

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200580026090.0

[51] Int. Cl.

C07K 14/165 (2006.01)

A61K 39/215 (2006.01)

C07K 16/10 (2006.01)

C12Q 1/70 (2006.01)

G01N 33/569 (2006.01)

A61K 39/42 (2006.01)

[43] 公开日 2007年9月19日

[11] 公开号 CN 101039955A

[51] Int. Cl. (续)

A61K 48/00 (2006.01)

[22] 申请日 2005.6.3

[21] 申请号 200580026090.0

[30] 优先权

[32] 2004.6.4 [33] US [31] 10/860,641

[32] 2004.6.10 [33] US [31] 60/578,348

[86] 国际申请 PCT/EP2005/006512 2005.6.3

[87] 国际公布 WO2005/118813 英 2005.12.15

[85] 进入国家阶段日期 2007.2.1

[71] 申请人 巴斯德研究院

地址 法国巴黎

共同申请人 香港大学 - 巴斯德研究中心有限公司

[72] 发明人 拉尔夫·阿尔特迈尔

贝亚特丽斯·纳尔 - 罗吉耶 陈澈曼

弗朗索瓦·基恩 甘耀永 肖雨岚

谢空山 伊莎贝尔·斯塔罗波利

让 - 克洛德·马努圭拉

[74] 专利代理机构 永新专利商标代理有限公司

代理人 林晓红

权利要求书 6 页 说明书 31 页 序列表 21 页
附图 29 页

[54] 发明名称

与 SARS 冠状病毒刺突蛋白相关的核酸、多肽、表达方法和免疫原性组合物

[57] 摘要

本发明披露了与 SARS 冠状病毒 (SARS CoV) 刺突蛋白相关的核酸分子、多肽、免疫原性组合物、疫苗和制备及使用所述核苷酸和所编码的多肽的方法。

1. 一种纯化的核酸分子，其包含 SEQ ID NO: 2 (Spike-Pasteur-modif)、SEQ ID NO: 3 (Spike-HKU-PRC)、或 SEQ ID NO: 6。

2. 一种纯化的核酸分子，其编码一种氨基酸序列，所述氨基酸序列包含 SEQ ID NO: 4 或 SEQ ID NO: 7 的序列，其中所述纯化的核酸分子与 SEQ ID NO: 1 相比显示出增强的刺突蛋白的表达。

3. 一种纯化的核酸分子，其与包含 SEQ ID NO: 3 或 SEQ ID NO: 6 的核酸序列的变性的双链 DNA 的任一链在高严格性条件下杂交。

4. 权利要求 3 的纯化的核酸分子，其中所述纯化的核酸分子与 SEQ ID NO: 1 相比显示出增强的刺突蛋白的表达。

5. 权利要求 4 的纯化的核酸分子，其包含至少一种负顺式作用信号的置换，且其中所编码的所述刺突蛋白的多肽序列没有改变。

6. 权利要求 5 的纯化的核酸分子，其中所述负顺式作用信号包含以下一组中的至少一种：

- (a) 富含 AU 的 RNA 不稳定性基序；
- (b) 重复序列；
- (c) 二级序列段(secondary stretch)；
- (d) 剪接供体和受体位点；和
- (e) 内部 poly (A)位点。

7. 权利要求 6 的纯化的核酸分子，其中所述纯化的核酸分子还包含至少一种额外的表达增强序列。

8. 权利要求 7 的纯化的核酸分子，其中所述额外的表达增强序列包含以下一组中的至少一种：

- (a) Kozak 共有序列；和
- (b) 额外的 STOP 密码子。

9. 权利要求 4 的纯化的核酸分子，其中密码子的使用已经根据中国

仓鼠(*Cricetulus griseus*)的偏好进行优化。

10. 权利要求 4 的纯化的核酸分子,其中所述纯化的核酸分子的编码所述刺突蛋白的部分与 SEQ ID NO: 1 相比,其 GC 含量百分比增加至少大约 10%。

11. 权利要求 7 的纯化的核酸分子,其中所述至少一种负顺式作用信号的置换和其中所述至少一种额外的表达增强序列不包括:

- (a) 内部 TATA 盒、chi 位点、和核糖体进入位点;
- (b) 富含 AT 或富含 GC 的序列段;
- (c) 重复序列和 RNA 二级结构; 和
- (d) 剪接供体和受体位点。

12. 一种重组载体,其指导选自权利要求 1 至 3 中任一项的纯化的核酸分子的核酸分子的表达。

13. 一种纯化的多肽,其包含 SEQ ID NO: 4、SEQ ID NO: 5、或 SEQ ID NO: 7。

14. 一种纯化的多肽,其由选自权利要求 1 至 3 中任一项的纯化的核酸分子的核酸分子编码。

15. 权利要求 14 的纯化的多肽,其中所述纯化的多肽包含高甘露糖 EndoH 敏感性 N-聚糖。

16. 纯化的抗体,所述抗体结合权利要求 14 的多肽。

17. 权利要求 16 的纯化的抗体,其中所述抗体是单克隆抗体。

18. 权利要求 16 的纯化的抗体,其中所述抗体包含中和抗体。

19. 用权利要求 12 的载体转染或者转导的宿主细胞。

20. 权利要求 19 的宿主细胞,其中所述宿主细胞选自 293T 细胞、BHK 细胞、和 FRHK4 细胞。

21. 增强由核酸分子表达 SARS CoV 刺突多肽的方法,包括降低所述核酸分子中的负顺式作用信号的数量,其中负顺式作用信号的数量降低不改变所述 SARS CoV 刺突多肽的序列。

22. 权利要求 21 的方法，其中所述负顺式作用信号包含以下一组中的至少一种：富含 AU 的 RNA 不稳定性基序；重复序列；二级序列段；剪接供体和受体位点；和内部 poly (A)位点。

23. 权利要求 22 的方法，其中所述方法还包括引入额外的信号而不改变所述 SARS CoV 刺突多肽的序列。

24. 权利要求 23 的方法，其中所述额外的信号包含 Kozak 共有序列和额外的 STOP 密码子中的至少一种。

25. 权利要求 24 的方法，还包括根据中国仓鼠的偏好优化密码子的使用。

26. 权利要求 24 的方法，还包括将刺突核苷酸的编码区内的 GC 含量增加至少大约 10%。

27. 一种分离的免疫复合物，其包含 SARS CoV 刺突多肽和特异性识别所述多肽的抗体。

28. 一种分离的免疫复合物，其包含 SARS CoV 刺突多肽和特异性识别所述多肽的抗体，其中所述抗体是针对权利要求 14 的纯化的多肽产生的。

29. 一种检测 SARS 病毒感染的方法，其中所述方法包括提供包含疑似感染了 SARS 病毒的生物学材料的组合物，以及测定刺突多肽的存在。

30. 权利要求 29 的方法，其中通过电泳或者通过使用与所述刺突多肽发生免疫学反应的抗体的免疫测定来测定所述刺突多肽。

31. 一种检测 SARS 病毒感染的方法，其中所述方法包括提供包含疑似感染了 SARS 病毒的生物学材料的组合物，以及测定刺突多肽的存在，其中所述抗体是针对权利要求 14 的纯化的多肽产生的。

32. 用于检测是否存在结合包含 SARS CoV 刺突多肽的抗原的抗体的体外诊断方法，其中所述方法包括将所述抗原与一种生物学液体相接触，接触的时间和条件足以使得所述抗原与生物学液体中的抗体形成抗原抗体复合物，以及检测所述复合物的形成。

33. 权利要求 32 的方法,其还包括测量所述抗原抗体复合物的形成。

34. 权利要求 33 的方法,其中通过基于蛋白印迹技术、ELISA、间接免疫荧光测定法、或免疫沉淀测定法的免疫测定法来检测抗原抗体复合物的形成。

35. 用于检测是否存在结合包含权利要求 14 的纯化的多肽的抗原的抗体的体外诊断方法,其中所述方法包括将所述抗原与一种生物学液体相接触,接触的时间和条件足以使得所述抗原与生物学液体中的抗体形成抗原抗体复合物,以及检测所述复合物的形成。

36. 用于检测是否存在结合 SARS CoV 刺突多肽或其混合物的抗体的诊断试剂盒,所述试剂盒包括包含 SARS CoV 刺突多肽或 SARS CoV 刺突多肽的混合物的抗原,以及用于检测所述抗原与抗体之间形成免疫复合物的试剂,其中所述试剂的量足以进行所述检测。

37. 用于检测是否存在结合 SARS CoV 刺突多肽或其混合物的抗体的诊断试剂盒,所述试剂盒包括包含权利要求 14 的纯化的多肽的抗原,以及用于检测所述抗原与抗体之间形成免疫复合物的试剂,其中所述试剂的量足以进行所述检测。

38. 一种免疫原性组合物,其包含至少一种 SARS CoV 刺突多肽以及药用可接受的载体,所述 SARS CoV 刺突多肽的量足以在体内诱导免疫原性或保护性应答。

39. 权利要求 38 的免疫原性组合物,其中所述组合物包含中和量的至少一种 SARS CoV 刺突多肽。

40. 权利要求 38 的免疫原性组合物,还包含 Alum 佐剂。

41. 一种免疫原性组合物,其包含权利要求 14 的纯化的多肽以及药用可接受的载体,所述多肽的量足以在体内诱导免疫原性或保护性应答。

42. 用权利要求 41 的免疫原性组合物处理宿主的方法,包括给宿主施用所述免疫原性组合物,所施用的免疫原性组合物的量足以在体内诱导免疫原性或保护性应答。

43. 权利要求 42 的方法, 其中通过腹腔内途径施用所述免疫原性组合物。

44. 权利要求 42 的方法, 其中所述免疫原性组合物还包括 Alum 佐剂。

45. 权利要求 42 的方法, 其中所述免疫原性组合物按照一剂量方案施用, 所述剂量方案包括两次或者多次施用 2-20 μg 的 SARS CoV 刺突肽。

46. 一种抗 SARS CoV 的疫苗组合物, 其包括权利要求 14 的多肽。

47. 一种抗 SARS CoV 的免疫接种方法, 其包括给有需要的动物施用权利要求 46 的疫苗组合物。

48. 权利要求 47 的方法, 其中所述免疫接种方法诱导增强的粘膜 IgA 和 IgG 抗体。

49. 权利要求 48 的方法, 其中所述疫苗组合物还包括 Alum 佐剂。

50. 权利要求 49 的方法, 其中所述疫苗组合物按照一剂量方案施用, 所述剂量方案包括两次或者多次施用 2-20 μg 的 SARS CoV 刺突肽以及 Alum 佐剂。

51. 一种检测是否存在 SARS CoV 的方法, 其包括:

(1) 将疑似含有 SARS CoV 病毒遗传物质的样品与至少一种核苷酸探针相接触, 和

(2) 检测所述核苷酸探针与样品中的病毒遗传物质之间的杂交, 其中所述核苷酸探针互补于 SEQ ID NO: 1、SEQ ID NO: 2、SEQ ID NO: 3、或 SEQ ID NO: 6 的纯化核酸的全长序列。

52. 一种质粒, 其保藏于 C.N.C.M, 保藏号为 I-3221、I-3222、或 I-3223。

53. 由权利要求 52 的质粒编码的 SARS CoV 刺突多肽。

54. 编码 SARS CoV 刺突多肽的片段的多核苷酸, 其与 SEQ ID NO: 1、SEQ ID NO: 2、SEQ ID NO: 3、或 SEQ ID NO: 6 相比具有至少一种突变。

55. 权利要求 54 的核苷酸序列或多核苷酸的片段, 其包含至少 10 个连续的核苷酸且最多 150 个连续的核苷酸。

56. 一种多核苷酸组合物, 其至少包含权利要求 54 或 55 的核苷酸序列。

57. 权利要求 1-11、13-15 和 53-55 中任一项的多肽或多核苷酸, 其能够诱导抗 SARS 感染的 T 细胞应答。

与 SARS 冠状病毒刺突蛋白相关的核酸、多肽、 表达方法和免疫原性组合物

相关申请

本申请要求申请日为 2004 年 6 月 4 日的美国专利申请 No. 10/860,641 和申请日为 2004 年 6 月 10 日的美国临时专利申请 No. 60/578,348 的优先权，在此通过引用将上述申请并入本申请。2005 年 5 月 23 日提交了将美国专利申请 No. 10/860,641 变更为临时申请的请求。

技术领域

本发明涉及纯化的和分离的核酸、多肽、纯化的和分离的多肽、编码此类多肽的核酸、产生重组形式的此类多肽的方法、针对这些多肽所产生的抗体、以及此类核酸和多肽在诊断方法、试剂盒、免疫原性组合物、疫苗或抗病毒治疗中的用途。

背景技术

2002 年，中国南方的广东省出现了一种新的感染性疾病，称为严重急性呼吸综合征 (SARS)。SARS 在 29 个国家中发生传播，据报道有 8,098 人受累，并造成 774 人死亡 (Stadler et al.)。虽然已经通过严格的隔离检疫措施遏制了 SARS 疫情，但并不知道 SARS 何时或是否将在人群中再次出现。

SARS 的主要特征为流感样综合征，包括高热超过 100.4°F、肌痛、干咳呼吸困难、低淋巴细胞血症、以及胸部 X 线上的浸润 (Stadler et al.)。所有病例中有 38% 因出现了导致急性呼吸障碍的肺炎而需要人工呼吸机。总体死亡率大约为 10%，但随年龄的不同而变化明显，因为 SARS

在儿童人群中似乎较轻，但在年长者中的死亡率则高达 50%。

引起 SARS 的是由一种以前未知的冠状病毒 (CoV)，冠状病毒是一多样化的大的、有包膜病毒群，在人和动物中引起呼吸道和肠道疾病。SARS CoV 分离自接种了来自患者的临床标本的 FRhK-4 细胞和 Vero E6 细胞，而接种了这种病毒的猕猴出现了类似于在 SARS 的人类病例中所观察到的那些症状。迄今为止，已经分离并测序了超过 30 种不同的 SARS CoV。

SARS CoV 含有大约 30 kB 的 RNA 基因组(Accession No. AY310120)，并具有冠状病毒的许多典型的特征。核苷酸 1-72 含有推定的 RNA 前导序列，其位于跨越 192 个核苷酸的非翻译区(UTR)之前。位于 UTR 下游的是两个重叠的可读框，它们覆盖了大约 2/3 的基因组 (核苷酸 265-21485)，并编码蛋白酶以及复制和转录所需的蛋白质 (综述可参见 Stadler et al., 2004)。基因组余下所 3'部分编码 4 种结构蛋白，这些蛋白在所有 CoV 中均以相同的顺序排列：刺突(Spike)蛋白、包膜蛋白、膜糖蛋白、和核壳蛋白。SARS CoV 基因组的结构蛋白区域还编码称为“附属基因(accessory gene)”的额外的非结构蛋白。尽管 SARS CoV 基因组的总体组织结构类似于其他冠状病毒，但所编码的蛋白质的氨基酸保守性通常较低。

刺突蛋白形成冠状病毒所特有的大的表面突出物。刺突被大量糖基化，具有 1,255 个氨基酸，含有位于氨基端的球形头部、紧邻的是茎部、单一的跨膜区、以及短的胞浆尾部 (参见 Stadler et al.)。

尽管有报道认为 β -干扰素在体外干扰 SARS 病毒的复制，但现在还没有批准任何药物或疫苗。此外，在现有的抗病毒剂或者大的化学剂文库中大规模筛查潜在的复制抑制剂的工作目前还没有成果。实际上，在首诊时要确诊 SARS 也几乎是不可能的，因为现有检测方法的敏感性和特异性会随着发生接触或出现症状之后的时间推移而变化 (参见 Rainer et al.)。目前还没有任何简便、快速、准确的检测方法能够在发病第一周

内诊断 SARS，也没有能够在采集样品后的数小时内给出结果的方法。鉴于这些原因，十分需要进一步详细了解 SARS CoV 蛋白。这样的了解能够为治疗或者控制这种感染提供有效的手段，且有助于对人类的 SARS CoV 感染进行诊断。

发明内容

因此，本发明的目的在于满足现有技术中的上述需求。本发明包括一种纯化的核酸分子，其包含 SEQ ID NO: 2、SEQ ID NO: 3、或 SEQ ID NO: 6 的 DNA 序列。本发明还包括与上述序列互补的核酸分子，例如全长的互补序列。

本发明包括那些包含 SEQ ID NO: 2、SEQ ID NO: 3、或 SEQ ID NO: 6 的 DNA 序列的双链核酸分子和编码 SEQ ID NO: 4 或 SEQ ID NO: 7 的氨基酸序列的纯化的核酸分子。本发明既包括单链的也包括双链的 RNA 和 DNA 核酸分子。这些分子可用于检测本发明所包括的单链和双链 RNA 和 DNA。双链 DNA 探针能够检测与所述核酸分子的任一链等价的核酸分子。

那些在高严格性条件下与包含 SEQ ID NO: 2、SEQ ID NO: 3、或 SEQ ID NO: 6 的 DNA 序列的变性的双链 DNA、或者与编码 SEQ ID NO: 4 或 SEQ ID NO: 7 的氨基酸序列的纯化的核酸分子杂交的纯化的核酸分子，包括在本发明的范围内。

本发明还包括通过体外诱变而自 SEQ ID NO: 1-3 和 6 衍生的纯化的核酸分子。体外诱变包括本领域已知的多种技术，包括但不限于，定点诱变、随机诱变和体外核酸合成。

本发明的核酸分子包括 DNA 和 RNA，这些核酸分子在此称为“刺突核酸(Spike nucleic acids)”或者“刺突 DNA(Spike DNA)”，由这些分子编码的氨基酸序列在此称为“刺突多肽”或者“刺突蛋白”。

本发明还包括因遗传密码简并性而自 SEQ ID NO: 1-3 和 6 简并得到

的纯化的核酸分子、作为刺突核酸的等位基因变体或者刺突核酸的同系物的纯化的核酸分子。

本发明包括与 SEQ ID NO: 1 相比显示出增强的刺突蛋白表达的纯化的核酸。

本发明还包括与 SEQ ID NO: 1 相比显示出增强的刺突蛋白表达的纯化的核酸，其中至少一种负顺式作用信号已经被置换，而这并未改变所编码的蛋白质的序列。本发明包括的负顺式作用信号包括但不限于富含 AU 的 RNA 不稳定性基序、重复序列、二级序列段(secondary stretch)、剪接供体和受体位点，和内部 poly (A)位点。

本发明还包括与 SEQ ID NO: 1 相比显示出增强的刺突蛋白表达的纯化的核酸分子，其中通过增加表达增强序列而增强表达。表达增强序列包括但不限于起始 ATG 上游的 Kozak 共有序列以及额外的 STOP 密码子。

本领域技术人员基于现有技术能够了解如何将 Kozak 共有序列置于合适的位置。

本发明还包括与 SEQ ID NO: 1 相比显示出增强的刺突蛋白表达的纯化的核酸分子，其中密码子的使用已经根据中国仓鼠(*Cricetulus griseus*)的偏好进行优化。

本发明还包括与 SEQ ID NO: 1 相比显示出增强的刺突蛋白表达的纯化的核酸分子，其中所述纯化的核酸分子的编码刺突蛋白的部分与 SEQ ID NO: 1 相比，其 GC 含量百分比增加至少大约 10%，且其中在可能的情况下已经避免出现 GC 含量极高(>80%)或者极低(<30%)的区域。

本发明还包括与 SEQ ID NO: 1 相比显示出增强的刺突蛋白表达的纯化的核酸分子，其中所述至少一种负顺式作用信号的置换和其中所述至少一种额外的表达增强序列不包括：内部 TATA 盒、chi 位点、和核糖体进入位点；富含 AT 或者富含 GC 的序列段(sequence stretches)；重复序列和 RNA 二级结构；和剪接供体和受体位点。

本发明还包括由这些核酸分子编码的纯化的多肽，包括所述纯化的

多肽的糖基化和非糖基化形式。

本发明还包括指导这些核酸分子表达的重组载体以及用这些载体转化或者转染的宿主细胞。

本发明包括与刺突多肽结合的纯化的多克隆或单克隆抗体，例如中和抗体。

本发明还包括制备刺突多肽的方法，包括在促进表达的条件下培养宿主细胞，并自培养基中收集所述多肽。本发明尤其包括在动物细胞中表达刺突多肽。

本发明还包括标记的刺突多肽。优选地，所述标记的多肽是纯化的形式。同样优选的是，未标记的或者标记的所述多肽能够被含有抗刺突多肽抗体的人的体液免疫学识别。所述多肽可以用例如选自以下一组的免疫测定标记物进行标记：放射性标记物、酶性标记物、荧光标记物、化学发光标记物、以及发色团。

本发明还提供了本发明的刺突多肽与识别所述多肽的抗体之间的免疫复合物。所述免疫复合物可以用选自以下一组的免疫测定标记物进行标记：放射性标记物、酶性标记物、荧光标记物、化学发光标记物、以及发色团。

此外，本发明提供用于检测 SARS CoV 感染的方法。所述方法包括提供包含疑似感染了 SARS CoV 的生物学材料的组合物，并测定 SARS CoV 刺突多肽的存在。所述多肽通常是通过电泳或者通过使用与本发明的刺突多肽具有免疫学反应性的抗体的免疫测定法进行测定。

本发明还提供体外诊断方法，其用于检测是否存在结合包含本发明的刺突多肽的抗原的抗体。所述方法包括将所述抗原与生物学液体相接触，接触的时间和条件足以使得所述抗原与生物学液体中的抗体形成抗原抗体复合物，然后检测所述复合物的形成。检测步骤可进一步包括测量所述抗原抗体复合物的形成。优选地通过基于蛋白印迹技术、ELISA (酶联免疫吸附测定法)、FACS、间接免疫荧光测定法、或免疫沉淀测定法的

免疫测定法测量抗原抗体复合物的形成。

本发明还包括诊断试剂盒，其用于检测是否存在与本发明的刺突多肽结合的抗体，所述试剂盒含有包含所述刺突多肽的抗原，以及用于检测所述抗原与抗体之间形成免疫复合物的试剂。其中所述抗原和所述试剂的量足以进行所述检测。

本发明还提供一种免疫原性组合物，其包含本发明的刺突多肽或其混合物以及药用可接受的载体，所述刺突多肽或其混合物的量足以在体内诱导免疫原性或保护性应答。免疫原性组合物可含有 Alum 佐剂。本发明的疫苗组合物包含中和量的刺突多肽以及药用可接受的载体。

因此，本发明的多肽可作为诊断组合物的一部分，以检测是否存在抗与 SARS CoV 相关的抗原性蛋白质的抗体。

此外，刺突多肽还可用于产生检测是否存在与 SARS CoV 相关的抗原性蛋白质的抗体。

本发明的多肽还可用于产生中和抗体，这些中和抗体要么灭活病毒，要么降低病毒在体内的活力，要么抑制或者阻断病毒的复制。当将本发明的多肽用于免疫组合物或者预防接种组合物以便在受者宿主体内激活免疫应答的 B 细胞分支或者诱导细胞毒 T 淋巴细胞应答(CTL)时，这种引发病毒中和抗体的能力是特别重要的。

本发明还涉及用于对人和哺乳动物进行抗 SARS CoV 免疫接种的疫苗组合物，其包含上述免疫原性组合物并组合一种或者多种药用可接受的赋形剂(例如盐水缓冲液)，并任选地组合至少一种佐剂，例如氢氧化铝或者属于胞壁酰肽家族的化合物。

本发明还包括用于检测是否存在 SARS CoV 的方法，包括：

(1) 将疑似含有 SARS CoV 的病毒遗传物质的样品与至少一种核苷酸探针相接触，和

(2) 检测所述核苷酸探针与所述样品中的病毒遗传物质之间的杂交，其中所述核苷酸探针互补于本发明的纯化的刺突核酸的全长序列。

本发明的其他特征和优势，其部分内容将在随后的说明书中进行阐述，而部分内容基于说明书将是显而易见的，或者可通过实践本发明而得知。本发明的目的和优越之处可通过如所附权利要求书中所指出的那些元素和组合而实现或获得。

应当理解的是，无论是前面的一般性描述还是随后的详细描述，均仅用于示例和解释本发明，而无意于对要求保护的发明进行限制。

附图说明

所附附图作为说明书的一部分，其举例说明了本发明的一些实施方式，这些附图与文字描述共同用于解释本发明的原则。

图 1 所示为 Spike-HKU-PRC 在转染的 293T 细胞中的表达。泳道 1 代表以 pcDNA-Spike-Pasteur 转染的细胞，泳道 2 代表以 pcDNA-HKU-PRC 转染的细胞，泳道 3 代表以 SFV-Spike-Pasteur-modif 转染的细胞，泳道 4 是空的，泳道 5 代表来自转染的 BHK 细胞的纯化的刺突(Spike)。

图 2 所示为用于 RNA 和蛋白质免疫接种的、具有 Flag 肽序列的 SARS CoV 刺突蛋白序列(SEQ ID NO: 5)，以标准的单字母缩写表示。阴影部分为相应于 SARS CoV 刺突蛋白的序列，下划线部分是包括 Flag 肽的序列。该代表序列在 Semliki Forest Virus (SFV) 载体中表达。

图 3 是以标记的 SFV-Spike 脉冲-追加式(pulse-chase)感染的 BHK 细胞经 M2 (Flag)抗体免疫沉淀后的 SDS-PAGE 图。在追加后的指定时间点收集细胞。“*”代表高甘露糖 N-聚糖 EndoH 敏感性刺突，“O”代表复合 N-聚糖 EndoH 抵抗性刺突，而“#”高甘露糖 N-聚糖 EndoH 敏感性去糖基化的刺突。

图 4 所示为感染了 SFV Spike 的 BHK 细胞质膜上的 Spike 的表达情况。以 M2 抗体对刺突蛋白进行标记，同时以 Erp72 单克隆抗体对内质网(ER) 进行染色。

图 5 是 Western 印迹分析,其显示 SARS CoV 蛋白结合 sACE2 受体。将以 Spike (泳道 1 和 4) 包被的 M2 珠或作为对照的以 BAP 包被的 M2 珠 (泳道 3 和 6) 与 sACE2 进行温育并进行 SDS 凝胶电泳,然后以抗 ACE2 抗体或者作为对照以 Mab M2 进行 Western 印迹。尽管反应中既存在 Spike 也存在 BAP 蛋白 (泳道 4 和 6),但仅有 Spike 结合 ACE2 (泳道 1 和 3)。

图 6 显示以重组的免疫纯化的 SARS CoV 刺突蛋白进行免疫接种的小鼠产生抗重组 Spike 的抗体。使用 1:100 稀释的混和小鼠血清 (n=5) 进行 Western 印迹。泳道 1 和 2 分别代表使用来自对照动物(CTRL)和接种 Spike 的动物(VACC)的免疫前血清进行的 Western 印迹。泳道 3 和 4 分别代表使用来自 CTRL 和 VACC 组的第 34 天的血清进行的 Western 印迹,而泳道 5 和 6 分别代表使用来自 CTRL 和 VACC 组的第 42 天的血清进行的 Western 印迹。泳道 7 代表人 SARS 患者的血清,泳道 8 代表一种商品化的来自以刺突蛋白免疫的兔的血清(1:50 稀释)。

图 7 显示以重组的免疫纯化的 SARS CoV 刺突蛋白进行免疫接种的小鼠产生抗重组 Spike 的抗体。使用 1:50 稀释的混和小鼠血清 (n=5) 进行 FACS 分析。右侧是作为对照的 SARS 病人血清。

图 8 显示以重组的免疫纯化的 SARS CoV 刺突蛋白进行免疫接种的小鼠产生抗 SARS CoV 的抗体。使用 1:50 稀释的混和小鼠血清 (n=5) 在感染了 SARS CoV 的 FRHK4 细胞或者假感染的 FRHK4 细胞上进行免疫荧光分析。

