(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 117957016 A
(43)申请公布日 2024.04.30

(21)申请号 202180054553.3
(22)申请日 2021.07.07
(30)优先权数据
2010425.3 2020.07.07 GB
(85)PCT国际申请进入国家阶段日
2023.03.03
(86)PCT国际申请的申请数据
PCT/IB2021/056102 2021.07.07
(87)PCT国际申请的公布数据
WO2022/009121 EN 2022.01.13
(71)申请人 斯皮科纳股份有限公司
地址 美国纽约州纽约市
(72)发明人 阿加塔·法齐奥
埃马努埃莱·蒙托莫利
卡斯帕·班齐格 卡洛斯·昆托
莱因哈德·格卢克
(74)专利代理机构 北京骥驰知识产权代理有限公司 11422
专利代理师 唐晓峰
(51)Int.Cl.
A61K 39/12(2006.01)
A61P 31/16(2006.01)
A61P 31/14(2006.01)

权利要求书3页 说明书75页
序列表59页 附图6页

(54)发明名称

SARS-COV-2与流感联合疫苗

(57)摘要

本发明涉及对抗流感和COVID-19的联合疫苗。特别地,本发明涉及包含一种或多种流感病毒抗原和一种或多种SARS-CoV-2(冠状病毒SARS-CoV-2)抗原,特别是一种或多种SARS-CoV-2刺突蛋白抗原的联合疫苗,以及包含编码所述抗原的多核苷酸的疫苗,以及这样的疫苗用于治疗或预防COVID-19(SARS-CoV-2感染)和流感感染。

CN 117957016 A

1.一种流感-COVID-19联合疫苗,其包含:

(a)流感血凝素(HA)或其免疫原性片段;和

(b)一种或多种源自SARS-CoV-2的抗原或其免疫原性片段;

其中该抗原能够引发对抗流感和COVID-19的免疫应答和保护。

2.权利要求1的流感-COVID-19联合疫苗,其还包含流感神经氨酸酶(NA)或其免疫原性片段。

3.权利要求1或2的流感-COVID-19联合疫苗,其中:

(a)流感HA或其免疫原性片段是:

(i)包含在灭活流感病毒粒子中;

(ii)重组HA或其免疫原性片段;

(iii)包含HA或其免疫原性片段的融合蛋白;或

(iv)由RNA或DNA疫苗编码;和/或

(b)流感NA或其免疫原性片段是:

(i)包含在灭活的流感病毒粒子中;

(ii)重组NA或其免疫原性片段;

(iii)包含NA或其免疫原性片段的融合蛋白;或

(iv)由RNA或DNA疫苗编码;和/或

(c)一种或多种源自SARS-CoV-2的抗原或其免疫原性片段是:

(i)至少一种重组SARS-CoV-2刺突蛋白或其免疫原性片段;

(ii)至少一种融合蛋白,其包含SARS-CoV-2刺突蛋白或其免疫原性片段;

(iii)至少一种病毒样颗粒(VLP),其包含SARS-CoV-2刺突蛋白或其免疫原性片段;

(iv)至少一种编码重组SARS-CoV-2刺突蛋白或其免疫原性片段的多核苷酸;或

(v)由至少一种RNA或DNA疫苗编码。

4.前述任一项权利要求的流感-CoVID-19联合疫苗,其中流感HA或其免疫原性片段和流感NA或其免疫原性片段包含在灭活流感病毒粒子中,并且一种或多种源自SARS-CoV-2的抗原或其免疫原性片段是:(i)至少一种包含SARS-CoV-2刺突蛋白或其免疫原性片段的融合蛋白或(ii)至少一种包含SARS-CoV-2刺突蛋白或其免疫原性片段的病毒样颗粒(VLP)。

5.权利要求1或2的流感-COVID-19联合疫苗,其中:

(a)流感HA或其免疫原性片段包含在减毒活流感病毒粒子中;

(b)流感NA或其免疫原性片段包含在减毒活流感病毒粒子中;和/或

(c)一种或多种源自SARS-CoV-2的抗原或其免疫原性片段包含在活病毒载体中。

6.权利要求5的流感-COVID-19联合疫苗,其中包含一种或多种源自SARS-CoV-2的抗原或其免疫原性片段的活病毒载体是:

(a)腺病毒载体;

(b)麻疹病毒载体;

(c)腮腺炎病毒载体;

(d)风疹病毒载体;

(e)水痘病毒载体;

(f)脊髓灰质炎病毒载体;或

(g)黄热病病毒载体。

7.前述任一项权利要求的流感-COVID-19联合疫苗,其还包含佐剂。

8.权利要求7的流感-COVID-19联合疫苗,其中所述佐剂是细胞(Th1)和体液(Th2)免疫应答的刺激剂。

9.前述任一项权利要求的流感-COVID-19联合疫苗,其中所述佐剂包含角鲨烯水包油乳液、铝盐或单磷酰脂质A(MPL)。

10.前述任一项权利要求的流感-COVID-19联合疫苗,其中一种或多种源自SARS-CoV-2的抗原选自:

(a)来自SARS-CoV-2的与SEQ ID NO:1具有至少90%同一性的刺突蛋白,或其与所述刺突蛋白具有共同抗原交叉反应性的片段;

(b)融合蛋白,其包含来自与SEQ ID NO:1具有至少90%同一性的SARS-CoV-2的刺突蛋白,或其与所述刺突蛋白具有共同抗原交叉反应性的片段;

(c)VLP,其包含来自SARS-CoV-2的与SEQ ID NO:1具有至少90%同一性的刺突蛋白,或其与所述刺突蛋白具有共同抗原交叉反应性的片段;

(d)编码来自SARS-CoV-2的与SEQ ID NO:1具有至少90%同一性的刺突蛋白或其与所述刺突蛋白具有共同抗原交叉反应性的片段的多核苷酸;或

(e)表达来自SARS-CoV-2的与SEQ ID NO:1具有至少90%同一性的刺突蛋白或其与所述刺突蛋白具有共同抗原交叉反应性的片段的病毒载体、RNA疫苗或DNA质粒,

其中任选地,SARS-CoV-2刺突蛋白的片段包含SARS-CoV-2刺突蛋白的受体结合结构域(RBD)或由其组成,优选与SEQ ID NO:15具有至少90%的同一性。

11.前述任一项权利要求的流感-COVID-19联合疫苗,其中一种或多种源自SARS-CoV-2的抗原是包含SARS-CoV-2刺突蛋白或其免疫原性片段的融合蛋白,并且还包含:

(a)乙型肝炎表面抗原,或与所述乙型肝炎表面抗原具有共同抗原交叉反应性的其片段;

(b)HPV 18L1蛋白,或其与所述HPV 18L1蛋白具有共同抗原交叉反应性的片段;

(c)戊型肝炎P239蛋白,或其与所述戊型肝炎P239蛋白具有共同抗原交叉反应性的片段;和/或

(d)HPV 16L1蛋白,或其与所述HPV 16L1蛋白具有共同抗原交叉反应性的片段。

12.权利要求11的流感-COVID-19联合疫苗,其中:

(a)融合蛋白由包含与SEQ ID NO:3、5、6、8、26、27、29、30或32中任一个具有至少90%同一性的核酸序列或由其组成的多核苷酸编码;和/或

(b)融合蛋白包含与SEQ ID NO:9、10、11、12、28、31或33中的任一个具有至少90%同一性的氨基酸序列或由其组成。

13.前述任一项权利要求的流感-CoVID-19联合疫苗,其中一种或多种源自SARS-CoV-2的抗原是包含SARS-CoV-2刺突蛋白或其免疫原性片段的VLP,其中所述VLP包含如权利要求11或12所定义的融合蛋白或由其组成。

14.前述任一项权利要求的流感-COVID-19联合疫苗,其中流感HA或其免疫原性片段和流感NA或其免疫原性片段包含在:

(a)季节性流感疫苗,特别是季节性3价流感疫苗或季节性4价流感疫苗;

(b)单价大流行性流感疫苗;或

(c)通用流感疫苗。

15.前述任一项权利要求的流感-COVID-19联合疫苗,其用于治疗和/或预防COVID-19和流感的方法中。

16.流感HA或其免疫原性片段;和源自SARS-CoV-2的抗原或其免疫原性片段,和任选的流感NA或其免疫原性片段在制备用于治疗和/或预防COVID-19和流感的药物中的用途,其中所述药物是如权利要求1-14任一项所定义的流感-COVID-19联合疫苗。

17.一种使受试者针对流感和COVID-19免疫的方法,包括将治疗有效量的权利要求1-14任一项所定义的流感-COVID-19联合疫苗施用于所述受试者。

18.权利要求15的流感-COVID-19联合疫苗、权利要求16的用途或权利要求17的方法,其中流感-COVID-19联合疫苗以10至14个月的间隔施用,任选其中流感-COVID-19联合疫苗以约12个月的间隔施用。

SARS-COV-2与流感联合疫苗

发明领域

[0001] 本发明涉及针对流感和COVID-19的组合疫苗。特别地,本发明涉及包含一种或多种流感病毒抗原和一种或多种SARS-CoV-2抗原(优选至少一种SARS-CoV-2(冠状病毒2019-nCoV)刺突蛋白抗原)的联合疫苗,以及包含编码所述抗原的多核苷酸的疫苗,以及用于治疗或预防COVID-19(SARS-CoV-2感染)和流感感染的此类疫苗。

[0002] 发明背景

[0003] 截至2020年6月29日,全世界共有超过10,000,000人被证实为COVID-19阳性(由严重急性呼吸综合征冠状病毒2(SARS-CoV-2或冠状病毒2019-nCoV)引起的疾病)。截至同日,全球已经记录了超过500,000例由于COVID-19导致的死亡。

[0004] 大多数感染SARS-CoV-2的患者经历轻度至中度症状,包括高温或发热、咳嗽、呼吸短促、疲劳以及个体嗅觉或味觉丧失或改变。一些患者进展为严重疾病,其可能涉及急性呼吸窘迫综合征(ARDS)、细胞因子风暴、多器官衰竭、脓毒性休克和血凝。此外,一些SARS-CoV-2感染检测呈阳性的患者无症状,或症状轻微,除非进行检测,否则诊断困难。迄今为止的证据表明,这些无症状患者脱落SARS-CoV-2病毒颗粒(通常比有症状感染的患者更长),因此仍然可以有效地传播SARS-CoV-2病毒。

[0005] 与SARS-CoV-2感染相关的症状范围广泛,以及无症状患者的存在,使得确定COVID-19的流行病学特征变得更加困难。此外,至少一项研究表明,大多数无症状和有症状的患者在康复后仅8周就降低了针对SARS-CoV-2的IgG和中和抗体水平。一些临床数据表明,显著比例的无症状患者(40%)以及较少数量的有症状性感染的患者(~13%)在恢复早期对IgG呈血清阴性(Long等Nat.Med.2020https://doi.org/10.1038/s41591-020-0965-6)。因此,尽管SARS-CoV-2疫苗的开发是一个全球研究巨大驱动力的主题,但现有证据表明任何对SARS-CoV-2感染产生的免疫力在本质上可能都是短期的。因此,目前需要开发COVID-19疫苗,其可用于产生和维持针对SARS-CoV-2感染和COVID-19疾病的保护性免疫。此外,需要提供可以容易地整合到现有公共健康疫苗接种程序和时间表中的疫苗(考虑与疫苗组分抑制相关的问题),并且需要大规模且廉价地生产此类疫苗。

[0006] 本发明通过提供流感-COVID-19联合疫苗解决了上述一个或多个需求。这些联合疫苗包含一种或多种流感病毒抗原和一种或多种SARS-CoV-2抗原,优选至少一种SARS-CoV-2(冠状病毒2019-nCoV)刺突蛋白抗原,或一种或多种编码所述抗原的多核苷酸,允许使用已经针对流感病毒制定的现有公共卫生计划每年加强针对SARS-CoV-2的免疫力。

[0007] 发明概述

[0008] 迄今为止,虽然有许多SARS-CoV-2疫苗正在开发和/或临床试验中,但还没有批准的可用于一般用途的疫苗。此外,现有证据表明,针对SARS-CoV-2的免疫可能是相对短暂的。

[0009] 本发明人先前已经开发了编码SARS-CoV-2刺突蛋白的多核苷酸,所述多核苷酸提供增加的SARS-CoV-2刺突蛋白的表达水平和持续时间,同时保留天然刺突蛋白的构象。

[0010] 本发明人现在已经证明了包含其SARS-CoV-2刺突蛋白的疫苗组合物可以成功地

与流感病毒疫苗联合,而没有预期的在联合疫苗产品的生产中常见的疫苗组分抑制的问题。此外,虽然标准流感疫苗不含佐剂,但佐剂Addavax®可以成功地掺入根据本发明的SARS-CoV-2/流感疫苗中。在现有的流感公共健康疫苗计划中,每年针对SARS-CoV-2感染接种疫苗,有可能增强对SARS-CoV-2的免疫力,同时实现良好的患者依从性。

[0011] 因此,本发明提供了一种流感-COVID-19联合疫苗,其包含:(a)流感血凝素(HA)或其免疫原性片段;和(b)一种或多种源自SARS-CoV-2的抗原或其免疫原性片段;其中抗原能够引发对抗流感和COVID-19两者的免疫应答和保护。

[0012] 所述流感-COVID-19联合疫苗还可包含流感神经氨酸酶(NA)或其免疫原性片段。该流感HA或其免疫原性片段可以:(i)包含在灭活的流感病毒粒子中;(ii)重组HA或其免疫原性片段;(iii)包含HA或其免疫原性片段的融合蛋白;或(iv)由RNA或DNA疫苗编码。流感NA或其免疫原性片段可以:(i)包含在灭活的流感病毒粒子中;(ii)重组NA或其免疫原性片段;(iii)包含NA或其免疫原性片段的融合蛋白;或(iv)由RNA或DNA疫苗编码。一种或多种源自SARS-CoV-2的抗原或其免疫原性片段可以是:(i)至少一种重组SARS-CoV-2刺突蛋白或其免疫原性片段;(ii)至少一种包含SARS-CoV-2刺突蛋白或其免疫原性片段的融合蛋白;(iii)至少一种包含SARS-CoV-2刺突蛋白或其免疫原性片段的病毒样颗粒(VLP);(iv)至少一种编码重组SARS-CoV-2刺突蛋白或其免疫原性片段的多核苷酸;或(v)由至少一种RNA或DNA疫苗编码。

[0013] 在本发明的流感-COVID-19联合疫苗中,(i)流感HA或其免疫原性片段和流感NA或其免疫原性片段可以包含在灭活的流感病毒粒子中;和(ii)源自SARS-CoV-2的一种或多种抗原或其免疫原性片段可以是:(i)至少一种包含SARS-CoV-2刺突蛋白或其免疫原性片段的融合蛋白或(ii)至少一种包含SARS-CoV-2刺突蛋白或其免疫原性片段的病毒样颗粒(VLP)。

[0014] 在本发明的流感-COVID-19联合疫苗中:(a)流感HA或其免疫原性片段可以包含在活的减毒流感病毒粒子中;(b)流感NA或其免疫原性片段可以包含在活的减毒流感病毒粒子中;和/或(c)源自SARS-CoV-2的一种或多种抗原或其免疫原性片段可以包含在活病毒载体中。所述包含一种或多种源自SARS-CoV-2的抗原或其免疫原性片段的活病毒载体可以是:腺病毒载体;麻疹病毒载体;腮腺炎病毒载体;风疹病毒载体;水痘病毒载体;脊髓灰质炎病毒载体;或黄热病病毒载体。

[0015] 本发明的流感-COVID-19联合疫苗还可以包含佐剂。所述佐剂通常是细胞(Th1)和/或体液(Th2)免疫应答的刺激物,优选两者。所述佐剂可包含角鲨烯水包油乳液、铝盐或单磷酰脂质A(MPL)。

[0016] 一种或多种源自SARS-CoV-2的抗原可以选自:(a)与SEQ ID NO:1具有至少90%同一性的来自SARS-CoV-2的刺突蛋白,或其与所述刺突蛋白具有共同抗原交叉反应性的片段;(b)融合蛋白,其包含与SEQ ID NO:1具有至少90%同一性的来自SARS-CoV-2的刺突蛋白或其与所述刺突蛋白具有共同的抗原交叉反应性的片段;(c)VLP,其包含与SEQ ID NO:1具有至少90%同一性的来自SARS-CoV-2的刺突蛋白,或其与所述刺突蛋白具有共同抗原交叉反应性的片段;(d)编码与SEQ ID NO:1具有至少90%同一性的来自SARS-CoV-2的刺突蛋白或其与所述刺突蛋白具有共同抗原交叉反应性的片段的多核苷酸;或(e)表达与SEQ ID NO:1具有至少90%同一性的来自SARS-CoV-2的刺突蛋白或其与所述刺突蛋白具有共同抗

原交叉反应性的片段的病毒载体、RNA疫苗或DNA质粒;其中任选地,SARS-CoV-2刺突蛋白的片段包含SARS-CoV-2刺突蛋白的受体结合结构域(RBD)或由SARS-CoV-2刺突蛋白的受体结合结构域(RBD)组成,优选与SEQ ID NO:15具有至少90%同一性。

[0017] 一种或多种源自SARS-CoV-2的抗原可以是融合蛋白,其包含SARS-CoV-2刺突蛋白或其免疫原性片段,并且还包含:(a)乙型肝炎表面抗原,或其与所述乙型肝炎表面抗原具有共同抗原交叉反应性的片段;(b)HPV 18L1蛋白或其与所述HPV 18L1蛋白具有共同抗原交叉反应性的片段;(c)戊型肝炎P239蛋白或其与所述戊型肝炎P239蛋白具有共同抗原交叉反应性的片段;和/或(e)HPV16L1蛋白或其与所述HPV 16L1蛋白具有共同抗原交叉反应性的片段。所述融合蛋白可以:(a)由包含与SEQ ID NO:3、5、6或8、26、27、29、30或32中的任一个具有至少90%同一性的核酸序列或由其组成的多核苷酸编码;和/或(b)包含与SEQ ID NO:9、10、11、12、28、31或33中的任一个具有至少90%同一性的氨基酸序列或由其组成。

[0018] 源自SARS-CoV-2的一种或多种抗原可以是包含SARS-CoV-2刺突蛋白或其免疫原性片段的VLP,其中所述VLP包含本发明的融合蛋白或由本发明的融合蛋白组成。

[0019] 流感HA或其免疫原性片段和流感NA或其免疫原性片段可以包含在:(a)季节性流感疫苗,特别是季节性3价流感疫苗或季节性4价流感疫苗;(b)单价大流行性流感疫苗;或(c)通用流感疫苗。

[0020] 本发明还提供了如本文所述的流感-COVID-19联合疫苗,其用于治疗和/或预防COVID-19和流感的方法中。

[0021] 本发明进一步提供了流感HA或其免疫原性片段;和源自SARS-CoV-2的抗原或其免疫原性片段,和任选的流感NA或其免疫原性片段,在制备用于治疗和/或预防COVID-19和流感的药物中的用途,其中所述药物是如本文所定义的流感-COVID-19联合疫苗。

[0022] 本发明还提供了一种使受试者针对流感和COVID-19免疫的方法,包括将治疗有效量的本文定义的流感-COVID-19联合疫苗施用于所述受试者。

[0023] 流感-COVID-19联合疫苗可以以10至14个月的间隔施用,任选其中流感-COVID-19联合疫苗以约12个月的间隔施用。

附图说明

[0024] 图1:冠状病毒结构和结构蛋白功能的示意图。

[0025] 图2:使用兔抗CoV-S(1:250,中)和小鼠抗HBSAg-$(EAAAK)_3$-RBD(1:1000,右)的HBsAg-$(EAAAK)_3$-CoV-S的SDS Page(左)和Western印迹(中和右)。

[0026] 图3:显示了通过ELISA测定免疫后14天的小鼠血清定量的抗HBsAg-$(EAAAK)_3$-CoV-S IgG滴度的图,所述小鼠用HBsAg-$(EAAAK)_3$-CoV-S蛋白单独免疫和与流感疫苗VAXIGRIP联合免疫,用两种不同的佐剂:ALU-280和Addavax配制。

[0027] 图4:A显示了通过ELISA测定免疫后14天的小鼠血清定量的抗HBsAg-$(EAAAK)_3$-RBD IgG滴度的图,所述小鼠用HBsAg-$(EAAAK)_3$-RBD免疫,用两种不同的佐剂:ALU-280和Addavax配制。B通过ELISA测定免疫后14天的小鼠血清定量的抗HBsAg-$(EAAAK)_3$-CoV-S IgG和抗HBsAg-$(EAAAK)_3$-RBD IgG滴度的比较,所述小鼠用HBsAg-$(EAAAK)_3$-CoV-S或HBsAg-$(EAAAK)_3$-RBD免疫,用两种不同的佐剂:ALU-280和Addavax配制。

[0028] 图5:显示了通过ELISA测定免疫后42天的小鼠血清定量的抗HBsAg-$(EAAAK)_3$-

CoV-S IgG滴度的图,所述小鼠用HBsAg-(EAAAK)$_3$-CoV-S蛋白单独免疫和与流感疫苗VAXIGRIP联合免疫,用两种不同的佐剂:ALU-280和Addavax配制。

[0029] 图6:A显示了通过ELISA测定免疫后42天的小鼠血清定量的抗HBsAg-(EAAAK)$_3$-RBD IgG滴度的图,所述小鼠用HBsAg-(EAAAK)$_3$-RBD免疫,用两种不同的佐剂:ALU-280和Addavax配制。B通过ELISA测定免疫后42天的小鼠血清定量的抗HBsAg-(EAAAK)$_3$-CoV-S IgG和抗HBsAg-(EAAAK)$_3$-RBD IgG滴度的比较,所述小鼠用HBsAg-(EAAAK)$_3$-CoV-S或HBsAg-(EAAAK)$_3$-RBD免疫,用两种不同的佐剂:ALU-280和Addavax配制。

[0030] 发明详述

[0031] 定义

[0032] 除非另有定义,否则本文使用的所有技术和科学术语具有与本公开所属领域的普通技术人员通常理解的含义相同的含义。Singleton等,DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY,第20版,John Wiley and Sons,New York(1994)和Hale&Marham,THE HARPER COLLINS DICTIONARY OF BIOLOGY,HARPER PERENNIAL,NY(1991)为技术人员提供了本公开中使用的许多术语的通用词典。术语的含义和范围应该是清楚的;然而,在任何潜在歧义的情况下,本文提供的定义优先于任何字典或外在定义。应当理解,本发明不限于本文所述的特定方法、方案和试剂等,且因此可以变化。

[0033] 本公开不受本文公开的示例性方法和材料的限制,并且与本文描述的那些类似或等同的任何方法和材料可用于实践或测试本公开的实施方案。本文使用的术语仅用于描述特定实施方案的目的,并不旨在限制本发明的范围,本发明的范围仅由权利要求限定。

[0034] 本公开的实施方案的描述并非旨在穷举或将本公开限制于所公开的精确形式。虽然出于说明性目的在本文中描述了本公开的具体实施例和示例,但是如相关领域的技术人员将认识到的,在本公开的范围内的各种等同修改是可能的。例如,虽然以给定顺序呈现方法步骤或功能,但是替代实施方案可以以不同的顺序执行功能,或者可以基本上同时执行功能。本文提供的本公开的教导可以适当地应用于其他程序或方法。可以组合本文所述的各种实施方案以提供另外的实施方案。如果需要,可以修改本公开的各个方面,以采用上述参考文献和申请的组合物、功能和概念来提供本公开的其他实施方案。此外,由于生物功能等效性考虑,可以在蛋白质结构中进行一些改变而不影响种类或数量的生物或化学作用。根据详细描述,可以对本公开进行这些和其他改变。所有这些修改旨在包括在所附权利要求的范围内。

[0035] 数值范围包括限定该范围的数字。除非另有说明,否则任何核酸序列均以5'至3'方向从左到右书写;氨基酸序列分别以氨基至羧基取向从左至右书写。

[0036] 本文提供的标题不是对本公开的各个方面或实施方案的限制。

[0037] 如本文所用,术语"能够"与动词一起使用时涵盖或意指对应动词的动作。例如,"能够相互作用"也意指相互作用,"能够切割"也意指切割,"能够结合"也意指结合,并且"能够特异性靶向…"也意指特异性靶向。

[0038] 术语的其他定义可以出现在整个说明书中。在更详细地描述示例性实施方案前,应当理解,本公开不限于所描述的特定实施方案,并且因此可以变化。还应理解,本文使用的术语仅用于描述特定实施方案的目的,并不旨在限制,因为本公开的范围将仅由所附权利要求限定。

[0039] 在提供数值范围的情况下,应当理解,除非上下文另有明确规定,否则还具体公开了该范围的上限和下限之间的每个中间值,至下限单位的十分之一。在所述范围内的任何所述值或中间值与所述范围内的任何其他所述值或中间值之间的每个较小范围都包含在本公开内。这些较小范围的上限和下限可以独立地包括在该范围内或排除在该范围之外,并且其中任一限值、没有限值或两个限值都包括在较小范围内的每个范围也包括在本公开内,受制于所述范围内的任何具体排除的限值。在所述范围包括限值中的一个或两个的情况下,排除那些所包括的限值中的任一个或两个的范围也包括在本公开中。

[0040] 如本文所用,冠词“一个(a)”和“一个(an)”可以指一个或多于一个(例如,至少一个)该冠词的语法对象。此外,除非上下文另有要求,否则单数术语应包括复数,并且复数术语应包括单数。在本申请中,除非另有说明,否则“或”的使用意指“和/或”。此外,术语“包括(including)”以及其他形式,如“包括(includes)”和“包括(included)”的使用不是限制性的。

[0041] “约”通常可以意指在给定测量的性质或精度的情况下测量的量的可接受的误差程度。示例性误差程度在给定值或值范围的百分之20(%)内,通常在10%内,更通常在5%内。优选地,术语“约”在本文中应理解为加或减(±)5%,优选±4%、±3%、±2%、±1%、±0.5%、±0.1%的与其一起使用的数字的数值。

[0042] 如本文所用,术语“包含(comprising)”或“包含(comprises)”用于指组合物、方法及其相应组分,其对于方法或组合物是必需的,但对于包括未指定的要素(无论是否必需)是开放的。

[0043] 术语“由……组成”是指如本文所述的组合物、方法及其各自的组分,其不包括该发明内容中未列举的任何要素。

[0044] 如本文所用,术语“基本上由……组成”是指给定发明所需的那些要素。该术语允许存在不会实质上影响本发明的基本和新的或功能性特征的要素。

[0045] 本文描述为“包含”一个或多个特征的实施方案也可以被认为是“由这些特征组成”和/或“基本上由这些特征组成”的相应实施方案的公开内容。

[0046] 如本文所用,术语“药学上可接受的”意指由联邦或州政府的管理机构批准,或在美国药典、欧洲药典或其他公认的药典中列出,用于动物,更特别是人。

[0047] 浓度、量、体积、百分比和其他数值在本文中可以以范围形式呈现。还应当理解,这样的范围格式仅为了方便和简洁而使用,并且应当灵活地解释为不仅包括明确列举为范围界限的数值,而且还包括该范围内涵盖的所有单独的数值或子范围,如同明确列举了每个数值和子范围一样。

[0048] 关于蛋白质使用时,术语“变体”意指蛋白质的肽或肽片段,其含有氨基酸(例如非天然氨基酸)的一个或多个类似物或置换的键。

[0049] 关于蛋白质使用时,术语“衍生物”意指包含所讨论的蛋白质和另外的肽序列的蛋白质。另外的肽序列应优选不干扰原始蛋白质的基本折叠并因此不干扰其构象结构。两个或更多个肽(或片段或变体)可以连接在一起以形成衍生物。或者,肽(或片段或变体)可以与不相关的分子(例如第二种不相关的肽)连接。衍生物可以化学合成,但通常通过重组核酸方法制备。可以包括另外的组分,如脂质和/或多糖和/或多肽组分。

[0050] 如本文所用,术语“蛋白质”和“多肽”在本文中可互换使用,以表示通过相邻残基

的α-氨基和羧基之间的肽键彼此连接的一系列氨基酸残基。术语“蛋白质”和“多肽”是指氨基酸的聚合物,包括修饰的氨基酸(例如,磷酸化、糖化、糖基化等)和氨基酸类似物,无论其大小或功能如何。“蛋白质”和“多肽”通常用于指相对大的多肽,而术语“肽”通常用于指小的多肽,但本领域中这些术语的使用重叠。当提及基因产物及其片段时,术语“蛋白质”和“多肽”在本文中可互换使用。因此,示例性多肽或蛋白质包括基因产物、天然存在的蛋白质、同源物、直系同源物、旁系同源物、片段和前述的其他等同物、变体、片段和类似物。

[0051] 本发明的蛋白质可以包括其中来自一个物种的氨基酸残基在保守或非保守位置处被另一物种中的相应残基置换的变体。本文公开的蛋白质分子的变体可以在本发明中产生和使用。在将多元数据分析技术应用于结构/性质-活性关系方面,遵循计算化学的领导[参见例如Wold等,Multivariate Data Analysis in Chemistry.Chemometrics-Mathematics and Statistics in Chemistry(编辑:B.Kowalski);D.Reidel Publishing Company,Dordrecht,Holland,1984(ISBN90-277-1846-6],蛋白质的定量活性-性质关系可以使用众所周知的数学技术,如统计回归、模式识别和分类来推导[参见例如Norman等,Applied Regression Analysis.Wiley-Interscience;第3版(1998年4月)ISBN:0471170828;Kandel,Abraham等,Computer-Assisted Reasoning in Cluster Analysis.Prentice Hall PTR,(1995年5月11日),ISBN:0133418847;Krzanowski,Wojtek.Principles of Multivariate Analysis:AUser's Perspective(Oxford Statistical Science Series,No 22(论文)).Oxford University Press;(2000年12月),ISBN:0198507089;Witten,Ian H.等,Data Mining:Practical Machine Learning Tools and Techniques with Java Implementations.Morgan Kaufmann;(1999年10月11日),ISBN:1558605525;Denison David G.T.(编辑)等,Bayesian Methods for Nonlinear Classification and Regression(Wiley Series in Probability and Statistics).John Wiley&Sons;(2002年7月),ISBN:0471490369;Ghose,Arup K.等,Combinatorial Library Design and Evaluation Principles,Software,Tools,and Applications in Drug Discovery.ISBN:0-8247-0487-8]。蛋白质的性质可以从蛋白质序列、功能和三维结构的经验和理论模型(例如,可能的接触残基的分析或计算的物理化学性质)导出,并且这些性质可以单独或组合考虑。

[0052] 氨基酸在本文中使用氨基酸名称、三字母缩写或单字母缩写来提及。如本文所用,术语“蛋白质”包括蛋白质、多肽和肽。如本文所用,术语“氨基酸序列”与术语“多肽”和/或术语“蛋白质”同义。在一些情况下,术语“氨基酸序列”与术语“肽”同义。在一些情况下,术语“氨基酸序列”与术语“酶”同义。术语“蛋白质”和“多肽”在本文中可互换使用。在本公开和权利要求中,可以使用氨基酸残基的常规单字母和三字母代码。氨基酸的3字母代码符合IUPACIUB生物化学命名联合委员会(JCBN)定义。还应当理解,由于遗传密码的简并性,多肽可以由多于一个的核苷酸序列来编码。

[0053] 非保守位置处的氨基酸残基可以用保守或非保守性残基置换。特别地,考虑保守性氨基酸置换。“保守性氨基酸置换”是其中氨基酸残基被具有相似侧链的氨基酸残基替换的置换。本领域已经定义了具有相似侧链的氨基酸残基家族,包括碱性侧链(例如赖氨酸、精氨酸或组氨酸)、酸性侧链(例如天冬氨酸或谷氨酸)、不带电荷的极性侧链(例如甘氨酸、天冬酰胺、谷氨酰胺、丝氨酸、苏氨酸、酪氨酸或半胱氨酸)、非极性侧链(例如丙氨酸、缬氨

酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、甲硫氨酸或色氨酸)、β-支链侧链(例如苏氨酸、缬氨酸、异亮氨酸)和芳族侧链(例如酪氨酸、苯丙氨酸、色氨酸或组氨酸)。因此,如果多肽中的氨基酸被来自相同侧链家族的另一个氨基酸替换,则该氨基酸置换被认为是保守性的。在本发明的抗体中包含保守修饰的变体不排除其他形式的变体,例如多态性变体、种间同源物和等位基因。

[0054] "非保守性氨基酸置换"包括其中(i)具有正电性侧链的残基(例如,Arg、His或Lys)置换负电性残基(例如,Glu或Asp)或被其置换,(ii)亲水性残基(例如,Ser或Thr)置换疏水性残基(例如,Ala、Leu、Ile、Phe或Val)或被其置换,(iii)半胱氨酸或脯氨酸置换任何其他残基或被其置换,或(iv)具有庞大疏水性或芳香族侧链的残基(例如,Val、His、Ile或Trp)置换具有较小侧链(例如Ala或Ser)或无侧链(例如Gly)的残基或被其置换。

[0055] 本说明书中提及的SARS-CoV-2多核苷酸和/或蛋白包括其片段和变体。

[0056] 如本文所用,在SARS-CoV-2刺突蛋白的上下文中,术语"片段"是指蛋白质的一部分,其可以包含全长SARS-CoV-2刺突蛋白的一个或多个结构域或部分结构域。根据本发明的SARS-CoV-2刺突蛋白片段通常可以是如本文所述的免疫原性片段。SARS-CoV-2刺突蛋白的片段的长度通常大于200个氨基酸。本发明的SARS-CoV-2刺突蛋白片段的长度可包含至少200、至少300、至少400、至少500、至少600、至少700、至少800、至少900、至少1000、至少1100或更多个氨基酸残基或由其组成。本发明的片段通常与SARS-CoV-2刺突蛋白具有共同的抗原交叉反应性(且因此称为免疫原性片段)。SARS-CoV-2刺突蛋白片段可包含以下或由以下组成:(i)SARS-CoV-2刺突蛋白的受体结合结构域(RBD);(ii)SARS-CoV-2刺突蛋白的N末端结构域(NTD);(iii)SARS-CoV-2刺突蛋白的C末端结构域(CTD),例如CTD1和/或CTD2,这些CTD也称为亚结构域(SD),其中CTD1也称为SD1和CTD2也称为SD2;和/或(iv)融合肽(FP);和/或(v)FPPR结构域;或其任何组合。特别地,根据本发明的SARS-CoV-2刺突蛋白的片段可包含RBD结构域或由RBD结构域组成。作为非限制性实例,根据本发明的SARS-CoV-2刺突蛋白的片段可以由RBD结构域组成,或者可以包含与NTD结构域组合的RBD结构域。

[0057] 变体SARS-CoV-2刺突蛋白保留了天然刺突蛋白的一个或多个构象表位以及引发产生中和抗体和/或免疫保护应答的能力。本发明的变体SARS-CoV-2刺突蛋白多核苷酸编码此类刺突蛋白。举例来说,变体可以与参考序列(例如本发明的SARS-CoV-2多核苷酸和/或蛋白,特别是本说明书中提供的定义SARS-CoV-2多核苷酸和/或蛋白的任何SEQ ID NO)具有至少80%,优选至少90%,更优选至少95%,最优选至少97%或至少99%的氨基酸序列同源性。因此,变体可以包括多核苷酸(例如非天然核酸)的一种或多种类似物,或置换的连接。此外,举例来说,当用于SARS-CoV-2多核苷酸和/或蛋白时,术语片段是指具有参考SARS-CoV-2多核苷酸和/或蛋白的至少10个,优选至少15个,更优选至少20个核酸残基的多核苷酸。术语片段还涉及上述变体。因此,举例来说,本发明的SARS-CoV-2多核苷酸和/或蛋白的片段可以包含具有至少10、20或30个核酸的核酸序列,其中多核苷酸序列与参考SARS-CoV-2多核苷酸和/或蛋白序列的(连续的)相应核酸序列具有至少80%的序列同源性。片段和变体的这些定义也适用于本发明的其他多核苷酸。在肽序列的上下文中,术语片段是指具有参考蛋白的至少10个,优选至少15个,更优选至少20个氨基酸残基。术语片段还涉及上述变体。因此,举例来说,片段可以包含具有至少10、20或30个氨基酸的氨基酸序列,其中氨基酸序列与参考序列的(连续的)氨基酸的相应氨基酸序列具有至少80%的序列同源性。

[0058] 优选地,变体是保守性置换变体。如本文所提及的“变体”是与天然或参考多肽基本上同源的多肽,但其由于一个或多个缺失、插入或置换而具有与天然或参考多肽的氨基酸序列不同的氨基酸序列。编码多肽的DNA序列涵盖与天然或参考DNA序列相比时包含核苷酸的一个或多个添加、缺失或置换,但编码相对于参考蛋白保留相关生物活性(例如野生型参考蛋白的至少50%)的变体蛋白或其片段的序列。关于氨基酸序列,技术人员将认识到,改变编码序列中的单个氨基酸或小百分比(即5%或更少,例如4%或更少,或3%或更少,或1%或更少)的氨基酸的核酸、肽、多肽或蛋白质序列的单个置换、缺失或添加是“保守修饰的变体”,其中所述改变导致氨基酸被化学上相似的氨基酸置换。预期一些变化可以潜在地改善相关活性,使得变体(无论是保守的还是非保守的)具有超过野生型活性100%,例如110%、125%、150%、175%、200%、500%、1000%或更多的活性。

[0059] 如本文所述的多肽可包含至少一个肽键置换。可以替换单个肽键或多个肽键,例如2个键、3个键、4个键、5个键或6个或更多个键,或所有肽键。如本文所述的分离的肽可包含一种类型的肽键置换或多种类型的肽键置换,例如2种类型、3种类型、4种类型、5种类型或更多种类型的肽键置换。肽键置换的非限制性实例包括脲、硫脲、氨基甲酸酯、磺酰脲、三氟乙胺、邻-(氨基烷基)-苯乙酸、对-(氨基烷基)-苯乙酸、间-(氨基烷基)-苯乙酸、硫代酰胺、四唑、硼酸酯、烯属基团及其衍生物。

[0060] 如本文所述的多肽可包含在由活生物体产生的多肽和/或蛋白质中常见的天然存在的氨基酸,例如Ala(A)、Val(V)、Leu(L)、Ile(I)、Pro(P)、Phe(F)、Trp(W)、Met(M)、Gly(G)、Ser(S)、Thr(T)、Cys(C)、Tyr(Y)、Asn(N)、Gln(Q)、Asp(D)、Glu(E)、Lys(K)、Arg(R)和His(H)。如本文所述的多肽可包含替代氨基酸。替代氨基酸的非限制性实例包括D氨基酸、β-氨基酸、高半胱氨酸、磷酸丝氨酸、磷酸苏氨酸、磷酸酪氨酸、羟脯氨酸、γ-羧基谷氨酸;马尿酸、八氢吲哚-2-羧酸、他汀、1,2,3,4,-四氢异喹啉-3-羧酸、青霉胺(3-巯基-D-缬氨酸)、鸟氨酸、瓜氨酸、α-甲基-丙氨酸、对苯甲酰基苯丙氨酸、对氨基苯丙氨酸、对氟苯丙氨酸、苯基甘氨酸、炔丙基甘氨酸、肌氨酸和叔丁基甘氨酸)、二氨基丁酸、7-羟基-四氢异喹啉羧酸、萘基丙氨酸、联苯丙氨酸、环己基丙氨酸、氨基-异丁酸、正缬氨酸、正亮氨酸、叔亮氨酸、四氢异喹啉羧酸、哌啶酸、苯基甘氨酸、高苯丙氨酸、环己基甘氨酸、脱氢亮氨酸、2,2-二乙基甘氨酸、I-氨基-1-环戊烷羧酸、I-氨基-1-环己烷羧酸、氨基-苯甲酸、氨基-萘甲酸、γ-氨基丁酸、二氟苯丙氨酸、3-哌啶甲酸、α氨基丁酸、噻吩-丙氨酸、叔-丁基甘氨酸、三氟缬氨酸;六氟亮氨酸;氟化类似物;叠氮化物修饰的氨基酸;炔修饰的氨基酸;氰基修饰的氨基酸;及其衍生物。

[0061] 可以修饰多肽,例如通过向包含肽的一个或多个氨基酸添加部分。如本文所述的多肽可以包含一个或多个部分分子,例如每个肽1个或更多个部分分子,每个肽2个或更多个部分分子,每个肽5个或更多个部分分子,每个肽10个或更多个部分分子或每个肽更多个部分分子。如本文所述的多肽可包含一种或多种类型的修饰和/或部分,例如1种类型的修饰、2种类型的修饰、3种类型的修饰或更多种类型的修饰。修饰和/或部分的非限制性实例包括PEG化;糖基化;HES化;ELP化;脂质化;乙酰化;酰胺化;封端修饰;氰基;磷酸化;白蛋白和环化。

[0062] 原始氨基酸序列的改变可以通过本领域技术人员已知的许多技术中的任一种来完成。例如,可以通过合成在编码待改变的氨基酸的核苷酸序列中含有密码子改变的寡核

苷酸,在特定位置引入氨基酸置换,所述核苷酸序列侧翼为允许与原始序列的片段连接的限制性位点。连接后,所得重建序列编码具有所需氨基酸插入、置换或缺失的类似物。或者,可以采用寡核苷酸指导的位点特异性诱变程序来提供具有根据所需的置换、缺失或插入而改变的特定密码子的改变的核苷酸序列。进行这种改变的技术包括Walder等(Gene 42:133,1986);Bauer等(Gene37:73,1985);Craik(Biotechniques,1985年1月,12-19);Smith等(Genetic Engineering:Principles and Methods,Plenum Press,1981);以及美国专利No.4,518,584和4,737,462公开的那些,其以整体按引用并入本文中。如本文所述的多肽可以是化学合成的,并且突变可以作为化学合成过程的一部分并入。

[0063] 如本文所用,术语“多核苷酸”、“核酸”和“核酸序列”是指掺入核糖核酸、脱氧核糖核酸或其类似物单元的任何分子,优选聚合分子。核酸可以是单链或双链的。单链核酸可以是变性双链DNA的一条核酸链,或者,它可以是不源自任何双链DNA的单链核酸。在一个方面中,核酸可以是DNA。在另一个方面中,核酸可以是RNA。合适的核酸分子是DNA,包括基因组DNA或cDNA。其他合适的核酸分子是RNA,包括mRNA。

[0064] 典型的抗体包含至少两条“轻链”(LC)和两条“重链”(HC)。此类抗体的轻链和重链是由若干结构域组成的多肽。每条重链包含重链可变区(本文缩写为“V_H”)和重链恒定区(本文缩写为“C_H”)。重链恒定区包含重链恒定结构域CH1、CH2和CH3(抗体类别IgA、IgD和IgG)和任选的重链恒定结构域CH4(抗体类别IgE和IgM)。每条轻链包含轻链可变结构域(本文缩写为“V_L”)和轻链恒定结构域(本文缩写为“C_L”)。可变区VH和VL可以进一步细分为高变区,称为互补决定区(CDR),其间散布有更保守的区域,称为框架区(FR)。每个VH和VL由三个CDR和四个FR组成,从氨基末端到羧基末端按以下顺序排列:FR1、CDR1、FR2、CDRs2、FR3、CDR3、FR4。重链和轻链的“恒定结构域”不直接参与抗体与靶标的结合,但表现出各种效应子功能。抗体与其靶抗原或表位之间的结合由互补决定区(CDR)介导。CDR是高序列变异性的区域,位于抗体重链和轻链的可变区内,在那里它们形成抗原结合位点。CDR是抗原特异性的主要决定簇。通常,抗体重链和轻链各自包含非连续排列的三个CDR。抗体重链和轻链CDR3区在根据本发明的抗体的结合特异性/亲和力中起着特别重要的作用,因此提供了本发明的再一个方面。因此,如本文所用的术语“抗原结合片段”包括包含一个、两个或三个轻链CDR和/或一个、两个或三个重链CDR的抗原结合多肽的任何天然存在的或人工构建的构型,其中所述多肽能够结合抗原。

[0065] CDR的序列可以通过参考本领域已知的任何数字系统来鉴定,例如Kabat系统(Kabat,E.A.,等,Sequences of Proteins of Immunological Interest,第5版,Public Health Service,National Institutes of Health,Bethesda,MD(1991);Chothia系统(Chothia&,Lesk,“Canonical Structures for the Hypervariable Regions of Immunoglobulins,”J.Mol.Biol.196,901-917(1987));或IMGT系统(Lefranc等,“IMGT unique numbering for immunoglobulin and cell receptor variable domains and Ig superfamily V-like domains,”Dev.Comp.Immunol.27,55-77(2003))。

[0066] 对于本发明中讨论的重链恒定区氨基酸位置,编号是根据Edelman,G.M.等,Proc.Natl.Acad.Sci.USA63(1969)78-85中首次描述的EU索引。Edelman的EU编号也在Kabat等(1991)(同上)中给出。因此,在重链的上下文中的术语“如Kabat所示的EU索引”、“EU索引”、“Kabat的EU索引”或“EU编”是指基于Edelman等的人IgG1 EU抗体的残基编号系

统,如Kabat等(1991)所述的。用于轻链恒定区氨基酸序列的编号系统类似于Kabat等(同上)所述。因此,如本文所用,“根据Kabat编号”是指Kabat等(同上)中所述的Kabat编号系统。

[0067] 术语“减少”、“降低的”、“降低”或“抑制”在本文中均用于意指减少统计学上显著的量。术语“降低(reduce)”、“降低(reduction)”或“减少(decrease)”或“抑制(inhibit)”通常意指与参考水平(例如,不存在给定的治疗)相比减少至少10%,,并且可以包括例如减少至少约10%、至少约20%、至少约25%、至少约30%、至少约35%、至少约40%、至少约45%、至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约98%、至少约99%或更多。如本文所用,“降低”或“抑制”不涵盖与参考水平相比的完全抑制或降低。“完全抑制”是与参考水平相比100%抑制。减少可以优选低至在没有给定病症的个体的正常范围内接受的水平。

