蛋白,它是最丰富的结构蛋白,支持病毒包膜,并以三个跨膜结构域嵌入包膜。M蛋白是病毒组装和出芽所必需的。

[0182] 包膜(E)蛋白是冠状病毒中的一种小跨膜蛋白,其也在包膜中以低于M蛋白的量存在。E蛋白也参与病毒的组装和排出。

[0183] 冠状病毒中的核衣壳(N)蛋白像串珠一样与RNA基因组结合,形成螺旋对称的核衣

壳。

[0184] 冠状病毒的病毒子表面被三聚的刺突(S)蛋白修饰。一些β冠状病毒也有二聚的血凝素酯酶(HE)蛋白,它们在病毒表面形成较短的突起。S和HE蛋白均为I型跨膜蛋白,具有大的外结构域和短的内结构域。

[0185] S蛋白包含两个亚基,S1和S2,并在其C端锚定于病毒包膜。例如,COVID-19的S1亚基包含N末端结构域(NTD)和受体结合结构域(RBD),而S2亚基包含融合肽(FP)、内部融合肽(IFP)、七肽重复序列1/2(HR1/2)和跨膜结构域(TM)。S蛋白的大外结构域三聚并在病毒表面形成特征性冠状病毒刺突。S蛋白负责受体结合和病毒子进入宿主细胞(Fehr和Perlman, *Coronaviruses: An Overview of Their Replication and Pathogenesis*. 收录于: Maier H., Bickerton E., Britton P. (编辑) *Coronaviruses. Methods in Molecular Biology*, vol1282. Humana Press, New York, NY; Wall等人, *Cell* 180:1-12 (2020))。

[0186] 许多病毒的融合蛋白需要在融合肽附近发生蛋白水解事件,以使病原体进入靶细胞。例如,来自COVID-19的S蛋白具有两个切割位点,其中第一个位于S1/S2边界,但与融合肽没有紧密连接。第二个切割位点(S2')暴露内部融合肽(IFP),这是紧挨S2'下游的基序,在所有测序的冠状病毒中高度保守。IFP的序列是SFIEDLLFNKVTADAGF(SEQ ID NO:7),其中加粗的LLF残基对于膜融合和感染性至关重要(Madu等人, *J. Virol.* 83(15):7411-7421 (2009))。COVID-19显示在S1/S2位点存在典型的弗林蛋白酶样切割基序,该基序在同一分支的其他冠状病毒中未发现,但在特别强毒的流感病毒(H5N1)中也有类似发现。通过S1/S2界面处的其他蛋白酶如弗林蛋白酶的切割可能会扩大病毒的向性(tropism),使动物至人的传播更可能(Coutard等人, *Antiviral Res.* 176:104742 (2020))。

[0187] 在一些方面,表达盒包含编码冠状病毒膜(M)蛋白、冠状病毒包膜(E)蛋白和重组蛋白的核酸序列,所述重组蛋白包含来自冠状病毒刺突(S)蛋白的保守氨基酸序列和免疫原性氨基酸序列。M、E和S蛋白在本文中可以互换地称为M、E和S糖蛋白。

[0188] 在一些方面,M蛋白包含与SEQ ID NO:1至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的氨基酸序列。在一些方面,M蛋白包含SEQ ID NO:1。在一些方面,M蛋白是SEQ ID NO:1。

[0189] 在一些方面,编码M蛋白的核酸序列与SEQ ID NO:2至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,编码M蛋白质的核酸序列包含SEQ ID NO:2。在一些方面,编码M蛋白的核酸序列是SEQ ID NO:2。

[0190] 在一些方面,E蛋白包含与SEQ ID NO:3至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的氨基酸序列。在一些方面,E蛋白包含SEQ ID NO:3。在一些方面,E蛋白是SEQ ID NO:3。在一些方面,E蛋白包含用另一个氨基酸替换位于SEQ ID NO:3中氨基酸编号71的脯氨酸(即,SEQ ID NO:3中的P71)。在一些方面,SEQ ID NO:3中P71处的替换是从脯氨酸到亮氨酸的改变(即P71L)。

[0191] 在一些方面,编码E蛋白的核酸序列与SEQ ID NO:4至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,编码E蛋白质的核酸序列包含SEQ ID NO:4。在一些方面,编码E

蛋白的核酸序列是SEQ ID NO:4。在一些方面,编码E蛋白的核酸序列包含

[0192] SEQ ID NO:4中核苷酸编号211-213处的脯氨酸密码子被另一氨基酸密码子替换。在一些方面,SEQ ID NO:4中核苷酸编号211-213处的脯氨酸密码子被亮氨酸密码子替换。

[0193] 在一些方面,保守氨基酸序列来自S蛋白的S1亚基或S2亚基、S蛋白的RBD、S蛋白S2'切割位点和S蛋白的内部融合肽(IFP)(本文中称为S2' IFP)、M蛋白或E蛋白。

[0194] 在一些方面,所述保守氨基酸序列与任一SEQ ID NO:12-54至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,保守氨基酸序列包含任一SEQ ID NO:12-54。在一些方面,保守氨基酸序列是任一SEQ ID NO:12-54。

[0195] 在一些方面,所述保守氨基酸序列与SEQ ID NO:7至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,保守氨基酸序列包含SEQ ID NO:7。在一些方面,保守氨基酸序列是SEQ ID NO:7。

[0196] 在一些方面,编码重组蛋白的保守氨基酸序列的核酸序列与SEQ ID NO:8至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,编码重组蛋白的保守氨基酸序列的核酸序列包含SEQ ID NO:8。在一些方面,编码重组蛋白的保守氨基酸序列的核酸序列是SEQ ID NO:8。

[0197] 在一些方面,免疫原性氨基酸序列来自S蛋白受体结合结构域(RBD)。

[0198] 在一些方面,所述免疫原性氨基酸序列与SEQ ID NO:11至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,免疫原性氨基酸序列包含SEQ ID NO:11。在一些方面,免疫原性氨基酸序列为SEQ ID NO:11。在一些方面,免疫原性蛋白质包含以下一个或多个替换:SEQ ID NO:11中的位于氨基酸编号88的赖氨酸(即K88)、位于氨基酸编号123的亮氨酸(即L123)、位于氨基酸编号155的谷氨酸(即E155)、或位于氨基酸编号172的天冬酰胺(即N172)(分别对应于SEQ ID NO:5中的K417,L452,E484,和N501)用另一个氨基酸替换。在一些方面,K88处的替换是K88N(即,从赖氨酸变为天冬酰胺)。在一些方面,K88处的替换是K88T(即,从赖氨酸变为苏氨酸)。在一些方面,L123处的替换是L123R(即,从亮氨酸变为精氨酸)。在一些方面,E155处的替换是E155K(即,从谷氨酸变为赖氨酸)。在一些方面,N172处的替换是N172Y(即,从天冬酰胺变为酪氨酸)。

[0199] 在一些方面,编码重组蛋白的免疫原性氨基酸序列的核酸序列与SEQ ID NO:101至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,编码重组蛋白的免疫原性氨基酸序列的核酸序列包含SEQ ID NO:101。在一些方面,编码重组蛋白的免疫原性氨基酸序列的核酸序列是SEQ ID NO:101。在一些方面,编码免疫原性蛋白的核酸序列包含包含以下一个或多个替换:SEQ ID NO:101的核苷酸编号262-264处的赖氨酸密码子用另一氨基酸的密码子替换、SEQ ID NO:101的核苷酸编号367-369处的亮氨酸密码子用另一氨基酸的密码子替换、SEQ ID NO:101的核苷酸编号463-465处的谷氨酸密码子用另一氨基酸的密码子替换或SEQ ID NO:101的核苷酸编号514-516处的天冬酰胺密码子用另一氨基酸的密码子替换。在

一些方面,核苷酸编号262-264处的赖氨酸密码子用天冬酰胺或苏氨酸密码子替换。在一些方面,核苷酸编号367-369处的亮氨酸密码子用精氨酸密码子替换。在一些方面,核苷酸编号463-465处的谷氨酸密码子用赖氨酸密码子替换。在一些方面,核苷酸编号514-516处的天冬酰胺密码子用酪氨酸密码子替换。

[0200] 在一些方面,重组蛋白进一步包含来自S蛋白的跨膜(TM)结构域序列。

[0201] 在一些方面,TM结构域序列包含与SEQ ID NO:102至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的氨基酸序列。在一些方面,TM结构域序列包含SEQ ID NO:102。在一些方面,TM结构域序列是SEQ ID NO:102。

[0202] 在一些方面,编码重组蛋白的TM结构域序列的核酸序列与SEQ ID NO:103至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,编码重组蛋白的TM结构域序列的核酸序列包含SEQ ID NO:103。在一些方面,编码重组蛋白TM结构域的核酸序列是SEQ ID NO:103。

[0203] 在一些方面,重组蛋白包含来自S2' IFP的保守氨基酸序列、来自RBD的免疫原性氨基酸序列和S蛋白的TM结构域序列。

[0204] 在一些方面,重组蛋白排除了S蛋白中刺激包含非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。

[0205] 在一些方面,重组蛋白的氨基酸序列与SEQ ID NO:55至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,重组蛋白质的氨基酸序列包含SEQ ID NO:55。在一些方面,重组蛋白的氨基酸序列是SEQ ID NO:55。在一些方面,重组蛋白包含用另一个氨基酸替换SEQ ID NO:55的K88,L123,E155或N172中的一个或多个。在一些方面,K88的替换是K88N。在一些方面,K88的替换是K88T。在一些方面,L123的替换是L123R。在一些方面,E155的替换是E155K。在一些方面,N172的替换是N172Y。

[0206] 在一些方面,编码重组蛋白的核酸序列与SEQ ID NO:56至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,编码重组蛋白的核酸序列包含SEQ ID NO:56。在一些方面,编码重组蛋白核酸序列是SEQ ID NO:56。在一些方面,编码重组蛋白的核酸序列包含以下一个或多个替换:SEQ ID NO:56的核苷酸编号262-264处的赖氨酸密码子用另一氨基酸的密码子替换、SEQ ID NO:56的核苷酸编号367-369处的亮氨酸密码子用另一氨基酸的密码子替换、SEQ ID NO:56的核苷酸编号463-465处的谷氨酸密码子用另一氨基酸的密码子替换、或者SEQ ID NO:56的核苷酸编号514-516处的天冬酰胺密码子用另一氨基酸的密码子替换。在一些方面,核苷酸编号262-264处的赖氨酸密码子用天冬酰胺或苏氨酸的密码子替换。在一些方面,核苷酸编号367-369处的亮氨酸密码子被精氨酸密码子替换。在一些方面,核苷酸编号463-465处的谷氨酸密码子被赖氨酸密码子替换。在一些方面,核苷酸编号514-516处的天冬酰胺密码子被酪氨酸密码子替换。

[0207] 在一些方面,表达盒包含翻译为与SEQ ID NO:57至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的氨基酸序列的单个可读框。在一些方面,表达盒包含翻译为包含

[0208] SEQ ID NO:57的氨基酸序列的单个可读框。在一些方面,表达盒包含翻译为SEQ IDNO:57的氨基酸序列的单个可读框。在一些方面,表达盒包含翻译为这样的氨基酸序列的单个可读框,该氨基酸序列包含用另一氨基酸替换SEQ ID NO:57的P71,K423,L458,E490或N507中的一个或多个。在一些方面,在P71处的替换是P71L。在一些方面,K423处的替换是K423N。在一些方面,在K423处的替换是K423T。在一些方面,L458处的替换是L458R。在一些方面,E490的替换是E490K。在一些方面,N507处的替换是N507Y。

[0209] 在一些方面,表达盒包含与SEQ ID NO:58至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的单个可读框。在一些方面,表达盒包含含有SEQ ID NO:58的单个可读框。在一些方面,表达盒包含SEQ ID NO:58的单个可读框。在一些方面,表达盒包含含有以下一个或多个替换的单个可读框:SEQ ID NO:58的核苷酸编号211-213处的脯氨酸密码子用另一个氨基酸的密码子替换、SEQ ID NO:58的核苷酸编号1267-1269处的赖氨酸密码子用另一个氨基酸的密码子替换、SEQ ID NO:58的核苷酸编号1372-1374处的亮氨酸密码子用另一个氨基酸的密码子替换、SEQ ID NO:58的氨基酸编号1468-1470处的谷氨酸密码子用另一个氨基酸的密码子替换、或者SEQ ID NO:58的核苷酸编号1519-1521处的天冬酰胺密码子用另一个氨基酸的密码子替换。在一些方面,SEQ ID NO:58的核苷酸编号211-213处的脯氨酸密码子被亮氨酸密码子替换。在一些方面,核苷酸编号1267-1269处的赖氨酸密码子被天冬酰胺或苏氨酸密码子替换。在一些方面,核苷酸编号1372-1374处的亮氨酸密码子被精氨酸密码子替换。在一些方面,核苷酸编号1468-1470的谷氨酸密码子被赖氨酸密码子替换。在一些方面,核苷酸编号1519-1521处的天冬酰胺密码子被酪氨酸密码子替换。

[0210] 在一些方面,表达盒与任一SEQ ID NO:59-62的核酸序列至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,表达盒包含任一SEQ ID NO:59-62的核酸序列。在一些方面,表达盒是任一SEQ ID NO:59-62的核酸序列。

[0211] 在一些方面,重组蛋白能够刺激针对COVID-19的免疫应答。

[0212] 在一些方面,重组蛋白能够刺激Th1细胞介导的针对COVID-19的免疫应答。

[0213] 在一些方面,针对COVID-19的免疫应答针对WHCV和/或一个或多个变体,例如但不限于英国变体B.1.1.7、南非变体B.1.351、巴西变体P.1或加利福尼亚变体B.1.427/429。

[0214] 在一些方面,免疫应答对其他冠状病毒具有交叉反应。在一些方面,免疫应答对其他严重急性呼吸综合征冠状病毒和/或人类β冠状病毒具有交叉反应。

[0215] 本文提供了编码与SEQ ID NO:57至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的氨基酸序列的多核苷酸。在一些方面,该多核苷酸编码包含SEQ ID NO:57的氨基酸序列。在一些方面,该多核苷酸编码SEQ ID NO:57的氨基酸序列。在一些方面,该多核苷酸编码包含SEQ ID NO:57中P71,K423,L458,E490或N507的一个或多个替换的氨基酸序列。在一些方面,在P71处的替换是P71L。在一些方面,K423处的替换是K423N。在一些方面,在K423处的替换是K423T。在一些方面,L458处的替换是L458R。在一些方面,E490的替换是E490K。在一些方面,N507处的替换是N507Y。

[0216] 本文提供了一种多核苷酸,其包含与SEQ ID NO:58至少约90%、至少约91%、至少

约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的核酸序列。在一些方面,所述多核苷酸包含SEQ ID NO:58。在一些方面,所述多核苷酸是SEQ ID NO:58。在一些方面,所述多核苷酸包含核酸序列,所述核酸序列包含以下一个或多个替换:SEQ ID NO:58中核苷酸编号211-213处的脯氨酸密码子用另一氨基酸的密码子替换、SEQ ID NO:58中核苷酸编号1267-1269处的赖氨酸密码子用另一氨基酸的密码子替换、SEQ ID NO:58中核苷酸编号1372-1374处的亮氨酸密码子用另一氨基酸的密码子替换、SEQ ID NO:58中核苷酸编号1468-1470处的谷氨酸密码子用另一氨基酸的密码子替换、或者SEQ ID NO:58中核苷酸编号1519-1521处的天冬酰胺密码子用另一氨基酸的密码子替换。在一些方面,SEQ ID NO:58中核苷酸编号211-213处的脯氨酸密码子被亮氨酸密码子替换。在一些方面,核苷酸编号1267-1269处的赖氨酸密码子被天冬酰胺或苏氨酸密码子取代。在一些方面,核苷酸编号1372-1374处的亮氨酸密码子被精氨酸密码子替换。在一些方面,核苷酸编号1468-1470的谷氨酸密码子被赖氨酸密码子替换。在一些方面,核苷酸编号1519-1521处的天冬酰胺密码子被酪氨酸密码子替换。

[0217] B. 无细菌序列载体

[0218] 本公开的无细菌序列载体可以包括本公开的任何表达盒。

[0219] 本文提供了一种无细菌序列载体,其包含表达盒,所述表达盒包含编码重组蛋白的核酸序列,所述重组蛋白包含与免疫原性氨基酸序列融合的来自病毒的保守氨基酸序列,其中从表达盒胞内表达的蛋白质能够形成VLP。

[0220] 在一些方面,免疫原性氨基酸序列与保守氨基酸序列来自同一病毒。在一些方面,保守氨基酸序列来自病毒糖蛋白。在一些方面,免疫原性氨基酸序列来自相同的病毒糖蛋白。

[0221] 在一些方面,表达盒还包含编码病毒包膜蛋白的核酸序列和/或编码病毒基质蛋白的核酸序列。在一些方面,病毒包膜蛋白和/或病毒基质蛋白来自与保守氨基酸序列相同的病毒。

[0222] 在一些方面,保守氨基酸序列、免疫原性氨基酸序列、病毒包膜蛋白和/或病毒基质蛋白是共有序列。

[0223] 在一些方面,重组蛋白能够刺激针对病毒的包含中和抗体的免疫应答。

[0224] 在一些方面,重组蛋白能够刺激Th1细胞介导的针对病毒的免疫应答。

[0225] 在一些方面,免疫应答对相关病毒或毒株具有交叉反应。

[0226] 在一些方面,重组蛋白排除了病毒中刺激包含非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。

[0227] 在一些方面,表达盒包含单个可读框,其包含编码蛋白质的每个核酸序列之间的自切割肽的编码核酸序列。表达盒和自切割肽包括上文关于表达载体所讨论的那些。

[0228] 在一些方面,病毒是冠状病毒、流感病毒、人类免疫缺陷病毒、人类乳头瘤病毒、肝炎病毒或溶瘤病毒。

[0229] 在一些方面,流感病毒是甲型流感病毒。在一些方面,甲型流感病毒是H1N1, H5N1或H3N2。

[0230] 在一些方面,流感病毒是乙型流感病毒。

[0231] 在一些方面,冠状病毒是人类冠状病毒,例如但不限于HCoV-229E, HCoV-NL63,

HCoV-OC43, HCoV-HKU1, SARS-CoV-1, SARS-CoV-2 (即, COVID-19) 和/或MERS-CoV。

[0232] 在一些方面, 冠状病毒是COVID-19 (即WHCV或其变体, 例如但不限于英国变体B.1.1.7、南非变体B.1.351、巴西变体P.1或加利福尼亚变体B.1.427/429)。

[0233] 在一些方面, 表达盒包含编码冠状病毒M蛋白、冠状病毒E蛋白和重组蛋白的核酸序列, 所述重组蛋白包含来自冠状病毒S蛋白的保守氨基酸序列和免疫原性氨基酸序列。

[0234] 在一些方面, M蛋白包含与SEQ ID NO:1至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的氨基酸序列。在一些方面, M蛋白包含SEQ ID NO:1。在一些方面, M蛋白是SEQ ID NO:1。

[0235] 在一些方面, 编码M蛋白的核酸序列与SEQ ID NO:2至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面, 编码M蛋白的核酸序列包含SEQ ID NO:2。在一些方面, 编码M蛋白的核酸序列是SEQ ID NO:2。

[0236] 在一些方面, E蛋白包含与SEQ ID NO:3至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的氨基酸序列。在一些方面, E蛋白包含SEQ ID NO:3。在一些方面, E蛋白是

[0237] SEQ ID NO:3。在一些方面, E蛋白包含用另一个氨基酸替换SEQ ID NO:3中的P71。在一些方面, SEQ ID NO:3中P71处的替换是P71L。

[0238] 在一些方面, 编码E蛋白的核酸序列与SEQ ID NO:4至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面, 编码E蛋白的核酸序列包含SEQ ID NO:4。在一些方面, 编码E蛋白的核酸序列是SEQ ID NO:4。在一些方面, 编码E蛋白的核酸序列包含

[0239] SEQ ID NO:4中核苷酸编号211-213处的脯氨酸密码子被另一氨基酸密码子替换。在一些方面, SEQ ID NO:4中核苷酸编号211-213处的脯氨酸密码子被亮氨酸密码子替换。

[0240] 在一些方面, 保守氨基酸序列来自S蛋白的S1亚基或S2亚基、S蛋白的RBD、S蛋白S2'切割位点和S蛋白的内部融合肽 (IFP) (本文中称为S2' IFP)、M蛋白或E蛋白。

[0241] 在一些方面, 所述保守氨基酸序列与任一SEQ ID NO:12-54至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面, 保守氨基酸序列包含任一SEQ ID NO:12-54。在一些方面, 保守氨基酸序列是任一SEQ ID NO:12-54。

[0242] 在一些方面, 所述保守氨基酸序列与SEQ ID NO:7至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面, 保守氨基酸序列包含SEQ ID NO:7。在一些方面, 保守氨基酸序列是SEQ ID NO:7。

[0243] 在一些方面, 编码重组蛋白的保守氨基酸序列的核酸序列与SEQ ID NO:8至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面, 编码重组蛋白的保守氨基酸序列的核酸序列包含SEQ ID NO:8。在一些方面, 编码重组蛋白的保守氨基酸序列的核酸序列是SEQ ID NO:8。

[0244] 在一些方面, 免疫原性氨基酸序列来自S蛋白受体结合结构域 (RBD)。

[0245] 在一些方面,所述免疫原性氨基酸序列与SEQ ID NO:11至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,免疫原性氨基酸序列包含SEQ ID NO:11。在一些方面,免疫原性氨基酸序列为SEQ ID NO:11。在一些方面,免疫原性蛋白质包含以下一个或多个替换:SEQ ID NO:11中的K88,L123,E155或N172用另一氨基酸替换。在一些方面,K88处的替换是K88N。在一些方面,K88处的替换是K88T。在一些方面,L123处的替换是L123R。在一些方面,E155处的替换是E155K。在一些方面,N172处的替换是N172Y。

[0246] 在一些方面,编码重组蛋白的免疫原性氨基酸序列的核酸序列与SEQ ID NO:101至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,编码重组蛋白的免疫原性氨基酸序列的核酸序列包含SEQ ID NO:101。在一些方面,编码重组蛋白的免疫原性氨基酸序列的核酸序列是SEQ ID NO:101。在一些方面,编码免疫原性蛋白的核酸序列包含包含以下一个或多个替换:SEQ ID NO:101的核苷酸编号262-264处的赖氨酸密码子用另一氨基酸的密码子替换、SEQ ID NO:101的核苷酸编号367-369处的亮氨酸密码子用另一氨基酸的密码子替换、SEQ ID NO:101的核苷酸编号463-465处的谷氨酸密码子用另一氨基酸的密码子替换或SEQ ID NO:101的核苷酸编号514-516处的天冬酰胺密码子用另一氨基酸的密码子替换。在一些方面,核苷酸编号262-264处的赖氨酸密码子用天冬酰胺或苏氨酸密码子替换。在一些方面,核苷酸编号367-369处的亮氨酸密码子用精氨酸密码子替换。在一些方面,核苷酸编号463-465处的谷氨酸密码子用赖氨酸密码子替换。在一些方面,核苷酸编号514-516处的天冬酰胺密码子用酪氨酸密码子替换。

[0247] 在一些方面,重组蛋白进一步包含来自S蛋白的跨膜(TM)结构域序列。

[0248] 在一些方面,TM结构域序列包含与SEQ ID NO:102至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的氨基酸序列。在一些方面,TM结构域序列包含SEQ ID NO:102。在一些方面,TM结构域序列是SEQ ID NO:102。

[0249] 在一些方面,编码重组蛋白的TM结构域序列的核酸序列与SEQ ID NO:103至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,编码重组蛋白的TM结构域序列的核酸序列包含SEQ ID NO:103。在一些方面,编码重组蛋白TM结构域的核酸序列是SEQ ID NO:103。

[0250] 在一些方面,重组蛋白包含来自S2' IFP的保守氨基酸序列、来自RBD的免疫原性氨基酸序列和S蛋白的TM结构域序列。

[0251] 在一些方面,重组蛋白排除了S蛋白中刺激包含非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。

[0252] 在一些方面,重组蛋白的氨基酸序列与SEQ ID NO:55至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,重组蛋白质的氨基酸序列包含SEQ ID NO:55。在一些方面,重组蛋白的氨基酸序列是SEQ ID NO:55。在一些方面,重组蛋白包含用另一个氨基酸替换SEQ ID NO:55的K88,L123,E155或N172中的一个或多个。在一些方面,K88的替换是K88N。在一些方面,K88的替换是K88T。在一些方面,L123的替换是L123R。在一些方面,E155的替换是

E155K。在一些方面，N172的替换是N172Y。

[0253] 在一些方面，编码重组蛋白的核酸序列与SEQ ID NO:56至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面，编码重组蛋白的核酸序列包含SEQ ID NO:56。在一些方面，编码重组蛋白核酸序列是SEQ ID NO:56。在一些方面，编码重组蛋白的核酸序列包含以下一个或多个替换：SEQ ID NO:56的核苷酸编号262-264处的赖氨酸密码子用另一氨基酸的密码子替换、SEQ ID NO:56的核苷酸编号367-369处的亮氨酸密码子用另一氨基酸的密码子替换、SEQ ID NO:56的核苷酸编号463-465处的谷氨酸密码子用另一氨基酸的密码子替换、或者SEQ ID NO:56的核苷酸编号514-516处的天冬酰胺密码子用另一氨基酸的密码子替换。在一些方面，核苷酸编号262-264处的赖氨酸密码子用天冬酰胺或苏氨酸的密码子替换。在一些方面，核苷酸编号367-369处的亮氨酸密码子被精氨酸密码子替换。在一些方面，核苷酸编号463-465处的谷氨酸密码子被赖氨酸密码子替换。在一些方面，核苷酸编号514-516处的天冬酰胺密码子被酪氨酸密码子替换。

[0254] 在一些方面，表达盒包含翻译为与SEQ ID NO:57至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的氨基酸序列的单个可读框。在一些方面，表达盒包含翻译为包含

[0255] SEQ ID NO:57的氨基酸序列的单个可读框。在一些方面，表达盒包含翻译为SEQ ID NO:57的氨基酸序列的单个可读框。在一些方面，表达盒包含翻译为这样的氨基酸序列的单个可读框，该氨基酸序列包含用另一氨基酸替换SEQ ID NO:57的P71, K423, L458, E490或N507中的一个或多个。在一些方面，在P71处的替换是P71L。在一些方面，K423处的替换是K423N。在一些方面，在K423处的替换是K423T。在一些方面，L458处的替换是L458R。在一些方面，E490的替换是E490K。在一些方面，N507处的替换是N507Y。

[0256] 在一些方面，表达盒包含与SEQ ID NO:58至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的单个可读框。在一些方面，表达盒包含含有SEQ ID NO:58的单个可读框。在一些方面，表达盒包含SEQ ID NO:58的单个可读框。在一些方面，表达盒包含含有以下一个或多个替换的单个可读框：SEQ ID NO:58的核苷酸编号211-213处的脯氨酸密码子用另一个氨基酸的密码子替换、SEQ ID NO:58的核苷酸编号1267-1269处的赖氨酸密码子用另一个氨基酸的密码子替换、SEQ ID NO:58的核苷酸编号1372-1374处的亮氨酸密码子用另一个氨基酸的密码子替换、SEQ ID NO:58的氨基酸编号1468-1470处的谷氨酸密码子用另一个氨基酸的密码子替换、或者SEQ ID NO:58的核苷酸编号1519-1521处的天冬酰胺密码子用另一个氨基酸的密码子替换。在一些方面，SEQ ID NO:58的核苷酸编号211-213处的脯氨酸密码子被亮氨酸密码子替换。在一些方面，核苷酸编号1267-1269处的赖氨酸密码子被天冬酰胺或苏氨酸密码子替换。在一些方面，核苷酸编号1372-1374处的亮氨酸密码子被精氨酸密码子替换。在一些方面，核苷酸编号1468-1470的谷氨酸密码子被赖氨酸密码子替换。在一些方面，核苷酸编号1519-1521处的天冬酰胺密码子被酪氨酸密码子替换。

[0257] 在一些方面，表达盒与任一SEQ ID NO:59-62的核酸序列至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面，表达盒包含任一SEQ ID NO:59-62的核酸序列。在一

些方面,表达盒是任一SEQ ID NO:59-62的核酸序列。

[0258] 在一些方面,重组蛋白能够刺激针对COVID-19的免疫应答。

[0259] 在一些方面,重组蛋白能够刺激Th1细胞介导的针对COVID-19的免疫应答。

[0260] 在一些方面,针对COVID-19的免疫应答针对WHCV和/或一个或多个变体,例如但不限于英国变体B.1.1.7、南非变体B.1.351、巴西变体P.1或加利福尼亚变体B.1.427/429。

[0261] 在一些方面,免疫应答对其他冠状病毒具有交叉反应。在一些方面,免疫应答对其他严重急性呼吸综合征冠状病毒和/或人类 β 冠状病毒具有交叉反应。

[0262] 在一些方面,无细菌序列载体还包含位于表达盒每侧的至少一个增强子序列。在一些方面,所述至少一个增强子序列是至少两个增强子序列。在一些方面,所述至少一个增强子序列是SV40增强子。

[0263] 在一些方面,无细菌序列载体包含环状共价闭合末端。

[0264] 在一些方面,无细菌序列载体包含线性共价闭合末端。在一些方面,无细菌序列载体是本文公开的msDNA。示例性msDNA的载体图如图4所示。

[0265] 在一些方面,无细菌序列载体与SEQ ID NO:104至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,无细菌序列载体包含SEQ ID NO:104。在一些方面,无细菌序列载体为SEQ ID NO:104。