图 9 (A)、9 (B)、和 9 (C) 显示了 Spike-Pasteur 的核酸序列 (SEQ ID NO: 1)。各个 Spe I 位点以下划线标示,阴影标示的是进行取代以形成 Spike-Pasteur-modif 的核酸残基。

图 10(A)、10(B)、和 10(C) 显示了 Spike-Pasteur-modif 的核酸序列 (SEQ ID NO: 2)。阴影标示的突变造成 Spe I 位点自 Spike-Pasteur 中消除。

图 11 (A)、11 (B)、11 (C)、11 (D)、11 (E)、和 11 (F) 显示了

Spike-HKU-PRC 的核酸序列 (SEQ ID NO: 3) 及其互补链。阴影标示的核酸序列编码刺突多肽。图 11 (A)、11 (B)、11 (C)、11 (D)、11 (E)、和 11 (F) 还显示了融合于 Flag 肽的 Spike 的氨基酸序列 (SEQ ID NO: 4)。STOP 密码子以星号标示。

图 12 (A)和 12(B) 显示了 Spike-HKU-PRC 内的编码 SARS CoV 刺突多肽的优化的核酸序列 (SEQ ID NO: 6)。SEQ ID NO: 6 与 SEQ ID NO: 3 的不同之处在于其不含编码 Flag 肽或上游或者下游序列的序列。

图 13 所示为由 Spike-HKU-PRC 编码的 SARS CoV 刺突多肽 (SEQ ID NO: 7) 的序列。

图 14 所示为本发明的质粒, 即标记的 040078pPCR-Script, 其含有编码 Spike-HKU-PRC 的序列。从合成的寡核苷酸组装成合成基因 040078。通过 KpnI 和 SacI 限制性位点将该片段克隆入 pPCR-Script Amp (Stratagene, LaJolla, CA, USA)。

图 15 所示为本发明的质粒, 即标记的 040086pcDNA3.1(+), 也称为 040078pcDNA3.1(+), 其含有克隆入 pcDNA3.1 (Invitrogen) 的 BamHI 限制性位点的编码 Spike-HKU-PRC 的序列。

图 16 所示为用于免疫接种的刺突蛋白的纯度, 并显示了银染着色的 SDS-PAGE 凝胶。样品包括: (M) 分子量标准; (1) 720 ng S 蛋白; (2) 360 ng S 蛋白; (3) 180 ng S 蛋白; 和 (4) 90 ng S 蛋白。标示了分子量, 箭头所示为复合糖基化单体刺突蛋白的位置(上方箭头)和高甘露糖单体刺突蛋白的位置(下方箭头)。

图 17 显示在以 TriSpike+ Alum 免疫的动物中血清抗体应答增强。分析了来自免疫接种的小鼠的血清与 TriSpike 的反应性。(A) 对照为 SARS 患者的高滴度中和抗血清、抗 S1 的兔血清、和抗 FLAG 肽的 M2 单克隆抗体。对来自以 TriSpike 加 Alum 佐剂(组 A)或者以 TriSpike 不加 Alum 佐剂(组 B)免疫的小鼠(n=3)的混和血清进行 Western 印迹分析。在所指定的时间点收集血清并以 1:1000 稀释, 用于 Western 印迹分析。所有血清

均与 FLAG 标记的对照蛋白(BAP-FLAG)进行反应以评估针对 FLAG 标签的抗体产生。使用 HRP 缀合的山羊抗小鼠、人、或兔的 IgG 多克隆抗体来检测免疫复合物。(B) 以 FRhk4 细胞在体外分析来自免疫接种的小鼠的血清的抗 SARS CoV 感染的中和活性。来自仅接种了 TriSpike 的小鼠的血清的中和活性(自第 49 天至第 87 天)迅速降低,但来自接种了 TriSpike + Alum 的小鼠的血清的中和活性在同一对比时间内保持稳定。

图 18 显示以 TriSpike + Alum 免疫接种小鼠诱导粘膜免疫应答。收集来自免疫小鼠(TriSpike 或者 TriSpike + Alum)的排泄物和鼻灌洗液样品并分析其与 TriSpike 的反应性(A-B)。以抗 FLAG 肽的 M2 单克隆抗体作为对照。(A)在第 44 天收集免疫接种小鼠的排泄物样品并按照 1:500 进行稀释,用于 Western 印迹分析。以 HRP 缀合的山羊抗小鼠 IgG 或 IgA 多克隆抗体检测免疫复合物。(B)所示为与(A)基本相同的实验,不同之处仅在于以在第 65 天收集自免疫接种小鼠的混合鼻灌洗液样品进行 Western 印迹分析。将鼻灌洗液样品按照 1:25 稀释,用于 Western 印迹分析。(C)收集来自免疫接种小鼠的排泄物样品,并以 FRhk4 细胞在体外分析抗 SARS CoV 感染的中和活性。仅在第三次免疫接种之后检测到弱中和活性。分析了来自免疫接种小鼠的鼻灌洗液样品,但在体外没有发现可观测到的中和活性水平。

图 19 显示了 TriSpike 在 Golden Syrian 仓鼠中的免疫原性。分析来自接种了所示浓度的 TriSpike+Alum 的仓鼠和来自对照仓鼠的血清与 TriSpike 的反应性和中和情况。(A)使用 FACS 来分析免疫血清与天然的 TriSpike 蛋白的反应性。将来自以 2、10、或者 20 μg 的 TriSpike 皮下免疫的仓鼠的血清(第 0、21 和 42 天)按照 1:100 稀释,与在质膜上表达 TriSpike 的活 BHK-21 细胞反应。使用 FITC 缀合的山羊抗仓鼠 IgG 多克隆抗体来鉴别免疫复合物。结果以 MFI (平均荧光强度) 值表示。MFI 值在第二次免疫后(剂量 2 之后)达到最大,并在第三次免疫后(剂量 3 之后)保持稳定。(B)中和活性得自在 FRhk4 细胞上进行的 SARS CoV 微中和测

定(microneutralization assay) (100TCID₅₀/孔终末)。

发明详述

已经发现了用于增强 SARS CoV 刺突蛋白表达的优化 DNA 序列, 包括 Spike-Pasteur-modif (SEQ ID NO: 2) 和 Spike-HKU-PRC (SEQ ID NO: 3 和 6)。

本发明的核酸序列包括分离的 DNA 和 RNA 序列, 所述 DNA 和 RNA 序列在中等或者高严格性条件下与在此所述的 SEQ ID NO: 2、3 和 6 杂交, 并编码刺突多肽。在此, 中等严格性条件是本领域人员已知的且记载在 Sambrook et al. *Molecular Cloning: A Laboratory Manual*, 2 ed. Vol. 1, pp. 1.101-104, Cold Spring Harbor Laboratory Press, (1989), 其包括对硝酸纤维素滤膜使用预洗涤溶液 5 X SSC、0.5% SDS、1.0 mM EDTA (pH 8.0), 杂交条件为 50%甲酰胺、6X SSC、42EC (或者其他类似的杂交溶液, 例如 Stark's 溶液、50% 甲酰胺, 42EC), 且洗涤条件为大约 60EC、0.5 X SSC、0.1 % SDS。高严格性条件被定义为: 杂交条件如上所述, 而洗涤条件为 68EC、0.2 X SSC、0.1 % SDS。本领域人员能够理解, 必要时可根据各种因素例如探针的长度来调整温度和洗涤溶液的盐浓度。

由这些新的核酸所编码的多肽在此称为“刺突多肽(Spike 多肽)”或者“刺突蛋白(Spike 蛋白)”。在此, 这些术语指的是一类多肽, 它们还包括具有 SEQ ID NO: 4 或 SEQ ID NO: 7 的氨基酸序列的蛋白质, 以及那些与此类氨基酸序列具有高度相似性(至少 90%相似性)的蛋白质和多肽, 并且所述的蛋白质和多肽具有免疫反应性。此外, “刺突多肽”和“刺突蛋白”还指在高严格性条件下与 SEQ ID NO: 3 或 SEQ ID NO: 6 的编码序列互补的核酸链杂交的核酸分子所编码的那些蛋白质。

术语“纯化的”在此是指刺突多肽基本上不与其他蛋白质或者多肽相结合, 例如是作为重组宿主细胞培养物的纯化产物, 或是作为非重组来源的纯化产物。术语“实质上纯化的”在此是指一种混合物, 其含有

刺突多肽，且若非存在能够使用特异性抗体而去除的已知的蛋白质，其基本上不与其他蛋白质或者多肽相结合，且所述实质上纯化的刺突多肽可用作抗原。

刺突多肽“变体”在此是指与天然的刺突多肽实质上同源的多肽，但由于一或多个缺失、插入或者取代而具有不同于天然刺突多肽的氨基酸序列。变体氨基酸序列优选地与天然刺突多肽的氨基酸序列具有至少80%的相同性，更加优选地至少90%的相同性。相同性百分比是可以确定的，例如可使用 Devereux et al. (Nucl. Acids Res. 12:387, 1984) 所述的 GAP 计算机程序 6.0 版比较序列信息而确定，该程序可自 University of Wisconsin Genetics Computer Group (UWGCG) 获得。GAP 程序使用 Needleman 和 Wunsch (J. Mol. Biol. 48:443, 1970) 的比对方法，Smith 和 Waterman (Adv. Appl. Math 2:482, 1981) 对该方法进行了修改。GAP 程序的优选默认参数包括：(1)用于核苷酸的一元比较矩阵(相同的值为 1，不相同的值为 0)，以及 Gribskov 和 Burgess, Nucl. Acids Res. 14: 6745, 1986 所述的加权比较矩阵，参见 Schwartz 和 Dayhoff, eds., Atlas of Protein Sequence and Structure, National Biomedical Research Foundation, pp. 353-358, 1979；(2)每一缺口的罚分为 3.0，且每一缺口中的每一符号有额外的 0.10 的罚分；以及(3)末端缺口无罚分。

变体可包含保守性取代的序列，保守性取代是指一种给定的氨基酸残基被具有相似生化特性的残基取代。保守性取代的实例包括脂肪族残基彼此间的置换，例如 Ile、Val、Leu、或 Ala 彼此之间的取代；或者是极性残基彼此之间的置换，例如 Lys 与 Arg 之间；Glu 与 Asp 之间；和 Gln 与 Asn 之间。其他此类保守性置换是已知的，例如具有类似疏水特性的整个区域的置换。本发明也包括天然存在的刺突多肽变体。此类变体的实例是因替代性 mRNA 剪接事件或因刺突多肽的蛋白酶裂解而产生的蛋白质。可归因于蛋白水解作用的变异包括例如在不同类型的宿主细胞内表达后出现的末端的差异，这是由于自刺突多肽蛋白水解去除了一或

多个末端氨基酸所致。可归因于移码的变异包括例如在不同类型的宿主细胞内表达后出现的末端的差异，这是由于不同的氨基酸所致。

如上所述，本发明提供分离并纯化的、或者均质的刺突多肽，可以是重组的也可以是非重组的。可用作抗原的天然刺突多肽的变体和衍生物可通过对编码天然刺突多肽的核苷酸序列进行突变而获得。可通过多种常规方法实现对天然氨基酸序列的改变。通过合成含有突变序列的寡核苷酸，其中突变序列两侧具有能够实现与天然序列的片段相连接的限制性位点，由此将突变引入特定的基因座。经连接后，所得到的重建序列编码具有所需氨基酸插入、取代或者缺失的类似物。

或者，可使用寡核苷酸指导的位点特异性诱变方法来提供一种改变的基因，其中可通过取代、缺失或者插入而改变其中预先确定的密码子。实现上述改变的示例性方法可参见 Walder et al. (Gene 42:133, 1986); Bauer et al. (Gene 37:73, 1985); Craik (BioTechniques, January 1985, 12-19); Smith et al. (Genetic Engineering: Principles and Methods, Plenum Press, 1981); Kunkel (Proc. Natl. Acad. Sci. USA 82:488, 1985); Kunkel et al. (Methods in Enzymol. 154: 367, 1987); 和美国专利 4,518,584 和 4,737,462, 所有这些文献均通过引用并入本申请。

在本发明的一个方面，刺突多肽可用于制备特异性结合刺突多肽的抗体。术语“抗体”包括多克隆抗体、单克隆抗体、其片段例如 F(ab')₂ 和 Fab 片段、以及任何重组方法产生的结合配偶体(binding partner)。如果抗体以高于或者等于大约 10^7 M^{-1} 的 K_a 值结合于刺突多肽，则所述抗体被认为能够发生特异性结合。可采用常规技术容易地确定结合配偶体或者抗体的亲和力，例如采用 Scatchard et al., Ann. N.Y Acad. Sci., 51:660 (1949) 所述的方法。可采用现有技术已知的方法从多种来源容易地产生多克隆抗体，例如来自马、牛、山羊、绵羊、狗、鸡、兔、小鼠或者大鼠。

本发明还包括衍生自 SEQ ID NO: 2-3 和 6 的核苷酸序列的分离的片段和寡核苷酸。本发明还包括由这些片段和寡核苷酸所编码的多肽。

由于已知遗传密码具有简并性，其中可有一个以上的密码子编码同一个氨基酸，因此不同于 SEQ ID NO: 2-3 和 6 所示序列的 DNA 序列仍可编码具有 SEQ ID NO: 7 的氨基酸序列的刺突多肽。此类变体 DNA 序列可通过沉默突变产生(例如，在 PCR 扩增过程中发生)，或者可以是对天然序列进行有目的的诱变的产物。

因此，本发明提供等价的分离的 DNA 序列，其编码刺突多肽，所述选自 DNA 序列：(a)包含 SEQ ID NO: 2-3 和 6 的核酸分子；(b)能够在高严格性条件下与 SEQ ID NO: 3 或 6 杂交的 DNA；(c)包含 SEQ ID NO: 2-3 和 6 的片段的核酸分子；和(d)因遗传密码子简并性而由(a)、(b)、或(c)的 DNA 产生的、且编码刺突多肽和刺突多肽片段的核酸分子。由此类核酸等价序列所编码的刺突多肽也属于本发明的范围内。

由与 SEQ ID NO: 3 或 6 等价的 DNA 编码的刺突多肽的实例包括但不限于上述包含灭活的 N-糖基化位点、蛋白酶加工位点、或者保守性氨基酸取代的刺突多肽片段和刺突多肽。

可采用熟知的方法制备含有编码刺突多肽的核酸序列的重组表达载体。表达载体包括可操纵地连接于合适的转录或者翻译调节核苷酸序列的 Spike DNA 序列，例如那些衍生自哺乳动物、微生物、病毒或者昆虫基因的调节核苷酸序列。调节序列的实例包括转录启动子、操纵子、或增强子、mRNA 核糖体结合位点、以及控制转录和翻译的起始和终止的适当序列。当调节序列与 Spike DNA 序列在功能上相关联时，则核苷酸序列是“可操纵地连接的”。因此，如果启动子核苷酸序列控制 Spike DNA 序列的转录，则启动子核苷酸序列是可操纵地连接于 Spike DNA 序列的。复制起始点通常赋予在所需宿主细胞中的复制能力，并可将用于鉴定转化体的选择基因掺入表达载体。

此外，还可将编码通常不与刺突多肽天然相关的合适的信号肽的序列掺入表达载体。

用于原核宿主细胞的表达载体通常包括一或多个可选择性表型标记

基因。可选择性表型标记基因例如是一种编码赋予抗生素抗性的蛋白质的基因或者是一种编码提供自养需求的蛋白质的基因。可用于原核宿主细胞的表达载体的实例包括那些来自已经商品化的质粒的表达载体。商品化的载体包括那些为表达蛋白质而特别设计的载体，包括 pMAL-p2 和 pMAL-c2 载体，它们用于表达融合于麦芽糖结合蛋白的蛋白质 (New England Biolabs, Beverly, MA, USA)。

包含优化的 SARS CoV Spike 基因的质粒的具体实例包括以下质粒：

pPCR-Script-040078	2004年6月8日保藏于C.N.C.M.，保藏号为I-3221；
pcDNA-Spike-HKUPRC-040086	2004年6月8日保藏于C.N.C.M.，保藏号为I-3222；
pcSFV-HKUPRC-040091	2004年6月8日保藏于C.N.C.M.，保藏号为I-3223。

常用于重组原核宿主细胞表达载体的启动子序列包括 β -内酰胺酶(青霉素酶)、乳糖启动子系统 (Chang et al., Nature 275:615, 1978; 和 Goeddel et al., Nature 281:544, 1979)、色氨酸(trp)启动子系统 (Goeddel et al., Nucl. Acids Res. 8:4057, 1980; 和 EP-A-36776)、tac 启动子 (Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, p. 412, 1982)。

用于表达刺突多肽的合适的宿主细胞包括原核细胞、酵母或者高等真核细胞。用于细菌、真菌、酵母、和哺乳动物细胞宿主的合适的克隆和表达载体可参见例如 Pouwels et al. Cloning Vectors: A Laboratory Manual, Elsevier, New York, (1985)。也可采用无细胞翻译系统，使用衍生自在此所述的 DNA 构建体的 RNA 来产生刺突多肽。

可以理解，本发明旨在包括上述蛋白质的分离的或者纯化的形式，无论其是使用在此所述的技术还是采用其他方法而获得的。在本发明的一个优选的实施方式中，刺突多肽基本上不含人体组织和人体组织成分、核酸、无关蛋白质和脂质、以及外来微生物例如细菌和病毒。可以理解，本发明还包括基本上具有所述同样生物学特性和免疫原性的等价蛋白质。因此，本发明旨在包括本发明的蛋白质的各种血清型变体。

根据本发明的刺突多肽的用途，可以对其进行标记。合适的标记物的实例为放射性标记物、酶标记物、荧光标记物、化学发光标记物、和发色团。用于标记本发明的蛋白质和糖蛋白的方法与标记免疫球蛋白所广泛采用的方法没有实质上的不同。如果采用针对本发明的抗原的标记的抗体，或者使用抗针对所述抗原的抗体的抗免疫球蛋白作为间接标记物，则无需再标记本发明的抗原。

获得本发明的刺突多肽之后，可用其产生与其具有反应性的多克隆和单克隆抗体。因此，可通过本领域已知的技术用本发明的蛋白质或者多肽来免疫动物。此类技术包括接种，不过也可包括其他的施用方式。要施用足量的蛋白质或者多肽以便在动物宿主中产生免疫应答。可使用能够产生针对本发明的抗原的抗体的任何宿主。在动物被免疫并经过足够常的时间以使其开始产生针对所述抗原的抗体后，即可收集多克隆抗体。常规的方法包括自动物采血，并自血液分离血清。含有针对所述抗原的抗体的血清可用作针对抗原的抗血清。或者，可自血清收集抗体。自血清收集纯化的针对抗原的多克隆抗体的优选技术是亲和纯化法。

也可以制备针对本发明的抗原的单克隆抗体。产生与所述抗原具有反应性的单克隆抗体的一种方法的步骤包括以抗原免疫动物；自宿主的脾脏收集产生抗体的细胞；将产生抗体的细胞与具有次黄嘌呤-鸟嘌呤磷酸核糖基转移酶缺陷的骨髓瘤细胞融合以形成杂交瘤；通过在包含次黄嘌呤、氨基蝶呤、和胸腺嘧啶的培养基中生长而选择出至少一种杂交瘤；鉴定出至少一种产生针对所述抗原的抗体的杂交瘤，培养所鉴定的杂交瘤以产生可回收量的抗体；以及收集由所述培养的杂交瘤所产生的抗体。

多克隆或者单克隆抗体可具有多种用途，其中之一是中和相应的蛋白质。它们也可用于检测生物制品中的病毒抗原，或者用于纯化相应的蛋白质、糖蛋白、或其混合物，例如使用亲和层析柱进行。

刺突多肽可用作抗原来鉴定生物学材料中的针对 SARS CoV 的抗体，并用来确定所述材料中抗体的浓度。因此，所述抗原可用于定性或

者定量确定材料中的病毒。此类材料包括人体组织和人体细胞、以及生物学液体，例如人的体液，包括人的血清。当作为试剂用于确定是否存在抗 SARS CoV 的抗体或其浓度的免疫测定法中时，本发明的抗原提供了一种简单、快捷、敏感、并且特异的测定法。

更具体地，本发明的抗原可用于通过免疫测定法检测 SARS CoV，这些免疫测定法已知可用于测定或者定量液体中的体液成分(humoral components in fluids)。这样，可以直接观察或者通过第二反应例如沉淀或者凝集来测定抗原抗体相互作用。此外，也可使用免疫电泳技术。例如，可以使用琼脂糖电泳继而与抗血清反应这种经典组合，以及双向电泳、火箭电泳、和聚丙烯酰胺凝胶电泳的免疫标记 (Western 印迹或者免疫印迹)。可以使用本发明的抗原的其他免疫测定法包括但不限于放射免疫测定法、竞争免疫沉淀测定法、酶免疫测定法、和免疫荧光测定法。可以理解，还可使用比浊法、比色法、和浊度分析技术。优选基于蛋白印迹技术的免疫测定法。

进行免疫测定法时可将免疫试剂之一(本发明的抗原或者针对所述抗原的本发明的抗体)固定化于一种载体表面而仍然保留该试剂的免疫反应性。另一方免疫试剂可以是非标记的或者标记后还仍然保留其免疫反应性。这些技术特别适合用于酶免疫测定法，例如酶联免疫吸附测定法 (ELISA) 竞争性抑制酶免疫测定法 (CIEIA)。

如果将本发明的抗原或者针对抗原的抗体附着于固相支持物，所述固相支持物通常是玻璃或者塑料材料。优选地将塑料材料成型为板型、管型、珠或者盘的形式。合适的塑料材料的形式为聚苯乙烯和聚氯乙烯。如果免疫试剂不易结合于固相支持物，则可在试剂与固相支持物之间设置一种载体材料。合适的载体材料的实例为蛋白质例如牛血清白蛋白，或者化学试剂例如戊二醛(gluteraldehyde)或者尿素。可采用常规技术包被固相支持物。

本发明提供免疫原性刺突多肽，更具体地是用于制备抗 SARS CoV

的疫苗组合物的保护性多肽。因此可将所述多肽作为病毒疫苗施用于对 SARS CoV 感染具有易感性的哺乳动物。可采用常规的施用方式。例如，可通过口服、呼吸道、或者胃肠外途径施用。当采用胃肠外施用时，优选地为皮内、皮下、和肌肉途径施用。

实现疫苗佐剂效应的各种方法包括使用例如氧化铝或者磷酸铝(alum)，通常使用 0.05%-0.1%的溶液，溶于磷酸缓冲盐中，并混和糖的合成聚合物(Carbopol)，其使用 0.25%的溶液。另一种合适的佐剂化合物包含 DDA (溴化二甲基双十八烷基铵(dimethyldioctadecyl-ammonium bromide))，以及免疫调节物例如淋巴因子(如 IFN- γ 、IL-1、IL-2、和 IL-12)或者 IFN- γ 诱导剂化合物如 poly I:C。

本发明的疫苗组合物优选地制备为可注射的形式(液态的溶液或者悬液)。不过，也可制备为适合在注射前在液体中配制成溶液或者悬液的固相形式。

此外，如果需要，疫苗组合物可含有少量的补充物质，例如湿化剂或者乳化剂、pH 缓冲剂、或者佐剂，以增强疫苗的效力。

本发明的疫苗组合物的施用方式与其剂量配制形式相适合，且其量能够产生治疗效果和免疫原性。施用的量取决于待治疗的个体，包括例如个体的免疫系统诱导产生免疫应答的能力。

疫苗的剂量取决于施用途径，且随着待接种的患者的年龄而变化，在较小程度上，随着待接种者的个头大小而变化。

感染 SARS CoV 的哺乳动物体内的免疫应答的主要目的是灭活游离 SARS CoV 并清除因感染 SARS CoV 而可能释放感染性病毒的细胞。B 细胞方面的免疫应答主要责任灭活 SARS CoV 病毒。实现这一目的的主要方式是对感染性进行中和。破坏感染 SARS CoV 的细胞的另一个重要机制是由细胞毒性 T 淋巴细胞(CTL)提供的，CTL 识别细胞表面上与 I 类组织相容性抗原联合表达的病毒 Spike 抗原。CTL 识别细胞内从刺突蛋白加工而来的刺突多肽，所述刺突蛋白例如是由感染的细胞所产生的，

或者是由吞噬细胞内在化的。因此，本发明可用于刺激抗刺突多肽的 B 细胞应答，以及病毒感染后由 CTL 介导的免疫。CTL 应答在介导原发 SARS CoV 感染的康复以及在随后的感染过程中加速康复等方面发挥重要作用。

可通过与佐剂进行乳化作用、掺入脂质体、偶联于适当载体、或者通过这些技术的组合，可增强本发明的刺突多肽和疫苗在宿主体内诱导保护性水平的中和抗体的能力。例如，本发明的刺突多肽可与常规佐剂一起施用，这些佐剂例如为磷酸铝和氧化铝凝胶，其量足以在宿主体内加强体液或细胞介导的免疫应答。类似地，可将刺突多肽结合于脂质膜或者掺入脂质膜以形成脂质体。为此可使用不含核酸和其他无关物质的非热源性脂质。

免疫接种方案取决于多种因素，例如宿主对感染的易感性以及宿主的年龄。可给宿主施用单一剂量的本发明的疫苗，或者先进行初始免疫接种，继而以一定时间间隔施用数个剂量。在初始接种后，根据需要可将随后的剂量作为加强接种而施用。

施用于宿主的本发明的刺突蛋白、多肽、和疫苗的量要足以预防或者抑制 SARS CoV 的感染或体内复制。在任何情况下，施用的量都应该至少足以保护宿主免于出现明显的免疫抑制，即使无法完全防止发生 SARS CoV 感染。可通过给宿主施用本发明的刺突蛋白或者糖蛋白而产生免疫原性应答，施用的量为大约 10 至大约 500 微克抗原/公斤体重，优选地为大约 50 至大约 100 微克抗原/公斤体重。本发明的蛋白质和疫苗可与生理学可接受的载体一起施用。例如，可使用稀释剂，例如水或者盐溶液。

本发明的另一个方面提供了 RNA 和/或 DNA 免疫接种的方法。该方法还包括给个体施用编码刺突多肽的核酸、所述蛋白质和多肽本身的任意组合，加或不加载体分子。在具体实施方式中，所述个体是动物，且优选地是哺乳动物。更优选地，哺乳动物选自人、小鼠、大鼠、兔、绵

羊、狗、猫、牛、猪、和马。在特别优选的实施方式中，哺乳动物是人。

治疗方法包括施用包含刺突多肽的免疫原性组合物，施用编码刺突多肽的核酸的组合物也可。本领域技术人员熟知核酸疫苗和核酸疫苗技术以及基于蛋白质和多肽的技术的概念、应用、以及有效性。基于核酸的技术能够将裸露的或者包囊化的编码刺突多肽的核酸直接施用至组织和细胞，而不需要在施用前产生所编码的蛋白质。该技术是基于这些核酸能够被受者生物体的细胞摄取并表达以产生免疫原性决定簇，而受者的免疫系统会对其产生应答。典型地，所表达的抗原展示于摄取并表达了所述核酸的细胞表面，但本发明也包括将所编码的抗原表达并输出至受者个体的循环系统中。此类核酸疫苗技术包括但不限于输送裸 DNA 和 RNA 以及输送编码刺突多肽的表达载体。尽管这种技术称为“疫苗”，但其也同等适用于并不产生保护性应答的免疫原性组合物。此类不诱导保护性的组合物和方法也属于本发明的范围。