[0068] 术语“增加的”、“增加”、“增强”或“激活”在本文中均用于意指统计学上显著量的增加。术语“增加的”、“增加”、“增强”或“激活”可以意指与参考水平相比至少10%的增加,例如与参考水平相比至少约20%、或至少约30%、或至少约40%、或至少约50%、或至少约60%、或至少约70%、或至少约80%、或至少约90%或至多并包括100%增加或10-100%之间的任何增加,或与参考水平相比至少约2倍,或至少约3倍,或至少约4倍,或至少约5倍或至少约10倍增加,或2倍和10倍之间的任何增加或更大的增加。在标志物或症状的上下文中,“增加”是此类水平的统计学上显著的增加。

[0069] 如本文所用,“受试者”意指人或动物。通常,动物是脊椎动物,如灵长类动物、啮齿动物、家养动物或猎用动物。灵长类动物包括黑猩猩、食蟹猴、蜘蛛猴和猕猴,例如恒河猴。啮齿动物包括小鼠、大鼠、土拨鼠、雪貂、兔和仓鼠。家畜和猎用动物包括牛、马、猪、鹿、野牛、水牛、猫科动物物种(例如家猫)、犬科动物物种(例如狗、狐狸、狼)、鸟类物种(例如鸡、鸸鹋、鸵鸟)和鱼(例如鳟鱼、鲶鱼和鲑鱼)。优选地,受试者是哺乳动物,例如灵长类动物,例如人。术语“个体”、“患者”和“受试者”在本文中可互换使用。

[0070] 优选地,受试者是哺乳动物。哺乳动物可以是人、非人灵长类动物、小鼠、大鼠、狗、猫、马或牛,但不限于这些实例。优选地,受试者是人。受试者可以是雄性或雌性、成年或青少年。

[0071] 受试者可以是先前已被诊断患有或被鉴定为患有或具有需要治疗的病症或与这种病症相关的一种或多种并发症,并且任选地,已经经历了针对如本文所定义的病症或与所述病症相关的一种或多种并发症的治疗的受试者。或者,受试者也可以是先前未被诊断为患有如本文所定义的病症或与所述病症相关的一种或多种并发症的受试者。例如,受试者可以是表现出病症或与所述病症相关的一种或多种并发症的一种或多种风险因素的受试者或不表现出风险因素的受试者。

[0072] “需要治疗特定病症的受试者”可以是患有该病症、被诊断为患有该病症或处于发展该病症的风险中的受试者。

[0073] 本文提及的特定分子的水平包括分子的实际量,如分子的质量、摩尔量、浓度或摩尔浓度。例如,在本发明的上下文中,提及特定分子的水平可以指分子的浓度。

[0074] 可以在任何合适的生理区室中测定分子的水平。优选的生理区室包括血浆、血液和/或血清。分子的水平可以从来自患者的任何合适的样品确定,例如血浆样品、血液样品、

血清样品、组织样品、支气管-肺泡灌洗(BAL)样品和/或CSF样品。可以测试的样品的其他非限制性实例是组织或流体样品、尿液和活检样品。因此,作为非限制性实例,本发明可以参考患者血浆和/或BAL中的分子水平(例如浓度)。用本发明的结合成员预处理的分子/生物标志物的水平可以互换地称为"基线"。

[0075] 可以将用本发明的疫苗治疗后的分子水平与用疫苗治疗前的患者中的分子水平进行比较。分子的水平可以直接或间接测量,并且可以使用任何合适的技术确定。合适的标准技术是本领域已知的,例如蛋白质印迹和酶联免疫吸附测定(ELISA)。

[0076] 如本文所用,术语SARS-CoV-2和2019-nCoV可互换使用,是指导致疾病COVID-19的病毒病原体。提及SARS-CoV-2感染是指疾病COVID-19。术语COVID-19疫苗(或针对COVID-19的疫苗)也与术语SARS-CoV-2疫苗(或针对SARS-CoV-2的疫苗)同义。

[0077] 如本文所用,术语"疫苗"用于指诱导免疫应答的组合物。例如,组合物可以在其所施用的患者中诱导免疫应答。

[0078] 减毒活疫苗包含能够在宿主细胞中感染和复制但已经以某种方式修饰使得它们不引起疾病的完整病毒颗粒或病毒粒子。

[0079] 活的载体疫苗包含活病毒载体,其通常是非致病性病毒,其已被修饰以表达来自病毒的一种或多种抗原,将引发针对所述病毒的免疫应答。通常,一种或多种抗原是关键抗原,如果患者暴露于野生型病毒(即感染疾病)或用减毒活疫苗或灭活疫苗接种,则将产生针对该关键抗原的免疫应答。抗原可以是蛋白质抗原或其片段,或多糖抗原或其片段。抗原可以重组表达或作为缀合物或融合蛋白表达。

[0080] 灭活疫苗包含已被杀死或灭活(例如通过热或化学处理)的完整病毒颗粒或病毒粒子。灭活的病毒粒子不能在宿主细胞中感染或复制,并且不会引起疾病。

[0081] 亚单位疫苗包含一种或多种病毒组分,针对所述病毒组分将产生免疫应答。通常,一种或多种组分是关键抗原,如果患者暴露于野生型病毒(即感染疾病)或用减毒活疫苗或灭活疫苗接种,则将产生针对该关键抗原的免疫应答。组合物可以是蛋白质抗原或其片段,或多糖抗原或其片段。组分可以重组表达或作为缀合物或融合蛋白表达。

[0082] 提供本文所讨论的出版物仅仅是因为它们在本申请的提交日之前的公开内容。本文中的任何内容都不应被解释为承认这些出版物构成所附权利要求的现有技术。

[0083] 联合疫苗

[0084] 当试图产生联合疫苗组合物时,常见的并发症是称为组分抑制(也称为抗原组合物)的现象。组分抑制描述了其中两种或更多种疫苗或疫苗抗原(通常来自不同病原体)同时施用并且与单独施用疫苗或疫苗抗原时引发的免疫应答相比,由一种或多种疫苗或疫苗抗原引发的免疫应答受损的情况。免疫应答可以以几种方式受损。例如,与单独施用疫苗或疫苗抗原时引发的免疫应答相比,由一种或多种疫苗或疫苗抗原引发的免疫应答可以降低。与单独施用疫苗或疫苗抗原时的血清转化和/或血清阳性相比,血清转化和/或血清阳性也可以降低。已经观察到与针对细菌病原体的疫苗(例如,用于百日咳-白喉-破伤风(DTaP)的疫苗和乙型流感嗜血杆菌(Hib)疫苗)和针对病毒病原体的疫苗(例如,黄热病疫苗和麻疹-腮腺炎-风疹(MMR)疫苗)相关的组分抑制现象。当在相同组合物中施用疫苗抗原时,且甚至当同时施用现有的有效疫苗组合物时,也观察到组分抑制。组分抑制的风险意味着不可能预测联合疫苗是否将是临床有效的,或者甚至是否可以一起施用两种单独的疫苗

组合物。组分抑制的风险在免疫学领域中是通常理解的,并且计入到疫苗时间表指定需要考虑的因素中,并且组分抑制的评估是医疗监管机构的要求。

[0085] 本发明人首次证明了可以施用包含流感抗原和源自SARS-CoV-2的抗原的疫苗,并实现对抗流感和SARS-CoV-2的良好免疫原性,即在流感和SARS-CoV-2的情况下不会发生组分抑制。

[0086] 因此,本发明提供了一种流感-COVID-19联合疫苗(在本文中也可互换地称为联合流感-COVID-19疫苗),其包含:(a)流感血凝素(或其免疫原性片段;和(b)一种或多种源自SARS-CoV-2的抗原或其免疫原性片段;其中抗原能够引发对抗流感和COVID-19的免疫应答和保护作用(如本文所述)。通常,所述流感-COVID-19联合疫苗还包含流感神经氨酸酶(NA)或其免疫原性片段。

[0087] 如本文所述,本发明的流感-COVID-19联合疫苗对于以下方面与组分抑制不相关,或具有最小的组分抑制:(i)流感HA或其免疫原性片段;(ii)一种或多种源自SARS-CoV-2的抗原(例如SARS-CoV-2刺突蛋白)或其免疫原性片段;(iii)任选的流感NA或其免疫原性片段;或其任何组合。优选地,本发明的流感-COVID-19联合疫苗对于以下的每一个方面与组分抑制不相关,或具有最小的组分抑制:(i)流感HA或其免疫原性片段;(ii)一种或多种源自SARS-CoV-2的抗原(例如SARS-CoV-2刺突蛋白)或其免疫原性片段;和(iii)任选的流感NA或其免疫原性片段;以及。

[0088] 如本文所用,术语“与组分抑制不相关”是指对作为本发明的流感-COVID-19联合疫苗的一部分施用的(i)流感HA或其免疫原性片段;(ii)一种或多种源自SARS-CoV-2的抗原(例如SARS-CoV-2刺突蛋白)或其免疫原性片段;(iii)任选的流感NA或其免疫原性片段;或其任意组合的免疫应答引发与分开施用(i)流感HA或其免疫原性片段;(ii)源自SARS-CoV-2的抗原(例如SARS-CoV-2刺突蛋白)或其免疫原性片段;和/或(iii)任选的流感NA或其免疫原性片段时基本相同的免疫应答。

[0089] 如本文所用,术语“具有最小的组分抑制”意指对作为本发明的流感-COVID-19联合疫苗的一部分施用的(i)流感HA或其免疫原性片段;(ii)一种或多种源自SARS-CoV-2的抗原(例如SARS-CoV-2刺突蛋白)或其免疫原性片段;(iii)任选的流感NA或其免疫原性片段;或其任何组合的免疫应答引发至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更多的分开施用(i)流感HA或其免疫原性片段;(ii)一种或多种源自SARS-CoV-2的抗原(例如SARS-CoV-2刺突蛋白)或其免疫原性片段;和/或(iii)任选的流感NA或其免疫原性片段时获得的免疫应答。

[0090] 本发明的流感-COVID-19联合疫苗的另一个优点是可以提高患者的依从性。本发明的流感-COVID-19联合疫苗可使患者接受单次疫苗施用,从而提供对流感和SARS-CoV-2两种感染的免疫力。减少所需的疫苗接种次数和所需的门诊就诊次数将增加疫苗摄取和患者依从性。此外,许多国家都有完善的公共卫生程序和年度流感疫苗接种计划时间表。本发明的流感-COVID-19联合疫苗可以利用这些现有的计划和程序进行对抗SARS-CoV-2感染的协调大规模疫苗接种,这也将促进针对SARS-CoV-2感染的大规模疫苗接种,而无需新的公共卫生计划或基础设施。此外,一些证据表明气候和季节性与COVID-19感染和传播有潜在的关联。因此,如本文所述,本发明有可能进行针对COVID-19的定期(例如季节性或年度)疫苗接种,从而减轻季节性感染和传播。此外,利用现有的公共卫生计划和程序,特别是那些

已经用于季节性流感疫苗接种的公共卫生计划和程序,促进COVID-19疫苗接种,可以实现这一点。

[0091] 流感HA或其免疫原性片段和任选的流感NA或其免疫原性片段可以各自由技术人员使用常规技术容易地选择。本文描述了流感HA(或其免疫原性片段)和流感NA(或其免疫原性片段)的非限制性实例。

[0092] 一种或多种SARS-CoV-2抗原或其免疫原性片段可以由技术人员使用常规技术容易地选择。本文描述了SARS-CoV-2抗原(或其免疫原性片段)的非限制性实例。通常,所述一种或多种SARS-CoV-2抗原包含至少一种SARS-CoV-2抗原刺突蛋白或其免疫原性片段,如本文所述的。

[0093] 流感HA或其免疫原性片段和/或任选的流感NA或其免疫原性片段可以包含在现有的流感疫苗组合物中。所述流感疫苗组合物可以与一种或多种SARS-CoV-2抗原(例如至少一种SARS-CoV-2刺突蛋白)或其免疫原性片段、或现有的COVID-19疫苗组合,以产生根据本发明的流感-CoVID-19联合疫苗。

[0094] 源自SARS-CoV-2的一种或多种抗原(例如,至少一种SARS-CoV-2刺突蛋白)或其免疫原性片段可以包含在现有的COVID-19疫苗组合物中。所述COVID-19疫苗组合物可以与流感HA或其免疫原性片段和/或任选的流感NA或其免疫原性片段组合,或现有的流感疫苗以产生本发明的流感-COVID-19联合疫苗。通常当使用活(减毒或载体)COVID-19疫苗时,使用活(减毒或载体)流感疫苗。通常当使用灭活或亚单位COVID-19疫苗时,使用灭活或亚单位流感疫苗。优选使用COVID-19亚单位(包括本文所述的融合蛋白和VLP)疫苗或组分和使用灭活流感疫苗。

[0095] 因此,包含在本发明的流感-COVID-19联合疫苗中的流感HA或其免疫原性片段可以是:(i)包含在灭活的流感病毒粒子中;(ii)重组HA或其免疫原性片段;(iii)包含HA或其免疫原性片段的融合蛋白;或(iv)由RNA或DNA疫苗编码。本文描述了流感HA、其免疫原性片段和包含HA的流感疫苗的非限制性实例。

[0096] 包含在本发明的流感-COVID-19联合疫苗中的(任选的)流感NA或其免疫原性片段可以是:(i)包含在灭活的流感病毒粒子中;(ii)重组NA或其免疫原性片段;(iii)包含NA或其免疫原性片段的融合蛋白;或(iv)由RNA或DNA疫苗编码。本文描述了流感NA、其免疫原性片段和包含NA的流感疫苗的非限制性实例。

[0097] 包含在本发明的流感-COVID-19联合疫苗中的一种或多种源自SARS-CoV-2的抗原或其免疫原性片段优选是:(i)至少一种重组SARS-CoV-2刺突蛋白或其免疫原性片段;(ii)至少一种融合蛋白,其包含SARS-CoV-2刺突蛋白或其免疫原性片段;(iii)至少一种病毒样颗粒(VLP),其包含SARS-CoV-2刺突蛋白或其免疫原性片段;(iv)至少一种编码重组SARS-CoV-2刺突蛋白或其免疫原性片段的多核苷酸;或(v)由RNA或DNA疫苗编码。本文描述了此类SARS-CoV-2抗原(特别是SARS-CoV-2刺突蛋白)及其免疫原性片段和COVID-19疫苗的非限制性实例。

[0098] (i)流感HA、其免疫原性片段和包含HA的流感疫苗;(ii)一种或多种SARS-CoV-2抗原,特别是SARS-CoV-2刺突蛋白及其免疫原性片段,以及COVID-19疫苗;和任选(iii)流感NA、其免疫原性片段和包含NA的流感疫苗的任何组合可以用于本发明的流感-COVID-19联合疫苗中,条件是HA、(任选的)NA和SARS-CoV-2抗原能够引发对抗流感和COVID-19的免疫

应答和保护。

[0099] 本发明的流感-COVID-19联合疫苗的流感成分可以包括活(减毒或载体)流感疫苗、灭活流感疫苗或亚单位流感疫苗。

[0100] 活的减毒流感疫苗的非限制性实例包括:季节性流感疫苗,如季节性四价(4价)流感疫苗。作为具体的非限制性实例,季节性四价流感疫苗(例如2019-2020季)可以包含减毒的甲型流感H1N1病毒、减毒的甲型流感H3N2病毒和两种乙型流感病毒(B/Colorado/06/2017样(Victoria谱系)病毒和B/Phuket/3073/2013样病毒(Yamagata谱系))。

[0101] 灭活流感疫苗的非限制性实例包括:季节性流感疫苗,如季节性三价(3价)流感疫苗和季节性四价(4价)流感疫苗。作为具体的非限制性实例,季节性三价流感疫苗(例如2019-2020季)可以包含减毒的甲型流感H1N1病毒、减毒的甲型流感H3N2病毒和乙型流感病毒(B/Colorado/06/2017样(Victoria谱系))。作为另一个具体的非限制性实例,季节性四价流感疫苗(例如2019-2020季)可以包含减毒的甲型流感H1N1病毒、减毒的甲型流感H3N2病毒和两种乙型流感病毒(B/Colorado/06/2017样(Victoria谱系)病毒和B/Phuket/3073/2013样病毒(Yamagata谱系))。

[0102] 可用于本发明的流感-COVID-19联合疫苗的其他流感疫苗实例包括单价大流行性流感疫苗(EMA预先批准的当前大流行性流感疫苗包括减毒活疫苗或灭活疫苗)和通用流感疫苗(正在开发的实例包括亚单位疫苗和包含引发DNA疫苗和活载体疫苗的两阶段疫苗)。

[0103] 优选,本发明的流感-COVID-19联合疫苗的流感成分是活的减毒流感疫苗或灭活流感疫苗。

[0104] 本发明的流感-COVID-19联合疫苗的SARS-CoV-2组分可以包含活(减毒或载体)SARS-CoV-2/COVID-19疫苗、灭活SARS-CoV-2/COVID-19疫苗或SARS-CoV-2/COVID-19亚单位疫苗。

[0105] 优选,本发明的流感-COVID-19联合疫苗的SARS-CoV-2成分是亚单位疫苗,其包含SARS-CoV-2刺突蛋白或其片段,或包含所述SARS-CoV-2刺突蛋白或其片段的融合蛋白或VLP。

[0106] 特别优选的是流感-COVID-19联合疫苗,其中流感组分是活的减毒流感疫苗或灭活流感疫苗,而SARS-CoV-2组分是包含SARS-CoV-2刺突蛋白或其片段的亚单位疫苗,或包含所述SARS-CoV-2刺突蛋白或其片段的融合蛋白或VLP。

[0107] 通常,本发明的流感-COVID-19联合疫苗的流感成分包含活(减毒或载体)流感疫苗时,SARS-CoV-2组分包含活(减毒或载体)SARS-CoV-2/COVID-19疫苗。

[0108] 通常,本发明的流感-COVID-19联合疫苗的流感组分包含灭活流感疫苗时,SARS-CoV-2成分包含灭活SARS-CoV-2/COVID-19疫苗。或者,本发明的流感-COVID-19联合疫苗的流感组分包含灭活流感疫苗时,SARS-CoV-2组分包含亚单位SARS-CoV-2/COVID-19疫苗,反之亦然。

[0109] 通常,本发明的流感-COVID-19联合疫苗的流感组分包含亚单位流感疫苗时,SARS-CoV-2组分包含亚单位SARS-CoV-2/COVID-19疫苗。或者,本发明的流感-COVID-19联合疫苗的流感组分包含亚单位流感疫苗时,SARS-CoV-2组分包含灭活的SARS-CoV-2/COVID-19疫苗,反之亦然。

[0110] 通常,本发明的流感-COVID-19联合疫苗的流感成分包含核酸(DNA或RNA,优选

DNA)流感疫苗时,SARS-CoV-2组分包含核酸(DNA或RNA,优选DNA)SARS-CoV-2/COVID-19疫苗。

[0111] 本发明提供了流感-CoVID-19联合疫苗,其中流感HA或其免疫原性片段和流感NA或其免疫原性片段包含在灭活流感病毒粒子中,并且一种或多种源自SARS-CoV-2的抗原(例如至少一种SARS-CoV-2刺突蛋白)或其免疫原性片段是:(i)至少一种融合蛋白,其包含SARS-CoV-2刺突蛋白或其免疫原性片段;(ii)至少一个病毒样颗粒(VLP),其包含SARS-CoV-2刺突蛋白或其免疫原性片段;或灭活的SARS-CoV-2病毒粒子。

[0112] 本发明提供了流感-CoVID-19联合疫苗,其中流感HA或其免疫原性片段和任选的流感NA或其免疫原性片段包含在亚单位疫苗中,并且一种或多种源自SARS-CoV-2的抗原(例如至少一种SARS-CoV-2刺突蛋白)或其免疫原性片段是:(i)至少一种融合蛋白,其包含SARS-CoV-2刺突蛋白或其免疫原性片段;(ii)至少一个病毒样颗粒(VLP),其包含SARS-CoV-2刺突蛋白或其免疫原性片段;或灭活的SARS-CoV-2病毒粒子。

[0113] 本发明提供了流感-COVID-19联合疫苗,其中:流感HA或其免疫原性片段包含在减毒活流感病毒粒子中;流感NA或其免疫原性片段包含在减毒活流感病毒粒子中;和/或一种或多种源自SARS-CoV-2的抗原或其免疫原性片段包含在活病毒载体中(即在活载体疫苗中)。包含一种或多种源自SARS-CoV-2的抗原或其免疫原性片段的活病毒载体可以是临床上用于疫苗的任何病毒载体。非限制性实例包括腺病毒载体、麻疹病毒载体、腮腺炎病毒载体、风疹病毒载体、水痘病毒载体、脊髓灰质炎病毒载体和黄热病病毒载体。

[0114] 冠状病毒抗原

[0115] 冠状病毒(CoV)在所有RNA病毒中具有最大的基因组,通常范围为27至32kb。CoV基因组编码至少四种主要结构蛋白:刺突(S)、膜(M)、包膜(E)、核衣壳(N)蛋白和其他辅助蛋白,其辅助复制过程并促进进入细胞。图1总结了冠状病毒的结构和结构蛋白的功能。简言之,CoV基因组包装在由核衣壳形成的螺旋衣壳内部,并进一步被包膜包围。与病毒包膜相关的是至少三种结构蛋白:参与病毒装配的膜和包膜蛋白,以及介导病毒进入宿主细胞中的刺突蛋白。一些冠状病毒还编码包膜相关的血凝素-酯酶蛋白(HE)。刺突蛋白从病毒表面形成大的突起,使冠状病毒具有冠状的外观,由此得名“冠状病毒”。除了介导病毒进入之外,刺突蛋白是病毒宿主范围和组织嗜性的关键决定因素,并且是宿主免疫应答的主要诱导物。

[0116] 2019-nCoV(正式命名为严重急性呼吸综合征冠状病毒2,SARS-CoV-2)是2019冠状病毒病(COVID-19)的病原体,在人类中具有传染性。认为SARS-CoV-2起源于动物,鉴于SARS-CoV-2与SARS-CoV(79.5%)和蝙蝠冠状病毒(96%)的遗传相似性,蝙蝠是可能的来源。本文关于CoV的任何公开内容也直接适用于且不限于SARS-CoVs-2。

[0117] 本发明的流感-COVID-19联合疫苗中的一种或多种源自SARS-CoV-2的抗原或其免疫原性片段可以是能够引发对抗SARS-CoV-2感染的免疫应答和/或保护的任何SARS-CoV-2抗原。优选,所述一种或多种抗原是:(i)至少一种重组SARS-CoV-2刺突蛋白或其免疫原性片段;(ii)至少一种包含SARS-CoV-2刺突蛋白的融合蛋白或其免疫原性片段;(iii)至少一种包含SARS-CoV-2刺突蛋白或其免疫原性片段的病毒样颗粒(VLP);(iv)至少一种编码重组SARS-CoV-2刺突蛋白或其免疫原性片段的多核苷酸;或(v)由至少一种RNA或DNA疫苗编码。

[0118] 本发明的流感-COVID-19联合疫苗的SARS-CoV-2组分可以包含至少一种、至少两种、至少三种、至少四种或更多种SARS-CoV-2抗原。作为非限制性实例,每种SARS-CoV-2抗原可以是不同的刺突蛋白抗原,如野生型刺突蛋白抗原和/或本文所述的变体刺突蛋白之一。可以包括在本发明的流感-COVID-19联合疫苗中的SARS-CoV-2抗原的其他非限制性实例包括来自2019-CoV衣壳、膜蛋白或包膜蛋白的抗原。一种或多种SARS-CoV-2抗原中的每一种可以以以下形式独立地提供:(i)重组抗原或其免疫原性片段;(ii)融合蛋白或其免疫原性片段;(iii)包含抗原或其免疫原性片段的病毒样颗粒(VLP);或(iv)编码抗原或其免疫原性片段的多核苷酸。本文关于包含SARS-CoV-2刺突蛋白抗原的重组、融合蛋白、VLP、多核苷酸和载体的公开内容同样适用于可包含在本发明的流感-COVID-19联合疫苗中的其他SARS-CoV-2抗原。

[0119] 刺突蛋白

[0120] CoV刺突蛋白包含三个结构域:(i)大的胞外域;(ii)跨膜结构域(其单程通过病毒包膜);和(iii)短的胞内尾部。胞外域由三个受体结合亚基(3×S1)和由三个膜融合亚基(3×S2)构成的三聚体茎组成。因此,SARS-CoV-2刺突蛋白是同型三聚体。在病毒进入期间,S1与宿主细胞表面上的受体结合用于病毒附着,而S2融合宿主和病毒膜,允许病毒基因组进入宿主细胞。受体结合和膜融合是冠状病毒感染周期的最初和关键步骤。不同CoV所靶向的受体存在显著差异。

[0121] SARS-CoV-2刺突蛋白的结构描述于例如Cai等(Science(2020)369:1586-1592)),其通过引用整体并入本文中。SARS-CoV-2刺突蛋白的每个S1亚单位包含N-末端结构域(NTD))、受体结合结构域(RBD)、两个C-末端结构域(CTD)。在与宿主细胞膜融合前,SARS-CoV-2刺突蛋白的S1亚单位保护S2亚单位。与ACE2结合时,SARS-CoV-2刺突蛋白以“折刀”方式重折叠,形成长中心卷曲螺旋并最终导致膜融合和病毒进入宿主细胞。

[0122] 本发明人先前已经表明了SARS-CoV-2刺突蛋白及其免疫原性片段作为对抗SARS-CoV-2/COVID-19感染的疫苗的抗原具有治疗潜力(包括预防潜力)。

[0123] 因此,如本文所述,本发明的流感-COVID-19联合疫苗中所含的一种或多种源自SARS-CoV-2的抗原优选是一种或多种SARS-CoV-2刺突蛋白或其免疫原性片段。通常,所述一种或多种SARS-CoV-2刺突蛋白与SEQ ID NO:1或其与所述刺突蛋白具有共同抗原交叉反应性的片段具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高的同一性。优选,来自SARS-CoV-2的一种或多种刺突蛋白与SEQ ID NO:1或其与所述刺突蛋白具有共同抗原交叉反应性的片段具有至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高的同一性。更优选,来自SARS-CoV-2的一种或多种刺突蛋白或其片段与SEQ ID NO:1或其与所述刺突蛋白具有共同抗原交叉反应性的片段具有至少98%、至少99%或更多的同一性。一种或多种来自SARS-CoV-2的刺突蛋白可包含SEQ ID NO:1或其与所述刺突蛋白具有共同抗原交叉反应性的片段(在本文中也称为免疫原性片段)或由其组成。

[0124] 根据本发明的SARS-CoV-2刺突蛋白或其免疫原性片段通常保留与天然SARS-CoV-2刺突蛋白相同的对其受体的结合亲和力。在本发明的上下文中,这可能意味着对SARS-CoV-2刺突蛋白受体的结合亲和力为天然SARS-CoV-2刺突蛋白的结合亲和力的至少80%、至少85%、至少90%、至少95%、至少99%或更高。优选,本发明的SARS-CoV-2刺突蛋白或其

免疫原性片段对SARS-CoV-2刺突蛋白的结合亲和力为天然SARS-CoV-2刺突蛋白的结合亲和力的至少90％、至少95％、至少99％或更高。

[0125] 在一些实施方案中,本发明的SARS-CoV-2刺突蛋白或其免疫原性片段对2019-nCoV刺突蛋白受体的结合亲和力大于全长蛋白的结合亲和力。例如,本发明的SARS-CoV-2刺突蛋白或其免疫原性片段可以具有天然SARS-CoV-2刺突蛋白的结合亲和力的至少100％、至少110％、至少120％、或至少150％或更高的结合亲和力。

[0126] 在其他实施方案中,本发明的SARS-CoV-2刺突蛋白或其免疫原性片段对SARS-CoV-2刺突蛋白受体的结合亲和力可以小于天然SARS-CoV-2刺突蛋白的结合亲和力。例如,本发明的SARS-CoV-2刺突蛋白或其免疫原性片段可以具有小于80％、小于70％、小于60％、小于50％或更小的天然SARS-CoV-2刺突蛋白结合亲和力的结合亲和力。

[0127] 由本发明的多核苷酸表达的SARS-CoV-2刺突蛋白或其免疫原性片段对其受体的结合亲和力可以根据解离常数(K_d)来定量。K_d可以使用任何合适的技术来确定,但在本发明的上下文中通常优选表面等离子体共振(SPR)。

[0128] 一种或多种SARS-CoV-2刺突蛋白的免疫原性片段的长度通常大于200个氨基酸。本发明的SARS-CoV-2刺突蛋白片段的长度可包含至少200、至少300、至少400、至少500、至少600、至少700、至少800、至少900、至少1000、至少1100或更多个氨基酸残基或由其组成。本发明的片段与SARS-CoV-2刺突蛋白具有共同的抗原交叉反应性(且因此称为免疫原性片段)。

[0129] 根据本发明,一种或多种SARS-CoV-2刺突蛋白或其片段维持天然(野生型)SARS-CoV-2刺突蛋白中存在的一个或多个构象表位。因此,一种或多种SARS-CoV-2刺突蛋白或其片段能够产生免疫保护作用。通常,所述免疫保护作用包括产生中和抗体(nAb),其特异性结合SARS-CoV-2刺突蛋白或其片段的一个或多个构象表位。CoV刺突蛋白的构象表位具有在CoV刺突蛋白的三级结构中发现的特定三维结构。所述一个或多个构象表位通常在刺突蛋白的胞外域内。优选,一种或多种SARS-CoV-2刺突蛋白或其片段保留天然SARS-CoV-2刺突蛋白中存在的所有构象表位。

[0130] SARS-CoV-2蛋白的免疫原性片段可包含RBD、NTD、CTD1、CDT2、FP和/或FPPR或其任何组合或由其组成。优选,SARS-CoV-2刺突蛋白的免疫原性片段包含SARS-CoV-2刺突蛋白的受体结合结构域(RBD)或由其组成。该RBD负责SARS-CoV-2与宿主细胞的结合,从而促进SARS-CoV-2颗粒进入宿主细胞中。RBD对应于SEQ ID NO:1的氨基酸残基319至529,如本文所述,称为SEQ ID NO:15。RBD由对应于SARS-CoV-2病毒(GenBank登录号No.MN908947,其版本3(MN908947.3)于2020年1月17日保藏)基因组中955至1597位的碱基编码。因此,如本文所述,本发明涉及SARS-CoV-2刺突蛋白的RBD,其与SEQ ID NO:15具有至少70％、至少75％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或更高的同一性。优选,SARS-CoV-2刺突蛋白的免疫原性片段包含SARS-CoV-2刺突蛋白的RBD或由其组成,所述RBD与SEQ ID NO:15具有至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或更高的同一性。更优选,SARS-CoV-2刺突蛋白的免疫原性片段包含SARS-CoV-2刺突蛋白的RBD或由其组成,所述RBD与SEQ ID NO:15具有至少98％、至少99％或更高的同一性。SARS-CoV-2刺突蛋白的RBD可包含SEQ ID NO:15或由其组成。本文关于SARS-CoV-2刺突蛋白的任何和所有公开内容(例如,关于多核苷酸、病毒载体、DNA质粒、RNA疫苗、病毒样颗粒

(VLP)、融合蛋白、抗体、组合物和药物组合物、制剂和治疗适应症)同样适用于且不限于SARS-CoV-2刺突蛋白的RBD。

[0131] CoV是大的有包膜的单正义RNA病毒。RNA病毒的突变率高于DNA病毒,表明生存的适应过程更有效。因此,与流感病毒观察到的类似,存在如下风险:抗原漂移也将成为SARS-CoV-2的特征,或者一旦大流行消退,SARS-CoV-2在人群中就成为地方性流行。实际上,迄今为止的研究已经鉴定了2019-nCoV的刺突蛋白的受体结合结构域(RBD)内的突变,特别是G476S和V483A/G,以及S1/S2位点附近的普遍D614G突变(Saha等,ChemRXIVhttp://doi.org/10.26434/chemrxiv.12320567.v1),证据表明这可以增强2019-nCoV病毒粒子的细胞进入,并且还拓宽宿主细胞向性。在2019-nCoV刺突蛋白中报道的其他突变包括S943(特别是S943P)、L5(特别是L5F)、L8(特别是L8F)、V367(特别是V367F)、H49(特别是H49Y)、Y145(特别是Y145H/del)、Q239(特别是Q239K)、A831(特别是A831V)、D839(特别是D839Y/N/E)和P1263(特别是P1263L),或其任何组合(Korber等,BioRxiv™ https://doi.org/10.1101/2020.04.29.069054)。

[0132] 因此,如果需要,本发明有利地允许修饰2019-nCoV疫苗抗原,以在它们产生时提供增强的对抗具有突变的刺突蛋白的毒株的免疫力。作为非限制性实例,根据本发明的任何2019-nCoV刺突蛋白或其片段可以在位置(i)D614、(ii)V483、(iii)G476、(iv)K417、(v)E484、(vi)N501、(vii)A570和(viii)P681或(i)至(viii)中的任何组合(包括任何两个、任何三个、任何四个、任何五个、任何六个、任何七个或全部八个)处进行修饰(特别是通过置换)。替换地或另外地,2019-nCoV刺突蛋白或其片段可包含缺失突变,包括在氨基酸残基69、70和/或144中的一个或多个处的缺失。如本文所述的,突变/修饰的位置通常对应于本发明的SEQ ID NO:1中的氨基酸编号。

[0133] 在位置D614处的修饰,特别是D614G置换是优选的。特别地,根据本发明的任何2019-nCoV刺突蛋白或其片段可包含以下置换(i)G476S、(ii)V483A/G、(iii)D614G、(iv)K417N/T、(v)E484K、(vi)N501Y、(vii)A570D和(viii)P681H,或(i)至(viii)的任何组合(包括任何两个、任何三个、任何四个、任何五个、任何六个、任何七个或全部八个)。

[0134] 本发明还涉及来自变体SARS-CoV-2的SARS-CoV-2刺突蛋白或其片段。特别地,本发明可以涉及来自B.1.1.7毒株(也称为201/501Y.V1,在英国首次检测到,现在称为α变体);B.1.351毒株(也称为20H/501.V2,首先在南非检测到,现在称为β变体)、P1毒株(也称为20J/501Y.v3,首先在日本和巴西检测到,现在称为γ变体)、B1.427和B1.429毒株(首先在加利福尼亚检测到,现在称为ε变体)和/或B.1.617.2毒株(首先在印度检测到,现在称为δ变体)的SARS-CoV-2刺突蛋白或其片段。根据CDC(SARS-CoV-2 Variant Classifications and Definitions(cdc.gov)),已经发现α变体包含以下突变:69缺失、70缺失、144缺失、(E484K*)、(S494P*)、N501Y、A570D、D614G、P681H、T716I、S982A、D1118H、和(K1191N*)。α变体的关键突变包括残基69/70和144Y的缺失,以及N501Y、A570D、D614G和P681H置换。根据CDC(SARS-CoV-2 Variant Classifications and Definitions(cdc.gov)),已经发现β变体包含以下突变:D80A、D215G、241缺失、242缺失、243缺失、K417N、E484K、N501Y、D614G和A701V。β变体的关键突变包括K417N、E484K、N501Y和D614G置换。根据CDC(SARS-CoV-2 Variant Classifications and Definitions(cdc.gov)),已经发现γ变体包含以下突变:L18F、T20N、P26S、D138Y、R190S、K417T、E484K、N501Y、D614G、H655Y、T1027I。γ变体的关键

突变包括E484K、K417N/T、N501Y和D614G。根据CDC(SARS-CoV-2 Variant Classifications and Definitions(cdc.gov)),已经发现δ变体包含以下突变:T19R、(G142D*)、156缺失、157缺失、R158G、L452R、T478K、D614G、P681R和D950N。δ变体的关键突变包括L452R、E484Q和T478K。根据CDC CDC(SARS-CoV-2 Variant Classifications and Definitions(cdc.gov)),已经发现ε变体包含以下突变:S13I、W152C、L452R、D614G。ε变体的关键突变是L452R。

[0135] 除非明确说明,否则本文关于联合疫苗、多核苷酸、刺突蛋白及其片段、VLP、融合蛋白和DNA/RNA疫苗的所有公开内容同样适用于SARS-CoV-2的不同变体和毒株。

[0136] 因此,开发可以安全重复施用的疫苗组合物不仅能够加强免疫应答以解决保护性免疫随时间丧失的问题(如本文所述和临床观察到的),而且还将有利地允许SARS-CoV-2疫苗抗原在需要时进行修饰,以提供增强的对抗具有突变刺突蛋白的毒株的免疫力。作为非限制性实例,用作根据本发明的一种或多种SARS-CoV-2抗原的任何SARS-CoV-2刺突蛋白或其片段可以在以下位置进行修饰:(i)417;(ii)452;(iii)478;(iv)484;(v)201;(vi)570;(vii)614;和/或(viii)681;或其任何组合。作为进一步的非限制性实例,用作根据本发明的一种或多种SARS-CoV-2抗原的任何SARS-CoV-2刺突蛋白或其片段可以在位置(i)D614、(ii)V483、(iii)G476、(iv)G476和V483、(v)G476和D614、(vi)V483和D614,或(vii)G476、V483和D614进行修饰(特别是通过置换)。在位置D614处的修饰,特别是D614G置换可以是优选的。在位置L452处的修饰,特别是L452R置换可以是优选的。特别地,用作根据本发明的一种或多种SARS-CoV-2抗原的任何SARS-CoV-2刺突蛋白或其片段可以包含以下置换(i)G476S,(ii)V483A/G,(iii)D614G,(iv)G476S和V483A/G,(v)G476S和D614G,(vi)V483A/G和D614G,(vii)G476S、V483A/G和D614G,(viii)L452R和E484Q,以及任选的T478K;或(ix)L452R。多种变体SARS-CoV-2刺突蛋白(本文所述的任何形式,特别是作为融合蛋白或VLP)可包含在本发明的流感-COVID-19联合疫苗中。

[0137] 多核苷酸

[0138] 一种或多种源自SARS-CoV-2的抗原或其免疫原性片段可以由一种或多种多核苷酸疫苗编码或表达(术语“编码”和“表达”在本文中可互换使用)以产生抗原或其免疫原性片段。术语多核苷酸包括DNA和RNA序列。在本文中,术语“核酸”、“核酸分子”和“多核苷酸”可互换使用。因此,源自SARS-CoV-2的抗原(例如SARS-CoV-2刺突蛋白)或其免疫原性片段可以由DNA或RNA疫苗编码或表达。

[0139] 在本发明的流感-COVID-19联合疫苗中,表达一种或多种SARS-CoV-2刺突蛋白或其免疫原性片段的一种或多种多核苷酸可以表达与SEQ ID NO:1具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的来自SARS-CoV-2的刺突蛋白,或其与所述刺突蛋白具有共同抗原交叉反应性的片段。优选,所述一种或多种多核苷酸表达与SEQ ID NO:1具有至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的来自SARS-CoV-2的一种或多种刺突蛋白,或其与所述刺突蛋白具有共同抗原交叉反应性的片段。更优选,所述一种或多种多核苷酸表达与SEQ ID NO:1具有至少98%、至少99%或更高同一性的来自SARS-CoV-2的一种或多种刺突蛋白或其与所述刺突蛋白具有共同抗原交叉反应性的片段。所述一种或多种多核苷酸可以表达包含SEQ ID NO:1或由SEQ ID NO:1组成的来自SARS-CoV-2的刺突蛋白或其与所述刺

突蛋白具有共同抗原交叉反应性的片段。多种SARS-CoV-2抗原(特别是一种或多种SARS-CoV-2刺突蛋白)可以由一种多核苷酸或多种多核苷酸或其组合表达。作为非限制性实例,所述一种或多种SARS-CoV-2抗原(特别是一种或多种SARS-CoV-2刺突蛋白)可以由单个多核苷酸表达,或者每种所述SARS-CoV-2抗原(特别是一种或多种SARS-CoV-2刺突蛋白)可以由单独的多核苷酸表达。

[0140] 通常,所述多核苷酸包含分离的多核苷酸,其编码与SEQ ID NO:1具有至少90%同一性的来自SARS-CoV-2的刺突蛋白,或其与所述刺突蛋白具有共同抗原交叉反应性的片段,或如本文所述的其任何变体。例如,多核苷酸可以编码SARS-CoV-2刺突蛋白的RBD,优选其中所述RBD与SEQ ID NO:15具有至少90%的同一性。编码RBD的示例性多核苷酸显示于SEQ ID NO:13和SEQ IDNO:14的密码子优化序列中。因此,本发明的多核苷酸可以包含与SEQ ID NO:13具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的核酸序列或由其组成。优选,本发明的多核苷酸可以包含与SEQ ID NO:13具有至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的核酸序列或由其组成。更优选,本发明的多核苷酸可以包含与SEQ ID NO:13具有至少98%、至少99%或更高同一性的核酸序列或由其组成。本发明的多核苷酸可以包含SEQ ID NO:13的核酸序列或由其组成。

[0141] 本发明还包括编码来自SARS-CoV-2的变体刺突蛋白的多核苷酸,如上所述,或其与所述变体刺突蛋白具有共同抗原交叉反应性的片段。所述变体刺突蛋白通常与SEQ ID NO:1或其片段(如SEQ ID NO:15的RBD)具有至少90%同一性。

[0142] 可以针对在患者中表达来优化编码一种或多种SARS-CoV-2刺突蛋白或其免疫原性片段的一种或多种多核苷酸(例如DNA或RNA疫苗)。如本文所用的术语“优化的”涉及针对一种或多种SARS-CoV-2刺突蛋白或其免疫原性片段的表达的优化,并且包括密码子优化和/或对多核苷酸的其他修饰(就核酸序列和其他修饰而言),其增加患者体内从多核苷酸表达一种或多种SARS-CoV-2刺突蛋白的水平和/或持续时间,或当从DNA或RNA疫苗表达一种或多种SARS-CoV-2刺突蛋白或其片段时以其他方式提供优势。发明人先前在英国专利申请No.2002166.3中描述了编码SARS-CoV-2刺突蛋白和片段的优化多核苷酸,其通过引用整体并入本文中。

[0143] 因此,一种或多种源自SARS-CoV-2的抗原或其免疫原性片段,特别是一种或多种SARS-CoV-2刺突蛋白或其免疫原性片段,可以由一种或多种多核苷酸(例如,DNA或RNA疫苗)来编码,所述多核苷酸包含与SEQ ID NO:2、3、4、5、6、7、8、13、14、26、27、29、30或32中的任一个具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的核酸序列。优选,所述一种或多种多核苷酸包含与SEQ ID NO:2、3、4、5、6、7、8、13、14、26、27、29、30或32中的任一个具有至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的核酸序列。更优选,所述一种或多种多核苷酸包含与SEQ ID NO:2、3、4、5、6、7、8、13、14、26、27、29、30或32中的任一个具有至少98%、至少99%或更高同一性的核酸序列。所述一种或多种多核苷酸可包含SEQ ID NO:2、3、4、5、6、7、8、13、14、26、27、29、30或32中任一个的核酸序列。此外,SEQ ID NO:2、3、4、5、6、7、8、13、14、26、27、29、30或32中的任一个中鉴定的5’克隆位点、3’克隆位点或5’和3’克隆位点可以在多核苷酸(例如DNA或RNA疫苗)中缺失。因此,一种或多种多核苷酸(例如DNA或

RNA疫苗)可包含SEQ IDNO:2、3、4、5、6、7、8、13、14、26、27、29、30或32中的任一个,但缺乏SEQ ID NO:2、3、4、5、6、7、8、13、14、26、27、29、30或32中的任一个中鉴定的5'克隆位点、3'克隆位点或5'和3'克隆位点。或者,SEQ ID NO:2、3、4、5、6、7、8、13、14、26、27、29、30或32中的任一个或如本文所述的其任何变体中鉴定的5'克隆位点、3'克隆位点或5'和3'克隆位点可以被另一个合适的克隆位点独立地替代。合适的可替换的克隆位点是本领域熟知的。

[0144] 本发明特别涉及源自SARS-CoV-2的抗原或免疫原性片段,其包含SARS-CoV-2刺突蛋白的RBD或由SARS-CoV-2刺突蛋白的RBD组成。因此,本发明的多核苷酸可以包含与SEQ ID NO:13或与SEQ ID NO:14的密码子优化序列具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的核酸序列或由其组成。优选,本发明的多核苷酸可以包含与SEQ ID NO:13或与SEQ ID NO:14的密码子优化序列具有至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的核酸序列或由其组成。更优选,本发明的多核苷酸可以包含与SEQ ID NO:13或与SEQ ID NO:14的密码子优化序列具有至少98%、至少99%或更高同一性的核酸序列或由其组成。本发明的多核苷酸可包含SEQ ID NO:13的核酸序列或SEQ ID NO:14的密码子优化序列或由其组成。

[0145] 根据本发明的一种或多种多核苷酸(例如DNA或RNA疫苗)通常编码至少一种SARS-CoV-2刺突蛋白或其免疫原性片段,其:(a)保留天然SARS-CoV-2刺突蛋白中存在的构象表位;和/或(b)当将所述核酸施用于患者时,导致产生对刺突蛋白或其片段特异性的中和抗体。

[0146] 一种或多种多核苷酸(例如DNA或RNA疫苗)通常表达至少一种来自SARS-CoV-2的刺突蛋白或其免疫原性片段,特别是至少一种来自SARS-CoV-2的刺突蛋白或其免疫原性片段,如本文所述(包括VLP或融合蛋白的形式)。

[0147] 根据本发明的一种或多种多核苷酸(例如DNA或RNA疫苗)可以包含在表达构建体中,以促进一种或多种SARS-CoV-2刺突蛋白或其片段的表达。通常,在这样的表达构建体中,所述一个或多个多核苷酸与合适的启动子可操作地连接。一个或多个多核苷酸可以与合适的终止子序列连接。一个或多个多核苷酸可以与启动子和终止子两者连接。合适的启动子和终止子序列是本领域熟知的。