[0266] III.VLP

[0267] 在一些方面,本文公开的VLP由本文所述的表达载体的表达盒和/或无细菌序列载体的表达盒产生。

[0268] 本文提供了一种重组细胞,其包含本文所述的表达载体或无细菌序列载体。

[0269] 在一些方面,重组细胞是酵母、细菌、古细菌、真菌、昆虫或动物细胞,包括哺乳动物细胞。在一些方面,重组细胞包括黑腹果蝇 (*Drosophila melanogaster*) 细胞、酿酒酵母 (*Saccharomyces cerevisiae*) 或其他酵母、大肠杆菌、枯草芽孢杆菌 (*Bacillus subtilis*)、Sf9细胞、C129细胞、HEK293细胞、链孢霉属 (*Neurospora*)、BHK、CHO、COS、HeLa细胞、Hep G2细胞和人细胞以及细胞系。

[0270] 在一些方面,表达载体用于在人细胞或细胞系中表达,所述载体如图2所示的示例性载体。

[0271] 在一些方面,表达载体是杆状病毒载体,如图5所示的示例性载体,且细胞类型是昆虫细胞(例如Sf9细胞)。

[0272] 在一些方面,本公开涉及一种生产VLP的方法,包括在合适的条件下培养包含表达载体或无细菌序列载体的重组细胞,以从表达载体或无细菌序列载体生产VLP。

[0273] 在一些方面,生产VLP的方法还包括分离该VLP。在一些方面,VLP由任何上述表达载体或任何上述无细菌序列载体产生,其中病毒是冠状病毒。

[0274] 在一些方面,从细胞裂解物中分离VLP。

[0275] 在一些方面,分离是通过亲和纯化。在一些方面,亲和纯化包括微流体法和/或层析法。

[0276] 在一些方面,亲和纯化包含血管紧张素转换酶2 (ACE2) 受体肽或抗S蛋白单克隆抗体。

[0277] 在一些方面,ACE2受体肽包含与SEQ ID NO:70至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的氨基酸序列。在一些方面,ACE2受体肽包含SEQ ID NO:70。在一些方面,ACE2受体肽是SEQ ID NO:70。

[0278] 在一些方面,ACE2受体肽在肽的C端或N端包含生物素受体肽(BAP)标签。在一些方面,所述BAP标签包含与SEQ ID NO:71至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的氨基酸序列。在一些方面,所述BAP标签包含SEQ ID NO:71。在一些方面,所述BAP标签是SEQ ID NO:71。

[0279] 在一些方面,ACE2受体肽或抗S蛋白单克隆抗体是生物素化的并固定在链霉亲和素包被的珠上。在一些方面,亲和纯化包括微流体法和/或层析法。

[0280] 在一些方面,本公开涉及通过该方法产生的VLP。

[0281] 本文提供了一种VLP,其包含融合至免疫原性氨基酸序列的来自病毒的保守氨基酸序列的重组蛋白。

[0282] 在一些方面,免疫原性氨基酸序列和保守氨基酸序列来自同一病毒。

[0283] 在一些方面,保守氨基酸序列来自病毒糖蛋白。在一些方面,免疫原性氨基酸序列来自相同的病毒糖蛋白。

[0284] 在一些方面,VLP还包含病毒包膜蛋白和/或病毒基质蛋白。在一些方面,病毒包膜蛋白和/或病毒基质蛋白来自与保守氨基酸序列相同的病毒。

[0285] 在一些方面,保守氨基酸序列、免疫原性氨基酸序列、病毒包膜蛋白和/或病毒基质蛋白是共有序列。

[0286] 在一些方面,重组蛋白能够刺激针对病毒的包含中和抗体的免疫应答。

[0287] 在一些方面,重组蛋白能够刺激针对病毒的Th1细胞介导的免疫应答。

[0288] 在一些方面,免疫应答对相关病毒或毒株具有交叉反应。

[0289] 在一些方面,重组蛋白排除了病毒中刺激包含非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。

[0290] 在一些方面,病毒是冠状病毒、流感病毒、人类免疫缺陷病毒、人类乳头瘤病毒、肝炎病毒或溶瘤病毒。

[0291] 在一些方面,流感病毒是甲型流感病毒。在一些方面,甲型流感病毒是H1N1,H5N1或H3N2。

[0292] 在一些方面,流感病毒是乙型流感病毒。

[0293] 在一些方面,冠状病毒是人类冠状病毒,例如但不限于HCoV-229E,HCoV-NL63,HCoV-OC43,HCoV-HKU1,SARS-CoV-1,SARS-CoV-2(即,COVID-19)和/或MERS-CoV。

[0294] 在一些方面,冠状病毒是COVID-19(即WHCV或其变体,例如但不限于英国变体B.1.1.7、南非变体B.1.351、巴西变体P.1或加利福尼亚变体B.1.427/429)。

[0295] 在一些方面,VLP包含冠状病毒M蛋白、冠状病毒E蛋白和重组蛋白,所述重组蛋白包含来自冠状病毒S蛋白的保守氨基酸序列和免疫原性氨基酸序列。

[0296] 在一些方面,M蛋白包含与SEQ ID NO:1至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同

一的氨基酸序列。在一些方面,M蛋白包含SEQ ID NO:1。在一些方面,M蛋白是SEQ ID NO:1。
[0297] 在一些方面,E蛋白包含与SEQ ID NO:3至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的氨基酸序列。在一些方面,E蛋白包含SEQ ID NO:3。在一些方面,E蛋白是SEQ ID NO:3。在一些方面,E蛋白包含用另一个氨基酸替换SEQ ID NO:3中的P71。在一些方面,SEQ ID NO:3中P71处的替换是P71L。

[0298] 在一些方面,保守氨基酸序列来自S蛋白的S1亚基或S2亚基、S蛋白的RBD、S蛋白S2'切割位点和S蛋白的内部融合肽(IFP)、M蛋白或E蛋白。

[0299] 在一些方面,所述保守氨基酸序列与任一SEQ ID NO:12-54至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,保守氨基酸序列包含任一SEQ ID NO:12-54。在一些方面,保守氨基酸序列是任一SEQ ID NO:12-54。

[0300] 在一些方面,所述保守氨基酸序列与SEQ ID NO:7至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,保守氨基酸序列包含SEQ ID NO:7。在一些方面,保守氨基酸序列是SEQ ID NO:7。

[0301] 在一些方面,免疫原性氨基酸序列来自S蛋白受体结合结构域(RBD)。

[0302] 在一些方面,所述免疫原性氨基酸序列与SEQ ID NO:11至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,免疫原性氨基酸序列包含SEQ ID NO:11。在一些方面,免疫原性氨基酸序列为SEQ ID NO:11。在一些方面,免疫原性蛋白质包含以下一个或多个替换:SEQ ID NO:11中的K88,L123,E155或N172用另一氨基酸替换。在一些方面,K88处的替换是K88N。在一些方面,K88处的替换是K88T。在一些方面,L123处的替换是L123R。在一些方面,E155处的替换是E155K。在一些方面,N172处的替换是N172Y。

[0303] 在一些方面,重组蛋白进一步包含来自S蛋白的跨膜(TM)结构域序列。

[0304] 在一些方面,TM结构域序列包含与SEQ ID NO:102至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的氨基酸序列。在一些方面,TM结构域序列包含SEQ ID NO:102。在一些方面,TM结构域序列是SEQ ID NO:102。

[0305] 在一些方面,重组蛋白包含来自S2' IFP的保守氨基酸序列、来自RBD的免疫原性氨基酸序列和S蛋白的TM结构域序列。

[0306] 在一些方面,重组蛋白排除了S蛋白中刺激包含非中和抗体的免疫应答和/或刺激Th2细胞介导的免疫应答的氨基酸序列。

[0307] 在一些方面,重组蛋白的氨基酸序列与SEQ ID NO:55至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一。在一些方面,重组蛋白质的氨基酸序列包含SEQ ID NO:55。在一些方面,重组蛋白的氨基酸序列是SEQ ID NO:55。在一些方面,重组蛋白包含用另一个氨基酸替换SEQ ID NO:55的K88,L123,E155或N172中的一个或多个。在一些方面,K88的替换是K88N。在一些方面,K88的替换是K88T。在一些方面,L123的替换是L123R。在一些方面,E155的替换是

E155K。在一些方面,N172的替换是N172Y。

[0308] 本文提供了一种VLP,其包含与SEQ ID NO:55至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的重组蛋白、与SEQ ID NO:1至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的M蛋白和与SEQ ID NO:3至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%

[0309] 同一的E蛋白。

[0310] 本文提供了一种VLP,其包含含有SEQ ID NO:55的重组蛋白、含有SEQ ID NO:1的M蛋白和含有SEQ ID:3的E蛋白。

[0311] 本文提供了一种VLP,其包含SEQ ID NO:55的重组蛋白、SEQ ID NO:1的M蛋白和SEQ ID NO:3的E蛋白。

[0312] 在一些方面,重组蛋白能够刺激针对COVID-19的免疫应答。

[0313] 在一些方面,重组蛋白能够刺激Th1细胞介导的针对COVID-19的免疫应答。

[0314] 在一些方面,针对COVID-19的免疫应答针对WHCV和/或一个或多个变体,例如但不限于英国变体B.1.1.7、南非变体B.1.351、巴西变体P.1或加利福尼亚变体B.1.427/429。

[0315] 在一些方面,免疫应答对其他冠状病毒具有交叉反应。

[0316] 在一些方面,免疫应答对其他严重急性呼吸综合征冠状病毒和/或人类β冠状病毒具有交叉反应。

[0317] IV.组合物

[0318] 本文提供的组合物包含本文所述的任何表达载体、无细菌序列载体或VLP。

[0319] 在一些方面,组合物进一步包含生理可接受的载体、赋形剂或稳定剂。例如,参见Remington:The Science and Practice of Pharmacy,第22版(2013)。可接受的载体、赋形剂或稳定剂可以包括对受试者无毒的那些。在一些方面,组合物是无菌的或组合物的一种或多种组分是无菌的。例如,可以通过过滤(例如,通过无菌滤膜)或通过辐照(例如,γ射线辐照)制备无菌组分。

[0320] 本发明的赋形剂当添加到药物组合物中时,可以被描述为“药学上可接受的”赋形剂,这意味着该赋形剂是一种化合物、材料、组合物、盐和/或剂型,在合理的医学判断中,其适于在以合理的益处/风险比相称的预期接触期间与人和动物的组织接触而没有过度毒性、刺激、过敏反应或其他有问题的并发症。在一些方面,术语“药学上可接受的”是指经联邦或州政府监管机构批准,或列入《美国药典》或其他公认的国际药典中,用于动物,尤其是人类。可以使用多种赋形剂。在一些方面,赋形剂可以是但不限于碱性剂、稳定剂、抗氧化剂、粘接剂、分离剂、包衣剂、外相组分、控释组分、溶剂、表面活性剂、保湿剂、缓冲剂、填料、润滑剂或其组合。除本文所讨论的赋形剂外,赋形剂还可包括(但不限于)Remington:The Science and Practice of Pharmacy,第22版(2013)中所列赋形剂。将赋形剂纳入本文的特定分类(例如“溶剂”)旨在说明而不是限制赋形剂的作用。特定赋形剂可分为多种类型。

[0321] 本公开的药物组合物被配制成与其预期给药途径施用。示例性施用途包括肠内、局部、肠外、口服、肺、鼻内、静脉内、表皮、经皮、皮下、肌内或腹膜内施用或吸入。本文所用的“肠胃外施用”是指除肠内和局部施用以外的施用方式,通常通过注射或输注,包括但

不限于静脉内、肌内、动脉内、鞘内、淋巴管内、病变内、囊内、眶内、心内、皮内、腹膜内、经气管、皮下、表皮下、关节内、包膜下、蛛网膜下腔、椎管内、硬膜外、胸膜内和胸骨内注射和输注,以及体内电穿孔。在一些方面,制剂通过非肠胃外途径施用,在一些方面,口服制剂。其他非肠胃外途径包括局部、表皮或粘膜施用途径,例如鼻内、阴道、直肠、舌下或局部施用。

[0322] 在一些方面,药物组合物是冻干的。

[0323] 本领域已知多种方法,适用于将核酸引入细胞。实例包括但不限于电穿孔、磷酸钙介导的转移、核转染、声穿孔、热休克、磁转染、脂质体介导的转运、微注射、微粒子介导的转移(纳米粒子)、阳离子聚合物介导的转移(DEAE-葡聚糖、聚乙烯亚胺、聚乙二醇(PEG)等)或细胞融合。

[0324] 已经研究了纳米粒子载体,如脂质体、胶束和聚合物纳米粒子,以通过各种机制,例如增强的渗透性和保留(EPR)效应,提高治疗剂的生物利用度和药代动力学特性。

[0325] 通过将靶向性配体缀合到纳米粒子上以实现靶向细胞的选择性递送,可以实现进一步的改进。例如,受体靶向的纳米粒子递送已被证明可以改善体外和体内的治疗反应。已研究的靶向性配体包括叶酸、转铁蛋白、抗体、肽和适体。此外,可以将多种功能结合到纳米粒子的设计中,例如,以能够成像和以触发细胞内药物释放。

[0326] 在一些方面,组合物还包含递送剂。在一些方面,递送剂是纳米粒子。在一些方面,递送剂选自脂质体、非脂质聚合物分子、内体和其任何组合。

[0327] 在一些方面,递送剂(例如,纳米粒子)包含靶向性配体。

[0328] 在一些方面,靶向性配体包含与ACE2受体具有结合亲和力的S蛋白肽(例如,用于递送表达载体、无细菌序列载体或包含冠状病毒序列的VLP)。

[0329] 在一些方面,S蛋白肽来自S蛋白的保守区。在一些方面,S蛋白肽的长度为3个氨基酸至100个氨基酸,包括其中的任何长度或长度范围,例如3个氨基酸到90、80、70、60、50、40、30、20或10个氨基酸。

[0330] 在一些方面,S蛋白肽包含与任一SEQ ID NO:76-99至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一的氨基酸序列。在一些方面,S蛋白肽包含任一SEQ ID NO:76-99。在一些方面,S蛋白肽是任一SEQ ID NO:76-99。

[0331] V. 治疗用途和方法

[0332] 如本文所述的表达载体、无细菌序列载体(例如msDNA)、VLP和组合物可用于对需要其的受试者进行预防或治疗,包括作为针对病毒感染(例如,冠状病毒感染,例如COVID-19)的疫苗或作为对感染病毒的个体的治疗。

[0333] 本文提供了一种用于病毒感染的疫苗,其包含如本文所述的表达载体、无细菌序列载体、VLP或组合物。

[0334] 本文提供了一种治疗受试者病毒感染的方法,包括向受试者施用本文所述的表达载体、无细菌序列载体、VLP或组合物,其中所述表达载体或无细菌序列的载体在受试者中的细胞内表达产生VLP。

[0335] 本文提供了一种本文所述的用于治疗受试者病毒感染的表达载体、无细菌序列载体、VLP或组合物,其中所述表达载体或无细菌序列载体在受试者中的细胞内表达产生VLP。

[0336] 本文提供了一种表达载体、无细菌序列载体、VLP或组合物在治疗受试者病毒感染

中的用途,其中所述表达载体或无细菌序列载体在受试者中的细胞内表达产生VLP。

[0337] 本文提供了表达载体、无细菌序列载体、VLP或组合物在制备用于治疗受试者病毒感染的药物中的用途,其中所述表达载体或无细菌序列载体在受试者中的细胞内表达产生VLP。

[0338] 表达载体、无细菌序列载体或组合物可通过对治疗病毒感染有效的任何施用途径施用至受试者。

[0339] 在一些方面,施用是通过肠内、局部、肠胃外、口服、肺、鼻内、静脉内、表皮、经皮、皮下、肌肉或腹膜内施用或吸入。

[0340] 在一些方面,施用是通过肠胃外或非肠胃外施用。

[0341] 在一些方面,通过注射或输注进行肠胃外施用。

[0342] 在一些方面,肠胃外施用是通过静脉内、肌肉、动脉内、鞘内、淋巴管内、病灶内、囊内、眶内、心内、皮内、腹膜内、经气管、皮下、表皮下、关节内、囊下、蛛网膜下腔、椎管内、硬膜外、胸膜内或胸骨内注射或输注,或通过体内电穿孔。

[0343] 在一些方面,非肠胃外施用是口服、局部、表皮、粘膜、鼻内、阴道、直肠或舌下施用。

[0344] 在一些方面,施用是通过口服、肺、鼻内、静脉内、表皮、经皮、皮下、肌肉或腹膜内施用或吸入施用。

[0345] 在一些方面,施用是通过病毒感染和传播途径施用。

[0346] 在一些方面,病毒感染和传播的途径是粘膜。

[0347] 在一些方面,对于呼吸道感染,施用是通过口服、鼻腔或肺施用。在一些方面,施用是通过鼻施用。

[0348] 除了如本文所述的将VLP高效、靶向性和无创地递送到下呼吸道组织之外,应用吸入和鼻内施用途径提供了一个强大的机会来通过肺和鼻咽相关淋巴组织(NALT)产生支持性免疫应答。

[0349] 在一些方面,针对性传播感染,施用是阴道施用。

[0350] 在一些方面,施用是通过肌肉、皮下或皮内施用,其中注射部位和深度都会影响免疫应答。肌肉注射为疫苗施用提供了一种强有力的备选和常用的技术,特别因为它是被认可的且易于再次施用。

[0351] 可以实施例如,一次、多次和/或经一个或多个延长期施用。在一些方面,施用是一次、两次(例如,第一施用之后约1周、约2周、约3周、约4周或更多周后的第二次施用)、约每周一次、约每月一次、约每2个月一次、约每3个月一次、约每4个月一次、约每6个月一次、每年一次或每十年一次。

[0352] 本文所述的表达盒提供了一种VLP,该VLP赋予强健的体液免疫应答,并具有DNA疫苗对细胞内病原体表位内部加工的益处,用于T细胞呈递和细胞介导的免疫。在一些方面,成功地将免疫优势赋予了重组蛋白的保守氨基酸序列,且疫苗产生了通用的冠状病毒免疫。

[0353] 在一些方面,从表达盒的翻译产物(无论是从本文所述的表达载体还是无细菌序列载体)细胞内自组装的VLP产生Th1细胞介导的应答,如以下所示:1) MHC-I环境,以启动针对病毒感染细胞的特异性细胞毒T细胞活性;2) 吞噬性抗原呈递细胞(APC)中的MHC-II环

境,用于补充体液和细胞介导的支持。

[0354] 在一些方面,从如本文所述的表达盒细胞内组装VLP消除了潜在的疫苗介导的TH2免疫病理学和辅助治疗的任何相关要求。

[0355] 在一些方面,VLP刺激受试者的免疫应答,包括针对病毒感染的中和抗体。

[0356] 在一些方面,VLP刺激受试者对病毒感染的Th1细胞介导的免疫应答。

[0357] 在一些方面,免疫应答对相关病毒或毒株具有交叉反应。

[0358] 在一些方面,VLP不刺激受试者中包含非中和抗体的免疫应答和/或不刺激受试者的Th2细胞介导的免疫应答。

[0359] 在一些方面,VLP诱导抗体阻断病毒受体结合、病毒基因组脱衣和/或基因组注入。

[0360] 在一些方面,VLP与感染病毒交叉竞争对病毒受体的结合。

[0361] 在一些方面,VLP与相关病毒或毒株交叉竞争对病毒受体的结合。

[0362] 在一些方面,病毒感染是冠状病毒、流感病毒、人类免疫缺陷病毒、人类乳头瘤病毒、肝炎病毒或溶瘤病毒感染。

[0363] 在一些方面,流感病毒是甲型流感病毒。在一些方面,甲型流感病毒是H1N1,H5N1或H3N2。

[0364] 在一些方面,流感病毒是乙型流感病毒。

[0365] 在一些方面,冠状病毒是人类冠状病毒,例如但不限于HCoV-229E,HCoV-NL63,HCoV-OC43,HCoV-HKU1,SARS-CoV-1,SARS-CoV-2(即,COVID-19)和/或MERS-CoV。

[0366] 在一些方面,冠状病毒是COVID-19(即WHCV或其变体,例如但不限于英国变体B.1.1.7、南非变体B.1.351、巴西变体P.1或加利福尼亚变体B.1.427/429)。

[0367] 在一些方面,VLP刺激受试者的免疫应答,包括针对COVID-19的中和抗体。

[0368] 在一些方面,VLP刺激受试者针对COVID-19的Th1细胞介导的免疫应答。

[0369] 在一些方面,针对COVID-19的免疫应答针对WHCV和/或一个或多个变体,例如但不限于英国变体B.1.1.7、南非变体B.1.351、巴西变体P.1或加利福尼亚变体B.1.427/429。

[0370] 在一些方面,免疫应答对其他冠状病毒具有交叉反应。

[0371] 在一些方面,免疫应答对其他严重急性呼吸综合征冠状病毒和/或人类 β 冠状病毒具有交叉反应。

[0372] 在一些方面,VLP不刺激受试者中包含非中和抗体的免疫应答和/或不刺激受试者的Th2细胞介导的免疫应答。

[0373] 在一些方面,施用是通过吸入施用。

[0374] COVID-19和许多其他冠状病毒的细胞配体是在人类下呼吸道中发现的ACE2受体,它调节跨物种和人与人之间的传播。ACE2受体与冠状病毒表面的S糖蛋白结合,融合后在双膜囊泡中形成复制转录复合物(Letko等人,Nat.Microbiol.5(4):562-569(2020);Wan等人,J.Virol.4(7)e00127-20(2020))。亚基因组RNA嵌套集合的连续复制和合成编码辅助蛋白和结构蛋白,使病毒颗粒芽出。这导致含有病毒子的囊泡与质膜融合,最终将病毒释放到宿主体内(Fehr和Perlman)。使用肾上腺素能阻滞剂(β -阻滞剂)控制血压的高血压患者特别容易感染,因为 β -阻滞剂刺激呼吸道中ACE2受体过量表达,由此促进病毒结合和感染。在如COPD、糖尿病和心血管疾病等潜在疾病的患者中也注意到了易感性(Guan等人,Eur.Resp.Journal,2000547;DOI:10.1183/13993003.00547-2020(2020))。

[0375] 在一些方面,如本文所述的抗冠状病毒(例如COVID-19)的VLP不仅递送治疗性DNA疫苗,而且还竞争呼吸组织中可用的冠状病毒受体位点,减轻进一步感染。

[0376] 在一些方面,从细胞中排出的功能性(表达表面RBD的)VLP进一步促进了对靶细胞上可用ACE2受体的竞争性干扰,并促进了与B细胞的相互作用,以确保稳健的中和体液应答。

[0377] 在一些方面,用于呈递的S2' IFP结构域暴露了高度保守的位点,并通过半抗原载体应答赋予该决定簇免疫优势。

[0378] 在一些方面,VLP与COVID-19交叉竞争结合ACE2受体、神经毡蛋白-1(neuropilin-1)和/或其他受体。

[0379] 在一些方面,VLP与其他冠状病毒交叉竞争结合ACE2受体、神经毡蛋白-1和/或其他受体。

[0380] 在一些方面,VLP与其他严重急性呼吸综合征冠状病毒和/或人类β冠状病毒交叉竞争结合ACE2受体、神经毡蛋白-1和/或其他受体。

[0381] 提供以下实施例仅供说明而非限制。

[0382] 实施例1

[0383] A.表达载体的产生,用于产生无细菌序列载体和VLP

[0384] 通过将COVID-19衍生的序列克隆到微链表达载体(Medipage Bioceuticals, Inc., Toronto, CA)中两个特化超序列(SS)位点之间的多克隆位点,如美国专利号9,290,778和9,862,954所述,产生四个表达载体,所述美国专利通过引用将其全部并入本文作为参考。

[0385] 衍生自COVID-19的序列包含编码包膜(E)蛋白的序列(GenBank登录号QHD43418.1;SEQ ID NO:3)和膜(M)蛋白的序列(GenBank登录号QHD43419.1;SEQ ID NO:1)。此外,产生了编码重组刺突(S)蛋白的序列,该序列包含与COVID-19S蛋白(GenBank登录号

[0386] QHD43416.1;SEQ ID NO:5)的受体结合域(RBD)、S2'切割位点和内部融合肽(S2' IFP)以及跨膜(TM)结构域相关的序列的融合(RBD::S2' IFP::TM;SEQ ID NO:55)。重组S蛋白被工程化,以从S蛋白中排除刺激包含非中和抗体的免疫应答的氨基酸序列,并排除刺激Th2细胞介导的免疫应答的氨基酸序列。

[0387] 三个表达载体的包含E蛋白、M蛋白和重组S蛋白的表达盒通过编码来自猪捷申病毒-1 2A的自切割肽P2A的序列融合到在巨细胞病毒(CMV)启动子控制下的单个多核苷酸(SEQ ID NO:58)中。图1展示了一个示例性表达盒。

[0388] 三个表达载体之一含有表达盒"CMV-E-P2A-M-P2A-RBD::S2' IFP::TM-bGHPolyA"

[0389] (SEQ ID NO:60),其含有牛生长激素(bGH)聚腺苷酸化(polyA)信号。含有表达盒的表达载体的图谱如图2所示(pGL2-SS-CMV-VLP-BGH-SS,SEQ ID NO:63)。

[0390] 三个表达载体中的另一表达载体含有表达盒"CMV-E-P2A-M-P2A-RBD::S2' IFP::TM-SV40polyA"(SEQ ID NO:59),其含有猴病毒40(SV40)polyA。

[0391] 三个表达载体中的另一表达载体含有表达盒"CMV-E-P2A-M-P2A-RBD::S2' IFP::TM-T2A-GFP-SV40polyA"(SEQ ID NO:61),其包含通过编码来自thossea asigna病毒2A的自切割肽T2A融合到COVID-19序列的绿色荧光蛋白(GFP)和SV40 polyA的序列。

[0392] 第四个表达载体包含表达盒“CMV-E-P2A-M-T2A-MCS-bGHpolyA” (SEQ ID NO:62), 该表达盒包含单个多核苷酸, 其中E蛋白和M蛋白通过编码P2A的序列彼此融合, 然后通过编码T2A的序列与多克隆位点 (MCS) 融合。该表达盒还包含CMV启动子和bGH polyA。MCS用于插入额外序列, 例如包含本文公开的保守序列和免疫原性序列的重组蛋白。

[0393] 含有SEQ ID NO:59-62的表达盒的表达载体与图2和SEQ ID NO:63的表达载体相同, 只是表达盒不同。

[0394] B.COVID-19基因的表达

[0395] 用1μg图2所示表达载体或无表达载体电穿孔人肺A549细胞 (1×10^6)。电穿孔48小时后提取总RNA并转化为cDNA文库。将1μL cDNA用作E、M和RBD::S2' IFP::TM转基因的实时qRT-PCR模板, 分别使用如下表1所示的E、M和RBD的基因特异性引物。将转基因的表达归一化为β-肌动蛋白表达。

[0396] 表1. 引物序列

基因	正向引物序列	反向引物序列
E	ACTGCTGCAACATCGTGAAC (SEQ ID NO:64)	TGCTAGAATTCAGGTTCTTC ACC (SEQ ID NO:65)
M	TTCCTGTGGCTGCTGTGG (SEQ ID NO:66)	ATGACCAGCTCGCTTTCCAG (SEQ ID NO:67)
RBD::S2'IFP ::TM	ATCAGCACAGAGATCTACCAGG (SEQ ID NO:68)	AGCACCACCACTCTGTAAGG (SEQ ID NO:69)

[0398] 如图3 (A) 所示, 在用表达载体电穿孔的细胞 (“VLP”) 的cDNA文库中检测到每个转基因, 但在对照细胞 (“CTL”) 的cDNA库中未检测到。图中所示的相对基因表达通过 $\Delta \Delta$ CT法计算。使用单因素方差分析ANOVA进行统计分析 (**=p<0.001, ***=p<0.0001)。

[0399] C. 重组刺突蛋白的表达

[0400] 将2μg图2的表达载体使用 Lipofectamine®3000 Reagent (Invitrogen) 转染HEK 293细胞 (1×10^6)。转染后48小时收集蛋白质样品。通过从转染细胞和未转染的对照细胞加载50μg全蛋白裂解物来制备蛋白质印迹物。兔多克隆抗RBD抗体用于检测重组S蛋白, 而兔多克隆抗β-肌动蛋白抗体用于检测β-肌动蛋白作为上样对照。抗兔辣根过氧化物酶 (HRP) 抗体和化学发光成像用于信号检测。代表性的蛋白质印迹如图3 (B) 所示, 显示从用表达载体转染的细胞 (“VLP”) 分离的蛋白质中检测到重组S蛋白, 但在从对照细胞分离的蛋白质中未检测到。

[0401] 转染细胞和对照细胞中重组S蛋白表达的相对平均蛋白强度通过蛋白质印迹图像的密度分析测定 (n=3)。见图3 (C)。

[0402] 实施例2

[0403] VLP表达载体刺激抗体产生

[0404] 将图2的表达载体包封在脂质纳米粒子 (Entos Pharmaceuticals) 中, 并在第0天通过肌肉注射以100μg的剂量施用给C57小鼠, 然后在第14天通过肌肉注射以100μg的加强剂量施用。在49天中每7天通过尾静脉收集血清。