尽管将编码刺突多肽的核酸和载体分子作为裸核酸而输送属于本发明的范围，但本发明也包括将核酸作为更大或者更复杂的组合物的一部分进行输送。这些输送系统包括含有编码刺突多肽的核酸的病毒、病毒样颗粒、或者细菌。同样，本发明的核酸和载体分子与细胞透过性化合物(cell permeabilizing compounds)如脂质体所形成的复合体也属于本发明的范围。用于核酸疫苗的其他化合物例如分子载体(EP 696,191, Samain et al.)和输送系统是本领域人员已知的，例如参见 WO 93 06223 和 WO 90 11092、U.S. 5,580,859、和 U.S. 5,589,466 (Vical 的专利)，通过引用将这些文献并入本文，上述这些均不需进行过多的实验即可制备并使用。

在单次或者两次注射刺突蛋白后，这种基于蛋白质的 SARS 疫苗可诱导中和性和保护性抗体依赖性免疫应答。基于蛋白质的疫苗与载体表达的病毒疫苗(即，质粒、MVA、Adeno)或者灭活全病毒疫苗相比表现出相当的安全性的优点。

为了进一步实现本发明的目的并与本发明的目的相一致，提供了一

种用于诊断 SARS CoV 感染的试剂盒。在一个实施方式中，所述试剂盒含有本发明的抗体，所述抗体能够结合 SARS CoV 刺突多肽。在另一个实施方式中，所述试剂盒含有本发明的多肽，所述多肽能够检测是否存在能够与刺突多肽相结合的抗体。在另一个实施方式中，所述试剂盒含有本发明的核酸分子，所述核酸分子能够与提示存在 SARS CoV 感染的病毒 RNA 或者类似的 DNA 序列杂交。可使用的各种不同的诊断技术包括但不限于：(1)用于鉴定可被或者不被限制性酶消化的细胞 DNA 的 Southern 印迹方法；(2)用于鉴定提取自细胞的 RNA 的 Northern 印迹；和(3)斑点杂交技术，即无需提前在琼脂糖凝胶上进行分离，直接将样品滤过通过 Ad Hoc 膜，例如硝酸纤维素膜或者尼龙膜；(4)基于蛋白印迹技术的免疫测定法；(5) ELISA (酶联免疫吸附测定法)；(6) FACS；(7)间接免疫荧光测定法；或(8)免疫沉淀测定法。用于斑点杂交技术的合适的材料可得自体液，包括但不限于，血清和血浆、培养细胞的上清液、或细胞裂解并通过离心去除细胞的膜和细胞核之后的胞浆提取物。

以下结合实施例进一步阐述本发明。

实施例 1

获得了来自感染 SARS CoV 的患者的刺突基因(Spike gene)，称为 Spike-Pasteur (SEQ ID NO:1)。将 Spike-Pasteur 克隆入 pcDNA 真核表达载体并转染入 293T 细胞。转染了 pcDNA-Spike-Pasteur 的细胞不表达 Spike-Pasteur 多肽，因为通过 FACS 或者 Western 印迹均没有观察到可检测水平的刺突蛋白。

随后在 SFV 病毒表达载体中表达 Spike-Pasteur，所述表达载体在转染的 BHK 细胞中能够有效地表达。不过，SFV 病毒颗粒的产率较低，这是因为在刺突基因中存在 2 个 Spe I 位点。Spe I 通常用于将质粒在 SFV 编码序列的末端进行线性化。由于不能使用 Spe I，因此使用了 Sph I 对 PSFV-Spike-Pasteur 进行线性化，这产生了 > 2000 个碱基的载体 RNA 的

额外的 3' RNA 序列。来自感染了 SFV-Spike-Pasteur 的细胞的代表表达水平很弱。

为了增强 Spike-Pasteur 的表达，将 Spike-Pasteur 序列中的两个 Spe I 位点消除，产生了 Spike-Pasteur-modif (SEQ ID NO: 2)。Spike-Pasteur-modif 使得能够以 Spe I 对 pSFV 进行标准的线性化，并发现其将 SFV 颗粒的产率最大提高了 100 倍。来自 Sph I 线性化的 pSFV-Spike-Pasteur 的 RNA 转染通常产生的 SFV 滴度为 2×10^7 IP/ml。来自 Spe I 线性化的 pSFV-Spike-Pasteur-modif 的 RNA 转染通常产生的 SFV 滴度为 $1-2 \times 10^9$ IP/ml。

实施例 2

进一步对 Spike-Pasteur 序列进行生物信息学分析。Spike-Pasteur 的 cDNA 含有许多顺式作用位点，这些位点可对表达造成负面影响。为了进一步增强表达，自 Spike-Pasteur 中消除了所鉴定的 33 个负顺式作用信号中的 32 个，并添加了刺激基因表达的额外的信号，由此产生了 Spike-HKU-PRC (SEQ ID NO: 3)。将 Spike-HKU-PRC 克隆入 pSC、pcDNA、和 pSFV 载体。

如表 1 所示，Spike-Pasteur 的 19 个富含 AU 的 RNA 不稳定性基序均被消除以产生 Spike-HKU-PRC。此外还去除了 12 个推定的剪接供体和受体位点中的 11 个、以及内部的 poly (A) 和重复序列和二级序列段。

表 1. 与优化的 Spike-HKU-PRC 相比，Spike-Pasteur 中的负顺式作用信号。

	Spike-Pasteur	Spike-HKU-PRC
富含 AU 的 RNA 不稳定性基序	19	0
重复序列和二级序列段	1	0
剪接供体和受体位点	12	1
内部 poly (A) 位点	1	0

添加到 Spike-HKU-PRC 上的额外的表达增强序列包括引入到起始 ATG 上游的 Kozak 共有序列以增强翻译的起始，并添加了 2 个 STOP 密码子以确保有效的终止。为了增加 mRNA 的半衰期，将 Spike-HKU-PRC 的 GC 含量由 38%提高至 49%，同时避免出现 GC 含量极高(>80%)或者极低(<30%)的区域。此外，使用的密码子偏向于中国仓鼠以提高翻译效率。表 2 所示为中国仓鼠的密码子使用，每一密码子的频率表示为数量/一千个密码子。

以下序列避免出现在 Spike-HKU-PRC 中：内部 TATA 盒、chi 位点、和核糖体进入位点；富含 AT 或者富含 GC 的序列段；重复序列和 RNA 二级结构；和剪接供体和受体位点以及剪接分枝点(splice branch points)。

表 2. 中国仓鼠的密码子使用，见 <http://www.kazusa.or.jp/codon/>。频率表示为数量/一千个密码子。

UUU	19.3	UCU	16.1	UAU	12.8	UGU	8.9
UUC	22.0	UCC	16.5	UAC	16.3	UGC	10.3
UUA	6.1	UCA	10.1	UAA	0.6	UGA	1.1
UUG	14.1	UCG	3.5	UAG	0.6	UGG	13.3
CUU	12.9	CCU	17.3	CAU	10.1	CGU	5.8
CUC	18.1	CCC	17.4	CAC	13.0	CGC	9.2
CUA	7.5	CCA	15.5	CAA	10.2	CGA	7.1
CUG	38.6	CCG	4.3	CAG	33.8	CGG	10.2
AUU	17.4	ACU	14.2	AAU	17.3	AGU	11.6
AUC	25.0	ACC	20.5	AAC	21.1	AGC	16.3
AUA	6.8	ACA	15.7	AAA	24.3	AGA	9.9
AUG	22.8	ACG	4.3	AAG	38.4	AGG	10.2
GUU	11.6	GCU	22.7	GAU	24.7	GGU	13.2
GUC	15.9	GCC	25.8	GAC	27.9	GGC	21.8
GUA	7.9	GCA	16.5	GAA	27.6	GGA	16.3
GUG	30.2	GCG	4.8	GAG	40.9	GGG	13.7

实施例 3

通过使用磷酸钙方法将质粒转染入 293T 细胞来比较 pcDNA-Spike-Pasteur 与 pcDNA-Spike-HKU-PRC 的刺突蛋白表达。转染了 pcDNA-Spike-Pasteur 的细胞中没有观察到刺突蛋白。相反，在转染了 pcDNA-Spike-HKU-PRC 的 293T 细胞中检测到高水平的刺突蛋白 (图 1)。刺突蛋白的迁移和寡聚化方式与以往得到的结果一致，提示这种质粒能够表达全长的、天然构象的 SARS CoV 蛋白质。这些结果证实，Spike 编码序列的密码子优化使得表达结果发生戏剧性的改善。

实施例 4

以 C 端 Flag 肽标记刺突蛋白，如图 2 所示。在先前所述的 SFV 载体系统中 (参见 Staropoli et al., Lozach et al., 和 Chanel et al., 均通过引用在此并入本文)，刺突蛋白表达为全长蛋白质，包括 C 端和跨膜结构域以及 C 端 Flag 标记物。

或者自转染了 SFV-Spike-RNA 的细胞或是感染了编码 SFV-Spike-RNA 的 SFV 颗粒的细胞制备刺突蛋白。制备 SFV 表达载体 RNA

通过一系列生化和免疫细胞化学分析法来分析刺突蛋白的正确折叠和预期特性。该蛋白质在进入内质网(ER)发生糖基化，获得高甘露糖 EndoH 敏感性 N-聚糖 (图 3)。通过 ER quality control exit、质膜表达(图 4)、可溶性 ACE2 受体结合(图 5)、和在 Western 印迹中被 SARS 患者血清所识别以及通过 FACS 分析(图 6-7)证实，刺突蛋白的折叠是正确的，即为其天然构象。

实施例 5

为进行 RNA 免疫接种，根据已公开的标准程序体外转录 RNA。

在 BHK 细胞中产生刺突蛋白，并使用抗 FLAG M2 抗体通过免疫亲和在天然条件下进行纯化。使用 Flag 肽在天然条件下洗脱结合 M2 的刺突蛋白。通过透析除去肽和残留的去垢剂。

肌肉注射 SFV Spike RNA 免疫小鼠，随后在第 14 天和第 35 天腹腔内(IP)注射刺突蛋白。在第 34、42、和 55 天自免疫的小鼠取得血清，所述血清在 Western 印迹(图 6)和 FACS (图 7)中显示存在重组 Spike 特异性抗体，且在感染了 SARS CoV 的 FRHK4 细胞上进行免疫荧光检测发现所述血清存在 SARS CoV 特异性抗体(图 8)。这些数据表明，SFV 载体中表达的刺突蛋白可以其天然构象而成功进行免疫纯化，并且这种纯化的蛋白质可在小鼠诱导产生高滴度的抗 SARS 抗体。

实施例 6

以下试剂和方法可用于实施本发明。

产生 SARS CoV Spike 亚单位疫苗

1/ 制备 SFV 表达载体 RNA

注：刺突蛋白可产生自，例如，转染了 SFV-Spike-RNA 的细胞或者以编码 SFV-Spike-RNA 的 SFV 颗粒感染的细胞。在此详细描述一种 Electroporaton 方法。

制备 1.2×10^7 细胞/ml 悬液，用于无菌条件下进行电穿孔 (Electroporation)

1. 制备无血清培养基，混和以下成分并滤过/除菌：
 - i. Hepes 5% 10 mL
 - ii. 胰蛋白胨-磷酸盐肉汤 (Tryptose-phosphate broth) 50 mL
 - iii. 青霉素 100 U/mL，链霉素 100 μ g/mL: 5mL
 - iv. GMEM QSP 500 mL
2. 预热试剂(GMEM、胰蛋白酶、PBS，不含 Ca^{2+} 或 Mg^{2+})至 37 $^{\circ}C$ ；
3. 轻柔地吸除原培养基，不要接触壁；

4. 以 10 ml PBS 冲洗细胞一次，弃洗液；
5. 加入 3 ml 胰蛋白酶并将平底培养瓶在罩(hood)中放置 4-5 分钟；
6. 加入 17 ml 新鲜完全培养基(GMEM 5% FCS)并悬浮细胞，转移至 50 ml 管中；
7. 细胞计数；取相当于 10^7 的细胞并在 1500 rpm 离心 5 分钟；
8. 将细胞重悬于 1 ml PBS (得到 10^7 细胞/ml)；
9. 将悬液置于冰上。

2/ 转染 SFV 表达载体 RNA

应当对样品和未转染的对照细胞均进行电穿孔。

1. 准备 2 个 75 ml 培养瓶，其中一个含有 20 ml GMEM，适当标记；
2. 使用无菌 P1000 滤器吸头转移 800 μ l 细胞悬液 (悬于无 Ca^{2+} 或 Mg^{2+} 的 PBS 中) 至含有 RNA 的管中，用移液器吸打两次；

3. 将混合物迅速转移至已经放置在电穿孔室中的电穿孔管 (Electroporator cuvette) 中；

4. 将电穿孔仪设定在 830 伏，25 μ Fd，并将阻抗设定在无限大，施加 2 个脉冲，各脉冲之间有 2-3 秒的延迟 (注：为达到合适的时间常数，在施加脉冲时保持电穿孔室是被遮蔽的)；

5. 注意电穿孔时间(应当在 0.4 ms 之内)；

6. 将细胞转移至含有 20 ml GMEM 的 75 ml 培养瓶中，轻柔地吸打以重悬细胞；

7. 在未转染的细胞上进行同样的操作；

8. 将细胞在 37 $^{\circ}$ C，5% CO_2 条件下温育过夜(约 16 小时)。

3/ 裂解细胞并制备蛋白

制备用于 Western 印迹和免疫纯化的蛋白裂解物

➤ 裂解缓冲液

Triton X-100 1%

Tris-HCl, pH 7.5	20 mM
NaCl	150 mM
EDTA	1 mM
PMSF	50 mg/ml

- 1 x PBS, 不含 Ca^{2+} 、 Mg^{2+}
- 蛋白样品缓冲液, 不含 DTT
- DTT , 1M, -20°C
- 细胞刮刀

制备新鲜的 20 ml 裂解缓冲液

1. 除去培养瓶中的培养基;
2. 以 10 ml 1 x PBS 洗涤细胞 (1 瓶);
3. 加入 500 μl 裂解液, 利用细胞刮刀自培养瓶中取出细胞;
4. 小心地将细胞裂解液转移至 1.5 ml 的 eppendorf 管中;
5. 以额外的 300 μl 裂解缓冲液尽可能多地取出残留细胞;
6. 将管在冰上放置 15 分钟;
7. 将管在 4°C 以 13,000 rpm 离心 15 分钟以去除细胞核物质;
8. 将上清液转移至置于冰上的新管中;
9. 将裂解液提取物放置于冰上。

4/ 自细胞裂解液中免疫亲和纯化 S-Flag 蛋白质

Triton X 100 裂解缓冲液 (Triton X 100 1%, Tris HCl pH7.5 20mM, NaCl 150mM, EGTA 1mM, PMSF 50 $\mu\text{g}/\text{ml}$)

1. 使用 200 μl 阔口移液头将 100 μl 的抗 flag M2 琼脂糖珠置于 1.5 ml eppendorf 中用于各个细胞裂解样品;
2. 用 1 ml 裂解缓冲液平衡各管的珠, 洗涤 3 次 (全速离心 15 秒将珠甩至底部, 轻柔地移去 90% 的 PBS, 注意避免将珠吸起);
3. 保存 50 μl 的细胞裂解物, 备用;
4. 充分混和(轻柔旋转)珠-裂解物混合物, 由此将洗涤后的珠与剩余

- 的细胞裂解物一起在 4°C 温育 4 小时；
5. 将珠甩至底部，并自各样品管去除上清液；
 6. 用 0.5 ml 1 X 洗涤缓冲液洗涤珠 3 次 (甩至底部，并添加新的洗涤缓冲液)；
 7. 将珠甩至底部，并去除绝大部分上清液，使得各个样品管中的残留体积是大约 100 μ l；
 8. 将 20 μ l 的各个珠样品分配至新的 1.5 ml eppendorf 管中，用于 Western 印迹检测；
 9. 将剩余的珠样品管存放于 -20°C，准备以后用于洗脱。

自免疫沉淀物中洗脱蛋白质

1. 方法参照 FLAGIPT-1 使用说明书；
2. 以 3 X FLAG 肽洗脱；
3. 通过将 3 μ l 浓度为 5 μ g/ μ l 的 3 X FLAG 肽添加至 100 μ l 的 1 X 洗涤缓冲液制备 3 X FLAG 肽工作液；
4. 将 100 μ l 的 3X FLAG 工作洗脱液加至树脂上；
5. 将混合物在 4°C 温育 1 小时，轻柔旋转；
6. 将树脂在 13000 rpm 离心 10 秒；
7. 保留上清液，并将步骤 4-6 再重复 3 次。

用 Amicon 滤器装置浓缩和纯化蛋白质

1. 温育后将各培养瓶的上清液收集至 50 mL 管中；
2. 2000 rpm 离心 5 分钟以去除细胞沉淀；
3. 将 15 mL 上清液转移至另一 50 mL 管中(含有 57 μ L 的 100 mM 的 PMSF，溶于 iPrOH) 任选
4. 置于冰上；
5. 将 20 ml 样品加至 Amicon Ultra-15 滤器装置；
6. 将盖好的滤器装置置于离心转子内，将体积刻度朝向上方，以类似的装置进行配重；

7. 在水平转子(swinging bucket rotor)中以最大 4000 x g 离心 20 分钟;
8. 通过自滤器装置吸取样品回收浓缩的溶质(500 μ l)。

实施例 7

经测定, 候选疫苗制剂, 即三聚体 S 蛋白(TriSpike, 与所述 Spike-HKU-PRC 相同的蛋白质), 纯度>90%。将用于对小鼠和仓鼠进行免疫接种研究的纯化 TriSpike 样品在 SDS/DTT 缓冲液(50mM DDT)中进行变性以使得三聚体蛋白质完全解离成为单体。以 4-12% SDS-PAGE 进行分离, 对凝胶进行银染(Current Protocols in Immunology Chapter 8, 9.1-9.10)以显示样品中所含有的所有蛋白质。图 16 显示只能够检测到单体 S 蛋白, 以其复合的糖基化和高甘露糖形式存在, 纯度> 90%。

实施例 8

以 TriSpike 加 Alum 佐剂免疫的动物产生了增强的血清 IgG 应答。以往对来自牙龈卟啉菌(*Porphyromonas gingivalis*)的重组 HagB 的粘膜和全身应答的研究显示, 与不加佐剂的 HagB 免疫相比, 以 HagB + Alum 免疫在 Balb/c 小鼠中诱导了更高的血清 IgG 应答和粘膜 IgA 应答 (Vaccine, 2003, 21, 4459-4471)。对 TriSpike 候选疫苗进行分析以确定其是否不仅能够诱导血清 IgG, 而且能够诱导对 SARS CoV 具有中和能力的粘膜 IgA。比较了 PBS 中的 TriSpike 制剂和 Alum 佐剂中的 TriSpike 制剂在诱导 SARS CoV 特异性血清 IgG 方面的能力。通过腹腔内途径对两组小鼠进行免疫: A 组代表单独接受 3 个剂量的 20 μ g TriSpike 蛋白的小鼠, 而 B 组代表接受 3 个剂量的预先混和了 1 mg Alum 佐剂的 20 μ g TriSpike 的小鼠。Western 印迹分析显示, 与单独使用 TriSpike 进行免疫的小鼠相比, 以 TriSpike + Alum 进行免疫的小鼠产生了更强的抗体应答(图 17)。TriSpike + Alum 组还显示出更高的中和滴度(图 17)。TriSpike + Alum 佐剂诱导了强烈的中和性和持久的血清 IgG 应答。

实施例 9

TriSpike + Alum 佐剂诱导了增强的粘膜 IgG 和 IgA 应答。在人类和实验性感染动物的上呼吸道和下呼吸道能够检测到 SARS CoV。除了呼吸道，在死亡病例的肠组织中也能检测到 SARS CoV (AJG, 2005, 100, 169-176)。为了研究 TriSpike 候选疫苗在粘膜部位诱导产生 SARS CoV 特异性 IgG 和 IgA 抗体的能力，发明人收集了来自以 TriSpike +/- Alum 佐剂通过腹腔内途径进行免疫的小鼠的粪便和鼻灌洗液样品。如前所述制备粪便样品 (PNAS, 2004, 101, 13584-13589)。简言之，在指定日收集粪块 (-100 mg)。在 40°C 加入 0.5 ml 的含有 0.02% Na-azide 的 PBS 轻柔旋转 30 分钟，并离心 (13,000 rpm)，由此制备到粪便提取物。通常，自一管粪块悬液能够制备到 0.2 ml 的上清液。Western 印迹分析 (图 18) 显示，仅在以 TriSpike + Alum 免疫的小鼠的粪便样品中存在粘膜 IgG 和 IgA 应答，而在单独以 TriSpike 免疫的小鼠中则没有。类似地，仅有来自以 TriSpike + Alum 免疫的小鼠的粪便样品中所含的 Ig 在微中和分析中显示出具有抗 SARS CoV 的中和活性。在来自以 TriSpike + Alum 免疫的小鼠的粪便样品中检测到了 IgG 和 IgA，这说明以 TriSpike 免疫的动物在胃肠道系统内建立了抗 SARS CoV 感染的第一线防御机制。

如 Current Protocols in Immunology (Chapter 19, 11. 15-16) 所述制备鼻灌洗液样品。简言之，在指定日收集鼻灌洗液样品。给原初小鼠和免疫的小鼠腹腔内注射两倍体积的氯胺酮/甲苯噻嗪溶液对小鼠实施麻醉。打开胸腔，插入 25G 针头，将 0.5 ml PBS/抑肽酶注射至阻塞部头部的气管腔内。可自各只小鼠收集大约 0.5 ml 的鼻清洗液样品。

对收集的鼻灌洗液样品 (n=3) 进行 Western 印迹分析发现，在用 TriSpike + Alum 免疫的小鼠的鼻灌洗液样品中存在粘膜 IgG，而单独以 TriSpike 免疫的小鼠则没有。不过，在用 TriSpike + Alum 免疫的小鼠或单独以 TriSpike 免疫的小鼠的鼻灌洗液样品中都没有检测到 IgA 应答，

这可能是由抗原施用途径所致。根据微中和分析结果，来自鼻部样品的粘膜 IgG 应答在体外没有产生任何抗 SARS CoV 感染的保护作用。

<110> INSTITUT PASTEUR
HONG KONG PASTEUR RESEARCH CENTRE LIMITED

<120> 与SARS冠状病毒刺突蛋白相关的核酸、多肽、表达方法和免疫原性组合物

<130> B6401AA 揭D/CAL

<140> PCT/EP 2005/006512
<141> 2005-06-03

<150> 10/860,641
<151> 2004-06-04

<150> 60/578,348
<151> 2004-06-10

<160> 7

<170> PatentIn Ver. 3.3

<210> 1
<211> 3795
<212> DNA
<213> SARS Coronavirus

<400> 1
atgttttattt tcttattatt tcttactctc actagtggtg gtgacctga ccgggtgcacc 60
acttttgatg atgttcaagc tcctaattac actcaacata cttcatctat gaggggggtt 120
tactatcctg atgaaatttt tagatcagac actctttatt taactcagga tttatttctt 180
ccatttttatt ctaatgttac agggtttcat actattaatc atacgtttg caaccctgtc 240
atacctttta aggatggtat ttattttgct gccacagaga aatcaaagt tgcccgtggt 300
tgggtttttg gttctaccat gaacaacaag tcacagtcgg tgattattat taacaattct 360
actaatgttg ttatacgagc atgtaacttt gaattgtgtg acaacccttt ctttgctggt 420
tctaaaccca tgggtacaca gacacatact atgatattcg ataatgcatt taattgcact 480
ttcgagtaca tatctgatgc cttttcgctt gatgtttcag aaaagtcagg taattttaa 540
cacttacgag agtttgtgtt taaaaataaa gatgggtttc tctatgttta taagggtat 600
caacctatag atgtagttcg tgatctacct tctggtttta acactttgaa acctattttt 660
aagttgcctc ttggtattaa cattacaaat tttagagcca ttcttacagc cttttcacct 720
gctcaagaca tttggggcac gtcagctgca gcctattttg ttggctattt aaagccaact 780
acatttatgc tcaagtatga tgaaaatggt acaatcacag atgctgttga ttgttctcaa 840
aatccacttg ctgaaactcaa atgctctggt aagagctttg agattgacaa aggaatttac 900
cagacctcta atttcagggt tgttccctca ggagatgttg tgagattccc taatattaca 960
aacttggtgc cttttggaga ggtttttaat gctactaaat tcccttctgt ctatgcatgg 1020
gagagaaaaa aaatttctaa ttgtgttgct gattactctg tgctctacaa ctcaacattt 1080
ttttcaacct ttaagtgcta tggcgtttct gccactaagt tgaatgatct ttgcttctcc 1140
aatgtctatg cagattcttt tgtagtcaag ggagatgatg taagacaaat agcggcagga 1200
caaactgggtg ttattgctga ttataattat aaattgccag atgatttcat gggttgtgtc 1260
cttgcttgga atactaggaa cattgatgct acttcaactg gtaattataa ttataaata 1320
aggatctta gacatggcaa gcttaggcc tttgagagag acatatctaa tgtgcctttc 1380
tcccctgatg gcaaaccttg caccccacct gctcttaatt gttattggcc attaaatgat 1440
tatggttttt acaccactac tggcattggc taccaacctt acagagttgt agtactttct 1500
tttgaacttt taaatgcacc ggccacggtt tgtggaccaa aattatccac tgaccttatt 1560
aagaaccagt gtgtcaattt taattttaat ggactcactg gtactgggtg gttactcct 1620
tcttcaaaga gatttcaacc atttcaacaa tttggccgtg atgtctctga tttcactgat 1680
tccgttcgag atcctaaaac atctgaaata ttagacattt caccttgctc ttttgggggt 1740
gtaagtgtaa ttacacctgg aacaaatgct tcatctgaag ttgctgttct atatcaagat 1800
gttaactgca ctgatgtttc tacagcaatc catgcagatc aactcacacc agcttggcgc 1860
atatattcta ctggaacaa tgtattccag actcaagcag gctgtcttat aggagctgag 1920
catgctgaca ctcttatga gtgcgacatt cctattggag ctggcatttg tgctagttag 1980
catacagttt ctttattacg tagtactagc caaaaatcta ttgtggctta tactatgtct 2040
ttaggtgctg atagttcaat tgcttactct aataacacca ttgctatacc tactaacttt 2100
tcaattagca ttactacaga agtaatgctt gtttctatgg ctaaaacctc cgtagattgt 2160
aatatgtaca tctgcggaga ttctactgaa tgtgctaatt tgcttctcca atatggtagc 2220

ttttgcacac	aactaaatcg	tgcactctca	ggatttgctg	ctgaacagga	tgcgaacaca	2280
cgtgaagtgt	tcgctcaagt	caaacaaatg	tacaaaacc	caactttgaa	atatttttgg	2340
ggttttaaatt	tttcacaaat	attacctgac	cctctaaagc	caactaagag	gtcttttatt	2400
gaggacttgc	tctttaataa	ggtgacactc	gctgatgctg	gcttcatgaa	gcaatatggc	2460
gaatgcctag	gtgataattaa	tgctagagat	ctcattttgtg	cgcagaagtt	caatgggctt	2520
acagtgttgc	cacctctgct	cactgatgat	atgattgctg	cctacactgc	tgctctagtt	2580
agtggtagt	ccactgctgg	atggacattt	ggtagctggc	ctgctcttca	aatacctttt	2640
gctatgcaaa	tgccatatag	gttcaatggc	attggagtta	cccaaaatgt	tctctatgag	2700
aaccaaaaaac	aaatcgccaa	ccaatttaac	aaggcgatta	gtcaaattca	agaatcactt	2760
acaacaacat	caactgcatt	gggcaagctg	caagacgttg	ttaaccagaa	tgctcaagca	2820
ttaaacaacac	ttgttaaaca	acttagctct	aattttgggtg	caatttcaag	tgtgctaaat	2880
gatatccttt	cgcgacttga	taaagtcgag	gcgagggtac	aaattgacag	gctaattaca	2940
ggcagacttc	aaagccttca	aacctatgta	acacaacaac	taatcagggc	tgctgaaatc	3000
agggcttctg	ctaactcttg	tgctactaaa	atgtctgagt	gtgttcttgg	acaatcaaaa	3060
agagttgact	tttgtggaaa	gggctaccac	cttatgtcct	tcccacaagc	agccccgcac	3120
gggtgtgtct	tctcatgtct	cacgtatgtg	ccatcccagg	agaggaactt	caccacagcg	3180
ccagcaatctt	gtcatgaagg	caaagcatac	ttccctcgtg	aagggtgttt	tgtgtttaat	3240
ggcacttctt	ggttttattac	acagaggaac	ttcttttctc	cacaaataat	tactacagac	3300
aatacatttg	tctcaggaaa	ttgtgatgtc	gttattggca	tcattaacaa	cacagtttat	3360
gatcctctgc	aacctgagct	tgactcattc	aaagaagagc	tggacaagta	cttcaaaaaat	3420
catacatcac	caagatgttga	tcttggcgac	atctcaggca	ttaacgcttc	tgctcgtcaac	3480
attcaaaaag	aaattgaccg	cctcaatgag	gtcgctaaaa	atttaaatga	atcactcatt	3540
gaccttcaag	aattgggaaa	atatgagcaa	tatattaaat	ggccttggtg	tgtttggctc	3600
ggcttcattg	ctggactaat	tgccatcgtc	atggttacaa	tcttgctttg	ttgcatgact	3660
agttgtttgca	gttgcoctcaa	gggtgcatgc	tcttgtgggt	cttgctgcaa	gtttgatgag	3720
gatgacgctg	agccagttct	caaggggtgc	aaattacatt	acacagacta	caaggatgac	3780
gatgacaata	agtaa					3795