[0148] 一种或多种多核苷酸(例如DNA或RNA疫苗)可以编码至少一种SARS-CoV-2刺突蛋白或其免疫原性片段,其另外包含前导序列,例如以帮助分泌至少一种SARS-CoV-2刺突蛋白或其免疫原性片段。可以使用任何合适的前导序列,包括本领域已知的常规前导序列。合适的前导序列包括人组织纤溶酶原激活物前导序列(tPA),其常规用于基于病毒和DNA的疫苗以及用于蛋白质疫苗以帮助从哺乳动物细胞分泌。

[0149] 至少一种SARS-CoV-2刺突蛋白或其免疫原性片段可另外包含N-或C-末端标签,例如以帮助重组生产和/或纯化所述至少一种SARS-CoV-2刺突蛋白或其免疫原性片段。可以使用任何N-或C-末端标签,包括本领域已知的常规标签。合适的标签序列包括C-末端六组氨酸标签和"C-标签"(在C-末端的四个氨基酸EPEA),其在本领域中通常用于帮助从异源表达系统(例如昆虫细胞、哺乳动物细胞、细菌或酵母)纯化。在其他实施方案中,本发明的至少一种SARS-CoV-2刺突蛋白或其免疫原性片段是从异源表达系统纯化而无需使用纯化标签。

[0150] 本发明的至少一种SARS-CoV-2刺突蛋白或其免疫原性片段可包含如本文所定义的前导序列和/或标签。

[0151] 病毒载体、DNA质粒和RNA疫苗

[0152] 在本发明的流感-COVID-19联合疫苗中,一种或多种源自SARS-CoV-2的抗原(例如SARS-CoV-2刺突蛋白)或其免疫原性片段可以由一种或多种病毒载体、DNA载体(或DNA质粒)或RNA疫苗编码或表达。如本文所用的术语"载体"是指病毒载体、DNA载体(或DNA质粒)或RNA疫苗。

[0153] 所述一种或多种病毒载体、DNA载体(或DNA质粒)或RNA疫苗可包含一种或多种编码至少一种源自SARS-CoV-2的抗原的多核苷酸,如本文所述。优选,所述一种或多种病毒载体、DNA载体(或DNA质粒)或RNA疫苗编码如本文所述的至少一种SARS-CoV-2刺突蛋白或其免疫原性片段。多种SARS-CoV-2抗原(特别是一种或多种SARS-CoV-2刺突蛋白)可以通过单一病毒载体、DNA载体(或DNA质粒)或RNA疫苗或通过多种病毒载体、DNA载体(或DNA质粒)或RNA疫苗或其组合来表达。作为非限制性实例,所述一种或多种SARS-CoV-2抗原(特别是一种或多种SARS-CoV-2刺突蛋白)可以由单一病毒载体、DNA载体(或DNA质粒)或RNA疫苗表达,或者所述SARS-CoV-2抗原(特别是一种或多种SARS-CoV-2刺突蛋白)的每种可以由单独的病毒载体、DNA载体(或DNA质粒)或RNA疫苗表达。

[0154] 在本发明的流感-COVID-19联合疫苗中表达一种或多种SARS-CoV-2刺突蛋白或其免疫原性片段的一种或多种病毒载体、DNA载体(或DNA质粒)或RNA疫苗可以表达至少一种与SEQ ID NO:1具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的来自SARS-CoV-2的刺突蛋白或其与所述刺突蛋白具有共同抗原交叉反应性的片段。优选,所述一种或多种病毒载体、DNA载体(或DNA质粒)或RNA疫苗表达至少一种与SEQ ID NO:1具有至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的来自SARS-CoV-2的刺突蛋白或其与所述刺突蛋白具有共同抗原交叉反应性的片段。更优选,所述一种或多种病毒载体、DNA载体(或DNA质粒)或RNA疫苗表达至少一种与SEQ ID NO:1具有至少98%、至少99%或更高同一性的来自SARS-CoV-2的刺突蛋白或其与所述刺突蛋白具有共同抗原交叉反应性的片段。所述一种或多种病毒载体、DNA载体(或DNA质粒)或RNA疫苗可以表达至少一种包含SEQ ID NO:1或由SEQ IDNO:1组成的来自SARS-CoV-2的刺突蛋白或其与所述刺突蛋白具有共同抗原交叉反应性的片段。在一些优选实施方案中,由本发明的载体表达的至少一种来自SARS-CoV-2的刺突蛋白或其免疫原性片段是如本文所定义的SARS-CoV-2刺突蛋白的RBD,优选其中所述RBD与SEQ ID NO:15具有至少90%的同一性。

[0155] 通常,所述一种或多种病毒载体、DNA载体(或DNA质粒)或RNA疫苗表达至少一种与SEQ ID NO:1具有至少90%同一性的来自SARS-CoV-2的刺突蛋白或其与所述刺突蛋白具有共同抗原交叉反应性的片段,或如本文所述的其任何变体。优选的片段是与SEQ ID NO:15具有至少90%同一性的RBD。

[0156] 在本发明的流感-COVID-19联合疫苗中表达至少一种SARS-CoV-2刺突蛋白或其免疫原性片段的一种或多种病毒载体、DNA载体(或DNA质粒)或RNA疫苗可以表达至少一种如本文定义的刺突蛋白或其免疫原性片段,其进一步包含信号肽。通常,所述信号肽指导所述至少一种SARS-CoV-2刺突蛋白或其片段从目标宿主细胞(例如待治疗患者的细胞)分泌。

[0157] 在本发明的流感-COVID-19联合疫苗中表达至少一种SARS-CoV-2刺突蛋白或其免疫原性片段的一种或多种病毒载体、DNA载体(或DNA质粒)或RNA疫苗可以进一步表达一种或多种其他抗原或其片段。刺突蛋白或其片段和一种或多种另外的抗原或其片段可以表达为融合蛋白。或者,可以使用表达SARS-CoV-2刺突蛋白或其片段和一种或多种另外的抗原或其片段的单独载体。在这种情况下,所述单独的载体可以组合使用,优选同时使用。所述一种或多种另外的抗原可以是与SARS-CoV-2或其片段相同的抗原或不同的抗原。更优选,所述一种或多种另外的抗原是来自SARS-CoV-2的不同抗原,如来自2019-CoV衣壳、膜蛋白或包膜蛋白的抗原。

[0158] 在本发明的流感-COVID-19联合疫苗中表达至少一种SARS-CoV-2刺突蛋白或其免疫原性片段的一种或多种病毒载体、DNA载体(或DNA质粒)或RNA疫苗可以包含本文定义的任意一种或多种多核苷酸或表达构建体,或其任意组合。

[0159] 一种或多种载体可以是病毒载体。这样的病毒载体可以是腺病毒(人血清型的,如AdHu5,猿猴血清型,如ChAd63、ChAdOX1或ChAdOX2,或另一种形式)、腺相关病毒(AAV)或痘病毒载体(如修饰的安卡拉痘苗(MVA))或腺相关病毒(AAV)。ChAdOx1和ChAdOx2公开于WO2012/172277中(通过引用整体并入本文)。ChAdOx2是BAC衍生的且是E4修饰的基于AdC68的病毒载体。优选,所述一种或多种病毒载体是AAV载体腺病毒。病毒载体的其他非限制性实例包括麻疹病毒载体、腮腺炎病毒载体、风疹病毒载体、水痘病毒载体、脊髓灰质炎病毒载体和黄热病病毒载体。

[0160] 病毒载体通常是非复制或复制受损的载体,这意味着病毒载体不能在正常细胞(例如正常人细胞)中复制至任何显著程度,如通过常规手段所测量的-例如通过测量DNA合成和/或病毒滴度。非复制或复制受损的载体可能已经变得如此天然(即它们已经这样从自然界中分离)或人工(例如通过体外育种或通过遗传操作)。通常存在至少一种细胞类型,其中复制受损的病毒载体可以生长-例如,修饰的安卡拉痘苗(MVA)可以在CEF细胞中生长。作为非限制性实例,载体可以选自人或猿猴腺病毒或痘病毒载体。

[0161] 通常,一种或多种病毒载体不能在动物受试者(通常哺乳动物受试者,如人或其他灵长类动物)中引起显著感染。

[0162] 一种或多种载体可以是DNA载体,如DNA质粒。一种或多种载体可以是RNA载体,例如mRNA载体或自扩增RNA载体。本发明的一种或多种DNA和/或RNA载体通常能够在真核细胞,特别是本文所述的任何宿主细胞类型中表达,或在待治疗的患者中表达。

[0163] 通常,DNA和/或RNA载体能够在人、大肠杆菌或酵母细胞中表达。

[0164] 一种或多种载体可以是噬菌体载体,如Hajitou等,Cell 2006;125(2)pp.385-398中所述的AAV/噬菌体杂合体;通过引用并入本文。

[0165] 本发明的核酸分子和载体可以使用本领域已知的任何合适的方法制备。因此,可以使用化学合成技术制备核酸分子。或者,可以使用分子生物学技术制备本发明的核酸分子和载体。

[0166] 本发明的载体可以在计算机上设计,然后通过常规的多核苷酸合成技术合成。

[0167] 病毒样颗粒

[0168] 在本发明的流感-COVID-19联合疫苗中,一种或多种源自SARS-CoV-2的抗原(例如,至少一种SARS-CoV-2刺突蛋白)或其免疫原性片段可以包含在病毒样颗粒(VLP)中。

[0169] 病毒样颗粒(VLP)是类似于病毒但不含病毒核酸的颗粒,且因此是非感染性的。它们通常含有一种或多种能够自组装形成VLP的病毒衣壳或包膜蛋白。VLP已经由多种病毒家族的组分产生(Noad和Roy(2003),Trends in Microbiology,11:43 8-444;Grgacic等,(2006),Methods,40:60-65)。一些VLP已被批准作为治疗性疫苗,例如Engerix-B(用于乙型肝炎)、Cervarix和Gardasil(用于人乳头瘤病毒)。

[0170] 多种SARS-CoV-2抗原(特别是一种或多种SARS-CoV-2刺突蛋白)可以包含在VLP或其组合中。作为非限制性实例,所述一种或多种SARS-CoV-2抗原(特别是一种或多种SARS-CoV-2刺突蛋白)可以包含在单一VLP中,或者每种所述SARS-CoV-2抗原(特别是一种或多种SARS-CoV-2刺突蛋白)可以包含在分开的VLP中。

[0171] 因此,一种或多种源自SARS-CoV-2的抗原(例如,至少一种SARS-CoV-2刺突蛋白)或其免疫原性片段可以包含在一种或多种VLP中。

[0172] 在本发明的流感-COVID-19联合疫苗中包含至少一种SARS-CoV-2刺突蛋白或其免疫原性片段的一种或多种VLP可以包含一种或多种与SEQ ID NO:1具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的来自SARS-CoV-2的刺突蛋白或其与所述刺突蛋白具有共同抗原交叉反应性的片段。优选,所述一种或多种VLP包含一种或多种与SEQ ID NO:1具有至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的来自SARS-CoV-2的刺突蛋白或其与所述刺突蛋白具有共同抗原交叉反应性的片段。更优选,所述一种或多种VLP包含一种或多种与SEQ ID NO:1具有至少98%、至少99%或更高同一性的来自SARS-CoV-2的刺突蛋白或其与所述刺突蛋白具有共同抗原交叉反应性的片段。所述一种或多种VLP可以包含至少一种包含SEQ ID NO:1或由SEQ ID NO:1组成的来自SARS-CoV-2的刺突蛋白或其与所述刺突蛋白具有共同抗原交叉反应性的片段。在一些优选实施方案中,包含在本发明的VLP中的SARS-CoV-2刺突蛋白的免疫原性片段是如本文所定义的SARS-CoV-2刺突蛋白的RBD,优选其中所述RBD与SEQ ID NO:15具有至少90%的同一性。

[0173] 通常,所述一种或多种VLP包含至少一种与SEQ ID NO:1具有至少90%同一性的来自SARS-CoV-2的刺突蛋白,或其与所述刺突蛋白具有共同抗原交叉反应性的片段,或其如本文所述的任何变体。优选的片段是与SEQ ID NO:15具有至少90%同一性的RBD。

[0174] 本领域技术人员将理解,可以通过病毒结构蛋白的单独表达来合成VLP,其然后可以自组装成病毒样结构。来自不同病毒的结构衣壳蛋白的组合可用于形成重组VLP。此外,抗原或其免疫原性片段可以与VLP的表面融合。作为非限制性实例,可以使用SpyCatcher-SpyTag系统(如Brune、Biswas、Howarth所述)将本发明的抗原或其免疫原性片段与VLP偶联。

[0175] 所述一种或多种VLP可包含一种或多种另外的蛋白质抗原。所述一种或多种另外的抗原可以是与SARS-CoV-2或其片段相同的抗原或不同的抗原。更优选,所述一种或多种另外的抗原是来自SARS-CoV-2的不同抗原,如来自SARS-CoV-2衣壳、膜蛋白或包膜蛋白的抗原。

[0176] 所述一种或多种VLP可包含至少一种如本文所述的融合蛋白。所述一种或多种VLP可包含SARS-CoV-2刺突蛋白或其免疫原性片段与乙型肝炎表面抗原(HBsAg)、人乳头瘤病毒(HPV)18L1蛋白、HPV 16L1蛋白和/或戊型肝炎P239(优选乙型肝炎表面抗原)的融合蛋

白。

[0177] 因此,所述一种或多种VLP可以由包含与SEQ ID NO:3、5、6或8中的任一个具有至少70％、至少75％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或更高同一性的核酸序列或由其组成的多核苷酸编码。优选,所述一种或多种VLP可以由包含与SEQ ID NO:3、5、6或8中的任一个具有至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或更高同一性的核酸序列或由其组成的多核苷酸编码。更优选,所述一种或多种VLP可以由包含与SEQ ID NO:3、5、6或8中的任一个具有至少98％、至少99％或更高同一性的核酸序列或由其组成的多核苷酸编码。所述一种或多种VLP可以由包含SEQ ID NO:3、5、6或8中任一个的核酸序列或由其组成的多核苷酸编码。

[0178] 本发明的VLP可以由包含与SEQ ID NO:26、27、29、30或32中的任一个具有至少70％、至少75％、至少80％、至少90％、至少95％、至少95％、至少96％、至少97％、至少98％、至少99％或更高同一性的核酸序列或由其组成的多核苷酸编码。优选地,本发明的VLP可以由包含与SEQ ID NO:26、27、29、30或32中的任一个具有至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或更高同一性的核酸序列或由其组成的多核苷酸编码。更优选,本发明的VLP可以由包含与SEQ ID NO:26、27、29、30或32中的任一个具有至少98％、至少99％或更高同一性的核酸序列或由其组成的多核苷酸编码。本发明的VLP可以由包含SEQ ID NO:26、27、29、30或32中任一个的核酸序列或由其组成的多核苷酸编码。

[0179] 所述一个或多个VLP可包含与SEQ ID NO:9、10、11或12中的任一个具有至少70％、至少75％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或更高同一性的氨基酸序列或由其组成。优选,所述VLP可包含与SEQ ID NO:9、10、11或12中的任一个具有至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或更高同一性的氨基酸序列或由其组成。更优选,所述一个或多个VLP包含与SEQ ID NO:9、10、11或12中的任一个具有至少98％、至少99％或更高同一性的氨基酸序列或由其组成。所述VLP可包含SEQ ID NO:9、10、11或12中的任一个的氨基酸序列或由其组成。

[0180] 本发明的VLP可包含与SEQ ID NO:28、31或33中的任一个具有至少70％、至少75％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或更高同一性的氨基酸序列或由其组成。优选,本发明的VLP可包含与SEQ ID NO:28、31或33中的任一个具有至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或更高同一性的氨基酸序列或由其组成。更优选,本发明的VLP可包含与SEQ ID NO:28、31或33中的任一个具有至少98％、至少99％或更高同一性的氨基酸序列或由其组成。本发明的VLP可包含SEQ IDNO:28、31或33中任一个的氨基酸序列或由其组成。

[0181] 使用一种或多种VLP可以增加由SARS-CoV-2刺突蛋白或免疫原性片段诱导的免疫保护应答的功效和/或可以增加如本文所定义的免疫保护应答的持续时间。

[0182] 融合蛋白

[0183] 在本发明的流感-COVID-19联合疫苗中,一种或多种源自SARS-CoV-2的抗原(例如一种或多种SARS-CoV-2刺突蛋白)或其免疫原性片段可以包含在融合蛋白中。

[0184] 因此,一种或多种源自SARS-CoV-2的抗原(例如一种或多种SARS-CoV-2刺突蛋白)或其免疫原性片段可以包含在一种或多种融合蛋白中。

[0185] 多种SARS-CoV-2抗原(特别是一种或多种SARS-CoV-2刺突蛋白)可以包含在融合

蛋白或其组合中。作为非限制性实例,所述一种或多种SARS-CoV-2抗原(特别是一种或多种SARS-CoV-2刺突蛋白)可以包含在单一融合蛋白中,或者每种所述SARS-CoV-2抗原(特别是一种或多种SARS-CoV-2刺突蛋白)可以包含在分开的融合蛋白中。

[0186] 在本发明的流感-COVID-19联合疫苗中包含至少一种SARS-CoV-2刺突蛋白或其免疫原性片段的一种或多种融合蛋白可以包含一种或多种与SEQ ID NO:1具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的来自SARS-CoV-2的刺突蛋白或其与所述刺突蛋白具有共同抗原交叉反应性的片段。优选,所述一种或多种融合蛋白包含一种或多种与SEQ ID NO:1具有至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的来自SARS-CoV-2的刺突蛋白或其与所述刺突蛋白具有共同抗原交叉反应性的片段。更优选,所述一种或多种融合蛋白包含一种或多种与SEQ ID NO:1具有至少98%、至少99%或更高同一性的来自SARS-CoV-2的刺突蛋白或其与所述刺突蛋白具有共同抗原交叉反应性的片段。所述一种或多种融合蛋白可以包含至少一种包含SEQ ID NO:1或由SEQ ID NO:1组成的来自SARS-CoV-2的刺突蛋白或其与所述刺突蛋白具有共同抗原交叉反应性的片段。

[0187] 通常,所述一种或多种融合蛋白包含至少一种与SEQ ID NO:1具有至少90%同一性的来自SARS-CoV-2的刺突蛋白,或其与所述刺突蛋白具有共同抗原交叉反应性的片段,或如本文所述的其任何变体。

[0188] 在一些优选实施方案中,包含在本发明融合蛋白中的SARS-CoV-2刺突蛋白的免疫原性片段是如本文所定义的SARS-CoV-2刺突蛋白的RBD,优选其中所述RBD与SEQ ID NO:15具有至少90%的同一性。

[0189] 本发明的融合蛋白通常还包含非SARS-CoV-2结构域或元件,通常为非SARS-CoV-2蛋白、多肽或肽结构域或元件。

[0190] 所述一种或多种融合蛋白可包含至少一种SARS-CoV-2刺突蛋白或其免疫原性片段和以下一种或多种:乙型肝炎表面抗原(HBsAg);人乳头瘤病毒(HPV)18L1蛋白;HPV 16L1蛋白;和/或戊型肝炎P239,优选乙型肝炎表面抗原。

[0191] 所述一种或多种融合蛋白可以由包含与SEQ ID NO:3、5、6或8中的任一个具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的核酸序列或由其组成的多核苷酸编码。优选,所述一种或多种融合蛋白可以由包含与SEQ ID NO:3、5、6或8中的任一个具有至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的核酸序列或由其组成的多核苷酸编码。更优选,所述一种或多种融合蛋白可以由包含与SEQ ID NO:3、5、6或8中的任一个具有至少98%、至少99%或更高同一性的核酸序列或由其组成的多核苷酸编码。所述一种或多种融合蛋白可以由包含SEQ ID NO:3、5、6或8中任一个的核酸序列或由其组成的多核苷酸编码。

[0192] 本发明的融合蛋白可以由包含与SEQ ID NO:26、27、29、30或32中的任一个具有至少70%、至少75%、至少80%、至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的核酸序列或由其组成的多核苷酸编码。优选,本发明的融合蛋白可以由包含与SEQ ID NO:26、27、29、30或32中的任一个具有至少90%、至少95%、至少96%、至少97%、至少98%、至少99%或更高同一性的核酸序列或由其组成的多核苷酸编码。更优选,本发明的融合蛋白可以由包含与SEQ ID NO:26、27、29、30或32中的任一个具有至少

98％、至少99％或更高同一性的核酸序列或由其组成的多核苷酸编码。本发明的融合蛋白可以由包含SEQ ID NO:26、27、29、30或32中任一个的核酸序列或由其组成的多核苷酸编码。

[0193] 所述一种或多种融合蛋白可包含与SEQ ID NO:9、10、11或12中的任一个具有至少70％、至少75％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或更高同一性的氨基酸序列或由其组成。优选,所述一种或多种融合蛋白可包含与SEQ ID NO:9、10、11或12中的任一个具有至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或更高同一性的氨基酸序列或由其组成。更优选,所述一种或多种融合蛋白可包含与SEQ IDNO:9、10、11或12中的任一个具有至少98％、至少99％或更高同一性的氨基酸序列或由其组成。所述一种或多种融合蛋白可包含SEQ ID NO:9、10、11或12中任一个的氨基酸序列或由其组成。

[0194] 本发明的融合蛋白可包含与SEQ ID NO:28、31或33中的任一个具有至少70％、至少75％、至少80％、至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或更高同一性的氨基酸序列或由其组成。优选,本发明的融合蛋白可包含与SEQ ID NO:28、31或33中的任一个具有至少90％、至少95％、至少96％、至少97％、至少98％、至少99％或更高同一性的氨基酸序列或由其组成。更优选,本发明的融合蛋白可包含与SEQ ID NO:28、31或33中的任一个具有至少98％、至少99％或更高同一性的氨基酸序列或由其组成。本发明的融合蛋白可包含SEQ ID NO:28、31或33中任一个的氨基酸序列或由其组成。

[0195] 所述一种或多种融合蛋白可优选采用VLP形式。不受理论束缚,这是因为已知HPSAg、HPV 18L1蛋白、HPB 16L1蛋白和戊型肝炎P239蛋白在重组表达时自发形成VLP,并且当HPsAg、HPV 18L1蛋白、HPB 16L1蛋白和/或戊型肝炎P239蛋白以与本发明的SARS-CoV-2刺突蛋白(或其免疫原性片段)组合的融合蛋白形式存在时,这种结构得以保留。

[0196] 本发明的融合蛋白可以包含接头(在本文中也可互换地称为接头肽、间隔物或间隔肽)。接头可用于连接本发明的融合蛋白的两个或更多个功能结构域。通常,存在接头时,其用于将融合蛋白的SARS-CoV-2刺突蛋白或其免疫原性片段结构域连接到融合蛋白的非SARS-CoV-2刺突蛋白结构域。在融合蛋白中使用接头在本领域中是常规的,并且任何常规接头蛋白都可以用于本发明的融合蛋白中,条件是所得融合蛋白保留SARS-CoV-2刺突蛋白或其免疫原性片段的所需功能特性和非SARS-CoV-2刺突蛋白结构域的所需功能特性。

[0197] 接头可以是长度多达约30个氨基酸,如约5-30个氨基酸、约5-25个氨基酸、约5-20个氨基酸、约10-20个氨基酸、约5-15个氨基酸或约10-15个氨基酸的短肽。在一些实施方案中,接头的长度为约10、约11、约12、约13、约14、约15、约16、约17、约18、约19或约20个氨基酸。

[0198] 在一些实施方案中,刚性接头可用于本发明的融合蛋白中。当需要保持融合蛋白的不同结构域/部分之间的固定距离并保持其独立功能时,通常使用刚性接头。当融合蛋白结构域的空间分离对于保持融合蛋白的稳定性或生物活性至关重要时,也可以使用刚性接头。具有序列A(EAAAK)$_n$A(n=2-5)SEQ ID NO:17)的经验刚性接头显示出α-螺旋构象,其通过Glu^--Lys^+盐桥稳定。刚性接头的非限制性实例是EAAAKEAAAKEAAAK(也称为(EAAAK)$_3$,SEQ ID NO:18),其可以由核酸序列(SEQ ID NO:17)编码。刚性接头可优选用于在哺乳动物细胞(如HEK 293细胞)中表达本发明的融合蛋白。

[0199] 在一些实施方案中,柔性接头可用于本发明的融合蛋白中。当连接的结构域需要一定程度的移动或相互作用时,通常使用柔性接头。柔性接头通常包含小氨基酸残基或由小氨基酸残基组成,如甘氨酸、苏氨酸、精氨酸、丝氨酸、天冬酰胺、谷氨酰胺、丙氨酸、天冬氨酸、脯氨酸、谷氨酸、赖氨酸、亮氨酸和/或缬氨酸,特别是甘氨酸、丝氨酸、丙氨酸、亮氨酸和/或缬氨酸。优选包含甘氨酸、丝氨酸和/或丙氨酸或由甘氨酸、丝氨酸和/或丙氨酸组成的柔性接头,其中特别优选甘氨酸和丝氨酸。因此,最常用的柔性接头具有主要由Gly和Ser残基的片段组成的序列("GS"接头),其包含$(Gly-Gly-Gly-Gly-Ser)_n$(SEQ ID NO:19)的序列。GS接头的非限制性实例包括GS5或$(GGGGS)_1$(SEQ ID NO:20);GS10或$(GGGGS)_2$(SEQ ID NO:21);GS15或$(GGGGS)_3$(SEQ ID NO:23);GS20或$(GGGGS)_4$(SEQ ID NO:24);和GS25或$(GGGGS)_5$(SEQ ID NO:25)。优选,可以使用GS15,其可以由(SEQ ID NO:22)编码。柔性接头可优选用于在细菌细胞(如大肠杆菌细胞)中表达本发明的融合蛋白。

[0200] 任何合适的接头,如本文所述的示例性接头,都可与本发明的任何融合蛋白(包含任何2019-NCoV刺突蛋白或免疫原性片段结构域和任何非219-NCoV刺突蛋白结构域)一起使用。作为非限制性实例,本发明的融合蛋白可以包含(或由其组成)$HBsAg-(EAAAK)_3$-RBD(SEQ ID NO:28),或与其具有至少90%序列同一性的变体(其可以由SEQ ID NO:26或27编码),或与其具有至少90%序列同一性的变体。作为另外的非限制性实例,本发明的融合蛋白可以包含(或由其组成)$HBsAg-(EAAAK)_3$-全长2019-nCoV刺突蛋白(SEQ ID NO:33),或与其具有至少90%序列同一性的变体(其可以由SEQ ID NO:32编码),或与其具有至少90%序列同一性的变体。作为另外的非限制性实例,本发明的融合蛋白可以包含(或由其组成)HEV-GS15-RBD(SEQ ID NO:31),或与其具有至少90%序列同一性的变体(其可以由(SEQ ID NO:29或30)编码),或与其具有至少90%序列同一性的变体。

[0201] 本发明的融合蛋白可以优选采用VLP的形式。不受理论的约束,这是因为已知HBsAg、HPV 18LI蛋白、HPB 16LI蛋白和戊型肝炎P239蛋白在重组表达时自发形成VLP,并且在HBsAg、HPV 18LI蛋白、HPB 16LI蛋白和/或戊型肝炎P239蛋白以与本发明的2019-nCoV刺突蛋白(或其免疫原性片段)组合的融合蛋白形式存在时,该结构得以保留。

[0202] 流感血凝素(HA)和神经氨酸酶(NA)抗原

[0203] 本发明的流感-COVID-19联合疫苗包含流感血凝素(HA)或其免疫原性片段。任选,本发明的流感-COVID-19联合疫苗还可以包含流感神经氨酸酶(NA)或其免疫原性片段。

[0204] HA的免疫原性片段与衍生其的HA具有共同的抗原交叉反应性。类似地,NA的免疫原性片段与衍生其的NA具有共同的抗原交叉反应性。

[0205] 流感HA或其免疫原性片段(和任选的流感NA或其免疫原性片段)可以以任何适当的形式存在于流感-COVID-19联合疫苗中。

[0206] 流感HA或其免疫原性片段和/或流感NA或其免疫原性片段通常由流感病毒粒子制备,但是作为替代方案,这些抗原可以以其他形式提供,如多核苷酸、病毒载体、DNA载体(或DNA质粒)或RNA疫苗、VLP和融合蛋白。

[0207] 本文关于多核苷酸、病毒载体、DNA载体(或DNA质粒)或RNA疫苗、VLP和融合蛋白的一般公开内容也适用于如本文所述的流感HA或其免疫原性片段和流感NA或其免疫原性片段。本文关于源自SARS-CoV-2的抗原(例如SARS-CoV-2刺突蛋白)的背景下的多核苷酸、病毒载体、DNA载体(或DNA质粒)或RNA疫苗、VLP和融合蛋白的任何一般公开内容同样适用于

但不限于本文所述的流感HA或其免疫原性片段和流感NA或其免疫原性片段。

[0208] 如本文所述,(a)流感HA或其免疫原性片段可以是(i)包含在灭活的流感病毒粒子中;(ii)重组HA或其免疫原性片段;(iii)包含HA或其免疫原性片段的融合蛋白;或(iv)由RNA或DNA疫苗编码。

[0209] 如本文所述,(a)流感NA或其免疫原性片段可以是(i)包含在灭活的流感病毒粒子中;(ii)重组NA或其免疫原性片段;(iii)包含NA或其免疫原性片段的融合蛋白;或(iv)由RNA或DNA疫苗编码。

[0210] 流感HA或其免疫原性片段和/或流感NA或其免疫原性片段可以采用现有流感疫苗的形式。流感HA或其免疫原性片段和/或流感NA或其免疫原性片段可采用活(减毒或载体的)疫苗、灭活疫苗或亚单位疫苗的形式。灭活流感疫苗包括灭活全病毒粒疫苗和灭活裂解病毒粒疫苗,优选全病毒粒灭活疫苗。通过用洗涤剂(例如乙醚、聚山梨醇酯80、脱氧胆酸盐、磷酸三正丁酯、Triton X-100、Triton N101、十六烷基三甲基溴化铵、Tergitol NP9等)处理病毒粒子以产生亚病毒粒制剂来获得裂解病毒粒。裂解流感病毒的方法是本领域熟知的。

[0211] 灭活疫苗可以通过任何合适的方式产生。用于灭活流感病毒粒的常规方法包括用有效量的一种或多种以下试剂处理:洗涤剂、甲醛、福尔马林、β-丙内酯或UV光。用于灭活的其他化学手段包括用亚甲基蓝、补骨脂素、羧基富勒烯(C60)或其任何组合处理。病毒灭活的其他方法是本领域已知的,如,例如二元乙胺、乙酰基乙烯亚胺或γ辐射。

[0212] 本发明的流感-COVID-19联合疫苗可以包含任何流感疫苗或使用任何流感疫苗生产,所述流感疫苗包括任何市售流感疫苗、通用流感疫苗和/或大流行性流感疫苗。

[0213] 通常,用于疫苗的流感病毒株因季节而异。在当前流感间大流行时期,疫苗通常包括两种甲型流感毒株(H1N1和H3N2)和一种乙型流感毒株(B/Colorado/06/2017样(Victoria谱系)病毒),并且针对季节性流感的三价疫苗(季节性三价流感疫苗)是典型的。针对季节性流感的四价疫苗(季节性四价流感疫苗)也是常用的。目前,季节性四价流感疫苗包括与季节性三价流感疫苗相同的毒株,包括另外的乙型流感毒株(B/Phuket/3073/2013样病毒(Yamagata谱系))。任何季节性流感疫苗,包括季节性三价和四价流感疫苗,都可以包含在本发明的流感-COVID-19联合疫苗中或用于生产本发明的流感-COVID-19联合疫苗。在疾病控制和预防中心(CDC)网站(此处提供CDC 2019-2020列表:https://www.cdc.gov/flu/professionals/acip/summary/summary-recommendations.htm#composition)上和欧洲药品管理局(EMA)鉴定监管批准的季节性流感疫苗。

[0214] 或者,大流行性流感疫苗可以包含在本发明的流感-COVID-19联合疫苗中或用于生产本发明的流感-COVID-19联合疫苗。大流行性流感疫苗是针对大流行性流感毒株产生的,所述大流行性流感毒株是疫苗接受者和一般人群在免疫上幼稚的毒株,如H2、H5、H7或H9亚型毒株(特别是甲型流感病毒)。大流行性流感病毒株通常出现在非人物种中,然后跨越物种屏障到达人类。潜在的大流行性流感毒株的最近实例是基因型4(G4)欧亚禽样(EA)H1N1猪流感毒株。本发明的流感-COVID-19联合疫苗可以包含针对此类跨物种大流行毒株(如G4 EA H1N1)的流感成分。大流行性流感疫苗可以是单价的,或者可以基于由大流行性毒株补充的三价疫苗。优选单价大流行性流感疫苗。

[0215] 通用流感疫苗可以包含在本发明的流感-COVID-19联合疫苗中或用于生产本发明

的流感-COVID-19联合疫苗。正在开发的通用流感疫苗的实例包括亚单位疫苗和包含引发DNA疫苗和活载体疫苗的两阶段疫苗。

[0216] 根据季节和疫苗中包括的HA和/或NA的性质,本发明的流感-COVID-19联合疫苗的流感成分可以针对甲型流感病毒血凝素亚型H1、H2、H3、H4、H5、H6、H7、H8、H9、H10、H11、H12、H13、H14、H15或H16中的一种或多种提供保护。本发明可以针对甲型流感病毒NA亚型N1、N2、N3、N4、N5、N6、N7、N8或N9中的一种或多种进行保护。

[0217] 本发明的流感-COVID-19联合疫苗的流感成分可以包括来自一种或多种(例如,1、2、3、4或更多种)流感毒株(包括甲型流感病毒和/或乙型流感病毒)的HA和/或NA(或其免疫原性片段)。

[0218] 用作流感HA和/或NA来源的病毒或形成流感-COVID-19联合疫苗的流感组分的流感疫苗可以在卵或细胞培养物上生长。目前用于流感病毒生长的标准方法使用无特定病原体(SPF)的含胚鸡卵,其中病毒从卵内容物(尿囊液)中纯化。然而,最近,病毒已经在动物细胞培养物中生长,并且由于速度和患者过敏的原因,这种生长方法是优选的。如果使用基于卵的病毒生长,则可以将一种或多种氨基酸与病毒一起引入卵的尿囊液中。当使用细胞培养物时,病毒生长基质通常是哺乳动物来源的细胞系。合适的哺乳动物来源细胞包括但不限于仓鼠、牛、灵长类动物(包括人和猴)和狗细胞。可以使用各种细胞类型,如肾细胞、成纤维细胞、视网膜细胞、肺细胞等。合适的细胞系包括但不限于:MDCK;CHO;293T;BHK;Vero;MRC-5;PER.C6;WI-38;等。用于使流感病毒生长的优选哺乳动物细胞系包括:源自Madin Darby犬肾的MDCK细胞,其可从例如美国典型细胞培养物(ATCC)保藏中心作为CCL-34获得。也可以使用MDCK细胞系的衍生物。

[0219] 在病毒已经在哺乳动物细胞系上生长的情况下,组合物将有利地不含卵蛋白(例如卵清蛋白和卵类粘蛋白)和鸡DNA,从而降低变应原性。

[0220] 组合物和治疗适应症

[0221] 如本文所述,本发明人已经证明了包含SARS-CoV-2抗原,特别是SARS-CoV-2刺突蛋白的疫苗组合物可以成功地与流感病毒疫苗组合,以产生稳健的对SARS-CoV-2和流感的抗体应答。因此,本发明令人惊讶地证明了可以生产流感-COVID-19联合疫苗,而没有预期的在联合疫苗产品生产中常见的疫苗成分抑制问题。

[0222] 因此,本发明提供了如本文所述的流感-COVID-19联合疫苗。本发明提供了一种组合物,其包含(i)流感HA抗原或其免疫原性片段;(ii)一种或多种源自SARS-CoV-2的抗原(特别是至少一种SARS-CoV-2刺突蛋白)或其免疫原性片段;和任选(iii)流感NA抗原或其免疫原性片段;其中所述组合物能够诱导对抗SARS-CoV-2(特别是针对SARS-CoV-2刺突蛋白)和流感(特别是流感HA和任选的NA)的免疫应答。本发明还提供了这种组合物作为疫苗的用途。

[0223] 本发明还提供了疫苗组合物,其包含(i)流感HA抗原或其免疫原性片段;(iii)一种或多种源自SARS-CoV-2的抗原(特别是至少一种SARS-CoV-2刺突蛋白)或其免疫原性片段;和任选(iii)流感NA抗原或其免疫原性片段。疫苗组合物可任选包含药学上可接受的赋形剂、稀释剂、载体、推进剂、盐和/或添加剂。

[0224] 疫苗组合物可包含至少两种不同的根据本发明的源自SARS-CoV-2或其免疫原性片段的抗原,和/或至少两种不同的编码至少两种不同的源自SARS-CoV-2或其免疫原性片

段的抗原的多核苷酸分子,如本文所述。作为非限制性实例,疫苗组合物可包含编码SARS-CoV-2刺突蛋白的多核苷酸和编码SARS-CoV-2膜蛋白的多核苷酸。

[0225] 疫苗组合物可包含至少两种不同的源自根据本发明的流感或其免疫原性片段的抗原,和/或编码至少两种不同的源自流感或免疫原性片段的抗原的至少两种不同的多核苷酸分子,如本文所述。通常,疫苗组合物包含流感HA抗原或其免疫原性片段和任选流感NA抗原或其免疫原性片段。由于本发明的流感-COVID-19联合疫苗的流感成分通常由包含完整或裂解流感病毒体的活(减毒或载体化)或灭活流感疫苗提供,因此也可以包含其他流感抗原。

[0226] 本发明还提供了一种使用本发明的流感-COVID-19联合疫苗或组合物(如上所述)刺激或诱导患者免疫应答的方法。本发明的疫苗和组合物通常刺激或诱导针对流感和COVID-19的免疫应答和/或保护。

[0227] 所述刺激或诱导个体免疫应答的方法可包括将本发明的流感-COVID-19联合疫苗或组合物(如上所述)施用于受试者。

[0228] 在治疗用途和方法的上下文中,“受试者”是将受益于对抗SARS-CoV-2和流感的免疫保护性应答的刺激或诱导的任何动物受试者。典型的动物受试者是哺乳动物,如灵长类动物,例如人。

[0229] 因此,本发明提供了一种用于治疗或预防SARS-CoV-2感染(COVID-19)和流感感染的方法。所述方法通常包括将本发明的流感-COVID-19联合疫苗或组合物施用于需要的受试者。

[0230] 本发明还提供了用于预防或治疗SARS-CoV-2感染的本发明的流感-CoVID-19联合疫苗或组合物。

[0231] 本发明还提供了(i)一种或多种表达一种或多种SARS-CoV-2刺突蛋白或其免疫原性片段的多核苷酸、表达构建体、病毒载体、DNA质粒或RNA疫苗,或本发明的一种或多种SARS-CoV-2刺突蛋白或其免疫原性片段、一种或多种SARS-CoV-2疫苗组合物的用途;和(ii)流感HA或其免疫原性片段(和任选的流感NA或其免疫原性片段)的用途,优选包含在如本文所述的流感疫苗中,用于制备预防或治疗SARS-CoV-2感染和流感感染的药物。

[0232] 如本文所用,术语“治疗(treatment)”或“治疗(treating)”包括治疗性或防止性/预防性措施,并且包括感染后治疗和改善SARS-CoV-2感染和流感感染。术语“治疗”和“治疗的”包括预防性治疗。

[0233] 如本文所用,术语“预防”包括预防SARS-CoV-2和流感感染的开始和/或降低SARS-CoV-2感染和流感的严重程度或强度。术语“预防”包括诱导或提供对抗SARS-CoV-2感染和流感感染的保护性免疫。对SARS-CoV-2感染和流感感染的免疫力可以使用任何合适的技术进行定量,其实例是本领域已知的。

[0234] 本发明的优选组合物满足1、2或3个CPMP功效标准。在成人(18-60岁)中,这些标准是:(1)≥70%血清保护;(2)≥40%血清转化;和/或(3)GMT增加≥2.5倍。在老年人(>60岁)中,这些标准是:(1)≥60%血清保护;(2)≥30%血清转化;和/或(3)GMT增加≥2倍。

[0235] 这些标准基于至少50名患者的开放标签研究。

[0236] 如本文所定义的本发明的流感-COVID-19联合疫苗或组合物可以施用于已经患有SARS-CoV-2感染和/或流感感染、与SARS-CoV-2感染和/或流感感染相关的病症或症状的受

试者(通常是哺乳动物受试者,例如人或其他灵长类),以治疗或预防SARS-CoV-2和/或流感感染。例如,受试者可能怀疑与SARS-CoV-2或流感接触,或已知与SARS-CoV-2或流感接触,但尚未显示暴露症状。

[0237] 施用于已经患有SARS-CoV-2感染和/或流感感染,或显示与SARS-CoV-2感染和/或流感感染相关的症状的受试者(如哺乳动物,如人或其他灵长类)时,如本文所定义的本发明的流感-COVID-19联合疫苗或组合物可以治愈、延迟、降低一种或多种症状的严重程度,或改善一种或多种症状,和/或延长受试者的生存期超过在没有这种治疗的情况下预期的。

[0238] 或者,本文所定义的本发明的流感-COVID-19联合疫苗或组合物可以施用于最终可能感染SARS-CoV-2和/或流感的受试者(例如,哺乳动物,如人或其他灵长类),以预防、治愈、延迟、降低所述SARS-CoV-2感染和/或流感的严重程度或改善所述SARS-CoV-2感染和/或流感的一种或多种症状,或延长受试者的存活期超过在没有这种治疗的情况下预期的,或帮助预防受试者传播SARS-CoV-2感染和/或流感感染。

[0239] 本发明的治疗和预防性疗法适用于不同年龄的各种不同受试者。在人类的情况下,所述疗法适用于儿童(例如婴儿、5岁以下儿童、较大儿童或青少年)和成人。在其他动物受试者(例如哺乳动物,如灵长类)的情况下,所述疗法适用于未成熟受试者和成熟/成年受试者。如本文所用,术语“预防”包括预防SARS-CoV-2感染和/或流感感染的开始;和/或降低SARS-CoV-2感染和/或流感感染的严重性或强度。术语“预防”包括诱导或提供对抗SARS-CoV-2感染和/或流感感染的保护性免疫。对SARS-CoV-2感染和/或流感感染的免疫力可以使用任何合适的技术进行定量,其实例是本领域已知的。

[0240] 如本文所用,“疫苗”是一种制剂,施用于动物受试者,如哺乳动物(例如人或其他灵长类动物)时,其刺激对抗SARS-CoV-2感染和/或流感感染的保护性免疫应答。免疫应答可以是体液和/或细胞介导的免疫应答。本发明的疫苗可用于例如保护受试者免受SARS-CoV-2感染和/或流感感染的影响。

[0241] 如本文所述,迄今为止可得到的证据表明,SARS-CoV-2感染后的免疫可能是相对短暂的。因此,本发明提供了通过定期重复施用COVID-19/SARS-CoV-2疫苗,特别是本发明的流感-COVID-19联合疫苗,来增强对SARS-CoV-2感染的免疫力的方法。这种重复施用可以使用或整合到现有的季节性流感疫苗接种的公共卫生计划/时间表中。

[0242] 因此,本发明提供了用于治疗和/或预防COVID-19和流感的本发明的流感-COVID-19联合疫苗,其中所述联合疫苗的给药间隔为约6个月、约7个月、约8个月、约9个月、约10个月、约11个月、约12个月、约13个月、约14个月或约15个月。优选,该联合疫苗以约11个月、约12个月、约13个月、最优选约12个月的间隔施用。本发明还提供了一种使受试者免疫以对抗流感和COVID-19的方法,包括以这些相同的间隔将治疗有效量的本发明的流感-COVID-19联合疫苗施用于所述受试者。本发明还提供了流感HA或其免疫原性片段,源自SARS-CoV-2的抗原或其免疫原性片段,和任选的流感NA或其免疫原性片段在制备用于治疗和/或预防COVID-19和流感的药物中的用途,其中所述药物用于以这些相同的间隔施用。

[0243] 该联合疫苗可以如本文所述的间隔施用至少两次、至少五次、至少十次、至少15次、至少20次或更多次。

[0244] 该联合疫苗可以如本文所述的间隔施用,持续至少两年、至少五年、至少十年或更长时间,直至患者的一生。

[0245] 药物组合物和制剂

[0246] 术语“疫苗”在本文中可与术语“治疗/预防组合物”、“制剂”或“药物”互换使用。

[0247] 除了药学上可接受的载体之外,本发明的疫苗(如上定义)可以组合或施用。可选地或另外地,本发明的疫苗可以进一步与盐、赋形剂、稀释剂、佐剂、免疫调节剂和/或抗微生物化合物中的一种或多种组合。

[0248] 药学上可接受的盐包括与无机酸(如,例如盐酸或磷酸)或与有机酸(如乙酸、草酸、酒石酸、马来酸等)形成的酸加成盐。与游离羧基形成的盐也可以衍生自无机碱,如,例如氢氧化钠、氢氧化钾、氢氧化铵、氢氧化钙或氢氧化铁,以及诸如异丙胺、三甲胺、2-乙氨基乙醇、组氨酸、普鲁卡因等的有机碱。

[0249] 免疫原性组合物、治疗制剂、药物和预防制剂(例如疫苗)的施用通常通过常规途径,例如静脉内、皮下、腹膜内或粘膜途径。施用可以通过肠胃外注射,例如皮下、皮内或肌内注射。

[0250] 因此,本发明的免疫原性组合物、治疗性制剂、药物和预防性制剂(例如疫苗)通常制成注射剂,作为液体溶液或悬浮液。或者,可以制备适于在注射前溶解或悬浮在液体中的固体形式。还可以将制剂乳化,或者将肽包封在脂质体或微胶囊中。

[0251] 活性免疫原性成分(如SARS-CoV-2刺突蛋白、其片段、编码所述刺突蛋白的核酸、表达载体、病毒载体、DNA质粒、RNA疫苗、融合蛋白和疫苗组合物以及如本文所述的流感HA和/或NA抗原或流感疫苗)通常与药学上可接受且与活性成分相容的载体、稀释剂、赋形剂或类似物混合。合适的赋形剂是例如水、盐水、右旋糖、甘油、乙醇等及其组合。此外,如果需要,疫苗可以含有少量辅助物质,如润湿剂或乳化剂、pH缓冲剂和/或增强疫苗有效性的佐剂。