[0405] 由间接ELISA通过结合纯化的S1蛋白 (Abclonal, Inc.) 评估小鼠血清中的抗体浓度。

[0406] 将血清在PBS中稀释至1%, 然后添加到含有S1蛋白的ELISA板中。通过抗小鼠IgG SULFO-TAG™缀合抗体 (Meso Scale Diagnostics, LLC) 检测结合S1蛋白的小鼠血清抗体。

[0407] 抗体浓度如图5A和5B所示。浓度在第21天达到峰值, 为约5000ng/mL, 至第49天保持一致表达, 为约3000ng/mL。

[0408] 实施例3

[0409] A. COVID-19基因组序列保守性的表征

[0410] 共有3928个代表性的完整COVID-19基因组从GISAID数据库下载 (<https://www.gisaid.org>)。基因组的收集日期为2019年12月至2021年2月, 包含所有主要变异株以及参考基因组 (NC_045512.2)。使用MAFFT多序列比对程序将基因组与参考基因组进行比对。进行序列保守性和核苷酸频率分析。

[0411] 图6显示了3928个代表性COVID-19基因组的序列保守性分析。(A) 水平轨迹表示根据参考基因组的所有COVID-19基因 (在y轴上描绘) 的基因组位置 (在x轴上指示)。(B) 柱状图中的柱高度对应于在每个给定基因组位置上与参考基因组不同的基因组百分数。柱状图和直方图是使用ggplot2软件包以R版本3.6.1生成的。

[0412] 如图6所示, COVID-19基因组具有相对高水平的序列保守性, 很少有关键的基因组变异。忽略可变的5'和3'末端区域, 仅发现三个基因组位置在>50%的序列中与参考基因组不同。这些单核苷酸多态性 (SNP) 中的两个单核苷酸多态性在ORF1ab中发现 (第一个 (C241T) 在基因间区, 第二个 (C14408T->L4715) 在编码区, 第三个 (D614G) 在刺突 (S) 蛋白中。

[0413] B. 人β冠状病毒基因组序列保守性的表征

[0414] 除了本实施例A部分讨论的3928个代表性完整COVID-19基因组外, 还从NCBI

[0415] GenBank®数据库下载了120个SARS CoV (导致SARS的病毒) 基因组和257个MERS-CoV (造成MERS的病毒) 基因组。使用MAFFT多序列比对程序将基因组与COVID-19参考基因组比对。由于这三种病毒基因组的基因组结构相似, 因此可以进行比较。进行序列保守性和核苷酸频率分析。

[0416] 图7显示了直方图, 其中柱高度对应于在每个给定基因组位置上与参考基因组不同的基因组百分数。该直方图是使用ggplot2软件包以R版本3.6.1生成的。

[0417] 如图7所示, 与COVID-19基因组相比, 其他著名的人β冠状病毒 (SARS-CoV和MERS-CoV) 的基因组也具有相对高水平的序列保守性。

[0418] C. 显著变异COVID-19株的功能相关突变的鉴定

[0419] 对A部分中讨论的3928个COVID-19序列筛选属于关键变异株 (英国变异株B.1.1.7 (n=233)、南非变异株B.1.351 (n=104)、巴西变异株P.1 (n=39) 和加利福尼亚变异株B.1.427/429 (n=62)) 的序列。使用MAFFT多序列比对程序, 将四个变异株的基因组与SARS-CoV-2参考基因组 (NC_045512.2) 独立比对。进行序列保守性和核苷酸频率分析。通过评估BLOSUM 62矩阵评分、表面暴露分析 (通过PyMol) 和文献综述确定功能重要性。

[0420] 图8A-D显示了直方图, 其中柱高度对应于在每个给定的基因组位置上与参考基因组不同的变异基因组的百分数 (图8A中的B.1.1.7、图8B中的B.1.351、图8C中的P.1和图8D

中的B.1.427/429)。该直方图是使用ggplot2软件包以R版本3.6.1生成的。

[0421] 表2显示了从位于图1所示表达盒中包含的COVID-19基因组区域的变异COVID-19株中鉴定的SNP的总结。

[0422] 表2. 鉴定的SNP的总结

[0423]	表达盒序列	COVID-19 变体			
		英国 (B.1.1.7)	南非(B.1.357)	巴西(P.1)	加利福尼亚 (B.1.427/429)
[0424]	RBD	AAT > TAT → N501Y	AAT > TAT → N501Y GAA > AAA → E484K AAG > AAT → K417N	AAT > TAT → N501Y GAA > AAA → E484K AAG > ACG → K417T	CTG > CGG → L452R
	S2'IFP	-	-	-	-
	E	-	CCT > CTT → P71L	-	-
	M	-	-	-	TTC > TTT → F53F

[0425] 在变异COVID-19株的刺突(S)蛋白的受体结合结构域(RBD)区域中鉴定的SNP被映射到参考蛋白质数据库(PDB)结构(PDB ID:6VXX)上,以评估表面暴露。N501、K417和L452残基被确定为表面暴露的,因此具有潜在更大的后果。确定E484残基未暴露于表面。

[0426] 变异COVID-19株包膜(E)蛋白中鉴定的SNP表面暴露是通过Bianchi等人,BioMed Research International, <https://doi.org/10.1155/2020/4389089> (2020) 的结构信息评估的。P71残基被确定为表面暴露,因此具有潜在更大的后果。

[0427] 膜(M)蛋白中鉴定的SNP导致同义突变,因此未进行功能分析。

[0428] 总的来说,分析表明,为图1所示的VLP表达盒选择的序列对COVID-19变异相对稳健,尤其是S2' IFP位点,该位点在所有关键变异株以及其他冠状病毒(SARS-CoV和MERS-CoV)中完全保守。

[0429] 实施例4

[0430] A. 生产VLP的无细菌序列载体的产生

[0431] 根据美国专利号9,290,778和9,862,954中描述的方法,从实施例1中描述的表达载体在可诱导的大肠杆菌细胞中产生用于生产VLP的DNA微链(msDNA-VLP)。

[0432] 纯化并浓缩msDNA-VLP,并对纯度和序列进行质量控制测试。

[0433] B. 无细菌序列载体与纳米粒子的复合

[0434] 将纯化的msDNA-VLP和表达标记蛋白(例如GFP)的对照msDNA(msDNA对照)与纳米粒子(例如脂质纳米粒子(LNP))复合。在其他研究中,商业LNP已证明msDNA在体内肺中的强

转染效率(未公开的数据)。商业LNP用作体外对照。商业JetPEI

[0435] (<https://www.polyplus-transfection.com/products/cgmp-grade-in-vivo-jetpei/>) 用作体内对照。

[0436] 将msDNA纳米粒子冻干,用于体外和体内试验。

[0437] C. 无细菌序列载体的体外VLP形成和免疫应答

[0438] 将msDNA纳米粒子(即,如本实施例B部分所述)以及裸msDNA(即,本实施例A部分所述的msDNA-VLP和msDNA对照,其未与纳米粒子复合)递送至表达ACE2受体的人细胞系(例如,A549细胞(ATCC CCL-185))、血管内皮细胞、或肺泡上皮细胞(Yen, T.-T. 等人, *Journal of Virology* 80(6):2684-2693 (2006); Qian, Z. 等人, *American Journal of Respiratory Cell and Molecular Biology* 48(6):742-748 (2013))。评估递送效率和平均荧光。

[0439] 通过透射电子显微镜评估细胞内VLP形成。

[0440] 使用免疫测定技术在细胞培养物中评估细胞因子风暴和炎症反应的过度活性。

[0441] D. 在真核表达系统中体外生产VLP

[0442] 产生了一种真核表达载体,其包含用于在真核细胞中产生VLP的启动子控制下的M-P2A-E和RBD::S2'::TM。用于在Sf9细胞中生产VLP的示例性杆状病毒表达载体如图9所示。VLP在体外产生并使用标准技术纯化。

[0443] E. 无细菌序列载体的体内VLP生产和免疫应答

[0444] 在动物模型中通过吸入、鼻内或肌内途径施用msDNA纳米粒子(即,如本实施例的B部分所述)。确定了细胞因子谱、免疫球蛋白谱和对COVID-19的保护作用。

[0445] 对于吸入和鼻内途径,进行以下施用:(1) 冻干的msDNA-VLP纳米粒子或msDNA对照纳米粒子通过吸入以一个或多个剂量施用(例如,以1、2、3和/或4周间隔施用;以2、3、4、5、6、7、8、9、10、11和/或12个月间隔施用;和/或每年间隔施用);(2) 冻干的msDNA-VLP纳米粒子或msDNA对照纳米粒子通过吸入以一个或多个剂量施用(例如,以1、2、3和/或4周间隔施用;以2、3、4、5、6、7、8、9、10、11和/或12个月间隔施用;和/或每年间隔施用),随后以一个或多个剂量鼻内施用(例如,以1、2、3和/或4周间隔施用;以2、3、4、5、6、7、8、9、10、11和/或12个月间隔施用;和/或每年间隔施用)纯化的VLP加强(即,如本实施例的D部分所述);或(3) 以一个或多个剂量鼻内施用纯化的VLP(例如,以1、2、3和/或4周间隔施用;以2、3、4、5、6、7、8、9、10、11和/或12个月间隔施用;和/或每年间隔施用)。

[0446] 对于肌内途径,进行以下施用:(1) msDNA-VLP纳米粒子或msDNA对照纳米粒子通过注射以一个或多个剂量施用(例如,以1、2、3和/或4周间隔施用;以2、3、4、5、6、7、8、9、10、11和/或12个月间隔施用;和/或每年间隔施用);(2) msDNA-VLP纳米粒子或msDNA对照纳米粒子通过注射以一个或多个剂量施用(例如,以1、2、3和/或4周间隔施用;以2、3、4、5、6、7、8、9、10、11和/或12个月间隔施用;和/或每年间隔施用),随后以一个或多个剂量注射施用(例如,以1、2、3和/或4周间隔施用;以2、3、4、5、6、7、8、9、10、11和/或12个月间隔施用;和/或每年间隔施用)纯化的VLP加强;或(3) 以一个或多个剂量注射纯化的VLP(例如,以1、2、3和/或4周间隔施用;以2、3、4、5、6、7、8、9、10、11和/或12个月间隔施用;和/或每年间隔施用)加强。

[0447] 实施例5

[0448] VLP的亲纯化

[0449] 在分析了S蛋白和ACE2受体的四种共结晶结构以及脂蛋白E和ACE2受体的一种共结晶结构后,鉴定了64个残基的ACE2受体肽(“ACE2-64”)是结合冠状病毒S蛋白的充分相互作用界面。ACE2-64的氨基酸序列为:

[0450] STIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQNMNAGDKWSAFLK EQSTLAQMY (SEQ ID NO:70)

[0451] 在ACE2-64的C端或N端编码生物素受体肽(BAP)标签(例如,GLNDIFEAQKIEWHE (SEQ ID NO:71)) (即,分别由SEQ ID NO:73编码的SEQ ID NO:72或由SEQ ID NO:75编码的SEQ ID NO:74)的表达质粒上编码该肽。将表达质粒转化入BirA阳性的大肠杆菌菌株,这导致ACE2-64的一步体内生物素化。裂解细胞,用市售试剂盒纯化生物素化的ACE2-64,并与包被有链霉亲和素的磁性微珠混合。

[0452] 一种针对COVID-19S蛋白的市售单克隆抗体(“S-Ab”)在体外生物素化,并与链霉亲和素包被的磁性微珠混合。

[0453] 在惰性Tris缓冲液(例如,20mM Tris pH 8.0,150mM NaCl)中洗涤并平衡具有固定的ACE2-64或固定的S-Ab的珠。

[0454] 裂解从msDNA-VLP表达VLP的重组细胞,例如实施例2(D)的真核细胞。

[0455] 将具有固定的ACE2-64或固定的S-Ab的珠和含有VLP的细胞裂解物添加到微流体装置中并混合。从细胞裂解液中分离由ACE2-64或S-Ab包被的珠捕获的VLP。然后用中等盐度的缓冲液(例如,20mM Tris pH 8.0,300mM NaCl)洗涤珠三次。然后在高盐度的缓冲液(例如,20mM Tris pH 8.0,1.5M NaCl)中纯化VLP,这导致VLP从珠解离。收集纯化的VLP。进行质量控制分析,如检测RNA和附加体DNA的琼脂糖凝胶电泳、评估gDNA水平的qPCR和电子显微术,以确认VLP的身份和纯度。

[0456] 实施例6

[0457] 用于纳米粒子制剂的靶向性配体的产生

[0458] 肽文库衍生于冠状病毒S蛋白的保守区,并通过肽合成产生。示例性肽是SEQ IDNO:76-99。

[0459] 重组ACE2蛋白购自商业来源。

[0460] COVID-19S蛋白的以下部分作为与ACE2结合的对照,粗体和下划线的残基直接参与ACE2的结合:

RVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRRKRISNCVADYSVLYNSASFST
FKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTG**K**IADYNYKLPDDFTGCVIA
[0461] WNSNNLDSKVGGN**Y**NYL**Y**R**L**FRKSNLKPFFERDISTEIYQAGSTPCNGV**E****G****F****N****C****Y****F****P****L****Q****S**
YGF**Q****P****T****N****G****V****G****Y****Q** (SEQ ID NO:100).

[0462] 根据标准程序进行体外荧光偏振(FP)测定法或类似技术,以确定每种肽对重组ACE2蛋白的亲和力。

[0463] 选择与ACE2受体亲和力最强的配体(即肽),并将其连接到纳米粒子(例如LNP)。

[0464] 确定了单配体和双配体纳米粒子靶向ACE2受体的能力。例如,将含有对ACE2受体具有最高亲和力的配体的纳米粒子的靶向能力与含有对ACE2受体具有最高亲和力的两种不同配体的纳米粒子进行比较。

[0465] 还使用具有一种靶向ACE2受体的配体(例如,以促进ACE2受体介导的内吞)和作为

核定位信号(NLS)的第二配体(例如,以通过核靶向促进适当的细胞内递送)的纳米粒子来测试多配体靶向。

[0466] 序列

[0467] SEQ ID NO:1膜蛋白,氨基酸序列

[0468] MADSNGTITVEELKKLLEQWNLVIGFLFTWICLLQFAYANRNRFLYIIKLIFLWLLWPVTLACFVLA
AVYRINWITGGIAIAMACLVGLMWLSYFIASFRLFARTRSMWSFNPETNILLNVPLHGTILTRPLLESELVIGAVI
LRGHLRIAGHHLGRCDIKDLPKEITVATSRTLSSYYKLGASQRVAGD SGFAAYSRYRIGNYKLNTDHSSSSDNIAL
LVQ

[0469] SEQ ID NO:2膜蛋白,核酸序列

[0470] atggcagattccaacggtactattaccgttgaagagcttaaaaagctccttgaacaatggaacctagt
aataggttttctattccttacatggatttgtcttctacaatttgcctatgccaacaggaataggtttttgtatata
attaagttaattttcctctggtgttatggccagtaacttttagcttgttttgtgcttgcctgtgtttacagaataa
attggatcacgggtggaattgctatcgcaatggcttgtctttagcttgcctgtgctgctgtttacagaataa
tttcagactgtttgcgcgtacgcgttccatgtggtcattcaatccagaaactaacattcttctcaacgtgccactc
catggcactattctgaccagaccgcttctagaaagtgaactcgtaatcggagctgtgatccttcgtggacatcttc
gtattgctggacaccatctaggacgctgtgacatcaaggacctgcctaaagaaatcactgttgctacatcacgaac
gctttcttattacaaattgggagcttcgcagcgtgtagcaggtgactcaggttttgcctacacagtcgctacagg
attggcaactataaattaaacacagaccattccagtagcagtgacaatattgctttgcttgctacagtaa

[0471] SEQ ID NO:3包膜蛋白,氨基酸序列

[0472] MYSFVSEETGTLIVNSVLLFLAFVVFLLVTLAILTALRLCAYCCNIVNVSLVKPSFYV
YSRVKNLNSSRVPDLLV

[0473] SEQ ID NO:4包膜蛋白,核酸序列

[0474] atgtactcattcgttttcggaagagacaggtacgttaatagttaatagcgtacttctttttcttgcttt
cgtggatattcttgctagttacactagccatccttactgcgcttcgattgtgtgctgactgctgcaatattgttaac
gtgagtccttgtaaaaccttctttttacgtttactctcgtgttaaaaatctgaattcttcttagagttcctgatcttc
tggtctaa

[0475] SEQ ID NO:5刺突蛋白,氨基酸序列

[0476] MFVFLVLLPLVSSQCVNLTTTTLQPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIH
VSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSTQSLIVNNATNVVIVKVEFQFCNDPFLGVYYH
KNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFNIDGYFKIYSKHTPINLVRDLPPQGSAL
EPLVDLPIGINITRFQTLALHRSYLTGPDSSSGWTAGAAAAYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK
CTLKSFTVEKGIIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLVNSASFSTF
KCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIADYNYKLPPDFTGCVIAWNSNNLDSKVGGNYNYLYR
LFRKSNLKPFERDISTEIIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPATVCGPKKST
NLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQV
AVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAHEVNNSECDIPGAGICASYQTQTNSPRRAR
SVASQSIIAYTMSLGAENSVAYSNNIAIPTNFTISVTTEILPVSMTKTSVDCMYICGDSTECSNLLLQYGSFCTQ
LNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTADAGFIKQYGD
CLGDIAARDLICAQKFNGLTVLPPLLTDEMAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVL

YENQKLIANQFNSAIGKIQDSLSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRDKVEAEVQI
DRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLSFPQSAPHGVVFLHVTYVPA
QEKNFHTTAPAICHGKAHFPREGVFVSNGTHWFTVQRNFYEPQIITDNTFVSGNCDVVIGIVNNTVYDPLQPELDS
FKEELDKYFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGL
IAIVMTIMLCCMTSCCCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

[0477] SEQ ID NO:6刺突蛋白,核酸序列

[0478] atgtttgtttttcttgttttattgccactagtctcttagtcagtggttaattcttacaaccagaactca
attacccctgcatacactaattctttcacacgtggtgtttattaccctgacaaagttttcagatcctcagtttta
cattcaactcaggacttgttcttacctttcttttccaatgttacttggttccatgctatacatgtctctgggacca
atggtactaagaggtttgataacctgtcctaccatttaaatgatggtgtttattttgcctccactgagaagtctaa
cataataagaggttggttttttggtactacttttagattcgaagaccagtcctacttattgttaataacgctact
aatgtttgttattaaagtctgtgaatttcaattttgtaaatgatccatttttgggtgtttattaccacaaaaacaaca
aaagttggatggaaagtgagttcagagtttattctagtgcgaataattgcacttttgaatatgtctctcagccttt
tcttatggaccttgaaggaaaacagggttaatttcaaaaatcttagggaatttgtgtttaagaatattgatggttat
tttaaaatataattctaagcacacgcctattaatttagtgcgtagatctccctcagggttttctcggttttagaaccat
tggtagatttgccaataggtattaacatcactaggtttcaaactttacttgccttacatagaagttatttgactcc
tggtgattcttcttcagggttgacagctggtgctgcagcttattatgtgggttatcttcaacctaggacttttcta
ttaaaatataatgaaaatggaaccattacagatgctgtagactgtgcacttgacctctctcagaaacaaagtgtga
cgttgaaatccttcactgtagaaaaggaatctatcaaacttctaacttttagagtccaaccaacagaatctattgt
tagatttccataatattacaaacttgtgcccttttgggtgaagtttttaacgccaccagatttgcattctgtttatgct
tggaacaggaagagaatcagcaactgtgttgctgattattctgtcctatataattccgcac

[0479] SEQ ID NO:7内部融合肽,氨基酸序列

[0480] SFIEDLLFNKVTLADAGF

[0481] SEQ ID NO:8内部融合肽,核酸序列

[0482] tcatttattgaagatctacttttcaacaaagtgcacttgcagatgctggcttc

[0483] SEQ ID NO:9受体结合结构域,氨基酸序列

[0484] PNITNLCPFGEVFNATRFASVYAWNKRKISNCVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYAD
SFVIRGDEVQRQIAPGQTGKIADYNYKLPPDFTGCVIAWNSNNLDSKVGNNYLYRLFRKSNLKPFERDISTEIQQA
GSTPCNGVEGFNCYFPLQSYGFQPTNGVGYPYRVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVL
TESNKKFLPFQQFGRDIADTTDAVRDPQTLE

[0485] SEQ ID NO:10受体结合域,核酸序列

[0486] cctaattattacaaacttgtgcccttttgggtgaagtttttaacgccaccagatttgcattctgtttatgc
ttggaacaggaagagaatcagcaactgtgttgctgattattctgtcctatataattccgcacattttccactttt
aagtgttatggagtgtctcctactaaattaaatgatctctgtcttactaatgtctatgcagattcatttgaatta
gaggtgatgaagtcagacaaatcgctccagggaactggaaagattgctgattataattataaattaccagatga
ttttacaggtgcgttatagcttggaattctaacaactttagattctaagggttggtggaattataattacctgtat
agattgttttaggaagtctaattctcaaacttttgagagagatatttcaactgaaatctatcaggccggtagcacac
cttgtaattggtgttgaaggttttaattgttacttttctttacaatcatatggtttccaaccactaatggtgttgg
ttaccaaccatacagagtagtagtactttcttttgaacttctacatgcaccagcaactgtttgtggacctaaaaag

tctactaatttgggttaaaaacaaatgtgtcaattttcaacttcaatgggtttaacaggcacaggtgttcttactgagt
ctaacaaaaagtttctgcctttccaacaatttggcagagacattgctgacactactgatgctgtccgtgatccaca
gacacttgag

[0487] SEQ ID NO:11免疫原性序列,氨基酸序列

[0488] PNITNLCPFGEVFNATRFASVYAWNKRKISNCVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYAD
SFVIRGDEVQRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNLYRLFRKSNLKPFERDISTEIYQA
GSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAP

[0489] SEQ ID NO:12保守氨基酸序列

[0490] SFIEDL

[0491] SEQ ID NO:13保守氨基酸序列

[0492] GVYYYP

[0493] SEQ ID NO:14保守氨基酸序列

[0494] FLPP

[0495] SEQ ID NO:15保守氨基酸序列

[0496] VLPF

[0497] SEQ ID NO:16保守氨基酸序列

[0498] SLLI

[0499] SEQ ID NO:17保守氨基酸序列

[0500] LPIGI

[0501] SEQ ID NO:18保守氨基酸序列

[0502] AAYYV

[0503] SEQ ID NO:19保守氨基酸序列

[0504] TFLI

[0505] SEQ ID NO:20保守氨基酸序列

[0506] AVDC

[0507] SEQ ID NO:21保守氨基酸序列IVRFP

[0508] SEQ ID NO:22保守氨基酸序列ISNC

[0509] SEQ ID NO:23保守氨基酸序列LCFT

[0510] SEQ ID NO:24保守氨基酸序列YNYKL

[0511] SEQ ID NO:25保守氨基酸序列IAWN

[0512] SEQ ID NO:26保守氨基酸序列VVVLSF

[0513] SEQ ID NO:27保守氨基酸序列CVNF

[0514] SEQ ID NO:28保守氨基酸序列GLTG

[0515] SEQ ID NO:29保守氨基酸序列VAVLY

[0516] SEQ ID NO:30保守氨基酸序列GCLI

[0517] SEQ ID NO:31保守氨基酸序列GICA

[0518] SEQ ID NO:32保守氨基酸序列FTIS

[0519] SEQ ID NO:33保守氨基酸序列SVDC

[0520] SEQ ID NO:34保守氨基酸序列YGSFC

- [0521] SEQ ID NO:35保守氨基酸序列FNFS
- [0522] SEQ ID NO:36保守氨基酸序列RDLICAQ
- [0523] SEQ ID NO:37保守氨基酸序列VLPPLL
- [0524] SEQ ID NO:38保守氨基酸序列IPFA
- [0525] SEQ ID NO:39保守氨基酸序列YRFN
- [0526] SEQ ID NO:40保守氨基酸序列KLQDVVN
- [0527] SEQ ID NO:41保守氨基酸序列GAISS
- [0528] SEQ ID NO:42保守氨基酸序列EVQIDRLI
- [0529] SEQ ID NO:43保守氨基酸序列YVTQQL
- [0530] SEQ ID NO:44保守氨基酸序列HLMSF
- [0531] SEQ ID NO:45保守氨基酸序列GVVHLF
- [0532] SEQ ID NO:46保守氨基酸序列WFVT
- [0533] SEQ ID NO:47保守氨基酸序列INAS
- [0534] SEQ ID NO:48保守氨基酸序列LLQF
- [0535] SEQ ID NO:49保守氨基酸序列LWLLWP
- [0536] SEQ ID NO:50保守氨基酸序列LMWL
- [0537] SEQ ID NO:51保守氨基酸序列
- [0538] SFRLF
- [0539] SEQ ID NO:52保守氨基酸序列
- [0540] FNPETN
- [0541] SEQ ID NO:53保守氨基酸序列
- [0542] ITVA
- [0543] SEQ ID NO:54保守氨基酸序列
- [0544] LRLC
- [0545] SEQ ID NO:55重组刺突蛋白,氨基酸序列
- [0546] PNITNLCPFGEVFNATRFASVYAWNKRISNCVADYSVLVNSASFSTFKCYGVSPTKLNDLCFTNVYAD
SFVIRGDEVQRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIQAGSTPCNGVEGFNCFYPLQSYGFQPTNGVGYPYRVVLSFELLHAPGGGGGGSFIEDLLFNKVTADAGFGGGGGGW
PWYIWLGFIAGLIAIVMTIML
- [0547] SEQ ID NO:56重组刺突蛋白,氨基酸序列
- [0548] ccaaacattaccaacctgtgcccccttcggcgaggtgttcaacgccacacggttcgccagcgtgtacgc
ctggaacagaaagcggatcagcaactgcgtggccgactacagtgtcctgtataactccgccagcttttctacattc
aagtgctacggcgtctccccctaccaagctgaacgacctgtgcttcaccaatgtgtacgccgattctttcgtgatca
gaggcgacgaggtgcggcagatcgccccctggccagaccggaaagatcgctgattacaactacaagctgcctgatga
cttcaccggctgcgtgatcgcttgaactccaacaacctggacagcaaggtggggggcaactacaactacctgtac
agactgttcagaaagagcaatctgaagcctttcgagagagatatcagcacagagatctaccaggccggcagcacc
cttgtaatggcgttgagggttcaattgctactttccactgcagagctatggctttcagcctacaaacggcgtggg
ctaccaaccttacagagtgggtggtgctgtcttttcgagctgctgcacgccccctggcggaggaggaggcggatctttc
atcgaggacctgctgttcaacaaggtgacctggccgacgcccgttttggcgggtggcggcgggcggctggccttgg

acatctggctgggcttcacgcccggactgatcgccatcgatggtcaccatcatgctgtga

[0549] SEQ ID NO:57冠状病毒VLP的单个可读框,氨基酸序列

[0550] MYSFVSEETGTLIVNSVLLFLAFVVFLVTLAILTALRLCAYCCNIVNVLKPSFYVYSRVKNLNSSR
VPDLLVATNFSLLKQAGDVEENPGPMADSNITVEELKKLLEQWNLVIGFLFTWICLLQFAYANRNRFLYIIKLI
FLWLLWPVTLACFVLAAYRINWITGGIAIAMACLVGLMWLSYFIASFRLFARTRSMWSFNPETNILLNVPLHGTL
TRPLLESELVIGAVILRGHLRIAGHHLGRCDIKDLPKEITVATSRTLSTYYKLGASQRVAGDSGFAAYSRYRIGNYKL
NTDHSSSSDNIALLVQATNFSLLKQAGDVEENPGPPNITNLCPFGEVFNATRFASVYAWNKRISNCVADYSVLVNS
ASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIADYNYKLDDFTGCVIAWNSNNLDSKVGGN
YNYLYRLFRKSNLKPFERDISTEYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVLSFELLHAPGGGG
GGSFIEDLLFNKVTADAGFGGGGGWPWYIWLGFIAGLIAIVMTIML

[0551] SEQ ID NO:58冠状病毒VLP的单个可读框,核酸序列

[0552] atgtactctttcgtgtctgaggaaaccggcaccctgatcgtaacagcgtgctgctgtttctggcctt
cgtggttttctcgtggtcaccctcgccatcctgaccgacctgaggctgtgcgctactgctgcaacatcgtaac
gtgtctctggtcaaacctagcttctacgtgtatagccgggtgaagaacctgaattctagcagggtgcccacctgc
tggtggccaccaacttcagcctgctgaaacaggctggcgatgtggaagagaacctggacctatggccgatagcaa
cggcaccattacagtggaggaaactcaaaaagctgctggaacagtggaatcttgatcggttctctgttctgacc
tggtctgacctgctgcagttcgccacgccaacagattcctgtacatcatcaaacctgatcttctctgtggc
tgctgtggcccgtgacctggcttgtttcgtgctggctgctgtttatagaatcaactggatcacaggcgcatcgc
aatcgccatggcctgtctggtggcctgatgtggctgagctacttcacgcccagcttttagactgttcgctagaaca
agaagcatgtggtcctttaaccccgagacaaacatcctcctgaatgtgccactgcatggcaccatcctgacaagac
ccctgctggaaagcgagctggtcatcggcgacctgatcctgccccggccacctgagaatcgctggccaccacctggg
cagatgtgacatcaaggacctgccaaggaaatcactgtggccacaagcagaacctcagctactacaagctggga
gcctctcagagagtggccggcgacagcggttcgcccctacagccggtaccggattggcaattacaaactgaaca
ccgaccacagctccagcagcgacaacatcgctctgctagtgcaggccaccaatttcagcctgctgaagcaagctgg
agatgtggaagaaaaccccgccctccaaacattaccaacctgtgcccccttcggcgagggtgttcaacgccacacgg
ttcgccagcgtgtacgcctggaacagaaagcggtacagcaactgcgtggccgactacagtgtcctgtataactccg
ccagcttttctacattcaagtgtacggcgtctcccctaccaagctgaacgacctgtgcttcaccaatgtgtacgc
cgattctttcgtgatcagaggcgacgaggtgcggcagatcgccccctggccagaccggaaagatcgctgattacaac
tacaagctgcctgatgacttcaccggtgctgctgatcgccctggaactccaacaacctggacagcaaggtggggggca
actacaactacctgtacagactgttcagaaagagcaatctgaagcctttcgagagagatatcagcacagagatcta
ccaggccggcgacaccccttgaatggcgttgagggttcaattgctactttccactgcagagctatggctttcag
cctacaaacggcgtgggctaccaaccttacagagtgggtggtgctgtctttcgagctgctgcagccccctggcgag
gaggaggcggtatctttcatcgaggacctgctgttcaacaaggtgacctggccgacgcccgttttggcggtggcg
cgcggttgcccttggtacatctggctgggcttcacgcccggactgatcgccatcgatggtcaccatcatgctg
tga