<210> 2

<211> 3795

<212> DNA

<213> SARS Coronavirus

<400> 2

atgttttattt	tcttattatt	tottactctc	acgagtggtg	gtgaccttga	ccgggtgcacc	60
actttttagtg	atgttcaagc	tcctaattac	actcaacata	cttcatctat	gaggggggtt	120
tactatcctg	atgaaatttt	tagatcagac	actctttatt	taactcagga	tttatttctt	180
ccatttttatt	ctaattgttac	agggtttcat	actattaatc	atacgttttg	caaccctgtc	240
atacctttta	aggatgggat	ttattttgct	gccacagaga	aatcaaagt	tgtccgtggg	300
tgggtttttg	gttctaccat	gaacaacaag	tcacagtcgg	tgattattat	taacaattct	360
actaatgttg	ttatacgagc	atgtaacttt	gaattgtgtg	acaacccttt	ctttgctgtt	420
tctaaaccca	tgggtacaca	gacacatact	atgatattcg	ataatgcatt	taattgcact	480
ttcgagtaca	tatctgatgc	cttttcgctt	gatgtttcag	aaaagtcagg	taattttaaa	540
cacttacagag	agtttgtggt	taaaaataaa	gatgggtttc	tctatgttta	taagggtcat	600
caacctatag	atgtagtctg	tgatctacct	tctggtttta	acactttgaa	acctattttt	660
aagttgcctc	ttggatttaa	cattacaaat	tttagagcca	ttcttacagc	cttttcacct	720
gctcaagaca	tttggggcac	gtcagctgca	gcctattttg	ttggctattt	aaagccaact	780
acatttatgc	tcaagtatga	tgaaaatggt	acaatcacag	atgctgttga	ttgttctcaa	840
aatccacttg	ctgaactcaa	atgctctgtt	aagagctttg	agattgacaa	aggaatttac	900
cagacctcta	atctcagggt	tgttccctca	ggagatgttg	tgagattccc	taatattaca	960
aacttgtgtc	cttttggaga	ggtttttaat	gctactaaat	tcccttctgt	ctatgcatgg	1020
gagagaaaaa	aaatttctaa	ttgtgttgct	gattactctg	tgctctacaa	ctcaacattt	1080
ttttcaacct	ttaagtgcta	tggcgtttct	gccactaagt	tgaatgatct	ttgcttctcc	1140
aatgtctatg	cagattcttt	tgtagtcaag	ggagatgatg	taagacaaat	agcgccagga	1200
caaactgggtg	ttattgctga	ttataattat	aaattgccag	atgatttcat	gggttgtgtc	1260
cttgctttgga	atactaggaa	cattgatgct	acttcaactg	gtaattataa	ttataaatat	1320
aggtatctta	gacatggcaa	gcttaggccc	tttgagagag	acatatctaa	tgtgcctttc	1380
tcccctgatg	gcaaaccttg	caccccacct	gctcttaatt	gttattggcc	attaaatgat	1440
tattggtttt	acaccactac	tggcattggc	taccaacctt	acagagttgt	agtactttct	1500
tttgaacttt	taaatgcacc	ggccacgggt	tgtggaccaa	aattatccac	tgaccttatt	1560
aagaaccagt	gtgtcaatct	taattttaat	ggactcactg	gtactgggtg	gttaactcct	1620
tcttcaaaga	gatttcaacc	atctcaacaa	tttggccgtg	atgtctctga	tttactgat	1680


```

tccgttcgag atcctaaaac atctgaaata ttagacattt caccttgctc ttttgggggt 1740
gtaagtgtaa ttacacctgg aacaaatgct tcactctgaag ttgctgttct atatcaagat 1800
gttaactgca ctgatgtttc tacagcaatc catgcagatc aactcacacc agcttggcgc 1860
atatattcta ctggaaacaa tgtattccag actcaagcag gctgtcttat aggagctgag 1920
catgtcgaca cttcttatga gtgcgacatt cctattggag ctggcatttg tgctagttac 1980
catacagttt ctttattacg tagtactagc caaaaatcta ttgtggctta tactatgtct 2040
ttaggtgctg atagttcaat tgcttactct aataacacca ttgctatacc tactaacttt 2100
tcaattagca ttactacaga agtaatgcct gtttctatgg ctaaaacctc cgtagattgt 2160
aatatgtaca tctgcgagga ttctactgaa tgtgctaatt tgcttctcca atatggtagc 2220
ttttgcacac aactaaatcg tgcactctca ggtattgctg ctgaacagga tcgcaacaca 2280
cgtgaagtgt tcgctcaagt caaacaatg tacaaaacc caactttgaa atattttgg 2340
ggttttaatt tttcacaat attacctgac cctctaaagc caactaagag gtcttttatt 2400
gaggacttgc tctttaataa ggtgacactc gctgatgctg gcttcatgaa gcaatatggc 2460
gaatgcctag gtgatattaa tgctagagat ctcatgttg cgcagaagtt caatgggctt 2520
acagtgttgc cacctctgct cactgatgat atgattgctg cctacactgc tgctctagtt 2580
agtggctact ccactgctgg atggacattt ggtgctggcg ctgctcttca aatacctttt 2640
gctatgcaaa tggcatatag gttcaatggc attggagtta cccaaaatgt tctctatgag 2700
aaccaaaaac aaatcgccaa ccaatttaac aaggcgatta gtcaaattca agaatacactt 2760
acaacaacat caactgcatt gggcaagctg caagacgttg ttaaccagaa tgctcaagca 2820
ttaaacacac ttgttaaaca acttagctct aattttggtg caatttcaag tgtgctaaat 2880
gatatacctt cgcgacttga taaagtcgag gcgagggtac aaattgacag gctaattaca 2940
ggcagacttc aaagccttca aacctatgta acacaacaac taatcagggc tgctgaaatc 3000
agggcttctg ctaactctgc tgctactaaa atgtctgagt gtgttcttgg acaatcaaaa 3060
agagttgact tttgtggaaa gggctaccac cttatgtcct tcccacaagc agccccgc 3120
ggtgttgtct tctacatgt cacgtatgtg ccatcccagg agaggaactt caccacagcg 3180
ccagcaatth gtcatgaagg caaagcatac tccctcgtg aagggtgttt tgtgtttaat 3240
ggcacttctt ggtttattac acagaggaac tcttttctc cacaaataat tactacagac 3300
aatacatttg tctcaggaaa ttgtgatgtc gttattggca tcattaacaa cacagtttat 3360
gatcctctgc aacctgagci igactcattc aaagaagagc tggacaagta cttcaaaaat 3420
catacatcac cagatgttga tcttggcgac atttcaggca ttaacgcttc tgcgtcaac 3480
attcaaaaag aaattgaccg cctcaatgag gtcgctaaaa atttaaatga atcactcatt 3540
gaccttcaag aattgggaaa atatgagcaa tatattaaat ggccttggtg tgtttggctc 3600
ggcttcattg ctggactaat tgccatcgtc atggttacaa tcttgctttg ttgcatgacg 3660
agttgttgca gttgcctcaa ggggtgatgc tcttgtgggt cttgctgcaa gtttgatgag 3720
gatgacgctg agccagttct caagggtgtc aaattacatt acacagacta caaggatgac 3780
gatgacaata agtaa 3795

```

```

<210> 3
<211> 3897
<212> DNA
<213> SARS Coronavirus

```

```

<220>
<221> CDS
<222> (44)..(3844)

```

```

<400> 3
ctatagggcg aattgggtac cgctagcggg tccgcgcgcc acc atg ttt att ttc 55
Met Phe Ile Phe
1

ctg ctg ttt ctg act ctg acc agc ggc agt gac ctg gac cgg tgc acc 103
Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu Asp Arg Cys Thr
5 10 15 20

act ttt gat gat gtg cag gct cct aat tac act cag cat act tcc tct 151
Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln His Thr Ser Ser
25 30 35

atg agg ggc gtg tac tat cct gat gaa att ttt aga tcc gac act ctg 199
Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg Ser Asp Thr Leu
40 45 50

tat ctg act cag gat ctg ttt ctg cca ttc tat tct aat gtg aca ggc 247

```

Tyr	Leu	Thr	Gln	Asp	Leu	Phe	Leu	Pro	Phe	Tyr	Ser	Asn	Val	Thr	Gly		
		55					60					65					
ttt	cat	act	att	aat	cat	acc	ttt	ggc	aac	cct	gtg	atc	cct	ttt	aag	295	
Phe	His	Thr	Ile	Asn	His	Thr	Phe	Gly	Asn	Pro	Val	Ile	Pro	Phe	Lys		
	70					75				80							
gat	ggc	atc	tat	ttt	gct	gcc	aca	gag	aag	tcc	aat	gtg	gtg	cgg	gga	343	
Asp	Gly	Ile	Tyr	Phe	Ala	Ala	Thr	Glu	Lys	Ser	Asn	Val	Val	Arg	Gly		
	85				90					95					100		
tgg	gtg	ttc	ggc	tct	acc	atg	aac	aac	aag	tcc	cag	tcc	gtg	att	att	391	
Trp	Val	Phe	Gly	Ser	Thr	Met	Asn	Asn	Lys	Ser	Gln	Ser	Val	Ile	Ile		
				105					110					115			
att	aac	aat	tct	act	aat	gtg	gtg	atc	cga	gcc	tgt	aac	ttt	gaa	ctg	439	
Ile	Asn	Asn	Ser	Thr	Asn	Val	Val	Ile	Arg	Ala	Cys	Asn	Phe	Glu	Leu		
			120					125					130				
tgt	gac	aac	cca	ttc	ttt	gct	gtg	tct	aag	ccc	atg	ggc	aca	cag	aca	487	
Cys	Asp	Asn	Pro	Phe	Phe	Ala	Val	Ser	Lys	Pro	Met	Gly	Thr	Gln	Thr		
		135				140						145					
cat	act	atg	atc	ttc	gat	aat	gcc	ttt	aat	tgc	act	ttc	gag	tac	atc	535	
His	Thr	Met	Ile	Phe	Asp	Asn	Ala	Phe	Asn	Cys	Thr	Phe	Glu	Tyr	Ile		
	150					155					160						
tct	gat	gcc	ttt	tcc	ctg	gat	gtg	tcc	gaa	aag	tcc	ggc	aac	ttt	aag	583	
Ser	Asp	Ala	Phe	Ser	Leu	Asp	Val	Ser	Glu	Lys	Ser	Gly	Asn	Phe	Lys		
	165				170					175					180		
cac	ctg	cga	gag	ttt	gtg	ttt	aag	aat	aag	gat	ggc	ttt	ctg	tat	gtg	631	
His	Leu	Arg	Glu	Phe	Val	Phe	Lys	Asn	Lys	Asp	Gly	Phe	Leu	Tyr	Val		
				185					190					195			
tat	aag	ggc	tat	cag	cct	atc	gac	gtg	gtg	cgc	gat	ctg	cct	tct	ggc	679	
Tyr	Lys	Gly	Tyr	Gln	Pro	Ile	Asp	Val	Val	Arg	Asp	Leu	Pro	Ser	Gly		
			200					205					210				
ttt	aac	act	ctg	aag	cct	att	ttt	aag	ctg	cct	ctg	ggc	att	aac	att	727	
Phe	Asn	Thr	Leu	Lys	Pro	Ile	Phe	Lys	Leu	Pro	Leu	Gly	Ile	Asn	Ile		
		215					220					225					
aca	aat	ttt	cgg	gcc	att	ctg	aca	gcc	ttt	agc	cct	gct	cag	gac	att	775	
Thr	Asn	Phe	Arg	Ala	Ile	Leu	Thr	Ala	Phe	Ser	Pro	Ala	Gln	Asp	Ile		
	230					235					240						
tgg	ggc	acc	tct	gct	gcc	gcc	tat	ttt	gtg	ggc	tat	ctg	aag	cca	act	823	
Trp	Gly	Thr	Ser	Ala	Ala	Ala	Tyr	Phe	Val	Gly	Tyr	Leu	Lys	Pro	Thr		
	245				250					255					260		
acc	ttt	atg	ctg	aag	tat	gat	gaa	aat	ggc	aca	atc	aca	gat	gct	gtg	871	
Thr	Phe	Met	Leu	Lys	Tyr	Asp	Glu	Asn	Gly	Thr	Ile	Thr	Asp	Ala	Val		
				265					270					275			
gat	tgt	tct	cag	aat	cca	ctg	gct	gaa	ctg	aag	tgc	tct	gtg	aag	agc	919	
Asp	Cys	Ser	Gln	Asn	Pro	Leu	Ala	Glu	Leu	Lys	Cys	Ser	Val	Lys	Ser		
			280					285					290				
ttt	gag	att	gac	aag	gga	atc	tac	cag	acc	tct	aat	ttc	cgc	gtg	gtg	967	
Phe	Glu	Ile	Asp	Lys	Gly	Ile	Tyr	Gln	Thr	Ser	Asn	Phe	Arg	Val	Val		
		295					300					305					

ccc tct gga gat gtg gtg aga ttc cct aat att aca aac ctg tgt cct Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro 310 315 320	1015
ttt gga gaa gtg ttt aat gct act aag ttc cct tct gtg tat gcc tgg Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser Val Tyr Ala Trp 325 330 335 340	1063
gag aga aag aag att tct aat tgt gtg gct gat tac tct gtg ctg tac Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr 345 350 355	1111
aac tcc aca ttt ttt agc acc ttt aag tgc tat ggc gtg tct gcc act Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Ala Thr 360 365 370	1159
aag ctg aat gat ctg tgc ttc tcc aat gtg tat gcc gat tct ttt gtg Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala Asp Ser Phe Val 375 380 385	1207
gtg aag gga gat gat gtg aga cag atc gcc cca gga cag act ggc gtg Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Val 390 395 400	1255
att gct gat tac aat tat aag ctg cca gat gat ttc atg ggc tgt gtg Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Met Gly Cys Val 405 410 415 420	1303
ctg gct tgg aat act agg aac att gat gct act tcc act ggc aat tat Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser Thr Gly Asn Tyr 425 430 435	1351
aat tac aag tat cgg tat ctg aga cat ggc aag ctg agg ccc ttt gag Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu Arg Pro Phe Glu 440 445 450	1399
aga gac atc tct aac gtg cct ttc agc cct gat ggc aag cct tgc acc Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly Lys Pro Cys Thr 455 460 465	1447
cca cct gct ctg aat tgt tat tgg cca ctg aat gat tat ggc ttt tac Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp Tyr Gly Phe Tyr 470 475 480	1495
acc act act ggc att ggc tac cag cct tac aga gtg gtg gtg ctg tct Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser 485 490 495 500	1543
ttt gaa ctg ctg aat gcc cct gcc aca gtg tgt gga cca aag ctg tcc Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly Pro Lys Leu Ser 505 510 515	1591
act gac ctg att aag aac cag tgt gtg aac ttt aac ttt aat gga ctg Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn Phe Asn Gly Leu 520 525 530	1639
act ggc act ggc gtg ctg act cct tct agc aag aga ttt cag cca ttt Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg Phe Gln Pro Phe 535 540 545	1687
cag cag ttt ggc cgg gat gtg tct gat ttc act gat tcc gtg cga gat Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp Ser Val Arg Asp 550 555 560	1735

cct aag aca tct gaa atc ctg gac att tcc cct tgc tct ttt ggc ggc	1783
Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys Ser Phe Gly Gly	
565 570 575 580	
gtg agc gtg att aca cct gga aca aat gct tcc tct gaa gtg gct gtg	1831
Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser Glu Val Ala Val	
585 590 595	
ctg tat cag gat gtg aac tgc act gat gtg tct aca gcc atc cat gcc	1879
Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr Ala Ile His Ala	
600 605 610	
gat cag ctg aca cca gct tgg cgc atc tat tct act gga aac aat gtg	1927
Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr Gly Asn Asn Val	
615 620 625	
ttc cag act cag gcc ggc tgt ctg atc gga gct gag cat gtg gac act	1975
Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu His Val Asp Thr	
630 635 640	
tct tat gag tgc gac att cct att gga gct ggc att tgt gct agt tac	2023
Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser Tyr	
645 650 655 660	
cat aca gtg tct ctg ctg cgg agt act agc cag aag tct att gtg gct	2071
His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys Ser Ile Val Ala	
665 670 675	
tat act atg tct ctg ggc gct gat agt tcc att gct tac tct aat aac	2119
Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala Tyr Ser Asn Asn	
680 685 690	
acc att gct atc cct act aac ttt tcc att agc att act aca gaa gtg	2167
Thr Ile Ala Ile Pro Thr Asn Phe Ser Ile Ser Ile Thr Thr Glu Val	
695 700 705	
atg cct gtg tct atg gct aag acc tcc gtg gat tgt aat atg tac atc	2215
Met Pro Val Ser Met Ala Lys Thr Ser Val Asp Cys Asn Met Tyr Ile	
710 715 720	
tgc gga gat tct acc gaa tgt gct aat ctg ctg ctg cag tat ggc agc	2263
Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu Gln Tyr Gly Ser	
725 730 735 740	
ttt tgc aca cag ctg aat cgg gct ctg tct ggc att gct gct gaa cag	2311
Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile Ala Ala Glu Gln	
745 750 755	
gat cgc aac aca cgg gaa gtg ttc gct caa gtg aag cag atg tat aag	2359
Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys Gln Met Tyr Lys	
760 765 770	
acc cca act ctg aag tat ttt ggc ggc ttt aat ttt tcc cag atc ctg	2407
Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe Ser Gln Ile Leu	
775 780 785	
cct gac cct ctg aag ccc act aag cgg tct ttt att gag gac ctg ctg	2455
Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile Glu Asp Leu Leu	
790 795 800	
ttt aac aaa gtg aca ctg gct gat gct ggc ttt atg aag cag tat ggc	2503
Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met Lys Gln Tyr Gly	

805	810	815	820	
gaa tgc ctg ggc gat att aat gct aga gat ctg att tgt gcc cag aag Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile Cys Ala Gln Lys 825 830 835				2551
ttc aat ggc ctg aca gtg ctg cct cct ctg ctg act gat gat atg att Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr Asp Asp Met Ile 840 845 850				2599
gct gcc tac act gct gct ctg gtg tct ggc act gcc act gct gga tgg Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala Thr Ala Gly Trp 855 860 865				2647
aca ttt ggc gct ggc gct gct ctg cag atc cct ttt gct atg cag atg Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met 870 875 880				2695
gcc tat cgg ttc aat ggc att gga gtg acc cag aat gtg ctg tat gag Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu 885 890 895 900				2743
aac cag aag cag att gcc aac cag ttt aac aag gcc att agt cag att Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala Ile Ser Gln Ile 905 910 915				2791
cag gaa tcc ctg aca aca aca tcc act gcc ctg ggc aag ctg cag gac Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly Lys Leu Gln Asp 920 925 930				2839
gtg gtg aac cag aat gct cag gcc ctg aac aca ctg gtg aag cag ctg Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu 935 940 945				2887
agc agc aat ttt ggc gcc att tcc agt gtg ctg aat gat atc ctg tcc Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser 950 955 960				2935
cga ctg gat aaa gtg gag gcc gaa gtg cag att gac agg ctg att aca Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp Arg Leu Ile Thr 965 970 975 980				2983
ggc aga ctg cag agc ctg cag acc tat gtg aca cag cag ctg atc agg Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg 985 990 995				3031
gct gct gaa atc agg gct tct gcc aat ctg gct gct act aag atg tct Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met Ser 1000 1005 1010				3079
gag tgt gtg ctg gga cag tcc aag aga gtg gac ttt tgt gga aag ggc Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly Lys Gly 1015 1020 1025				3127
tac cac ctg atg tcc ttc cca cag gct gcc cct cat gga gtg gtg ttc Tyr His Leu Met Ser Phe Pro Gln Ala Ala Pro His Gly Val Val Phe 1030 1035 1040				3175
ctg cat gtg acc tat gtg cca tcc cag gag agg aac ttc acc aca gcc Leu His Val Thr Tyr Val Pro Ser Gln Glu Arg Asn Phe Thr Thr Ala 1045 1050 1055 1060				3223
cca gcc att tgt cat gaa ggc aag gcc tac ttc cct cgg gaa ggc gtg				3271

Pro Ala Ile Cys His Glu Gly Lys Ala Tyr Phe Pro Arg Glu Gly Val
1065 1070 1075

ttc gtg ttt aat ggc act tct tgg ttt att aca cag cgg aac ttc ttt 3319
Phe Val Phe Asn Gly Thr Ser Trp Phe Ile Thr Gln Arg Asn Phe Phe
1080 1085 1090

agc cca cag atc atc act aca gac aat aca ttt gtg tcc gga aat tgt 3367
Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr Phe Val Ser Gly Asn Cys
1095 1100 1105

gat gtg gtg att ggc atc att aac aac aca gtg tat gat cct ctg cag 3415
Asp Val Val Ile Gly Ile Ile Asn Asn Thr Val Tyr Asp Pro Leu Gln
1110 1115 1120

cct gag ctg gac tcc ttc aag gaa gag ctg gac aag tac ttc aag aat 3463
Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn
1125 1130 1135 1140

cat aca tcc cca gat gtg gat ctg ggc gac att tcc ggc att aac gct 3511
His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala
1145 1150 1155

tct gtg gtg aac att cag aag gaa att gac cgc ctg aat gaa gtg gct 3559
Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala
1160 1165 1170

aag aat ctg aat gaa tcc ctg att gac ctg cag gaa ctg ggc aag tat 3607
Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr
1175 1180 1185

gag cag tat att aag tgg cct tgg tat gtg tgg ctg ggc ttc att gct 3655
Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Val Trp Leu Gly Phe Ile Ala
1190 1195 1200

gga ctg att gcc atc gtg atg gtg aca atc ctg ctg tgt tgc atg acc 3703
Gly Leu Ile Ala Ile Val Met Val Thr Ile Leu Leu Cys Cys Met Thr
1205 1210 1215 1220

tcc tgt tgc agt tgc ctg aag ggc gct tgc tct tgt gga tct tgc tgc 3751
Ser Cys Cys Ser Cys Leu Lys Gly Ala Cys Ser Cys Gly Ser Cys Cys
1225 1230 1235

aag ttt gat gag gat gac tct gag cca gtg ctg aag ggc gtg aag ctg 3799
Lys Phe Asp Glu Asp Asp Ser Glu Pro Val Leu Lys Gly Val Lys Leu
1240 1245 1250

cat tac aca ggg ccc ggc ggc gac tac aag gac gat gac gac aag 3844
His Tyr Thr Gly Pro Gly Gly Asp Tyr Lys Asp Asp Asp Asp Lys
1255 1260 1265

tgatagatcg atgcatggat ccgtttaaac cgagctccag ctttgttccc tta 3897

<210> 4
<211> 1267
<212> PRT
<213> SARS Coronavirus

<400> 4
Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
1 5 10 15

Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 20 25 30
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 35 40 45
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 50 55 60
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 65 70 75 80
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 85 90 95
 Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 100 105 110
 Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
 115 120 125
 Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
 130 135 140
 Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 145 150 155 160
 Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 165 170 175
 Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 180 185 190
 Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
 195 200 205
 Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 210 215 220
 Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 225 230 235 240
 Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 245 250 255
 Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 260 265 270
 Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
 275 280 285
 Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
 290 295 300
 Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
 305 310 315 320
 Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 325 330 335
 Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 340 345 350

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
 355 360 365

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
 370 375 380

Asp Ser Phe Val Val Lys Gly Asp Asp Val Arg Gln Ile Ala Pro Gly
 385 390 395 400

Gln Thr Gly Val Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe
 405 410 415

Met Gly Cys Val Leu Ala Trp Asn Thr Arg Asn Ile Asp Ala Thr Ser
 420 425 430

Thr Gly Asn Tyr Asn Tyr Lys Tyr Arg Tyr Leu Arg His Gly Lys Leu
 435 440 445

Arg Pro Phe Glu Arg Asp Ile Ser Asn Val Pro Phe Ser Pro Asp Gly
 450 455 460

Lys Pro Cys Thr Pro Pro Ala Leu Asn Cys Tyr Trp Pro Leu Asn Asp
 465 470 475 480

Tyr Gly Phe Tyr Thr Thr Thr Gly Ile Gly Tyr Gln Pro Tyr Arg Val
 485 490 495

Val Val Leu Ser Phe Glu Leu Leu Asn Ala Pro Ala Thr Val Cys Gly
 500 505 510

Pro Lys Leu Ser Thr Asp Leu Ile Lys Asn Gln Cys Val Asn Phe Asn
 515 520 525

Phe Asn Gly Leu Thr Gly Thr Gly Val Leu Thr Pro Ser Ser Lys Arg
 530 535 540

Phe Gln Pro Phe Gln Gln Phe Gly Arg Asp Val Ser Asp Phe Thr Asp
 545 550 555 560

Ser Val Arg Asp Pro Lys Thr Ser Glu Ile Leu Asp Ile Ser Pro Cys
 565 570 575

Ser Phe Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Ala Ser Ser
 580 585 590

Glu Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Asp Val Ser Thr
 595 600 605

Ala Ile His Ala Asp Gln Leu Thr Pro Ala Trp Arg Ile Tyr Ser Thr
 610 615 620

Gly Asn Asn Val Phe Gln Thr Gln Ala Gly Cys Leu Ile Gly Ala Glu
 625 630 635 640

His Val Asp Thr Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile
 645 650 655