[0252] 通常,载体、稀释剂、赋形剂或类似物是药学上可接受的载体。药学上可接受的载体的非限制性实例包括水、盐水和磷酸盐缓冲盐水。然而,在一些实施方案中,组合物是冻干形式,在这种情况下,其可以包括稳定剂,如BSA。在一些实施方案中,可能需要用防腐剂(如硫柳汞或叠氮化钠)配制组合物,以便于长期储存。

[0253] 缓冲剂的实例包括但不限于琥珀酸钠(pH 6.5)和磷酸盐缓冲盐水(PBS,pH 6.5和7.5)。

[0254] 适用于其他给药方式的其他制剂包括栓剂,并且在一些情况下,口服制剂或适于作为气溶胶分布的制剂。对于栓剂,传统的粘合剂和载体可包括例如聚亚烷基二醇或甘油三酯;这种栓剂可以由含有0.5%至10%、优选1%-2%范围内的活性成分的混合物形成。

[0255] 口服制剂包括通常使用的赋形剂,如,例如药物级甘露醇、乳糖、淀粉、硬脂酸镁、糖精钠、纤维素、碳酸镁等。这些组合物采取溶液、悬浮液、片剂、丸剂、胶囊、持续释放制剂或粉末的形式。

[0256] 佐剂

[0257] 虽然常规流感疫苗不包含佐剂,但是本发明的流感-COVID-19联合疫苗还可以包含佐剂。所述佐剂可以是细胞(Th1)和/或体液(Th2)免疫应答的刺激物。

[0258] 可能有效的其他佐剂的实例包括但不限于:完全弗氏佐剂(CFA)、不完全弗氏佐剂(IFA)、皂苷、皂苷的纯化提取物级分,如Quil A,皂苷的衍生物,如QS-21、基于皂苷的脂质颗粒,如ISCOM/ISCOMATRIX、大肠杆菌热不稳定性毒素(LT)突变体,如LTK63和/或LTK72、氢

氧化铝、N-乙酰基-胞壁酰-L-苏氨酰-D-异谷氨酰胺(thr-MDP)、N-乙酰基-去甲-胞壁酰-L-丙氨酰-D-异谷氨酰胺(CGP11637,称为去甲-MDP)、N-乙酰胞壁酰-L-丙氨酰-D-异谷氨酰胺酰-L-丙氨酸-2-(1'-2'-二棕榈酰-sn-甘油基-3-羟基磷酰氧基)-乙胺(CGP 19835A,称为MTP-PE)和RIBI,其含有从细菌提取的三种组分,在2%角鲨烯/吐温80乳液中的单磷酰脂质A、海藻糖二霉菌酸酯和细胞壁骨架(MPL+TDM+CWS)、由Novartis开发的MF59制剂以及由GSK Biologicals(Rixensart,比利时)开发的AS02、AS01、AS03和AS04佐剂制剂。通常存在于本发明的流感-COVID-19联合疫苗中的佐剂选自角鲨烯水包油乳液、铝盐和单磷酰脂质A(MPL)。特别优选的佐剂包括Addavax®、5%角鲨烯(MF59)、MPL和氢氧化铝和磷酸铝凝胶。

[0259] 试剂盒

[0260] 本发明提供了包含本发明的流感-COVID-19联合疫苗的试剂盒,可选地还包含使用说明书。任何佐剂都可以与联合疫苗分开包含在试剂盒内,或者可以与联合疫苗组合。试剂盒中的联合疫苗可以准备好待用(例如包括佐剂),或者可以准备好在递送时临时制备(例如掺入佐剂)。这种临时布置允许佐剂和抗原在使用前分开保存,这在使用水包油乳液佐剂时特别有用。

[0261] 本发明还提供了包含联合疫苗的SARS-CoV-2组分和联合疫苗的流感组分的试剂盒。这两种组分在试剂盒内可以是分开的。任何佐剂都可以单独包含在试剂盒内,或者可以与SARS-CoV-2组分或流感组分组合。在这种情况下,组分可以在施用于患者之前混合,或者组分可以保持分开但基本上同时或同时施用于患者。

[0262] 本发明还提供了几个部分的试剂盒,其包含联合疫苗的SARS-CoV-2组分和佐剂,佐剂优选为角鲨烯水包油乳液、铝盐或MPL,更优选Addavax®、MF59、MPL或氢氧化铝和磷酸铝凝胶。任选地,几个部分的试剂盒可以包括关于将SARS-CoV-2组分和佐剂与现有流感疫苗(其实例在本文中描述)组合,并将流感-CoVID-19联合疫苗作为单一单位施用,或将混合的SARS-CoV-2和佐剂与流感疫苗基本上同时或同时施用于患者的说明书。

[0263] 试剂盒中的SARS-CoV-2组分和/或流感组分可以随时使用,或者可以在递送时临时制备。这种临时安排允许佐剂和SARS-CoV-2和/或流感组分在使用前分开保存,这在使用水包油乳液佐剂时特别有用。

[0264] 在临时制备疫苗的情况下,其组分在试剂盒内彼此物理分离,并且这种分离可以以各种方式实现。例如,两种组分可以在两个单独的容器如小瓶中。然后可以混合两个小瓶的内容物,例如通过取出一个小瓶的内容物并将它们添加到另一个小瓶中,或通过分别取出两个小瓶的内容物并在第三个容器中混合。作为非限制性实例,试剂盒组分之一在注射器中,另一种在容器(如小瓶)中。可以使用注射器(例如,与针一起)将其内容物插入第二容器中进行混合,然后可以将混合物抽出到注射器中。然后可以将注射器的混合内容物施用给患者,通常通过新的无菌针。将一种组分包装在注射器中消除了使用单独的注射器用于患者给药的需要。作为进一步的非限制性实例,疫苗的两种组分在同一注射器(例如双室注射器)中保持在一起但分开。注射器被致动时(例如,在向患者给药期间),则两个室的内容物被混合。这种布置避免了在使用时需要单独的混合步骤。

[0265] 在临时制备疫苗的情况下(通过将组合疫苗与佐剂混合,或通过将SARS-CoV-2组分和流感组分混合,任选与佐剂混合),其组分通常为水性形式。在一些布置中,组分(通常是联合疫苗或所述疫苗的SARS-CoV-2组分和/或流感组分,而不是佐剂组分)是干燥形式

(例如冻干形式),一种或多种其他组分是水性形式。可以将组分混合以使干燥组分再活化并得到用于向患者施用的水性组合物。

[0266] 序列同源性

[0267] 可以使用多种序列比对方法中的任一种来确定同一性百分比,包括但不限于全局方法、局部方法和杂交方法,如,例如,区段逼近方法。确定同一性百分比的方案是本领域技术人员范围内的常规程序。全局方法从分子的开始到结束比对序列,并通过将各个残基对的得分相加并通过施加空位罚分来确定最佳比对。非限制性方法包括例如CLUSTAL W,参见,例如Julie D.Thompson等,CLUSTAL W:Improving the Sensitivity of Progressive Multiple Sequence Alignment Through Sequence Weighting,Position-Specific Gap Penalties and Weight Matrix Choice,22(22)Nucleic Acids Research 4673-4680(1994);和迭代精修,参见,例如Osamu Gotoh,Significant Improvement in Accuracy of Multiple Protein.Sequence Alignments by Iterative Refinement as Assessed by Reference to Structural Alignments,264(4)J.MoI.Biol.823-838(1996)。局部方法通过鉴定由所有输入序列共享的一个或多个保守基序来比对序列。非限制性方法包括例如匹配框,参见,例如Eric Depiereux和Ernest Feytmans,Match-Box:A Fundamentally New Algorithm for the Simultaneous Alignment of Seumber Protein Sequences,8(5)CABIOS 501-509(1992);Gibbs取样,参见,例如C.E.Lawrence等,Detecting Subtle Sequence Signals:A Gibbs Sampling Strategy for Multiple Alignment,262(5131)Science 208-214(1993);Align-M,参见,例如Ivo van Walle等,Align-M-A New Algorithm for Multiple Alignment of Highly Divergent Sequences,20(9)Bioinformatics:1428-1435(2004)。

[0268] 因此,通过常规方法测定序列同一性百分比。参见,例如,Altschul等,Bull.Math.Bio.48:603-16,1986以及Henikoff和Henikoff,Proc.Natl.Acad.Sci.USA89:10915-19,1992。简而言之,使用空位开放罚分10、空位延伸罚分1和如下所示的Henikoff和Henikoff(同上)的“blosum 62”评分矩阵(氨基酸由标准单字母代码表示)比对两个氨基酸序列以优化比对评分。

[0269] 用于确定序列同一性的比对评分

[0270] BLOSUM62表

[0271]

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
A	4																			
R	-1	5																		
N	-2	0	6																	
D	-2	-2	1	6																
C	0	-3	-3	-3	9															
Q	-1	1	0	0	-3	5														
E	-1	0	0	2	-4	2	5													

G 0 -2 0 -1 -3 -2 -2 6

H -2 0 1 -1 -3 0 0 -2 8

I -1 -3 -3 -3 -1 -3 -3 -4 -3 4

L -1 -2 -3 -4 -1 -2 -3 -4 -3 2 4

K -1 2 0 -1 -3 1 1 -2 -1 -3 -2 5

M -1 -1 -2 -3 -1 0 -2 -3 -2 1 2 -1 5

[0272] F -2 -3 -3 -3 -2 -3 -3 -3 -1 0 0 -3 0 6

P -1 -2 -2 -1 -3 -1 -1 -2 -2 -3 -3 -1 -2 -4 7

S 1 -1 1 0 -1 0 0 0 -1 -2 -2 0 -1 -2 -1 4

T 0 -1 0 -1 -1 -1 -1 -2 -2 -1 -1 -1 -1 -2 -1 1 5

W -3 -3 -4 -4 -2 -2 -3 -2 -2 -3 -2 -3 -1 1 -4 -3 -2 11

Y -2 -2 -2 -3 -2 -1 -2 -3 2 -1 -1 -2 -1 3 -3 -2 -2 2 7

V 0 -3 -3 -3 -1 -2 -2 -3 -3 3 1 -2 1 -1 -2 -2 0 -3 -1 4

[0273] 然后将同一性百分比计算为：

[0274] 相同匹配的总数/[较长序列的长度加上为了比对两个序列引入较长序列中的空位数]×100

[0275] 基本上同源的多肽的特征在于具有一个或多个氨基酸置换、缺失或添加。这些变化优选具有次要的性质，即保守性氨基酸置换(参见下文)和不显著影响多肽的折叠或活性的其他置换；小的缺失，通常为1至约30个氨基酸；和小的氨基末端或羧基末端延伸，如氨基末端或羧基末端延伸，如氨基末端甲硫氨酸残基，多达约20-25个残基的小接头肽，或亲和标签。

[0276] 保守性氨基酸置换：

[0277] 碱性： 精氨酸

[0278] 赖氨酸

[0279] 组氨酸

[0280] 酸性： 谷氨酸

[0281] 天冬氨酸

[0282] 极性： 谷氨酰胺

[0283] 天冬酰胺

[0284] 疏水性：亮氨酸

[0285] 异亮氨酸

[0286] 缬氨酸

[0287] 芳香族：苯丙氨酸

[0288] 色氨酸

[0289] 酪氨酸

[0290] 小的：甘氨酸

[0291] 丙氨酸

[0292] 丝氨酸

[0293] 苏氨酸

[0294] 甲硫氨酸

[0295] 除了20种标准氨基酸之外,非标准氨基酸(如4-羟基脯氨酸、6-N-甲基赖氨酸、2-氨基异丁酸、异缬氨酸和α-甲基丝氨酸)可以置换本发明多肽的氨基酸残基。有限数量的非保守性氨基酸、不由遗传密码编码的氨基酸和非天然氨基酸可以置换多肽氨基酸残基。本发明的多肽还可以包含非天然存在的氨基酸残基。

[0296] 非天然存在的氨基酸包括但不限于反式-3-甲基脯氨酸、2,4-亚甲基-脯氨酸、顺式-4-羟基脯氨酸、反式-4-羟基-脯氨酸、N-甲基甘氨酸、别苏氨酸、甲基-苏氨酸、羟基-乙基半胱氨酸、羟基乙基高半胱氨酸、硝基-谷氨酰胺、高谷氨酰胺、哌啶酸、叔亮氨酸、正缬氨酸、2-氮杂苯丙氨酸、3-氮杂苯丙氨酸、4-氮杂苯丙氨酸和4-氟苯丙氨酸。用于将非天然存在的氨基酸残基掺入蛋白质中的几种方法是本领域已知的。例如,可以采用体外系统,其中使用化学氨基酰化的抑制子tRNA抑制无义突变。用于合成氨基酸和氨酰化tRNA的方法是本领域已知的。含有无义突变的质粒的转录和翻译在包含大肠杆菌S30提取物和市售酶和其他试剂的无细胞系统中进行。通过色谱法纯化蛋白质。参见,例如,Robertson等,J.Am.Chem.Soc.113,2722,1991;Ellman等,Methods Enzymol.202:301,1991;Chung等,Science 259:806-9,193;和Chung等,Proc.Natl.Acad.Sci.USA 90:10145-9,1993)。在第二种方法中,通过显微注射突变的mRNA和化学氨基酰化的抑制子tRNAs在非洲爪蟾卵母细胞中进行翻译(Turcatti等,J.Biol.Chem.271:19991-8,1996)。在第三种方法中,在不存在待置换的天然氨基酸(例如苯丙氨酸)和存在所需的非天然存在的氨基酸(例如2-氮杂苯丙氨酸、3-氮杂苯丙氨酸、4-氮杂苯丙氨酸或4-氟苯丙氨酸)的情况下培养大肠杆菌细胞。将非天然存在的氨基酸掺入多肽中以代替其天然对应物。参见,Koide等,Biochem.33:7470-6,1994。天然存在的氨基酸残基可以通过体外化学修饰转化成非天然存在的种类。化学修饰可以与定点诱变组合以进一步扩大置换范围(Wynn和Richards,Protein Sci.2:395-403,1993)。

[0297] 有限数量的非保守性氨基酸、不由遗传密码编码的氨基酸、非天然存在的氨基酸和非天然氨基酸可以置换本发明多肽的氨基酸残基。

[0298] 本发明的多肽中的必需氨基酸可以根据本领域已知的程序鉴定,如定点诱变或丙氨酸扫描诱变(Cunningham和Wells,Science 244:1081-5,1989)。生物相互作用的位点也可以通过结构的物理分析来确定,如通过诸如核磁共振、晶体学、电子衍射或光亲和标记的技术结合推定的接触位点氨基酸的突变来确定。参见,例如de Vos等,Science 255:306-12,199 2;Smith等,J.Mol.Biol.224:899-904,1992;Wlodaver等,FEBS Lett.309:59-64,1992。还可以从与本发明多肽的相关组分(例如,易位或蛋白酶组分)的同源性分析推断必需氨基酸的身份。

[0299] 可以使用已知的诱变和筛选方法进行和测试多个氨基酸置换,如Reidhaar-Olson和Sauer(Science241:53-7,1988)或Bowie和Sauer(Proc.Natl.Acad.Sci.USA 86:2152-6,1989)公开的那些。简而言之,这些作者公开了用于同时随机化多肽中的两个或更多个位

置,选择功能性多肽,然后对诱变的多肽进行测序以确定每个位置处可允许的置换谱的方法。可以使用的其他方法包括噬菌体展示(例如,Lowman等,Biochem.30:10832-7,1991;Ladner等,美国专利号5,223,409;Huse,WIPO公开WO 92/06204)和区域定向诱变(Derbyshire等,Gene 46:145,1986;Ner等,DNA 7:127,1988)。

[0300] 以下实施例说明本发明。

实施例

[0301] 实施例1:比较单独的三价商业流感疫苗(Addavax佐剂化的)和单独的COVID-19疫苗(在HEK细胞中产生的RBD-HB缀合的和Addavax佐剂化的)与组合的流感-COVID-19疫苗(Addavax佐剂化的)的免疫原性

[0302] 制备了三种疫苗制剂:

[0303] 1.商业流感疫苗3μg/mL(分裂型)Addavax佐剂化的(20μL/mL)

[0304] 2.COVID-19疫苗(RBD-HBs缀合的,在HEK细胞中产生的)3μg/mL Addavax佐剂化的(20μL/mL)

[0305] 3.Flu-Covid-19联合疫苗(每种成分3μg/ml)Addavax佐剂化的(20μl/

[0306] 给三组各5只Balb/c小鼠接种0.5ml上述每种疫苗(第0天)。在第0天和第14天从小鼠采集血清样品。

[0307] 通过ELISA针对SARS-CoV-2刺突蛋白(COVID-19抗原)的受体结合结构域(RBD)以及针对流感病毒的H1N1、H3N2和B抗原测量了抗体滴度。针对流感抗原的抗体滴度显示于表1中。针对SARS-CoV-2刺突蛋白的抗体滴度显示于表2中。所有疫苗均引发强烈的抗体应答。使用含有流感-COVID-19联合疫苗的佐剂能够引发针对流感和SARS-CoV-2刺突蛋白的强烈抗体应答,但没有成分抑制的迹象。

[0308] 表1:针对流感抗原的抗体滴度

[0309]

疫苗组 (5 Balb/c 小鼠/组)	针对流感抗原的 ELISA 抗体滴度
PBS 对照	0
COVID-19 第 0 天	0
COVID-19 第 14 天	0
Flu H1N1 第 0 天	0
Flu H1N1 第 14 天	67.1
Flu H3N2 第 0 天	0
Flu H3N2 第 14 天	43.1
Flu B 第 0 天	0
Flu B 第 14 天	40.5

[0310]

COVID-19 + Flu H1N1 第 0 天	0
COVID-19 + Flu H1N1 第 14 天	69.3
COVID-19 + Flu H3N2 第 0 天	0
COVID-19 + Flu H3N2 第 14 天	50.3
COVID-19 + Flu B 第 0 天	0
COVID-19 + Flu B 第 14 天	39.4

[0311] 表2:针对SARS-CoV-2刺突蛋白的抗体滴度

[0312]

疫苗组 （5 Balb/c 小鼠/组）	针对 SARS-CoV-2 刺突蛋白的 ELISA 抗体滴度
PBS 对照	0
COVID-19 第 0 天	0
COVID-19 第 14 天	3.2
Flu H1N1 第 0 天	0
Flu H1N1 第 14 天	0
Flu H3N2 第 0 天	0
Flu H3N2 第 14 天	0
Flu B 第 0 天	0
Flu B 第 14 天	0
COVID-19 + Flu H1N1 第 0 天	0
COVID-19 + Flu H1N1 第 14 天	3.5
COVID-19 + Flu H3N2 第 0 天	0
COVID-19 + Flu H3N2 第 14 天	3.6
COVID-19 + Flu B 第 0 天	0
COVID-19 + Flu B 第 14 天	3.4

[0313] 实施例2:比较单独的市售流感疫苗(VaxiGrip)和单独的COVID-19疫苗(与HBSAg缀合的全长刺突蛋白)与组合的流感-COVID-19疫苗的免疫原性

[0314] 在HEK细胞中重组表达HBSAg和全长SARS-CoV-2刺突蛋白(具有$(EAAAK)_3$接头)的融合蛋白。在两个独立的实验中进行重组表达,合并来自5个克隆(实验1)和4个克隆(实验2)的培养基并评估融合蛋白表达,如图2所示。

[0315] 将来自5个克隆(实验1)的合并培养基命名为HBsAg-$(EAAAK)_3$-COV-SD8-SA01-02-

01(5x)HBsAg。将来自4个克隆(实验2)的合并培养基命名为HBsAg-(EAAAK)$_3$-CoV-S D8-SA01-01-01(4x)HBsAg。

[0316] 通过Bradford测定法测定两种融合蛋白池的总蛋白含量,并在100ml的总体积中调节至1mg/ml。

[0317] 用单独的HBsAg-(EAAAK)$_3$-CoV-S D8-SA01-02-01(5x)HBSAg或HBsAg-(EAAAK)$_3$-CoV-S D8-SA01-01-01(4x)HBSAg或与Vaxigrip流感疫苗组合免疫Balb/c小鼠。COVID-19/'flu联合疫苗在不用佐剂、用ALU-280佐剂或Adda-Vax佐剂情况下给药,如下表3所示。

[0318] 用HBSAg-(EAAAK)$_3$-CoV-S D8-SA01-02-01(5x)或HBSAg-(EAAAK)$_3$-CoV-S D8-SA01-01-01(4x)进行免疫,使用50μg/剂(体积100μl)。使用1.5μg/剂(体积50μl)进行流感疫苗的免疫。在使用任一佐剂的情况下,使用1:1v/v疫苗:佐剂比率(对于佐剂+1疫苗,总计100μL;或对于佐剂+2疫苗,总计150μL)。在第0天免疫小鼠,在第7、14和28天加强免疫。在第14天和第42天处死后获得血清样品。在处死后还分离了免疫小鼠的脾脏用于测试。

[0319] 通过ELISA针对SARS-CoV-2刺突蛋白(COVID-19抗原)的受体结合结构域(RBD)测量了抗体滴度。如下图3所示,在所有实验组(第1、3-9和11组)中,与PBS对照组(第10组)或单独的流感疫苗(第2组)相比,在初免免疫后14天存在可观察的抗-HBSAg-(EAAAK)$_3$-CoV-S IgG滴度。值得注意的是,HBSAg-(EAAAK)$_3$-CoV-S融合蛋白与流感疫苗一起施用时,没有观察到明显的组分抑制,这支持了COVID-19/流感疫苗的潜在临床效用。如图3所示,佐剂,特别是Adda-Vax的使用进一步增加了IgG的产生,特别是对于HBSAg-(EAAAK)$_3$-CoV-S D8-SA01-02-01(5x)和HBSAg-(EAAAK)$_3$-CoV-SD8-SA01-01-01(4x)与Vaxigrip的组合。

[0320] 将使用ELISA定量的抗COVID刺突蛋白IgG(单独或与Vaxigrip组合)的滴度与针对仅含SARS-CoV-2刺突蛋白的受体结合结构域(RBD)的类似融合蛋白HBSAg-(EAAAK)$_3$-COV-S产生的IgG进行比较。单独的HBSAg-(EAAAK)$_3$-CoV-S的数据显示在图4A中,并与图4B中的HBsAg-(EAAAK)$_3$-CoV-S D8-SA01-02-01(5x)进行比较。使用RBD融合物获得了更高的滴度(图4B)。在引发免疫后42天再次测量抗体滴度。再次,如在第14天,在所有实验组(第1、3-9和11组)中,与PBS对照组(第10组)或单独的流感疫苗(第2组)相比,在初免免疫后14天存在可观察到的抗-HBSAg-(EAAAK)$_3$-CoV-S IgG滴度。值得注意的是,HBSAg-(EAAAK)$_3$-CoV-S融合蛋白与流感疫苗一起施用时,没有观察到明显的组分抑制,这支持了COVID-19/流感联合疫苗的潜在临床效用。实际上,第3组的抗-HBSAg-(EAAAK)$_3$-CoV-S IgG滴度(用HBSAg-(EAAAK)$_3$-CoV-SD8-SA01-02-01(5x)和Vaxigrip免疫的)大于第1组的(用单独的HBSAg-(EAAAK)$_3$-CoV-S D8-SA01-02-01(5x)免疫的)。

[0321] 表3:HBSAg-(EAAAK)$_3$-CoV-S、流感和HBSAg-(EAAAK)$_3$-CoV-S/流感免疫

[0322]

组	动物 N° Balb/c	笼	疫苗	佐剂	注射体积/途径
1	5	A/B	HBSAg-(EAAAK)3 - CoV-S (HEK) D8-SA01-02-01 (5x)	无	50µl (i.p.)
2	5	C/D	流感(VAXIGRIP 0.5 ml)	无	50µl (i.p.)
3	5	E/F	HBSAg-(EAAAK)3 -CoV-S (HEK) D8-SA01-02-01 (5x) + 流感(VAXIGRIP 0.5 ml)	无	100µl (i.p.)
4	5	G/H	HBSAg-(EAAAK)3 -CoV-S (HEK) D8-SA01-02-01 (5x)	Alu-280	100µl (i.p.)
5	5	I/L	HBSAg-(EAAAK)3 -CoV-S (HEK) D8-SA01-02-01 (5x)	Adda-Vax	100µl (i.p.)
6	5	M/N	HBSAg-(EAAAK)3 -CoV-S (HEK) D8-SA01-02-01 (5x) + 流感 (VAXIGRIP 0.5 ml)	Alu-280	150µl (i.p.) (50µl + 50µl + 50µl)
7	5	O/P	HBSAg-(EAAAK)3 -CoV-S (HEK) D8-SA01-02-01 (5x) + 流感(VAXIGRIP 0.5 ml)	Adda-Vax	150µl (i.p.) (50µl + 50µl + 50µl)
8	5	Q/R	HBSAg-CoV-S (HEK) D8-SA01-01-01 (4x)	无	50µl (i.p.)
9	5	S/T	HBSAg-CoV-S (HEK) D8-SA01-01-01 (4x) + 流感(VAXIGRIP 0.5 ml)	无	100µl (i.p.)
10	5	U/V	PBS	无	50µl (i.p.)
11	4	Z	HBSAg-CoV-S (HEK) D8-SA01-01-01 (4x) + 流感(VAXIGRIP 0.5 ml)	Adda-Vax	150µl (i.p.) (50µl + 50µl + 50µl)

[0323] 如图5还显示了佐剂(特别是Adda-Vax)的使用进一步增加了IgG产生,特别是对于单独的HBSAg-(EAAAK)$_3$-CoV-S D8-SA01-02-01(5x),或其与Vaxigrip的组合。

[0324] 将使用ELISA定量的抗COVID刺突蛋白IgG(单独或与Vaxigrip组合)的滴度与针对

仅含SARS-CoV-2刺突蛋白的受体结合结构域(RBD)的类似融合蛋白HBSAg-(EAAAK)$_3$-CoV-S产生的IgG进行比较。在第42天针对单独的HBSAg-(EAAAK)$_3$-CoV-S的数据显示在图5A中,并与图5B中的HBSAg-(EAAAK)$_3$-CoV-S D8-SA01-02-01(5x)和HBSAg-(EAAAK)$_3$-CoV-SD8-SA01-01-01(4x)进行比较。使用RBD融合物获得了最高滴度(图5B),然而,用Adda-Vax配制时,用HBSAg-(EAAAK)$_3$-CoV-S D8-SA01-02-01(4x)与Vaxigrip组合维持了高滴度。

[0325] 这些实验证明包含SARS-CoV-2刺突蛋白融合体的疫苗组合物可以成功地与流感病毒疫苗组合,没有预期的在组合疫苗产品生产中常见的疫苗组分抑制问题。因此,使用所述联合疫苗计划了中和测定。

[0326] 实施例3:比较单独的市售流感疫苗(Vaxigrip)和单独的COVID-19疫苗(与HBSAg缀合的全长刺突蛋白)与流感-Covid-19联合疫苗的中和试验

[0327] 本发明的SARS-CoV-2融合蛋白、'流感疫苗和COVID-19-'流感联合疫苗产生针对其各自抗体的中和抗体的能力可以使用基于致细胞病变效应(MN-CPE)的微中和试验进行测试。

[0328] 用0.5ml每种上述疫苗接种5只Balb/c小鼠的组(第0天)。在第0、14和42天从小鼠采集血清样品。

[0329] 1.商业流感疫苗(例如VaxiGrip)

[0330] 2.COVID-19疫苗(例如,HBSAg-(EAAAK)$_3$-CoV-S)

[0331] 3.联合Flu-Covid-19疫苗

[0332] 这些疫苗可以在用或未用佐剂(例如Addavax)的情况下重复。

[0333] 将Vero E6细胞接种在96孔板中并培养以实现亚汇合。

[0334] 使用标准滴定测定法计算SARS-CoV-2的滴度,并制备SARS-CoV-2的10倍系列稀释液(log10)。或者,可以进行3.16倍连续稀释(0.5log10)。

[0335] 将系列稀释的SARS-CoV-2施加于96孔板中的汇合Vero细胞。一列板未用SARS-CoV-2处理,作为细胞对照。此外,含有已知SARS-CoV-2特异性中和抗体的样品可用作阳性对照,人或动物耗尽的样品可用作阴性对照(例如人血清减IgA/IgM/IgG)。

[0336] 加入SARS-CoV-2后,将板在37℃、5%CO_2下孵育3天(孵育时间可以根据SARS-CoV-2株和变体而变化)。孵育后,在倒置显微镜下观察平板,并将孔评分为SARS-CoV-2阳性(即,观察到CPE)或SARS-CoV-2阴性(即,细胞存活且没有CPE)。

[0337] 一旦计算出50%组织培养感染剂量(TCID50),就可以进行MN-CPE测定。

[0338] 对于MN-CPE,如前所述培养Vero E6细胞并接种在96孔板中。将来自接种小鼠的血清样品在56±1℃下热处理30分钟±10分钟。首先以1:10系列稀释来自经处理小鼠的血清样品。然后在板的各行上进行2倍系列稀释。将所需的病毒滴度(一个板用于SARS-CoV-2,一个板用于流感)加入板的每个孔中,然后将板在37℃±1℃,5±1%CO_2下孵育1小时。然后将病毒-血清混合物施加于亚汇合的预培养的Vero E6细胞,并将板在37°±1C,5%±1CO_2下孵育3天(孵育时间可根据SARS-CoV-2毒株和变体而变化)。

[0339] 微量中和滴度(MNt)是保护至少50%的细胞免于CPE的最高样品稀释度的倒数。如果没有观察到中和,则假设MNt<10,这低于检测下限。

[0340] 来自用HBSAg-(EAAAK)$_3$-Cov-S处理的小鼠的血清表现出Vero细胞中有效的中和和CPE的抑制。类似地,用流感疫苗处理的小鼠产生具有对抗流感的中和活性的血清。用

HBSAg-(EAAAK)$_3$-Cov-S和流感疫苗的组合处理小鼠时,实现了对抗SARS-CoV-2和流感两者的中和,证明了使用SARS-CoV-2和流感联合疫苗时没有组分抑制。

[0341] 使用SARS-CoV-2RBD片段疫苗和流感疫苗的组合重复实验。同样,没有观察到组分抑制。

[0342] 序列信息

[0343] SEQ ID NO:1-SARS-CoV-2刺突蛋白氨基酸序列

[0344] MFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIH
VSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYH
KNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSAL
EPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETK
CTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTF
KCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYR
LFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKST
NLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQV
AVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRAR
SVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQ
LNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGD
CLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVL
YENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQI
DRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPA
QEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDS
FKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGL
IAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

[0345] 刺突蛋白的RDB结构域(残基319至529)是带下划线的。

[0346] SEQ ID NO:2-SARS-CoV-2刺突蛋白核酸序列-针对在大肠杆菌中表达优化并含有SacI和NotI单克隆位点。

[0347]
```
GAGCTCatgt ttgtttttct ggttctgctg ccgctggtta gcagccagtg tgttaatctg
accacacgta cccagctgcc tccggcatat accaatagct ttacccgtgg tgtttattat
ccggacaaag tttttcgtag cagcgttctg catagcaccc aggacctgtt tctgccgttt
tttagcaatg ttacctggtt tcatgccatt catgttagcg gcaccaatgg caccaaacgt
tttgataatc cggtgctgcc gtttaatgat ggtgtgtatt ttgcaagcac cgaaaaaagc
aacattattc gcggttggat ttttggtaca accctggata gcaaaaccca gagcctgctg
attgttaata atgccaccaa tgtggtgatc aaagtgtgcg aatttcagtt ttgcaatgat
ccgtttctgg gcgtgtatta ccacaaaaat aacaagagct ggatggaaag cgaatttcgt
```

gtttatagca gcgccaataa ttgcaccttt gaatatgtta gccagccgtt tctgatggat

ctggaaggta aacagggtaa ctttaaaaac ctgcgcgagt tcgtgttcaa aaacatcgat

ggttacttca aaatctatag caaacacacc ccgattaatc tggttcgtga tctgccgcag

ggttttagcg cactggaacc gctggttgat ctgccaattg gtattaacat tacccgtttt

cagaccctgc tggcactgca tcgtagctat ctgacaccgg gtgatagcag cagcggttgg

accgcaggcg cagcagcata ttatgttggt tatctgcagc ctcgtacctt tctgctgaaa

tataacgaaa atggcacaat taccgatgcc gttgattgtg ccctggatcc gctgagcgaa

accaaatgta ccctgaaaag ctttaccgtt gagaaaggta tttatcagac cagcaatttt

cgtgtgcagc cgaccgaaag cattgttcgt tttccgaata tcaccaatct gtgtccgttt

ggcgaagttt ttaatgcaac ccgttttgcc agcgtttatg catggaatcg taaacgtatt

agcaattgcg ttgccgatta tagcgttctg tataatagcg caagcttcag cacctttaaa

tgctatggtg ttagcccgac caaactgaat gatctgtgtt ttaccaatgt gtatgccgat

agctttgtga ttcgtggtga tgaagttcgt cagattgcac cgggtcagac cggtaaaatt

gcagattata actataaact gccggatgat tttacgggtt gtgttattgc ctggaatagc

aataatctgg acagcaaagt tggtggcaac tataactatc tgtatcgcct gtttcgtaag

[0348]

agcaatctga aaccgtttga acgtgatatt agcaccgaga tttatcaggc aggtagcacc

ccgtgtaatg gtgttgaagg ttttaattgc tattttccgc tgcagagcta tggttttcag

ccgacaaatg gtgtgggtta tcagccgtat cgtgttgttg ttctgtcatt tgaactgctg

catgcaccgg caaccgtttg tggtccgaaa aaaagtacca atctggtgaa aaataagtgc

gtgaacttta actttaatgg tctgaccggc accggtgttc tgaccgaaag taacaaaaaa

ttcctgccgt ttcagcagtt tggccgtgat attgcagata ccaccgatgc agttcgcgat

ccgcagacac tggaaattct ggatattacc ccgtgcagct ttggtggtgt ttcagttatt

acaccgggta caaataccag caatcaggtt gcagttctgt atcaggatgt taattgtacc

gaagttccgg ttgcaattca tgcagatcag ctgaccccga cctggcgtgt gtatagcacc

ggtagcaatg tgtttcagac acgtgcaggt tgtctgattg gtgcagaaca tgtgaataat

agctatgaat gcgatattcc gattggtgcg ggtatttgtg ccagctatca gacccagacc

aatagtccgc gtcgtgcacg tagcgttgca agccagagca ttattgccta taccatgagc

ctgggtgcag aaaatagcgt tgcctatagt aataacagca ttgccattcc gaccaacttt

accattagcg ttaccaccga aattctgccg gttagcatga ccaaaaccag cgttgattgc

accatgtata tttgtggtga tagtaccgaa tgtagcaatc tgctgctgca gtatggtagc

ttttgcaccc agctgaatcg tgcactgacc ggtattgcag ttgaacagga taaaaacacg

caagaagttt ttgcacaggt caagcagatc tataaaaccc ctccgattaa agattttggc

ggtttcaatt ttagccagat cctgccggat ccgagcaaac cgagtaaacg tagctttatt

gaagatctgc tgttcaacaa agtgaccctg gcagatgcag gttttatcaa acagtatggt

gattgcctgg gcgatattgc cgcacgtgat ctgatttgtg cacagaaatt taacggcctg

accgttctgc ctccgctgct gaccgatgaa atgattgcac agtataccag cgcactgctg

gcaggcacca ttaccagtgg ttggaccttt ggtgccggtg ccgcactgca gattccgttt

gcaatgcaga tggcatatcg ttttaatggt attggtgtta cccagaacgt gctgtatgaa

aaccagaaac tgattgccaa ccagtttaat agcgccattg gcaaaattca ggatagcctg

agcagcaccg caagtgcact gggtaaactg caggacgttg ttaatcagaa tgcacaggca

ctgaataccc tggttaaaca gctgagcagt aattttggtg caatttcaag cgtgctgaac

gatattctga gccgtctgga taaagttgaa gcagaagttc agattgatcg tctgattacc

ggtcgtctgc aaagcctgca gacctatgtg acccagcagc tgattcgcgc agcagaaatt

[0349]

cgtgcaagcg caaatctggc agccaccaaa atgagcgaat gtgttctggg tcagagcaaa

cgtgttgatt tttgcggcaa aggttatcac ctgatgagct ttccgcagag cgcaccgcat

ggtgttgtgt ttctgcatgt tacctatgtt ccggcacaag aaaaaaactt tacaaccgct

ccggcaattt gccatgatgg taaagcacat tttccgcgtg aaggtgtttt tgttagtaat

ggcacccatt ggtttgttac acagcgcaac ttttatgaac cgcagattat tacaaccgac

aacacctttg ttagcggtaa ctgtgatgtt gtgattggca ttgtgaataa caccgtttat

gatccactgc agccggaact ggatagcttt aaagaagaac tggacaaata tttcaaaaac

cacaccagtc cggatgttga tctgggtgat atttcaggta ttaatgccag cgtggtgaac

atccagaaag aaattgatcg cctgaatgaa gtggccaaaa atctgaatga aagcctgatt

gatctgcaag aactggggaa atatgagcag tatatcaaat ggccgtggta tatttggctg

ggttttattg caggcctgat tgcaattgtt atggtgacca ttatgctgtg ttgtatgacc

agctgttgta gctgtctgaa aggttgttgc agctgcggta gctgttgcaa atttgatgaa

gatgatagcg aaccggtgct gaaaggtgtt aaactgcatt atacctaatg aGCGGCCGC

[0350] 5’SacI单克隆位点是带单下划线的

[0351] 3’NotI单克隆位点是带短下划线的

[0352] ATG起始密码子是粗体和斜体的

[0353] SEQ ID NO:2的核酸序列翻译得到SEQ ID NO:1的天然SARS-CoV-2刺突蛋白。

[0354] SEQ ID NO:3-编码融合蛋白HEV-SARS-CoV-2刺突蛋白的核酸-针对在大肠杆菌中

表达优化并含有SacI和NotI单克隆位点。

gagctcATGA TTGCACTGAC CCTGTTTAAT CTGGCAGATA CCCTGTTAGG TGGTCTGCCG

ACCGAACTGA TTAGCAGTGC CGGTGGTCAG CTGTTTTATA GCCGTCCGGT TGTTAGCGCA

AATGGTGAAC CGACCGTTAA ACTGTATACC AGCGTTGAAA ATGCACAGCA GGATAAAGGT

ATTGCAATTC CGCATGATAT TGATCTGGGT GAAAGCCGTG TTGTGATTCA GGATTATGAT

AATCAGCATG AACAGGATCG TCCGACACCG AGTCCGGCAC CGAGCCGTCC GTTTAGCGTT

CTGCGTGCAA ATGATGTTCT GTGGCTGAGC CTGACCGCAG CAGAATATGA TCAGAGCACC

TATGGTAGCA GCACCGGTCC GGTTTATGTT AGCGATAGCG TTACCCTGGT TAATGTTGCA

ACCGGTGCAC AGGCAGTTGC ACGTAGCCTG GATTGGACCA AAGTGACCCT GGATGGTCGT

CCGCTGAGCA CCATTCAGCA GTATAGCAAA ACCTTTTTTG TTCTGCCGCT GCGTGGTAAA

CTGAGCTTTT GGGAAGCAGG CACCACCAAA GCAGGTTATC CGTATAACTA TAATACCACC

GCAAGCGATC AGCTGCTGGT TGAAAACGCA GCAGGTCATC GTGTTGCAAT TAGCACCTAT

ACCACCAGTT TAGGTGCAGG TCCGGTTAGC ATTAGCGCAG TTGCAGTTCT GGCACCGCAT

[0355]

AGCGCAtttg tttttctggt tctgctgccg ctggttagca gccagtgtgt taatctgacc

acacgtaccc agctgcctcc ggcatatacc aatagcttta cccgtggtgt ttattatccg

gacaaagttt ttcgtagcag cgttctgcat agcacccagg acctgtttct gccgtttttt

agcaatgtta cctggtttca tgccattcat gttagcggca ccaatggcac caaacgtttt

gataatccgg tgctgccgtt taatgatggt gtgtattttg caagcaccga aaaaagcaac

attattcgcg gttggatttt tggtacaacc ctggatagca aaacccagag cctgctgatt

gttaataatg ccaccaatgt ggtgatcaaa gtgtgcgaat ttcagttttg caatgatccg

tttctgggcg tgtattacca caaaaataac aagagctgga tggaaagcga atttcgtgtt

tatagcagcg ccaataattg cacctttgaa tatgttagcc agccgtttct gatggatctg

gaaggtaaac agggtaactt taaaaacctg cgcgagttcg tgttcaaaaa catcgatggt

tacttcaaaa tctatagcaa acacaccccg attaatctgg ttcgtgatct gccgcagggt

tttagcgcac tggaaccgct ggttgatctg ccaattggta ttaacattac ccgttttcag

[0356]

accctgctgg cactgcatcg tagctatctg acaccgggtg atagcagcag cggttggacc

gcaggcgcag cagcatatta tgttggttat ctgcagcctc gtacctttct gctgaaatat

aacgaaaatg gcacaattac cgatgccgtt gattgtgccc tggatccgct gagcgaaacc

aaatgtaccc tgaaaagctt taccgttgag aaaggtattt atcagaccag caattttcgt

gtgcagccga ccgaaagcat tgttcgtttt ccgaatatca ccaatctgtg tccgtttggc

gaagttttta atgcaacccg ttttgccagc gtttatgcat ggaatcgtaa acgtattagc

aattgcgttg ccgattatag cgttctgtat aatagcgcaa gcttcagcac ctttaaatgc

tatggtgtta gcccgaccaa actgaatgat ctgtgtttta ccaatgtgta tgccgatagc

tttgtgattc gtggtgatga agttcgtcag attgcaccgg gtcagaccgg taaaattgca

gattataact ataaactgcc ggatgatttt acgggttgtg ttattgcctg gaatagcaat

aatctggaca gcaaagttgg tggcaactat aactatctgt atcgcctgtt tcgtaagagc

aatctgaaac cgtttgaacg tgatattagc accgagattt atcaggcagg tagcaccccg

tgtaatggtg ttgaaggttt taattgctat tttccgctgc agagctatgg ttttcagccg

acaaatggtg tgggttatca gccgtatcgt gttgttgttc tgtcatttga actgctgcat

gcaccggcaa ccgtttgtgg tccgaaaaaa agtaccaatc tggtgaaaaa taagtgcgtg

aactttaact ttaatggtct gaccggcacc ggtgttctga ccgaaagtaa caaaaaattc

ctgccgtttc agcagtttgg ccgtgatatt gcagatacca ccgatgcagt tcgcgatccg

cagacactgg aaattctgga tattaccccg tgcagctttg gtggtgtttc agttattaca

ccgggtacaa ataccagcaa tcaggttgca gttctgtatc aggatgttaa ttgtaccgaa

gttccggttg caattcatgc agatcagctg accccgacct ggcgtgtgta tagcaccggt

agcaatgtgt ttcagacacg tgcaggttgt ctgattggtg cagaacatgt gaataatagc

tatgaatgcg atattccgat tggtgcgggt atttgtgcca gctatcagac ccagaccaat

agtccgcgtc gtgcacgtag cgttgcaagc cagagcatta ttgcctatac catgagcctg

ggtgcagaaa atagcgttgc ctatagtaat aacagcattg ccattccgac caactttacc

attagcgtta ccaccgaaat tctgccggtt agcatgacca aaaccagcgt tgattgcacc

atgtatattt gtggtgatag taccgaatgt agcaatctgc tgctgcagta tggtagcttt

tgcacccagc tgaatcgtgc actgaccggt attgcagttg aacaggataa aaacacgcaa

gaagtttttg cacaggtcaa gcagatctat aaaaccccctc cgattaaaga ttttggcggt

ttcaatttta gccagatcct gccggatccg agcaaaccga gtaaacgtag ctttattgaa

gatctgctgt tcaacaaagt gaccctggca gatgcaggtt ttatcaaaca gtatggtgat

tgcctgggcg atattgccgc acgtgatctg atttgtgcac agaaatttaa cggcctgacc

gttctgcctc cgctgctgac cgatgaaatg attgcacagt ataccagcgc actgctggca

ggcaccatta ccagtggttg gacctttggt gccggtgccg cactgcagat tccgtttgca

atgcagatgg catatcgttt taatggtatt ggtgttaccc agaacgtgct gtatgaaaac

cagaaactga ttgccaacca gtttaatagc gccattggca aaattcagga tagcctgagc

agcaccgcaa gtgcactggg taaactgcag gacgttgtta atcagaatgc acaggcactg

aataccctgg ttaaacagct gagcagtaat tttggtgcaa tttcaagcgt gctgaacgat

attctgagcc gtctggataa agttgaagca gaagttcaga ttgatcgtct gattaccggt

cgtctgcaaa gcctgcagac ctatgtgacc cagcagctga ttcgcgcagc agaaattcgt

gcaagcgcaa atctggcagc caccaaaatg agcgaatgtg ttctgggtca gagcaaacgt

gttgattttt gcggcaaagg ttatcacctg atgagctttc cgcagagcgc accgcatggt

[0357]

gttgtgtttc tgcatgttac ctatgttccg gcacaagaaa aaaactttac aaccgctccg

gcaatttgcc atgatggtaa agcacatttt ccgcgtgaag gtgtttttgt tagtaatggc

acccattggt ttgttacaca gcgcaacttt tatgaaccgc agattattac aaccgacaac

acctttgtta gcggtaactg tgatgttgtg attggcattg tgaataacac cgtttatgat

ccactgcagc cggaactgga tagctttaaa gaagaactgg acaaatattt caaaaaccac

accagtccgg atgttgatct gggtgatatt tcaggtatta atgccagcgt ggtgaacatc

cagaaagaaa ttgatcgcct gaatgaagtg gccaaaaatc tgaatgaaag cctgattgat

ctgcaagaac tggggaaata tgagcagtat atcaaatggc cgtggtatat ttggctgggt

tttattgcag gcctgattgc aattgttatg gtgaccatta tgctgtgttg tatgaccagc

tgttgtagct gtctgaaagg ttgttgcagc tgcggtagct gttgcaaatt tgatgaagat

gatagcgaac cggtgctgaa aggtgttaaa ctgcattata cctaatgagc ggccgc

[0358] 5’SacI单克隆位点是带单下划线的

[0359] HEV(p239片段)序列以大写字母显示

[0360] SARS-CoV-2刺突蛋白编码序列以小写字母显示

[0361] 3’NotI单克隆位点是带短下划线的

[0362] SEQ ID NO:4-SARS-CoV-2刺突蛋白核酸序列-针对在Kamagataella pastoris中的表达优化并含有BstB1和NotI单克隆位点。