[0553] SEQ ID NO:59VLP表达盒,核酸序列

[0554] cgttacataacttacggtaaattggccgacctggtgaccgccaacgacccccgccattgacgtcaa
taatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaac
tgcccacttggcagtacatcaagtgtatcatatgccaaagtacgccccctattgacgtcaatgacggtaaattggccc

gcctggcattatgcccagtacatgaccttatgggactttcctacttggcagtacatctacgtattagtcacgcta
ttaccatggtgatgcggttttggcagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtct
ccacccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtacaactcc
gccccattgacgcaaattgggcggtaggcgtgtacggtgggaggtctatataagcagagctggtttagtgaaccgtc
agatccgctagcgccaccatgtactctttcgtgtctgaggaaaccggcaccctgatcgtgaacagcgtgctgctgt
ttctggccttcgtgggttttctgctggtcaccctcgccatcctgaccgacctgcggtgtgcgccctactgctgcaa
catcgtgaacgtgtctctggtcaaacctagcttctacgtgtatagccgggtgaagaacctgaattctagcagggtg
cccgaacctgctgggtggccaccaacttcagcctgctgaaacagcgtggcgatgtggaagagaacctggacctatgg
ccgatagcaacggcaccattacagtggaggaactcaaaaagctgctggaacagtggaatcttgtgatcggttctct
gttcttgacctggatctgcctgctgcagttcgccctacgccaaccgcaacagattcctgtacatcatcaaactgatc
ttcctgtggctgctgtggcccgtagacctggcttgtttcgtgctggctgctgtttatagaatcaactggatcacag
gcgcatcgcaatcgccatggcctgtctggtgggcctgatgtggctgagctacttcatcgccagcttttagactgtt
cgctagaacaagaagcatgtgggtcctttaaccccagacaaaacatcctcctgaatgtgccactgcatggcaccatc
ctgacaagacccctgctggaaagcgagctggctcatcggcgcctgatcctgcggggccacctgagaatcgctggcc
accacctgggcagatgtgacatcaaggacctgccaaggaaatcactgtggccacaagcagaaccttcagctacta
caagctgggagcctctcagagagtggccggcgacagcggttcgcccctacagccggtaccggattggcaattac
aaactgaacaccgaccacagctccagcagcgacaacatcgctctgctagtgcaggccaccaatttcagcctgctga
agcaagctggagatgtggaagaaaaccccgccctccaaacattaccaacctgtgcccccttcggcgaggtgttcaa
cgccacacggttcgccagcgtgtacgcctggaacagaaagcggtatcagcaactgcgtggccgactacagtgtcctg
tataactccgccagcttttctacattcaagtgtacggcgtctcccctaccaagctgaacgacctgtgcttcacca
atgtgtacgccgattctttcgtgatcagaggcgacgaggtgcggcagatcgccccctggccagaccggaaagatcgc
tgattacaactacaagctgcctgatgacttcaccggctgcgtgatcgccctggaactccaacaacctggacagcaag
gtggggggcaactacaactacctgtacagactgttcagaaagagcaatctgaagccttctgagagagatatcgca
cagagatctaccaggccggcagcacccttgtaattggcgttgagggttcaattgctactttccactgcagagcta
tggctttcagcctacaaacggcgtgggttaccaaccttacagagtgggtggtgctgtctttcgagctgctgcacgcc
cctggcggaggaggaggcggtatctttcatcgaggacctgctgttcaacaaggtgacctggccgacgccggttttg
gcggtggcggcgggcggtggccttggtacatctggctgggcttcatcgccgactgatcgccatcgtgatggtcac
catcatgctgtgaacggccggctgatcataatcagccataccacattttagtagggttttacttgccttaaaaaacc
tcccacacctccccctgaacctgaaacataaaatgaatgcaattgttgttgaacttgtttattgcagcttataa
tggttacaaataaagcaatagcatcacaatttcacaaataaagcatttttttactgcattctagttgtggtttg
tccaaactcatcaatgtatctta

[0555] SEQ ID NO:60VLP表达盒,核酸序列

[0556] cgttacataacttacgtaaatggccgcctggctgaccgccaacgacccccgccattgacgtcaa
taatgacgtatgttcccatagtaacgccaatagggaactttccattgacgtcaatgggtggagtatttacggtaaac
tgcccacttggcagtacatcaagtgtatcatatgccaaagtacgccccctattgacgtcaatgacggtaaatggccc
gcttggcattatgcccagtacatgaccttatgggactttcctacttggcagtacatctacgtattagtcacgcta
ttaccatggtgatgcggttttggcagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtct
ccacccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtacaactcc
gccccattgacgcaaattgggcggtaggcgtgtacggtgggaggtctatataagcagagctggtttagtgaaccgtc

agatccgctagcggccaccatgtactctttcgtgtctgaggaaaccggcaccctgatcgtgaacagcgtgctgctgt
ttctggccttcgtgggttttctgctgggtaccctcgccatcctgaccgccctgcggtgtgcgccactgctgcaa
catcgtgaacgtgtctctgggtcaaacctagctttctacgtgtatagccgggtgaagaacctgaattctagcagggtg
cccgaacctgctgggtggccaccaacttcagcctgctgaaacaggctggcgatgtggaagagaacctggacctatgg
ccgatagcaacggcaccattacagtggaggaactcaaaaagctgctggaacagtggaatcttgtgatcggcttct
gttcttgacctggatctgcctgctgcagttcgccctacgccaaaccgcaacagattcctgtacatcatcaaactgatc
ttcctgtggctgctgtggcccgtgacctggcttgtttcgtgctggctgctgtttatagaatcaactggatcacag
gcggcatcgcaatcgccatggcctgtctgggtgggctgatgtggctgagctacttcacgcccagcttttagactgtt
cgctagaacaagaagcatgtgggtcctttaaccccagacaaaacatcctcctgaatgtgccactgcatggcaccatc
ctgacaagacccctgctggaagcgagctgggtcatcgggcgccgtgatcctgcggggccacctgagaatcgctggcc
accacctgggcagatgtgacatcaaggacctgcccaaggaaatcaactgtggccacaagcagaacctcagctacta
caagctgggagcctctcagagagtggccggcgacagcggttcgccgcctacagccgtaccggattggcaattac
aaactgaacaccgaccacagctccagcagcgacaacatcgctctgctagtgcaggccaccaatttcagcctgctga
agcaagctggagatgtggaagaaaaccccgccctccaaacattaccaacctgtgcccccttcggcgaggtgttcaa
cgccacacgggttcgccagcgtgtacgcctggaacagaaagcggtacgcaactgcgtggccgactacagtgtcctg
tataactccgccagcttttctacattcaagtgtacggcgtctccccctaccaagctgaacgacctgtgcttcacca
atgtgtacgccgattctttcgtgatcagaggcgacgaggtgcggcgatcgccccctggccagaccggaaagatcgc
tgattacaactacaagctgcctgatgacttcaccggctgcgtgatcgccctggaactccaacaacctggacagcaag
gtgggggggcaactacaactacctgtacagactgttcagaaagagcaatctgaagcctttcgagagagatatcgca
cagagatctaccaggccggcagcacccttgtaatggcgttgagggttcaattgctactttccactgcagagcta
tggttttcagcctacaaacggcgtgggttaccaaccttacagagtgggtgctgtctttcgagctgctgcacgcc
cctggcggaggaggaggcggtatctttcatcgaggacctgctgtttcaacaaggtgacctggccgacgccggttttg
gcggtggcggcgccggtggccttggtacatctggctgggcttcacgcccggactgatcgccatcgatgggtcac
catcatgctgtgactgtgccttctagttgccagccatctgttgtttgccccctccccctgccttccttgacctgg
aaggtgccactcccactgtcctttcctaataaaatgaggaaattgcatcgcaattgtctgagtaggtgtcattctat
tctggggggtggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgcggtg
ggctctatgg

[0557] SEQ ID NO:61VLP表达盒,核酸序列

[0558] cgttacataacttacggtaaatggcccgctggctgaccgccaacgacccccgcccattgacgtcaa
taatgacgtatgttcccatagtaacgccaatagggtacgtttccattgacgtcaatgggtggagtatttacggtaa
tgcccacttgccagttacatcaagtgtatcatatgccaaagtacgccccctattgacgtcaatgacggtaaatggccc
gcctggcattatgccagttacatgaccttatgggactttcctacttgccagttacatctacgtattagtcacgcta
ttaccatgggtgatgcggttttggcagttacatcaatgggcgtggatagcggtttgactcacggggatttccaagtct
ccacccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtacaactcc
gcccattgacgcaaat

[0559] gggcggtaggcgtgtacgggtgggaggtctatataagcagagctgggttagtgaaccgtcagatccgct
agcggccaccatgtactctttcgtgt

[0560] ctgaggaaaccggcaccctgatcgtgaacagcgtgctgctgtttctggccttcgtgggttttctgctg
gtcaccctcgccatcctgaccgcct

[0561] gcggtgtgctgcctactgctgcaacatcgtgaacgtgtctcttggtcaaacctagcttctacgtgtata
gccgggtgaagaacctgaattctag

[0562] cagggtgcccgaacctgctgggtggccaccaacttcagcctgctgaaacaggctggcgatgtggaagaga
acctggacctgccgatagca

[0563] acggcaccattacagtggaggaaactcaaaaagctgctggaacagtggaaatcttgtgatcggcttcctg
ttcctgacctggatctgcctgctgc

[0564] agtttgcctacgccaaccgcaacagattcctgtacatcatcaaactgatcttcctgtggctgctgtgg
cccgtgacctggcttgtttcgtgctg

[0565] gctgctgtttatagaatcaactggatcacaggcggcatcgcaatcgccatggcctgtctgggtgggcct
gatgtggctgagctacttcacgccc

[0566] agcttttagactgttcgctagaacaagaagcatgttggtcctttaaccccagagacaaacatcctcctgaa
tgtgccaactgcatggcaccatcctg

[0567] acaagacccctgctggaaagcgagctggtcatcggcgcctgatcctgcggggccacctgagaatcgc
tgccaccacctgggcagatg

[0568] tgacatcaaggacctgcccaggaaatcactgtggccacaagcagaacctcagctactacaagctgg
gagcctctcagagagtggccg

[0569] gcgacagcggcttcgccgctacagccggtaccggattggcaattacaaactgaacaccgaccacagc
tccagcagcgacaacatcgct

[0570] ctgctagtgcaggccaccaatcttcagcctgctgaagcaagctggagatgtggaagaaaaccccgcccc
tccaaacattaccaacctgtgcc

[0571] ccttcggcgaggtgttcaacgccacacggttcgccagcgtgtacgcctggaacagaaagcggatcagc
aactgcgtggccgactacagtg

[0572] tcctgtataactccgccagcttttctacattcaagtgtacggcgtctcccctaccaagctgaacgac
ctgtgcttcaccaatgtgtacgccgatt

[0573] ctttcgtgatcagaggcgacgaggtgcggcagatcgccccctggccagaccggaaagatcgctgattac
aactacaagctgcctgatgactt

[0574] caccggctgcgtgatcgctggaactccaacaacctggacagcaaggtggggggcaactacaactacc
tgtacagactgttcagaaagag

[0575] caatctgaagcctttcgagagagatatcagcacagagatctaccaggccggcagcacccttgtaatg
gcgttgagggttcaattgctactt

[0576] tccactgcagagctatggctttcagcctacaaacggcgtgggctaccaaccttacagagtgggtggc
tgtcttttcgagctgctgcacgcccc

[0577] tggcggaggaggaggcgatctttcatcaggacctgctgttcaacaagtgacctggccgacgcgcg
gttttggcggtggcgcgccg

[0578] gctggccttggtacatctggctgggcttcacgcggactgatcgccatcgtgatggtcaccatcatg
ctggagggcagggaagtcttcta

[0579] acatgcggggacgtggaggaaaatcccggcccagagagcgacgagagcggcctgcccgcctggagat
cgagtgcgcacatcccgcc

[0580] acctgaacggcgtggagttcgagctggtgggcggcgagagggcacccccgagcagggccgcatgac

caacaagatgaagagcac

[0581] caaaggcgccctgaccttcagcccctacgtgctgagccacgtgatgggctacggcttctaccacttcg
gcacctaccccagcggtacgag

[0582] aacccttctctgcacgccatcaacaacggcggttacaccaacacccgcacgcagaagtacgaggacgg
cggcgtgctgcacgtgagctt

[0583] cagctaccgctacgaggccggccgctgatcggcgacttcaaggtgatgggcaccggcttccccgagg
acagcgtgatcttcaccgaca

[0584] agatcatccgcagcaacgccaccgtggagcacctgcaccccatgggcgataacgatctggatggcagc
ttcacccgcaccttcagcctgc

[0585] gcgacggcggtactacagctccgtggtggacagccacatgcacttcaagagcgccatccaccccagc
atcctgcagaacggggggcccc

[0586] atgttcgccttccgcccgtggaggaggatcacagcaacaccgagctgggcacgtggagtaccagca
cgcttcaagaccccggatgc

[0587] agatgccggtgaagaaagagtttaaacggccggtgatcataatcagccataccacattttagaggt
tttacttgccttataaaaaacctccac

[0588] acctccccctgaacctgaaacataaaatgaatgcaattgttgtttaaactgtttattgcagcttat
aatggttacaataaagcaatagcatcac

[0589] aaatttcacaaataaagcatttttttactgcattctagttgtggtttgtccaaactcatcaatgtat
ctta

[0590] SEQ ID NO:62VLP表达盒,核酸序列

[0591] cgttacataacttacggtaaatggcccgctggctgaccgccaacgacccccgccattgacgtcaa
taatgacgtatgttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaac
tgcccacttggcagttacatcaagtgtatcatatgccaaagtacgccccctattgacgtcaatgacggtaaatggccc
gcctggcatttatgccagttacatgaccttatgggactttcctacttggcagttacatctacgtatttagtcatcgcta
ttaccatggtgatgcggttttggcagttacatcaatggcggtgtagcggtttgactcacggggatttccaagtct
ccacccattgacgtcaatgggagtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtacaactcc
gccccattgacgcaaatggcggttagcggtgtacggtgggaggtctatataagcagagctggttttagtgaaccgtc
agatccgctagcgccaccatgtactctttcgtgtctgaggaaaccggcaccctgatcgtgaacagcgtgctgctgt
ttctggccttctggttttctgctggtcaccctcgccatcctgaccgcccctgcggctgtgcgcctactgctgcaa
catcgtgaacgtgtctctggtcaaacctagcttctacgtgtatagccgggtgaagaacctgaattctagcagggtg
cccgacctgctggtggccaccaacttcagcctgctgaaacaggtggcgatgtggaagagaacctggacctgccg
atagcaacggcaccattacagtggaggaactcaaaaagctgctggaacagtggaaatcttgtgatcggttctctgtt
cctgacctggatctgcctgctgcagttcgctacgccaaccgcaacagattcctgtacatcatcaactgatcttc
ctgtggctgctgtggccgtgacctggttgtttcgtgctggctgctgtttatagaatcaactggatcacaggcg
gcacgcaatcgccatggcctgtctggtggcctgatgtggctgagctacttcatcgccagcttttagactgttcgc
tagaacaagaagcatgtggtcctttaaccccgagacaaacatcctcctgaatgtgccactgcatggcaccatcctg
acaagacccctgctggaaagcgagctggtcatcgcgccgtgatcctgcggggccacctgagaatcgctggccacc
acctgggcagatgtgacatcaaggacctgcccgaagaaatcactgtggccacaagcagaacctcagctactacaa
gctgggagcctctcagagagtggccggcgacagcggttccgcccctacagccggtaccggatttgcaattacaaa

ctgaacaccgaccacagctccagcagcgacaacatcgctctgctagtgaggagggcaggggaagtcttctaacaat
gcggggacgtggaggaaaatcccgcccaagaccaagctggctagcctcgagcttagagggcccggtttaaacccg
ctgatcagcctcgaggtaccggatccgcggccgcgatatctctagactgtgccttctagttgccagccatctgttg
tttgcacctccccgtgccttccttgacctggaaggtgccactcccactgtcctttcctaataaaatgaggaaat
tgcatcgcatgtgtctgagtaggtgtcattctattctgggggtgggggtggggcaggacagcaagggggaggattgg
gaagacaatagcaggcatgctggggatgcggtgggctctatgg

[0592] SEQ ID NO:63带有VLP表达盒的表达载体,核酸序列

[0593] cccgggaggtaccgagctcttacgcgtgctagaattaaagtaacccaatcagcacacaattgccatta
tacgcgcgtataatggactattgtgtgctgataaacctatttcagcatactacgcgcgtagtatgctgaaataggt
gactagaagttcctatactttctagagaataggaacttcataacttcgtataatgtatgctatacgaagttatggg
ttacttttaatttgggtgctgactaattgagatgcatgctttgcatacttctgcctgctggggagcctggggacttt
ccacacctgggtgctgactaattgagatgcatgctttgcatacttctgcctgctggggagcctggggactttccac
accctgattctgtggataaccgtattaccgccatgcattagttattaatagtaatcaattacggggtcattagtt
catagcccatatatggagttccgcgtt

[0594] acataacttacggtaaattggccgcttggtgaccgccaacgacccccgccattgacgtcaataat
gacgtatgttcccatagtaacgcc

[0595] aatagggactttccattgacgtcaatgggtggagtatttacggtaaactgccacttggcagtacatc
aagtgtatcatatgccaaagtacgcc

[0596] cctattgacgtcaatgacggtaaattggccgcttggtcattatgccagtcacatgaccttatgggactt
tctacttggcagtcacatctacgtatta

[0597] gtcacgctattaccatgggtgatgcggttttggcagtcacatcaatgggcgtggatagcggtttgactc
acggggattttccaagtctccaccca

[0598] ttgacgtcaatgggagtttgttttggcacaaaatcaacgggactttccaaaatgtcgtacaactcc
gccccattgacgcaaatgggcggtta

[0599] ggcgtgtacggtgggaggtctatataagcagagctgggttagtgaaccgtcagatccgctagcgccac
catgtactctttcgtgtctgaggaa

[0600] accggcacccctgatcgtgaacagcgtgctgctgtttctggccttcgtgggttttctgctggtcaccct
cgccatcctgaccgccctgcggctgt

[0601] gcgcctactgctgcaacatcgtgaacgtgtctctggtcaaacctagcttctacgtgtatagccgggtg
aagaacctgaattctagcaggggtgc

[0602] ccgacctgctgggtggccaccaacttcagcctgctgaaacaggctggcgatgtggaagagaaccctgga
cctatggccgatagcaacggc

[0603] accattacagtggaggaactcaaaaagctgctggaacagtggaatcttgtgatcggttctctgttctc
gacctggatctgcctgctgcagttcg

[0604] cctacgccaaccgcaacagattcctgtacatcatcaactgatcttctgtggctgctgtggccccgtg
accctggcttgtttcgtgctggctgct

[0605] gtttatagaatcaactggatcacaggcggcatcgcaatcgccatggcctgtctggtgggcctgatgtg
gctgagctacttcacgccagcttta

[0606] gactgttcgctagaacaagaagcatgtggtcctttaaccccgagacaaacatcctcctgaatgtgcc

ctgcatggcaccatcctgacaaga

[0607] cccctgctggaaagcgagctgggtcatcggcgcctgatcctgcggggccacctgagaatcgctggcca
ccacctgggcagatgtgacatc

[0608] aaggacctgcccaggaaatcactgtggccacaagcagaaccctcagctactacaagctgggagcctc
tcagagagtggccggcgacag

[0609] cggcttcgcgcctacagccggtaccggattggcaattacaaactgaacaccgaccacagctccagca
gcgacaacatcgctctgctagt

[0610] caggccaccaatttcagcctgctgaagcaagctggagatgtggaagaaaaccccgccctccaaacat
taccaacctgtgcccccttcggcg

[0611] aggtgttcaacgccacacggttcgccagcgtgtacgcctggaacagaaagcggatcagcaactgcgtg
gccgactacagtgtcctgtataa

[0612] ctccgccagctttttctacattcaagtgtacggcgtctccccctaccaagctgaacgacctgtgcttca
ccaatgtgtacgccgattcttttcgtgat

[0613] cagaggcgacgaggtgcggcagatcgccccctggccagaccggaaagatcgctgattacaactacaagc
tgctgatgacttcaccgct

[0614] gcgtgatcgcttggaaactccaacaacctggacagcaaggtggggggcaactacaactacctgtacaga
ctgttcagaaagagcaatctga

[0615] agcctttcgagagagatatcagcacagagatctaccaggccggcagcacccttgtaatggcgttgag
ggcttcaattgctactttccactgc

[0616] agagctatggctttcagcctacaaacggcgtgggctaccaaccttacagagtgggtggtgctgtctttc
gagctgctgcacgccccctggcgga

[0617] ggaggaggcggatctttcatcgaggacctgctgttcaacaaggtgacctggccgacgccggttttg
cggtggcgggcgggcggtggcct

[0618] tggtacatctggctgggcttcacgcgcgactgatcgccatcgctgatggtcaccatcatgctgtgact
gtgccttctagttagccagccatctgtt

[0619] gtttgccccctcccccgctgccttccttgacctggaaggtgccactcccactgtcctttcctaataaaa
tgaggaaattgcatcgcatgtgtgagt

[0620] aggtgtcattctattctggggggtggggtggggcaggacagcaagggggaggattgggaagacaatag
caggcatgctgggggatgcggt

[0621] gggctctatggaagcttacgcgtggccgctcgagacgcaattcggttggtgtggaaagtccccaggc
tccccagcaggcagaagtatgc

[0622] aaagcatgcatctcaattagtcagcaaccaggtgtggaaagtccccaggctccccagcaggcagaagt
atgcaaagcatgcatctcaatta

[0623] gtcagcaaccaaattaaagtaaccataacttcgtatagcatacattatacgaagttatgaagttcct
attctctagaaagtataggaacttctagt

[0624] cacctatttcagcatactacgcgctagtatgctgaaataggtttatcagcacacaatagtccattat
acgcgcgtataatggcaattgtgtgctg

[0625] attgggttactttaatttggtaccgctcgaccgatgcccttgagagccttcaaccagtcagctccttc
cggtgggcgcggggcatgactatcgt

[0626] cgccgcacttatgactgtcttctttatcatgcaactcgtaggacaggtgccggcagcgctcttccgct
tcctcgctcactgactcgctgcgctcg
[0627] gtcgttcggctgcggcgagcggatcagctcactcaaaggcggtaatacggttatccacagaatcagg
ggataacgcaggaaagaacatg
[0628] tgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcttgctggcggttttccataggct
ccgccccctgacgagcatcac
[0629] aaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcggttcccc
tggaagctccctcgctgcgctctc
[0630] ctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttct
catagctcacgctgtaggtatctcagtt
[0631] cgggtgtaggtcggttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgaccgctgcgc
ttatccggttaactatcgctttgagt
[0632] ccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagcagagcgagg
tatgtaggcggtgctacagagt
[0633] tcttgaagtggtagcctaactacggctacactagaagaacagtatgttggtatctgcgctctgctgaag
ccagttaccttcggaaaaagagttgg
[0634] tagctcttgatccggcaaaacaaccaccgctggtagcggtaggttttttgtttgcaagcagcagatta
cgcgcaaaaaaaaaaggatctcaaga
[0635] agatcctttgatcttttctacggggtctgacgctcagtggaaacgaaaactcacgttaagggattttgg
tcatgagattatcaaaaaggatcttcac
[0636] ctagatccttttaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctg
acagttaccaatgcttaatcagtgaggcac
[0637] ctatctcagegatctgtctatcttcgttcacatagttgcctgactccccgtcgtgtagataactacg
atacgggagggcttaccatctggcccca
[0638] gtgctgcaatgataccgcgagaccacgctcaccggctccagatttatcagcaataaaccagccagcc
ggaagggccgagcgcaagt
[0639] ggtcctgcaactttatccgcctccatccagctctattaattgttgccgggaagctagagtaagtagttc
gccagttaatagttttgcgcaacgttggt
[0640] gccattgctacaggcatcggtgtcacgctcgctgtttggtatggcttcattcagctccggttccca
acgatcaaggcgagttacatgatecc
[0641] ccatgttggtgcaaaaaagcggtagctccttcggctcctccgatcgttgtcagaagtaagttggccgca
gtgttatcactcatggttatggcagca
[0642] ctgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaa
gtcattctgagaatagtgtatgcggcgac
[0643] cgagttgctcttgcccggtcaatacgggataataccgcgccacatagcagaactttaaaagtgtctc
atcattggaaaacgttcttcggggc
[0644] gaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcaccaactga
tcttcagcatcttttactttcaccagcg
[0645] tttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataaggcgcacacggaaatgt

tgaataactcatactcttccctttttc

[0646] aatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgatgtatttagaaa
aataaacaatataggggttcgcgcacattt

[0647] ccccgaaaagtgccacctgacgcgccctgtagcggcgcathtaagcgcggcggtgtggtggttacgcg
cagcgtgaccgctacacttgc

[0648] cagcgccctagcgcgccgctcctttcgctttcttcccttccctttctcgccacgttcgccggctttcccc
gtcaagctctaaatcgggggctcccttt

[0649] agggttccgatttagtgctttacggcacctcgacccccaaaaaacttgattagggtgatggttcacgta
gtgggccatcgccctgatagacggtt

[0650] tttcgccctttgacgttggagtccacgttctttaatagtgactcttggttccaaactggaacaacact
caaccctatctcggtctattcttttgatttat

[0651] aagggattttgccgatttcggcctattggttaaaaaatgagctgatttaacaaaaatttaacgcgaat
tttaacaaaatattaacgcttacaatttgc

[0652] cattcgccattcaggctgcgcaactgttggaagggcgatcggtgcgggcctcttcgctattacgcca
gccaagctaccatgataagtaag

[0653] taatattaaggtacgtggaggttttacttgctttaaaaaacctcccacacctccccctgaacctgaaa
cataaaatgaatgcaattgttggtttaa

[0654] cttgtttattgcagcttataatggttacaataaagcaatagcatcacaatttcacaaataaagcat
ttttttcactgcattctagttgtggtttgtcc

[0655] aaactcatcaatgtatcttatggtactgtaactgagctaacataa

[0656] SEQ ID NO:64包膜蛋白正向引物,核酸序列

[0657] actgctgcaacatcgtgaac

[0658] SEQ ID NO:65包膜蛋白反向引物,核酸序列

[0659] tgctagaattcaggttcttcacc

[0660] SEQ ID NO:66膜蛋白正向引物,核酸序列

[0661] ttcctgtggctgctgtgg

[0662] SEQ ID NO:67膜蛋白反向引物,核酸序列

[0663] atgaccagctcgctttccag

[0664] SEQ ID NO:68受体结合域正向引物,核酸序列

[0665] atcagcacagagatctaccagg

[0666] SEQ ID NO:69受体结合域反向引物,核酸序列

[0667] agcaccaccactctgtaagg

[0668] SEQ ID NO:70ACE2受体肽,氨基酸序列

[0669] STIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQNMNAGDKWSAFLK EQSTLAQMY

[0670] SEQ ID NO:71BAP标签,氨基酸序列

[0671] GLNDIFEAQKIEWHE

[0672] SEQ ID NO:72具有C端BAP标签的ACE2受体肽,氨基酸序列

[0673] STIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQNMNAGDKWSAFLK
EQSTLAQMYGLNDIFEAQKIEWHE

[0674] SEQ ID NO:73带有C端BAP标签的ACE2受体肽,核酸序列

[0675] tccactattgaagaacaggcaagactttcttggacaaattcaaccacgaggccgaagacttggttcta
tcaaagttcccttgcgagttggaattacaatacgaatatcaccgaagaaaacggttcagaatatgaacaatgcaggc
gacaaatggtccgcctttttgaaagaacaaagtaccctggcccagatgtacggctcttaatgacatctttgaagcgc
aaaagatcgagtggcacgaa