Cys Ala Ser Tyr His Thr Val Ser Leu Leu Arg Ser Thr Ser Gln Lys
 660 665 670

Ser Ile Val Ala Tyr Thr Met Ser Leu Gly Ala Asp Ser Ser Ile Ala
 675 680 685

Tyr	Ser	Asn	Asn	Thr	Ile	Ala	Ile	Pro	Thr	Asn	Phe	Ser	Ile	Ser	Ile
	690					695					700				
Thr	Thr	Glu	Val	Met	Pro	Val	Ser	Met	Ala	Lys	Thr	Ser	Val	Asp	Cys
705					710					715					720
Asn	Met	Tyr	Ile	Cys	Gly	Asp	Ser	Thr	Glu	Cys	Ala	Asn	Leu	Leu	Leu
				725					730					735	
Gln	Tyr	Gly	Ser	Phe	Cys	Thr	Gln	Leu	Asn	Arg	Ala	Leu	Ser	Gly	Ile
			740					745					750		
Ala	Ala	Glu	Gln	Asp	Arg	Asn	Thr	Arg	Glu	Val	Phe	Ala	Gln	Val	Lys
		755					760					765			
Gln	Met	Tyr	Lys	Thr	Pro	Thr	Leu	Lys	Tyr	Phe	Gly	Gly	Phe	Asn	Phe
	770					775					780				
Ser	Gln	Ile	Leu	Pro	Asp	Pro	Leu	Lys	Pro	Thr	Lys	Arg	Ser	Phe	Ile
785					790					795					800
Glu	Asp	Leu	Leu	Phe	Asn	Lys	Val	Thr	Leu	Ala	Asp	Ala	Gly	Phe	Met
				805					810					815	
Lys	Gln	Tyr	Gly	Glu	Cys	Leu	Gly	Asp	Ile	Asn	Ala	Arg	Asp	Leu	Ile
			820					825					830		
Cys	Ala	Gln	Lys	Phe	Asn	Gly	Leu	Thr	Val	Leu	Pro	Pro	Leu	Leu	Thr
		835					840					845			
Asp	Asp	Met	Ile	Ala	Ala	Tyr	Thr	Ala	Ala	Leu	Val	Ser	Gly	Thr	Ala
	850					855					860				
Thr	Ala	Gly	Trp	Thr	Phe	Gly	Ala	Gly	Ala	Ala	Leu	Gln	Ile	Pro	Phe
865					870					875					880
Ala	Met	Gln	Met	Ala	Tyr	Arg	Phe	Asn	Gly	Ile	Gly	Val	Thr	Gln	Asn
				885					890					895	
Val	Leu	Tyr	Glu	Asn	Gln	Lys	Gln	Ile	Ala	Asn	Gln	Phe	Asn	Lys	Ala
			900					905					910		
Ile	Ser	Gln	Ile	Gln	Glu	Ser	Leu	Thr	Thr	Thr	Ser	Thr	Ala	Leu	Gly
		915					920					925			
Lys	Leu	Gln	Asp	Val	Val	Asn	Gln	Asn	Ala	Gln	Ala	Leu	Asn	Thr	Leu
	930					935					940				
Val	Lys	Gln	Leu	Ser	Ser	Asn	Phe	Gly	Ala	Ile	Ser	Ser	Val	Leu	Asn
945					950					955					960
Asp	Ile	Leu	Ser	Arg	Leu	Asp	Lys	Val	Glu	Ala	Glu	Val	Gln	Ile	Asp
				965					970					975	
Arg	Leu	Ile	Thr	Gly	Arg	Leu	Gln	Ser	Leu	Gln	Thr	Tyr	Val	Thr	Gln
			980					985					990		
Gln	Leu	Ile	Arg	Ala	Ala	Glu	Ile	Arg	Ala	Ser	Ala	Asn	Leu	Ala	Ala
		995					1000					1005			
Thr	Lys	Met	Ser	Glu	Cys	Val	Leu	Gly	Gln	Ser	Lys	Arg	Val	Asp	Phe
	1010					1015					1020				
Cys	Gly	Lys	Gly	Tyr	His	Leu	Met	Ser	Phe	Pro	Gln	Ala	Ala	Pro	His

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

<210> 5
 <211> 1267
 <212> PRT
 <213> SARS Coronavirus

<400> 5
 Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
 1 5 10 15
 Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
 20 25 30
 His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 35 40 45

Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 50 55 60

Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 65 70 75 80

Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 85 90 95

Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 100 105 110

Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
 115 120 125

Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
 130 135 140

Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 145 150 155 160

Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 165 170 175

Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 180 185 190

Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
 195 200 205

Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 210 215 220

Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 225 230 235 240

Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 245 250 255

Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 260 265 270

Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
 275 280 285

Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
 290 295 300

Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
 305 310 315 320

Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 325 330 335

Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 340 345 350

Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
 355 360 365

Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala
 370 375 380

Asp 385	Ser	Phe	Val	Val	Lys 390	Gly	Asp	Asp	Val	Arg 395	Gln	Ile	Ala	Pro	Gly 400
Gln	Thr	Gly	Val	Ile 405	Ala	Asp	Tyr	Asn	Tyr 410	Lys	Leu	Pro	Asp	Asp	Phe 415
Met	Gly	Cys	Val 420	Leu	Ala	Trp	Asn	Thr	Arg 425	Asn	Ile	Asp	Ala	Thr	Ser 430
Thr	Gly	Asn 435	Tyr	Asn	Tyr	Lys	Tyr 440	Arg	Tyr	Leu	Arg	His 445	Gly	Lys	Leu
Arg 450	Pro	Phe	Glu	Arg	Asp	Ile 455	Ser	Asn	Val	Pro	Phe 460	Ser	Pro	Asp	Gly
Lys 465	Pro	Cys	Thr	Pro	Pro 470	Ala	Leu	Asn	Cys	Tyr 475	Trp	Pro	Leu	Asn	Asp 480
Tyr	Gly	Phe	Tyr	Thr 485	Thr	Thr	Gly	Ile	Gly 490	Tyr	Gln	Pro	Tyr	Arg	Val 495
Val	Val	Leu	Ser 500	Phe	Glu	Leu	Leu	Asn 505	Ala	Pro	Ala	Thr	Val 510	Cys	Gly
Pro	Lys	Leu 515	Ser	Thr	Asp	Leu	Ile 520	Lys	Asn	Gln	Cys	Val 525	Asn	Phe	Asn
Phe 530	Asn	Gly	Leu	Thr	Gly	Thr 535	Gly	Val	Leu	Thr 540	Pro	Ser	Ser	Lys	Arg
Phe 545	Gln	Pro	Phe	Gln	Gln 550	Phe	Gly	Arg	Asp	Val 555	Ser	Asp	Phe	Thr	Asp 560
Ser	Val	Arg	Asp	Pro 565	Lys	Thr	Ser	Glu	Ile 570	Leu	Asp	Ile	Ser	Pro	Cys 575
Ser	Phe	Gly	Gly 580	Val	Ser	Val	Ile	Thr 585	Pro	Gly	Thr	Asn	Ala	Ser	Ser 590
Glu	Val	Ala 595	Val	Leu	Tyr	Gln	Asp 600	Val	Asn	Cys	Thr	Asp 605	Val	Ser	Thr
Ala 610	Ile	His	Ala	Asp	Gln	Leu 615	Thr	Pro	Ala	Trp	Arg 620	Ile	Tyr	Ser	Thr
Gly 625	Asn	Asn	Val	Phe	Gln 630	Thr	Gln	Ala	Gly	Cys 635	Leu	Ile	Gly	Ala	Glu 640
His	Val	Asp	Thr	Ser 645	Tyr	Glu	Cys	Asp	Ile 650	Pro	Ile	Gly	Ala	Gly	Ile 655
Cys	Ala	Ser	Tyr 660	His	Thr	Val	Ser	Leu 665	Leu	Arg	Ser	Thr	Ser	Gln	Lys 670
Ser	Ile	Val 675	Ala	Tyr	Thr	Met	Ser 680	Leu	Gly	Ala	Asp	Ser	Ser	Ile	Ala
Tyr 690	Ser	Asn	Asn	Thr	Ile	Ala 695	Ile	Pro	Thr	Asn	Phe 700	Ser	Ile	Ser	Ile
Thr 705	Thr	Glu	Val	Met	Pro 710	Val	Ser	Met	Ala	Lys 715	Thr	Ser	Val	Asp	Cys 720
Asn	Met	Tyr	Ile	Cys	Gly	Asp	Ser	Thr	Glu	Cys	Ala	Asn	Leu	Leu	Leu

725					730					735					
Gln	Tyr	Gly	Ser	Phe	Cys	Thr	Gln	Leu	Asn	Arg	Ala	Leu	Ser	Gly	Ile
			740					745					750		
Ala	Ala	Glu	Gln	Asp	Arg	Asn	Thr	Arg	Glu	Val	Phe	Ala	Gln	Val	Lys
		755					760					765			
Gln	Met	Tyr	Lys	Thr	Pro	Thr	Leu	Lys	Tyr	Phe	Gly	Gly	Phe	Asn	Phe
	770					775					780				
Ser	Gln	Ile	Leu	Pro	Asp	Pro	Leu	Lys	Pro	Thr	Lys	Arg	Ser	Phe	Ile
785					790					795					800
Glu	Asp	Leu	Leu	Phe	Asn	Lys	Val	Thr	Leu	Ala	Asp	Ala	Gly	Phe	Met
				805					810					815	
Lys	Gln	Tyr	Gly	Glu	Cys	Leu	Gly	Asp	Ile	Asn	Ala	Arg	Asp	Leu	Ile
			820					825					830		
Cys	Ala	Gln	Lys	Phe	Asn	Gly	Leu	Thr	Val	Leu	Pro	Pro	Leu	Leu	Thr
		835					840					845			
Asp	Asp	Met	Ile	Ala	Ala	Tyr	Thr	Ala	Ala	Leu	Val	Ser	Gly	Thr	Ala
	850					855					860				
Thr	Ala	Gly	Trp	Thr	Phe	Gly	Ala	Gly	Ala	Ala	Leu	Gln	Ile	Pro	Phe
865					870					875					880
Ala	Met	Gln	Met	Ala	Tyr	Arg	Phe	Asn	Gly	Ile	Gly	Val	Thr	Gln	Asn
				885					890					895	
Val	Leu	Tyr	Glu	Asn	Gln	Lys	Gln	Ile	Ala	Asn	Gln	Phe	Asn	Lys	Ala
			900					905					910		
Ile	Ser	Gln	Ile	Gln	Glu	Ser	Leu	Thr	Thr	Thr	Ser	Thr	Ala	Leu	Gly
		915					920					925			
Lys	Leu	Gln	Asp	Val	Val	Asn	Gln	Asn	Ala	Gln	Ala	Leu	Asn	Thr	Leu
	930					935					940				
Val	Lys	Gln	Leu	Ser	Ser	Asn	Phe	Gly	Ala	Ile	Ser	Ser	Val	Leu	Asn
945					950					955					960
Asp	Ile	Leu	Ser	Arg	Leu	Asp	Lys	Val	Glu	Ala	Glu	Val	Gln	Ile	Asp
				965					970					975	
Arg	Leu	Ile	Thr	Gly	Arg	Leu	Gln	Ser	Leu	Gln	Thr	Tyr	Val	Thr	Gln
			980					985					990		
Gln	Leu	Ile	Arg	Ala	Ala	Glu	Ile	Arg	Ala	Ser	Ala	Asn	Leu	Ala	Ala
		995				1000						1005			
Thr	Lys	Met	Ser	Glu	Cys	Val	Leu	Gly	Gln	Ser	Lys	Arg	Val	Asp	Phe
	1010					1015					1020				
Cys	Gly	Lys	Gly	Tyr	His	Leu	Met	Ser	Phe	Pro	Gln	Ala	Ala	Pro	His
1025					1030					1035					1040
Gly	Val	Val	Phe	Leu	His	Val	Thr	Tyr	Val	Pro	Ser	Gln	Glu	Arg	Asn
			1045						1050					1055	

Phe Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys Ala Tyr Phe Pro
 1060 1065 1070
 Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser Trp Phe Ile Thr Gln
 1075 1080 1085
 Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr Phe Val
 1090 1095 1100
 Ser Gly Asn Cys Asp Val Val Ile Gly Ile Ile Asn Asn Thr Val Tyr
 1105 1110 1115 1120
 Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys
 1125 1130 1135
 Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser
 1140 1145 1150
 Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu
 1155 1160 1165
 Asn Glu Val Ala Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu
 1170 1175 1180
 Leu Gly Lys Tyr Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Val Trp Leu
 1185 1190 1195 1200
 Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile Leu Leu
 1205 1210 1215
 Cys Cys Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Ala Cys Ser Cys
 1220 1225 1230
 Gly Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro Val Leu Lys
 1235 1240 1245
 Gly Val Lys Leu His Tyr Thr Gly Pro Gly Gly Asp Tyr Lys Asp Asp
 1250 1255 1260
 Asp Asp Lys
 1265

<210> 6

<211> 3765

<212> DNA

<213> SARS Coronavirus

<400> 6

atgtttatatt tcttgctggt tctgactctg accagcggca gtgacctgga cgggtgcacc 60
 acttttgatg atgtgcaggc tcctaattac actcagcata cttcctctat gaggggcgtg 120
 tactatcctg atgaaatttt tagatccgac actctgtatc tgactcagga tctgtttctg 180
 ccattctatt ctaatgtgac aggccttcat actattaatc atacctttgg caaccctgtg 240
 atccctttta aggatggcat ctattttgct gccacagaga agtccaatgt ggtgcgggga 300
 tgggtgttcg gctctacat gaacaacaag tcccagtcg tgattattat taacaattct 360
 actaatgtgg tgatccgagc ctgtaacttt gaactgtgtg acaaccatt ctttgctgtg 420
 tctaagccca tgggcacaca gacacatact atgatcttcg ataatgcctt taattgcact 480
 ttcgagtaca tctctgatgc cttttccctg gatgtgtccg aaaagtccgg caactttaag 540
 cacctgogag agtttggtgt taagaataag gatggctttc tgtatgtgta taagggctat 600
 cagcctatcg acgtgggtgc cgatctgcct tctggcttta aactctgaa gcctatattt 660
 aagctgcctc tgggcattaa cattacaaa tttcgggcca ttctgacagc ctttagcct 720
 gctcaggaca tttggggcac ctctgctgcc gcctattttg tgggctatct gaagccaact 780
 acctttatgc tgaagtatga tgaaaatggc acaatcacag atgctgtgga ttgttctcag 840

```

aatccactgg ctgaactgaa gtgctctgtg aagagctttg agattgacaa gggaaatctac 900
cagacctcta atttccgcgt ggtgccctct ggagatgtgg tgagattccc taatattaca 960
aacctgtgtc cttttggaga agtgtttaat gctactaagt tcccttctgt gtatgcctgg 1020
gagagaaaga agatttctaa ttgtgtggct gattactctg tgctgtacaa ctccacattt 1080
tttagcacct ttaagtgcta tggcgtgtct gccactaagc tgaatgatct gtgcttctcc 1140
aatgtgtatg ccgattcctt tgtggtgaag ggagatgatg tgagacagat cgccccagga 1200
cagactggcg tgattgctga ttacaattat aagctgccag atgatttcat gggctgtgtg 1260
ctggcttggg atactaggaa cattgatgct acttccactg gcaattataa ttacaagtat 1320
cggtatctga gacatggcaa gctgaggccc tttgagagag acatctctaa cgtgccttct 1380
agccctgatg gcaagccttg caccacacct gctctgaatt gttattggcc actgaatgat 1440
tatggctttt acaccactac tggcattggc taccagcctt acagagtggg ggtgctgtct 1500
tttgaactgc tgaatgcccc tgccacagtg tgtggaccaa agctgtccac tgacctgatt 1560
aagaaccagt gtgtgaactt taactttaat ggactgactg gcaactggcg gctgactcct 1620
tctagcaaga gatttcagcc atttcagcag tttggccggg atgtgtctga tttcactgat 1680
tccgtgcbag atcctaagac atctgaaatc ctggacattt ccccttgctc ttttggcggc 1740
gtgagcgtga ttacacctgg aacaaatgct tccttgaag tggctgtgct gtatcaggat 1800
gtgaactgca ctgagtgtgc tacagccatc catgccgac agctgacacc agcttggcgc 1860
atctattcta ctggaacaa tgtgttccag actcaggccg gctgtctgat cggagctgag 1920
catgtggaca cttcttatga gtgacacatt cctattggag ctggcatttg tgctagttac 1980
catacagtgt ctctgctgcg gactactagc cagaagtcta ttgtggctta tactatgtct 2040
ctgggcgctg atagttccat tgcttactct aataacacca ttgctatccc tactaacttt 2100
tccattagca ttactacaga agtgatgctt gtgtctatgg ctaagacctc cgtggattgt 2160
aatatgtaca tctgcbgaga ttctaccgaa tgtgctaate tgctgtgca gtatggcagc 2220
ttttgcacac agctgaatcg ggctctgtct ggcattgctg ctgaacagga tcgcaacaca 2280
cgggaagtgt tcgctcaagt gaagcagatg tataagacc caactctgaa gtattttggc 2340
ggctttaatt tttcccagat cctgcctgac cctctgaagc ccactaagcg gtcttttatt 2400
gaggacctgc tgtttaacaa agtgacactg gctgatgctg gctttatgaa gcagtatggc 2460
gaatgcctgg gcgataattaa tgctagagat ctgattttgt cccagaagtt caatggcctg 2520
acagtgtctg ctctctgtct gactgatgat atgattgctg cctacactgc tgctctggtg 2580
tctggcactg ccaactgctg atggacattt ggcgctggcg ctgctctgca gatccctttt 2640
gctatgcaag tggcctatcg gttcaatggc attggagtga cccagaatgt gctgtatgag 2700
aaccagaagc agattgcca ccagtttaac aaggcatta gtcagattca ggaatccctg 2760
acaacaacat ccaactgccct gggcaagctg caggacgtgg tgaaccagaa tgctcaggcc 2820
ctgaacacac tgggtaagca gctgagcagc aattttggcg ccatttccag tgtgctgaat 2880
gatatcctgt cccgactgga taaagtggag gccgaagtgc agattgacag gctgattaca 2940
ggcagactgc agagcctgca gacctatgtg acacagcagc tgatcagggc tgctgaaatc 3000
agggcttctg ccaatctggc tgctactaag atgtctgagt gtgtgctggg acagtccaag 3060
agagtggact tttgtggaaa gggctaccac ctgatgtcct tcccacaggc tgcccctcat 3120
ggagtgggtg tctgcatgt gacctatgt ccatcccagg agaggaactt caccacagcc 3180
ccagccattt gtcataaggg caaggcctac ttccctcggg aaggcgtgtt cgtgtttaat 3240
ggcacttctt ggtttattac acagcggaac ttctttagcc cacagatcat cactacagac 3300
aatacatttg tgtccggaaa ttgtgatgtg gtgattggca tcattaacaa cacagtgtat 3360
gatectctgc agcctgagct ggactccttc aaggaagagc tggacaagta cttcaagaat 3420
catacatccc cagatgtgga tctgggcgac atttccggca ttaacgcttc tgtggtgaac 3480
attcagaagg aaattgaccg cctgaatgaa gtggctaaga atctgaatga atccctgatt 3540
gacctgcagg aactgggcaa gtatgagcag tatattaagt ggccttggta tgtgtggctg 3600
ggcttcattg ctggactgat tgccatcgtg atggtgacaa tctgctgtg ttgcatgacc 3660
tctgtttgca gttccctgaa gggcgttgc tctgtggat cttgctgcaa gtttgatgag 3720
gatgactctg agccagtgtc gaaggcgtg aagctgcatt acaca 3765

```

<210> 7

<211> 1255

<212> PRT

<213> SARS Coronavirus

<400> 7

```

Met Phe Ile Phe Leu Leu Phe Leu Thr Leu Thr Ser Gly Ser Asp Leu
  1                   5                   10                   15

```

```

Asp Arg Cys Thr Thr Phe Asp Asp Val Gln Ala Pro Asn Tyr Thr Gln
                   20                   25                   30

```

His Thr Ser Ser Met Arg Gly Val Tyr Tyr Pro Asp Glu Ile Phe Arg
 35 40 45
 Ser Asp Thr Leu Tyr Leu Thr Gln Asp Leu Phe Leu Pro Phe Tyr Ser
 50 55 60
 Asn Val Thr Gly Phe His Thr Ile Asn His Thr Phe Gly Asn Pro Val
 65 70 75 80
 Ile Pro Phe Lys Asp Gly Ile Tyr Phe Ala Ala Thr Glu Lys Ser Asn
 85 90 95
 Val Val Arg Gly Trp Val Phe Gly Ser Thr Met Asn Asn Lys Ser Gln
 100 105 110
 Ser Val Ile Ile Ile Asn Asn Ser Thr Asn Val Val Ile Arg Ala Cys
 115 120 125
 Asn Phe Glu Leu Cys Asp Asn Pro Phe Phe Ala Val Ser Lys Pro Met
 130 135 140
 Gly Thr Gln Thr His Thr Met Ile Phe Asp Asn Ala Phe Asn Cys Thr
 145 150 155 160
 Phe Glu Tyr Ile Ser Asp Ala Phe Ser Leu Asp Val Ser Glu Lys Ser
 165 170 175
 Gly Asn Phe Lys His Leu Arg Glu Phe Val Phe Lys Asn Lys Asp Gly
 180 185 190
 Phe Leu Tyr Val Tyr Lys Gly Tyr Gln Pro Ile Asp Val Val Arg Asp
 195 200 205
 Leu Pro Ser Gly Phe Asn Thr Leu Lys Pro Ile Phe Lys Leu Pro Leu
 210 215 220
 Gly Ile Asn Ile Thr Asn Phe Arg Ala Ile Leu Thr Ala Phe Ser Pro
 225 230 235 240
 Ala Gln Asp Ile Trp Gly Thr Ser Ala Ala Ala Tyr Phe Val Gly Tyr
 245 250 255
 Leu Lys Pro Thr Thr Phe Met Leu Lys Tyr Asp Glu Asn Gly Thr Ile
 260 265 270
 Thr Asp Ala Val Asp Cys Ser Gln Asn Pro Leu Ala Glu Leu Lys Cys
 275 280 285
 Ser Val Lys Ser Phe Glu Ile Asp Lys Gly Ile Tyr Gln Thr Ser Asn
 290 295 300
 Phe Arg Val Val Pro Ser Gly Asp Val Val Arg Phe Pro Asn Ile Thr
 305 310 315 320
 Asn Leu Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Lys Phe Pro Ser
 325 330 335
 Val Tyr Ala Trp Glu Arg Lys Lys Ile Ser Asn Cys Val Ala Asp Tyr
 340 345 350
 Ser Val Leu Tyr Asn Ser Thr Phe Phe Ser Thr Phe Lys Cys Tyr Gly
 355 360 365
 Val Ser Ala Thr Lys Leu Asn Asp Leu Cys Phe Ser Asn Val Tyr Ala

370				375				380							
Asp 385	Ser	Phe	Val	Val	Lys	Gly	Asp	Asp	Val	Arg	Gln	Ile	Ala	Pro	Gly 400
Gln	Thr	Gly	Val	Ile 405	Ala	Asp	Tyr	Asn	Tyr 410	Lys	Leu	Pro	Asp	Asp 415	Phe
Met	Gly	Cys	Val 420	Leu	Ala	Trp	Asn	Thr 425	Arg	Asn	Ile	Asp	Ala 430	Thr	Ser
Thr	Gly	Asn 435	Tyr	Asn	Tyr	Lys	Tyr 440	Arg	Tyr	Leu	Arg	His 445	Gly	Lys	Leu
Arg 450	Pro	Phe	Glu	Arg	Asp	Ile 455	Ser	Asn	Val	Pro	Phe 460	Ser	Pro	Asp	Gly
Lys 465	Pro	Cys	Thr	Pro	Pro	Ala 470	Leu	Asn	Cys	Tyr 475	Trp	Pro	Leu	Asn	Asp 480
Tyr	Gly	Phe	Tyr	Thr 485	Thr	Thr	Gly	Ile	Gly 490	Tyr	Gln	Pro	Tyr	Arg 495	Val
Val	Val	Leu	Ser 500	Phe	Glu	Leu	Leu	Asn 505	Ala	Pro	Ala	Thr	Val 510	Cys	Gly
Pro	Lys	Leu 515	Ser	Thr	Asp	Leu	Ile 520	Lys	Asn	Gln	Cys	Val 525	Asn	Phe	Asn
Phe 530	Asn	Gly	Leu	Thr	Gly	Thr 535	Gly	Val	Leu	Thr	Pro 540	Ser	Ser	Lys	Arg
Phe 545	Gln	Pro	Phe	Gln	Gln 550	Phe	Gly	Arg	Asp	Val 555	Ser	Asp	Phe	Thr	Asp 560
Ser	Val	Arg	Asp	Pro 565	Lys	Thr	Ser	Glu	Ile 570	Leu	Asp	Ile	Ser	Pro	Cys 575
Ser	Phe	Gly	Gly 580	Val	Ser	Val	Ile	Thr 585	Pro	Gly	Thr	Asn	Ala 590	Ser	Ser
Glu	Val	Ala 595	Val	Leu	Tyr	Gln	Asp 600	Val	Asn	Cys	Thr	Asp 605	Val	Ser	Thr
Ala 610	Ile	His	Ala	Asp	Gln	Leu 615	Thr	Pro	Ala	Trp	Arg 620	Ile	Tyr	Ser	Thr
Gly 625	Asn	Asn	Val	Phe	Gln 630	Thr	Gln	Ala	Gly	Cys 635	Leu	Ile	Gly	Ala	Glu 640
His	Val	Asp	Thr	Ser 645	Tyr	Glu	Cys	Asp	Ile 650	Pro	Ile	Gly	Ala	Gly 655	Ile
Cys	Ala	Ser	Tyr 660	His	Thr	Val	Ser	Leu 665	Leu	Arg	Ser	Thr	Ser 670	Gln	Lys
Ser	Ile	Val 675	Ala	Tyr	Thr	Met	Ser 680	Leu	Gly	Ala	Asp	Ser 685	Ser	Ile	Ala
Tyr 690	Ser	Asn	Asn	Thr	Ile	Ala 695	Ile	Pro	Thr	Asn	Phe 700	Ser	Ile	Ser	Ile
Thr 705	Thr	Glu	Val	Met	Pro 710	Val	Ser	Met	Ala	Lys 715	Thr	Ser	Val	Asp	Cys 720

Asn Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ala Asn Leu Leu Leu
 725 730 735
 Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Ser Gly Ile
 740 745 750
 Ala Ala Glu Gln Asp Arg Asn Thr Arg Glu Val Phe Ala Gln Val Lys
 755 760 765
 Gln Met Tyr Lys Thr Pro Thr Leu Lys Tyr Phe Gly Gly Phe Asn Phe
 770 775 780
 Ser Gln Ile Leu Pro Asp Pro Leu Lys Pro Thr Lys Arg Ser Phe Ile
 785 790 795 800
 Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Met
 805 810 815
 Lys Gln Tyr Gly Glu Cys Leu Gly Asp Ile Asn Ala Arg Asp Leu Ile
 820 825 830
 Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu Leu Thr
 835 840 845
 Asp Asp Met Ile Ala Ala Tyr Thr Ala Ala Leu Val Ser Gly Thr Ala
 850 855 860
 Thr Ala Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe
 865 870 875 880
 Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn
 885 890 895
 Val Leu Tyr Glu Asn Gln Lys Gln Ile Ala Asn Gln Phe Asn Lys Ala
 900 905 910
 Ile Ser Gln Ile Gln Glu Ser Leu Thr Thr Thr Ser Thr Ala Leu Gly
 915 920 925
 Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu
 930 935 940
 Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
 945 950 955 960
 Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp
 965 970 975
 Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln
 980 985 990
 Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala
 995 1000 1005
 Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe
 1010 1015 1020
 Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ala Ala Pro His
 1025 1030 1035 1040
 Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ser Gln Glu Arg Asn
 1045 1050 1055

Phe Thr Thr Ala Pro Ala Ile Cys His Glu Gly Lys Ala Tyr Phe Pro
 1060 1065 1070

Arg Glu Gly Val Phe Val Phe Asn Gly Thr Ser Trp Phe Ile Thr Gln
 1075 1080 1085

Arg Asn Phe Phe Ser Pro Gln Ile Ile Thr Thr Asp Asn Thr Phe Val
 1090 1095 1100

Ser Gly Asn Cys Asp Val Val Ile Gly Ile Ile Asn Asn Thr Val Tyr
 1105 1110 1115 1120

Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys
 1125 1130 1135

Tyr Phe Lys Asn His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser
 1140 1145 1150

Gly Ile Asn Ala Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu
 1155 1160 1165

Asn Glu Val Ala Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu
 1170 1175 1180

Leu Gly Lys Tyr Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Val Trp Leu
 1185 1190 1195 1200

Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile Leu Leu
 1205 1210 1215

Cys Cys Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Ala Cys Ser Cys
 1220 1225 1230

Gly Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro Val Leu Lys
 1235 1240 1245

Gly Val Lys Leu His Tyr Thr
 1250 1255

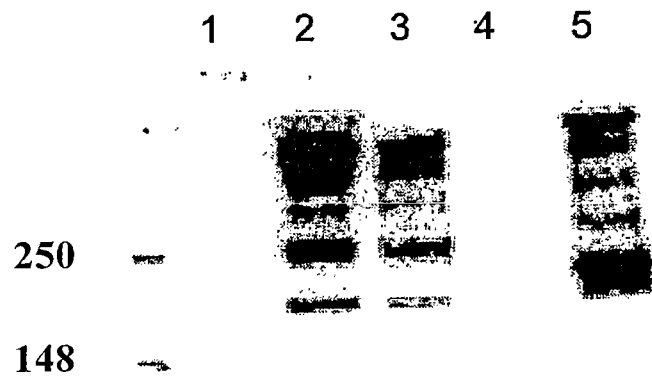


图 1

标记了 C 端 Flag 肽的刺突蛋白

SEQ ID NO: 5

MFIFLLFLTLTSGSDLDRCTTFDDVQAPNYTQHTSSMRGVVYYPDEIFRSDTLYLTQDLF
LPFYSNVTGFHTINHTFGNPVIFPKDGIYFAATEKSNVVRGWVFGSTMNKSQSVIIIN
NSTNVVIRACNFELCDNPFFAVSKPMGTQHTMIFDNAFNCTFEYISDAFSLDVSEKSG
NFKHLREFVFKNKDGFLYVYKGYQPIDVVRDLPSGFNTLKPIFKLPLGINITNFRAILT
AFSPAQDIWGTSAAYFVGYLKPTTFMLKYDENGITITDAVDCSQNPLAELKCSVKSFEI
DKGIYQTSNFRVVPDGDVVRFPNITNLCPFGEVFNATKFPVYAWERKKISNCVADYSV
LYNSTFFSTFKCYGVSATKLNLDLCSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLP
DDFMGCVLAWNTRNIDATSTGNVYKYRYLRHGKLRPFERDISNVPFSPDGKPCPPAL
NCYWPLNDYGFYTTTGIGYQPYRVVLSFELLNAPATVCGPKLSTDLIKNQCVNFNFNNG
LTGTGVLTPSSKRFQPFQFGRDVSDFTDSVRDPKTSEILDISPCSFGGVSVITPGTNA
SSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNVVFQTOAGCLIGAEHVDTSYECD
IPIGAGICASYHTVSLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTEV
MPVSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQV
KQMYKTPTLKYFGGFNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQYGECLGDI
NARDLICAQKFNGLTVLPPLLTDDMIAAYTAALVSGTATAGWTFGAGAALQIPFAMQMA
YRFNGIGVTQNVLYENQKQIANQFNKAI SQIQESLTTTSTALGKLQDVVNQNAQALNTL
VKQLSSNFGAISSVLNDILSRDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRAS
ANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTVVPSQERNFTTAPA
ICHEGKAYFPREGVVFVNGTSWFITQRNFFSPQIITTDNTFVSGNCDVVIGIINNTVYD
PLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLI
DLQELGKYEQYIKWPWYVWLGFIAGLIAIVMVTILLCCMTSCCSCLKGACSCGSCCKFD
EDDSEPVLKGVKLHYTGPGGDYKDDDDK

图 2

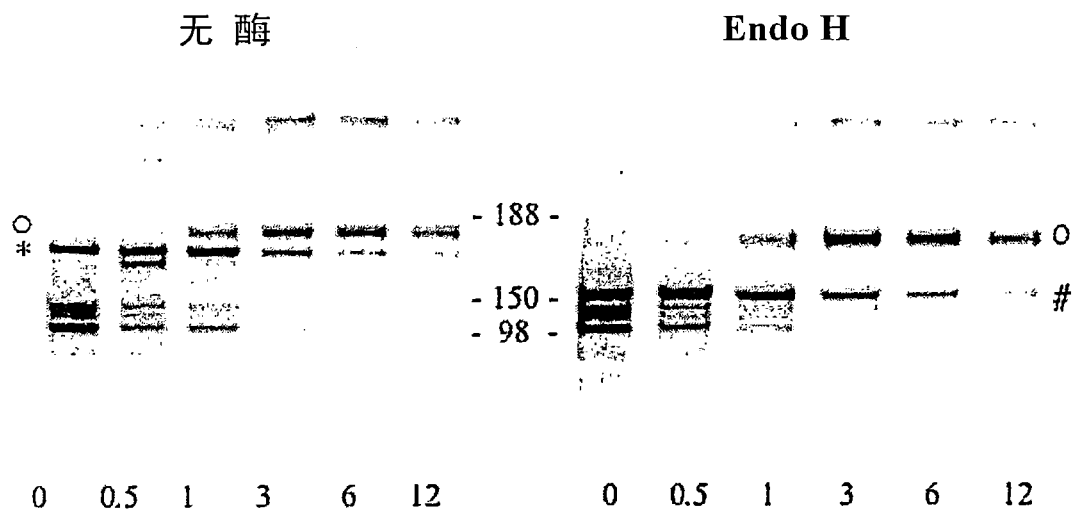


图 3

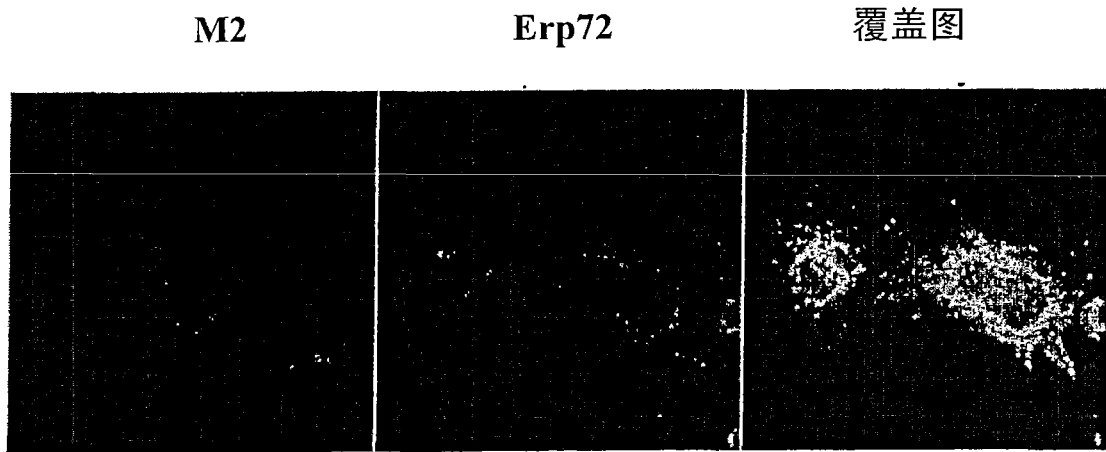


图 4

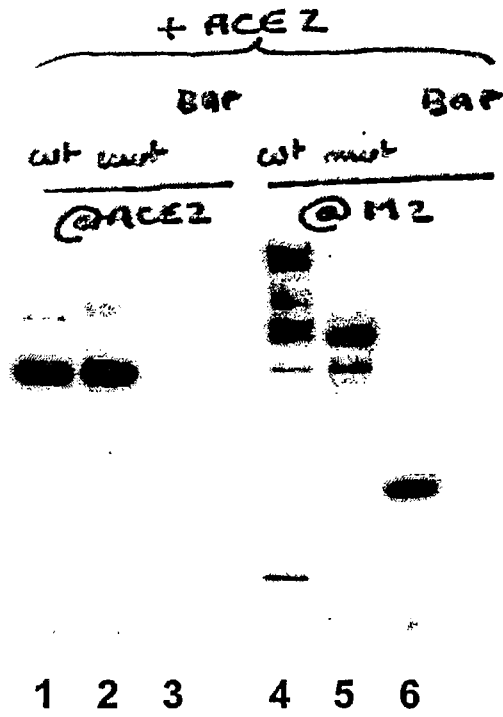


图 5

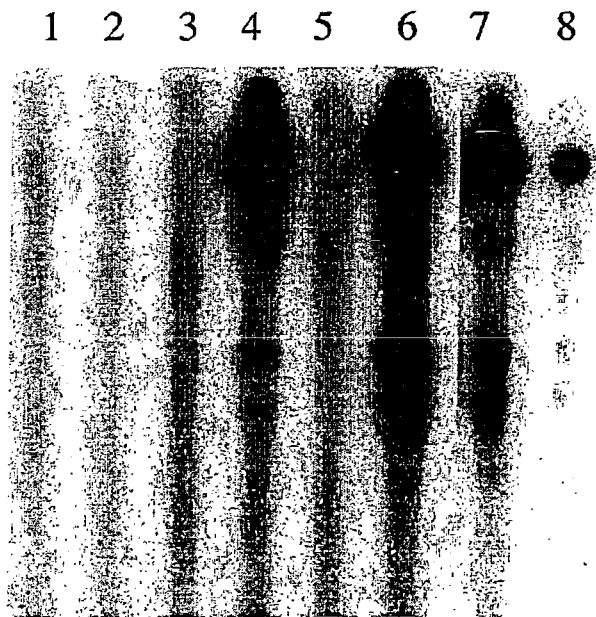


图 6

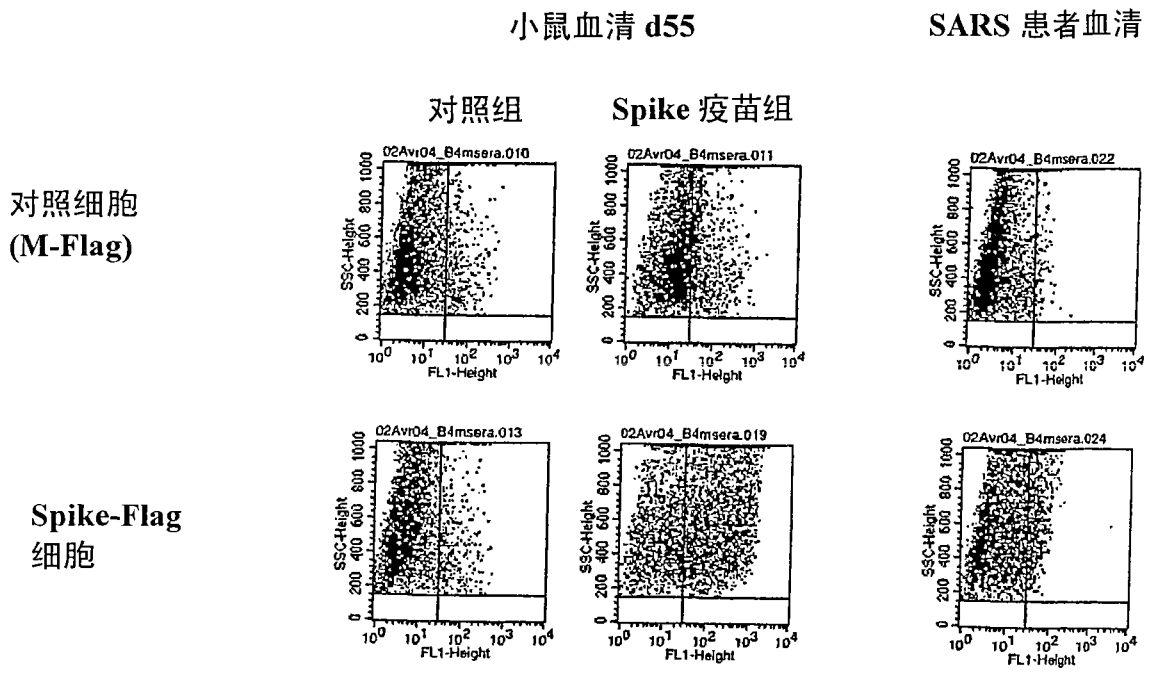
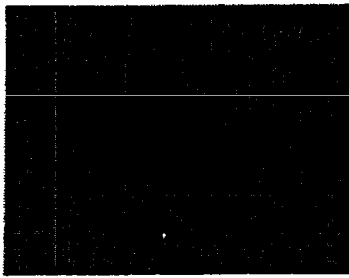
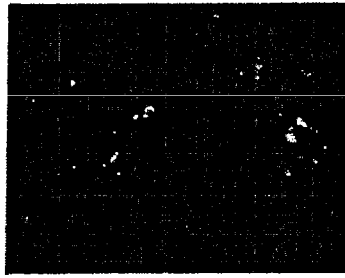


图 7

接种 Spike 的小鼠血清 d55
假感染的 FRhk4 细胞



接种 Spike 的小鼠血清 d34
感染 SARS-CoV 的 FRhk4 细胞



接种 Spike 的小鼠血清 d55
感染 SARS-CoV 的 FRhk4 细胞

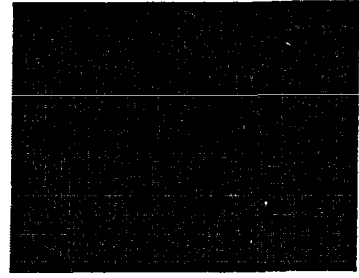


图 8

SEQ ID NO: 1 Spike-Pasteur

```

1  ATGTTTATTT TCTTAT TATT TCTTACTCTC ACTAGTGGTA GTGACCTTGA
51  CCGGTGCACC ACTTTTGATG ATGTTCAAGC TCCTAAT TAC ACTCAACATA
101 CTTTCATCTAT GAGGGGGGTT TAC TATCCTG ATGAAAT TTT TAGATCAGAC
151 ACTCTTTTAT TAACTCAGGA TTTATTCTT CCATTTT ATT CTAATGTTAC
201 AGGGTTTCAT ACTATTAATC ATA CGTTGG CAACCC TGC ATACCTTTTA
251 AGGATGGTAT TTATTTTGCT GCCACAGAGA AATCAAA TGT TGTCCGTGGT
301 TGGGTTTTTG GTTCTACCAT GAA CAACAAG TCACAGT CGG TGATTATTAT
351 TAACAATTCT ACTAATGTTG TTA TACGAGC ATGTAA CTTT GAATTGTGTG
401 ACAACCCTTT CTTTGCTGTT TCT AAACCCA TGGGTAC ACA GACACATACT
451 ATGATATTTC ATAATGCATT TAA TTGCACT TTCGAGT ACA TATCTGATGC
501 CTTTTGCTT GATGTTTCAG AAA AGTCAGG TAATTTT AAA CACTTACGAG
551 AGTTTGTGTT TAAAAATAAA GATGGGTTTC TCTATGT TTA TAAGGGCTAT
601 CAACCTATAG ATGTAGTTCG TGA TCTACCT TCTGGTT TTA ACACTTTGAA
651 ACCTATTTTT AAGTTGCCTC TTG GTATTAA CATTACA AAT TTAGAGCCA
701 TTCTTACAGC CTTTTCACCT GCT CAAGACA TTTGGGG CAC GTCAGCTGCA
751 GCCTATTTTG TTGGCTATTT AAA GCCAACT ACATTTA TGC TCAAGTATGA
801 TGAAAATGGT ACAATCACAG ATG CTGTTGA TTGTCTT CAA AATCCACTTG
851 CTGAACTCAA ATGCTCTGTT AAGAGCTTTG AGATTGA CAA AGGAATTTAC
901 CAGACCTCTA ATTTTCAGGT TGT TCCCTCA GGAGATG TTG TGAGATTCCC
951 TAATATTACA AACTTGTGTC CT TTTGGAGA GGTTTT TAAT GCTACTAAAT
1001 TCCCTTCTGT CTATGCATGG GAGAGAAAAA AAATTTT TAA TTGTGTTGCT
1051 GATTACTCTG TGCTCTACAA CTC AACATTT TTTCAA CCT TTAAGTGCTA
1101 TGGCGTTTCT GCCACTAAGT TGA ATGATCT TTGCTTC TCC AATGTCTATG
1151 CAGATTCTTT TGTAGTCAAG GGA GATGATG TAAGACA AAT AGCGCCAGGA
1201 CAAACTGGTG TTATTGCTGA TTA TAATTAT AAATTGC CAG ATGATTTTCT
1251 GGGTTGTGTC CTTGCTTGG AATA CTAGGAA CATTGAT GCT ACTTCAACTG
1301 GTAATTATAA TTATAAATAT AGG TATCTTA GACATGG CAA GCTTAGGCC
1351 TTTGAGAGAG ACATATCTAA TGT GCCTTTC TCCCCTG ATG GCAAACCTTG
1401 CACCCACCT GCTCTTAATT GTT ATTGGCC ATTAAAT GAT TATGGTTTTT
1451 ACACCACTAC TGGCATTGGC TAC CAACCTT ACAGAGT TGT AGTACTTTCT
1501 TTTGAACTTT TAAATGCACC GGC CACGGTT TGTGGA CCAA AATTATCCAC

```

图 9 (A)

1551 TGACCTTATT AAGAACCAGT GTG TCAATTT TAATTTT AAT GGA CTCACTG
 1601 GTACTGGTGT GTTAACTCCT TCT TCAAGA GATTTCACC ATTTCAACAA
 1651 TTTGGCCGTG ATGTCTCTGA TTT CACTGAT TCCGTTCCAG ATCCTAAAAC
 1701 ATCTGAAATA TTAGACATTT CA CCTTGCTC TTTTGGGGT GTAAGTGTA
 1751 TTACACCTGG AACAAATGCT TCA TCTGAAG TTGCTGT TCT ATATCAAGAT
 1801 GTTAACTGCA CTGATGTTTC TAC AGCAATC CATGCAGATC AACTCACACC
 1851 AGCTTGGCGC ATATATTCTA CTG GAAACAA TGTATTCCAG ACTCAAGCAG
 1901 GCTGTCTTAT AGGAGCTGAG CAT GTCGACA CTTCTTA TGA GTGCGACATT
 1951 CCTATTGGAG CTGGCATTG TGC TAGTTAC CATAACAG TTT CTTTATTACG
 2001 TAGTACTAGC CAAAATCTA TTG TGGCTTA TACTATG TCT TTAGGTGCTG
 2051 ATAGTTCAAT TGCTTACTCT AAT AACACCA TTGCTAT ACC TACTAACTTT
 2101 TCAATTAGCA TTACTACAGA AGT AATGCCT GTTTCTA TGG CTA AACCTC
 2151 CGTAGATTGT AATATGTACA TCT GCGGAGA TTCTACT GAA TGTGCTAATT
 2201 TGCTTCTCCA ATATGGTAGC TTT TGCACAC AACTAAA TCG TGCACCTCA
 2251 GGTATTGCTG CTGAACAGGA TCG CAACACA CGTGAAG TGT TCGCTCAAGT
 2301 CAAACAAATG TACAAAACCC CAA CTTTGAA ATATTTT GGT GGTTTAAATT
 2351 TTTCAAAAT ATTACCTGAC CCT CTAAGC CAACTAA GAG GTCTTTTATT
 2401 GAGGACTTGC TCTTTAATAA GGT GACACTC GCTGATG CTG GCTTCATGAA
 2451 GCAATATGGC GAATGCCTAG GTG ATATTAA TGCTAGA GAT CTCATTTGTG
 2501 CGCAGAAGTT CAATGGGCTT ACA GTGTTGC CACCTCT GCT CACTGATGAT
 2551 ATGATTGCTG CCTACACTGC TGC TCTAGTT AGTGGTA CTG CCACTGCTGG
 2601 ATGGACATTT GGTGCTGGCG CTG CTCTTCA AATACCT TTT GCTATGCAA
 2651 TGGCATATAG GTTCAATGGC ATT GGAGTTA CCCAAAA TGT TCTCTATGAG
 2701 AACCAAAAAC AAATCGCCAA CCA ATTTAAC AAGGCGA TTA GTCAAATTCA
 2751 AGAATCACTT ACAACAACAT CAA CTGCATT GGGCAAG CTG CAAGACGTTG
 2801 TTAACCAGAA TGCTCAAGCA TTA AACACAC TTGTTAA ACA ACTTAGCTCT
 2851 AATTTTGGT G CAATTTCAAG TG TGCTAAAT GATATC CTTT CGCGACTTGA
 2901 TAAAGTCGAG GCGGAGGTAC AAA TTGACAG GCTAATT ACA GGCAGACTTC
 2951 AAAGCCTTCA AACCTATGTA ACA CAACAAC TAATCAG GGC TGCTGAAATC
 3001 AGGGCTTCTG CTAATCTTGC TGC TACTAAA ATGTCTG AGT GTGTTCTTGG
 3051 ACAATCAAAA AGAGTTGACT TT TGFGAAA GGGCTA CCAC CTTATGTCCT
 3101 TCCCACAAGC AGCCCCGCAT GGT GTTGTCT TCCTACA TGT CACGTATGTG
 3151 CCATCCCAGG AGAGGAACTT CAC CACAGCG CCAGCAA TTT GTCATGAAGG

图 9 (B)

3 201 CAAAGCATA C TTCCTCGTG AAG GTGTTT TGTGTTT AAT GGCACCTTCTT
3 251 GGTTTATTAC ACAGAGGAAC TTC TTTTCTC CACAAAT AAT TACTACAGAC
3 301 AATACATTTG TCTCAGGAAA TTG TGATGTC GTTATTG GCA TCATTAACAA
3 351 CACAGTTTAT GATCCTCTGC AAC CTGAGCT TGACTCA TTC AAAGAAGAGC
3 401 TGGACAAGTA CTTCAAAAAT CAT ACATCAC CAGATGT TGA TCTTGGCGAC
3 451 ATTTCAGGCA TTAACGCTTC TGT CGTCAAC ATTCAAA AAG AAATTGACCG
3 501 CCTCAATGAG GTCGCTAAAA ATT TAAATGA ATCACTC ATT GACCTTCAAG
3 551 AATTGGGAAA ATATGAGCAA TAT ATTAAAT GGCCITG GTA TGTTTGGCTC
3 601 GGCTTCATTG CTGGACTAAT TGC CATCGTC ATGGTT ACAA TCTTGCTTTG
3 651 TTGCATGACT AGTTGTTGCA GTGCCTCAA GGTGC ATGC TCTTGTTGTT
3 701 CTTGCTGCAA GTTTGATGAG GAT GACGCTG AGCCAGT TCT CAAGGGTGTC
3 751 AAATTACATT ACACAGACTA CAA GGATGAC GATGACA ATA AGTAA

图 9 (C)

SEQ ID NO: 2 Spike-Pasteur-modif

1 ATGTTTATTT TCTTAT TATT TCTTACTCTC ACGAGTGGTA GTGACCTTGA
 51 CCGGTGCACC ACTTTTGATG ATG TTCAAGC TCCTAAT TAC ACTCAACATA
 101 CTTTATCTAT GAGGGGGGTT TAC TATCCTG ATGAAAT TTT TAGATCAGAC
 151 ACTCTTTATT TAACTCAGGA TTT ATTTCTT CCATTTT ATT CTAATGTTAC
 201 AGGGTTTCAT ACTATTAATC ATA CGTTTGG CAACCCT GTC ATACCTTTTA
 251 AGGATGGTAT TTATTTTGCT GCC ACAGAGA AATCAAA TGT TGTCCGTGGT
 301 TGGGTTTTTG GTTCTACCAT GAA CAACAAG TCACAGT CGG TGATTATTAT
 351 TAACAATTCT ACTAATGTTG TTA TACGAGC ATGTAAC TTT GAATTGTGTG
 401 ACAACCCCTT CTTTGCTGTT TCT AAACCCA TGGGTAC ACA GACACATACT
 451 ATGATATTCG ATAATGCATT TAA TTGCACT TTCGAGT ACA TATCTGATGC
 501 CTTTTGCTT GATGTTTCAG AAA AGTCAGG TAATTTT AAA CACTTACGAG
 551 AGTTTGTGTT TAAAAATAAA GAT GGGTTTC TCTATG TTA TAAGGGCIAT
 601 CAACCTATAG ATGTAGTTCG TGA TCTACCT TCTGGTT TTA ACACTTTGAA
 651 ACCTATTTTT AAGTTGCTC TTG GTATTAA CATTACA AAT TTAGAGCCA
 701 TTCTTACAGC CTTTTACCT GCT CAAGACA TTTGGGG CAC GTCAGCTGCA
 751 GCCTATTTTG TTGGCTATTT AA AGCCAAC ACATTT ATGC TCAAGTATGA
 801 TGAAAATGGT ACAATCACAG ATG CTGTGA TTGTTCT CAA AATCCACTTG
 851 CTGAACTCAA ATGCTCTGTT AAG AGCTTTG AGATTGA CAA AGGAATTCAC
 901 CAGACCTCTA ATTTACAGGT TGT TCCCTCA GGAGATG TTG TGAGATTCCC
 951 TAATATTACA AACTTGTGTC CTT TTGGAGA GGTTTTT AAT GCTACTAAAT
 1001 TCCCTTCTGT CTATGCATGG GAG AGAAAA AAATTC TAA TTGTGTTGCT
 1051 GATTACTCTG TGCTCTACAA CTC AACATTT TTTTCAA CCT TTAAGTGCTA
 1101 TGGCGTTTCT GCCACTAAGT TGA ATGATCT TTGCTTC TCC AATGTCTATG
 1151 CAGATCTTT TGTAGTCAAG GGA GATGATG TAAGACA AAT AGCGCCAGGA
 1201 CAAACTGGTG TTATTGCTGA TTA TAATTAT AAATTGC CAG ATGATTCAT
 1251 GGGTTGTGTC CTTGCTTGA ATA CTAGGAA CATTGAT GCT ACTTCAACTG
 1301 GTAATTATAA TTATAAATAT AGG TATCTTA GACATGG CAA GCTTAGGCC
 1351 TTTGAGAGAG ACATATCTAA TGT GCCTTTC TCCCCTG ATG GCAAACCTTG
 1401 CACCCACCT GCTCTTAATT GTT ATTGGCC ATTAAAT GAT TATGGTTTTT
 1451 ACACCACTAC TGGCATTGGC TAC CAACCTT ACAGAGT TGT AGTACTTTCT
 1501 TTTGAACTTT TAAATGCACC GGC CACGGT TGTGGAC CAA AATTATCCAC

图 10 (A)