TTCGAA*acga tg*ttcgtgtt cttggtcctg ttgccattgg tttcttccca gtgtgttaac

ctgaccacta gaactcaatt gcctccagcc tacaccaatt ccttcaccag aggtgtttac

tacccagaca aggtgttcag atcttccgtc ttgcactcca ctcaggactt gttcttgcca

ttcttctcca acgttacctg gttccacgct attcacgttt ccggaactaa cggtactaag

agattcgaca acccagtcct gccattcaac gatggtgtct acttcgcttc taccgagaag

tccaacatca tcagaggttg gatcttcggt actaccctgg actctaagac tcagtccttg

ctgatcgtta acaacgccac caacgttgtc atcaaggttt gcgagttcca gttctgcaac

gacccattct tgggtgtgta ctaccacaag aacaacaagt cttggatgga atccgagttc

agagtttact cctccgccaa caactgtacc ttcgagtacg tttcccagcc attcttgatg

gacttggagg gtaagcaggg taacttcaag aacctgagag agttcgtttt caagaacatc

gacggttact tcaagatcta ctccaagcac accccaatca acctggttag agatttgcca

caaggtttct ccgctttgga gcctttggtt gacttgccaa tcggtatcaa catcaccaga

ttccagacct tgttggcctt gcacagatcc tacttgactc caggtgattc ttcttccggt

tggactgctg gtgctgctgc ttactatgtt ggttacttgc agccaagaac cttcctgctg

[0363] aagtacaacg agaacggaac tatcactgac gctgttgact gtgctttgga cccattgtct

gagactaagt gcaccttgaa gtccttcacc gttgagaagg gtatctacca gacctccaac

ttcagagttc agccaactga gtccatcgtc agattcccaa acatcactaa cttgtgccca

ttcggtgagg tgttcaacgc tactagattc gcttctgttt acgcctggaa cagaaagaga

atctccaact gcgttgctga ctactccgtc ttgtacaact ctgcttcatt ctccaccttc

aagtgctacg gtgtttcccc aactaagttg aacgacctgt gtttcactaa cgtctacgcc

gactccttcg ttattagagg tgacgaggtt agacagatcg ctccaggtca aactggtaag

atcgctgact acaactacaa gctgccagac gacttcaccg gttgtgttat tgcttggaac

tccaacaacc tggactccaa ggttggtggt aactacaatt acctgtaccg tctgttcaga

aagtccaact tgaagccatt cgagagagac atctccaccg agatctacca agctggttct

actccatgta acggtgtcga gggtttcaac tgctacttcc cattgcaatc ctacggtttc

caacctacca acggtgttgg ataccagcca tacagagttg tcgttttgtc cttcgagttg

ttgcacgctc cagctactgt ttgtggtcca aagaagtcca ccaacttggt caagaacaaa

tgcgtcaact ttaacttcaa cggcctgacc ggtactggtg ttttgactga atccaacaag

aagttcctgc ctttccagca gttcggtaga gacattgctg acactactga cgccgttaga

[0364]

gatccacaga ctttggagat cttggacatc accccatgtt ccttcggtgg tgtttccgtt

attacccctg gaactaacac ctccaatcag gtcgctgtct tgtaccagga cgttaactgt

actgaggttc cagttgctat ccacgctgac caattgactc caacttggag agtctactcc

accggttcca acgttttcca aactagagcc ggttgtttga tcggtgctga acacgtcaac

aactcctacg agtgtgacat tccaattggt gctggtatct gtgcctccta ccaaactcaa

actaactccc caagaagggc tagatccgtt gcttcccaat ccattatcgc ttacaccatg

tctttgggtg ccgagaactc tgttgcctac tctaacaact ctatcgctat ccctaccaac

ttcaccatct ccgttaccac tgagatcttg ccagtctcca tgaccaagac ttccgttgac

tgtaccatgt acatctgtgg tgactccact gagtgttcca acttgttgct gcaatacggt

tccttctgca cccagttgaa cagagctttg actggtattg ctgtcgagca agacaagaac

actcaagagg ttttcgccca ggtgaagcag atctacaaga ctccacctat taaggacttc

ggtggcttca acttctccca gattttgcca gatccatcta agccctccaa gagatccttc

attgaggacc tgctgttcaa caaggttact ttggctgacg ccggtttcat caagcagtac

ggtgattgct tgggtgacat tgcagctaga gacttgatct gtgcccagaa gttcaacggt

ttgaccgttt tgccaccttt gttgaccgac gagatgatcg ctcagtacac ttctgctttg

ttggccggta ctatcacttc tggttggaca tttggagctg gtgccgcatt gcaaattcca

ttcgctatgc aaatggccta cagattcaac ggtatcggtg ttacccagaa cgtcctgtac

gagaaccaga agcttatcgc caaccagttc aactccgcta tcggtaagat tcaggactcc

ttgtcctcta ctgcttctgc cttgggaaag ttgcaggatg ttgttaacca gaatgcccag

gctttgaaca ccctggttaa gcaactgtcc tctaacttcg gtgctatctc ctccgttttg

aacgacatct tgtcccgttt ggacaaggtt gaggctgagg ttcagatcga cagattgatc

actggtagat tgcagtccct gcagacttac gttactcagc agttgattag agctgccgag

attagagcct ctgctaactt ggctgctact aagatgtccg agtgtgtttt gggtcagtcc

aagagagttg acttctgcgg taagggttac cacctgatgt ctttcccaca atctgctcca

cacggtgtcg ttttcttgca cgttacttac gttccagctc aagagaagaa cttcactact

gctccagcca tttgtcacga tggtaaggct cactttcctc gtgagggtgt tttcgtttcc

aacggtactc actggttcgt cacccagaga aacttttacg agccacagat catcaccacc

gacaacactt tcgtttctgg taactgtgac gtcgtcatcg gtatcgtgaa caacactgtc

tacgatccat tgcagccaga attggactcc ttcaaagagg aactggacaa gtactttaag

aaccacactt ccccagacgt tgacctgggt gatatttccg gtattaacgc ctccgttgtc

aacatccaaa aagagatcga ccgtttgaac gaggtcgcca agaacttgaa cgagtccttg

attgacttgc aagagctggg caagtacgag cagtacatta agtggccatg gtacatttgg

[0365] ctgggtttca ttgctggttt gatcgccatc gttatggtca ccatcatgtt gtgctgtatg

acctcctgtt gctcctgttt gaagggttgt tgttcctgcg gttcctgttg taagttcgac

gaagatgact ccgagccagt cttgaagggt gttaagttgc actacactta a**GCGGCCGC**

[0366] 5'BstBI单克隆位点是带单下划线的

[0367] 3'NotI单克隆位点是带短下划线的

[0368] 在5'SacI之后紧接着是ACG密码子(为编码序列与紧接着ACG之后的ATG起始密码子在框内所需)。这两个密码子以粗体和斜体显示。

[0369] SEQ ID NO:4的核酸序列翻译得到SEQ ID NO:1的天然SARS-CoV-2刺突蛋白。

[0370] SEQ ID NO:5-编码融合蛋白HPV18L1/SARS-CoV-2刺突蛋白的核酸-针对在K.pastoris中表达优化并含有BstB1和NotI单克隆位点。

TTCGAA*acgatg*gctctttggagaccatccgacaacactgtttacttgcc

accaccatccgttgctagagttgttaacactgacgactacgttactagaa

cttccatcttctaccacgctggttcttccagattgttgactgttggtaac

ccatacttcagagttccagctggaggtggtaacaagcaagacatcccaaa

ggtttccgcttaccagtacagagttttcagagttcagttgccagacccaa

acaagtttggattgccagacacttccatctacaacccagagactcagaga

cttgtttgggcttgtgctggtgttgaaatcggtagaggacagccattggg

[0371] tgttggtttgtctggtcacccattctacaacaagttggacgacactgaat

cttctcacgctgctacttctaacgtttccgaggatgttagagacaacgtt

tccgttgactacaagcagactcagttgtgtatcttgggttgtgctccagc

tattggtgaacattgggctaagggtactgcttgtaagtccagaccattgt

ctcagggagattgtccaccattggagttgaagaacactgttttggaggac

ggtgatatggttgatactggttacggtgctatggacttctctactttgca

ggacactaagtgtgaagttccattggacatctgtcagtccatctgtaagt

acccagactacttgcaaatgtccgctgatccatacggtgactctatgttc

ttctgtttgagaagagagcagttgttcgctagacacttctggaacagagc

tggtactatgggtgacactgttccacaatccttgtacatcaagggtactg

gaatgagagcttctcctggttcttgtgtttactctccatctccatccggt

tccattgttacttccgactcccagttgttcaacaagccatactggttgca

taaggctcaaggtcacaacaacggtgtttgttggcacaaccagttgttcg

ttactgttgttgacactactagatccactaacttgactatctgtgcttcc

actcaatctccagttccaggacaatacgacgctactaagttcaagcagta

ctccagacacgttgaagagtacgacttgcagttcatcttccagttgtgta

ctatcactttgactgctgatgttatgtcctacatccactctatgaactcc

tccattttggaggattggaacttcggtgttccaccaccaccaactacttc

attggttgacacttacagattcgttcagtccgttgctatcacttgtcaaa

aggacgctgctccagctgaaaacaaggacccatacgacaagttgaagttc

tggaacgttgacttgaaagagaagttctccttggacttggaccaataccc

attgggtagaaagtttttggttcaggctggattgagaagaaagccaacta

tcggtccaagaaagagatcagctccatccgctactacttcatccaagcca

[0372]

gctaagagagttagagttagagctagaaagtTCGTGTTCTTGGTCCTGTT

GCCATTGGTTTCTTCCCAGTGTGTTAACCTGACCACTAGAACTCAATTGC

CTCCAGCCTACACCAATTCCTTCACCAGAGGTGTTTACTACCCAGACAAG

GTGTTCAGATCTTCCGTCTTGCACTCCACTCAGGACTTGTTCTTGCCATT

CTTCTCCAACGTTACCTGGTTCCACGCTATTCACGTTTCCGGAACTAACG

GTACTAAGAGATTCGACAACCCAGTCCTGCCATTCAACGATGGTGTCTAC

TTCGCTTCTACCGAGAAGTCCAACATCATCAGAGGTTGGATCTTCGGTAC

TACCCTGGACTCTAAGACTCAGTCCTTGCTGATCGTTAACAACGCCACCA

ACGTTGTCATCAAGGTTTGCGAGTTCCAGTTCTGCAACGACCCATTCTTG

GGTGTGTACTACCACAAGAACAACAAGTCTTGGATGGAATCCGAGTTCAG

AGTTTACTCCTCCGCCAACAACTGTACCTTCGAGTACGTTTCCCAGCCAT

TCTTGATGGACTTGGAGGGTAAGCAGGGTAACTTCAAGAACCTGAGAGAG

TTCGTTTTCAAGAACATCGACGGTTACTTCAAGATCTACTCCAAGCACAC

CCCAATCAACCTGGTTAGAGATTTGCCACAAGGTTTCTCCGCTTTGGAGC

CTTTGGTTGACTTGCCAATCGGTATCAACATCACCAGATTCCAGACCTTG

TTGGCCTTGCACAGATCCTACTTGACTCCAGGTGATTCTTCTTCCGGTTG

GACTGCTGGTGCTGCTGCTTACTATGTTGGTTACTTGCAGCCAAGAACCT

TCCTGCTGAAGTACAACGAGAACGGAACTATCACTGACGCTGTTGACTGT

GCTTTGGACCCATTGTCTGAGACTAAGTGCACCTTGAAGTCCTTCACCGT

TGAGAAGGGTATCTACCAGACCTCCAACTTCAGAGTTCAGCCAACTGAGT

CCATCGTCAGATTCCCAAACATCACTAACTTGTGCCCATTCGGTGAGGTG

TTCAACGCTACTAGATTCGCTTCTGTTTACGCCTGGAACAGAAAGAGAAT

CTCCAACTGCGTTGCTGACTACTCCGTCTTGTACAACTCTGCTTCATTCT

CCACCTTCAAGTGCTACGGTGTTTCCCCAACTAAGTTGAACGACCTGTGT

TTCACTAACGTCTACGCCGACTCCTTCGTTATTAGAGGTGACGAGGTTAG

ACAGATCGCTCCAGGTCAAACTGGTAAGATCGCTGACTACAACTACAAGC

TGCCAGACGACTTCACCGGTTGTGTTATTGCTTGGAACTCCAACAACCTG

GACTCCAAGGTTGGTGGTAACTACAATTACCTGTACCGTCTGTTCAGAAA

GTCCAACTTGAAGCCATTCGAGAGAGACATCTCCACCGAGATCTACCAAG

CTGGTTCTACTCCATGTAACGGTGTCGAGGGTTTCAACTGCTACTTCCCA

[0373]

TTGCAATCCTACGGTTTCCAACCTACCAACGGTGTTGGATACCAGCCATA

CAGAGTTGTCGTTTTGTCCTTCGAGTTGTTGCACGCTCCAGCTACTGTTT

GTGGTCCAAAGAAGTCCACCAACTTGGTCAAGAACAAATGCGTCAACTTT

AACTTCAACGGCCTGACCGGTACTGGTGTTTTGACTGAATCCAACAAGAA

GTTCCTGCCTTTCCAGCAGTTCGGTAGAGACATTGCTGACACTACTGACG

CCGTTAGAGATCCACAGACTTTGGAGATCTTGGACATCACCCCATGTTCC

TTCGGTGGTGTTTCCGTTATTACCCCTGGAACTAACACCTCCAATCAGGT

CGCTGTCTTGTACCAGGACGTTAACTGTACTGAGGTTCCAGTTGCTATCC

ACGCTGACCAATTGACTCCAACTTGGAGAGTCTACTCCACCGGTTCCAAC

GTTTTCCAAACTAGAGCCGGTTGTTTGATCGGTGCTGAACACGTCAACAA

CTCCTACGAGTGTGACATTCCAATTGGTGCTGGTATCTGTGCCTCCTACC

AAACTCAAACTAACTCCCCAAGAAGGGCTAGATCCGTTGCTTCCCAATCC

ATTATCGCTTACACCATGTCTTTGGGTGCCGAGAACTCTGTTGCCTACTC

TAACAACTCTATCGCTATCCCTACCAACTTCACCATCTCCGTTACCACTG

AGATCTTGCCAGTCTCCATGACCAAGACTTCCGTTGACTGTACCATGTAC

ATCTGTGGTGACTCCACTGAGTGTTCCAACTTGTTGCTGCAATACGGTTC

CTTCTGCACCCAGTTGAACAGAGCTTTGACTGGTATTGCTGTCGAGCAAG

ACAAGAACACTCAAGAGGTTTTCGCCCAGGTGAAGCAGATCTACAAGACT

CCACCTATTAAGGACTTCGGTGGCTTCAACTTCTCCCAGATTTTGCCAGA

TCCATCTAAGCCCTCCAAGAGATCCTTCATTGAGGACCTGCTGTTCAACA

AGGTTACTTTGGCTGACGCCGGTTTCATCAAGCAGTACGGTGATTGCTTG

GGTGACATTGCAGCTAGAGACTTGATCTGTGCCCAGAAGTTCAACGGTTT

GACCGTTTTGCCACCTTTGTTGACCGACGAGATGATCGCTCAGTACACTT

CTGCTTTGTTGGCCGGTACTATCACTTCTGGTTGGACATTTGGAGCTGGT

GCCGCATTGCAAATTCCATTCGCTATGCAAATGGCCTACAGATTCAACGG

TATCGGTGTTACCCAGAACGTCCTGTACGAGAACCAGAAGCTTATCGCCA

ACCAGTTCAACTCCGCTATCGGTAAGATTCAGGACTCCTTGTCCTCTACT

GCTTCTGCCTTGGGAAAGTTGCAGGATGTTGTTAACCAGAATGCCCAGGC

TTTGAACACCCTGGTTAAGCAACTGTCCTCTAACTTCGGTGCTATCTCCT

CCGTTTTGAACGACATCTTGTCCCGTTTGGACAAGGTTGAGGCTGAGGTT

[0374]

CAGATCGACAGATTGATCACTGGTAGATTGCAGTCCCTGCAGACTTACGT

TACTCAGCAGTTGATTAGAGCTGCCGAGATTAGAGCCTCTGCTAACTTGG

CTGCTACTAAGATGTCCGAGTGTGTTTTGGGTCAGTCCAAGAGAGTTGAC

TTCTGCGGTAAGGGTTACCACCTGATGTCTTTCCCACAATCTGCTCCACA

CGGTGTCGTTTTCTTGCACGTTACTTACGTTCCAGCTCAAGAGAAGAACT

TCACTACTGCTCCAGCCATTTGTCACGATGGTAAGGCTCACTTTCCTCGT

GAGGGTGTTTTCGTTTCCAACGGTACTCACTGGTTCGTCACCCAGAGAAA

CTTTTACGAGCCACAGATCATCACCACCGACAACACTTTCGTTTCTGGTA

ACTGTGACGTCGTCATCGGTATCGTGAACAACACTGTCTACGATCCATTG

CAGCCAGAATTGGACTCCTTCAAAGAGGAACTGGACAAGTACTTTAAGAA

CCACACTTCCCCAGACGTTGACCTGGGTGATATTTCCGGTATTAACGCCT

CCGTTGTCAACATCCAAAAAGAGATCGACCGTTTGAACGAGGTCGCCAAG

AACTTGAACGAGTCCTTGATTGACTTGCAAGAGCTGGGCAAGTACGAGCA

GTACATTAAGTGGCCATGGTACATTTGGCTGGGTTTCATTGCTGGTTTGA

TCGCCATCGTTATGGTCACCATCATGTTGTGCTGTATGACCTCCTGTTGC

TCCTGTTTGAAGGGTTGTTGTTCCTGCGGTTCCTGTTGTAAGTTCGACGA

[0375] AGATGACTCCGAGCCAGTCTTGAAGGGTGTTAAGTTGCACTACACTTAAG

CGGCCGC

[0376] 5'BstBI单克隆位点是带单下划线的

[0377] HPV18L1序列以小写字母显示

[0378] SARS-CoV-2刺突蛋白编码序列以大写字母显示

[0379] 3'NotI单克隆位点是带短下划线的

[0380] 紧接着5'BstBI是ACG密码子(为编码序列与紧接着ACG的ATG起始密码子在框内所需)。这两个密码子以粗体和斜体显示。

[0381] SEQ ID NO:6-编码融合蛋白HPV16L1/SARS-CoV-2刺突蛋白的核酸-针对在K.pastoris中表达并含有BstB1和NotI单一克隆位点。

TTCGAA*acgatg*tctttgtggttgccatctgaagctactgtttacttgcc

accagttccagtttctaaagttgtttccactgacgaatacgttgctagaa

ctaacatctactaccacgctggtacttctagattgttggctgttggtcat

ccatacttcccaattaagaagccaaacaacaacaagattttggttccaaa

ggtttccggattgcaatacagagttttcagaatccatttgccagatccaa

acaagtttggtttcccagatacttctttctacaacccagacactcaaaga

cttgtttgggcttgtgttggtgttgaagttggtagaggtcaaccattggg

tgttggtatttctggtcacccattgttgaacaagttggacgatactgaaa

[0382] acgcttctgcttacgctgctaacgctggtgttgataacagagaatgtatt

tctatggactacaagcaaactcaattgtgtttgattggttgtaagccacc

aattggtgaacattggggaaagggttctccatgtactaatgttgctgtta

accctggtgattgtccaccattggaattgattaacactgttattcaagac

ggtgatatggttgatactggtttcggtgctatggatttcactactttgca

agctaacaagtctgaagttccattggacatttgtacttccatctgtaagt

acccagactacattaagatggtttctgaaccatacggtgattctttgttc

ttctacttgagaagagaacaaatgtttgttagacacttgttcaacagagc

tggtgctgttggtgaaaacgttccagatgacttgtacattaagggttctg

gttctactgctaacttggcttcttctaactactttccaactccatctggt

tctatggttacttctgacgctcaaattttcaacaagccatactggttgca

aagagcacaaggtcataacaacggtatttgttggggtaaccaattgttcg

ttactgttgttgacactactagatccactaacatgtccttgtgtgctgct

atttctacttctgaaactacttacaagaacactaacttcaaagagtactt

gagacacggagaagaatacgacttgcaattcattttccaattgtgtaaga

ttactttgactgctgacgttatgacttacattcactctatgaactctact

attttggaagattggaacttcggattgcaaccaccaccaggtggtacttt

ggaagatacttacagattcgttacttctcaagctattgcttgtcaaaagc

atactccacctgctccaaaagaagatccattgaagaagtacactttctgg

gaagttaacttgaaagaaaagttctctgctgatttggatcaattcccatt

gggtagaaagtttttgttgcaagctggattgaaggctaaaccaaagttca

ctttgggaaagagaaaggctactccaactacttcttctacttctactact

gctaagagaaagaagagaaaattgtTCGTGTTCTTGGTCCTGTTGCCATT

[0383] GGTTTCTTCCCAGTGTGTTAACCTGACCACTAGAACTCAATTGCCTCCAG

CCTACACCAATTCCTTCACCAGAGGTGTTTACTACCCAGACAAGGTGTTC

AGATCTTCCGTCTTGCACTCCACTCAGGACTTGTTCTTGCCATTCTTCTC

CAACGTTACCTGGTTCCACGCTATTCACGTTTCCGGAACTAACGGTACTA

AGAGATTCGACAACCCAGTCCTGCCATTCAACGATGGTGTCTACTTCGCT

TCTACCGAGAAGTCCAACATCATCAGAGGTTGGATCTTCGGTACTACCCT

GGACTCTAAGACTCAGTCCTTGCTGATCGTTAACAACGCCACCAACGTTG

TCATCAAGGTTTGCGAGTTCCAGTTCTGCAACGACCCATTCTTGGGTGTG

TACTACCACAAGAACAACAAGTCTTGGATGGAATCCGAGTTCAGAGTTTA

CTCCTCCGCCAACAACTGTACCTTCGAGTACGTTTCCCAGCCATTCTTGA

TGGACTTGGAGGGTAAGCAGGGTAACTTCAAGAACCTGAGAGAGTTCGTT

TTCAAGAACATCGACGGTTACTTCAAGATCTACTCCAAGCACACCCCAAT

CAACCTGGTTAGAGATTTGCCACAAGGTTTCTCCGCTTTGGAGCCTTTGG

TTGACTTGCCAATCGGTATCAACATCACCAGATTCCAGACCTTGTTGGCC

TTGCACAGATCCTACTTGACTCCAGGTGATTCTTCTTCCGGTTGGACTGC

TGGTGCTGCTGCTTACTATGTTGGTTACTTGCAGCCAAGAACCTTCCTGC

TGAAGTACAACGAGAACGGAACTATCACTGACGCTGTTGACTGTGCTTTG

GACCCATTGTCTGAGACTAAGTGCACCTTGAAGTCCTTCACCGTTGAGAA

GGGTATCTACCAGACCTCCAACTTCAGAGTTCAGCCAACTGAGTCCATCG

TCAGATTCCCAAACATCACTAACTTGTGCCCATTCGGTGAGGTGTTCAAC

GCTACTAGATTCGCTTCTGTTTACGCCTGGAACAGAAAGAGAATCTCCAA

CTGCGTTGCTGACTACTCCGTCTTGTACAACTCTGCTTCATTCTCCACCT

TCAAGTGCTACGGTGTTTCCCCAACTAAGTTGAACGACCTGTGTTTCACT

AACGTCTACGCCGACTCCTTCGTTATTAGAGGTGACGAGGTTAGACAGAT

CGCTCCAGGTCAAACTGGTAAGATCGCTGACTACAACTACAAGCTGCCAG

ACGACTTCACCGGTTGTGTTATTGCTTGGAACTCCAACAACCTGGACTCC

AAGGTTGGTGGTAACTACAATTACCTGTACCGTCTGTTCAGAAAGTCCAA

CTTGAAGCCATTCGAGAGAGACATCTCCACCGAGATCTACCAAGCTGGTT

CTACTCCATGTAACGGTGTCGAGGGTTTCAACTGCTACTTCCCATTGCAA

TCCTACGGTTTCCAACCTACCAACGGTGTTGGATACCAGCCATACAGAGT

TGTCGTTTTGTCCTTCGAGTTGTTGCACGCTCCAGCTACTGTTTGTGGTC

[0384]

CAAAGAAGTCCACCAACTTGGTCAAGAACAAATGCGTCAACTTTAACTTC

AACGGCCTGACCGGTACTGGTGTTTTGACTGAATCCAACAAGAAGTTCCT

GCCTTTCCAGCAGTTCGGTAGAGACATTGCTGACACTACTGACGCCGTTA

GAGATCCACAGACTTTGGAGATCTTGGACATCACCCCATGTTCCTTCGGT

GGTGTTTCCGTTATTACCCCTGGAACTAACACCTCCAATCAGGTCGCTGT

CTTGTACCAGGACGTTAACTGTACTGAGGTTCCAGTTGCTATCCACGCTG

ACCAATTGACTCCAACTTGGAGAGTCTACTCCACCGGTTCCAACGTTTTC

CAAACTAGAGCCGGTTGTTTGATCGGTGCTGAACACGTCAACAACTCCTA

CGAGTGTGACATTCCAATTGGTGCTGGTATCTGTGCCTCCTACCAAACTC

AAACTAACTCCCCAAGAAGGGCTAGATCCGTTGCTTCCCAATCCATTATC

GCTTACACCATGTCTTTGGGTGCCGAGAACTCTGTTGCCTACTCTAACAA

CTCTATCGCTATCCCTACCAACTTCACCATCTCCGTTACCACTGAGATCT

TGCCAGTCTCCATGACCAAGACTTCCGTTGACTGTACCATGTACATCTGT

GGTGACTCCACTGAGTGTTCCAACTTGTTGCTGCAATACGGTTCCTTCTG

CACCCAGTTGAACAGAGCTTTGACTGGTATTGCTGTCGAGCAAGACAAGA

ACACTCAAGAGGTTTTCGCCCAGGTGAAGCAGATCTACAAGACTCCACCT

ATTAAGGACTTCGGTGGCTTCAACTTCTCCCAGATTTTGCCAGATCCATC

TAAGCCCTCCAAGAGATCCTTCATTGAGGACCTGCTGTTCAACAAGGTTA

CTTTGGCTGACGCCGGTTTCATCAAGCAGTACGGTGATTGCTTGGGTGAC

ATTGCAGCTAGAGACTTGATCTGTGCCCAGAAGTTCAACGGTTTGACCGT

TTTGCCACCTTTGTTGACCGACGAGATGATCGCTCAGTACACTTCTGCTT

TGTTGGCCGGTACTATCACTTCTGGTTGGACATTTGGAGCTGGTGCCGCA

TTGCAAATTCCATTCGCTATGCAAATGGCCTACAGATTCAACGGTATCGG

TGTTACCCAGAACGTCCTGTACGAGAACCAGAAGCTTATCGCCAACCAGT

TCAACTCCGCTATCGGTAAGATTCAGGACTCCTTGTCCTCTACTGCTTCT

GCCTTGGGAAAGTTGCAGGATGTTGTTAACCAGAATGCCCAGGCTTTGAA

CACCCTGGTTAAGCAACTGTCCTCTAACTTCGGTGCTATCTCCTCCGTTT

TGAACGACATCTTGTCCCGTTTGGACAAGGTTGAGGCTGAGGTTCAGATC

GACAGATTGATCACTGGTAGATTGCAGTCCCTGCAGACTTACGTTACTCA

GCAGTTGATTAGAGCTGCCGAGATTAGAGCCTCTGCTAACTTGGCTGCTA

[0385]

CTAAGATGTCCGAGTGTGTTTTGGGTCAGTCCAAGAGAGTTGACTTCTGC

GGTAAGGGTTACCACCTGATGTCTTTCCCACAATCTGCTCCACACGGTGT

CGTTTTCTTGCACGTTACTTACGTTCCAGCTCAAGAGAAGAACTTCACTA

CTGCTCCAGCCATTTGTCACGATGGTAAGGCTCACTTTCCTCGTGAGGGT

GTTTTCGTTTCCAACGGTACTCACTGGTTCGTCACCCAGAGAAACTTTTA

CGAGCCACAGATCATCACCACCGACAACACTTTCGTTTCTGGTAACTGTG

ACGTCGTCATCGGTATCGTGAACAACACTGTCTACGATCCATTGCAGCCA

GAATTGGACTCCTTCAAAGAGGAACTGGACAAGTACTTTAAGAACCACAC

TTCCCCAGACGTTGACCTGGGTGATATTTCCGGTATTAACGCCTCCGTTG

TCAACATCCAAAAAGAGATCGACCGTTTGAACGAGGTCGCCAAGAACTTG

AACGAGTCCTTGATTGACTTGCAAGAGCTGGGCAAGTACGAGCAGTACAT

TAAGTGGCCATGGTACATTTGGCTGGGTTTCATTGCTGGTTTGATCGCCA

TCGTTATGGTCACCATCATGTTGTGCTGTATGACCTCCTGTTGCTCCTGT

TTGAAGGGTTGTTGTTCCTGCGGTTCCTGTTGTAAGTTCGACGAAGATGA

CTCCGAGCCAGTCTTGAAGGGTGTTAAGTTGCACTACACTTAA**GCGGCCG**

[0386] **C**

[0387] 5’BstBI单克隆位点是带单下划线的

[0388] HPV16L1序列以小写字母显示

[0389] SARS-CoV-2刺突蛋白编码序列以大写字母显示

[0390] 3'NotI单克隆位点是带短下划线的

[0391] 紧接着5'BstBI是ACG密码子(为编码序列与紧接着ACG的ATG起始密码子在框内所需)。这两个密码子以粗体和斜体显示。

[0392] SEQ ID NO:7-SARS-CoV-2刺突蛋白核酸序列-针对在人类中表达(293F)优化并含有NheI和NotI单克隆位点。

GCTAGC*gaca tg*ttcgtgtt tctggtgctg ctgcctctgg tgtccagcca gtgtgtgaac

ctgaccacca gaacacagct gcctccagcc tacaccaata gcttcaccag gggcgtgtac

taccccgaca aggtgttcag atctagcgtg ctgcacagca cccaggacct gtttctgccc

ttcttcagca acgtgacctg gttccacgcc atccacgtgt ccggcaccaa tggcaccaag

agattcgaca accccgtgct gcccttcaac gatggggtgt actttgccag caccgagaag

tccaacatca tcagaggctg gatcttcggc accacactgg acagcaagac ccagagcctg

ctgatcgtga acaacgccac caacgtggtc atcaaagtgt gcgagttcca gttctgcaac

gacccattcc tgggagtcta ctaccacaag aacaacaaga gctggatgga aagcgagttc

cgggtgtaca gcagcgccaa caactgcacc ttcgagtacg tgtcccagcc tttcctgatg

[0393] gacctggaag gcaagcaggg caacttcaag aacctgcgcg agttcgtgtt caagaacatc

gacggctact tcaagatcta cagcaagcac acccctatca acctcgtgcg ggatctgcct

cagggctttt ctgctctgga acctctggtg gacctgccta tcggcatcaa catcacccgg

tttcagaccc tgctggccct gcacagatct tacctgacac ctggcgatag cagctctgga

tggacagctg gcgccgctgc ctattatgtg ggctacctgc agcctcggac cttcctgctg

aagtacaacg agaacggcac catcaccgac gccgtggatt gtgctctgga tcccctgagc

gagacaaagt gcaccctgaa gtccttcacc gtggaaaagg gcatctacca gaccagcaac

ttcagagtgc agcccaccga gagcatcgtg cggttcccca atatcaccaa tctgtgcccc

ttcggcgagg tgttcaatgc cacaagattt gccagcgtgt acgcctggaa ccggaagaga

atcagcaact gcgtggccga ctacagcgtg ctgtacaata gcgccagctt cagcaccttc

aagtgctacg gcgtgtcccc taccaagctg aacgacctgt gcttcaccaa tgtgtacgcc

gacagcttcg tgatcagagg cgacgaagtt cggcagatcg ctcctggaca gacaggcaag

atcgccgatt acaactacaa gctgcccgac gacttcaccg gctgcgtgat cgcctggaat

agcaacaacc tggactccaa agtcggcggc aactacaact acctgtaccg gctgttccgg

aagtccaatc tgaagccctt cgagcgggac atctccaccg aaatctatca ggccggcagc

acccccttgta acggcgtgga aggcttcaac tgctacttcc cactgcagtc ctacggcttt

cagcctacca atggcgtggg ctatcagccc tatagagtgg tggtgctgag cttcgaactg

ctgcatgccc ctgctaccgt gtgcggccct aagaagtcta ccaacctggt caagaacaaa

tgcgtgaact tcaacttcaa cggcctgacc ggcacaggcg tgctgacaga gagcaacaag

aagttcctgc ctttccagca gtttggccgg gatatcgccg ataccacaga cgccgttaga

gatccccaga cactggaaat cctggacatc accccatgca gctttggcgg agtgtctgtg

atcacccctg gcaccaatac cagcaatcag gtggccgtgc tgtatcagga cgtgaactgt

acagaggtgc ccgtggccat tcacgccgat caactgacac ccacttggag agtgtactcc

accggctcca acgtgttcca gactagagcc ggatgtctga tcggagccga gcacgtgaac

aatagctacg agtgcgacat ccccatcggc gctggcatct gtgccagcta ccagacacag

[0394]

acaaatagcc ccagacgggc cagaagcgtg gcctctcaga gcatcattgc ctacacaatg

agcctgggcg ccgagaattc tgtggcctac agcaacaact ctatcgctat ccccaccaac

ttcaccatca gcgtgaccac cgagatcctg cctgtgtcca tgaccaagac cagcgtggac

tgcaccatgt acatctgcgg cgattccacc gagtgcagca acctgctgct gcagtacggc

agcttctgca cccagctgaa tagagccctg acagggatcg ccgtggaaca ggacaagaac

acccaagagg tgttcgccca agtgaagcag atctacaaga cccctcctat caaggacttc

ggcggcttca atttcagcca gattctgccc gatcctagca agcccagcaa gcggagcttt

atcgaggacc tgctgttcaa caaagtgaca ctggccgacg ccggcttcat caagcagtat

ggcgattgcc tgggcgacat tgccgccaga gatctgattt gcgcccagaa gtttaacgga

ctgacagtgc tgcctcctct gctgaccgat gagatgatcg cccagtacac atctgctctg

ctggccggca caatcaccag cggatggaca tttggagctg gcgcagccct gcagatcccc

tttgctatgc agatggccta ccggttcaac ggcatcggag tgacccagaa tgtgctgtac

gagaaccaga agctgatcgc caaccagttc aacagcgcca tcggcaagat ccaggatagc

ctgtctagca cagccagcgc tctgggcaaa ctgcaggacg tggtcaatca gaacgctcag

gccctgaaca ccctcgtgaa gcagctgagc agcaatttcg gcgccatcag ctccgtgctg

aacgatatcc tgagccggct ggataaggtg gaagccgagg tgcagatcga cagactgatc

acaggcagac tgcagagcct ccagacatac gtgacccagc agctgatcag agccgccgag

attagagcct ctgccaatct ggccgccacc aagatgtctg agtgtgtgct gggccagagc

aagagagtgg atttctgcgg caagggctac cacctgatga gctttccaca gtctgctcct

cacggcgtgg tgtttctgca cgtgacctat gtgcccgctc aagagaagaa cttcacaaca

gcccctgcca tctgccacga cggaaaggcc cattttccta gagaaggcgt gttcgtgtcc

aacggcaccc attggttcgt gacacagcgg aacttctacg agccccagat catcaccacc

[0395] gacaacacct tcgtgtctgg caactgtgac gtcgtgatcg gcattgtgaa caacaccgtg

tacgaccctc tgcagcccga gctggacagc ttcaaagagg aactggacaa gtactttaag

aaccacacaa gccccgacgt ggacctgggc gatattagcg gcatcaatgc ctccgtggtc

aacatccaga aagagatcga ccggctgaac gaggtggcca agaatctgaa cgagagcctg

atcgacctgc aagaactggg gaagtacgag cagtacatca agtggccctg gtacatctgg

ctgggcttta tcgccggact gattgccatc gtgatggtca caatcatgct gtgctgcatg

accagctgct gtagctgcct gaagggctgt tgcagctgtg gcagctgctg caagttcgac

gaggatgata gcgagcctgt gctgaagggc gtgaaactgc actacacc**GC GGCCGC**

[0396] 5'BstBI单克隆位点是带单下划线的

[0397] 3'NotI单克隆位点是带短下划线的

[0398] 紧接着5'NheI是GAC密码子(为编码序列与紧接着GAC的ATG起始密码子在框内所需)。这两个密码子以粗体和斜体显示。

[0399] SEQ ID NO:7的核酸序列翻译得到SEQ ID NO:1的天然SARS-CoV-2刺突蛋白。

[0400] SEQ ID NO:8-编码融合蛋白HBSAg/SARS-CoV-2刺突蛋白的核酸-针对在人(293F)中的表达优化并含有NheI和NotI单克隆位点。

GCTAGCGACatgaactttctgggcggtacgacagtatgccttggacaaaattcacaatctccgacgtct

aatcactcccctacaagttgtccaccgacttgccccggctataggtggatgtgtctcagacgattcata

[0401] atctttctcttcattcttcttctgtgcctgatattcttgctggtccttctggattaccagggaatgctt

cccgtgtgtcctctgattcctggttcatccactacatctacgggtccctgtagaacatgcaccacacct

gcacagggcacctccatgtatccgtcatgctgctgcacgaaaccatcagatggtaactgcacgtgcata

ccgatcccctcatcatgggcgtttgggaaatttctgtgggagtgggcctcagcccggttttccTTCGTG

TTTCTGGTGCTGCTGCCTCTGGTGTCCAGCCAGTGTGTGAACCTGACCACCAGAACACAGCTGCCTCCA

GCCTACACCAATAGCTTCACCAGGGGCGTGTACTACCCCGACAAGGTGTTCAGATCTAGCGTGCTGCAC

AGCACCCAGGACCTGTTTCTGCCCTTCTTCAGCAACGTGACCTGGTTCCACGCCATCCACGTGTCCGGC

ACCAATGGCACCAAGAGATTCGACAACCCCGTGCTGCCCTTCAACGATGGGGTGTACTTTGCCAGCACC

GAGAAGTCCAACATCATCAGAGGCTGGATCTTCGGCACCACACTGGACAGCAAGACCCAGAGCCTGCTG

ATCGTGAACAACGCCACCAACGTGGTCATCAAAGTGTGCGAGTTCCAGTTCTGCAACGACCCATTCCTG

GGAGTCTACTACCACAAGAACAACAAGAGCTGGATGGAAAGCGAGTTCCGGGTGTACAGCAGCGCCAAC

AACTGCACCTTCGAGTACGTGTCCCAGCCTTTCCTGATGGACCTGGAAGGCAAGCAGGGCAACTTCAAG

AACCTGCGCGAGTTCGTGTTCAAGAACATCGACGGCTACTTCAAGATCTACAGCAAGCACACCCCTATC

AACCTCGTGCGGGATCTGCCTCAGGGCTTTTCTGCTCTGGAACCTCTGGTGGACCTGCCTATCGGCATC

AACATCACCCGGTTTCAGACCCTGCTGGCCCTGCACAGATCTTACCTGACACCTGGCGATAGCAGCTCT

GGATGGACAGCTGGCGCCGCTGCCTATTATGTGGGCTACCTGCAGCCTCGGACCTTCCTGCTGAAGTAC

AACGAGAACGGCACCATCACCGACGCCGTGGATTGTGCTCTGGATCCCCTGAGCGAGACAAAGTGCACC

CTGAAGTCCTTCACCGTGGAAAAGGGCATCTACCAGACCAGCAACTTCAGAGTGCAGCCCACCGAGAGC

[0402]

ATCGTGCGGTTCCCCAATATCACCAATCTGTGCCCCTTCGGCGAGGTGTTCAATGCCACAAGATTTGCC

AGCGTGTACGCCTGGAACCGGAAGAGAATCAGCAACTGCGTGGCCGACTACAGCGTGCTGTACAATAGC

GCCAGCTTCAGCACCTTCAAGTGCTACGGCGTGTCCCCTACCAAGCTGAACGACCTGTGCTTCACCAAT

GTGTACGCCGACAGCTTCGTGATCAGAGGCGACGAAGTTCGGCAGATCGCTCCTGGACAGACAGGCAAG

ATCGCCGATTACAACTACAAGCTGCCCGACGACTTCACCGGCTGCGTGATCGCCTGGAATAGCAACAAC

CTGGACTCCAAAGTCGGCGGCAACTACAACTACCTGTACCGGCTGTTCCGGAAGTCCAATCTGAAGCCC

TTCGAGCGGGACATCTCCACCGAAATCTATCAGGCCGGCAGCACCCCTTGTAACGGCGTGGAAGGCTTC

AACTGCTACTTCCCACTGCAGTCCTACGGCTTTCAGCCTACCAATGGCGTGGGCTATCAGCCCTATAGA

GTGGTGGTGCTGAGCTTCGAACTGCTGCATGCCCCTGCTACCGTGTGCGGCCCTAAGAAGTCTACCAAC

CTGGTCAAGAACAAATGCGTGAACTTCAACTTCAACGGCCTGACCGGCACAGGCGTGCTGACAGAGAGC

AACAAGAAGTTCCTGCCTTTCCAGCAGTTTGGCCGGGATATCGCCGATACCACAGACGCCGTTAGAGAT

CCCCAGACACTGGAAATCCTGGACATCACCCCATGCAGCTTTGGCGGAGTGTCTGTGATCACCCCTGGC

ACCAATACCAGCAATCAGGTGGCCGTGCTGTATCAGGACGTGAACTGTACAGAGGTGCCCGTGGCCATT

CACGCCGATCAACTGACACCCACTTGGAGAGTGTACTCCACCGGCTCCAACGTGTTCCAGACTAGAGCC

GGATGTCTGATCGGAGCCGAGCACGTGAACAATAGCTACGAGTGCGACATCCCCATCGGCGCTGGCATC

TGTGCCAGCTACCAGACACAGACAAATAGCCCCAGACGGGCCAGAAGCGTGGCCTCTCAGAGCATCATT

GCCTACACAATGAGCCTGGGCGCCGAGAATTCTGTGGCCTACAGCAACAACTCTATCGCTATCCCCACC

AACTTCACCATCAGCGTGACCACCGAGATCCTGCCTGTGTCCATGACCAAGACCAGCGTGGACTGCACC

ATGTACATCTGCGGCGATTCCACCGAGTGCAGCAACCTGCTGCTGCAGTACGGCAGCTTCTGCACCCAG

CTGAATAGAGCCCTGACAGGGATCGCCGTGGAACAGGACAAGAACACCCAAGAGGTGTTCGCCCAAGTG

AAGCAGATCTACAAGACCCCTCCTATCAAGGACTTCGGCGGCTTCAATTTCAGCCAGATTCTGCCCGAT

CCTAGCAAGCCCAGCAAGCGGAGCTTTATCGAGGACCTGCTGTTCAACAAAGTGACACTGGCCGACGCC

GGCTTCATCAAGCAGTATGGCGATTGCCTGGGCGACATTGCCGCCAGAGATCTGATTTGCGCCCAGAAG

TTTAACGGACTGACAGTGCTGCCTCCTCTGCTGACCGATGAGATGATCGCCCAGTACACATCTGCTCTG

CTGGCCGGCACAATCACCAGCGGATGGACATTTGGAGCTGGCGCAGCCCTGCAGATCCCCTTTGCTATG

CAGATGGCCTACCGGTTCAACGGCATCGGAGTGACCCAGAATGTGCTGTACGAGAACCAGAAGCTGATC

GCCAACCAGTTCAACAGCGCCATCGGCAAGATCCAGGATAGCCTGTCTAGCACAGCCAGCGCTCTGGGC

AAACTGCAGGACGTGGTCAATCAGAACGCTCAGGCCCTGAACACCCTCGTGAAGCAGCTGAGCAGCAAT

[0403] TTCGGCGCCATCAGCTCCGTGCTGAACGATATCCTGAGCCGGCTGGATAAGGTGGAAGCCGAGGTGCAG

ATCGACAGACTGATCACAGGCAGACTGCAGAGCCTCCAGACATACGTGACCCAGCAGCTGATCAGAGCC

GCCGAGATTAGAGCCTCTGCCAATCTGGCCGCCACCAAGATGTCTGAGTGTGTGCTGGGCCAGAGCAAG

AGAGTGGATTTCTGCGGCAAGGGCTACCACCTGATGAGCTTTCCACAGTCTGCTCCTCACGGCGTGGTG

TTTCTGCACGTGACCTATGTGCCCGCTCAAGAGAAGAACTTCACAACAGCCCCTGCCATCTGCCACGAC

GGAAAGGCCCATTTTCCTAGAGAAGGCGTGTTCGTGTCCAACGGCACCCATTGGTTCGTGACACAGCGG

AACTTCTACGAGCCCCAGATCATCACCACCGACAACACCTTCGTGTCTGGCAACTGTGACGTCGTGATC

GGCATTGTGAACAACACCGTGTACGACCCTCTGCAGCCCGAGCTGGACAGCTTCAAAGAGGAACTGGAC

AAGTACTTTAAGAACCACACAAGCCCCGACGTGGACCTGGGCGATATTAGCGGCATCAATGCCTCCGTG

GTCAACATCCAGAAAGAGATCGACCGGCTGAACGAGGTGGCCAAGAATCTGAACGAGAGCCTGATCGAC

CTGCAAGAACTGGGGAAGTACGAGCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGGCTTTATCGCC

GGACTGATTGCCATCGTGATGGTCACAATCATGCTGTGCTGCATGACCAGCTGCTGTAGCTGCCTGAAG

GGCTGTTGCAGCTGTGGCAGCTGCTGCAAGTTCGACGAGGATGATAGCGAGCCTGTGCTGAAGGGCGTG

AAACTGCACTACACC**GCGGCCGC**

[0404] 5’NheI单克隆位点是带单下划线的

[0405] HBSAg序列以小写字母显示

[0406] SARS-CoV-2刺突蛋白编码序列以大写字母显示

[0407] 3’NotI单克隆位点是带短下划线的

[0408] 紧接着5’NheI是GAC密码子(为编码序列与紧接着GAC的ATG起始密码子在框内所需)。这两个密码子以粗体和斜体显示。

[0409] SEQ ID NO:9-对应于SEQ ID NO:3的氨基酸序列

[0410] (融合蛋白HEV-SARS-CoV-2刺突蛋白-针对在大肠杆菌中的表达优化并含有SacI和NotI单一克隆位点。)