[0676] SEQ ID NO:74具有N-末端BAP标签的ACE2受体肽,氨基酸序列

[0677] GLNDIFEAQKIEWHESTIEEQAKTFLDKFNHEAEDLFYQSSLASWNYNTNITEENVQ
NMNAGDKWSAFLKEQSTLAQMY

[0678] SEQ ID NO:75具有N端BAP标签的ACE2受体肽,核酸序列

[0679] ggtcttaatgacatctttgaagcgcaaaagatcgagtggcacgaatccactattgaagaacaggcaaa
gactttcttggacaaattcaaccacgaggccgaagacttggttctatcaaagttcccttgcgagttggaattacaat
acgaatatcaccgaagaaaacggttcagaatatgaacaatgcaggcgacaaatggtccgcctttttgaaagaacaaa
gtaccctggcccagatgtac

[0680] SEQ ID NO:76ACE2结合肽,氨基酸序列

[0681] QSYGFQPTN

[0682] SEQ ID NO:77ACE2结合肽,氨基酸序列

[0683] LQSYGFQPTN

[0684] SEQ ID NO:78ACE2结合肽,氨基酸序列

[0685] QSYGFQPTNGVGY

[0686] SEQ ID NO:79ACE2结合肽,氨基酸序列

[0687] QPTNGVGY

[0688] SEQ ID NO:80ACE2结合肽,氨基酸序列

[0689] FQPTNGVGY

[0690] SEQ ID NO:81ACE2结合肽,氨基酸序列

[0691] QPTN

[0692] SEQ ID NO:82ACE2结合肽,氨基酸序列

[0693] FQPTN

[0694] SEQ ID NO:83ACE2结合肽,氨基酸序列

[0695] FQPTNGV

[0696] SEQ ID NO:84ACE2结合肽,氨基酸序列

[0697] TNGVGY

[0698] SEQ ID NO:85ACE2结合肽,氨基酸序列

[0699] FNCYFPLQ

[0700] SEQ ID NO:86ACE2结合肽,氨基酸序列

[0701] GFNCYFPLQ

[0702] SEQ ID NO:87ACE2结合肽,氨基酸序列

[0703] EGFN

[0704] SEQ ID NO:88ACE2结合肽,氨基酸序列

[0705] VEGFNCY

- [0706] SEQ ID NO:89ACE2结合肽,氨基酸序列
- [0707] EGFNCYFPLQ
- [0708] SEQ ID NO:90ACE2结合肽,氨基酸序列
- [0709] YNYLY
- [0710] SEQ ID NO:91ACE2结合肽,氨基酸序列
- [0711] NYNYLYR
- [0712] SEQ ID NO:92ACE2结合肽,氨基酸序列
- [0713] SFIEDLLFNKVTADAGF
- [0714] SEQ ID NO:93ACE2结合肽,氨基酸序列
- [0715] SFIEDLLFNKVTADAGFMKQYGCCKKKK
- [0716] SEQ ID NO:94ACE2结合肽,氨基酸序列
- [0717] SFIEDLLF
- [0718] SEQ ID NO:95ACE2结合肽,氨基酸序列
- [0719] SFIEDLLFGCKKKK
- [0720] SEQ ID NO:96ACE2结合肽,氨基酸序列
- [0721] SFIEDLLFNKVTADAGFMKQY
- [0722] SEQ ID NO:97ACE2结合肽,氨基酸序列
- [0723] SFIEDAAAGCKKKK
- [0724] SEQ ID NO:98ACE2结合肽,氨基酸序列
- [0725] SFIEDAAA
- [0726] SEQ ID NO:99ACE2结合肽,氨基酸序列
- [0727] TRYYYLNYNYTTGY
- [0728] SEQ ID NO:100ACE2结合对照肽,氨基酸序列
- [0729] RVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRRKISNCVADYSVLNYSASFSTFKCYGVSP TKL
NDLCFTNVYADSFVIRGDEVQRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNLYNYRLFRKSNL KPF
ERDISTEIQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQ
- [0730] SEQ ID NO:101免疫原性序列、核酸序列
- [0731] cctaataattacaaacttggtgcccttttggtgaagtttttaacgccaccagatttgcatctgtttatgc
ttggaacaggaagagaatcagcaactgtgttgctgattattctgtcctatataattccgcatcattttccactttt
aagtgttatggagtgtctcctactaaattaaatgatctctgtctttactaatgtctatgcagattcatttgtaatta
gaggtgatgaagtcagacaaatcgctccagggaactggaaagattgctgattataattataaattaccagatga
ttttacaggctgcgttatagcttggaattctaacaatcttgattctaaggttggtggaattataattacctgtat
agattgttttaggaagtctaattctcaaaccttttgagagagatatttcaactgaaatctatcaggccggtacacac
cttgtaattggtgttggaagttttaattgttacttttcttttacaatcatatggtttccaaccactaatggtgttgg
ttaccaaccatacagagtagtagtactttcttttgaacttctacatgcacca
- [0732] SEQ ID NO:102跨膜结构域,氨基酸序列
- [0733] WPWYIWLGFIAL
- [0734] SEQ ID NO:103跨膜结构域,核酸序列
- [0735] tggccatggtacattttggttaggttttatagctggcttga

[0736] SEQ ID NO:104无细菌序列载体,核酸序列

[0737] cgcgcgtagtatgctgaaataggtgactagaagttcctatactttctagagaataggaacttcataac
ttcgtataatgtatgctatacgaagtattgggttactttaatttggttgctgactaattgagatgcatgctttgca
tacttctgcctgctggggagcctggggactttccacacctggttgctgactaattgagatgcatgctttgcatact
tctgcctgctggggagcctggggactttccacacccctgattctgtggataaccgtattaccgccatgcattagtt
attaatagtaatacaattacggggtcattagttcatagcccatatatggagttccgcgttacataacttacggtaaa
tgccccgcctggctgaccgcccacgacccccgccattgacgtcaataatgacgtatgttcccatagtaacgcc
atagggactttccattgacgtcaatgggtggagtatttacggtaaactgcccacttggcagtacatcaagtgtatc
atatgccaaagtacgccccctattgacgtcaatgacggtaaattggccccgcctggcattatgccagtacatgacct
atgggacttttctacttggcagtacatctacgtattagtcacgtattaccatgggtgatgcggttttggcagtac
atcaatggcgctggatagcggtttgactcacggggatttccaagtctccaccccatgacgtcaatgggagtttgt
tttggcaccaaaatcaacgggactttccaaaatgtcgtacaactccgccccattgacgcaaatggcggttaggcg
tgtacggtgggaggtctatataagcagagctgggttagtgaaccgtcagatccgctagcgccaccatgtactcttt
cgtgtctgaggaaaccggcacccctgatcgtgaacagcgtgctgctgtttctggccttcgtgggttttctgctggtc
accctcgccatcctgaccgcccctgcggctgtgcgcctactgctgcaacatcgtgaacgtgtctctggtcaaacct
gcttctacgtgtatagccgggtgaagaacctgaattctagcagggtgcccgaacctgctggtggccaccaacttcag
cctgctgaaacaggctggcgatgtggaagagaacctggacctatggccgatagcaacggcaccattacagtggag
gaactcaaaaagctgctggaacagtggaaatcttgtgatcggcttctctgttctgacctggatctgcctgctgcagt
tcgcctacgccaaccgcaacagattcctgtacatcatcaaactgatcttctgtggctgctgtggcccgtagacct
ggcttggttctgctgctggctgctgtttatagaatcaactggatcacaggcgcatcgcaatcgccatggcctgtctg
gtgggcctgatgtggctgagctacttcatcgccagcttttagactgttcgctagaacaagaagcatgtggtccttta
accccgagacaaacatcctcctgaatgtgccactgcatggcaccatcctgacaagacccctgctggaaagcgagct
ggctcatcggcgcccgtgatcctgcggggccacctgagaatcgctggccaccacctgggcagatgtgacatcaaggac
ctgcccgaaggaaatcactgtggccacaagcagaacctcagctactacaagctgggagcctctcagagagtggccg
gcgacagcggttcgcccctacagccggtaccggattggcaattacaaactgaacaccgaccacagctccagcag
cgacaacatcgctctgctagtgcaggccaccaatttcagcctgctgaagcaagctggagatgtggaagaaaacccc
ggccctccaaacattaccaacctgtgccccttcggcgaggtgttcaacgccacacggttcgccagcgtgtacgcct
ggaacagaaagcggatcagcaactgcgtggccgactacagtgtcctgtataactccgccagcttttctacattcaa
gtgctacggcgtctccccctaccaagctgaacgacctgtgcttcaccaatgtgtacgccgattctttcgtgatcaga
ggcgacgaggtgcggcagatcgccctggccagaccggaaagatcgctgattacaactacaagctgcctgatgact
tcaccggctgcgtgatcgcttggaaactccaacaacctggacagcaaggtggggggcaactacaactacctgtacag
actgttcagaaagagcaatctgaagcctttcgagagagatatcagcacagagatctaccaggccggcagcacccct
tgtaatggcggttaggggttcaattgctactttccactgcagagctatggctttcagcctacaaacggcgtgggct
accaaccttacagagtgggtggtgctgtctttcagactgctgcacgcccctggcgaggaggaggcggtatctttcat
cgaggacctgctgttcaacaaggtgacctggccgacgcgggttttggcggtggcgggcggtggccttggttac
atctggctgggcttcatcgccgactgatcgccatcgctgatgggtaccatcatgctgtgactgtgccttctagttg
ccagccatctgttggttggccctccccctgccttcttgacctggaaggtgccactcccactgtcctttcttaa
taaaatgaggaaattgcatcgcatgtctgagtaggtgtcattctattctggggggtggggtggggcaggacagca
agggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggaagcttacgcgtggccgct

cgagacgcaattcggcttggtgtggaaagtccccaggctccccagcaggcagaagtatgcaaagcatgcatctcaa
ttagtcagcaaccagggtgtggaaagtccccaggctccccagcaggcagaagtatgcaaagcatgcatctcaattag
tcagcaaccaaattaaagtaacccataacttcgtatagcatacattatacgaagttatgaagttcctattctctag
aaagtataggaacttctagtcacctatttcagcatactacgcgcg

[0738] 本公开的范围不受本文描述的具体实施方案的限制。事实上,从上述描述和附图中,本领域技术人员将清楚除了所描述的那些之外,对本公开的各种修改是明显的。这种修改旨在落入所附权利要求的范围内。

[0739] 其他实施方案在以下权利要求中。

序列表

- <110> 医学噬菌体生物制药公司
 <120> 用于产生病毒样颗粒的载体及其用途
 <130> 4471.005PC02
 <150> US 63/124,397
 <151> 2020-12-11
 <150> US 63/003,281
 <151> 2020-03-31
 <160> 104
 <170> PatentIn版本3.5
 <210> 1
 <211> 222
 <212> PRT
 <213> 人工序列
 <220>
 <223> 膜蛋白
 <400> 1

```

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

```

	165		170		175	
Ser Tyr Tyr Lys Leu Gly Ala Ser Gln Arg Val Ala Gly Asp Ser Gly						
	180		185		190	
Phe Ala Ala Tyr Ser Arg Tyr Arg Ile Gly Asn Tyr Lys Leu Asn Thr						
	195		200		205	
Asp His Ser Ser Ser Ser Asp Asn Ile Ala Leu Leu Val Gln						
	210		215		220	
<210>	2					
<211>	669					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	膜蛋白					
<400>	2					
atggcagatt ccaacggtac tattaccgtt gaagagctta aaaagctcct tgaacaatgg						60
aacctagtaa taggtttcct attccttaca tggatttgct ttctacaatt tgcctatgcc						120
aacaggaata ggtttttgta tataattaag ttaattttcc tctggctggt atggccagta						180
acttttagctt gttttgtgct tgctgctggt tacagaataa attggatcac cgggtggaatt						240
gctatcgcaa tggcttgtct tgtaggcttg atgtggctca gctacttcat tgctttcttc						300
agactgtttg cgcgtacgcg ttccatgtgg tcattcaatc cagaaactaa cattcttctc						360
aacgtgccac tccatggcac tattctgacc agaccgcttc tagaaagtga actcgtaatc						420
ggagctgtga tccttcgtgg acatcttcgt attgctggac accatctagg acgctgtgac						480
atcaaggacc tgcctaaaga aatcactggt gctacatcac gaacgctttc ttattacaaa						540
ttgggagctt cgcagcgtgt agcaggtgac tcaggttttg ctgcatacag tcgctacagg						600
attggcaact ataaattaaa cacagaccat tccagtagca gtgacaatat tgctttgctt						660
gtacagtaa						669
<210>	3					
<211>	75					
<212>	PRT					
<213>	人工序列					
<220>						
<223>	包膜蛋白					
<400>	3					
Met Tyr Ser Phe Val Ser Glu Glu Thr Gly Thr Leu Ile Val Asn Ser						
1	5		10		15	
Val Leu Leu Phe Leu Ala Phe Val Val Phe Leu Leu Val Thr Leu Ala						
	20		25		30	
Ile Leu Thr Ala Leu Arg Leu Cys Ala Tyr Cys Cys Asn Ile Val Asn						
	35		40		45	

Val Ser Leu Val Lys Pro Ser Phe Tyr Val Tyr Ser Arg Val Lys Asn	
50	55 60
Leu Asn Ser Ser Arg Val Pro Asp Leu Leu Val	
65	70 75
<210> 4	
<211> 228	
<212> DNA	
<213> 人工序列	
<220>	
<223> 包膜蛋白	
<400> 4	
atgtactcat tcgtttcgga agagacaggt acgttaatag ttaatagcgt acttcttttt	60
cttgcttttcg tggatttctt gctagttaca ctagccatcc ttactgcgct tcgattgtgt	120
gcgtactgct gcaatatgtg taacgtgagt cttgtaaaac cttcttttta cgtttactct	180
cgtgttaaaa atctgaattc ttctagagtt cctgatcttc tggctctaa	228
<210> 5	
<211> 1273	
<212> PRT	
<213> 人工序列	
<220>	
<223> 刺突蛋白	
<400> 5	
Met Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val	
1	5 10 15
Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe	
20	25 30
Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu	
35	40 45
His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp	
50	55 60
Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp	
65	70 75 80
Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu	
85	90 95
Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser	
100	105 110
Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile	
115	120 125
Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr	

130	135	140
Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr		
145	150	155
Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu		
165	170	175
Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe		
180	185	190
Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr		
195	200	205
Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu		
210	215	220
Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr		
225	230	235
Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser		
245	250	255
Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro		
260	265	270
Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala		
275	280	285
Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys		
290	295	300
Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val		
305	310	315
Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys		
325	330	335
Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala		
340	345	350
Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu		
355	360	365
Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro		
370	375	380
Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe		
385	390	395
Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly		
405	410	415
Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys		
420	425	430
Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn		
435	440	445

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe			
450	455	460	
Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys			
465	470	475	480
Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly			
	485	490	495
Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val			
	500	505	510
Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys			
	515	520	525
Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn			
	530	535	540
Gly Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu			
545	550	555	560
Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val			
	565	570	575
Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe			
	580	585	590
Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val			
	595	600	605
Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile			
	610	615	620
His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly Ser			
625	630	635	640
Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val			
	645	650	655
Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala			
	660	665	670
Ser Tyr Gln Thr Gln Thr Asn Ser Pro Arg Arg Ala Arg Ser Val Ala			
	675	680	685
Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Ala Glu Asn Ser			
	690	695	700
Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr Asn Phe Thr Ile			
705	710	715	720
Ser Val Thr Thr Glu Ile Leu Pro Val Ser Met Thr Lys Thr Ser Val			
	725	730	735
Asp Cys Thr Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ser Asn Leu			
	740	745	750
Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Thr			

755	760	765
Gly Ile Ala Val Glu Gln Asp Lys Asn Thr Gln Glu Val Phe Ala Gln		
770	775	780
Val Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly Phe		
785	790	795
Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg Ser		800
805	810	815
Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly		
820	825	830
Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala Arg Asp		
835	840	845
Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu		
850	855	860
Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser Ala Leu Leu Ala Gly		
865	870	875
Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile		
885	890	895
Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr		
900	905	910
Gln Asn Val Leu Tyr Glu Asn Gln Lys Leu Ile Ala Asn Gln Phe Asn		
915	920	925
Ser Ala Ile Gly Lys Ile Gln Asp Ser Leu Ser Ser Thr Ala Ser Ala		
930	935	940
Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn		
945	950	955
Thr Leu Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val		
965	970	975
Leu Asn Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln		
980	985	990
Ile Asp Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val		
995	1000	1005
Thr Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn		
1010	1015	1020
Leu Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys		
1025	1030	1035
Arg Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro		
1040	1045	1050
Gln Ser Ala Pro His Gly Val Val Phe Leu His Val Thr Tyr Val		
1055	1060	1065

Pro Ala	Gln Glu Lys Asn Phe	Thr Thr Ala Pro Ala	Ile Cys His
1070	1075	1080	
Asp Gly	Lys Ala His Phe Pro	Arg Glu Gly Val Phe	Val Ser Asn
1085	1090	1095	
Gly Thr	His Trp Phe Val Thr	Gln Arg Asn Phe Tyr	Glu Pro Gln
1100	1105	1110	
Ile Ile	Thr Thr Asp Asn Thr	Phe Val Ser Gly Asn	Cys Asp Val
1115	1120	1125	
Val Ile	Gly Ile Val Asn Asn	Thr Val Tyr Asp Pro	Leu Gln Pro
1130	1135	1140	
Glu Leu	Asp Ser Phe Lys Glu	Glu Leu Asp Lys Tyr	Phe Lys Asn
1145	1150	1155	
His Thr	Ser Pro Asp Val Asp	Leu Gly Asp Ile Ser	Gly Ile Asn
1160	1165	1170	
Ala Ser	Val Val Asn Ile Gln	Lys Glu Ile Asp Arg	Leu Asn Glu
1175	1180	1185	
Val Ala	Lys Asn Leu Asn Glu	Ser Leu Ile Asp Leu	Gln Glu Leu
1190	1195	1200	
Gly Lys	Tyr Glu Gln Tyr Ile	Lys Trp Pro Trp Tyr	Ile Trp Leu
1205	1210	1215	
Gly Phe	Ile Ala Gly Leu Ile	Ala Ile Val Met Val	Thr Ile Met
1220	1225	1230	
Leu Cys	Cys Met Thr Ser Cys	Cys Ser Cys Leu Lys	Gly Cys Cys
1235	1240	1245	
Ser Cys	Gly Ser Cys Cys Lys	Phe Asp Glu Asp Asp	Ser Glu Pro
1250	1255	1260	
Val Leu	Lys Gly Val Lys Leu	His Tyr Thr	
1265	1270		

<210> 6

<211> 1118

<212> DNA

<213> 人工序列

<220>

<223> 刺突蛋白

<400> 6

atgtttgttt ttcttgtttt attgccacta gtctctagtc agtgtgttaa tcttacaacc	60
agaactcaat taccctctgc atacactaat tctttcacac gtggtgttta ttaccctgac	120
aaagttttca gatcctcagt ttacattca actcaggact tggttcttacc tttcttttcc	180
aatgttactt ggttccatgc tatacatgtc tctgggacca atggtactaa gaggtttgat	240

aaccctgtcc taccatttaa tgatggtggt tatttttgctt ccactgagaa gtctaacata 300
ataagaggct ggatttttgg tactacttta gattcgaaga cccagtcctt acttattggt 360
aataacgcta ctaatgttgt tattaaagtc tgtgaatttc aattttgtaa tgatccattt 420
ttgggtgttt attaccacaa aaacaacaaa agttggatgg aaagtgagtt cagagtttat 480
tctagtgcga ataattgcac ttttgaatat gtctctcagc cttttcttat ggaccttgaa 540
ggaaaacagg gtaatttcaa aaatcctagg gaatttgtgt ttaagaatat tgatggttat 600
tttaaaatat attctaagca cacgcctatt aatttagtgc gtgatctccc tcagggtttt 660
tcggcttttag aaccattggt agatttgcca ataggtatta acatcactag gtttcaaact 720
ttacttgctt tacatagaag ttatttgact cctgggtgatt cttcttcagg ttggacagct 780
gggtgctgcag cttattatgt gggttatctt caacctagga cttttctatt aaaatataat 840
gaaaatggaa ccattacaga tgctgtagac tgtgcacttg accctctctc agaaacaaag 900
tgtacgttga aatccttcac tgtagaaaaa ggaatctatc aaacttctaa ctttagagtc 960
caaccaacag aatctattgt tagatttcct aatattacaa acttggtccc ttttggtgaa 1020
gtttttaacg ccaccagatt tgcactgtgt tatgcttgga acaggaagag aatcagcaac 1080
tgtgttgctg attattctgt cctatataat tccgcac 1118

<210> 7

<211> 18

<212> PRT

<213> 人工序列

<220>

<223> 内部融合肽

<400> 7

Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala

1

5

10

15

Gly Phe

<210> 8

<211> 54

<212> DNA

<213> 人工序列

<220>

<223> 内部融合肽

<400> 8

tcattttattg aagatctact tttcaacaaa gtgacacttg cagatgctgg ctte 54

<210> 9

<211> 254

<212> PRT

<213> 人工序列

<220>

<223> 受体结合结构域

<400> 9

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

<210> 10

<211> 762

<212> DNA

<213> 人工序列

<220>

<223> 受体结合结构域

<400> 10
 cctaataatta caaacttgtg cccttttggg gaagttttta acgccaccag atttgcattct 60
 gtttatgctt ggaacaggaa gagaatcagc aactgtgttg ctgattattc tgcctatat 120
 aattccgcat cattttccac ttttaagtgt tatggagtgt ctctactaa attaaatgat 180
 ctctgcttta ctaatgtcta tgcagattca tttgtaatta gaggtgatga agtcagacaa 240
 atcgctccag ggcaaaactgg aaagattgct gattataatt ataaattacc agatgatttt 300
 acaggctgcg ttatagcttg gaattctaac aatcttgatt ctaaggttgg tggtaattat 360
 aattacctgt atagattggt taggaagtct aatctcaaac cttttgagag agatatttca 420
 actgaaatct atcaggccgg tagcacacct tgtaatgggtg ttgaagggtt taattgttac 480
 tttcctttac aatcatatgg tttccaaccc actaatgggtg ttggttacca accatacaga 540
 gtagtagtac tttcttttga acttctacat gcaccagcaa ctgtttgttg acctaaaaag 600
 tctactaatt tggtaaaaaa caaatgtgtc aatttcaact tcaatgggtt aacaggcaca 660
 ggtgttctta ctgagtctaa caaaaagttt ctgcctttcc aacaatttgg cagagacatt 720
 gctgacacta ctgatgctgt ccgtgatcca cagacacttg ag 762

<210> 11

<211> 192

<212> PRT

<213> 人工序列

<220>

<223> 免疫原性序列

<400> 11

Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr
 1 5 10 15
 Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser Asn Cys
 20 25 30
 Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser Thr Phe
 35 40 45
 Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu Cys Phe Thr
 50 55 60
 Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val Arg Gln
 65 70 75 80
 Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr Lys Leu
 85 90 95
 Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn Asn Leu
 100 105 110
 Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg
 115 120 125
 Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu Ile Tyr
 130 135 140

Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn Cys Tyr
145 150 155 160
Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val Gly Tyr
165 170 175
Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His Ala Pro
180 185 190

<210> 12

<211> 6

<212> PRT

<213> 人工序列

<220>

<223> 保守的氨基酸序列

<400> 12

Ser Phe Ile Glu Asp Leu

1 5

<210> 13

<211> 5

<212> PRT

<213> 人工序列

<220>

<223> 保守的氨基酸序列

<400> 13

Gly Val Tyr Tyr Pro

1 5

<210> 14

<211> 4

<212> PRT

<213> 人工序列

<220>

<223> 保守的氨基酸序列

<400> 14

Phe Leu Pro Phe

1

<210> 15

<211> 4

<212> PRT

<213> 人工序列

<220>

<223> 保守的氨基酸序列

<400> 15
Val Leu Pro Phe
1
<210> 16
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 16
Ser Leu Leu Ile
1
<210> 17
<211> 5
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 17
Leu Pro Ile Gly Ile
1 5
<210> 18
<211> 5
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 18
Ala Ala Tyr Tyr Val
1 5
<210> 19
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 19
Thr Phe Leu Leu
1

<210> 20
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 20
Ala Val Asp Cys
1
<210> 21
<211> 5
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 21
Ile Val Arg Phe Pro
1 5
<210> 22
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 22
Ile Ser Asn Cys
1
<210> 23
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 23
Leu Cys Phe Thr
1
<210> 24
<211> 5
<212> PRT

<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 24
Tyr Asn Tyr Lys Leu
1 5
<210> 25
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 25
Ile Ala Trp Asn
1
<210> 26
<211> 6
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 26
Val Val Val Leu Ser Phe
1 5
<210> 27
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 27
Cys Val Asn Phe
1
<210> 28
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列

<400> 28
Gly Leu Thr Gly
1
<210> 29
<211> 5
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 29
Val Ala Val Leu Tyr
1 5
<210> 30
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 30
Gly Cys Leu Ile
1
<210> 31
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 31
Gly Ile Cys Ala
1
<210> 32
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 32
Phe Thr Ile Ser
1

<210> 33
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 33
Ser Val Asp Cys
1
<210> 34
<211> 5
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 34
Tyr Gly Ser Phe Cys
1 5
<210> 35
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 35
Phe Asn Phe Ser
1
<210> 36
<211> 7
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 36
Arg Asp Leu Ile Cys Ala Gln
1 5
<210> 37
<211> 6
<212> PRT

<213> 人工序列

<220>

<223> 保守的氨基酸序列

<400> 37

Val Leu Pro Pro Leu Leu

1 5

<210> 38

<211> 4

<212> PRT

<213> 人工序列

<220>

<223> 保守的氨基酸序列

<400> 38

Ile Pro Phe Ala

1

<210> 39

<211> 4

<212> PRT

<213> 人工序列

<220>

<223> 保守的氨基酸序列

<400> 39

Tyr Arg Phe Asn

1

<210> 40

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> 保守的氨基酸序列

<400> 40

Lys Leu Gln Asp Val Val Asn

1 5

<210> 41

<211> 5

<212> PRT

<213> 人工序列

<220>

<223> 保守的氨基酸序列

<400> 41

Gly Ala Ile Ser Ser

1 5

<210> 42

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> 保守的氨基酸序列

<400> 42

Glu Val Gln Ile Asp Arg Leu Ile

1 5

<210> 43

<211> 6

<212> PRT

<213> 人工序列

<220>

<223> 保守的氨基酸序列

<400> 43

Tyr Val Thr Gln Gln Leu

1 5

<210> 44

<211> 5

<212> PRT

<213> 人工序列

<220>

<223> 保守的氨基酸序列

<400> 44

His Leu Met Ser Phe

1 5

<210> 45

<211> 6

<212> PRT

<213> 人工序列

<220>

<223> 保守的氨基酸序列

<400> 45

Gly Val Val His Leu Phe

1 5

<210> 46
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 46
Trp Phe Val Thr
1
<210> 47
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 47
Ile Asn Ala Ser
1
<210> 48
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 48
Leu Leu Gln Phe
1
<210> 49
<211> 6
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 49
Leu Trp Leu Leu Trp Pro
1 5
<210> 50
<211> 4
<212> PRT

<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 50
Leu Met Trp Leu
1
<210> 51
<211> 5
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 51
Ser Phe Arg Leu Phe
1 5
<210> 52
<211> 6
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 52
Phe Asn Pro Glu Thr Asn
1 5
<210> 53
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列
<400> 53
Ile Thr Val Ala
1
<210> 54
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> 保守的氨基酸序列

<400> 54

Leu Arg Leu Cys

1

<210> 55

<211> 245

<212> PRT

<213> 人工序列

<220>

<223> 重组刺突蛋白

<400> 55

Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr

1

5

10

15

Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser Asn Cys

20

25

30

Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser Thr Phe

35

40

45

Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu Cys Phe Thr

50

55

60

Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val Arg Gln

65

70

75

80

Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr Lys Leu

85

90

95

Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn Asn Leu

100

105

110

Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg

115

120

125

Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu Ile Tyr

130

135

140

Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn Cys Tyr

145

150

155

160

Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val Gly Tyr

165

170

175

Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His Ala Pro

180

185

190

Gly Gly Gly Gly Gly Gly Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys

195

200

205

Val Thr Leu Ala Asp Ala Gly Phe Gly Gly Gly Gly Gly Gly Trp Pro

210

215

220

Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met

225	230	235	240
Val Thr Ile Met Leu			
	245		
<210> 56			
<211> 738			
<212> DNA			
<213> 人工序列			
<220>			
<223> 重组刺突蛋白			
<400> 56			
ccaacatta ccaacctgtg ccccttcggc gaggtgttca acgccacacg gttcgccagc			60
gtgtacgcct ggaacagaaa gcggatcagc aactgcgtgg ccgactacag tgtcctgtat			120
aactccgcca gcttttctac attcaagtgc tacggcgtct cccctacca gctgaacgac			180
ctgtgcttca ccaatgtgta cgccgattct ttcgtgatca gaggcgacga ggtgcggcag			240
atcgccccctg gccagaccgg aaagatcgct gattacaact acaagctgcc tgatgacttc			300
accggctgcg tgatcgccctg gaactccaac aacctggaca gcaaggtggg gggcaactac			360
aactacctgt acagactgtt cagaaagagc aatctgaagc ctttcgagag agatatcagc			420
acagagatct accaggccgg cagcaccctt tgtaatggcg ttgagggtt caattgctac			480
tttccactgc agagctatgg ctttcagcct acaaacggcg tgggctacca accttacaga			540
gtggtggtgc tgtctttcga gctgctgcac gccctggcg gaggaggagg cggatctttc			600
atcgaggacc tgctgttcaa caagtgacc ctggccgacg ccggttttgg cggtggcggc			660
ggcggctggc cttggtacat ctggctgggc ttcacgccc gactgatcgc catcgtgatg			720
gtcaccatca tgctgtga			738
<210> 57			
<211> 580			
<212> PRT			
<213> 人工序列			
<220>			
<223> 冠状病毒VLP的单个可读框			
<400> 57			
Met Tyr Ser Phe Val Ser Glu Glu Thr Gly Thr Leu Ile Val Asn Ser			
1 5 10 15			
Val Leu Leu Phe Leu Ala Phe Val Val Phe Leu Leu Val Thr Leu Ala			
20 25 30			
Ile Leu Thr Ala Leu Arg Leu Cys Ala Tyr Cys Cys Asn Ile Val Asn			
35 40 45			
Val Ser Leu Val Lys Pro Ser Phe Tyr Val Tyr Ser Arg Val Lys Asn			
50 55 60			
Leu Asn Ser Ser Arg Val Pro Asp Leu Leu Val Ala Thr Asn Phe Ser			