1551 TGACCTTATT AAGAACCAGT GTG TCAATTT TAATTTT AAT GGACTCACTG
1601 GTACTGGTGT GTTAACTCCT TCT TCAAAGA GATTTCACC ATTTCAACAA
1651 TTTGGCCGTG ATGTCTCTGA TTT CACTGAT TCCGTTC GAG ATCCTAAAAC
1701 ATCTGAAATA TTAGACATTT CAC CTTGCTC TTTTGGG GGT GTAAGTGTA
1751 TTACACCTGG AACAAATGCT TCA TCTGAAG TTGCTGT TCT ATATCAAGAT
1801 GTTAACTGCA CTGATGTTTC TAC AGCAATC CATGCAG ATC AACTCACACC
1851 AGCTTGGCGC ATATATTCTA CTG GAAACAA TGTATTC CAG ACTCAAGCAG
1901 GCTGTCTTAT AGGAGCTGAG CA TGTCGACA CTTCTT ATGA GTGCGACATT
1951 CCTATTGGAG CTGGCATTTG TGC TAGTTAC CATAACAG TTT CTTTATTACG
2001 TAGTACTAGC CAAAAATCTA TTG TGGCTTA TACTATG TCT TTAGGTGCTG
2051 ATAGTTCAAT TGCTTACTCT AAT AACACCA TTGCTAT ACC TACTAATTT
2101 TCAATTAGCA TTACTACAGA AG TAATGCCT GTTTCT ATGG CTA AACCTC
2151 CGTAGATTGT AATATGTACA TCT GCGGAGA TTCTACT GAA TGTGCTAATT
2201 TGCTTCTCCA ATATGGTAGC TTT TGACAC AACTAAA TCG TGCACCTCA
2251 GGTATTGCTG CTGAACAGGA TCG CAACACA CGTGAAG TGT TCGCTCAAGT
2301 CAAACAAATG TACAAAACCC CAA CTTTGAA ATATTTT GGT GGTTTTAATT
2351 TTTCAAAAT ATTACCTGAC CCT CTAAGC CAACTAA GAG GTCTTTTATT
2401 GAGGACTTGC TCTTTAATAA GGT GACACTC GCTGATG CTG GCTTCATGAA
2451 GCAATATGGC GAATGCCTAG GTG ATATTAA TGCTAGA GAT CTCATTTGTG
2501 CGCAGAAGTT CAATGGGCTT ACA GTGTTGC CACCTCT GCT CACIGATGAT
2551 ATGATTGCTG CCTACACTGC TGC TCTAGTT AGTGGTA CTG CCACTGCTGG
2601 ATGGACATTT GGTGCTGGCG CTG CTCTTCA AATACCT TTT GCTATGCAAA
2651 TGGCATATAG GTTCAATGGC ATT GGAGTTA CCCAAA ATGT TCTCTATGAG
2701 AACCAAAAAC AAATCGCCAA CCA ATTTAAC AAGGCGA TTA GTCAAATTCA
2751 AGAATCACTT ACAACAACAT CAA CTGCATT GGGCAAG CTG CAAGACGTTG
2801 TTAACCAGAA TGCTCAAGCA TTA AACACAC TTGTTAA ACA ACTTAGCTCT
2851 AATTTTGGTG CAATTTCAAG TG TGCTAAAT GATATC CTTT CGCGACTTGA
2901 TAAAGTCGAG GCGGAGGTAC AAA TTGACAG GCTAATT ACA GGCAGACTTC
2951 AAAGCCTTCA AACCTATGTA ACA CAACAAC TAATCAG GGC TGCTGAAATC
3001 AGGGCTTCTG CTAATCTTGC TGC TACTAAA ATGTCAG AGT GTGTTCTTGG
3051 ACAATCAAAA AGAGTTGACT TTT GTGGAAA GGGCTAC CAC CTTATGTCCT
3101 TCCCACAAGC AGCCCCGCAT GGT GTTGTCT TCCTACA TGT CACGTATGTG

图 10(B)

3151 CCATCCCAGG AGAGGAACTT CAC CACAGCG CCAGCAA TTT GTCATGAAGG
3201 CAAAGCATA C TTCCTCGTG AAG GTGTTTT TGTGTTT AAT GGCACCTTCTT
3251 GGTTTATTAC ACAGAGGAAC TTC TTTTCTC CACAAAT AAT TACTACAGAC
3301 AATACATTG TCTCAGGAAA TTG TGATGTC GTTATTG GCA TCATTAACAA
3351 CACAGTTTAT GATCCTCTGC AAC CTGAGCT TGACTCA TTC AAAGAAGAGC
3401 TGGACAAGTA CTTCAAAAAT CAT ACATCAC CAGATGT TGA TCTTGGCGAC
3451 ATTTCAGGCA TTAACGCTTC TGT CGTCAAC ATTCAAA AAG AAATTGACCG
3501 CCTCAATGAG GTCGCTAAAA ATT TAAATGA ATCACTC ATT GACCTTCAAG
3551 AATTGGGAAA ATATGAGCAA TAT ATTAAAT GGCCTTG GTA TGTTTGGCTC
3601 GGCTTCATFG CTGGACTAAT TGC CATCGTC ATGGTTA CAA TCTTGCTTTG
3651 TTGCATGAC G AGTTGTTGCA GTTGCCTCAA GGGTGC ATGC TCTTGTGTT
3701 CTTGCTGCAA GTTTGATGAG GAT GACGCTG AGCCAGT TCT CAAGGGTGTC
3751 AAATTACATT ACACAGACTA CAA GGATGAC GATGACA ATA AGTAA

图 10 (C)

SEQ ID NOS: 3-4 Spike-HKU-PRC

```

CTATAGGGCGAATTGGGTA CCGCTAGCGGATCCG CGCGCCACCATGTTA TTTTCCTGCT
1 -----+-----+-----+-----+-----+-----+-----+
GATATCCCGCTTAACCCATG GCGATCGCCTAGGCG CGCGGTGGTACAAAT AAAAGGACGA
                                     M F I F L L
GTTTCTGACTCTGACCAGCG GCAGTGACCTGGACC GGTGCACCACCTTTG ATGATGTGCA
61 -----+-----+-----+-----+-----+-----+-----+
CAAAGACTGAGACTGGTCCG CGTCACTGGACCTGG CCACGTGGTGAAAAC TACTACACGT
F L T L T S G S D L D R C T T F D D V Q
GGCTCCTAATTACACTCAGC ATACTTCCTCTATGA GGGGCGTGTACTATC CTGATGAAAT
121 -----+-----+-----+-----+-----+-----+-----+
CCGAGGATTAATGTGAGTCG TATGAAGGAGATACT CCCCACATGATAG GACTACTTTA
A P N Y T Q H T S S M R G V Y Y P D E I
TTTTAGATCCGACACTCTGT ATCTGACTCAGGATC TGTTCCTGCCATTCT ATTCTAATGT
181 -----+-----+-----+-----+-----+-----+-----+
AAAATCTAGGCTGTGAGACA TAGACTGAGTCCTAG ACAAAGACGGTAAGA TAAGATTACA
F R S D T L Y L T Q D L F L P F Y S N V
GACAGGCTTTCATACTATTA ATCATACCTTTGGCA ACCCTGTGATCCCTT TTAAGGATGG
241 -----+-----+-----+-----+-----+-----+-----+
CTGTCCGAAAGTATGATAAT TAGTATGGAAACCGT TGGGACACTAGGGAA AATTCCTACC
T G F H T I N H T F G N P V I P F K D G
CATCTAATTTGCTGCCACAG AGAAGTCCAATGTGG TCGGGGATGGGTGT TCGGCTCTAC
301 -----+-----+-----+-----+-----+-----+-----+
GTAGATAAAAACGACGGTGC TCTTCAGGTACACC ACGCCCCTACCCACA AGCCGAGATG
I Y F A A T E K S N V V R G W V F G S T
CATGAACAACAAGTCCCAGT CCGTGATTATTATTA ACAATCTACTAATG TGGTGATCCG
361 -----+-----+-----+-----+-----+-----+-----+
GTACTTGTGTTCAGGGTCA GGCCTAATAATAAT TGTTAAGATGATTAC ACCACTAGGC
M N N K S O S V I I I N N S T N V V I R
AGCCTGTAACCTTGAAGTGT GTGACAACCCATTCT TTGCTGTGTCTAAGC CCATGGGCAC
421 -----+-----+-----+-----+-----+-----+-----+
TCGGACATTGAAACTTGACA CACTGTTGGGTAAGA AACGACACAGATTCG GGTACCCGTG
A C N F E L C D N P F F A V S K P M G T
ACAGACACATACTATGATCT TCGATAATGCCTTTA ATTGCACCTTCGAGT ACATCTCTGA
481 -----+-----+-----+-----+-----+-----+-----+
TGTCTGTGTATGATACTAGA AGCTATTACGGAAAT TAACGTGAAAGCTCA TGTAGAGACT
Q T H T M I F D N A F N C T F E Y I S D
TGCCTTTCCCTGGATGTGT CCGAAAAGTCCGGCA ACTTTAAGCACCTGC GAGAGTTTGT
541 -----+-----+-----+-----+-----+-----+-----+
ACGGAAAAGGGACCTACACA GGCTTTTCAGGCCGT TGAATTCGTGGACG CTCTCAAACA
A F S L D V S E K S G N F K H L R E F V
GTTTAAGAATAAGGATGGCT TTCTGTATGTGTATA AGGGCTATCAGCCTA TCGACGTGGT
601 -----+-----+-----+-----+-----+-----+-----+
CAAATTCCTTATCCTACCGA AAGACATACACATAT TCCCGATAGTCGGAT AGCTGCACCA

```

图 11 (A)

```

F K N K D G F L Y V Y K G Y Q P I D V V
GCGGATCTGCCTTCTGGCT TTAACACTCTGAAGC CTATTTTAAAGCTGC CTCTGGGCAT
661 -----+-----+-----+-----+-----+-----+
CGCGCTAGACGGAAGACCG AAATTGTGAGACTTC GGATAAAAATTCGAC GGAGACCCGTA
R D L P S G F N T L K P I F K L P L G I
TAACATTACAAATTTTCGGG CCATTCTGACAGCCT TTAGCCCTGCTCAGG ACATTTGGGG
721 -----+-----+-----+-----+-----+-----+
ATTGTAATGTTTAAAAGCCC GGTAAGACTGTTCGGA AATCGGGACGAGTCC TGTAAACCCC
N I T N F R A I L T A F S P A Q D I W G
CACCTCTGCTGCCGCTATT TTGTGGGCTATCTGA AGCCAACTACCTTTA TGCTGAAGTA
781 -----+-----+-----+-----+-----+-----+
GTGGAGACGACGGCGGATAA AACACCCGATAGAC TTCGGTTGATGGAAA TAGCACTTCAT
T S A A A Y F V G Y L K P T T F M L K Y
TGATGAAAATGGCACAATCA CAGATGCTGTGGATT GTTCTCAGAATCCAC TGGCTGAACT
841 -----+-----+-----+-----+-----+-----+
ACTACTTTTACCGTGTAGT GTCTACGACACCTAA CAAGAGTCTTAGGTG ACCGACTTGA
D E N G T I T D A V D C S Q N P L A E L
GAAGTGCTCTGTGAAGAGCT TTGAGATTGACAAGG GAATCTACCAGACCT CTAATTTCCG
901 -----+-----+-----+-----+-----+-----+
CTTCACGAGACACTTCTCGA AACTCTAACTGTCC CTTAGATGGTCTGG AGATTAAAGGC
K C S V K S F E I D K G I Y Q T S N F R
CGTGGTGCCCTCTGGAGATG TGGTGAGATTCCCTAATATTACAAACCTGT GTCCTTTTGG
961 -----+-----+-----+-----+-----+-----+
GCACCACGGGAGACCTCTAC ACCACTCTAAGGGAT TATAATGTTTGACA CAGGAAAACC
V V P S G D V V R F P N I T N L C P F G
AGAAGTGTTTAAATGCTACTA AGTTCCTTCTGTGT ATGCCTGGGAGAGAA AGAAGATTC
1021 -----+-----+-----+-----+-----+-----+
TCTTCACAAATTACGATGAT TCAAGGGAAGACACA TACGGACCCTCTCTT TCTTCTAAAG
E V F N A T K F P S V Y A W E R K K I S
TAATTGTGTGGCTGATTACT CTGTGCTGTACAACCT CCACATTTTITAGCA CCTTTAAGTG
1081 -----+-----+-----+-----+-----+-----+
ATTAACACACCGACTAATGA GACACGACATGTTGA GGTGTAAAAAATCGT GGAAATTCAC
N C V A D Y S V L Y N S T F F S T F K C
CTATGGCGTGTCTGCCACTA AGCTGAATGATCTGT GCTTCTCCAATGTGT ATGCCGATTCT
1141 -----+-----+-----+-----+-----+-----+
GATACCGCACAGACGGTGAT TCGACTTACTAGACA CGAAGAGGTTACACA TACGGCTAAG
Y G V S A T K L N D L C F S N V Y A D S
TTTTGTGGTGAAGGGAGATG ATGTGAGACAGATCG CCCAGGACAGACTG GCGTGATTGC
1201 -----+-----+-----+-----+-----+-----+
AAAACACCACTTCCCTCTAC TACACTCTGTCTAGC GGGGTCTGTCTGAC CGCACTAACG
F V V K G D D V R Q I A P G Q T G V I A
TGATTACAATTATAAGCTGC CAGATGATTCATGG GCTGTGTGCTGGCTT GGAATACTAG
1261 -----+-----+-----+-----+-----+-----+
ACTAATGTTAATATTCGACG GTCTACTAAAGTACC CGACACACGACCGAA CCTTATGATC
D Y N Y K L P D D F M G C V L A W N T R

```

图 11 (B)

```

GAACATTGATGCTACTTCCA CTGGCAATTATAATT ACAAGTATCGGTATC TGAGACATGG
1321 -----+-----+-----+-----+-----+-----+
CTTGTAACTACGATGAAGGT GACCGTTAATATTAA TGTTCATAGCCATAG ACTCTGTACC
N I D A T S T G N Y N Y K Y R Y L R H G
CAAGCTGAGGCCCTTTGAGA GAGACATCTCTAACG TGCCTTTCAGCCCTG ATGGCAAGCC
1381 -----+-----+-----+-----+-----+
GTTTCGACTCCGGGAAACTCT CTCTGTAGAGATTGC ACGGAAAGTCGGGAC.TACCGTTCCG
K L R P F E R D I S N V P F S P D G K P
TTGCACCCCACTGCTCTGA ATGTATTGGCCAC TGAATGATTATGGCT TTTACACCAC
1441 -----+-----+-----+-----+-----+
AACGTGGGGTGGACGAGACT TAACAATAACCGGTG ACTTACTAATACCGA AAATGTGGTG
C T P P A L N C Y W P L N D Y G F Y T T
TACTGGCATTGGCTACCAGC CTTACAGAGTGGTGG TGCTGTCTTTTGAAC TGCTGAATGC
1501 -----+-----+-----+-----+-----+
ATGACCGTAACCGATGGTCG GAATGTCTCACCACC ACGACAGAAAACCTG ACGACTTACG
T G I G Y Q P Y R V V V L S F E L L N A
CCCTGCCACAGTGTGGGAC CAAAGCTGTCCACTG ACCTGATTAAGAACC AGTGTGTGAA
1561 -----+-----+-----+-----+-----+
GGGACGGTGTACACACCTG GTTTCGACAGGTGAC TGGACTAATCTTGG TCACACACTT
P A T V C G P K L S T D L I K N Q C V N
CTTTAACTTTAATGGACTGA CTGGCACTGGCGTGC TGACTCCTTCTAGCA AGAGATTTCA
1621 -----+-----+-----+-----+-----+
GAAATTGAAATTACCTGACT GACCGTGACCGCAGC ACTGAGGAAGATCGT TCTCTAAAGT
F N F N G L T G T G V L T P S S K R F Q
GCCATTTACAGCAGTTTGGCC GGGATGTGTCTGATT TCACTGATTCCTGTC GAGATCCTAA
1681 -----+-----+-----+-----+-----+
CGGTAAAGTCGTCAAACCGC CCCTACACAGACTAA AGTACTAAGGCACG CTCTAGGATT
P F Q Q F G R D V S D F T D S V R D P K
GACATCTGAAATCCTGGAC ATTTCCCCTTGCTCT TTTGGCGCGTGAGC GTGATTACACC
1741 -----+-----+-----+-----+-----+
CTGTAGACTTTAGGACCTGT AAAGGGGAACGAGAA AACCGCCGCACTCGC ACTAATGTGG
T S E I L D I S P C S F G G V S V I T P
TGGAACAAATGCTTCTCTG AAGTGGCTGTGCTGT ATCAGGATGTGAACT GCACTGATGT
1801 -----+-----+-----+-----+-----+
ACCTTGTTTACGAAGGAGAC TTCACCGACACGACA TAGTCCTACACTTGA CGTGACTACA
G T N A S S E V A V L Y Q D V N C T D V
GTCTACAGCCATCCATGCCG ATCAGCTGACACCA GCTTGGCGCATCTAT TCTACTGGAAA
1861 -----+-----+-----+-----+-----+
CAGATGTCCGTAGGTACGGC TAGTCCACTGTGGTC GAACCGGTAGATAA GATGACCTTT
S T A I H A D Q L T P A W R I Y S T G N
CAATGTGTTCAGACTCAGG CCGGCTGTCTGATCG GAGCTGAGCATGTGG ACACTTCTTA
1921 -----+-----+-----+-----+-----+
GTTACACAAGTCTGAGTCC GGCCGACAGACTAGC CTCGACTCGTACACC TGTGAAGAAT
N V F Q T Q A G C L I G A E H V D T S Y
TGAGTGCACATTCCTATTG GAGCTGGCATTGTG CTAGTACCATACA GTGTCTCTGCT

```

图 11 (C)

```

1981 -----+-----+-----+-----+-----+-----+
ACTCACGCTGTAAGGATAAC CTCGACCGTAAACAC GATCAATGGTATGTC ACAGAGACGA
E C D I P I G A G I C A S Y H T V S L L
GCGGAGTACTAGCCAGAAGT CTATTGTGGCTTATA CTATGTCTCTGGGCG CTGATAGTTC
2041 -----+-----+-----+-----+-----+-----+
CGCCTCATGATCGGTCTTCA GATAACACCGAATAT GATACAGAGACCCGC GACTATCAAG
R S T S Q K S I V A Y T M S L G A D S S
CATTGCTTACTCTAATAACA CCATTGCTATCCCTA CTAACCTTTCCATTA GCATTACTAC
2101 -----+-----+-----+-----+-----+-----+
GTAACGAATGAGATTATTGT GGTAACGATAGGGAT GATTGAAAAGGTAAT CGTAATGATG
I A Y S N N T I A I P T N F S I S I T T
AGAAGTGATGCCTGTGTCTA TGGCTAAGACCTCCG TGGATTGTAATATGT ACATCTGCGG
2161 -----+-----+-----+-----+-----+-----+
TCCTCACTACGGACACAGAT ACCGATTCTGGAGGC ACCTAACATTATACA TGTAGACGCC
E V M P V S M A K T S V D C N M Y I C G
AGATTCTACCGAATGTGCTA ATCTGCTGCTGCAGT ATGGCAGCTTTTGCA CACAGCTGAA
2221 -----+-----+-----+-----+-----+-----+
TCTAAGATGGCTTACACGAT TAGACGACGACGTCA TACCGTCGAAAACGT GTGTCGACTT
D S T E C A N L L L Q Y G S F C T Q L N
TCGGGCTCTGTCTGGCATTG CTGCTGAACAGGATC GCAACACACGGGAAG TGTTCCGCTCA
2281 -----+-----+-----+-----+-----+-----+
AGCCCGAGACAGACCGTAAC GACGACTTGTCTTAG CGTTGTGTGCCCTTC ACAAGCGAGT
R A L S G I A A E Q D R N T R E V F A Q
AGTGAAGCAGATGTATAAGA CCCCACTCTGAAGT ATTTTGGCGGCTTTA ATTTTCCCA
2341 -----+-----+-----+-----+-----+-----+
TCACTTCGTCTACATATTCT GGGGTTGAGACTTCA TAAAACCGCCGAAAT TAAAAAGGGT
V K Q M Y K T P T L K Y F G G F N F S Q
GATCCTGCCTGACCCTCTGA AGCCCACTAAGCGGT CTTTATTGAGGACC TGCTGTTAA
2401 -----+-----+-----+-----+-----+-----+
CTAGGACGGACTGGGAGACT TCGGGTGATTGCGCA GAAATAACTCCTGG ACGACAAATT
I L P D P L K P T K R S F I E D L L F N
CAAAGTGACACTGGCTGATG CTGGCTTATGAAGC AGTATGGCGAATGCC TGGGCGATAT
2461 -----+-----+-----+-----+-----+-----+
GTTTCACTGTGACCGACTAC GACCGAAATACTTCG TCATACCGCTTACGG ACCCGCTATA
K V T L A D A G F M K Q Y G E C L G D I
TAATGCTAGAGATCTGATTT GTGCCCAGAAGTCA ATGGCCTGACAGTGC TGCCTCCTCT
2521 -----+-----+-----+-----+-----+-----+
ATTACGATCTCTAGACTAAA CACGGGTCTTCAAGT TACCGGACTGTACCG ACGGAGGAGA
N A R D L I C A Q K F N G L T V L P P L
GCTGACTGATGATATGATTG CTGCCTACACTGCTG CTCTGGTGTCTGGCA CTGCCACTGC
2581 -----+-----+-----+-----+-----+-----+
CGACTGACTACTATACTAAC GACGGATGTGACGAC GAGACCACAGACCGT GACGGTGACG
L T D D M I A A Y T A A L V S G T A T A
TGGATGGACATTTGGCGCTG GCGCTGCTCTGCAGA TCCCTTTTGCTATGC AGATGGCCTA
2641 -----+-----+-----+-----+-----+-----+

```

图 11 (D)

```

ACCTA CCTGTAAACCGCGAC CGCGACGAGACGTCT AGGAAAACGATACG TCTACCGGAT
G W T F G A G A A L Q I P F A M Q M A Y
TCGGTTCAATGGCATTGGAG TGACCCAGAATGTGC TGTATGAGAACCAGA AGCAGATTGC
2701 -----+-----+-----+-----+-----+
AGCCAAGTTACCGTAACCTC ACTGGGTCTTACACG ACATACTCTTGGTCT TCGTCTAACG
R F N G I G V T Q N V L Y E N Q K Q I A
CAACCAGTTTAAACAAGGCCA TTAGTCAGATT CAGG AATCCCTGACAACAA CATCCACTGC
2761 -----+-----+-----+-----+-----+
GTTGGTCAAATTGTCCGGT AATCAGTCTAAGTCC TTAGGGACTGTTGTT GTAGGTGACG
N Q F N K A I S Q I Q E S L T T T S T A
CCTGGGCAAGCTGCAGGACG TGGTGAACCAGAATG CTCAGGCCCTGAACA CACTGGTGAA
2821 -----+-----+-----+-----+-----+
GGACCCGTTTCGACGTCCTG CACCCTTGGTCTTA CGAGTCCGGGACTTG TGTGACCACTT
L G K L Q D V V N Q N A Q A L N T L V K
GCAGCTGAGCAGCAATTTTG GCGCCATTTCCAGTG TGCTGAATGATATCC TGTCCCGACT
2881 -----+-----+-----+-----+-----+
CGTCGACTCGTCGTTAAAAC CGCGGTAAGGTCAC ACGACTTACTATAGG ACAGGGCTGA
Q L S S N F G A I S S V L N D I L S R L
GGATAAAGTGGAGGCCGAAG TGCAGATTGACAGGC TGATTACAGGCAGAC TGCAGAGCCT
2941 -----+-----+-----+-----+-----+
CCTATTTACCTCCGGCTTC ACGTCTAACTGTCC GACTAATGTCCGTCT GACGTCTCGGA
D K V E A E V Q I D R L I T G R L Q S L
GCAGACCTATGTGACACAGC AGCTGATCAGGGCTG CTGAAATCAGGGCTT CTGCCAATCT
3001 -----+-----+-----+-----+-----+
CGTCTGGATACACTGTGTGC TCGACTAGTCCCGAC GACTTTAGTCCCGAA GACCGTTAGA
Q T Y V T Q Q L I R A A E I R A S A N L
GGCTGCTACTAAGATGTCTG AGTGTGTGCTGGGAC AGTCCAAGAGAGTGG ACTTTTGTGG
3061 -----+-----+-----+-----+-----+
CCGACGATGATTCTACAGAC TCACACACGACCCTG TCAGTTCTCTCAC CTGAAAACACC
A A T K M S E C V L G Q S K R V D F C G
AAAGGCTACCACTGATGT CCTTCCCACAGGCTG CCCCTCATGGAGTGG TGTTCCTGCA
3121 -----+-----+-----+-----+-----+
TTTCCGATGGTGGACTACA GGAAGGGTGTCCGAC GGGGAGTACCTCACC ACAAGGACGT
K G Y H L M S F P Q A A P H G V V F L H
TGTGACCTATGTGCCATCCC AGGAGAGGAACTTCA CCACAGCCCCAGCCA TTTGTCTATGA
3181 -----+-----+-----+-----+-----+
ACACTGGATACACGGTAGGG TCCTCTCCTTGAAGT GGTGTGCGGGTCCGT AAACAGTACT
V T Y V P S Q E R N F T T A P A I C H E
AGGCAAGGCCTACTTCCCTC GGAAGGCGTGTTCG TGTTAATGGCACTT CTGCGTTTAT
3241 -----+-----+-----+-----+-----+
TCCGTTCGGATGAAGGGAG CCCTTCCGCACAAGC ACAAATTACCGTGAA GAACCAAATA
G K A Y F P R E G V F V F N G T S W F I
TACACAGCGGAACCTCTTTA GCCCACAGATCATCA CTACAGACAATACAT TTGTGTCCGG
3301 -----+-----+-----+-----+-----+
ATGTGTGCGCCTGAAGAAAT CGGGTGTCTAGTAGT GATGTCTGTATGTA AACACAGGCC

```

图 11 (E)

```

T Q R N F F S P Q I I T T D N T F V S G
AAATTGTGATGTGGTGATTG GCATCATTAAACAACA CAGTGTATGATCCTC TGCAGCCTGA
3361 -----+-----+-----+-----+-----+-----+
TTAACACTACACCACTAAC CGTAGTAATTGTTGT GTCACATACTAGGAG ACGTCGGACT
N C D V V I G I I N N T V Y D P L Q P E
GCTGGACTCCTTCAAGGAAG AGCTGGACAAGTACT TCAAGAATCATACAT CCCCAGATGT
3421 -----+-----+-----+-----+-----+
CGACCTGAGGAAGTTCCTTC TCGACCTGTTTCATGA AGTTCTTAGTATGTA GGGGTCTACA
L D S F K E E L D K Y F K N H T S P D V
GGATCTGGGCGACATTTCCG GCATTAACGCTTCTG TGGTGAACATTCAGA AGGAAATTGA
3481 -----+-----+-----+-----+-----+
CCTAGACCCGCTGTAAGGC CGTAATTGCGAAGAC ACCACTTGTAAGTCT TCCTTTAACT
D L G D I S G I N A S V V N I Q K E I D
CCGCTGAATGAAGTGGCTA AGAATCTGAATGAAT CCCTGATGACCTGC AGGAACTGGG
3541 -----+-----+-----+-----+-----+
GGCGGACTTACTTACCGAT TCTTAGACTTACTTA GGGACTAACTGGACG TCCTTGACCC
R L N E V A K N L N E S L I D L Q E L G
CAAGTATGAGCAGTATATTA AGTGGCCTTGGTATG TGTGGCTGGGCTTCA TTGCTGGACT
3601 -----+-----+-----+-----+-----+
GTTCACTACTCGTCATATAAT TCACCGGAACCATAC ACACCGACCCGAAGT AACGACCTGA
K Y E Q Y I K W P W Y V W L G F I A G L
GATTGCCATCGTATGGTGA CAATCCTGCTGTGTI GCATGACCTCCTGTT GCAGTTGCCT
3661 -----+-----+-----+-----+-----+
CTAACGGTAGCACTACCACT GTTAGGACGACACAA CGTACTGGAGGACAA CGTCAACGGA
I A I V M V T I L L C C M T S C C S C L
GAAGGGCGCTTGCTCTTGTG GATCTTGCTGCAAGT TTGATGAGGATGACT CTGAGCCAGT
3721 -----+-----+-----+-----+-----+
CTTCCC GCGAACGAGAACAC CTAGAACGACGTTCA AACTACTCCTACTGA GACTCGGTCA
K G A C S C G S C C K F D E D D S E P V
GCTGAAGGGCGTGAAGCTGC ATTACACAGGGCCCG GCGGCGACTACAAGG ACGATGACGA
3781 -----+-----+-----+-----+-----+
CGACTTCCCGCACTTCGACG TAATGTGTCCCGGGC CGCCGCTGATGTTCC TGCTACTGCT
L K G V K L H Y T G P G G D Y K D D D D
CAAGTGATAGATCGATGCAT GGATCCGTTTAAACC GAGCTCCAGCTTTGT TCCCTTA
3841 -----+-----+-----+-----+-----+
GTTCACTATCTAGCTACGTA CCTAGGCAAATTTGG CTCGAGGTGCAAACA AGGGAAT
K * * *