[0411] MIALTLFNLADTLLGGLPTELISSAGGQLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDLGE
SRVVIQDYDNQHEQDRPTPSPAPSRPFSVLRANDVLWLSLTAAEYDQSTYGSSTGPVYVSDSVTLVNVATGAQAVAR
SLDWTKVTLDGRPLSTIQQYSKTFFVLPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAGHRVAISTYTTSLG
AGPVSISAVAVLAPHSAFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSN
VTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCND
PFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRD
LPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCA
LDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLY
NSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVG
GNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPAT
VCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITP
GTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQ
TNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLL
QYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADA
GFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNG
IGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLD
KVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVF
LHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYD
PLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYI
WLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

[0412] SEQ ID NO:10-对应于SEQ ID NO:5的氨基酸序列

[0413] (融合蛋白HPV18L1/SARS-CoV-2刺突蛋白-针对在K.pastoris中的表达优化并含有BstB1和NotI单一克隆位点。)

[0414] MALWRPSDNTVYLPPPSVARVVNTDDYVTRTSIFYHAGSSRLLTVGNPYFRVPAGGGNKQDIPKVSAYQ
YRVFRVQLPDPNKFGLPDTSIYNPETQRLVWACAGVEIGRGQPLGVGLSGHPFYNKLDDTESSHAATSNVSEDVRDN
VSVDYKQTQLCILGCAPAIGEHWAKGTACKSRPLSQGDCPPLELKNTVLEDGDMVDTGYGAMDFSTLQDTKCEVPLD
ICQSICKYPDYLQMSADPYGDSMFFCLRREQLFARHFWNRAGTMGDTVPQSLYIKGTGMRASPGSCVYSPSPSGSIV
TSDSQLFNKPYWLHKAQGHNNGVCWHNQLFVTVVDTTRSTNLTICASTQSPVPGQYDATKFKQYSRHVEEYDLQFIF
QLCTITLTADVMSYIHSMNSSILEDWNFGVPPPPTTSLVDTYRFVQSVAITCQKDAAPAENKDPYDKLKFWNVDLKE
KFSLDLDQYPLGRKFLVQAGLRRKPTIGPRKRSAPSATTSSKPAKRVRVRARKFVFLVLLPLVSSQCVNLTTRTQLP
PAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIR
GWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDL
EGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSS
GWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNIT

NLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQI
APGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFN
CYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQ
FGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSN
VFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNF
TISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIK
DFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQ
YTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQ
DVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAA
TKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWF
VTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQ
KEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFD
EDDSEPVLKGVKLHYT

[0415] SEQ ID ID NO:11-对应于SEQ ID NO:6的氨基酸序列

[0416] (融合蛋白HPV16L1/SARS-CoV-2刺突蛋白-针对在K.pastoris中的表达优化并含有BstB1和NotI单一克隆位点。)

[0417] MSLWLPSEATVYLPPVPVSKVVSTDEYVARTNIYYHAGTSRLLAVGHPYFPIKKPNNNKILVPKVSGLQ
YRVFRIHLPDPNKFGFPDTSFYNPDTQRLVWACVGVEVGRGQPLGVGISGHPLLNKLDDTENASAYAANAGVDNREC
ISMDYKQTQLCLIGCKPPIGEHWGKGSPCTNVAVNPGDCPPLELINTVIQDGDMVDTGFGAMDFTTLQANKSEVPLD
ICTSICKYPDYIKMVSEPYGDSLFFYLRREQMFVRHLFNRAGAVGENVPDDLYIKGSGSTANLASSNYFPTPSGSMV
TSDAQIFNKPYWLQRAQGHNNGICWGNQLFVTVVDTTRSTNMSLCAAISTSETTYKNTNFKEYLRHGEEYDLQFIFQ
LCKITLTADVMTYIHSMNSTILEDWNFGLQPPPGGTLEDTYRFVTSQAIACQKHTPPAPKEDPLKKYTFWEVNLKEK
FSADLDQFPLGRKFLLQAGLKAKPKFTLGKRKATPTTSSTSTTAKRKKRKLFVFLVLLPLVSSQCVNLTTRTQLPPA
YTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGW
IFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEG
KQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGW
TAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNL
CPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAP
GQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCY
FPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFG
RDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVF
QTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTI
SVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDF
GGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYT
SALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDV
VNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATK
MSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVT
QRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKE
IDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDED

DSEPVLKGVKLHYT

[0418] SEQ ID NO:12-对应于SEQ ID NO:8的氨基酸序列

[0419] (融合蛋白HBSAg/SARS-CoV-2刺突蛋白-针对在人类(293F)中的表达优化并含有NheI和NotI单一克隆位点。)

[0420] MNFLGGTTVCLGQNSQSPTSNHSPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVC
PLIPGSSTTSTGPCRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSFVFLVLLPLVSSQ
CVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFSNVTWFHAIHVSGTNGTKRFDNPVLPFNDGVY
FASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCNDPFLGVYYHKNNKSWMESEFRVYSSANNCTF
EYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVRDLPQGFSALEPLVDLPIGINITRFQTLLALH
RSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDCALDPLSETKCTLKSFTVEKGIYQTSNFRVQP
TESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADS
FVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAG
STPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKKSTNLVKNKCVNFNFNGLTGTGVLT
ESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVITPGTNTSNQVAVLYQDVNCTEVPVAIHADQLT
PTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVASQSIIAYTMSLGAENSVAY
SNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLLLQYGSFCTQLNRALTGIAVEQDKNTQEVFAQ
VKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLADAGFIKQYGDCLGDIAARDLICAQKFNGLTVL
PPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFNGIGVTQNVLYENQKLIANQFNSAIGKIQDSL
SSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRLDKVEAEVQIDRLITGRLQSLQTYVTQQLIRA
AEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVVFLHVTYVPAQEKNFTTAPAICHDGKAHFPRE
GVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVYDPLQPELDSFKEELDKYFKNHTSPDVDLGDI
SGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWYIWLGFIAGLIAIVMVTIMLCCMTSCCSCLKG
CCSCGSCCKFDEDDSEPVLKGVKLHYTAA

[0421] SEQ ID NO:13-RBD SARS-CoV-2刺突蛋白核酸序列

[0422]
```
GCTAGCGACgccaccATGAGAGTCCAACCAACAGAATCTATTGTTAGATT
TCCTAATATTACAAACTTGTGCCCTTTTGGTGAAGTTTTTAACGCCACCA
GATTTGCATCTGTTTATGCTTGGAACAGGAAGAGAATCAGCAACTGTGTT
GCTGATTATTCTGTCCTATATAATTCCGCATCATTTTCCACTTTTAAGTG
TTATGGAGTGTCTCCTACTAAATTAAATGATCTCTGCTTTACTAATGTCT
ATGCAGATTCATTTGTAATTAGAGGTGATGAAGTCAGACAAATCGCTCCA
GGGCAAACTGGAAAGATTGCTGATTATAATTATAAATTACCAGATGATTT
TACAGGCTGCGTTATAGCTTGGAATTCTAACAATCTTGATTCTAAGGTTG
GTGGTAATTATAATTACCTGTATAGATTGTTTAGGAAGTCTAATCTCAAA
CCTTTTGAGAGAGATATTTCAACTGAAATCTATCAGGCCGGTAGCACACC
TTGTAATGGTGTTGAAGGTTTTAATTGTTACTTTCCTTTACAATCATATG
GTTTCCAACCCACTAATGGTGTTGGTTACCAACCATACAGAGTAGTAGTA
CTTTCTTTTGAACTTCTACATGCACCAGCAACTGTTTGTGGACCTAAAAA
GtgataaGCGGCCGC
```

[0423]

[0424] 在起始ATG(粗体)前添加了KOZAC序列(gcc acc,带下划线的)。

[0425] 在NotI前添加了分泌形式的tga taa(双下划线的)-该tga taa序列是“两个终止

密码子”基序,其中断蛋白质合成,促进分泌到胞外培养基中(也包括在其他序列中,如下所述)。

[0426] SEQ ID NO:14-RBD SARS-CoV-2刺突蛋白核酸序列-针对293F(HEK)细胞表达优化的人密码子。

[0427] GCTAGCGACgccaccATGAGAGTGCAGCCTACAGAGTCTATCGTGCGGTTCCCCAACATCACCAATCTGTGCC
CTTTCGGCGAGGTGTTCAACGCCACAAGATTTGCCAGCGTGTACGCCTGGAACCGGAAGAGAATCAGCAA
CTGCGTGGCCGACTACAGCGTGCTGTACAATAGCGCCAGCTTCAGCACCTTCAAGTGCTACGGCGTGTCCCC
TACCAAGCTGAACGACCTGTGCTTCACCAATGTGTACGCCGACAGCTTCGTGATCAGAGGCGACGAAGTTC
GGCAGATCGCTCCTGGACAGACAGGCAAGATCGCCGATTACAACTACAAGCTGCCCGACGACTTCACCGGC
TGCGTGATCGCCTGGAATAGCAACAACCTGGACAGCAAAGTCGGCGGCAACTACAACTACCTGTACCGGCT
GTTCCGGAAGTCCAACCTGAAGCCTTTCGAGCGGGACATCAGCACCGAGATCTATCAGGCCGGCAGCACCC
CTTGTAATGGCGTGGAAGGCTTCAACTGCTACTTCCCACTGCAGTCCTACGGCTTCCAGCCTACAAACGGCG
TGGGCTACCAGCCTTATAGAGTGGTGGTGCTGAGCTTCGAACTGCTGCATGCCCCTGCTACAGTGTGCGGCC
CCAAGAAGtgataaGCGGCCGC

[0428] 在起始ATG(粗体)前添加了KOZAC序列(gcc acc,带下划线的)。

[0429] 在NotI前添加了分泌形式的tga taa

[0430] 已经分别在5’端NheI和3’端NotI(带短下划线的)添加了独特的限制位点

[0431] SEQ ID NO:15-对应于SEQ ID NO:13和14的RBD SARS-CoV-2刺突蛋白氨基酸序列

[0432] MRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTK
LNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKP
FERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKK

[0433] SEQ ID NO:16-刚性EAAAK接头共有氨基酸序列

[0434] $A(EAAAK)_nA$(n=2-5)

[0435] SEQ ID NO:17-刚性$(EAAAK)_3$接头核酸序列

[0436] GAA GCC GCC GCT AAA GAG GCC GCT GCC AAA GAA GCT GCT GCT AAG

[0437] SEQ ID NO:18-刚性$(EAAAK)_3$接头氨基酸序列

[0438] EAAAKEAAAKEAAAK

[0439] SEQ ID NO:19-柔性GS_n接头共有氨基酸序列

[0440] $(Gly-Gly-Gly-Gly-Ser)_n$(n=1-6)

[0441] SEQ ID NO:20-柔性GS5($(GGGGS)_1$)接头氨基酸序列

[0442] GGGGS

[0443] SEQ ID NO:21-柔性GS10($(GGGGS)_2$)接头氨基酸序列

[0444] GGGGSGGGGS

[0445] SEQ ID NO:22-柔性GS15($(GGGGS)_3$)接头核酸序列

[0446] GGT GGT GGT GGT AGC GGT GGT GGC GGT TCA GGT GGC GGT GGT TCA

[0447] SEQ ID NO:23-柔性GS15($(GGGGS)_3$)接头氨基酸序列

[0448] GGGGSGGGGSGGGGS

[0449] SEQ ID NO:24-柔性GS20($(GGGGS)_4$)接头氨基酸序列

[0450] GGGGSGGGGSGGGGSGGGGS

[0451] SEQ ID NO:25-柔性GS25($(GGGGS)_5$)接头氨基酸序列

[0452] GGGGSGGGGSGGGGSGGGGSGGGGS

[0453] SEQ ID NO:26-HBsAg-(EAAAK)$_3$-RBD核酸序列

GCTAGCGACgccacc**ATG**ATTGCACTGACCCTGTTTAATCTGGCAGATAC
CCTGTTAGGTGGTCTGCCGACCGAACTGATTAGCAGTGCCGGTGGTCAGC
TGTTTTATAGCCGTCCGGTTGTTAGCGCAAATGGTGAACCGACCGTTAAA
CTGTATACCAGCGTTGAAAATGCACAGCAGGATAAAGGTATTGCAATTCC
GCATGATATTGATCTGGGTGAAAGCCGTGTTGTGATTCAGGATTATGATA
ATCAGCATGAACAGGATCGTCCGACACCGAGTCCGGCACCGAGCCGTCCG
TTTAGCGTTCTGCGTGCAAATGATGTTCTGTGGCTGAGCCTGACCGCAGC
AGAATATGATCAGAGCACCTATGGTAGCAGCACCGGTCCGGTTTATGTTA
GCGATAGCGTTACCCTGGTTAATGTTGCAACCGGTGCACAGGCAGTTGCA
CGTAGCCTGGATTGGACCAAAGTGACCCTGGATGGTCGTCCGCTGAGCAC
CATTCAGCAGTATAGCAAAACCTTTTTTGTTCTGCCGCTGCGTGGTAAAC
TGAGCTTTTGGGAAGCAGGCACCACCAAAGCAGGTTATCCGTATAACTAT
AATACCACCGCAAGCGATCAGCTGCTGGTTGAAAACGCAGCAGGTCATCG
TGTTGCAATTAGCACCTATACCACCAGTTTAGGTGCAGGTCCGGTTAGCA

[0454] TTAGCGCAGTTGCAGTTCTGGCACCGCATTCAGCC**gaagcagccgctaaa**
gaagcagccgctaaagaagcagccgctaaaAGAGTCCAACCAACAGAATC
TATTGTTAGATTTCCTAATATTACAAACTTGTGCCCTTTTGGTGAAGTTT
TTAACGCCACCAGATTTGCATCTGTTTATGCTTGGAACAGGAAGAGAATC
AGCAACTGTGTTGCTGATTATTCTGTCCTATATAATTCCGCATCATTTTC
CACTTTTAAGTGTTATGGAGTGTCTCCTACTAAATTAAATGATCTCTGCT
TTACTAATGTCTATGCAGATTCATTTGTAATTAGAGGTGATGAAGTCAGA
CAAATCGCTCCAGGGCAAACTGGAAAGATTGCTGATTATAATTATAAATT
ACCAGATGATTTTACAGGCTGCGTTATAGCTTGGAATTCTAACAATCTTG
ATTCTAAGGTTGGTGGTAATTATAATTACCTGTATAGATTGTTTAGGAAG
TCTAATCTCAAACCTTTTGAGAGAGATATTTCAACTGAAATCTATCAGGC
CGGTAGCACACCTTGTAATGGTGTTGAAGGTTTTAATTGTTACTTTCCTT
TACAATCATATGGTTTCCAACCCACTAATGGTGTTGGTTACCAACCATAC
AGAGTAGTAGTACTTTCTTTTGAACTTCTACATGCACCAGCAACTGTTTG

TGGACCTAAAAAGtgataaGCGGCCGC

[0455] 在起始ATG(粗体)前添加了KOZAC序列(gcc acc,带下划线的)。

[0456] 在NotI前添加了分泌形式的tga taa(带双下划线的)

[0457] 已经分别在5’端NheI和3’端NotI(带短下划线的)添加了独特的限制位点

[0458] 粗体和带短下划线的序列对应于(EAAAK)$_3$接头

[0459] SEQ ID NO:27-针对293f(HEK)细胞表达人密码子优化的HBSAg-(EAAAK)$_3$-RBD核酸序列

[0460]

GCTAGCGACgccaccATGAATTTTCTCGGCGGCACAACAGTGTGCCTGGGCCAGAATAGCCAGTCTCCTACCA

GCAATCACAGCCCCACCAGCTGTCCTCCAACCTGTCCTGGCTACAGATGGATGTGCCTGCGGCGGTTCATCA

TCTTTCTGTTCATCCTGCTGCTGTGCCTGATCTTCCTGCTGGTGCTGCTGGATTACCAGGGAATGCTGCCTGT

GTGTCCTCTGATCCCTGGCAGCAGCACAACAAGCACAGGCCCTTGCAGAACCTGCACAACACCAGCTCAGG

GCACCAGCATGTACCCTAGCTGCTGTTGTACCAAGCCTAGCGACGGCAACTGCACATGCATCCCCATTCCTA

GCAGCTGGGCCTTCGGCAAGTTTCTGTGGGAATGGGCCAGCGCCAGATTTTCCGAAGCCGCCGCTAAAGA

GGCCGCTGCCAAAGAAGCTGCTGCTAAGAGAGTGCAGCCCACCGAGTCTATCGTGCGGTTCCCCAACATCA

CCAATCTGTGCCCTTTCGGCGAGGTGTTCAACGCCACAAGATTTGCCAGCGTGTACGCCTGGAACCGGAAG

AGAATCAGCAACTGCGTGGCCGACTACAGCGTGCTGTACAATAGCGCCAGCTTCAGCACCTTCAAGTGCTAC

GGCGTGTCCCCTACCAAGCTGAACGACCTGTGCTTCACCAATGTGTACGCCGACAGCTTCGTGATCAGAGGC

GACGAAGTTCGGCAGATCGCTCCTGGACAGACAGGCAAGATCGCCGATTACAACTACAAGCTGCCCGACGA

CTTCACCGGCTGCGTGATCGCCTGGAATAGCAACAACCTGGACAGCAAAGTCGGCGGCAACTACAACTACC

TGTACCGGCTGTTCCGGAAGTCCAACCTGAAGCCTTTCGAGCGGGACATCAGCACCGAAATCTACCAGGCC

GGCAGCACCCCTTGTAATGGCGTGGAAGGCTTCAACTGCTACTTCCCACTGCAGTCCTACGGCTTCCAGCCT

ACAAACGGCGTGGGCTACCAGCCTTATAGAGTGGTGGTGCTGAGCTTCGAACTGCTGCATGCCCCTGCTAC

AGTGTGCGGCCCCAAGAAGtgataaGCGGCCGC

[0461] 在起始ATG(粗体)前添加了KOZAC序列(gcc acc,带下划线的)。

[0462] 在NotI前添加了分泌形式的tga taa

[0463] 已经分别在5’端NheI和3’端NotI(带短下划线的)添加了独特的限制位点

[0464] 粗体和带短下划线的序列对应于$(EAAAK)_3$接头

[0465] SEQ ID NO:28-对应于SEQ ID NO:26和27的HBSAg-$(EAAAK)_3$-RBD氨基酸序列

[0466] MNFLGGTTVCLGQNSQSPTSNHSPTSCPPTCPGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVC
PLIPGSSTTSTGPCRTCTTPAQGTSMYPSCCCTKPSDGNCTCIPIPSSWAFGKFLWEWASARFSEAAAKEAAAKEAA
AKRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVLYNSASFSTFKCYGVSPTKLNDLCFT
NVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKVGGNYNYLYRLFRKSNLKPFERDIST
EIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPATVCGPKK

[0467] $(EAAAK)_3$接头是带下划线的。

SEQ ID NO:29-HEV-GS15-RBD 核酸序列

[0468]

GAGCTC**ATG**ATTGCACTGACCCTGTTTAATCTGGCAGATACCCTGCTGGG
TGGTCTGCCGACCGAACTGATTAGCAGTGCCGGTGGTCAGCTGTTTTATA
GCCGTCCGGTTGTTAGCGCAAATGGTGAACCGACCGTTAAACTGTATACC
AGCGTTGAAAATGCACAGCAGGATAAAGGTATTGCAATTCCGCATGATAT
TGATCTGGGTGAAAGCCGTGTTGTGATTCAGGATTATGATAATCAGCATG
AACAGGATCGTCCGACCCCGAGTCCGGCACCGAGCCGTCCGTTTAGCGTT
CTGCGTGCAAATGATGTTCTGTGGCTGAGCCTGACCGCAGCAGAATATGA
TCAGAGCACCTATGGTAGCAGCACCGGTCCGGTTTATGTTAGCGATAGCG
TTACCCTGGTTAATGTTGCAACCGGTGCACAGGCAGTTGCACGTAGCCTG
GATTGGACCAAAGTGACCCTGGATGGTCGTCCGCTGAGCACCATTCAGCA
GTATAGCAAAACCTTTTTTGTTCTGCCGCTGCGTGGTAAACTGAGCTTTT
GGGAAGCAGGCACCACCAAAGCAGGTTATCCGTATAACTATAATACCACC
GCAAGCGATCAGCTGCTGGTTGAAAACGCAGCAGGTCATCGTGTTGCAAT
TAGCACCTATACCACCAGTCTGGGTGCAGGTCCGGTTAGCATTAGCGCAG
TTGCAGTTCTGGCACCGCATAGCGCA**ggtggaggaggttctggaggcggt**
ggaagtggtggcggaggtagcAGAgtccaaccaacagaatctattgttag
atttcctaatattacaaacttgtgccctttggtgaagtttttaacgcca
ccagatttgcatctgtttatgcttggaacaggaagagaatcagcaactgt
gttgctgattattctgtcctatataattccgcatcattttccacttttaa
gtgttatggagtgtctcctactaaattaaatgatctctgctttactaatg
tctatgcagattcatttgtaattagaggtgatgaagtcagacaaatcgct
ccagggcaaactggaaagattgctgattataattataaattaccagatga
ttttacaggctgcgttatagcttggaattctaacaatcttgattctaagg
ttggtggtaattataattacctgtatagattgtttaggaagtctaatctc
aaaccttttgagagagatatttcaactgaaatctatcaggccggtagcac
accttgtaatggtgttgaaggttttaattgttactttcctttacaatcat
atggtttccaacccactaatggtgttggttaccaaccatacagagtagta
gtactttcttttgaacttctacatgcaccagcaactgtttgtggacctaa
aaag**tgataa**GCGGCCGC

[0469] 起始ATG(粗体)

[0470] 已经分别在5’端SacI和3’端NotI(带短下划线的)添加了独特的限制位点

[0471] 在NotI前添加了分泌形式的tga taa(带双下划线的)

[0472] 粗体和带短下划线的序列对应于GS15接头

[0473] SEQ ID NO:30-针对大肠杆菌表达优化的HEV-GS15-RBD核酸序列

[0474] GAGCTCATGATTGCACTGACCCTGTTTAATCTGGCAGATACCCTGTTAGGTGGTCTGCCGACCGAACTGATTA
GCAGTGCCGGTGGTCAGCTGTTTTATAGCCGTCCGGTTGTTAGCGCAAATGGTGAACCGACCGTTAAACTGT
ATACCAGCGTTGAAAATGCACAGCAGGATAAAGGTATTGCAATTCCGCATGATATTGATCTGGGTGAAAGCC
GTGTTGTGATTCAGGATTATGATAATCAGCATGAACAGGATCGTCCGACACCGAGTCCGGCACCGAGCCGTC
CGTTTAGCGTTCTGCGTGCAAATGATGTTCTGTGGCTGAGCCTGACCGCAGCAGAATATGATCAGAGCACCT
ATGGTAGCAGCACCGGTCCGGTTTATGTTAGCGATAGCGTTACCCTGGTTAATGTTGCAACCGGTGCACAGG
CAGTTGCACGTAGCCTGGATTGGACCAAAGTGACCCTGGATGGTCGTCCGCTGAGCACCATTCAGCAGTATA
GCAAAACCTTTTTTGTTCTGCCGCTGCGTGGTAAACTGAGCTTTTGGGAAGCAGGCACCACCAAAGCAGGT
TATCCGTATAACTATAATACCACCGCAAGCGATCAGCTGCTGGTTGAAAACGCAGCAGGTCATCGTGTTGCAA
TTAGCACCTATACCACCAGTTTAGGTGCAGGTCCGGTTAGCATTAGCGCAGTTGCAGTTCTGGCACCGCATT
CAGCCGGTGGTGGTGGTAGCGGTGGTGGCGGTTCAGGTGGCGGTGGTTCACGTGTTCAGCCGACAGAAA
GCATTGTTCGTTTTCCGAATATCACCAATCTGTGTCCGTTTGGCGAAGTTTTTAATGCAACCCGTTTTGCAAG
CGTTTATGCCTGGAATCGTAAACGTATTAGCAATTGCGTTGCCGATTATAGCGTGCTGTATAATAGCGCAAGCT
TTAGCACCTTTAAATGCTATGGTGTTAGCCCGACCAAACTGAATGATCTGTGTTTTACCAATGTGTATGCCGAT
AGCTTTGTGATTCGTGGTGATGAAGTTCGTCAGATTGCACCGGGTCAGACCGGTAAAATTGCAGATTATAAC
TACAAACTGCCGGATGATTTTACGGGTTGTGTTATTGCATGGAATAGCAATAACCTGGATAGCAAAGTTGGTG
GCAACTATAACTATCTGTATCGCCTGTTTCGTAAGAGCAATCTGAAACCGTTTGAACGTGATATTAGCACCGAA
ATTTATCAGGCAGGTAGCACCCCGTGCAATGGTGTTGAAGGTTTTAATTGTTATTTTCCGCTGCAGAGCTATG
GTTTTCAGCCTACCAATGGTGTGGGTTATCAGCCGTATCGTGTTGTTGTTCTGTCATTTGAACTGCTGCATGC
ACCGGCAACCGTTTGTGGTCCGAAAAAAtgataaGCGGCCGC

[0475] 起始ATG(粗体)

[0476] 在NotI前添加了分泌形式的tga taa(带双下划线的)

[0477] 已经分别在5’端SacI和3’端NotI(带短下划线的)添加了独特的限制位点

[0478] 粗体和带短下划线的序列对应于GS15接头

[0479] SEQ ID NO:31-对应于SEQ ID NO:29和30的HEV-GS15-RBD氨基酸序列

[0480] MIALTLFNLADTLLGGLPTELISSAGGQLFYSRPVVSANGEPTVKLYTSVENAQQDKGIAIPHDIDLGE
SRVVIQDYDNQHEQDRPTPSPAPSRPFSVLRANDVLWLSLTAAEYDQSTYGSSTGPVYVSDSVTLVNVATGAQAVAR
SLDWTKVTLDGRPLSTIQQYSKTFFVLPLRGKLSFWEAGTTKAGYPYNYNTTASDQLLVENAAGHRVAISTYTTSLG
AGPVSISAVAVLAPHSAGGGGSGGGGSGGGGSRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVA
DYSVLYNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNN
LDSKVGGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFEL
LHAPATVCGPKK

[0481] GS15接头是带下划线的。

[0482] SEQ ID NO:32-针对293f(HEK)细胞表达人密码子优化的HBsAg-(EAAAK)$_3$-全长SARS-CoV-2刺突蛋白核酸序列

AAGCTTGCCgccacc**ATG**GAGAACATCACATCAGGATTCCTAGGACCCCTGCTCGTGTTACAGGC

GGGGTTTTTCTTGTTGACAAGAATCCTCACAATACCACAGAGTCTAGACTCGTGGTGGACTTCTC

TCAATTTTCTAGGGGGATCACCCGTGTGTCTGGGCCAAAATTCGCAGTCCCCAACCTCCAATCAC

TCACCAACCTCTTGTCCTCCAATTTGTCCTGGCTATCGCTGGATGTGTCTGCGGCGTTTTATCAT

ATTCCTCTTCATCCTGCTGCTATGCCTCATCTTCTTGTTGGTTCTTCTGGACTACCAGGGTATGT

TGCCCGTTTGTCCTCTAATTCCAGGATCAACAACTACCAACACGGGACCATGCAAGACCTGCACG

ACTCCTGCTCAAGGAAACTCTATGTTTCCCTCTTGTTGCTGTACAAAACCTACCGACGGAAACTG

CACTTGTATTCCCATCCCATCATCCTGGGCTTTCGCAAAATACCTATGGGAGTGGGCCTCAGTCC

GTTTCTCCTGGCTCAGTTTACTAGTGCCATTTGTTCAGTGGTTCGTAGGGCTTTCCCCCACTGTT

TGGCTTTCCGCTATATGGATGATGTGGTATTGGGGGCCAAGTCTGTACAGCATCGTGAGTCCCTT

[0483] TATACCTCTATTACCAATTTTCTTTTGTCTTTGGGTATACATT**GAGGCTGCCGCAAAGGAAGCCG**

CAGCTAAAGAGGCAGCTGCCAAGTTCGTGTTCCTGGTTCTGCTGCCCCTGGTGTCTAGCCAGTGC

GTGAACCTGACCACCAGAACACAGCTGCCTCCAGCCTACACCAACAGCTTCACCAGAGGCGTGTA

CTACCCCGACAAGGTGTTCCGGTCCTCCGTGCTGCATTCTACCCAGGACCTGTTCCTGCCTTTCT

TCTCCAACGTGACCTGGTTCCACGCCATCCATGTGTCTGGCACCAACGGCACCAAGAGATTCGAC

AACCCCGTGCTGCCTTTCAACGACGGGGTGTACTTTGCCTCCACCGAGAAGTCCAACATCATCAG

AGGCTGGATCTTCGGCACAACCCTGGACAGCAAGACCCAGAGCCTGCTGATCGTGAACAACGCCA

CCAACGTGGTCATCAAAGTGTGCGAGTTCCAGTTCTGCAACGACCCCTTCCTGGGCGTCTACTAC

CACAAGAACAACAAGTCCTGGATGGAATCCGAGTTCCGGGTGTACTCCTCCGCCAACAACTGCAC

CTTCGAATACGTGTCCCAGCCTTTCCTGATGGACCTGGAAGGCAAGCAGGGCAACTTCAAGAACC

TGCGCGAGTTCGTGTTCAAGAACATCGACGGCTACTTCAAGATCTACTCCAAGCACACCCCTATC

AACCTCGTGCGGGATCTGCCTCAGGGCTTCTCTGCTCTGGAACCCCTGGTGGATCTGCCCATCGG

CATCAACATCACCCGGTTTCAGACCCTGCTGGCCCTGCACCGGTCTTATTTGACCCCTGGCGACT

CCTCTTCTGGCTGGACTGCTGGCGCCGCTGCTTACTATGTGGGCTACCTGCAGCCTCGGACCTTT

CTGCTGAAGTACAACGAGAATGGCACCATCACCGACGCCGTGGACTGTGCTCTGGATCCTCTGTC

CGAGACAAAGTGCACCCTGAAGTCCTTCACCGTGGAAAAGGGCATCTACCAGACCTCCAACTTCC

GGGTGCAGCCCACCGAGTCTATCGTGCGGTTCCCTAACATCACCAACCTGTGTCCTTTCGGCGAG

GTGTTCAATGCCACCAGATTCGCCTCTGTGTACGCCTGGAACCGGAAGCGGATCTCTAACTGCGT

GGCCGACTACAGCGTGCTGTACAACTCCGCCTCCTTCAGCACCTTCAAGTGCTACGGCGTGTCCC

CTACAAAGCTGAACGACCTGTGCTTCACAAACGTGTACGCCGACAGCTTCGTGATCCGGGGAGAT

GAAGTGCGGCAGATCGCTCCTGGACAGACCGGCAAGATCGCCGATTACAACTACAAGCTGCCCGA

CGACTTCACCGGCTGTGTGATCGCTTGGAACTCCAACAACCTGGACTCCAAAGTCGGCGGCAACT

ACAACTACCTGTACCGGCTGTTCCGGAAGTCTAACCTGAAGCCTTTCGAGCGGGACATCAGCACC

GAGATCTACCAGGCTGGCAGCACCCCTTGTAACGGCGTGGAAGGCTTCAACTGCTACTTCCCACT

GCAGTCCTACGGCTTTCAGCCTACCAATGGCGTGGGCTATCAGCCCTACAGAGTGGTGGTGCTGT

CCTTCGAGCTGCTGCATGCTCCTGCTACCGTGTGCGGCCCTAAGAAATCTACCAACCTGGTCAAG

AACAAATGCGTGAACTTCAACTTCAACGGCCTGACCGGCACCGGCGTGCTGACAGAGTCCAACAA

[0484]

GAAGTTCCTGCCATTCCAGCAGTTCGGCCGGGATATCGCCGATACCACAGATGCCGTCAGGGACC

CTCAGACACTGGAAATCCTGGACATCACCCCTTGCTCCTTCGGCGGAGTGTCTGTGATCACCCCA

GGCACCAACACCTCTAACCAGGTGGCCGTGCTGTATCAGGACGTGAACTGTACCGAGGTGCCCGT

GGCTATCCATGCCGATCAGCTGACCCCTACATGGCGCGTGTACTCCACCGGCTCTAACGTGTTCC

AGACAAGAGCTGGCTGTCTGATCGGCGCTGAGCACGTGAACAATTCCTACGAGTGCGACATCCCC

ATCGGAGCCGGAATCTGCGCCTCTTATCAGACCCAGACCAACTCTCCCAGACGGGCCAGATCTGT

GGCCAGCCAGTCTATCATTGCTTACACCATGAGCCTGGGCGCCGAGAACTCTGTGGCCTACAGCA

ACAACTCTATCGCTATCCCCACCAACTTCACCATCTCCGTGACCACAGAGATCCTGCCAGTGTCC

ATGACCAAGACCAGCGTGGACTGCACCATGTACATCTGCGGCGACTCTACCGAGTGCTCCAACCT

GCTGCTCCAGTACGGCTCCTTCTGCACCCAGCTGAATAGAGCCCTGACCGGAATCGCCGTGGAAC

AGGACAAGAACACCCAAGAGGTGTTCGCCCAAGTGAAGCAGATCTACAAGACCCCTCCTATCAAG

GACTTCGGCGGCTTCAATTTCTCCCAGATTCTGCCCGATCCTAGCAAGCCCTCCAAGCGGTCTTT

CATCGAGGACCTGCTGTTCAACAAAGTGACACTGGCCGACGCCGGCTTCATCAAGCAGTACGGCG

ACTGTCTGGGCGACATTGCCGCTAGGGATCTGATCTGCGCCCAGAAGTTTAACGGACTGACAGTG

CTGCCTCCTCTGCTGACCGATGAGATGATCGCCCAGTACACCTCCGCACTGCTGGCTGGCACAAT

CACCTCTGGATGGACATTTGGCGCTGGCGCTGCTCTGCAAATCCCATTCGCTATGCAAATGGCCT

ACCGGTTCAACGGCATCGGCGTGACCCAGAATGTGCTGTACGAGAACCAGAAGCTGATCGCCAAC

CAGTTCAACAGCGCCATCGGAAAGATCCAGGACAGCCTGTCCAGCACCGCTTCTGCCCTGGGAAA

GCTGCAGGATGTGGTCAACCAGAACGCTCAGGCCCTGAACACCCTCGTGAAGCAGCTGTCTAGCA

ACTTCGGCGCCATCTCCTCTGTGCTGAACGATATCCTGAGCCGGCTGGACAAGGTGGAAGCCGAG

GTGCAGATCGACAGACTGATCACCGGACGGCTGCAGTCCCTGCAGACCTATGTTACCCAGCAGCT

GATCCGGGCTGCCGAGATTAGAGCCTCTGCCAATCTGGCCGCAACCAAGATGTCTGAGTGTGTGC

TGGGACAGTCCAAGAGAGTGGACTTCTGCGGCAAGGGCTACCACCTGATGAGCTTCCCTCAGTCT

GCTCCTCACGGCGTGGTGTTTCTGCACGTGACCTACGTGCCCGCTCAAGAGAAGAACTTTACCAC

[0485] CGCTCCTGCCATCTGCCACGACGGCAAGGCTCACTTTCCTAGAGAAGGCGTGTTCGTGTCTAACG

GCACCCATTGGTTCGTGACACAGCGGAACTTCTACGAGCCCCAGATCATCACCACCGACAACACC

TTCGTGTCCGGCAACTGCGACGTCGTGATCGGAATTGTGAACAATACCGTGTACGACCCTCTGCA

GCCCGAGCTGGACTCCTTCAAAGAGGAACTGGACAAGTACTTTAAGAACCACACAAGCCCCGACG

TGGACCTGGGAGACATCTCTGGCATCAACGCCTCCGTGGTCAACATCCAGAAAGAGATCGACCGG

CTGAACGAGGTGGCCAAGAATCTGAACGAGTCCCTGATCGACCTGCAAGAACTGGGGAAGTACGA

GCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGGCTTTATCGCTGGCCTGATCGCTATCGTGA

TGGTCACAATCATGCTGTGCTGTATGACCTCCTGTTGCTCCTGCCTGAAGGGCTGCTGCTCTTGC

GGCTCTTGCTGCAAGTTCGACGAGGACGACTCTGAGCCCGTGCTGAAAGGCGTGAAGCTGCACTA

TACCTGATGACTCGAG

[0486] 在起始ATG(粗体)前添加了KOZAC序列(gcc acc,下划线的)。

[0487] 粗体和短下划线的序列对应于$(EAAAK)_3$接头

[0488] SEQ ID NO:33-对应于SEQ ID NO:32的HBsAg-$(EAAAK)_3$-2019-nCoV刺突蛋白氨基酸序列

[0489] MENITSGFLGPLLVLQAGFFLLTRILTIPQSLDSWWTSLNFLGGSPVCLGQNSQSPTSNHSPTSCPPIC
PGYRWMCLRRFIIFLFILLLCLIFLLVLLDYQGMLPVCPLIPGSTTTNTGPCKTCTTPAQGNSMFPSCCCTKPTDGN
CTCIPIPSSWAFAKYLWEWASVRFSWLSLLVPFVQWFVGLSPTVWLSAIWMMWYWGPSLYSIVSPFIPLLPIFFCLW
VYIEAAAKEAAAKEAAAKFVFLVLLPLVSSQCVNLTTRTQLPPAYTNSFTRGVYYPDKVFRSSVLHSTQDLFLPFFS
NVTWFHAIHVSGTNGTKRFDNPVLPFNDGVYFASTEKSNIIRGWIFGTTLDSKTQSLLIVNNATNVVIKVCEFQFCN
DPFLGVYYHKNNKSWMESEFRVYSSANNCTFEYVSQPFLMDLEGKQGNFKNLREFVFKNIDGYFKIYSKHTPINLVR
DLPQGFSALEPLVDLPIGINITRFQTLLALHRSYLTPGDSSSGWTAGAAAYYVGYLQPRTFLLKYNENGTITDAVDC
ALDPLSETKCTLKSFTVEKGIYQTSNFRVQPTESIVRFPNITNLCPFGEVFNATRFASVYAWNRKRISNCVADYSVL
YNSASFSTFKCYGVSPTKLNDLCFTNVYADSFVIRGDEVRQIAPGQTGKIADYNYKLPDDFTGCVIAWNSNNLDSKV
GGNYNYLYRLFRKSNLKPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQPYRVVVLSFELLHAPA

TVCGPKKSTNLVKNKCVNFNFNGLTGTGVLTESNKKFLPFQQFGRDIADTTDAVRDPQTLEILDITPCSFGGVSVIT
PGTNTSNQVAVLYQDVNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQT
QTNSPRRARSVASQSIIAYTMSLGAENSVAYSNNSIAIPTNFTISVTTEILPVSMTKTSVDCTMYICGDSTECSNLL
LQYGSFCTQLNRALTGIAVEQDKNTQEVFAQVKQIYKTPPIKDFGGFNFSQILPDPSKPSKRSFIEDLLFNKVTLAD
AGFIKQYGDCLGDIAARDLICAQKFNGLTVLPPLLTDEMIAQYTSALLAGTITSGWTFGAGAALQIPFAMQMAYRFN
GIGVTQNVLYENQKLIANQFNSAIGKIQDSLSSTASALGKLQDVVNQNAQALNTLVKQLSSNFGAISSVLNDILSRL
DKVEAEVQIDRLITGRLQSLQTYVTQQLIRAAEIRASANLAATKMSECVLGQSKRVDFCGKGYHLMSFPQSAPHGVV
FLHVTYVPAQEKNFTTAPAICHDGKAHFPREGVFVSNGTHWFVTQRNFYEPQIITTDNTFVSGNCDVVIGIVNNTVY
DPLQPELDSFKEELDKYFKNHTSPDVDLGDISGINASVVNIQKEIDRLNEVAKNLNESLIDLQELGKYEQYIKWPWY
IWLGFIAGLIAIVMVTIMLCCMTSCCSCLKGCCSCGSCCKFDEDDSEPVLKGVKLHYT

[0490] $(EAAAK)_3$-接头是带下划线的。

序列表

<110> 斯皮科纳股份有限公司

<120> SARS-COV-2 与流感联合疫苗

<130> P67254WO

<150> GB 2010425.3

<151> 2020-07-07

<160> 33

<170> PatentIn version 3.5

<210> 1

<211> 1273

<212> PRT

<213> SARS-CoV-2

<400> 1

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

195 200 205

Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu

210 215 220

Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr

225 230 235 240

Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser

245 250 255

Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro

260 265 270

Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala

275 280 285

Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys

290 295 300

Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val

305 310 315 320

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

325 330 335

Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala

340 345 350

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

355 360 365

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

370 375 380

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

385 390 395 400

Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly

405 410 415

Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys

420 425 430

Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn

435 440 445

Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe

450 455 460

Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys

465 470 475 480

Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly

485 490 495

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

500 505 510

Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
515 520 525
Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn
530 535 540
Gly Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu
545 550 555 560
Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val
565 570 575
Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe
580 585 590
Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val
595 600 605
Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile
610 615 620
His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly Ser
625 630 635 640
Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val
645 650 655
Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala
660 665 670
Ser Tyr Gln Thr Gln Thr Asn Ser Pro Arg Arg Ala Arg Ser Val Ala
675 680 685
Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Ala Glu Asn Ser
690 695 700
Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr Asn Phe Thr Ile
705 710 715 720
Ser Val Thr Thr Glu Ile Leu Pro Val Ser Met Thr Lys Thr Ser Val
725 730 735
Asp Cys Thr Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ser Asn Leu
740 745 750
Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Thr
755 760 765
Gly Ile Ala Val Glu Gln Asp Lys Asn Thr Gln Glu Val Phe Ala Gln
770 775 780
Val Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly Phe
785 790 795 800
Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg Ser
805 810 815
Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly

820 825 830

Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala Arg Asp

835 840 845

Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu

850 855 860

Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser Ala Leu Leu Ala Gly

865 870 875 880

Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu Gln Ile

885 890 895

Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile Gly Val Thr

900 905 910

Gln Asn Val Leu Tyr Glu Asn Gln Lys Leu Ile Ala Asn Gln Phe Asn

915 920 925

Ser Ala Ile Gly Lys Ile Gln Asp Ser Leu Ser Ser Thr Ala Ser Ala

930 935 940

Leu Gly Lys Leu Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn

945 950 955 960

Thr Leu Val Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val

965 970 975

Leu Asn Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln

980 985 990

Ile Asp Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val

995 1000 1005

Thr Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn

1010 1015 1020

Leu Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys

1025 1030 1035

Arg Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro

1040 1045 1050

Gln Ser Ala Pro His Gly Val Val Phe Leu His Val Thr Tyr Val

1055 1060 1065

Pro Ala Gln Glu Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His

1070 1075 1080

Asp Gly Lys Ala His Phe Pro Arg Glu Gly Val Phe Val Ser Asn

1085 1090 1095

Gly Thr His Trp Phe Val Thr Gln Arg Asn Phe Tyr Glu Pro Gln

1100 1105 1110

Ile Ile Thr Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val

1115 1120 1125

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

<210> 2
<211> 3839
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> SARS-CoV2刺突蛋白核酸-针对在大肠杆菌(E. coli)中的表达优化
<400> 2

gagctcatgt ttgtttttct ggttctgctg ccgctggtta gcagccagtg tgttaatctg 60
accacacgta cccagctgcc tccggcatat accaatagct ttacccgtgg tgtttattat 120
ccggacaaag tttttcgtag cagcgttctg catagcaccc aggacctgtt tctgccgttt 180
tttagcaatg ttacctggtt tcatgccatt catgttagcg gcaccaatgg caccaaacgt 240
tttgataatc cggtgctgcc gtttaatgat ggtgtgtatt ttgcaagcac cgaaaaaagc 300
aacattattc gcggttggat ttttggtaca accctggata gcaaaaccca gagcctgctg 360
attgttaata atgccaccaa tgtggtgatc aaagtgtgcg aatttcagtt ttgcaatgat 420
ccgtttctgg gcgtgtatta ccacaaaaat aacaagagct ggatggaaag cgaatttcgt 480
gtttatagca gcgccaataa ttgcaccttt gaatatgtta gccagccgtt tctgatggat 540
ctggaaggta aacagggtaa ctttaaaaac ctgcgcgagt tcgtgttcaa aaacatcgat 600
ggttacttca aaatctatag caaacacacc ccgattaatc tggttcgtga tctgccgcag 660
ggttttagcg cactggaacc gctggttgat ctgccaattg gtattaacat tacccgtttt 720

cagaccctgc tggcactgca tcgtagctat ctgacaccgg gtgatagcag cagcggttgg 780
accgcaggcg cagcagcata ttatgttggt tatctgcagc ctcgtacctt tctgctgaaa 840
tataacgaaa atggcacaat taccgatgcc gttgattgtg ccctggatcc gctgagcgaa 900
accaaatgta ccctgaaaag ctttaccgtt gagaaaggta tttatcagac cagcaatttt 960
cgtgtgcagc cgaccgaaag cattgttcgt tttccgaata tcaccaatct gtgtccgttt 1020
ggcgaagttt ttaatgcaac ccgttttgcc agcgtttatg catggaatcg taaacgtatt 1080
agcaattgcg ttgccgatta tagcgttctg tataatagcg caagcttcag cacctttaaa 1140
tgctatggtg ttagcccgac caaactgaat gatctgtgtt ttaccaatgt gtatgccgat 1200
agctttgtga ttcgtggtga tgaagttcgt cagattgcac cgggtcagac cggtaaaatt 1260
gcagattata actataaact gccggatgat tttacgggtt gtgttattgc ctggaatagc 1320
aataatctgg acagcaaagt tggtggcaac tataactatc tgtatcgcct gtttcgtaag 1380
agcaatctga aaccgtttga acgtgatatt agcaccgaga tttatcaggc aggtagcacc 1440
ccgtgtaatg gtgttgaagg ttttaattgc tattttccgc tgcagagcta tggttttcag 1500
ccgacaaatg gtgtgggtta tcagccgtat cgtgttgttg ttctgtcatt tgaactgctg 1560
catgcaccgg caaccgtttg tggtccgaaa aaaagtacca atctggtgaa aaataagtgc 1620
gtgaacttta actttaatgg tctgaccggc accggtgttc tgaccgaaag taacaaaaaa 1680
ttcctgccgt ttcagcagtt tggccgtgat attgcagata ccaccgatgc agttcgcgat 1740
ccgcagacac tggaaattct ggatattacc ccgtgcagct ttggtggtgt ttcagttatt 1800
acaccgggta caaataccag caatcaggtt gcagttctgt atcaggatgt taattgtacc 1860
gaagttccgg ttgcaattca tgcagatcag ctgaccccga cctggcgtgt gtatagcacc 1920
ggtagcaatg tgtttcagac acgtgcaggt tgtctgattg gtgcagaaca tgtgaataat 1980
agctatgaat gcgatattcc gattggtgcg ggtatttgtg ccagctatca gacccagacc 2040
aatagtccgc gtcgtgcacg tagcgttgca agccagagca ttattgccta taccatgagc 2100
ctgggtgcag aaaatagcgt tgcctatagt aataacagca ttgccattcc gaccaacttt 2160
accattagcg ttaccaccga aattctgccg gttagcatga ccaaaaccag cgttgattgc 2220
accatgtata tttgtggtga tagtaccgaa tgtagcaatc tgctgctgca gtatggtagc 2280
ttttgcaccc agctgaatcg tgcactgacc ggtattgcag ttgaacagga taaaaacacg 2340
caagaagttt ttgcacaggt caagcagatc tataaaaccc ctccgattaa agattttggc 2400
ggtttcaatt ttagccagat cctgccggat ccgagcaaac cgagtaaacg tagctttatt 2460
gaagatctgc tgttcaacaa agtgaccctg gcagatgcag gttttatcaa acagtatggt 2520
gattgcctgg gcgatattgc cgcacgtgat ctgatttgtg cacagaaatt taacggcctg 2580
accgttctgc ctccgctgct gaccgatgaa atgattgcac agtataccag cgcactgctg 2640
gcaggcacca ttaccagtgg ttggaccttt ggtgccggtg ccgcactgca gattccgttt 2700
gcaatgcaga tggcatatcg ttttaatggt attggtgtta cccagaacgt gctgtatgaa 2760
aaccagaaac tgattgccaa ccagtttaat agcgccattg gcaaaattca ggatagcctg 2820
agcagcaccg caagtgcact gggtaaactg caggacgttg ttaatcagaa tgcacaggca 2880
ctgaataccc tggttaaaca gctgagcagt aattttggtg caatttcaag cgtgctgaac 2940
gatattctga gccgtctgga taaagttgaa gcagaagttc agattgatcg tctgattacc 3000
ggtcgtctgc aaagcctgca gacctatgtg acccagcagc tgattcgcgc agcagaaatt 3060

cgtgcaagcg caaatctggc agccaccaaa atgagcgaat gtgttctggg tcagagcaaa 3120
cgtgttgatt tttgcggcaa aggttatcac ctgatgagct ttccgcagag cgcaccgcat 3180
ggtgttgtgt ttctgcatgt tacctatgtt ccggcacaag aaaaaaactt tacaaccgct 3240
ccggcaattt gccatgatgg taaagcacat tttccgcgtg aaggtgtttt tgttagtaat 3300
ggcacccatt ggtttgttac acagcgcaac ttttatgaac cgcagattat tacaaccgac 3360
aacacctttg ttagcggtaa ctgtgatgtt gtgattggca ttgtgaataa caccgtttat 3420
gatccactgc agccggaact ggatagcttt aaagaagaac tggacaaata tttcaaaaac 3480
cacaccagtc cggatgttga tctgggtgat atttcaggta ttaatgccag cgtggtgaac 3540
atccagaaag aaattgatcg cctgaatgaa gtggccaaaa atctgaatga aagcctgatt 3600
gatctgcaag aactggggaa atatgagcag tatatcaaat ggccgtggta tatttggctg 3660
ggttttattg caggcctgat tgcaattgtt atggtgacca ttatgctgtg ttgtatgacc 3720
agctgttgta gctgtctgaa aggttgttgc agctgcggta gctgttgcaa atttgatgaa 3780
gatgatagcg aaccggtgct gaaaggtgtt aaactgcatt atacctaatg agcggccgc 3839

<210> 3

<211> 4556

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 编码融合蛋白HEV-SARS-CoV-2刺突蛋白的核酸-针对在大肠杆菌(E. coli)中的表达优化

<400> 3

gagctcatga ttgcactgac cctgtttaat ctggcagata ccctgttagg tggtctgccg 60
accgaactga ttagcagtgc cggtggtcag ctgttttata gccgtccggt tgttagcgca 120
aatggtgaac cgaccgttaa actgtatacc agcgttgaaa atgcacagca ggataaaggt 180
attgcaattc cgcatgatat tgatctgggt gaaagccgtg ttgtgattca ggattatgat 240
aatcagcatg aacaggatcg tccgacaccg agtccggcac cgagccgtcc gtttagcgtt 300
ctgcgtgcaa atgatgttct gtggctgagc ctgaccgcag cagaatatga tcagagcacc 360
tatggtagca gcaccggtcc ggtttatgtt agcgatagcg ttaccctggt taatgttgca 420
accggtgcac aggcagttgc acgtagcctg gattggacca aagtgaccct ggatggtcgt 480
ccgctgagca ccattcagca gtatagcaaa acctttttg ttctgccgct gcgtggtaaa 540
ctgagctttt gggaagcagg caccaccaaa gcaggttatc cgtataacta taataccacc 600
gcaagcgatc agctgctggt tgaaaacgca gcaggtcatc gtgttgcaat tagcacctat 660
accaccagtt taggtgcagg tccggttagc attagcgcag ttgcagttct ggcaccgcat 720
agcgcatttg tttttctggt tctgctgccg ctggttagca gccagtgtgt taatctgacc 780
acacgtaccc agctgcctcc ggcatatacc aatagcttta cccgtggtgt ttattatccg 840
gacaaagttt ttcgtagcag cgttctgcat agcacccagg acctgtttct gccgtttttt 900
agcaatgtta cctggtttca tgccattcat gttagcggca ccaatggcac caaacgtttt 960
gataatccgg tgctgccgtt taatgatggt gtgtattttg caagcaccga aaaaagcaac 1020
attattcgcg gttggatttt tggtacaacc ctggatagca aaacccagag cctgctgatt 1080

gttaataatg ccaccaatgt ggtgatcaaa gtgtgcgaat ttcagttttg caatgatccg 1140
tttctgggcg tgtattacca caaaaataac aagagctgga tggaaagcga atttcgtgtt 1200
tatagcagcg ccaataattg cacctttgaa tatgttagcc agccgtttct gatggatctg 1260
gaaggtaaac agggtaactt taaaaacctg cgcgagttcg tgttcaaaaa catcgatggt 1320
tacttcaaaa tctatagcaa acacaccccg attaatctgg ttcgtgatct gccgcagggt 1380
tttagcgcac tggaaccgct ggttgatctg ccaattggta ttaacattac ccgttttcag 1440
accctgctgg cactgcatcg tagctatctg acaccgggtg atagcagcag cggttggacc 1500
gcaggcgcag cagcatatta tgttggttat ctgcagcctc gtacctttct gctgaaatat 1560
aacgaaaatg gcacaattac cgatgccgtt gattgtgccc tggatccgct gagcgaaacc 1620
aaatgtaccc tgaaaagctt taccgttgag aaaggtattt atcagaccag caattttcgt 1680
gtgcagccga ccgaaagcat tgttcgtttt ccgaatatca ccaatctgtg tccgtttggc 1740
gaagttttta atgcaacccg ttttgccagc gtttatgcat ggaatcgtaa acgtattagc 1800
aattgcgttg ccgattatag cgttctgtat aatagcgcaa gcttcagcac ctttaaatgc 1860
tatggtgtta gcccgaccaa actgaatgat ctgtgtttta ccaatgtgta tgccgatagc 1920
tttgtgattc gtggtgatga agttcgtcag attgcaccgg gtcagaccgg taaaattgca 1980
gattataact ataaactgcc ggatgatttt acgggttgtg ttattgcctg gaatagcaat 2040
aatctggaca gcaaagttgg tggcaactat aactatctgt atcgcctgtt tcgtaagagc 2100
aatctgaaac cgtttgaacg tgatattagc accgagattt atcaggcagg tagcaccccg 2160
tgtaatggtg ttgaaggttt taattgctat tttccgctgc agagctatgg ttttcagccg 2220
acaaatggtg tgggttatca gccgtatcgt gttgttgttc tgtcatttga actgctgcat 2280
gcaccggcaa ccgtttgtgg tccgaaaaaa agtaccaatc tggtgaaaaa taagtgcgtg 2340
aactttaact ttaatggtct gaccggcacc ggtgttctga ccgaaagtaa caaaaaattc 2400
ctgccgtttc agcagtttgg ccgtgatatt gcagatacca ccgatgcagt tcgcgatccg 2460
cagacactgg aaattctgga tattaccccg tgcagctttg gtggtgtttc agttattaca 2520
ccgggtacaa ataccagcaa tcaggttgca gttctgtatc aggatgttaa ttgtaccgaa 2580
gttccggttg caattcatgc agatcagctg accccgacct ggcgtgtgta tagcaccggt 2640
agcaatgtgt ttcagacacg tgcaggttgt ctgattggtg cagaacatgt gaataatagc 2700
tatgaatgcg atattccgat tggtgcgggt atttgtgcca gctatcagac ccagaccaat 2760
agtccgcgtc gtgcacgtag cgttgcaagc cagagcatta ttgcctatac catgagcctg 2820
ggtgcagaaa atagcgttgc ctatagtaat aacagcattg ccattccgac caactttacc 2880
attagcgtta ccaccgaaat tctgccggtt agcatgacca aaaccagcgt tgattgcacc 2940
atgtatattt gtggtgatag taccgaatgt agcaatctgc tgctgcagta tggtagcttt 3000
tgcacccagc tgaatcgtgc actgaccggt attgcagttg aacaggataa aaacacgcaa 3060
gaagtttttg cacaggtcaa gcagatctat aaaaccccctc cgattaaaga ttttggcggt 3120
ttcaatttta gccagatcct gccggatccg agcaaaccga gtaaacgtag ctttattgaa 3180
gatctgctgt tcaacaaagt gaccctggca gatgcaggtt ttatcaaaca gtatggtgat 3240
tgcctgggcg atattgccgc acgtgatctg atttgtgcac agaaatttaa cggcctgacc 3300
gttctgcctc cgctgctgac cgatgaaatg attgcacagt ataccagcgc actgctggca 3360
ggcaccatta ccagtggttg gacctttggt gccggtgccg cactgcagat tccgtttgca 3420

atgcagatgg catatcgttt taatggtatt ggtgttaccc agaacgtgct gtatgaaaac 3480
cagaaactga ttgccaacca gtttaatagc gccattggca aaattcagga tagcctgagc 3540
agcaccgcaa gtgcactggg taaactgcag gacgttgtta atcagaatgc acaggcactg 3600
aataccctgg ttaaacagct gagcagtaat tttggtgcaa tttcaagcgt gctgaacgat 3660
attctgagcc gtctggataa agttgaagca gaagttcaga ttgatcgtct gattaccggt 3720
cgtctgcaaa gcctgcagac ctatgtgacc cagcagctga ttcgcgcagc agaaattcgt 3780
gcaagcgcaa atctggcagc caccaaaatg agcgaatgtg ttctgggtca gagcaaacgt 3840
gttgattttt gcggcaaagg ttatcacctg atgagctttc cgcagagcgc accgcatggt 3900
gttgtgtttc tgcatgttac ctatgttccg gcacaagaaa aaaactttac aaccgctccg 3960
gcaatttgcc atgatggtaa agcacatttt ccgcgtgaag gtgtttttgt tagtaatggc 4020
acccattggt ttgttacaca gcgcaacttt tatgaaccgc agattattac aaccgacaac 4080
acctttgtta gcggtaactg tgatgttgtg attggcattg tgaataacac cgtttatgat 4140
ccactgcagc cggaactgga tagctttaaa gaagaactgg acaaatattt caaaaaccac 4200
accagtccgg atgttgatct gggtgatatt tcaggtatta atgccagcgt ggtgaacatc 4260
cagaaagaaa ttgatcgcct gaatgaagtg gccaaaaatc tgaatgaaag cctgattgat 4320
ctgcaagaac tggggaaata tgagcagtat atcaaatggc cgtggtatat ttggctgggt 4380
tttattgcag gcctgattgc aattgttatg gtgaccatta tgctgtgttg tatgaccagc 4440
tgttgtagct gtctgaaagg ttgttgcagc tgcggtagct gttgcaaatt tgatgaagat 4500
gatagcgaac cggtgctgaa aggtgttaaa ctgcattata cctaatgagc ggccgc 4556

<210> 4
<211> 3839
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> SARS-CoV-2刺突蛋白核酸序列- 针对在Komagataella pastoris中的表达优化

<400> 4
ttcgaaacga tgttcgtgtt cttggtcctg ttgccattgg tttcttccca gtgtgttaac 60
ctgaccacta gaactcaatt gcctccagcc tacaccaatt ccttcaccag aggtgtttac 120
tacccagaca aggtgttcag atcttccgtc ttgcactcca ctcaggactt gttcttgcca 180
ttcttctcca acgttacctg gttccacgct attcacgttt ccggaactaa cggtactaag 240
agattcgaca acccagtcct gccattcaac gatggtgtct acttcgcttc taccgagaag 300
tccaacatca tcagaggttg gatcttcggt actaccctgg actctaagac tcagtccttg 360
ctgatcgtta acaacgccac caacgttgtc atcaaggttt gcgagttcca gttctgcaac 420
gacccattct tgggtgtgta ctaccacaag aacaacaagt cttggatgga atccgagttc 480
agagtttact cctccgccaa caactgtacc ttcgagtacg tttcccagcc attcttgatg 540
gacttggagg gtaagcaggg taacttcaag aacctgagag agttcgtttt caagaacatc 600
gacggttact tcaagatcta ctccaagcac accccaatca acctggttag agatttgcca 660
caaggtttct ccgctttgga gcctttggtt gacttgccaa tcggtatcaa catcaccaga 720

ttccagacct tgttggcctt gcacagatcc tacttgactc caggtgattc ttcttccggt 780
tggactgctg gtgctgctgc ttactatgtt ggttacttgc agccaagaac cttcctgctg 840
aagtacaacg agaacggaac tatcactgac gctgttgact gtgctttgga cccattgtct 900
gagactaagt gcaccttgaa gtccttcacc gttgagaagg gtatctacca gacctccaac 960
ttcagagttc agccaactga gtccatcgtc agattcccaa acatcactaa cttgtgccca 1020
ttcggtgagg tgttcaacgc tactagattc gcttctgttt acgcctggaa cagaaagaga 1080
atctccaact gcgttgctga ctactccgtc ttgtacaact ctgcttcatt ctccaccttc 1140
aagtgctacg gtgtttcccc aactaagttg aacgacctgt gtttcactaa cgtctacgcc 1200
gactccttcg ttattagagg tgacgaggtt agacagatcg ctccaggtca aactggtaag 1260
atcgctgact acaactacaa gctgccagac gacttcaccg gttgtgttat tgcttggaac 1320
tccaacaacc tggactccaa ggttggtggt aactacaatt acctgtaccg tctgttcaga 1380
aagtccaact tgaagccatt cgagagagac atctccaccg agatctacca agctggttct 1440
actccatgta acggtgtcga gggtttcaac tgctacttcc cattgcaatc ctacggtttc 1500
caacctacca acggtgttgg ataccagcca tacagagttg tcgttttgtc cttcgagttg 1560
ttgcacgctc cagctactgt ttgtggtcca aagaagtcca ccaacttggt caagaacaaa 1620
tgcgtcaact ttaacttcaa cggcctgacc ggtactggtg ttttgactga atccaacaag 1680
aagttcctgc ctttccagca gttcggtaga gacattgctg acactactga cgccgttaga 1740
gatccacaga ctttggagat cttggacatc accccatgtt ccttcggtgg tgtttccgtt 1800
attacccctg gaactaacac ctccaatcag gtcgctgtct tgtaccagga cgttaactgt 1860
actgaggttc cagttgctat ccacgctgac caattgactc caacttggag agtctactcc 1920
accggttcca acgttttcca aactagagcc ggttgtttga tcggtgctga acacgtcaac 1980
aactcctacg agtgtgacat tccaattggt gctggtatct gtgcctccta ccaaactcaa 2040
actaactccc caagaagggc tagatccgtt gcttcccaat ccattatcgc ttacaccatg 2100
tctttgggtg ccgagaactc tgttgcctac tctaacaact ctatcgctat ccctaccaac 2160
ttcaccatct ccgttaccac tgagatcttg ccagtctcca tgaccaagac ttccgttgac 2220
tgtaccatgt acatctgtgg tgactccact gagtgttcca acttgttgct gcaatacggt 2280
tccttctgca cccagttgaa cagagctttg actggtattg ctgtcgagca agacaagaac 2340
actcaagagg ttttcgccca ggtgaagcag atctacaaga ctccacctat taaggacttc 2400
ggtggcttca acttctccca gattttgcca gatccatcta agccctccaa gagatccttc 2460
attgaggacc tgctgttcaa caaggttact ttggctgacg ccggtttcat caagcagtac 2520
ggtgattgct tgggtgacat tgcagctaga gacttgatct gtgcccagaa gttcaacggt 2580
ttgaccgttt tgccaccttt gttgaccgac gagatgatcg ctcagtacac ttctgctttg 2640
ttggccggta ctatcacttc tggttggaca tttggagctg gtgccgcatt gcaaattcca 2700
ttcgctatgc aaatggccta cagattcaac ggtatcggtg ttacccagaa cgtcctgtac 2760
gagaaccaga agcttatcgc caaccagttc aactccgcta tcggtaagat tcaggactcc 2820
ttgtcctcta ctgcttctgc cttgggaaag ttgcaggatg ttgttaacca gaatgcccag 2880
gctttgaaca ccctggttaa gcaactgtcc tctaacttcg gtgctatctc ctccgttttg 2940
aacgacatct tgtcccgttt ggacaaggtt gaggctgagg ttcagatcga cagattgatc 3000
actggtagat tgcagtccct gcagacttac gttactcagc agttgattag agctgccgag 3060

attagagcct ctgctaactt ggctgctact aagatgtccg agtgtgtttt gggtcagtcc 3120
aagagagttg acttctgcgg taagggttac cacctgatgt ctttcccaca atctgctcca 3180
cacggtgtcg ttttcttgca cgttacttac gttccagctc aagagaagaa cttcactact 3240
gctccagcca tttgtcacga tggtaaggct cactttcctc gtgagggtgt tttcgtttcc 3300
aacggtactc actggttcgt cacccagaga aacttttacg agccacagat catcaccacc 3360
gacaacactt tcgtttctgg taactgtgac gtcgtcatcg gtatcgtgaa caacactgtc 3420
tacgatccat tgcagccaga attggactcc ttcaaagagg aactggacaa gtactttaag 3480
aaccacactt ccccagacgt tgacctgggt gatatttccg gtattaacgc ctccgttgtc 3540
aacatccaaa aagagatcga ccgtttgaac gaggtcgcca agaacttgaa cgagtccttg 3600
attgacttgc aagagctggg caagtacgag cagtacatta agtggccatg gtacatttgg 3660
ctgggtttca ttgctggttt gatcgccatc gttatggtca ccatcatgtt gtgctgtatg 3720
acctcctgtt gctcctgttt gaagggttgt tgttcctgcg gttcctgttg taagttcgac 3780
gaagatgact ccgagccagt cttgaagggt gttaagttgc actacactta agcggccgc 3839

<210> 5

<211> 5357

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> 编码融合蛋白HPV18L1/SARS-CoV-2刺突蛋白的核酸-针对在K. pastoris中的表达优化

<400> 5

ttcgaaacga tggctctttg gagaccatcc gacaacactg tttacttgcc accaccatcc 60
gttgctagag ttgttaacac tgacgactac gttactagaa cttccatctt ctaccacgct 120
ggttcttcca gattgttgac tgttggtaac ccatacttca gagttccagc tggaggtggt 180
aacaagcaag acatcccaaa ggtttccgct taccagtaca gagttttcag agttcagttg 240
ccagacccaa acaagtttgg attgccagac acttccatct acaacccaga gactcagaga 300
cttgtttggg cttgtgctgg tgttgaaatc ggtagaggac agccattggg tgttggtttg 360
tctggtcacc cattctacaa caagttggac gacactgaat cttctcacgc tgctacttct 420
aacgtttccg aggatgttag agacaacgtt tccgttgact acaagcagac tcagttgtgt 480
atcttgggtt gtgctccagc tattggtgaa cattgggcta agggtactgc ttgtaagtcc 540
agaccattgt ctcagggaga ttgtccacca ttggagttga agaacactgt tttggaggac 600
ggtgatatgg ttgatactgg ttacggtgct atggacttct ctactttgca ggacactaag 660
tgtgaagttc cattggacat ctgtcagtcc atctgtaagt acccagacta cttgcaaatg 720
tccgctgatc catacggtga ctctatgttc ttctgtttga gaagagagca gttgttcgct 780
agacacttct ggaacagagc tggtactatg ggtgacactg ttccacaatc cttgtacatc 840
aagggtactg gaatgagagc ttctcctggt tcttgtgttt actctccatc tccatccggt 900
tccattgtta cttccgactc ccagttgttc aacaagccat actggttgca taaggctcaa 960
ggtcacaaca acggtgtttg ttggcacaac cagttgttcg ttactgttgt tgacactact 1020
agatccacta acttgactat ctgtgcttcc actcaatctc cagttccagg acaatacgac 1080

gctactaagt tcaagcagta ctccagacac gttgaagagt acgacttgca gttcatcttc 1140
cagttgtgta ctatcacttt gactgctgat gttatgtcct acatccactc tatgaactcc 1200
tccattttgg aggattggaa cttcggtgtt ccaccaccac caactacttc attggttgac 1260
acttacagat tcgttcagtc cgttgctatc acttgtcaaa aggacgctgc tccagctgaa 1320
aacaaggacc catacgacaa gttgaagttc tggaacgttg acttgaaaga gaagttctcc 1380
ttggacttgg accaataccc attgggtaga aagtttttgg ttcaggctgg attgagaaga 1440
aagccaacta tcggtccaag aaagagatca gctccatccg ctactacttc atccaagcca 1500
gctaagagag ttagagttag agctagaaag ttcgtgttct tggtcctgtt gccattggtt 1560
tcttcccagt gtgttaacct gaccactaga actcaattgc ctccagccta caccaattcc 1620
ttcaccagag gtgtttacta cccagacaag gtgttcagat cttccgtctt gcactccact 1680
caggacttgt tcttgccatt cttctccaac gttacctggt tccacgctat tcacgtttcc 1740
ggaactaacg gtactaagag attcgacaac ccagtcctgc cattcaacga tggtgtctac 1800
ttcgcttcta ccgagaagtc caacatcatc agaggttgga tcttcggtac taccctggac 1860
tctaagactc agtccttgct gatcgttaac aacgccacca acgttgtcat caaggtttgc 1920
gagttccagt tctgcaacga cccattcttg ggtgtgtact accacaagaa caacaagtct 1980
tggatggaat ccgagttcag agtttactcc tccgccaaca actgtacctt cgagtacgtt 2040
tcccagccat tcttgatgga cttggagggt aagcagggta acttcaagaa cctgagagag 2100
ttcgttttca agaacatcga cggttacttc aagatctact ccaagcacac cccaatcaac 2160
ctggttagag atttgccaca aggtttctcc gctttggagc ctttggttga cttgccaatc 2220
ggtatcaaca tcaccagatt ccagaccttg ttggccttgc acagatccta cttgactcca 2280
ggtgattctt cttccggttg gactgctggt gctgctgctt actatgttgg ttacttgcag 2340
ccaagaacct tcctgctgaa gtacaacgag aacggaacta tcactgacgc tgttgactgt 2400
gctttggacc cattgtctga gactaagtgc accttgaagt ccttcaccgt tgagaagggt 2460
atctaccaga cctccaactt cagagttcag ccaactgagt ccatcgtcag attcccaaac 2520
atcactaact tgtgcccatt cggtgaggtg ttcaacgcta ctagattcgc ttctgtttac 2580
gcctggaaca gaaagagaat ctccaactgc gttgctgact actccgtctt gtacaactct 2640
gcttcattct ccaccttcaa gtgctacggt gtttccccaa ctaagttgaa cgacctgtgt 2700
ttcactaacg tctacgccga ctccttcgtt attagaggtg acgaggttag acagatcgct 2760
ccaggtcaaa ctggtaagat cgctgactac aactacaagc tgccagacga cttcaccggt 2820
tgtgttattg cttggaactc caacaacctg gactccaagg ttggtggtaa ctacaattac 2880
ctgtaccgtc tgttcagaaa gtccaacttg aagccattcg agagagacat ctccaccgag 2940
atctaccaag ctggttctac tccatgtaac ggtgtcgagg gtttcaactg ctacttccca 3000
ttgcaatcct acggtttcca acctaccaac ggtgttggat accagccata cagagttgtc 3060
gttttgtcct tcgagttgtt gcacgctcca gctactgttt gtggtccaaa gaagtccacc 3120
aacttggtca agaacaaatg cgtcaacttt aacttcaacg gcctgaccgg tactggtgtt 3180
ttgactgaat ccaacaagaa gttcctgcct ttccagcagt tcggtagaga cattgctgac 3240
actactgacg ccgttagaga tccacagact ttggagatct tggacatcac cccatgttcc 3300
ttcggtggtg tttccgttat taccctgga actaacacct ccaatcaggt cgctgtcttg 3360
taccaggacg ttaactgtac tgaggttcca gttgctatcc acgctgacca attgactcca 3420

acttggagag tctactccac cggttccaac gttttccaaa ctagagccgg ttgtttgatc 3480
ggtgctgaac acgtcaacaa ctcctacgag tgtgacattc caattggtgc tggtatctgt 3540
gcctcctacc aaactcaaac taactcccca agaagggcta gatccgttgc ttcccaatcc 3600
attatcgctt acaccatgtc tttgggtgcc gagaactctg ttgcctactc taacaactct 3660
atcgctatcc ctaccaactt caccatctcc gttaccactg agatcttgcc agtctccatg 3720
accaagactt ccgttgactg taccatgtac atctgtggtg actccactga gtgttccaac 3780
ttgttgctgc aatacggttc cttctgcacc cagttgaaca gagctttgac tggtattgct 3840
gtcgagcaag acaagaacac tcaagaggtt ttcgcccagg tgaagcagat ctacaagact 3900
ccacctatta aggacttcgg tggcttcaac ttctcccaga ttttgccaga tccatctaag 3960
ccctccaaga gatccttcat tgaggacctg ctgttcaaca aggttacttt ggctgacgcc 4020
ggtttcatca agcagtacgg tgattgcttg ggtgacattg cagctagaga cttgatctgt 4080
gcccagaagt tcaacggttt gaccgttttg ccacctttgt tgaccgacga gatgatcgct 4140
cagtacactt ctgctttgtt ggccggtact atcacttctg gttggacatt tggagctggt 4200
gccgcattgc aaattccatt cgctatgcaa atggcctaca gattcaacgg tatcggtgtt 4260
acccagaacg tcctgtacga gaaccagaag cttatcgcca accagttcaa ctccgctatc 4320
ggtaagattc aggactcctt gtcctctact gcttctgcct tgggaaagtt gcaggatgtt 4380
gttaaccaga atgcccaggc tttgaacacc ctggttaagc aactgtcctc taacttcggt 4440
gctatctcct ccgttttgaa cgacatcttg tcccgtttgg acaaggttga ggctgaggtt 4500
cagatcgaca gattgatcac tggtagattg cagtccctgc agacttacgt tactcagcag 4560
ttgattagag ctgccgagat tagagcctct gctaacttgg ctgctactaa gatgtccgag 4620
tgtgttttgg gtcagtccaa gagagttgac ttctgcggta agggttacca cctgatgtct 4680
ttcccacaat ctgctccaca cggtgtcgtt ttcttgcacg ttacttacgt tccagctcaa 4740
gagaagaact tcactactgc tccagccatt tgtcacgatg gtaaggctca ctttcctcgt 4800
gagggtgttt tcgtttccaa cggtactcac tggttcgtca cccagagaaa cttttacgag 4860
ccacagatca tcaccaccga caacactttc gtttctggta actgtgacgt cgtcatcggt 4920
atcgtgaaca acactgtcta cgatccattg cagccagaat tggactcctt caaagaggaa 4980
ctggacaagt actttaagaa ccacacttcc ccagacgttg acctgggtga tatttccggt 5040
attaacgcct ccgttgtcaa catccaaaaa gagatcgacc gtttgaacga ggtcgccaag 5100
aacttgaacg agtccttgat tgacttgcaa gagctgggca agtacgagca gtacattaag 5160
tggccatggt acatttggct gggtttcatt gctggtttga tcgccatcgt tatggtcacc 5220
atcatgttgt gctgtatgac ctcctgttgc tcctgtttga agggttgttg ttcctgcggt 5280
tcctgttgta agttcgacga agatgactcc gagccagtct tgaagggtgt taagttgcac 5340
tacacttaag cggccgc 5357

<210> 6
<211> 5351
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 编码融合蛋白HPV16L1/SARS-CoV-2刺突蛋白的核酸-针对在K. pastoris中的

表达优化

〈400〉 6

ttcgaaacga tgtctttgtg gttgccatct gaagctactg tttacttgcc accagttcca 60
gtttctaaag ttgtttccac tgacgaatac gttgctagaa ctaacatcta ctaccacgct 120
ggtacttcta gattgttggc tgttggtcat ccatacttcc caattaagaa gccaaacaac 180
aacaagattt tggttccaaa ggtttccgga ttgcaataca gagttttcag aatccatttg 240
ccagatccaa acaagtttgg tttcccagat acttctttct acaacccaga cactcaaaga 300
cttgtttggg cttgtgttgg tgttgaagtt ggtagaggtc aaccattggg tgttggtatt 360
tctggtcacc cattgttgaa caagttggac gatactgaaa acgcttctgc ttacgctgct 420
aacgctggtg ttgataacag agaatgtatt tctatggact acaagcaaac tcaattgtgt 480
ttgattggtt gtaagccacc aattggtgaa cattggggaa agggttctcc atgtactaat 540
gttgctgtta accctggtga ttgtccacca ttggaattga ttaacactgt tattcaagac 600
ggtgatatgg ttgatactgg tttcggtgct atggatttca ctactttgca agctaacaag 660
tctgaagttc cattggacat ttgtacttcc atctgtaagt acccagacta cattaagatg 720
gtttctgaac catacggtga ttctttgttc ttctacttga gaagagaaca aatgtttgtt 780
agacacttgt tcaacagagc tggtgctgtt ggtgaaaacg ttccagatga cttgtacatt 840
aagggttctg gttctactgc taacttggct tcttctaact actttccaac tccatctggt 900
tctatggtta cttctgacgc tcaaattttc aacaagccat actggttgca aagagcacaa 960
ggtcataaca acggtatttg ttggggtaac caattgttcg ttactgttgt tgacactact 1020
agatccacta acatgtcctt gtgtgctgct atttctactt ctgaaactac ttacaagaac 1080
actaacttca aagagtactt gagacacgga gaagaatacg acttgcaatt cattttccaa 1140
ttgtgtaaga ttactttgac tgctgacgtt atgacttaca ttcactctat gaactctact 1200
attttggaag attggaactt cggattgcaa ccaccaccag gtggtacttt ggaagatact 1260
tacagattcg ttacttctca agctattgct tgtcaaaagc atactccacc tgctccaaaa 1320
gaagatccat tgaagaagta cactttctgg gaagttaact tgaaagaaaa gttctctgct 1380
gatttggatc aattcccatt gggtagaaag tttttgttgc aagctggatt gaaggctaaa 1440
ccaaagttca ctttgggaaa gagaaaggct actccaacta cttcttctac ttctactact 1500
gctaagagaa agaagagaaa attgttcgtg ttcttggtcc tgttgccatt ggtttcttcc 1560
cagtgtgtta acctgaccac tagaactcaa ttgcctccag cctacaccaa ttccttcacc 1620
agaggtgttt actacccaga caaggtgttc agatcttccg tcttgcactc cactcaggac 1680
ttgttcttgc cattcttctc caacgttacc tggttccacg ctattcacgt ttccggaact 1740
aacggtacta agagattcga caacccagtc ctgccattca acgatggtgt ctacttcgct 1800
tctaccgaga agtccaacat catcagaggt tggatcttcg gtactaccct ggactctaag 1860
actcagtcct tgctgatcgt taacaacgcc accaacgttg tcatcaaggt ttgcgagttc 1920
cagttctgca acgacccatt cttgggtgtg tactaccaca agaacaacaa gtcttggatg 1980
gaatccgagt tcagagttta ctcctccgcc aacaactgta ccttcgagta cgtttcccag 2040
ccattcttga tggacttgga gggtaagcag ggtaacttca agaacctgag agagttcgtt 2100
ttcaagaaca tcgacggtta cttcaagatc tactccaagc acaccccaat caacctggtt 2160
agagatttgc cacaaggttt ctccgctttg gagcctttgg ttgacttgcc aatcggtatc 2220

aacatcacca gattccagac cttgttggcc ttgcacagat cctacttgac tccaggtgat 2280
tcttcttccg gttggactgc tggtgctgct gcttactatg ttggttactt gcagccaaga 2340
accttcctgc tgaagtacaa cgagaacgga actatcactg acgctgttga ctgtgctttg 2400
gacccattgt ctgagactaa gtgcaccttg aagtccttca ccgttgagaa gggtatctac 2460
cagacctcca acttcagagt tcagccaact gagtccatcg tcagattccc aaacatcact 2520
aacttgtgcc cattcggtga ggtgttcaac gctactagat tcgcttctgt ttacgcctgg 2580
aacagaaaga gaatctccaa ctgcgttgct gactactccg tcttgtacaa ctctgcttca 2640
ttctccacct tcaagtgcta cggtgtttcc ccaactaagt tgaacgacct gtgtttcact 2700
aacgtctacg ccgactcctt cgttattaga ggtgacgagg ttagacagat cgctccaggt 2760
caaactggta agatcgctga ctacaactac aagctgccag acgacttcac cggttgtgtt 2820
attgcttgga actccaacaa cctggactcc aaggttggtg gtaactacaa ttacctgtac 2880
cgtctgttca gaaagtccaa cttgaagcca ttcgagagag acatctccac cgagatctac 2940
caagctggtt ctactccatg taacggtgtc gagggtttca actgctactt cccattgcaa 3000
tcctacggtt tccaacctac caacggtgtt ggataccagc catacagagt tgtcgttttg 3060
tccttcgagt tgttgcacgc tccagctact gtttgtggtc caaagaagtc caccaacttg 3120
gtcaagaaca aatgcgtcaa ctttaacttc aacggcctga ccggtactgg tgttttgact 3180
gaatccaaca agaagttcct gcctttccag cagttcggta gagacattgc tgacactact 3240
gacgccgtta gagatccaca gactttggag atcttggaca tcaccccatg ttccttcggt 3300
ggtgtttccg ttattacccc tggaactaac acctccaatc aggtcgctgt cttgtaccag 3360
gacgttaact gtactgaggt tccagttgct atccacgctg accaattgac tccaacttgg 3420
agagtctact ccaccggttc caacgttttc caaactagag ccggttgttt gatcggtgct 3480
gaacacgtca acaactccta cgagtgtgac attccaattg gtgctggtat ctgtgcctcc 3540
taccaaactc aaactaactc cccaagaagg gctagatccg ttgcttccca atccattatc 3600
gcttacacca tgtctttggg tgccgagaac tctgttgcct actctaacaa ctctatcgct 3660
atccctacca acttcaccat ctccgttacc actgagatct tgccagtctc catgaccaag 3720
acttccgttg actgtaccat gtacatctgt ggtgactcca ctgagtgttc caacttgttg 3780
ctgcaatacg gttccttctg cacccagttg aacagagctt tgactggtat tgctgtcgag 3840
caagacaaga acactcaaga ggttttcgcc caggtgaagc agatctacaa gactccacct 3900
attaaggact tcggtggctt caacttctcc cagattttgc cagatccatc taagccctcc 3960
aagagatcct tcattgagga cctgctgttc aacaaggtta ctttggctga cgccggtttc 4020
atcaagcagt acggtgattg cttgggtgac attgcagcta gagacttgat ctgtgcccag 4080
aagttcaacg gtttgaccgt tttgccacct ttgttgaccg acgagatgat cgctcagtac 4140
acttctgctt tgttggccgg tactatcact tctggttgga catttggagc tggtgccgca 4200
ttgcaaattc cattcgctat gcaaatggcc tacagattca acggtatcgg tgttacccag 4260
aacgtcctgt acgagaacca gaagcttatc gccaaccagt tcaactccgc tatcggtaag 4320
attcaggact ccttgtcctc tactgcttct gccttgggaa agttgcagga tgttgttaac 4380
cagaatgccc aggctttgaa caccctggtt aagcaactgt cctctaactt cggtgctatc 4440
tcctccgttt tgaacgacat cttgtcccgt ttggacaagg ttgaggctga ggttcagatc 4500
gacagattga tcactggtag attgcagtcc ctgcagactt acgttactca gcagttgatt 4560

agagctgccg agattagagc ctctgctaac ttggctgcta ctaagatgtc cgagtgtgtt 4620
ttgggtcagt ccaagagagt tgacttctgc ggtaagggtt accacctgat gtctttccca 4680
caatctgctc cacacggtgt cgttttcttg cacgttactt acgttccagc tcaagagaag 4740
aacttcacta ctgctccagc catttgtcac gatggtaagg ctcactttcc tcgtgagggt 4800
gttttcgttt ccaacggtac tcactggttc gtcacccaga gaaactttta cgagccacag 4860
atcatcacca ccgacaacac tttcgtttct ggtaactgtg acgtcgtcat cggtatcgtg 4920
aacaacactg tctacgatcc attgcagcca gaattggact ccttcaaaga ggaactggac 4980
aagtacttta agaaccacac ttccccagac gttgacctgg gtgatatttc cggtattaac 5040
gcctccgttg tcaacatcca aaaagagatc gaccgtttga acgaggtcgc caagaacttg 5100
aacgagtcct tgattgactt gcaagagctg ggcaagtacg agcagtacat taagtggcca 5160
tggtacattt ggctgggttt cattgctggt ttgatcgcca tcgttatggt caccatcatg 5220
ttgtgctgta tgacctcctg ttgctcctgt ttgaagggtt gttgttcctg cggttcctgt 5280
tgtaagttcg acgaagatga ctccgagcca gtcttgaagg gtgttaagtt gcactacact 5340
taagcggccg c 5351

<210> 7
<211> 3836
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> SARS-CoV-2刺突蛋白核酸序列-针对在人类中的表达优化
<400> 7

gctagcgaca tgttcgtgtt tctggtgctg ctgcctctgg tgtccagcca gtgtgtgaac 60
ctgaccacca gaacacagct gcctccagcc tacaccaata gcttcaccag gggcgtgtac 120
taccccgaca aggtgttcag atctagcgtg ctgcacagca cccaggacct gtttctgccc 180
ttcttcagca acgtgacctg gttccacgcc atccacgtgt ccggcaccaa tggcaccaag 240
agattcgaca accccgtgct gcccttcaac gatggggtgt actttgccag caccgagaag 300
tccaacatca tcagaggctg gatcttcggc accacactgg acagcaagac ccagagcctg 360
ctgatcgtga acaacgccac caacgtggtc atcaaagtgt gcgagttcca gttctgcaac 420
gacccattcc tgggagtcta ctaccacaag aacaacaaga gctggatgga aagcgagttc 480
cgggtgtaca gcagcgccaa caactgcacc ttcgagtacg tgtcccagcc tttcctgatg 540
gacctggaag gcaagcaggg caacttcaag aacctgcgcg agttcgtgtt caagaacatc 600
gacggctact tcaagatcta cagcaagcac acccctatca acctcgtgcg ggatctgcct 660
cagggctttt ctgctctgga acctctggtg gacctgccta tcggcatcaa catcacccgg 720
tttcagaccc tgctggccct gcacagatct tacctgacac ctggcgatag cagctctgga 780
tggacagctg gcgccgctgc ctattatgtg ggctacctgc agcctcggac cttcctgctg 840
aagtacaacg agaacggcac catcaccgac gccgtggatt gtgctctgga tcccctgagc 900
gagacaaagt gcaccctgaa gtccttcacc gtggaaaagg gcatctacca gaccagcaac 960
ttcagagtgc agcccaccga gagcatcgtg cggttcccca atatcaccaa tctgtgcccc 1020
ttcggcgagg tgttcaatgc cacaagattt gccagcgtgt acgcctggaa ccggaagaga 1080

atcagcaact gcgtggccga ctacagcgtg ctgtacaata gcgccagctt cagcaccttc 1140
aagtgctacg gcgtgtcccc taccaagctg aacgacctgt gcttcaccaa tgtgtacgcc 1200
gacagcttcg tgatcagagg cgacgaagtt cggcagatcg ctcctggaca gacaggcaag 1260
atcgccgatt acaactacaa gctgcccgac gacttcaccg gctgcgtgat cgcctggaat 1320
agcaacaacc tggactccaa agtcggcggc aactacaact acctgtaccg gctgttccgg 1380
aagtccaatc tgaagccctt cgagcgggac atctccaccg aaatctatca ggccggcagc 1440
acccccttgta acggcgtgga aggcttcaac tgctacttcc cactgcagtc ctacggcttt 1500
cagcctacca atggcgtggg ctatcagccc tatagagtgg tggtgctgag cttcgaactg 1560
ctgcatgccc ctgctaccgt gtgcggccct aagaagtcta ccaacctggt caagaacaaa 1620
tgcgtgaact tcaacttcaa cggcctgacc ggcacaggcg tgctgacaga gagcaacaag 1680
aagttcctgc ctttccagca gtttggccgg gatatcgccg ataccacaga cgccgttaga 1740
gatccccaga cactggaaat cctggacatc accccatgca gctttggcgg agtgtctgtg 1800
atcacccctg gcaccaatac cagcaatcag gtggccgtgc tgtatcagga cgtgaactgt 1860
acagaggtgc ccgtggccat tcacgccgat caactgacac ccacttggag agtgtactcc 1920
accggctcca acgtgttcca gactagagcc ggatgtctga tcggagccga gcacgtgaac 1980
aatagctacg agtgcgacat ccccatcggc gctggcatct gtgccagcta ccagacacag 2040
acaaatagcc ccagacgggc cagaagcgtg gcctctcaga gcatcattgc ctacacaatg 2100
agcctgggcg ccgagaattc tgtggcctac agcaacaact ctatcgctat ccccaccaac 2160
ttcaccatca gcgtgaccac cgagatcctg cctgtgtcca tgaccaagac cagcgtggac 2220
tgcaccatgt acatctgcgg cgattccacc gagtgcagca acctgctgct gcagtacggc 2280
agcttctgca cccagctgaa tagagccctg acagggatcg ccgtggaaca ggacaagaac 2340
acccaagagg tgttcgccca agtgaagcag atctacaaga cccctcctat caaggacttc 2400
ggcggcttca atttcagcca gattctgccc gatcctagca agcccagcaa gcggagcttt 2460
atcgaggacc tgctgttcaa caaagtgaca ctggccgacg ccggcttcat caagcagtat 2520
ggcgattgcc tgggcgacat tgccgccaga gatctgattt gcgcccagaa gtttaacgga 2580
ctgacagtgc tgcctcctct gctgaccgat gagatgatcg cccagtacac atctgctctg 2640
ctggccggca caatcaccag cggatggaca tttggagctg gcgcagccct gcagatcccc 2700
tttgctatgc agatggccta ccggttcaac ggcatcggag tgacccagaa tgtgctgtac 2760
gagaaccaga agctgatcgc caaccagttc aacagcgcca tcggcaagat ccaggatagc 2820
ctgtctagca cagccagcgc tctgggcaaa ctgcaggacg tggtcaatca gaacgctcag 2880
gccctgaaca ccctcgtgaa gcagctgagc agcaatttcg gcgccatcag ctccgtgctg 2940
aacgatatcc tgagccggct ggataaggtg gaagccgagg tgcagatcga cagactgatc 3000
acaggcagac tgcagagcct ccagacatac gtgacccagc agctgatcag agccgccgag 3060
attagagcct ctgccaatct ggccgccacc aagatgtctg agtgtgtgct gggccagagc 3120
aagagagtgg atttctgcgg caagggctac cacctgatga gctttccaca gtctgctcct 3180
cacggcgtgg tgtttctgca cgtgacctat gtgcccgctc aagagaagaa cttcacaaca 3240
gcccctgcca tctgccacga cggaaaggcc cattttccta gagaaggcgt gttcgtgtcc 3300
aacggcaccc attggttcgt gacacagcgg aacttctacg agccccagat catcaccacc 3360
gacaacacct tcgtgtctgg caactgtgac gtcgtgatcg gcattgtgaa caacaccgtg 3420

tacgaccctc tgcagcccga gctggacagc ttcaaagagg aactggacaa gtactttaag 3480
aaccacacaa gccccgacgt ggacctgggc gatattagcg gcatcaatgc ctccgtggtc 3540
aacatccaga aagagatcga ccggctgaac gaggtggcca agaatctgaa cgagagcctg 3600
atcgacctgc aagaactggg gaagtacgag cagtacatca agtggccctg gtacatctgg 3660
ctgggcttta tcgccggact gattgccatc gtgatggtca caatcatgct gtgctgcatg 3720
accagctgct gtagctgcct gaagggctgt tgcagctgtg gcagctgctg caagttcgac 3780
gaggatgata gcgagcctgt gctgaagggc gtgaaactgc actacaccgc ggccgc 3836