65	70	75	80
Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn Pro Gly Pro Met Ala			
	85	90	95
Asp Ser Asn Gly Thr Ile Thr Val Glu Glu Leu Lys Lys Leu Leu Glu			
	100	105	110
Gln Trp Asn Leu Val Ile Gly Phe Leu Phe Leu Thr Trp Ile Cys Leu			
	115	120	125
Leu Gln Phe Ala Tyr Ala Asn Arg Asn Arg Phe Leu Tyr Ile Ile Lys			
	130	135	140
Leu Ile Phe Leu Trp Leu Leu Trp Pro Val Thr Leu Ala Cys Phe Val			
145	150	155	160
Leu Ala Ala Val Tyr Arg Ile Asn Trp Ile Thr Gly Gly Ile Ala Ile			
	165	170	175
Ala Met Ala Cys Leu Val Gly Leu Met Trp Leu Ser Tyr Phe Ile Ala			
	180	185	190
Ser Phe Arg Leu Phe Ala Arg Thr Arg Ser Met Trp Ser Phe Asn Pro			
	195	200	205
Glu Thr Asn Ile Leu Leu Asn Val Pro Leu His Gly Thr Ile Leu Thr			
	210	215	220
Arg Pro Leu Leu Glu Ser Glu Leu Val Ile Gly Ala Val Ile Leu Arg			
225	230	235	240
Gly His Leu Arg Ile Ala Gly His His Leu Gly Arg Cys Asp Ile Lys			
	245	250	255
Asp Leu Pro Lys Glu Ile Thr Val Ala Thr Ser Arg Thr Leu Ser Tyr			
	260	265	270
Tyr Lys Leu Gly Ala Ser Gln Arg Val Ala Gly Asp Ser Gly Phe Ala			
	275	280	285
Ala Tyr Ser Arg Tyr Arg Ile Gly Asn Tyr Lys Leu Asn Thr Asp His			
	290	295	300
Ser Ser Ser Ser Asp Asn Ile Ala Leu Leu Val Gln Ala Thr Asn Phe			
305	310	315	320
Ser Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn Pro Gly Pro Pro			
	325	330	335
Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Arg			
	340	345	350
Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile Ser Asn Cys Val			
	355	360	365
Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys			
	370	375	380

Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn	
385	390 395 400
Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu Val Arg Gln Ile	
	405 410 415
Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro	
	420 425 430
Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser Asn Asn Leu Asp	
	435 440 445
Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys	
	450 455 460
Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln	
465	470 475 480
Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe	
	485 490 495
Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln	
	500 505 510
Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His Ala Pro Gly	
	515 520 525
Gly Gly Gly Gly Gly Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val	
	530 535 540
Thr Leu Ala Asp Ala Gly Phe Gly Gly Gly Gly Gly Gly Trp Pro Trp	
545	550 555 560
Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met Val	
	565 570 575
Thr Ile Met Leu	
	580

<210> 58

<211> 1743

<212> DNA

<213> 人工序列

<220>

<223> 冠状病毒VLP的单个可读框

<400> 58

atgtactctt tcgtgtctga ggaaaccggc accctgatcg tgaacagcgt gctgctgttt	60
ctggccttcg tggttttcct gctggtcacc ctcgccatcc tgaccgccct gcggctgtgc	120
gcctactgct gcaacatcgt gaacgtgtct ctgggtcaaac ctagcttcta cgtgtatagc	180
cgggtgaaga acctgaattc tagcagggtg cccgacctgc tgggtggccac caacttcagc	240
ctgctgaaac aggcctggcga tgtggaagag aaccctggac ctatggccga tagcaacggc	300
accattacag tggaggaact caaaaagctg ctggaacagt ggaatcttgt gatcggttc	360

ctgttcctga cctggatctg cctgctgcag ttgcctacg ccaaccgcaa cagattcctg	420
tacatcatca aactgatctt cctgtggctg ctgtggcccc tgaccctggc ttgtttcgtg	480
ctggctgctg tttatagaat caactggatc acaggcggca tcgcaatcgc catggcctgt	540
ctgggtgggcc tgatgtggct gagctacttc atcgccagct ttagactgtt cgctagaaca	600
agaagcatgt ggtcctttaa ccccagaca aacatcctcc tgaatgtgcc actgcatggc	660
accatcctga caagaccctt gctggaaagc gagctggcca tcggcgccgt gatcctgcgg	720
ggccacctga gaatcgctgg ccaccacctg ggcagatgtg acatcaagga cctgccccag	780
gaaatcactg tggccacaag cagaaccctc agctactaca agctgggagc ctctcagaga	840
gtggccggcg acagcggtt cgccgcctac agccggtacc ggattggcaa ttacaaactg	900
aacaccgacc acagctccag cagcgacaac atcgctctgc tagtgcaggc caccaatttc	960
agcctgctga agcaagctgg agatgtggaa gaaaacccc gccctccaaa cattaccaac	1020
ctgtgcccc tggcgaggt gttcaacgcc acacggttcg ccagcgtgta cgctggaac	1080
agaaagcgga tcagcaactg cgtggccgac tacagtgtcc tgtataactc cgccagcttt	1140
tctacattca agtgctacgg cgtctcccct accaagctga acgacctgtg cttaccaat	1200
gtgtacgccg attctttcgt gatcagaggc gacgaggtgc ggcagatcgc ccctggccag	1260
accggaaaga tcgctgatta caactacaag ctgcctgatg acttcaccgg ctgcgtgatc	1320
gcctggaact ccaacaacct ggacagcaag gtggggggca actacaacta cctgtacaga	1380
ctgttcagaa agagcaatct gaagcctttc gagagagata tcagcacaga gatctaccag	1440
gccgagcaga ccccttgtaa tggcgttgag ggcttcaatt gctactttcc actgcagagc	1500
tatggctttc agcctacaaa cggcgtgggc taccaacctt acagagtggg ggtgctgtct	1560
ttcagactgc tgcacgcccc tggcggagga ggaggcggat ctttcatcga ggacctgctg	1620
ttcaacaagg tgaccctggc cgacgccggt tttggcgggt gcggcgggcg ctggccttgg	1680
tacatctggc tgggcttcat cgccggactg atcgccatcg tgatggtcac catcatgctg	1740
tga	1743

<210> 59

<211> 2523

<212> DNA

<213> 人工序列

<220>

<223> VLP表达盒

<400> 59

cgttacataa cttacggtaa atggcccgcc tggtgaccg cccaacgacc cccgccatt	60
gacgtcaata atgacgtatg ttcccatagt aacgccaata gggactttcc attgacgtca	120
atgggtggag tatttacggt aaactgcca cttggcagta catcaagtgt atcatatgcc	180
aagtacgcc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta	240
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac	300
catggtgatg cggttttggc agtacatcaa tgggcgtgga tagcggtttg actcacgggg	360
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	420
ggactttcca aaatgtcgta acaactccgc ccattgacg caaatgggcg gtaggcgtgt	480

acggtgggag	gtctatataa	gcagagctgg	tttagtgaac	cgtcagatcc	gctagcgcca	540
ccatgtactc	tttcgtgtct	gaggaaaccg	gcaccctgat	cgatgaacagc	gtgctgctgt	600
ttctggcctt	cgtggttttc	ctgctgggtca	ccctcgccat	cctgaccgcc	ctgcggctgt	660
gcgcctactg	ctgcaacatc	gtgaacgtgt	ctctgggtcaa	acctagcttc	tacgtgtata	720
gccgggtgaa	gaacctgaat	tctagcaggg	tgcccgaacct	gctgggtggcc	accaacttca	780
gcctgctgaa	acaggctggc	gatgtggaag	agaaccctgg	acctatggcc	gatagcaacg	840
gcaccattac	agtggaggaa	ctcaaaaagc	tgctggaaca	gtggaatctt	gtgatcggct	900
tcctgttcct	gacctggatc	tgccctgtgc	agttcgccta	cgccaaccgc	aacagattcc	960
tgtacatcat	caaactgac	ttcctgtggc	tgctgtggcc	cgtgaccctg	gcttgtttcg	1020
tgctggctgc	tgtttataga	atcaactgga	tcacaggcgg	catcgcaatc	gccatggcct	1080
gtctgggtgg	cctgatgtgg	ctgagctact	tcacgccag	ctttagactg	ttcgctagaa	1140
caagaagcat	gtggtccttt	aaccccgaga	caaacatcct	cctgaatgtg	ccactgcatg	1200
gcaccatcct	gacaagaccc	ctgctggaaa	gcgagctggt	catcggcgcc	gtgatcctgc	1260
ggggccacct	gagaatcgct	ggccaccacc	tgggcagatg	tgacatcaag	gacctgcccc	1320
aggaaatcac	tgtggccaca	agcagaaccc	tcagctacta	caagctggga	gcctctcaga	1380
gagtggccgg	cgacagcggc	ttcgccgcct	acagccggta	ccggattggc	aattacaaac	1440
tgaacaccga	ccacagctcc	agcagcgaca	acatcgctct	gctagtgcag	gccaccaatt	1500
tcagcctgct	gaagcaagct	ggagatgtgg	aagaaaaccc	cggccctcca	aacattacca	1560
acctgtgccc	cttcggcgag	gtgttcaacg	ccacacggtt	cgccagcgtg	tacgcctgga	1620
acagaaagcg	gatcagcaac	tgctgtggccg	actacagtgt	cctgtataac	tccgccagct	1680
tttctacatt	caagtgtctac	ggcgtctccc	ctaccaagct	gaacgacctg	tgcttcacca	1740
atgtgtacgc	cgattctttc	gtgatcagag	gcgacgaggt	gcggcagatc	gcccctggcc	1800
agaccggaaa	gatcgtgat	tacaactaca	agctgcctga	tgacttcacc	ggctgcgtga	1860
tcgcctggaa	ctccaacaac	ctggacagca	aggtgggggg	caactacaac	tacctgtaca	1920
gactgttcag	aaagagcaat	ctgaagcctt	tcgagagaga	tatcagcaca	gagatctacc	1980
aggccggcag	caccccttgt	aatggcggtg	agggttcaa	ttgctacttt	ccactgcaga	2040
gctatggctt	tcagcctaca	aacggcgtgg	gctaccaacc	ttacagagtg	gtgggtgctgt	2100
ctttcgagct	gctgcacgcc	cctggcggag	gaggaggcgg	atctttcatc	gaggacctgc	2160
tgttcaacaa	ggtgaccctg	gccgacgccg	gttttggcgg	tggcggcgcc	ggctggcctt	2220
ggtacatctg	gctgggcttc	atcgccggac	tgatcgccat	cgtgatggtc	accatcatgc	2280
tgtgaacggc	cggctgatca	taatcagcca	taccacattt	gtagaggttt	tacttgcttt	2340
aaaaaacctc	ccacacctcc	ccctgaacct	gaaacataaa	atgaatgcaa	ttgttgttgt	2400
taacttgttt	attgcagctt	ataatggtta	caaataaagc	aatagcatca	caaatttcac	2460
aaataaagca	tttttttcac	tgcattctag	ttgtggtttg	tccaaactca	tcaatgtatc	2520
tta						2523
<210>	60					
<211>	2510					
<212>	DNA					
<213>	人工序列					

<220>

<223> VLP表达盒

<400> 60

cgttacataa	cttacggtaa	atggccccgcc	tggctgaccg	cccaacgacc	cccgccatt	60
gacgtcaata	atgacgtatg	ttcccatagt	aacgccaata	gggactttcc	attgacgtca	120
atgggtggag	tatttacggt	aaactgcccc	cttggcagta	catcaagtgt	atcatatgcc	180
aagtacgccc	cctattgacg	tcaatgacgg	taaatggccc	gcctggcatt	atgccagta	240
catgacctta	tgggactttc	ctacttggca	gtacatctac	gtattagtca	tcgctattac	300
catggtgatg	cggtttttggc	agtacatcaa	tgggcgtgga	tagcggtttg	actcacgggg	360
atttccaagt	ctccacccca	ttgacgtcaa	tgggagtttg	ttttggcacc	aaaatcaacg	420
ggactttcca	aaatgtcgta	acaactccgc	cccattgacg	caaatgggcg	gtaggcgtgt	480
acggtgggag	gtctatataa	gcagagctgg	tttagtgaac	cgtcagatcc	gctagcgcca	540
ccatgtactc	tttcgtgtct	gaggaaaccg	gcacctgat	cgtgaacagc	gtgctgctgt	600
ttctggcctt	cgtggttttc	ctgctggtea	ccctcgccat	cctgaccgcc	ctgcggctgt	660
gcgcctactg	ctgcaacatc	gtgaacgtgt	ctctggtea	acctagcttc	tacgtgtata	720
gccgggtgaa	gaacctgaat	tctagcaggg	tgcccacct	gctggtggcc	accaacttca	780
gcctgctgaa	acaggctggc	gatgtggaag	agaaccctgg	acctatggcc	gatagcaacg	840
gcaccattac	agtggaggaa	ctcaaaaagc	tgctggaaca	gtggaatctt	gtgatcggct	900
tcctgttcct	gacctggatc	tgctgtctgc	agttcgctca	cgccaaccgc	aacagattcc	960
tgtacatcat	caaactgac	ttcctgtggc	tgctgtggcc	cgtgaccctg	gcttgtttcg	1020
tgctggctgc	tgtttataga	atcaactgga	tcacaggcgg	catcgcaatc	gccatggcct	1080
gtctgggtggg	cctgatgtgg	ctgagctact	tcacgccag	ctttagactg	ttcgctagaa	1140
caagaagcat	gtggtccttt	aaccccgaga	caaacatcct	cctgaatgtg	ccactgcatg	1200
gcaccatcct	gacaagaccc	ctgctggaaa	gcgagctggt	catcggcgcc	gtgatcctgc	1260
ggggccacct	gagaatcgct	ggccaccacc	tgggcagatg	tgacatcaag	gacctgcccc	1320
aggaaatcac	tgtggccaca	agcagaaccc	tcagctacta	caagctggga	gcctctcaga	1380
gagtggccgg	cgacagcggc	ttcgccgcct	acagccggta	ccggattggc	aattacaaac	1440
tgaacaccga	ccacagctcc	agcagcgaca	acatcgctct	gctagtgcag	gccaccaatt	1500
tcagcctgct	gaagcaagct	ggagatgtgg	aagaaaaccc	cggccctcca	aacattacca	1560
acctgtgccc	cttcggcgag	gtgttcaacg	ccacacggtt	cgccagcgtg	tacgcctgga	1620
acagaaaagc	gatcagcaac	tgcgtggccg	actacagtgt	cctgtataac	tcgccagct	1680
tttctacatt	caagtgtac	ggcgtctccc	ctaccaagct	gaacgacctg	tgcttcacca	1740
atgtgtacgc	cgattctttc	tgatcagag	gcgacgaggt	gcggcagatc	gcccctggcc	1800
agaccgga	gatcgctgat	tacaactaca	agctgcctga	tgacttcacc	ggctgcgtga	1860
tcgcctggaa	ctccaacaac	ctggacagca	aggtgggggg	caactacaac	tacctgtaca	1920
gactgttcag	aaagagcaat	ctgaagcctt	tcgagagaga	tatcagcaca	gagatctacc	1980
aggccggcag	caccccttgt	aatggcggtg	agggttcaa	ttgctacttt	ccactgcaga	2040
gctatggctt	tcagcctaca	aacggcgtgg	gctaccaacc	ttacagagtg	gtggtgctgt	2100
ctttcgagct	gctgcacgcc	cctggcgagg	gaggaggcgg	atctttcatc	gaggacctgc	2160

tgttcaacaa	ggtgaccctg	gccgacgccg	gttttggcgg	tggcggcggc	ggctggcctt	2220
ggtacatctg	gctgggcttc	atcgccggac	tgatcgccat	cgtgatggtc	accatcatgc	2280
tgtgactgtg	ccttctagtt	gccagccatc	tgttgtttgc	ccctcccccg	tgcttccctt	2340
gaccttgaa	ggtgccactc	ccactgtcct	ttcctaataa	aatgaggaaa	ttgcatcgca	2400
ttgtctgagt	aggtgtcatt	ctattctggg	gggtgggggtg	gggcaggaca	gcaaggggga	2460
ggattgggaa	gacaatagca	ggcatgctgg	ggatgcgggtg	ggctctatgg		2510
<210>	61					
<211>	3273					
<212>	DNA					
<213>	人工序列					
<220>						
<223>	VLP表达盒					
<400>	61					
cgttacataa	cttacggtaa	atggccccgc	tggtgaccg	cccaacgacc	cccgccatt	60
gacgtcaata	atgacgtatg	ttcccatagt	aacgccaata	gggactttcc	attgacgtca	120
atgggtggag	tatttacggt	aaactgcccc	cttggcagta	catcaagtgt	atcatatgcc	180
aagtacgccc	cctattgacg	tcaatgacgg	taaatggccc	gcctggcatt	atgccagta	240
catgacctta	tgggactttc	ctacttggca	gtacatctac	gtattagtca	tcgctattac	300
catggtgatg	cggttttggc	agtacatcaa	tgggcgtgga	tagcggtttg	actcacgggg	360
atttccaagt	ctccacccca	ttgacgtcaa	tgggagtttg	ttttggcacc	aaaatcaacg	420
ggactttcca	aaatgtcgta	acaactccgc	cccattgacg	caaatgggcg	gtaggcgtgt	480
acgggtgggag	gtctatataa	gcagagctgg	tttagtgaac	cgtcagatcc	gctagcgcca	540
ccatgtactc	tttcgtgtct	gaggaaaccg	gcaccctgat	cgtgaacagc	gtgctgtgt	600
ttctggcctt	cgtggttttc	ctgctgggtca	ccctcgccat	cctgaccgcc	ctgcggctgt	660
gcgcctactg	ctgcaacatc	gtgaacgtgt	ctctgggtcaa	acctagcttc	tacgtgtata	720
gccgggtgaa	gaacctgaat	tctagcaggg	tgcccgacct	gctggtggcc	accaacttca	780
gcctgctgaa	acaggctggc	gatgtggaag	agaaccctgg	acctgccgat	agcaacggca	840
ccattacagt	ggaggaactc	aaaaagctgc	tggaacagtg	gaatcttgtg	atcggttcc	900
tgttcctgac	ctggatctgc	ctgctgcagt	tcgcctacgc	caaccgcaac	agattcctgt	960
acatcatcaa	actgatcttc	ctgtggctgc	tgtggcccg	gacctggct	tgtttcgtgc	1020
tggctgctgt	ttatagaatc	aactggatca	caggcggcat	cgcaatcgcc	atggcctgtc	1080
tgggtggcct	gatgtggctg	agctacttca	tcgccagctt	tagactgttc	gctagaacaa	1140
gaagcatgtg	gtcctttaac	cccagacaa	acatcctcct	gaatgtgcca	ctgcatggca	1200
ccatcctgac	aagaccctg	ctggaaagcg	agctggctcat	cggcgccgtg	atcctgcggg	1260
gccacctgag	aatcgctggc	caccacctgg	gcagatgtga	catcaaggac	ctgcccagg	1320
aaatcaactgt	ggccacaagc	agaaccctca	gtactataca	gctgggagcc	tctcagagag	1380
tggccggcga	cagcggttc	gccgcctaca	gccgtaccg	gattggcaat	tacaaactga	1440
acaccgacca	cagctccagc	agcgacaaca	tcgctctgct	agtgcaggcc	accaatttca	1500
gcctgctgaa	gcaagctgga	gatgtggaag	aaaaccccg	ccctccaaac	attaccaacc	1560

tgtgccccctt	cggcgaggtg	ttcaacgcca	cacggttcgc	cagcgtgtac	gcctggaaca	1620
gaaagcggat	cagcaactgc	gtggccgact	acagtgtcct	gtataactcc	gccagctttt	1680
ctacattcaa	gtgctacggc	gtctccccta	ccaagctgaa	cgacctgtgc	ttcaccaatg	1740
tgtacgccga	ttctttcgtg	atcagaggcg	acgaggtgcg	gcagatcgcc	cctggccaga	1800
ccggaaagat	cgctgattac	aactacaagc	tgctgatga	cttcaccggc	tgctgatcg	1860
cctggaactc	caacaacctg	gacagcaagg	tggggggcaa	ctacaactac	ctgtacagac	1920
tgttcagaaa	gagcaatctg	aagcctttcg	agagagatat	cagcacagag	atctaccagg	1980
ccggcagcac	cccttgtaat	ggcgttgagg	gcttcaattg	ctactttcca	ctgcagagct	2040
atggctttca	gcctacaaac	ggcgtgggct	accaacctta	cagagtgggtg	gtgctgtctt	2100
tcgagctgct	gcacgcccct	ggcggaggag	gaggcggatc	tttcatcgag	gacctgctgt	2160
tcaacaaggt	gaccctggcc	gacgccggtt	ttggcggtgg	cggcgggcgc	tggccttggt	2220
acatctggct	gggcttcac	gccggactga	tcgccatcgt	gatggtcacc	atcatgctgg	2280
agggcagggg	aagtcttcta	acatgcgggg	acgtggagga	aatcccgcgc	ccagagagcg	2340
acgagagcgg	cctgccccgc	atggagatcg	agtccgcgat	caccggcacc	ctgaacggcg	2400
tggagttcga	gctggtgggc	ggcggagagg	gcacccccga	gcagggccgc	atgaccaaca	2460
agatgaagag	caccaaaggc	gccctgacct	tcagccccta	cctgctgagc	cacgtgatgg	2520
gctacggctt	ctaccacttc	ggcacctacc	ccagcggcta	cgagaacccc	ttcctgcacg	2580
ccatcaacaa	cggcggctac	accaacaccc	gcatcgagaa	gtacaggagc	ggcggcgtgc	2640
tgcacgtgag	cttcagctac	cgctacgagg	ccggccgcgt	gatcggcgac	ttcaaggtga	2700
tgggcaccgg	cttccccgag	gacagcgtga	tcttcaccga	caagatcatc	cgcagcaacg	2760
ccaccgtgga	gcacctgcac	cccatgggcg	ataacgatct	ggatggcagc	ttcacccgca	2820
ccttcagcct	gcgcgacggc	ggctactaca	gctccgtggt	ggacagccac	atgcacttca	2880
agagcgccat	ccaccccagc	atcctgcaga	acggggggccc	catgttcgcc	ttccgcccgcg	2940
tggaggagga	tcacagcaac	accgagctgg	gcatcgtgga	gtaccagcac	gccttcaaga	3000
ccccggatgc	agatgccggt	gaagaaagag	tttaaacggc	cggctgatca	taatcagcca	3060
taccacattt	gtagaggttt	tacttgcttt	aaaaaacctc	ccacacctcc	ccctgaacct	3120
gaaacataaa	atgaatgcaa	ttgttggtgt	taacttggtt	attgcagctt	ataatggtta	3180
caaataaagc	aatagcatca	caaatttcac	aaataaagca	tttttttcac	tgcatcttag	3240
ttgtggtttg	tccaaactca	tcaatgtatc	tta			3273

<210> 62

<211> 1859

<212> DNA

<213> 人工序列

<220>

<223> VLP表达盒

<400> 62

cgttacataa	cttacggtaa	atggccccgc	tggtgaccg	cccaacgacc	cccgccatt	60
gacgtcaata	atgacgtatg	ttcccatagt	aacgccaata	gggactttcc	attgacgtca	120
atgggtggag	tatttacggt	aaactgcccc	cttggcagta	catcaagtgt	atcatatgcc	180

aagtacgccc cctattgacg tcaatgacgg taaatggccc gcctggcatt atgcccagta	240
catgacctta tgggactttc ctacttggca gtacatctac gtattagtca tcgctattac	300
catggtgatg cggtttttggc agtacaatcaa tgggcgtgga tagcggtttg actcacgggg	360
atttccaagt ctccacccca ttgacgtcaa tgggagtttg ttttggcacc aaaatcaacg	420
ggactttcca aaatgtcgtg acaactccgc ccattgacg caaatgggcg gtaggcgtgt	480
acgggtgggag gtctatataa gcagagctgg tttagtgaac cgtcagatcc gctagcgcca	540
ccatgtactc tttcgtgtct gaggaaccg gcaccctgat cgtgaacagc gtgctgctgt	600
ttctggcctt cgtggttttc ctgctgggtc cctcgccat cctgaccgcc ctgcggtgt	660
gcgcctactg ctgcaacatc gtgaacgtgt ctctgggtcaa acctagcttc tacgtgtata	720
gccgggtgaa gaacctgaat tctagcaggg tgcccacct gctgggtggc accaacttca	780
gcctgctgaa acaggctggc gatgtggaag agaacctgg acctgccgat agcaacggca	840
ccattacagt ggaggaaactc aaaaagctgc tggaacagt gaatcttgtg atcggcttcc	900
tgttcctgac ctggatctgc ctgctgcagt tcgcctacgc caaccgcaac agattcctgt	960
acatcatcaa actgatcttc ctgtggctgc tgtggcccg gaccctggct tgtttcgtgc	1020
tggtgctgtg ttatagaatc aactggatca caggcggcat cgcaatcgcc atggcctgtc	1080
tggtgggcct gatgtggctg agctacttca tcgccagctt tagactgttc gctagaacaa	1140
gaagcatgtg gtcctttaac cccgagacaa acatcctcct gaatgtgcca ctgcatggca	1200
ccatcctgac aagaccctg ctggaaagcg agctggctcat cggcgccgtg atcctgcggg	1260
gccacctgag aatcgtggc caccacctgg gcagatgtga catcaaggac ctgccaagg	1320
aaatcactgt ggccacaagc agaacctca gctactacaa gctgggagcc tctcagagag	1380
tggccggcga cagcggcttc gccgcctaca gccggtaccg gattggcaat taaaactga	1440
acaccgacca cagctccagc agcgacaaca tcgctctgct agtgcaggag ggcaggggaa	1500
gtcttctaac atgcggggac gtggaggaaa atcccggccc aagaccaag ctggctagcc	1560
tcgagtctag agggcccgtt taaaccgct gatcagcctc gaggtaccgg atccgcggcc	1620
gcgatatctc tagactgtgc cttctagtgt ccagccatct gttgtttgcc cctccccgt	1680
gccttccttg accctggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat	1740
tgcacgcgat tgtctgagta ggtgtcattc tattctgggg ggtgggggtg ggcaggacag	1800
caagggggag gattgggaag acaatagcag gcatgctggg gatgcggtgg gctctatgg	1859