```

SEQ ID NO: 3

SEQ ID NO: 4

图 11 (F)

Spike-HKU-PRC 内所含的 Spike 编码序列

```

SEQ ID NO: 6                               ATGTTTA TTTTCCTGCT
GTTTC T GACTCTG ACCAGCG GCAGTGACCTGGACC GGTGCACCAC TTTTG ATGATGTGCA
GGCTCCTAATTACACTCAGC ATACTTCCTCTATGA GGGGCGTG TACTATC CTGATGAAAT
TTTTAGATCCGACACTCTGT ATCTGACTCAGGATC TGTTTCTGCCAT TCT ATTCTAATGT
GACAGGCTTTCATACTATTA ATCATACCTTTGGCA ACCCTGTGATCCCTT TTAAGGATGG
CATCTATTTTGCTGCCACAG AGAAGTCCAATGTGG TCGGGGATGGGTGT TCGGCTCTAC
CATGAACAACAAGTCCCAGT CCGTGATTATATTA ACAATCTACTAATG TGGTGATCCG
AGCCTGTAAC TTTGAACTGT GTGACAACCCATTCT TTGCTGTGTCTAAGC CCATGGGCAC
ACAGACACATACTATGATCT TCGATAATGCCTTTA ATTGCAC TTTCCGAGT ACATCTCTGA
TGCCTTTTCCCTGGATGTGT CCGAAAAGTCCGGCA ACTTTAAGCACCTGC GAGAGTTTGT
GTTTAAGAATAAGGATGGCT TCTGTATGTGTATA AGGGCTATCAGCCTA TCGACGTGGT
GCGCGATCTGCC TTTCTGGCT TTAACACTCTGAAGC CTATTTTAAAGCTGC CTCTGGGCAT
TAACATTACAAATTTTCGG GCCATTCTGACAGCC TTTAGCCCTGCTCAG GACATTTGGGG
CACCTCTGCTGCCGCTATT TTGTGGGCTATCTGA AGCCAACTACCTTTA TGCTGAAGTA
TGATGAAAATGGCACAATCA CAGATGCTGTGGATT GTTCTCAGAATCCAC TGGCTGAACT
GAAGTGCTCTGTGAAGAGCT TTGAGATTGACAAGG GAATCTACCAGACCT CTAATTTCCG
CGTGGTGCCCTCTGGAGATG TGGTGAGATTCCCTA ATATTACAAACCTGT GTCCTTTTGG
AGAAGTGTTAATGCTACTA AGTTCCTTCTGTGT ATGCCTGGGAGAGAA AGAAGATTTC
TAATTGTGTGGCTGATTACT CTGTGCTGTACAAC CCACATTTT TAGCA CCTTTAAGTG
CTATGGCTGTCTGCCACTA AGCTGAATGATCTGT GCTTCTCCAATGTGT ATGCCGATTC
TTTTGTGGTGAAGGAGATG ATGTGAGACAGATCG CCCAGGACAGACTG GCGTGATGTC
TGATTACAATTATAAGCTGC CAGATGATTT CATGG GCTGTGTGCTGGCTT GGAATACTAG
GAACATTTGATGCTACTTCCA CTGGCAATTATAATT ACAAGTATCGGTATC TGAGACATGG
CAAGCTGAGGCCCTTTGAGA GAGACATCTCTAACG TGCCTTTCAGCCCTG ATGGCAAGCC
TTGCACCCCACTGCTCTGA ATTGTTATTGGCCAC TGAATGATTATGGCT TTTACACCAC
TACTGGCATTTGGCTACCAGC CTTACAGAGTGGTGG TGCTGTCTTTGAAC TGCTGAATGC
CCCTGCCACAGTGTGTGGAC CAAAGCTGTCCACTG ACCTGATTAAGAACC AGTGTGTGAA
CTTTAACTTTAATGGACTGA CTGGCACTGGCGTGC TGA CTCTCTAGCA AGAGATTTCA
GCCATTT CAGCAGTTTGGCC GGGATGTGTCTGATT TCACTGATTC CGTGC GAGATCCTAA
GACATCTGAAATCCTGGACA TTTCCCTTGCTCTT TTGGCGGCGTGAGCG TGATTACACC
TGGAA CAATGCTTCCTCTG AAGTGGCTGTGTGT ATCAGGATGTGAACT GCACTGATGT
GTCTA CAGCCATCCATGCCG ATCAGCTGACACCAG CTTGGCGCATCTATT CTACTGGAAA
CAATGTGTCCAGACTCAGG CCGGCTGTCTGATCG GAGCTGAGCATGTGG ACACTTCTTA
TGAGT GCGACATTCTATT GGAGCTGGCATT TGT GCTAGTTACCATA CA GTGTCTCTGCT
GCGGAGTACTAGCCAGAAGT CTATTGTGGCTTATA CTATGTCTCTGGGCG CTGATAGTTC
CATTGCTTACTCTAATAACA CCATTGCTATCCCTA CTA ACTTTTCCATTA GCATTACTAC
AGAAGTGATGCC TGTGTCTA TGGCTAAGACCTCCG TGGATTGTAATATGT ACATCTGCGG
AGATTCTACCGAATGTGCTA ATCTGCTGCTGCAGT ATGGCAGCTTTTGCA CACAGCTGAA
TCGGGCTCTGTCTGGCATTG CTGCTGAACAGGATC GCAACACACGGGAAG TGTTGCTCA
AGTGAAGCAGATGTATAAGA CCCCAACTCTGAAGT ATTTTGGCGGCTTTA ATTTTCCCA
GATCCTGCCTGACCCTCTGA AGCCCACTAAGCGGT CTTTATTGAGGACC TGCTGTTTAA
CAAAGTGACACTGGCTGATG CTGGCTTTATGAAG AGTATGGCGAATGCC TGGGCGATAT
TAATGCTAGAGATCTGATTT GTGCCAGAAGT TCA ATGGCCTGACAGTGC TGCCTCCTCT

```

图 12 (A)

GCTGACTGATGATATGATTG CTGCCTACACTGCTG CTCTGGTGTCTGGCA CTGCCACTGC
TGGATGGACATTTGGCGCTG GCGCTGCTCTGCAGA TCCCTTTTGCTATGC AGATGGCCTA
TCCGTTCAATGGCATTGGAG TGACCCAGAATGTGC TGTATGAGAAACCAG AAGCAGATTGC
CAACCAGTTTAAACAAGGCCA TTAGTCAGATTCCAGG AATCCCTGACAACAA CATCCACTGC
CCTGGGCAAGCTGCAGGACG TGGTGAACCAGAATG CTCAGGCCCTGAACA CACTGGTGAA
GCAGCTGAGCAGCAATTTTG GCGCCATTTCCAGTG TGCTGAATGATATCC TGTCCCAGCT
GGATAAAGTGGAGGCCGAAG TGCAGATTGACAGGC TGATTACAGGCAGAC TGCAGAGCCT
GCAGACCTATGTGACACAGC AGCTGATCAGGGCTG CTGAAATCAGGGCTT CTGCCAATCT
GGCTGCTACTAAGATGTCTG AGTGTGTGCTGGGAC AGTCCAAGAGAGTGG ACTTTTGTGG
AAAGGGCTACCACCTGATGT CCTTCCACAGGCTG CCCCTCATGGAGTGG TGTTCCTGCA
TGTGA CCTATGTGCCATCCC AGGAGAGGAACTTCA CCACAGCCCCAGCCA TTTGTCATGA
AGGCAAGGCCTACTTCCCT CGGGAAGGCGTGTTC GTGTTAATGGCACT TCTTGGTTTAT
TACACAGCGGAACTTCTTTA GCCCAGATCATCA CTACAGACAATACAT TTGTGTCCGG
AAATTTGTGATGTGGTGATTG GCATCATTAAACAACA CAGTGTATGATCCTC TGCAGCCTGA
GCTGGACTCCTTCAAGGAAG AGCTGGACAAGTACT TCAAGAATCATAAT CCCCAGATGT
GGATCTGGGCGACATTTCCG GCATTAACGCTTCTG TGGTGAACATTGAGA AGGAAATTGA
CCGCCTGAATGAAGTGGCTA AGAATCTGAATGAAT CCCTGATGACCTGC AGGAACTGGG
CAAGTATGAGCAGTATATTA AGTGGCCTTGGTATG TGTGGCTGGGCTTCA TTGCTGGACT
GATTGCCATCGTGATGGTGA CAATCCTGCTGTGTT GCATGACCTCCTGTT GCAGTTGCCT
GAAGGGCGCTTGCTCTTGTG GATCTTGCTGCAAGT TTGATGAGGATGACT CTGAGCCAGT
GCTGAAGGGCGTGAAGCTGC ATTACACA

图 12 (B)

Spike-HKU-PRC 内所含的 SARS CoV Spike 多肽序列

MFIFLLFLTLTSGSDLRCTTFDDVQAPNYTQHTSSMRGVYYPDEIFRSDTLYLTQDLFLPFYSNVTGFHT
 INHTFGNPVI PFKDGIYFAATEKSNVVRGWVFGSTMNKSQSVIIINNSTNVVIRACNFELCDNPFPAVSK
 PMGTQTHMIFDNAFNCTFEYISDAFSLDVSEKSGNFKHLREFVFKNKDGFYVYKGYQPIDVVRDLPSGF
 NTLKPIFKLPLGINITNFRAILTAFSPAQDIWGTSAAYFVGYLKPTTFMLKYDENGTTDAVDCSQNPLA
 ELKCSVKSEIDKGIYQTSNFRVVPVSGDVVRFNITNLCPFGEVFNATKFPSVYAWERKKISNCVADYSVL
 YNSTFFSTFKCYGVSATKLNLCFSNVYADSFVVKGDDVRQIAPGQTGVIADYNYKLPDDFMGCVLWNTN
 NIDATSTGNYNKYRYLRHGKLRPFERDISNVVFPSPDGKCTPPALNCYWPLNDYGFYTTTGIGYQPYRVV
 VLSFELLNAPATVCGPKLSTDLIKNQCVNFNFNGLTGTGVLTTPSSKRFQPFQFGRDVSDFIDSVRDPKTS
 EILDISPCSFGGVSVITPGTNASSEVAVLYQDVNCTDVSTAIHADQLTPAWRIYSTGNVVFQTAGCLIGA
 EHVDTSEYCDIPIGAGICASYHTVSLLRSTSQKSIVAYTMSLGADSSIAYSNNTIAIPTNFSISITTEVMP
 VSMAKTSVDCNMYICGDSTECANLLLQYGSFCTQLNRALSGIAAEQDRNTREVFAQVKQMYKPTLKYFGG
 FNFSQILPDPLKPTKRSFIEDLLFNKVTLADAGFMKQYGECLGDINARDLICAQKFNGLTVLPPLLTDDMI
 AAYTAALVSGTATAGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKQIANQFNKAI SQIQESLTTT
 STALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLOSLQTYVTQQL
 IRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQAAPHGVVFLHVTVVPSQERNFTTAPAICH
 EGKAYFPREGVVFVNGTSWFITQRNFFSPQIITDNTFVSGNCDVVIGIINNVTYDPLQPELDSFKEELDK
 YFKNHTSPDVLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYVWLGFIAGLI
 AIVMVTILLCCMTSCCSCLKGACSCGSCCKFDEDDSEPVLLKGVKLHYT
 SEQ ID NO: 7

图 13

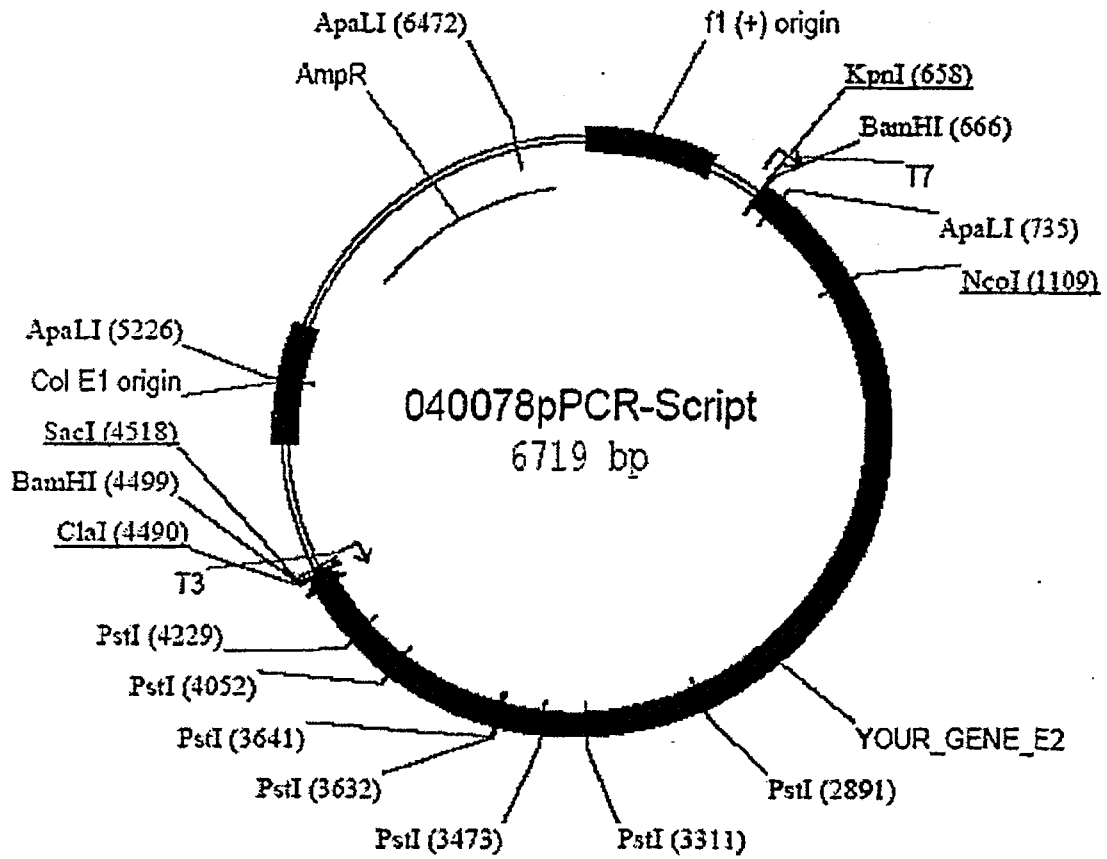


图 14

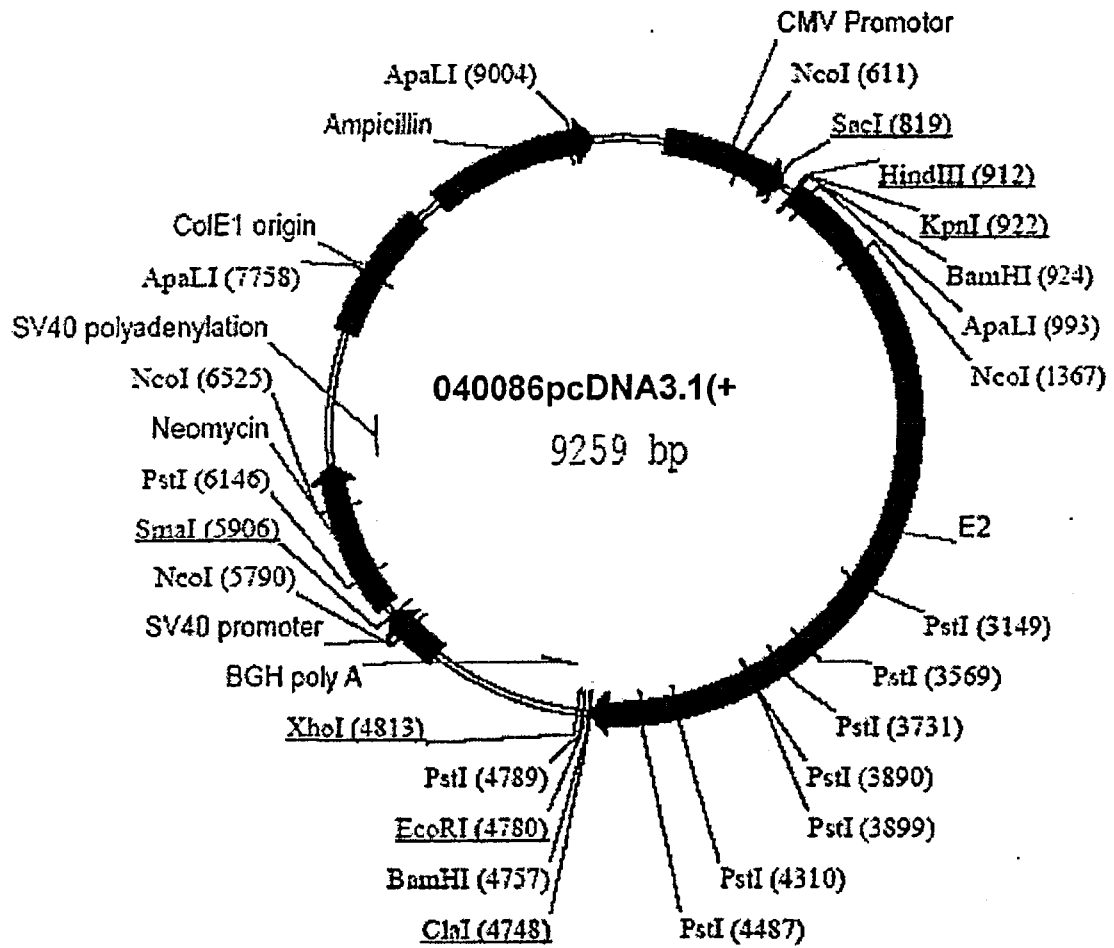


图 15

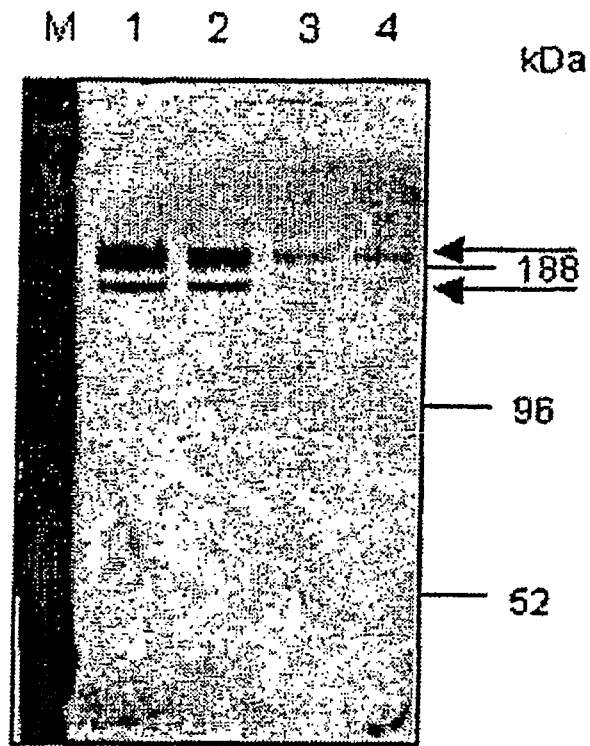


图 16

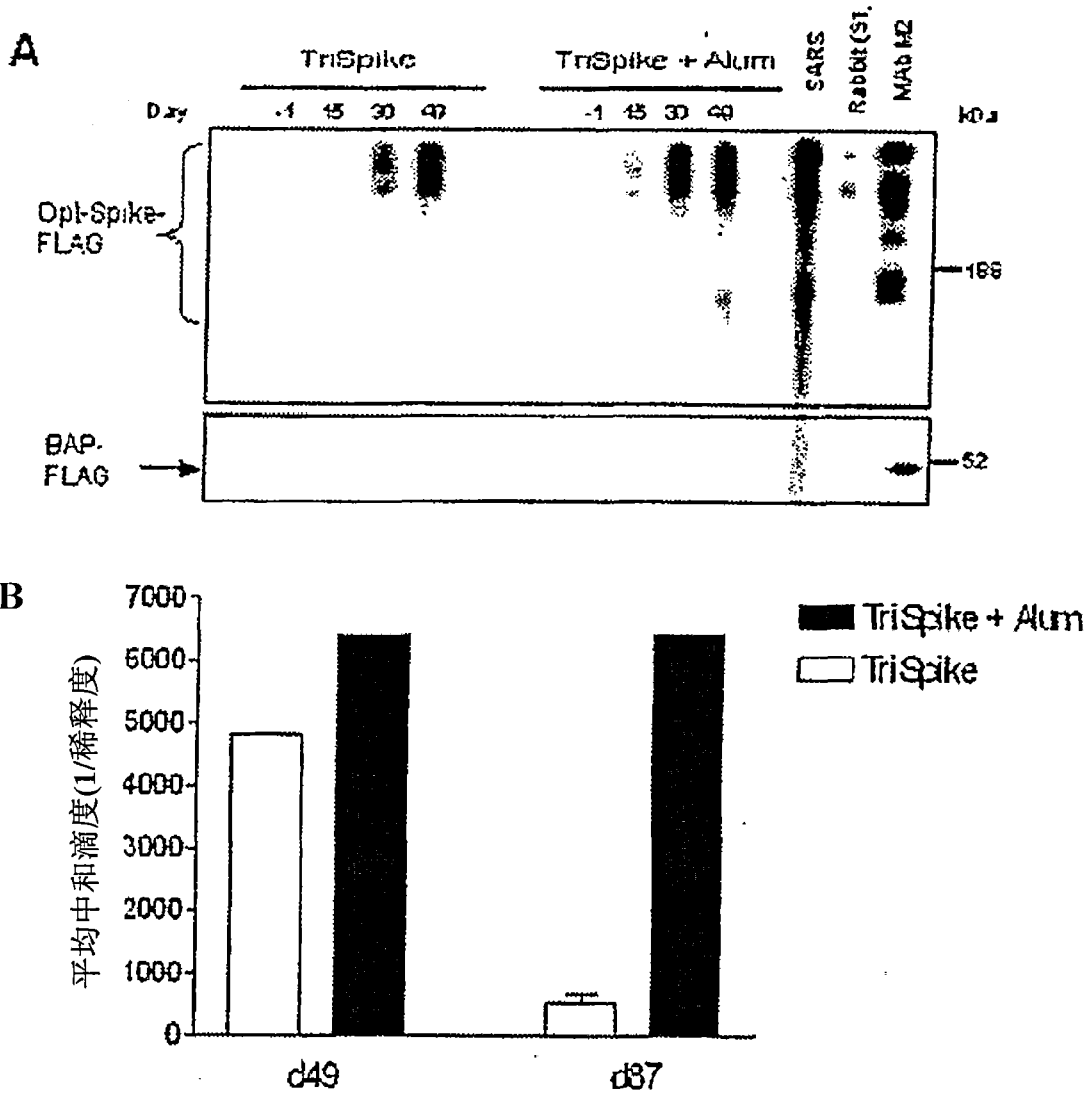


图 17

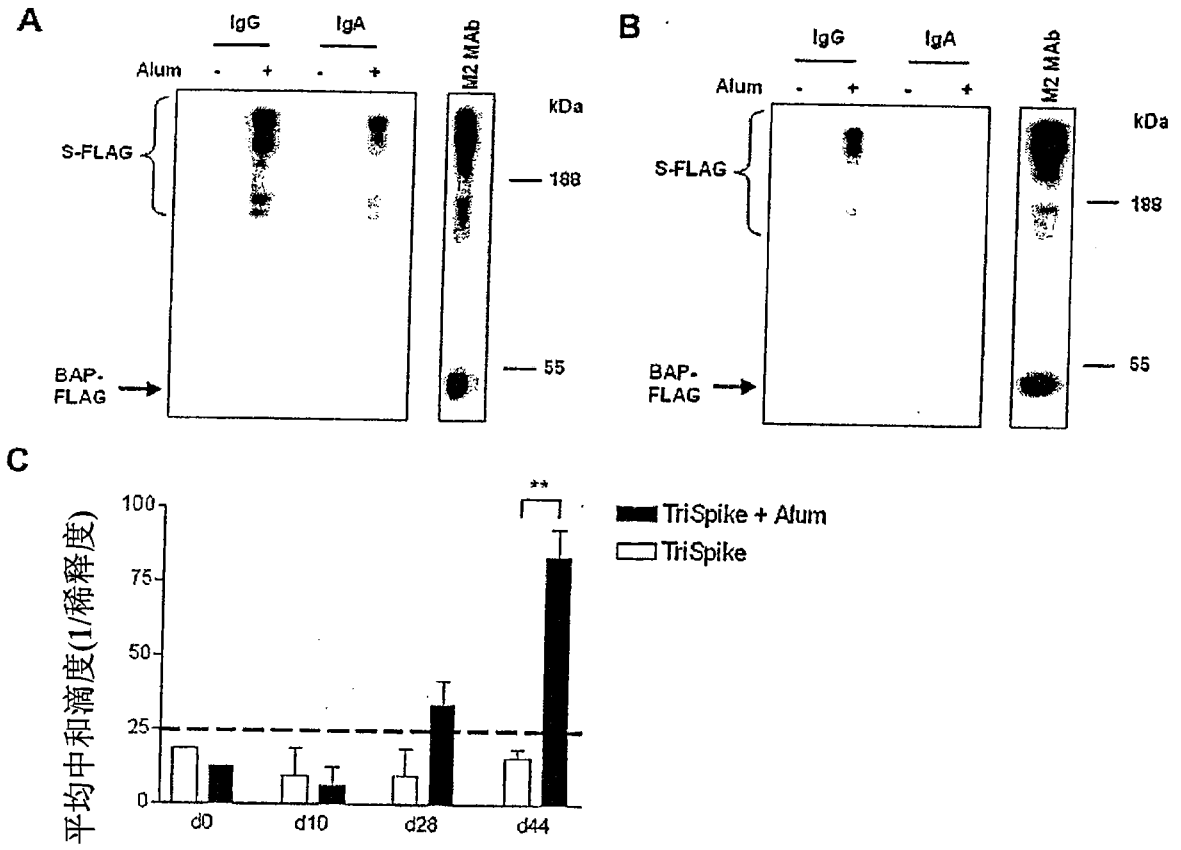


图 18

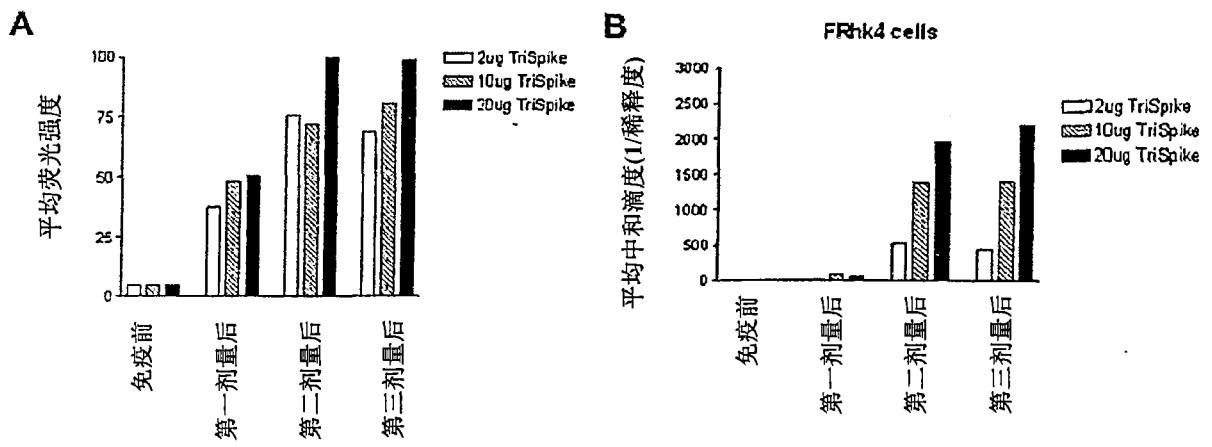


图 19