<210> 63

<211> 6222

<212> DNA

<213> 人工序列

<220>

<223> 具有 VLP表达盒的表达载体

<400> 63

cccgggaggt accgagctct tacgcgtgct agaattaaag taaccaatc agcacacaat	60
tgccattata cgcgcgtata atggactatt gtgtgctgat aaacctattt cagcatacta	120
cgcgcgtagt atgctgaaat aggtgactag aagttcctat actttctaga gaataggaac	180
ttcataactt cgtataatgt atgctatacg aagttatggg ttactttaat ttggttgctg	240

actaattgag	atgcatgctt	tgcatacttc	tgcttgctgg	ggagcctggg	gactttccac	300
acctgggtgc	tgactaattg	agatgcatgc	tttgcatact	tctgcctgct	ggggagcctg	360
gggactttcc	acacccctga	ttctgtggat	aaccgtatta	ccgccatgca	ttagttatta	420
atagtaatca	attacggggg	cattagttca	tagcccatat	atggagttcc	gcgttacata	480
acttacggta	aatggcccg	ctggctgacc	gccaacgac	ccccgccc	tgacgtcaat	540
aatgacgtat	gttcccatag	taacgccaat	agggactttc	cattgacgtc	aatgggtgga	600
gtattttacgg	taaactgccc	acttggcagt	acatcaagt	tatcatatgc	caagtacgcc	660
ccctattgac	gtcaatgacg	gtaaatggcc	cgctggcat	tatgcccagt	acatgacctt	720
atgggacttt	cctacttggc	agtacatcta	cgtattagtc	atcgctatta	ccatgggtgat	780
gcggttttgg	cagtacatca	atgggcgtgg	atagcggttt	gactcacggg	gatttccaag	840
tctccacccc	attgacgtca	atgggagttt	gttttggcac	caaaatcaac	gggactttcc	900
aaaatgtcgt	aacaactccg	ccccattgac	gcaaattgggc	ggtaggcgtg	tacggtggga	960
ggtctatata	agcagagctg	gttttagtgaa	ccgtcagatc	cgctagcgcc	accatgtact	1020
ctttcgtgtc	tgaggaaacc	ggcaccctga	tcgtgaacag	cgtgctgctg	tttctggcct	1080
tcgtggtttt	cctgctggtc	accctcgcca	tcctgaccgc	cctgcggctg	tgcgctact	1140
gctgcaacat	cgtgaacgtg	tctcttgtca	aacctagctt	ctacgtgtat	agccgggtga	1200
agaacctgaa	ttctagcagg	gtgcccagcc	tgctgggtgg	caccaacttc	agcctgctga	1260
aacaggctgg	cgatgtggaa	gagaaccctg	gacctatggc	cgatagcaac	ggcaccatta	1320
cagtggagga	actcaaaaag	ctgctggaac	agtggaatct	tgtgatcggc	ttcctgttcc	1380
tgacctggat	ctgcctgctg	cagttcgctt	acgccaaccg	caacagattc	ctgtacatca	1440
tcaaaactgat	cttcctgtgg	ctgctgtggc	ccgtgaccct	ggcttgtttc	gtgctggctg	1500
ctgtttatag	aatcaactgg	atcacaggcg	gcatcgcaat	cgccatggcc	tgtctggtgg	1560
gcctgatgtg	gctgagctac	ttcatcgcca	gcttagact	gttcgctaga	acaagaagca	1620
tgtggtcctt	taaccccag	acaaacatcc	tcctgaatgt	gccactgcat	ggcaccatcc	1680
tgacaagacc	cctgctggaa	agcgagctgg	tcctcgccgc	cgtgatcctg	cggggccacc	1740
tgagaatcgc	tggccaccac	ctgggcagat	gtgacatcaa	ggacctgccc	aaggaaatca	1800
ctgtggccac	aagcagaacc	ctcagctact	acaagctggg	agcctctcag	agagtggccg	1860
gcgacagcgg	cttcgccgcc	tacagccggt	accggattgg	caattacaaa	ctgaacaccg	1920
accacagctc	cagcagcgac	aacatcgctc	tgctagtgea	ggccaccaat	ttcagcctgc	1980
tgaagcaagc	tggagatgtg	gaagaaaacc	ccggccctcc	aaacattacc	aacctgtgcc	2040
ccttcggcga	ggtgttcaac	gccacacggg	tcgccagcgt	gtacgcctgg	aacagaaagc	2100
ggatcagcaa	ctgcgtggcc	gactacagt	tcctgtataa	ctccgccagc	ttttctacat	2160
tcaagtgcta	cggcgtctcc	cctaccaagc	tgaacgacct	gtgcttcacc	aatgtgtacg	2220
ccgattcttt	cgtgatcaga	ggcgacgagg	tgccgcagat	cggccctggc	cagaccggaa	2280
agatcgctga	ttacaactac	aagctgcctg	atgacttcac	cggctgcgtg	atcgctgga	2340
actccaacaa	cctggacagc	aaggtggggg	gcaactacaa	ctacctgtac	agactgttca	2400
gaaagagcaa	tctgaagcct	ttcgagagag	atatcagcac	agagatctac	caggccggca	2460
gcaccccttg	taatggcgtt	gagggttca	attgctactt	tccactgcag	agctatggct	2520
ttcagcctac	aaacggcgtg	ggctaccaac	cttacagagt	ggtggtgctg	tctttcgagc	2580

tgctgcacgc	ccctggcgga	ggaggaggcg	gatctttcat	cgaggacctg	ctgttcaaca	2640
aggtgaccct	ggccgacgcc	ggttttggcg	gtggcgggcg	cggctggcct	tggtacatct	2700
ggctgggctt	catcgccgga	ctgatcgcca	tcgtgatggg	caccatcatg	ctgtgactgt	2760
gccttctagt	tgccagccat	ctgttggttg	cccctcccc	gtgccttcct	tgaccctgga	2820
aggtgccact	cccactgtcc	tttcctaata	aaatgaggaa	attgcatcgc	attgtctgag	2880
taggtgtcat	tctattctgg	ggggtgggg	ggggcaggac	agcaaggggg	aggattggga	2940
agacaatagc	aggcatgctg	gggatgcggg	gggctctatg	gaagcttacg	cgtggccgct	3000
cgagacgcaa	ttcggcttgg	tgtggaaagt	ccccaggctc	cccagcaggc	agaagtatgc	3060
aaagcatgca	tctcaattag	tcagcaacca	ggtgtggaaa	gtccccaggc	tccccagcag	3120
gcagaagtat	gcaaagcatg	catctcaatt	agtcagcaac	caaattaaag	taacccataa	3180
cttcgtatag	catacattat	acgaagtatt	gaagtctcta	ttctctagaa	agtataggaa	3240
cttctagtca	cctatttcag	catactacgc	gcgtagtatg	ctgaaatagg	tttatcagca	3300
cacaatagtc	cattatacgc	gcgtataatg	gcaatttgtg	gctgattggg	ttactttaat	3360
ttggatccgt	cgaccgatgc	ccttgagagc	cttcaaccca	gtcagctcct	tccggtgggc	3420
gcggggcatg	actatcgtcg	ccgcacttat	gactgtcttc	tttatcatgc	aactcgtagg	3480
acaggtgccg	gcagcgctct	tccgcttcct	cgctcactga	ctcgctgcgc	tcggtcgttc	3540
ggctgcggcg	agcggtatca	gctcactcaa	aggcggtaat	acggttatcc	acagaatcag	3600
gggataacgc	aggaaagaac	atgtgagcaa	aaggccagca	aaaggccagg	aaccgtaaaa	3660
aggccgcgtt	gctggcggtt	ttccataggc	tccgcccccc	tgacgagcat	cacaaaaatc	3720
gacgctcaag	tcagagggtg	cgaaacccga	caggactata	aagataccag	gcgtttcccc	3780
ctggaagctc	cctcgtgcgc	tctcctgttc	cgaccctgcc	gcttaccgga	tacctgtccg	3840
cctttctccc	ttcggaagc	gtggcgcttt	ctcatagctc	acgctgtagg	tatctcagtt	3900
cgggtgtaggt	cgttcgctcc	aagctgggct	gtgtgcacga	acccccggtt	cagccccgacc	3960
gctgcgcctt	atccggtaac	tatcgtcttg	agtccaaccc	ggtaaagacac	gacttatcgc	4020
cactggcagc	agccactggg	aacaggatta	gcagagcgag	gtatgtaggc	ggtgctacag	4080
agttcttgaa	gtgggtggcct	aactacggct	acactagaag	aacagtattt	ggtatctgcg	4140
ctctgctgaa	gccagttacc	ttcggaaaaa	gagttggtag	ctcttgatcc	ggcaaaaaaa	4200
ccaccgctgg	tagcgggtgg	ttttttgttt	gcaagcagca	gattacgcgc	agaaaaaaag	4260
gatctcaaga	agatcctttg	atcttttcta	cggggtctga	cgctcagtgg	aacgaaaact	4320
cacgttaagg	gatttttggtc	atgagattat	caaaaaggat	cttcacctag	atccttttaa	4380
attaaaaatg	aagtttttaa	tcaatctaaa	gtatatatga	gtaaacttgg	tctgacagtt	4440
accaatgctt	aatcagtgag	gcacctatct	cagcgatctg	tctatttcgt	tcatccatag	4500
ttgcctgact	ccccgtcgtg	tagataacta	cgatacggga	gggcttacca	tctggcccca	4560
gtgctgcaat	gataccgcga	gaccacgct	caccggtcc	agatttatca	gcaataaacc	4620
agccagccgg	aagggccgag	cgcagaagtg	gtcctgcaac	tttatccgcc	tccatccagt	4680
ctattaattg	ttgccgggaa	gctagagtaa	gtagttcgcc	agttaatagt	ttgcgcaacg	4740
ttgttgccat	tgtacaggc	atcgtggtgt	cacgctcgtc	gtttggtatg	gcttcattca	4800
gctccggttc	ccaacgatca	aggcgagtta	catgatcccc	catgttgtgc	aaaaaagcgg	4860
ttagctcctt	cggtcctccg	atcgttgtca	gaagtaagtt	ggccgcagtg	ttatcactca	4920

tggttatggc agcactgcat aattctctta ctgtcatgcc atccgtaaga tgcttttctg	4980
tgactgggtga gtactcaacc aagtcattct gagaatagtg tatgcggcga ccgagttgct	5040
cttgcccggc gtcaatacgg gataataccg cgccacatag cagaacttta aaagtgtca	5100
tcattggaaa acgttcttcg gggcgaaaac tctcaaggat cttaccgctg ttgagatcca	5160
gttcgatgta acccactcgt gcacccaact gatcttcagc atcttttact ttcaccagcg	5220
tttctgggtg agcaaaaaca ggaaggcaaa atgccgcaaa aaagggaata agggcgacac	5280
ggaaatgttg aatactcata ctcttccttt ttcaatatta ttgaagcatt tatcagggtt	5340
attgtctcat gagcggatac atatttgaat gtatttagaa aaataaaca ataggggttc	5400
cgcgcacatt tccccgaaaa gtgccacctg acgcgccctg tagcggcgca ttaagcgcg	5460
cgggtgtggt ggttacgcgc agcgtgaccg ctacacttgc cagcgcccta gcgcccgtc	5520
ctttcgcttt cttcccttcc tttctcgcca cgttcgccgg ctttccccgt caagctctaa	5580
atcgggggct ccctttaggg ttccgattta gtgctttacg gcacctcgac ccaaaaaaac	5640
ttgattaggg tgatggttca cgtagtgggc catcgccctg atagacggtt tttcgccctt	5700
tgacgttgga gtccacgttc tttaatagtg gactcttggt ccaaactgga acaacactca	5760
accctatctc ggtctattct tttgatttat aagggatttt gccgatttcg gcctattggt	5820
taaaaaatga gctgatttaa caaaaattta acgcgaattt taacaaaata ttaacgtta	5880
caatttgcca ttcgccattc aggctgcgca actgttgga agggcgatcg gtgcgggcct	5940
cttcgctatt acgccagccc aagctaccat gataagtaag taatattaag gtacgtggag	6000
gttttacttg ctttaaaaaa cctcccacac ctccccctga acctgaaaca taaaatgaat	6060
gcaattgttg ttgttaactt gtttattgca gcttataatg gttacaaata aagcaatagc	6120
atcacaaatt tcacaaataa agcatttttt tctactgcatt ctagttgtgg tttgtccaaa	6180
ctcatcaatg tatcttatgg tactgtaact gagctaacat aa	6222
<210> 64	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<223> 正向引物	
<400> 64	
actgctgcaa catcgtgaac	20
<210> 65	
<211> 23	
<212> DNA	
<213> 人工序列	
<220>	
<223> 反向引物	
<400> 65	
tgctagaatt caggttcttc acc	23
<210> 66	

<211>	18	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	正向引物	
<400>	66	
ttcctgtggc tgctgtgg		18
<210>	67	
<211>	20	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	反向引物	
<400>	67	
atgaccagct cgctttccag		20
<210>	68	
<211>	22	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	正向引物	
<400>	68	
atcagcacag agatctacca gg		22
<210>	69	
<211>	20	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	反向引物	
<400>	69	
agcaccacca ctctgtaagg		20
<210>	70	
<211>	65	
<212>	PRT	
<213>	人工序列	
<220>		
<223>	ACE2受体肽	
<400>	70	
Ser Thr Ile Glu Glu Gln Ala Lys Thr Phe Leu Asp Lys Phe Asn His		

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

Tyr

65

<210> 71

<211> 15

<212> PRT

<213> 人工序列

<220>

<223> BAP标签

<400> 71

Gly Leu Asn Asp Ile Phe Glu Ala Gln Lys Ile Glu Trp His Glu

1	5	10	15
---	---	----	----

<210> 72

<211> 80

<212> PRT

<213> 人工序列

<220>

<223> 具有C端BAP标签的ACE2受体肽

<400> 72

Ser Thr Ile Glu Glu Gln Ala Lys Thr Phe Leu Asp Lys Phe Asn His

1	5	10	15
---	---	----	----

Glu Ala Glu Asp Leu Phe Tyr Gln Ser Ser Leu Ala Ser Trp Asn Tyr

20	25	30
----	----	----

Asn Thr Asn Ile Thr Glu Glu Asn Val Gln Asn Met Asn Asn Ala Gly

35	40	45
----	----	----

Asp Lys Trp Ser Ala Phe Leu Lys Glu Gln Ser Thr Leu Ala Gln Met

50	55	60
----	----	----

Tyr Gly Leu Asn Asp Ile Phe Glu Ala Gln Lys Ile Glu Trp His Glu

65	70	75	80
----	----	----	----

<210> 73

<211> 240

<212> DNA

<213> 人工序列

〈213〉 人工序列

<220>

<223> ACE2结合肽

<400> 76

Gln Ser Tyr Gly Phe Gln Pro Thr Asn

1 5

<210> 77

<211> 10

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 77

Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn

1 5 10

<210> 78

<211> 13

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 78

Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val Gly Tyr

1 5 10

<210> 79

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 79

Gln Pro Thr Asn Gly Val Gly Tyr

1 5

<210> 80

<211> 9

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 80

Phe Gln Pro Thr Asn Gly Val Gly Tyr

1 5

<210> 81

<211> 4

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 81

Gln Pro Thr Asn

1

<210> 82

<211> 5

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 82

Phe Gln Pro Thr Asn

1 5

<210> 83

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 83

Phe Gln Pro Thr Asn Gly Val

1 5

<210> 84

<211> 6

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 84

Thr Asn Gly Val Gly Tyr

1 5

<210> 85

<211> 8
<212> PRT
<213> 人工序列
<220>
<223> ACE2结合肽
<400> 85
Phe Asn Cys Tyr Phe Pro Leu Gln
1 5
<210> 86
<211> 9
<212> PRT
<213> 人工序列
<220>
<223> ACE2结合肽
<400> 86
Gly Phe Asn Cys Tyr Phe Pro Leu Gln
1 5
<210> 87
<211> 4
<212> PRT
<213> 人工序列
<220>
<223> ACE2结合肽
<400> 87
Glu Gly Phe Asn
1
<210> 88
<211> 7
<212> PRT
<213> 人工序列
<220>
<223> ACE2结合肽
<400> 88
Val Glu Gly Phe Asn Cys Tyr
1 5
<210> 89
<211> 10
<212> PRT
<213> 人工序列

<220>

<223> ACE2结合肽

<400> 89

Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln

1 5 10

<210> 90

<211> 5

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 90

Tyr Asn Tyr Leu Tyr

1 5

<210> 91

<211> 7

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 91

Asn Tyr Asn Tyr Leu Tyr Arg

1 5

<210> 92

<211> 18

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 92

Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala

1 5 10 15

Gly Phe

<210> 93

<211> 29

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 93

Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala

1 5 10 15

Gly Phe Met Lys Gln Tyr Gly Cys Gly Lys Lys Lys Lys

20 25

<210> 94

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 94

Ser Phe Ile Glu Asp Leu Leu Phe

1 5

<210> 95

<211> 15

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 95

Ser Phe Ile Glu Asp Leu Leu Phe Gly Cys Gly Lys Lys Lys Lys

1 5 10 15

<210> 96

<211> 22

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 96

Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala

1 5 10 15

Gly Phe Met Lys Gln Tyr

20

<210> 97

<211> 15

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 97

Ser Phe Ile Glu Asp Ala Ala Ala Gly Cys Gly Lys Lys Lys Lys

1 5 10 15

<210> 98

<211> 8

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 98

Ser Phe Ile Glu Asp Ala Ala Ala

1 5

<210> 99

<211> 14

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合肽

<400> 99

Thr Arg Tyr Tyr Tyr Leu Asn Tyr Asn Tyr Thr Thr Gly Tyr

1 5 10

<210> 100

<211> 188

<212> PRT

<213> 人工序列

<220>

<223> ACE2结合对照肽

<400> 100

Arg Val Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn

1 5 10 15

Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val

20 25 30

Tyr Ala Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser

35 40 45

Val Leu Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val

50 55 60

Ser Pro Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp

65 70 75 80

Ser Phe Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln
85 90 95
Thr Gly Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr
100 105 110
Gly Cys Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly
115 120 125
Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys
130 135 140
Pro Phe Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr
145 150 155 160
Pro Cys Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser
165 170 175
Tyr Gly Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln
180 185

<210> 101

<211> 576

<212> DNA

<213> 人工序列

<220>

<223> 免疫原性序列

<400> 101

cctaataatta caaacttggt cccttttggt gaagttttta acgccaccag atttgcattt 60
gtttatgctt ggaacaggaa gagaatcagc aactgtgttg ctgattattc tgcctatat 120
aattccgcat cattttccac ttttaagtgt tatggagtgt ctctactaa attaaatgat 180
ctctgcttta ctaatgtcta tgcagattca tttgtaatta gaggtgatga agtcagacaa 240
atcgcctccag ggcaaactgg aaagattgct gattataatt ataaattacc agatgatttt 300
acaggctgctg ttatagcttg gaattctaac aatcttgatt ctaaggttgg tggtaattat 360
aattacctgt atagattggt taggaagtct aatctcaaac cttttgagag agatattttca 420
actgaaatct atcaggccgg tagcacacct tgtaatgggtg ttgaagggtt taattgttac 480
tttcctttac aatcatatgg tttccaaccc actaatgggtg ttggttacca accatacaga 540
gtagtagtac tttcttttga acttctacat gcacca 576

<210> 102

<211> 13

<212> PRT

<213> 人工序列

<220>

<223> 跨膜结构域

<400> 102

Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu

1	5	10	
<210>	103		
<211>	40		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	跨膜结构域		
<400>	103		
tggccatggt acatttggct aggttttata gctggcttga			40
<210>	104		
<211>	3153		
<212>	DNA		
<213>	人工序列		
<220>			
<223>	无细菌序列载体		
<400>	104		
cgcgcgtagt atgctgaaat aggtgactag aagttcctat actttctaga gaataggaac			60
ttcataactt cgtataatgt atgctatacg aagttagggg ttactttaat ttggttgctg			120
actaattgag atgcatgctt tgcatacttc tgcctgctgg ggagcctggg gactttccac			180
acctggttgc tgactaattg agatgcatgc tttgcatact tctgcctgct ggggagcctg			240
gggactttcc acacccctga ttctgtggat aaccgtatta ccgccatgca ttagttatta			300
atagtaatca attacgggggt cattagtcca tagcccatat atggagttcc gcgttacata			360
acttacggta aatggccccg ctggctgacc gccaacgac ccccgcccat tgacgtcaat			420
aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc aatgggtgga			480
gtattttacgg taaactgccc acttggcagt acatcaagt tatcatatgc caagtacgcc			540
ccctattgac gtcaatgacg gtaaattggc cgctggcat tatgcccagt acatgacctt			600
atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta ccatggtgat			660
gcggttttgg cagtacatca atgggcgtgg atagcggttt gactcacggg gatttccaag			720
tctccacccc attgacgtca atgggagttt gttttggcac caaatcaac gggactttcc			780
aaaatgtcgt aacaactccg ccccatgac gcaaatgggc ggtaggcgtg tacggtggga			840
ggtctatata agcagagctg gtttagtgaa ccgtcagatc cgctagcgcc accatgtact			900
ctttcgtgtc tgaggaaacc ggacacctga tcgtgaacag cgtgctgctg tttctggcct			960
tcgtggtttt cctgctggtc accctcgcca tctgaccgc cctgcggctg tgcgcctact			1020
gctgcaacat cgtgaacgtg tctctgggtca aacctagctt ctacgtgtat agccgggtga			1080
agaacctgaa ttctagcagg gtgcccgacc tgctgggtggc caccaacttc agcctgctga			1140
aacaggctgg cgatgtggaa gagaacctg gacctatggc cgatagcaac ggcaccatta			1200
cagtggagga actcaaaaag ctgctggaac agtggaatct tgtgatcggc ttctgtttcc			1260
tgacctggat ctgcctgctg cagttcgctt acgccaaccg caacagattc ctgtacatca			1320
tcaaaactgat cttcctgtgg ctgctgtggc ccgtgacctt ggcttgtttc gtgctggctg			1380

ctgtttatag	aatcaactgg	atcacaggcg	gcatcgcaat	cgccatggcc	tgtctggtgg	1440
gcctgatgtg	gctgagctac	ttcatcgcca	gctttagact	gttcgctaga	acaagaagca	1500
tgtggtcctt	taaccccag	acaaacatcc	tcctgaatgt	gccactgcat	ggcaccatcc	1560
tgacaagacc	cctgctggaa	agcgagctgg	tcacggcgcg	cgtgatcctg	cggggccacc	1620
tgagaatcgc	tggccaccac	ctgggcagat	gtgacatcaa	ggacctgccc	aaggaaatca	1680
ctgtggccac	aagcagaacc	ctcagctact	acaagctggg	agcctctcag	agagtggccg	1740
gcgacagcgg	cttcgccgcc	tacagccggt	accggattgg	caattacaaa	ctgaacaccg	1800
accacagctc	cagcagcgac	aacatcgctc	tgctagtgcg	ggccaccaat	ttcagcctgc	1860
tgaagcaagc	tggagatgtg	gaagaaaacc	ccggccctcc	aaacattacc	aacctgtgcc	1920
ccttcggcga	ggtgttcaac	gccacacggg	tcgccagcgt	gtacgcctgg	aacagaaagc	1980
ggatcagcaa	ctgcgtggcc	gactacagtg	tcctgtataa	ctccgccagc	ttttctacat	2040
tcaagtgcta	cggcgtctcc	cctaccaagc	tgaacgacct	gtgcttcacc	aatgtgtacg	2100
ccgattcttt	cgtgatcaga	ggcgacgagg	tgcggcagat	cggccctggc	cagaccggaa	2160
agatcgctga	ttacaactac	aagctgcctg	atgacttcac	cggctgcgtg	atcgccctgga	2220
actccaacaa	cctggacagc	aaggtggggg	gcaactacaa	ctacctgtac	agactgttca	2280
gaaagagcaa	tctgaagcct	ttcgagagag	atatcagcac	agagatctac	caggccggca	2340
gcaccccttg	taatggcgtt	gagggttca	attgctactt	tccactgcag	agctatggct	2400
ttcagcctac	aaacggcgtg	ggctaccaac	cttacagagt	ggtggtgctg	tctttcgagc	2460
tgtgcacgc	ccctggcgga	ggaggaggcg	gatctttcat	cgaggacctg	ctgttcaaca	2520
aggtgaccct	ggccgacgcc	ggttttggcg	gtggcggcgg	cggctggcct	tggtacatct	2580
ggctgggctt	catcgccgga	ctgatcgcca	tcgtgatggt	caccatcatg	ctgtgactgt	2640
gccttctagt	tgccagccat	ctgttgtttg	cccctcccc	gtgccttcct	tgaccctgga	2700
aggtgccact	cccactgtcc	tttcctaata	aaatgaggaa	attgcatcgc	attgtctgag	2760
taggtgtcat	tctattctgg	gggggtgggt	ggggcaggac	agcaaggggg	aggattggga	2820
agacaatagc	aggcatgctg	gggatgcggt	gggctctatg	gaagcttacg	cgtggccgct	2880
cgagacgcaa	ttcggccttg	tgtggaaagt	ccccaggctc	cccagcaggc	agaagtatgc	2940
aaagcatgca	tctcaattag	tcagcaacca	ggtgtggaaa	gtccccaggc	tccccagcag	3000
gcagaagtat	gcaaagcatg	catctcaatt	agtcagcaac	caaattaaag	taaccataa	3060
cttcgtatag	catacattat	acgaagttat	gaagttccta	ttctctagaa	agtataggaa	3120
cttctagtca	cctatttcag	catactacgc	gcg			3153

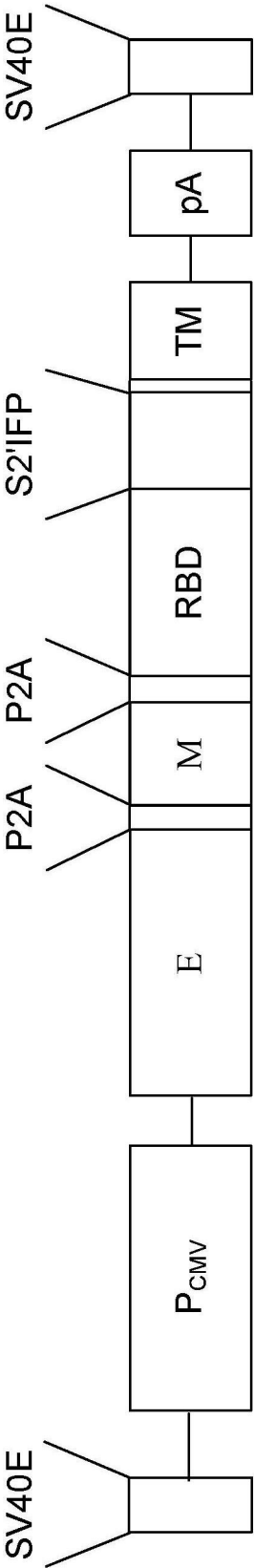


图1

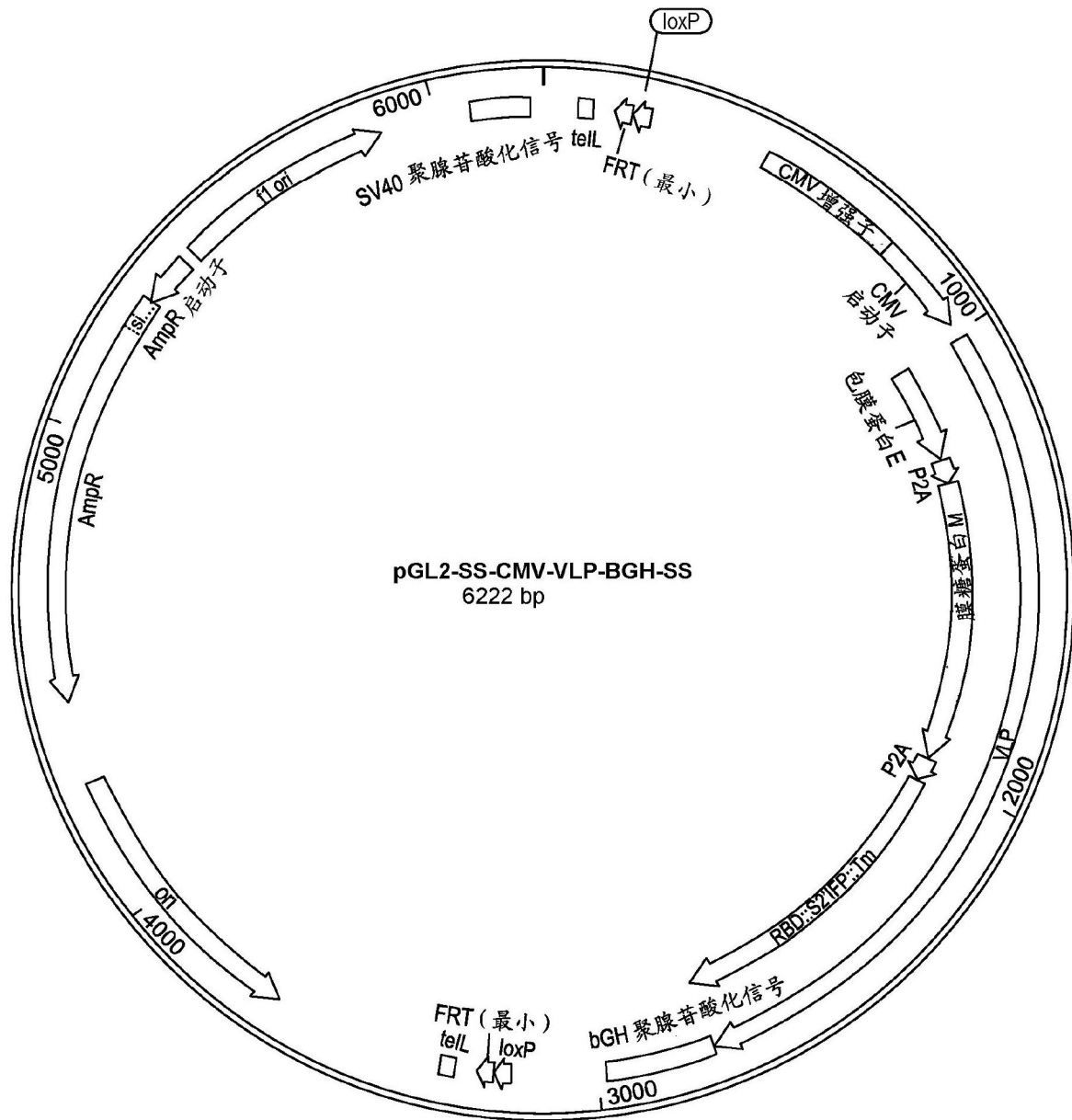


图2

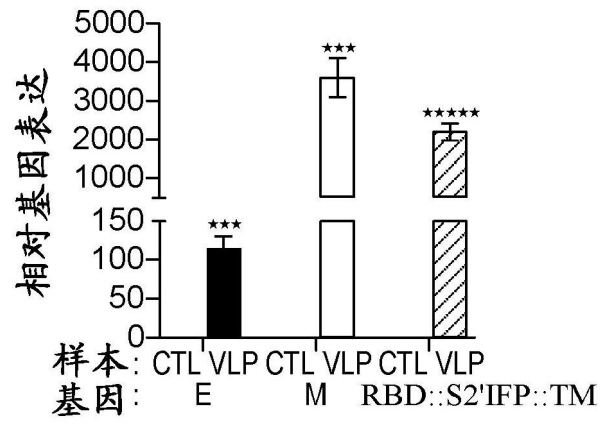


图3A

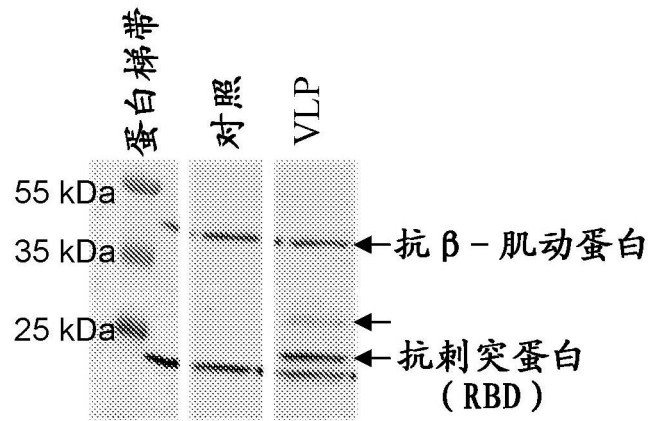


图3B

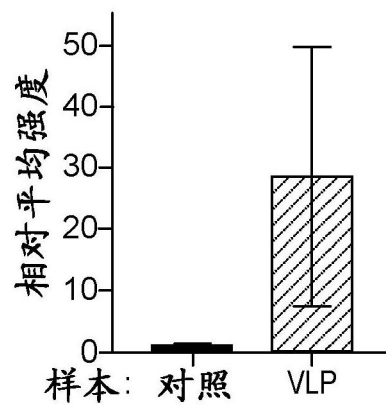


图3C

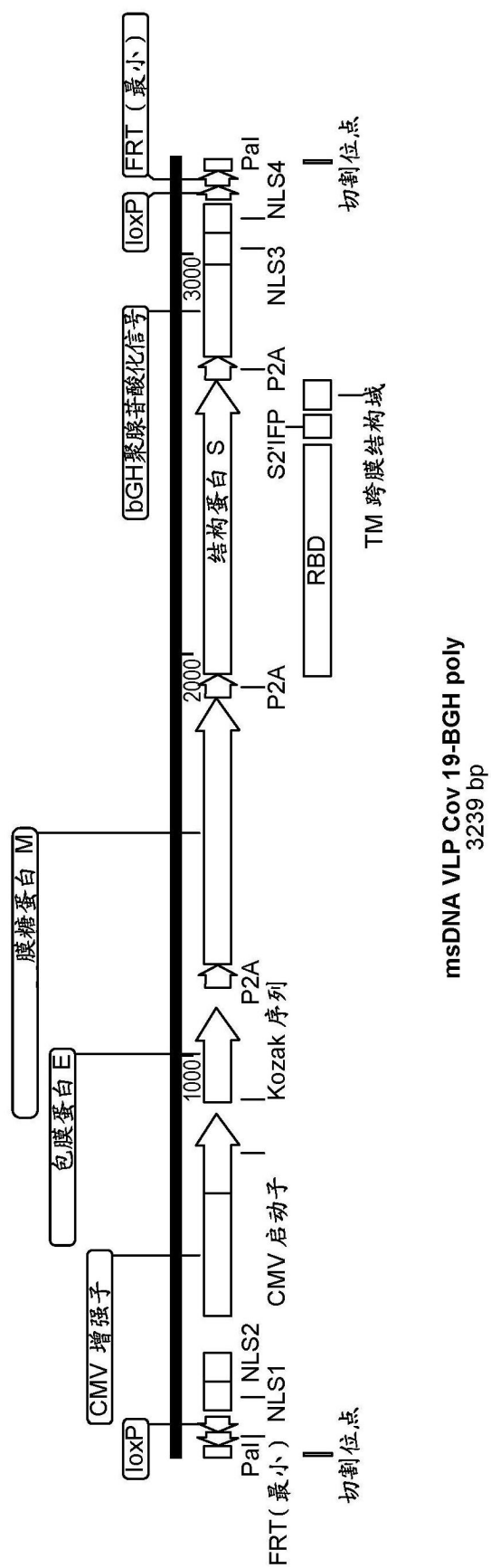


图4

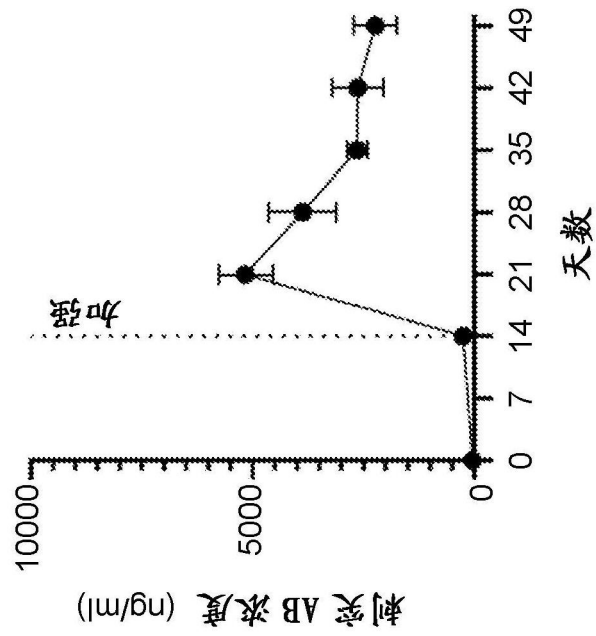


图5A

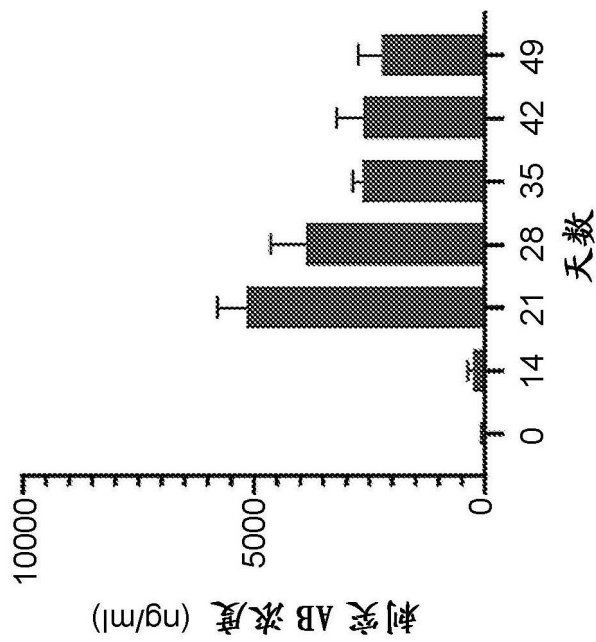


图5B

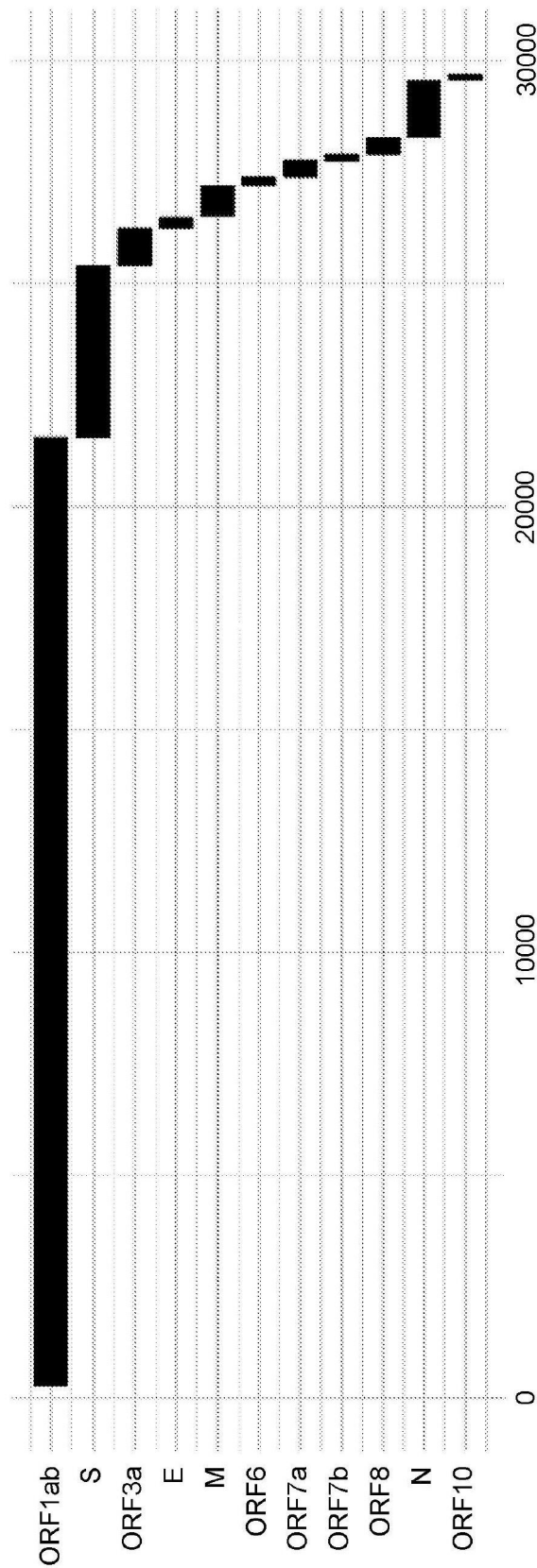


图6A

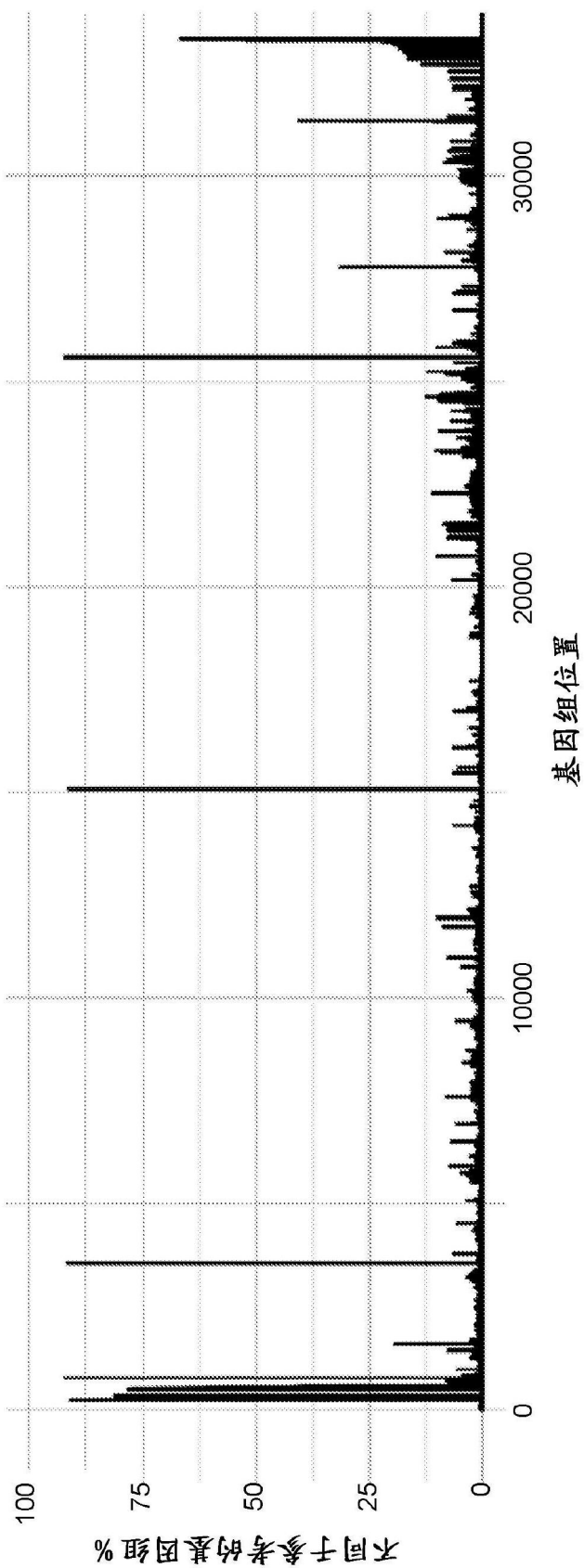


图6B

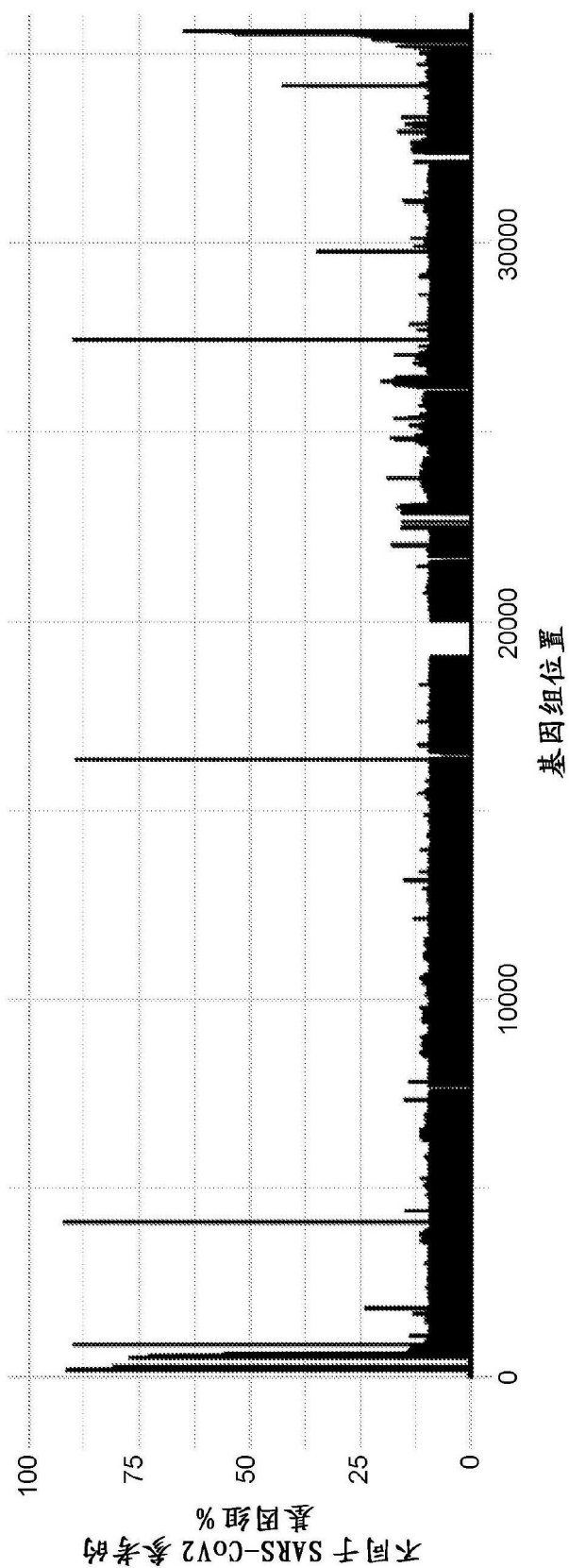


图7

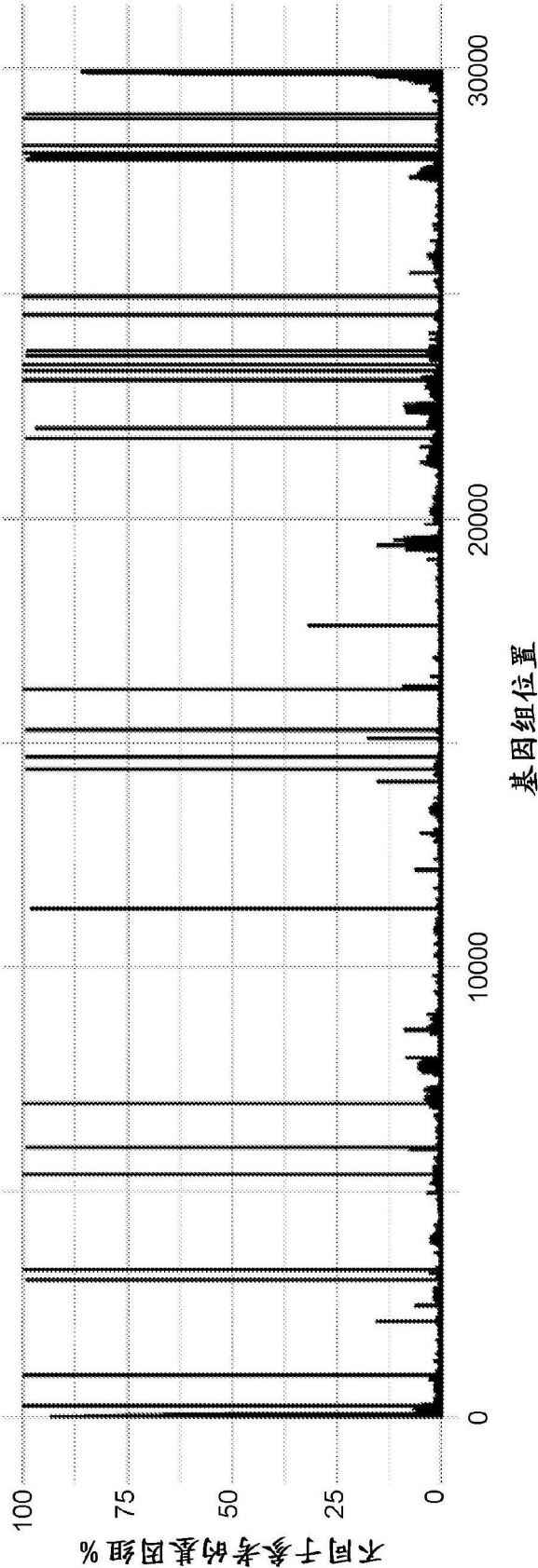


图8A

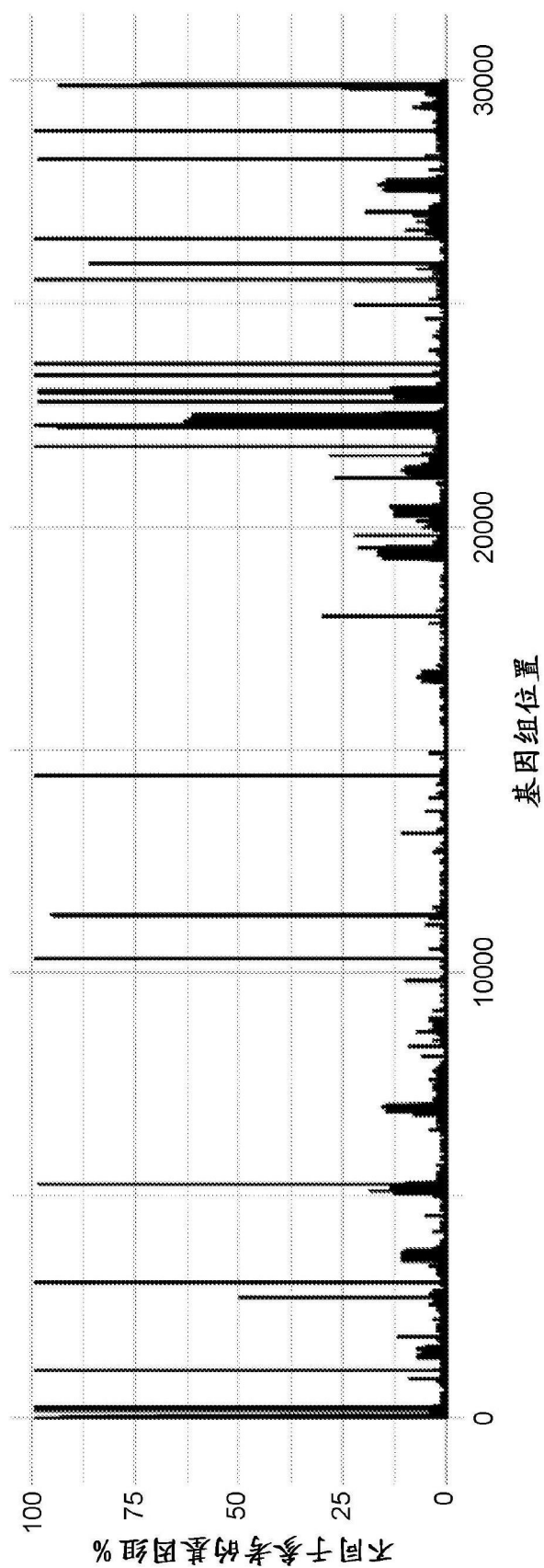


图8B

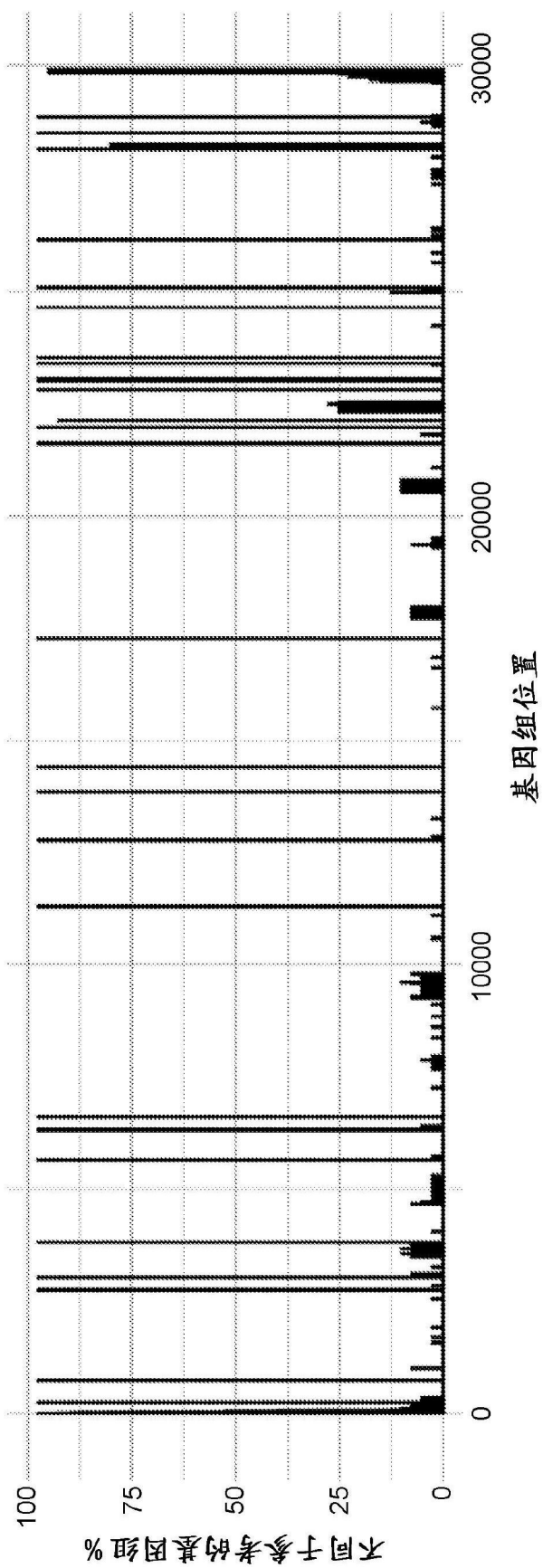


图8C

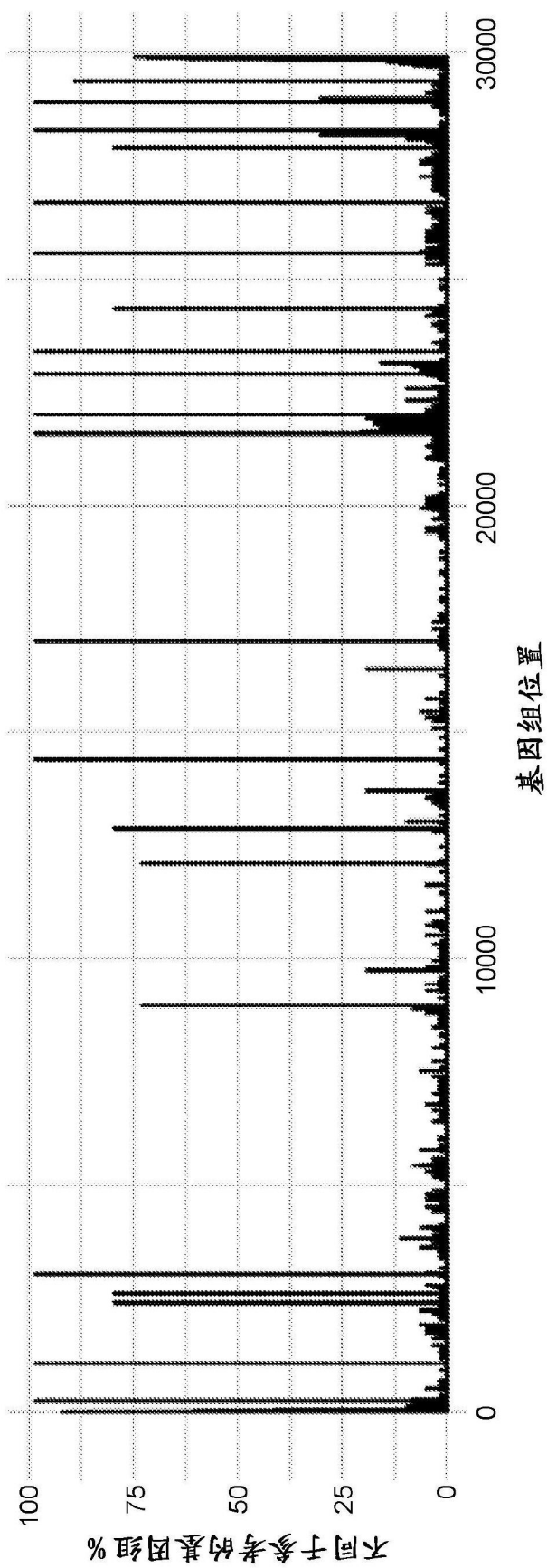


图8D

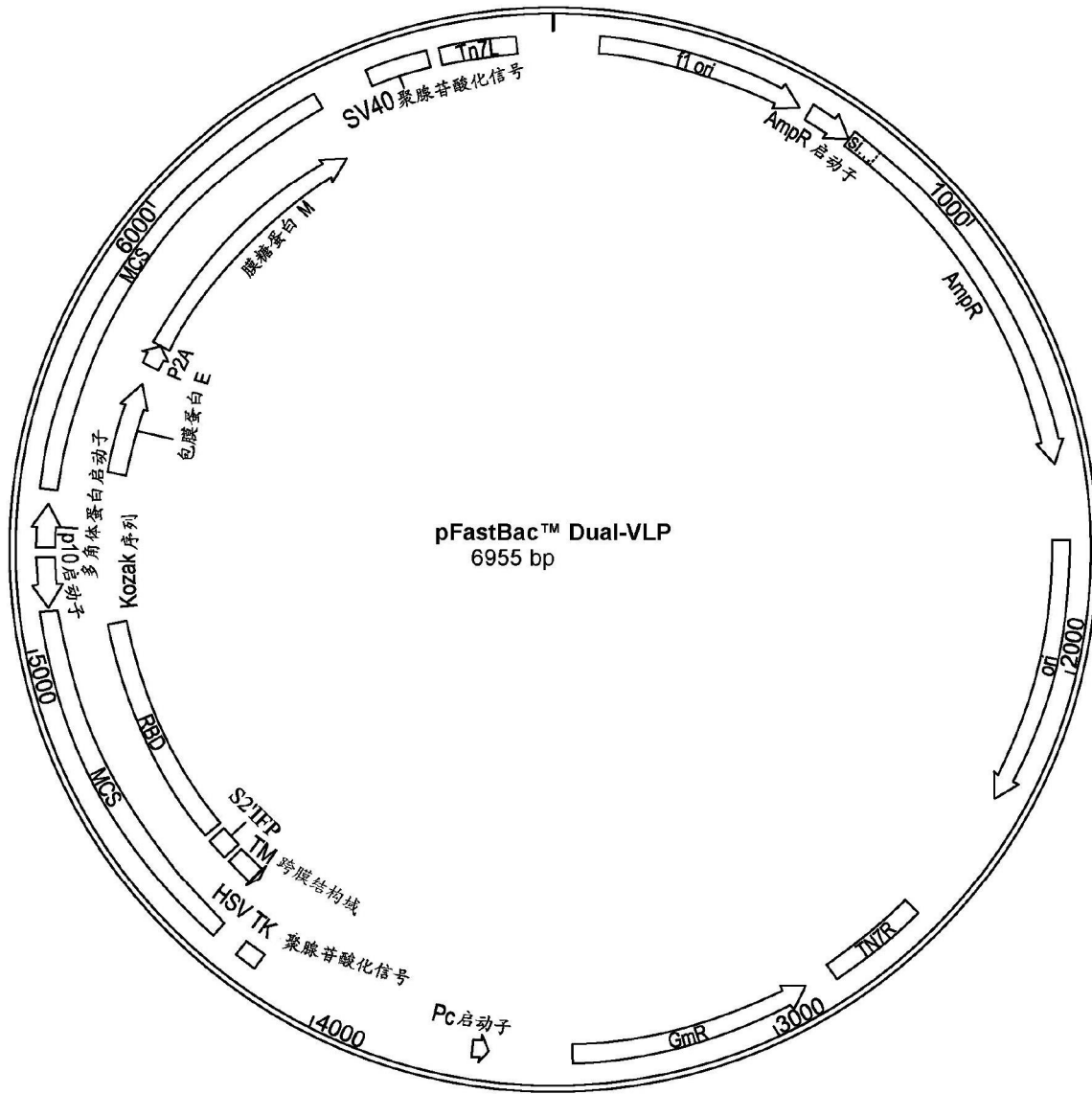


图9