〈210〉 8

〈211〉 4232

〈212〉 DNA

〈213〉 人工序列(Artificial Sequence)

〈220〉

〈223〉 编码融合蛋白HBSAg/SARS-CoV-2刺突蛋白的核酸序列-针对在人类中的表达优化

〈400〉 8

gctagcgaca tgaactttct gggcggtacg acagtatgcc ttggacaaaa ttcacaatct 60
ccgacgtcta atcactcccc tacaagttgt ccaccgactt gccccggcta taggtggatg 120
tgtctcagac gattcataat ctttctcttc attcttcttc tgtgcctgat attcttgctg 180
gtccttctgg attaccaggg aatgcttccc gtgtgtcctc tgattcctgg ttcatccact 240
acatctacgg gtccctgtag aacatgcacc acacctgcac agggcacctc catgtatccg 300
tcatgctgct gcacgaaacc atcagatggt aactgcacgt gcataccgat cccctcatca 360
tgggcgtttg ggaaatttct gtgggagtgg gcctcagccc ggttttcctt cgtgtttctg 420
gtgctgctgc ctctggtgtc cagccagtgt gtgaacctga ccaccagaac acagctgcct 480
ccagcctaca ccaatagctt caccaggggc gtgtactacc ccgacaaggt gttcagatct 540
agcgtgctgc acagcaccca ggacctgttt ctgcccttct tcagcaacgt gacctggttc 600
cacgccatcc acgtgtccgg caccaatggc accaagagat tcgacaaccc cgtgctgccc 660
ttcaacgatg gggtgtactt tgccagcacc gagaagtcca acatcatcag aggctggatc 720
ttcggcacca cactggacag caagacccag agcctgctga tcgtgaacaa cgccaccaac 780
gtggtcatca aagtgtgcga gttccagttc tgcaacgacc cattcctggg agtctactac 840
cacaagaaca acaagagctg gatggaaagc gagttccggg tgtacagcag cgccaacaac 900
tgcaccttcg agtacgtgtc ccagcctttc ctgatggacc tggaaggcaa gcagggcaac 960
ttcaagaacc tgcgcgagtt cgtgttcaag aacatcgacg gctacttcaa gatctacagc 1020
aagcacaccc ctatcaacct cgtgcgggat ctgcctcagg gcttttctgc tctggaacct 1080
ctggtggacc tgcctatcgg catcaacatc acccggtttc agaccctgct ggccctgcac 1140
agatcttacc tgacacctgg cgatagcagc tctggatgga cagctggcgc cgctgcctat 1200
tatgtgggct acctgcagcc tcggaccttc ctgctgaagt acaacgagaa cggcaccatc 1260
accgacgccg tggattgtgc tctggatccc ctgagcgaga caaagtgcac cctgaagtcc 1320
ttcaccgtgg aaaagggcat ctaccagacc agcaacttca gagtgcagcc caccgagagc 1380
atcgtgcggt tccccaatat caccaatctg tgcccttcg gcgaggtgtt caatgccaca 1440

agatttgcca gcgtgtacgc ctggaaccgg aagagaatca gcaactgcgt ggccgactac 1500
agcgtgctgt acaatagcgc cagcttcagc accttcaagt gctacggcgt gtcccctacc 1560
aagctgaacg acctgtgctt caccaatgtg tacgccgaca gcttcgtgat cagaggcgac 1620
gaagttcggc agatcgctcc tggacagaca ggcaagatcg ccgattacaa ctacaagctg 1680
cccgacgact tcaccggctg cgtgatcgcc tggaatagca acaacctgga ctccaaagtc 1740
ggcggcaact acaactacct gtaccggctg ttccggaagt ccaatctgaa gcccttcgag 1800
cgggacatct ccaccgaaat ctatcaggcc ggcagcaccc cttgtaacgg cgtggaaggc 1860
ttcaactgct acttcccact gcagtcctac ggctttcagc ctaccaatgg cgtgggctat 1920
cagccctata gagtggtggt gctgagcttc gaactgctgc atgcccctgc taccgtgtgc 1980
ggccctaaga agtctaccaa cctggtcaag aacaaatgcg tgaacttcaa cttcaacggc 2040
ctgaccggca caggcgtgct gacagagagc aacaagaagt tcctgccttt ccagcagttt 2100
ggccgggata tcgccgatac cacagacgcc gttagagatc cccagacact ggaaatcctg 2160
gacatcaccc catgcagctt tggcggagtg tctgtgatca cccctggcac caataccagc 2220
aatcaggtgg ccgtgctgta tcaggacgtg aactgtacag aggtgcccgt ggccattcac 2280
gccgatcaac tgacacccac ttggagagtg tactccaccg gctccaacgt gttccagact 2340
agagccggat gtctgatcgg agccgagcac gtgaacaata gctacgagtg cgacatcccc 2400
atcggcgctg gcatctgtgc cagctaccag acacagacaa atagccccag acgggccaga 2460
agcgtggcct ctcagagcat cattgcctac acaatgagcc tgggcgccga gaattctgtg 2520
gcctacagca acaactctat cgctatcccc accaacttca ccatcagcgt gaccaccgag 2580
atcctgcctg tgtccatgac caagaccagc gtggactgca ccatgtacat ctgcggcgat 2640
tccaccgagt gcagcaacct gctgctgcag tacggcagct tctgcaccca gctgaataga 2700
gccctgacag ggatcgccgt ggaacaggac aagaacaccc aagaggtgtt cgcccaagtg 2760
aagcagatct acaagacccc tcctatcaag gacttcggcg gcttcaattt cagccagatt 2820
ctgcccgatc ctagcaagcc cagcaagcgg agctttatcg aggacctgct gttcaacaaa 2880
gtgacactgg ccgacgccgg cttcatcaag cagtatggcg attgcctggg cgacattgcc 2940
gccagagatc tgatttgcgc ccagaagttt aacggactga cagtgctgcc tcctctgctg 3000
accgatgaga tgatcgccca gtacacatct gctctgctgg ccggcacaat caccagcgga 3060
tggacatttg gagctggcgc agccctgcag atccccttg ctatgcagat ggcctaccgg 3120
ttcaacggca tcggagtgac ccagaatgtg ctgtacgaga accagaagct gatcgccaac 3180
cagttcaaca gcgccatcgg caagatccag gatagcctgt ctagcacagc cagcgctctg 3240
ggcaaactgc aggacgtggt caatcagaac gctcaggccc tgaacaccct cgtgaagcag 3300
ctgagcagca atttcggcgc catcagctcc gtgctgaacg atatcctgag ccggctggat 3360
aaggtggaag ccgaggtgca gatcgacaga ctgatcacag gcagactgca gagcctccag 3420
acatacgtga cccagcagct gatcagagcc gccgagatta gagcctctgc caatctggcc 3480
gccaccaaga tgtctgagtg tgtgctgggc cagagcaaga gagtggattt ctgcggcaag 3540
ggctaccacc tgatgagctt tccacagtct gctcctcacg gcgtggtgtt tctgcacgtg 3600
acctatgtgc ccgctcaaga gaagaacttc acaacagccc ctgccatctg ccacgacgga 3660
aaggcccatt ttcctagaga aggcgtgttc gtgtccaacg gcacccattg gttcgtgaca 3720
cagcggaact tctacgagcc ccagatcatc accaccgaca acaccttcgt gtctggcaac 3780

tgtgacgtcg tgatcggcat tgtgaacaac accgtgtacg accctctgca gcccgagctg 3840
gacagcttca aagaggaact ggacaagtac tttaagaacc acacaagccc cgacgtggac 3900
ctgggcgata ttagcggcat caatgcctcc gtggtcaaca tccagaaaga gatcgaccgg 3960
ctgaacgagg tggccaagaa tctgaacgag agcctgatcg acctgcaaga actggggaag 4020
tacgagcagt acatcaagtg gccctggtac atctggctgg gctttatcgc cggactgatt 4080
gccatcgtga tggtcacaat catgctgtgc tgcatgacca gctgctgtag ctgcctgaag 4140
ggctgttgca gctgtggcag ctgctgcaag ttcgacgagg atgatagcga gcctgtgctg 4200
aagggcgtga aactgcacta caccgcggcc gc 4232

<210> 9
<211> 1512
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> 融合蛋白HEV-SARS-CoV-2刺突蛋白
<400> 9

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

Tyr Asn Tyr Asn Thr Thr Ala Ser Asp Gln Leu Leu Val Glu Asn Ala
195 200 205
Ala Gly His Arg Val Ala Ile Ser Thr Tyr Thr Thr Ser Leu Gly Ala
210 215 220
Gly Pro Val Ser Ile Ser Ala Val Ala Val Leu Ala Pro His Ser Ala
225 230 235 240
Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser Ser Gln Cys Val Asn
245 250 255
Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe Thr
260 265 270
Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg Ser Ser Val Leu His
275 280 285
Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser Asn Val Thr Trp Phe
290 295 300
His Ala Ile His Val Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp Asn
305 310 315 320
Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys
325 330 335
Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys
340 345 350
Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile Lys
355 360 365
Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr Tyr
370 375 380
His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr Ser
385 390 395 400
Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu Met
405 410 415
Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe Val
420 425 430
Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr Pro
435 440 445
Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu Pro
450 455 460
Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr Leu
465 470 475 480
Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser Gly
485 490 495
Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro Arg

500 505 510

Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala Val
515 520 525

Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys Ser
530 535 540

Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln
545 550 555 560

Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro
565 570 575

Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp
580 585 590

Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr
595 600 605

Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr
610 615 620

Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val
625 630 635 640

Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys
645 650 655

Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys Val
660 665 670

Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr
675 680 685

Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu
690 695 700

Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn
705 710 715 720

Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe
725 730 735

Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val Leu
740 745 750

Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys Lys
755 760 765

Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn Gly
770 775 780

Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu Pro
785 790 795 800

Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val Arg
805 810 815

Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly
820 825 830
Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val Ala
835 840 845
Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile His
850 855 860
Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly Ser Asn
865 870 875 880
Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val Asn
885 890 895
Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala Ser
900 905 910
Tyr Gln Thr Gln Thr Asn Ser Pro Arg Arg Ala Arg Ser Val Ala Ser
915 920 925
Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Ala Glu Asn Ser Val
930 935 940
Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr Asn Phe Thr Ile Ser
945 950 955 960
Val Thr Thr Glu Ile Leu Pro Val Ser Met Thr Lys Thr Ser Val Asp
965 970 975
Cys Thr Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu
980 985 990
Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly
995 1000 1005
Ile Ala Val Glu Gln Asp Lys Asn Thr Gln Glu Val Phe Ala Gln
1010 1015 1020
Val Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly
1025 1030 1035
Phe Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys
1040 1045 1050
Arg Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala
1055 1060 1065
Asp Ala Gly Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile
1070 1075 1080
Ala Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr
1085 1090 1095
Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr
1100 1105 1110
Ser Ala Leu Leu Ala Gly Thr Ile Thr Ser Gly Trp Thr Phe Gly

1115 1120 1125

Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala Tyr

1130 1135 1140

Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn

1145 1150 1155

Gln Lys Leu Ile Ala Asn Gln Phe Asn Ser Ala Ile Gly Lys Ile

1160 1165 1170

Gln Asp Ser Leu Ser Ser Thr Ala Ser Ala Leu Gly Lys Leu Gln

1175 1180 1185

Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys

1190 1195 1200

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

1205 1210 1215

Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp

1220 1225 1230

Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr

1235 1240 1245

Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu

1250 1255 1260

Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg

1265 1270 1275

Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln

1280 1285 1290

Ser Ala Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro

1295 1300 1305

Ala Gln Glu Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His Asp

1310 1315 1320

Gly Lys Ala His Phe Pro Arg Glu Gly Val Phe Val Ser Asn Gly

1325 1330 1335

Thr His Trp Phe Val Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile

1340 1345 1350

Ile Thr Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val

1355 1360 1365

Ile Gly Ile Val Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu

1370 1375 1380

Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His

1385 1390 1395

Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala

1400 1405 1410

Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu Val
1415 1420 1425
Ala Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly
1430 1435 1440
Lys Tyr Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly
1445 1450 1455
Phe Ile Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile Met Leu
1460 1465 1470
Cys Cys Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Cys Cys Ser
1475 1480 1485
Cys Gly Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro Val
1490 1495 1500
Leu Lys Gly Val Lys Leu His Tyr Thr
1505 1510

<210> 10
<211> 1779
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> 融合蛋白HPV18L1/SARS-CoV-2刺突蛋白
<400> 10

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

Asp Asn Val Ser Val Asp Tyr Lys Gln Thr Gln Leu Cys Ile Leu Gly
145 150 155 160
Cys Ala Pro Ala Ile Gly Glu His Trp Ala Lys Gly Thr Ala Cys Lys
165 170 175
Ser Arg Pro Leu Ser Gln Gly Asp Cys Pro Pro Leu Glu Leu Lys Asn
180 185 190
Thr Val Leu Glu Asp Gly Asp Met Val Asp Thr Gly Tyr Gly Ala Met
195 200 205
Asp Phe Ser Thr Leu Gln Asp Thr Lys Cys Glu Val Pro Leu Asp Ile
210 215 220
Cys Gln Ser Ile Cys Lys Tyr Pro Asp Tyr Leu Gln Met Ser Ala Asp
225 230 235 240
Pro Tyr Gly Asp Ser Met Phe Phe Cys Leu Arg Arg Glu Gln Leu Phe
245 250 255
Ala Arg His Phe Trp Asn Arg Ala Gly Thr Met Gly Asp Thr Val Pro
260 265 270
Gln Ser Leu Tyr Ile Lys Gly Thr Gly Met Arg Ala Ser Pro Gly Ser
275 280 285
Cys Val Tyr Ser Pro Ser Pro Ser Gly Ser Ile Val Thr Ser Asp Ser
290 295 300
Gln Leu Phe Asn Lys Pro Tyr Trp Leu His Lys Ala Gln Gly His Asn
305 310 315 320
Asn Gly Val Cys Trp His Asn Gln Leu Phe Val Thr Val Val Asp Thr
325 330 335
Thr Arg Ser Thr Asn Leu Thr Ile Cys Ala Ser Thr Gln Ser Pro Val
340 345 350
Pro Gly Gln Tyr Asp Ala Thr Lys Phe Lys Gln Tyr Ser Arg His Val
355 360 365
Glu Glu Tyr Asp Leu Gln Phe Ile Phe Gln Leu Cys Thr Ile Thr Leu
370 375 380
Thr Ala Asp Val Met Ser Tyr Ile His Ser Met Asn Ser Ser Ile Leu
385 390 395 400
Glu Asp Trp Asn Phe Gly Val Pro Pro Pro Pro Thr Thr Ser Leu Val
405 410 415
Asp Thr Tyr Arg Phe Val Gln Ser Val Ala Ile Thr Cys Gln Lys Asp
420 425 430
Ala Ala Pro Ala Glu Asn Lys Asp Pro Tyr Asp Lys Leu Lys Phe Trp
435 440 445
Asn Val Asp Leu Lys Glu Lys Phe Ser Leu Asp Leu Asp Gln Tyr Pro

450 455 460
Leu Gly Arg Lys Phe Leu Val Gln Ala Gly Leu Arg Arg Lys Pro Thr
465 470 475 480
Ile Gly Pro Arg Lys Arg Ser Ala Pro Ser Ala Thr Thr Ser Ser Lys
485 490 495
Pro Ala Lys Arg Val Arg Val Arg Ala Arg Lys Phe Val Phe Leu Val
500 505 510
Leu Leu Pro Leu Val Ser Ser Gln Cys Val Asn Leu Thr Thr Arg Thr
515 520 525
Gln Leu Pro Pro Ala Tyr Thr Asn Ser Phe Thr Arg Gly Val Tyr Tyr
530 535 540
Pro Asp Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu
545 550 555 560
Phe Leu Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val
565 570 575
Ser Gly Thr Asn Gly Thr Lys Arg Phe Asp Asn Pro Val Leu Pro Phe
580 585 590
Asn Asp Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg
595 600 605
Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu
610 615 620
Ile Val Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln
625 630 635 640
Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr Tyr His Lys Asn Asn Lys
645 650 655
Ser Trp Met Glu Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys
660 665 670
Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu Met Asp Leu Glu Gly Lys
675 680 685
Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp
690 695 700
Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg
705 710 715 720
Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu Pro Leu Val Asp Leu Pro
725 730 735
Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr Leu Leu Ala Leu His Arg
740 745 750
Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser Gly Trp Thr Ala Gly Ala
755 760 765

Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro Arg Thr Phe Leu Leu Lys
770 775 780
Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala Val Asp Cys Ala Leu Asp
785 790 795 800
Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys Ser Phe Thr Val Glu Lys
805 810 815
Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val Gln Pro Thr Glu Ser Ile
820 825 830
Val Arg Phe Pro Asn Ile Thr Asn Leu Cys Pro Phe Gly Glu Val Phe
835 840 845
Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala Trp Asn Arg Lys Arg Ile
850 855 860
Ser Asn Cys Val Ala Asp Tyr Ser Val Leu Tyr Asn Ser Ala Ser Phe
865 870 875 880
Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro Thr Lys Leu Asn Asp Leu
885 890 895
Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe Val Ile Arg Gly Asp Glu
900 905 910
Val Arg Gln Ile Ala Pro Gly Gln Thr Gly Lys Ile Ala Asp Tyr Asn
915 920 925
Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys Val Ile Ala Trp Asn Ser
930 935 940
Asn Asn Leu Asp Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg
945 950 955 960
Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr
965 970 975
Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe
980 985 990
Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly
995 1000 1005
Val Gly Tyr Gln Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu
1010 1015 1020
Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys Lys Ser Thr Asn
1025 1030 1035
Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn Gly Leu Thr
1040 1045 1050
Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu Pro Phe
1055 1060 1065
Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val Arg

1070 1075 1080
Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe
1085 1090 1095
Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln
1100 1105 1110
Val Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val
1115 1120 1125
Ala Ile His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser
1130 1135 1140
Thr Gly Ser Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly
1145 1150 1155
Ala Glu His Val Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly
1160 1165 1170
Ala Gly Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn Ser Pro Arg
1175 1180 1185
Arg Ala Arg Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr Thr Met
1190 1195 1200
Ser Leu Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser Ile
1205 1210 1215
Ala Ile Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu
1220 1225 1230
Pro Val Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile
1235 1240 1245
Cys Gly Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln Tyr Gly
1250 1255 1260
Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val
1265 1270 1275
Glu Gln Asp Lys Asn Thr Gln Glu Val Phe Ala Gln Val Lys Gln
1280 1285 1290
Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly Phe Asn Phe
1295 1300 1305
Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe
1310 1315 1320
Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly
1325 1330 1335
Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala Arg
1340 1345 1350
Asp Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro
1355 1360 1365

Pro Leu Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser Ala Leu
1370 1375 1380
Leu Ala Gly Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala Gly Ala
1385 1390 1395
Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn
1400 1405 1410
Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn Gln Lys Leu
1415 1420 1425
Ile Ala Asn Gln Phe Asn Ser Ala Ile Gly Lys Ile Gln Asp Ser
1430 1435 1440
Leu Ser Ser Thr Ala Ser Ala Leu Gly Lys Leu Gln Asp Val Val
1445 1450 1455
Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu Ser
1460 1465 1470
Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser
1475 1480 1485
Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp Arg Leu Ile
1490 1495 1500
Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu
1505 1510 1515
Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr
1520 1525 1530
Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe
1535 1540 1545
Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala Pro
1550 1555 1560
His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu
1565 1570 1575
Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala
1580 1585 1590
His Phe Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp
1595 1600 1605
Phe Val Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr
1610 1615 1620
Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile
1625 1630 1635
Val Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser
1640 1645 1650
Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro

1655 1660 1665
Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val
1670 1675 1680
Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys Asn
1685 1690 1695
Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu
1700 1705 1710
Gln Tyr Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala
1715 1720 1725
Gly Leu Ile Ala Ile Val Met Val Thr Ile Met Leu Cys Cys Met
1730 1735 1740
Thr Ser Cys Cys Ser Cys Leu Lys Gly Cys Cys Ser Cys Gly Ser
1745 1750 1755
Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro Val Leu Lys Gly
1760 1765 1770
Val Lys Leu His Tyr Thr
1775

<210> 11
<211> 1777
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> 融合蛋白HPV16L1/SARS-CoV-2刺突蛋白
<400> 11

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

115 120 125

Thr Glu Asn Ala Ser Ala Tyr Ala Ala Asn Ala Gly Val Asp Asn Arg
130 135 140

Glu Cys Ile Ser Met Asp Tyr Lys Gln Thr Gln Leu Cys Leu Ile Gly
145 150 155 160

Cys Lys Pro Pro Ile Gly Glu His Trp Gly Lys Gly Ser Pro Cys Thr
165 170 175

Asn Val Ala Val Asn Pro Gly Asp Cys Pro Pro Leu Glu Leu Ile Asn
180 185 190

Thr Val Ile Gln Asp Gly Asp Met Val Asp Thr Gly Phe Gly Ala Met
195 200 205

Asp Phe Thr Thr Leu Gln Ala Asn Lys Ser Glu Val Pro Leu Asp Ile
210 215 220

Cys Thr Ser Ile Cys Lys Tyr Pro Asp Tyr Ile Lys Met Val Ser Glu
225 230 235 240

Pro Tyr Gly Asp Ser Leu Phe Phe Tyr Leu Arg Arg Glu Gln Met Phe
245 250 255

Val Arg His Leu Phe Asn Arg Ala Gly Ala Val Gly Glu Asn Val Pro
260 265 270

Asp Asp Leu Tyr Ile Lys Gly Ser Gly Ser Thr Ala Asn Leu Ala Ser
275 280 285

Ser Asn Tyr Phe Pro Thr Pro Ser Gly Ser Met Val Thr Ser Asp Ala
290 295 300

Gln Ile Phe Asn Lys Pro Tyr Trp Leu Gln Arg Ala Gln Gly His Asn
305 310 315 320

Asn Gly Ile Cys Trp Gly Asn Gln Leu Phe Val Thr Val Val Asp Thr
325 330 335

Thr Arg Ser Thr Asn Met Ser Leu Cys Ala Ala Ile Ser Thr Ser Glu
340 345 350

Thr Thr Tyr Lys Asn Thr Asn Phe Lys Glu Tyr Leu Arg His Gly Glu
355 360 365

Glu Tyr Asp Leu Gln Phe Ile Phe Gln Leu Cys Lys Ile Thr Leu Thr
370 375 380

Ala Asp Val Met Thr Tyr Ile His Ser Met Asn Ser Thr Ile Leu Glu
385 390 395 400

Asp Trp Asn Phe Gly Leu Gln Pro Pro Pro Gly Gly Thr Leu Glu Asp
405 410 415

Thr Tyr Arg Phe Val Thr Ser Gln Ala Ile Ala Cys Gln Lys His Thr
420 425 430

Pro Pro Ala Pro Lys Glu Asp Pro Leu Lys Lys Tyr Thr Phe Trp Glu
435 440 445
Val Asn Leu Lys Glu Lys Phe Ser Ala Asp Leu Asp Gln Phe Pro Leu
450 455 460
Gly Arg Lys Phe Leu Leu Gln Ala Gly Leu Lys Ala Lys Pro Lys Phe
465 470 475 480
Thr Leu Gly Lys Arg Lys Ala Thr Pro Thr Thr Ser Ser Thr Ser Thr
485 490 495
Thr Ala Lys Arg Lys Lys Arg Lys Leu Phe Val Phe Leu Val Leu Leu
500 505 510
Pro Leu Val Ser Ser Gln Cys Val Asn Leu Thr Thr Arg Thr Gln Leu
515 520 525
Pro Pro Ala Tyr Thr Asn Ser Phe Thr Arg Gly Val Tyr Tyr Pro Asp
530 535 540
Lys Val Phe Arg Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe Leu
545 550 555 560
Pro Phe Phe Ser Asn Val Thr Trp Phe His Ala Ile His Val Ser Gly
565 570 575
Thr Asn Gly Thr Lys Arg Phe Asp Asn Pro Val Leu Pro Phe Asn Asp
580 585 590
Gly Val Tyr Phe Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly Trp
595 600 605
Ile Phe Gly Thr Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile Val
610 615 620
Asn Asn Ala Thr Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe Cys
625 630 635 640
Asn Asp Pro Phe Leu Gly Val Tyr Tyr His Lys Asn Asn Lys Ser Trp
645 650 655
Met Glu Ser Glu Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr Phe
660 665 670
Glu Tyr Val Ser Gln Pro Phe Leu Met Asp Leu Glu Gly Lys Gln Gly
675 680 685
Asn Phe Lys Asn Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly Tyr
690 695 700
Phe Lys Ile Tyr Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp Leu
705 710 715 720
Pro Gln Gly Phe Ser Ala Leu Glu Pro Leu Val Asp Leu Pro Ile Gly
725 730 735
Ile Asn Ile Thr Arg Phe Gln Thr Leu Leu Ala Leu His Arg Ser Tyr

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

Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu Pro Phe Gln Gln
1055 1060 1065
Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val Arg Asp Pro
1070 1075 1080
Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe Gly Gly
1085 1090 1095
Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val Ala
1100 1105 1110
Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile
1115 1120 1125
His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly
1130 1135 1140
Ser Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu
1145 1150 1155
His Val Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly
1160 1165 1170
Ile Cys Ala Ser Tyr Gln Thr Gln Thr Asn Ser Pro Arg Arg Ala
1175 1180 1185
Arg Ser Val Ala Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu
1190 1195 1200
Gly Ala Glu Asn Ser Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile
1205 1210 1215
Pro Thr Asn Phe Thr Ile Ser Val Thr Thr Glu Ile Leu Pro Val
1220 1225 1230
Ser Met Thr Lys Thr Ser Val Asp Cys Thr Met Tyr Ile Cys Gly
1235 1240 1245
Asp Ser Thr Glu Cys Ser Asn Leu Leu Leu Gln Tyr Gly Ser Phe
1250 1255 1260
Cys Thr Gln Leu Asn Arg Ala Leu Thr Gly Ile Ala Val Glu Gln
1265 1270 1275
Asp Lys Asn Thr Gln Glu Val Phe Ala Gln Val Lys Gln Ile Tyr
1280 1285 1290
Lys Thr Pro Pro Ile Lys Asp Phe Gly Gly Phe Asn Phe Ser Gln
1295 1300 1305
Ile Leu Pro Asp Pro Ser Lys Pro Ser Lys Arg Ser Phe Ile Glu
1310 1315 1320
Asp Leu Leu Phe Asn Lys Val Thr Leu Ala Asp Ala Gly Phe Ile
1325 1330 1335
Lys Gln Tyr Gly Asp Cys Leu Gly Asp Ile Ala Ala Arg Asp Leu

1340 1345 1350

Ile Cys Ala Gln Lys Phe Asn Gly Leu Thr Val Leu Pro Pro Leu

1355 1360 1365

Leu Thr Asp Glu Met Ile Ala Gln Tyr Thr Ser Ala Leu Leu Ala

1370 1375 1380

Gly Thr Ile Thr Ser Gly Trp Thr Phe Gly Ala Gly Ala Ala Leu

1385 1390 1395

Gln Ile Pro Phe Ala Met Gln Met Ala Tyr Arg Phe Asn Gly Ile

1400 1405 1410

Gly Val Thr Gln Asn Val Leu Tyr Glu Asn Gln Lys Leu Ile Ala

1415 1420 1425

Asn Gln Phe Asn Ser Ala Ile Gly Lys Ile Gln Asp Ser Leu Ser

1430 1435 1440

Ser Thr Ala Ser Ala Leu Gly Lys Leu Gln Asp Val Val Asn Gln

1445 1450 1455

Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu Ser Ser Asn

1460 1465 1470

Phe Gly Ala Ile Ser Ser Val Leu Asn Asp Ile Leu Ser Arg Leu

1475 1480 1485

Asp Lys Val Glu Ala Glu Val Gln Ile Asp Arg Leu Ile Thr Gly

1490 1495 1500

Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln Leu Ile Arg

1505 1510 1515

Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala Thr Lys Met

1520 1525 1530

Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp Phe Cys Gly

1535 1540 1545

Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala Pro His Gly

1550 1555 1560

Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln Glu Lys Asn

1565 1570 1575

Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys Ala His Phe

1580 1585 1590

Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His Trp Phe Val

1595 1600 1605

Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr Thr Asp Asn

1610 1615 1620

Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly Ile Val Asn

1625 1630 1635

Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp Ser Phe Lys
1640 1645 1650
Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser Pro Asp Val
1655 1660 1665
Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val Val Asn Ile
1670 1675 1680
Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys Asn Leu Asn
1685 1690 1695
Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr Glu Gln Tyr
1700 1705 1710
Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile Ala Gly Leu
1715 1720 1725
Ile Ala Ile Val Met Val Thr Ile Met Leu Cys Cys Met Thr Ser
1730 1735 1740
Cys Cys Ser Cys Leu Lys Gly Cys Cys Ser Cys Gly Ser Cys Cys
1745 1750 1755
Lys Phe Asp Glu Asp Asp Ser Glu Pro Val Leu Lys Gly Val Lys
1760 1765 1770
Leu His Tyr Thr
1775

<210> 12
<211> 1407
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> 融合蛋白HBSAg/SARS-CoV-2刺突蛋白
<400> 12

Met Asn Phe Leu Gly Gly Thr Thr Val Cys Leu Gly Gln Asn Ser Gln
1 5 10 15
Ser Pro Thr Ser Asn His Ser Pro Thr Ser Cys Pro Pro Thr Cys Pro
20 25 30
Gly Tyr Arg Trp Met Cys Leu Arg Arg Phe Ile Ile Phe Leu Phe Ile
35 40 45
Leu Leu Leu Cys Leu Ile Phe Leu Leu Val Leu Leu Asp Tyr Gln Gly
50 55 60
Met Leu Pro Val Cys Pro Leu Ile Pro Gly Ser Ser Thr Thr Ser Thr
65 70 75 80
Gly Pro Cys Arg Thr Cys Thr Thr Pro Ala Gln Gly Thr Ser Met Tyr
85 90 95

Pro Ser Cys Cys Cys Thr Lys Pro Ser Asp Gly Asn Cys Thr Cys Ile
100 105 110
Pro Ile Pro Ser Ser Trp Ala Phe Gly Lys Phe Leu Trp Glu Trp Ala
115 120 125
Ser Ala Arg Phe Ser Phe Val Phe Leu Val Leu Leu Pro Leu Val Ser
130 135 140
Ser Gln Cys Val Asn Leu Thr Thr Arg Thr Gln Leu Pro Pro Ala Tyr
145 150 155 160
Thr Asn Ser Phe Thr Arg Gly Val Tyr Tyr Pro Asp Lys Val Phe Arg
165 170 175
Ser Ser Val Leu His Ser Thr Gln Asp Leu Phe Leu Pro Phe Phe Ser
180 185 190
Asn Val Thr Trp Phe His Ala Ile His Val Ser Gly Thr Asn Gly Thr
195 200 205
Lys Arg Phe Asp Asn Pro Val Leu Pro Phe Asn Asp Gly Val Tyr Phe
210 215 220
Ala Ser Thr Glu Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr
225 230 235 240
Thr Leu Asp Ser Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr
245 250 255
Asn Val Val Ile Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe
260 265 270
Leu Gly Val Tyr Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu
275 280 285
Phe Arg Val Tyr Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser
290 295 300
Gln Pro Phe Leu Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn
305 310 315 320
Leu Arg Glu Phe Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr
325 330 335
Ser Lys His Thr Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe
340 345 350
Ser Ala Leu Glu Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr
355 360 365
Arg Phe Gln Thr Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly
370 375 380
Asp Ser Ser Ser Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly
385 390 395 400
Tyr Leu Gln Pro Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr

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

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

1025 1030 1035

Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu Asn Gln Lys

1040 1045 1050

Leu Ile Ala Asn Gln Phe Asn Ser Ala Ile Gly Lys Ile Gln Asp

1055 1060 1065

Ser Leu Ser Ser Thr Ala Ser Ala Leu Gly Lys Leu Gln Asp Val

1070 1075 1080

Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val Lys Gln Leu

1085 1090 1095

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

1100 1105 1110

Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile Asp Arg Leu

1115 1120 1125

Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val Thr Gln Gln

1130 1135 1140

Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn Leu Ala Ala

1145 1150 1155

Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys Arg Val Asp

1160 1165 1170

Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro Gln Ser Ala

1175 1180 1185

Pro His Gly Val Val Phe Leu His Val Thr Tyr Val Pro Ala Gln

1190 1195 1200

Glu Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His Asp Gly Lys

1205 1210 1215

Ala His Phe Pro Arg Glu Gly Val Phe Val Ser Asn Gly Thr His

1220 1225 1230

Trp Phe Val Thr Gln Arg Asn Phe Tyr Glu Pro Gln Ile Ile Thr

1235 1240 1245

Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val Val Ile Gly

1250 1255 1260

Ile Val Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro Glu Leu Asp

1265 1270 1275

Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn His Thr Ser

1280 1285 1290

Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn Ala Ser Val

1295 1300 1305

Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu Val Ala Lys

1310 1315 1320

Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu Leu Gly Lys Tyr
1325 1330 1335
Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Ile Trp Leu Gly Phe Ile
1340 1345 1350
Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile Met Leu Cys Cys
1355 1360 1365
Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Cys Cys Ser Cys Gly
1370 1375 1380
Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro Val Leu Lys
1385 1390 1395
Gly Val Lys Leu His Tyr Thr Ala Ala
1400 1405

<210> 13
<211> 665
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> RBD SARS-CoV-2刺突蛋白核酸
<400> 13
gctagcgacg ccaccatgag agtccaacca acagaatcta ttgttagatt tcctaatatt 60
acaaacttgt gccctttttgg tgaagttttt aacgccacca gatttgcatc tgtttatgct 120
tggaacagga agagaatcag caactgtgtt gctgattatt ctgtcctata taattccgca 180
tcattttcca cttttaagtg ttatggagtg tctcctacta aattaaatga tctctgcttt 240
actaatgtct atgcagattc atttgtaatt agaggtgatg aagtcagaca aatcgctcca 300
gggcaaactg gaaagattgc tgattataat tataaattac cagatgattt tacaggctgc 360
gttatagctt ggaattctaa caatcttgat tctaaggttg gtggtaatta taattacctg 420
tatagattgt ttaggaagtc taatctcaaa ccttttgaga gagatatttc aactgaaatc 480
tatcaggccg gtagcacacc ttgtaatggt gttgaaggtt ttaattgtta ctttccttta 540
caatcatatg gtttccaacc cactaatggt gttggttacc aaccatacag agtagtagta 600
ctttcttttg aacttctaca tgcaccagca actgtttgtg gacctaaaaa gtgataagcg 660
gccgc 665
<210> 14
<211> 665
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> RBD SARS-CoV-2刺突蛋白核酸序列-人密码子优化的
<400> 14
gctagcgacg ccaccatgag agtgcagcct acagagtcta tcgtgcggtt ccccaacatc 60

accaatctgt gccctttcgg cgaggtgttc aacgccacaa gatttgccag cgtgtacgcc 120
tggaaccgga agagaatcag caactgcgtg gccgactaca gcgtgctgta caatagcgcc 180
agcttcagca ccttcaagtg ctacggcgtg tcccctacca agctgaacga cctgtgcttc 240
accaatgtgt acgccgacag cttcgtgatc agaggcgacg aagttcggca gatcgctcct 300
ggacagacag gcaagatcgc cgattacaac tacaagctgc ccgacgactt caccggctgc 360
gtgatcgcct ggaatagcaa caacctggac agcaaagtcg gcggcaacta caactacctg 420
taccggctgt tccggaagtc caacctgaag cctttcgagc gggacatcag caccgagatc 480
tatcaggccg gcagcacccc ttgtaatggc gtggaaggct tcaactgcta cttcccactg 540
cagtcctacg gcttccagcc tacaaacggc gtgggctacc agccttatag agtggtggtg 600
ctgagcttcg aactgctgca tgcccctgct acagtgtgcg gccccaagaa gtgataagcg 660
gccgc 665

<210> 15
<211> 212
<212> PRT
<213> SARS-CoV-2
<400> 15

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

180 185 190

Val Val Val Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys

195 200 205

Gly Pro Lys Lys

210

<210> 16

<211> 7

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> EAAAK接头共有序列

<400> 16

Ala Glu Ala Ala Ala Lys Ala

1 5

<210> 17

<211> 45

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> (EAAAK)3接头

<400> 17

gaagccgccg ctaaagaggc cgctgccaaa gaagctgctg ctaag 45

<210> 18

<211> 15

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> (EAAAK)3接头

<400> 18

Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys Glu Ala Ala Ala Lys

1 5 10 15

<210> 19

<211> 5

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> GS接头共有序列

<400> 19

Gly Gly Gly Gly Ser

1 5

<210> 20

<211> 5

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> GS5接头

<400> 20

Gly Gly Gly Gly Ser

1 5

<210> 21

<211> 10

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> GS10接头

<400> 21

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10

<210> 22

<211> 45

<212> DNA

<213> 人工序列(Artificial Sequence)

<220>

<223> GS15接头

<400> 22

ggtggtggtg gtagcggtgg tggcggttca ggtggcggtg gttca 45

<210> 23

<211> 15

<212> PRT

<213> 人工序列(Artificial Sequence)

<220>

<223> GS15接头

<400> 23

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10 15

<210> 24

<211> 20

<212> PRT

<213> 人工序列(Artificial Sequence)
<220>
<223> GS20接头
<400> 24
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1 5 10 15
Gly Gly Gly Ser
20
<210> 25
<211> 25
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> GS25接头
<400> 25
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1 5 10 15
Gly Gly Gly Ser Gly Gly Gly Gly Ser
20 25
<210> 26
<211> 1427
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> HBSAg-(EAAAK)3-RBD
<400> 26
gctagcgacg ccaccatgat tgcactgacc ctgtttaatc tggcagatac cctgttaggt 60
ggtctgccga ccgaactgat tagcagtgcc ggtggtcagc tgttttatag ccgtccggtt 120
gttagcgcaa atggtgaacc gaccgttaaa ctgtatacca gcgttgaaaa tgcacagcag 180
gataaaggta ttgcaattcc gcatgatatt gatctgggtg aaagccgtgt tgtgattcag 240
gattatgata atcagcatga acaggatcgt ccgacaccga gtccggcacc gagccgtccg 300
tttagcgttc tgcgtgcaaa tgatgttctg tggctgagcc tgaccgcagc agaatatgat 360
cagagcacct atggtagcag caccggtccg gtttatgtta gcgatagcgt taccctggtt 420
aatgttgcaa ccggtgcaca ggcagttgca cgtagcctgg attggaccaa agtgaccctg 480
gatggtcgtc cgctgagcac cattcagcag tatagcaaaa ccttttttgt tctgccgctg 540
cgtggtaaac tgagcttttg ggaagcaggc accaccaaag caggttatcc gtataactat 600
aataccaccg caagcgatca gctgctggtt gaaaacgcag caggtcatcg tgttgcaatt 660
agcacctata ccaccagttt aggtgcaggt ccggttagca ttagcgcagt tgcagttctg 720
gcaccgcatt cagccgaagc agccgctaaa gaagcagccg ctaaagaagc agccgctaaa 780

agagtccaac caacagaatc tattgttaga tttcctaata ttacaaactt gtgccctttt 840
ggtgaagttt ttaacgccac cagatttgca tctgtttatg cttggaacag gaagagaatc 900
agcaactgtg ttgctgatta ttctgtccta tataattccg catcattttc acttttaag 960
tgttatggag tgtctcctac taaattaaat gatctctgct ttactaatgt ctatgcagat 1020
tcatttgtaa ttagaggtga tgaagtcaga caaatcgctc cagggcaaac tggaaagatt 1080
gctgattata attataaatt accagatgat tttacaggct gcgttatagc ttggaattct 1140
aacaatcttg attctaaggt tggtggtaat tataattacc tgtatagatt gtttaggaag 1200
tctaatctca aacctttga gagagatatt tcaactgaaa tctatcaggc cggtagcaca 1260
ccttgtaatg gtgttgaagg ttttaattgt tactttcctt tacaatcata tggtttccaa 1320
cccactaatg gtgttggtta ccaaccatac agagtagtag tactttcttt tgaacttcta 1380
catgcaccag caactgtttg tggacctaaa aagtgataag cggccgc 1427

<210> 27
<211> 1106
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 人密码子优化的HBSAg-(EAAAK)3-RBD核酸序列
<400> 27

gctagcgacg ccaccatgaa ttttctcggc ggcacaacag tgtgcctggg ccagaatagc 60
cagtctccta ccagcaatca cagccccacc agctgtcctc caacctgtcc tggctacaga 120
tggatgtgcc tgcggcggtt catcatcttt ctgttcatcc tgctgctgtg cctgatcttc 180
ctgctggtgc tgctggatta ccagggaatg ctgcctgtgt gtcctctgat ccctggcagc 240
agcacaacaa gcacaggccc ttgcagaacc tgcacaacac cagctcaggg caccagcatg 300
taccctagct gctgttgtac caagcctagc gacggcaact gcacatgcat ccccattcct 360
agcagctggg ccttcggcaa gtttctgtgg gaatgggcca gcgccagatt ttccgaagcc 420
gccgctaaag aggccgctgc caaagaagct gctgctaaga gagtgcagcc caccgagtct 480
atcgtgcggt tccccaacat caccaatctg tgcccctttcg gcgaggtgtt caacgccaca 540
agatttgcca gcgtgtacgc ctggaaccgg aagagaatca gcaactgcgt ggccgactac 600
agcgtgctgt acaatagcgc cagcttcagc accttcaagt gctacggcgt gtcccctacc 660
aagctgaacg acctgtgctt caccaatgtg tacgccgaca gcttcgtgat cagaggcgac 720
gaagttcggc agatcgctcc tggacagaca ggcaagatcg ccgattacaa ctacaagctg 780
cccgacgact tcaccggctg cgtgatcgcc tggaatagca acaacctgga cagcaaagtc 840
ggcggcaact acaactacct gtaccggctg ttccggaagt ccaacctgaa gcctttcgag 900
cgggacatca gcaccgaaat ctaccaggcc ggcagcaccc cttgtaatgg cgtggaaggc 960
ttcaactgct acttcccact gcagtcctac ggcttccagc ctacaaacgg cgtgggctac 1020
cagccttata gagtggtggt gctgagcttc gaactgctgc atgcccctgc tacagtgtgc 1080
ggccccaaga agtgataagc ggccgc 1106

<210> 28
<211> 359

<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> HBSAg-(EAAAK)3-RBD
<400> 28

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

Ser Lys Val Gly Gly Asn Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys
275 280 285
Ser Asn Leu Lys Pro Phe Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln
290 295 300
Ala Gly Ser Thr Pro Cys Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe
305 310 315 320
Pro Leu Gln Ser Tyr Gly Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln
325 330 335
Pro Tyr Arg Val Val Val Leu Ser Phe Glu Leu Leu His Ala Pro Ala
340 345 350
Thr Val Cys Gly Pro Lys Lys
355

<210> 29
<211> 1418
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> HEV-GS15-RBD
<400> 29

gagctcatga ttgcactgac cctgtttaat ctggcagata ccctgctggg tggtctgccg 60
accgaactga ttagcagtgc cggtggtcag ctgttttata gccgtccggt tgttagcgca 120
aatggtgaac cgaccgttaa actgtatacc agcgttgaaa atgcacagca ggataaaggt 180
attgcaattc cgcatgatat tgatctgggt gaaagccgtg ttgtgattca ggattatgat 240
aatcagcatg aacaggatcg tccgaccccg agtccggcac cgagccgtcc gtttagcgtt 300
ctgcgtgcaa atgatgttct gtggctgagc ctgaccgcag cagaatatga tcagagcacc 360
tatggtagca gcaccggtcc ggtttatgtt agcgatagcg ttaccctggt taatgttgca 420
accggtgcac aggcagttgc acgtagcctg gattggacca aagtgaccct ggatggtcgt 480
ccgctgagca ccattcagca gtatagcaaa acctttttg ttctgccgct gcgtggtaaa 540
ctgagctttt gggaagcagg caccaccaaa gcaggttatc cgtataacta taataccacc 600
gcaagcgatc agctgctggt tgaaaacgca gcaggtcatc gtgttgcaat tagcacctat 660
accaccagtc tgggtgcagg tccggttagc attagcgcag ttgcagttct ggcaccgcat 720
agcgcaggtg gaggaggttc tggaggcggt ggaagtggtg gcggaggtag cagagtccaa 780
ccaacagaat ctattgttag atttcctaat attacaaact tgtgcccttt tggtgaagtt 840
tttaacgcca ccagatttgc atctgtttat gcttggaaca ggaagagaat cagcaactgt 900
gttgctgatt attctgtcct atataattcc gcatcatttt ccacttttaa gtgttatgga 960
gtgtctccta ctaaattaaa tgatctctgc tttactaatg tctatgcaga ttcatttgta 1020
attagaggtg atgaagtcag acaaatcgct ccagggcaaa ctggaaagat tgctgattat 1080
aattataaat taccagatga ttttacaggc tgcgttatag cttggaattc taacaatctt 1140
gattctaagg ttggtggtaa ttataattac ctgtatagat tgtttaggaa gtctaatctc 1200

aaacctttg agagagatat ttcaactgaa atctatcagg ccggtagcac accttgtaat 1260
ggtgttgaag gttttaattg ttactttcct ttacaatcat atggtttcca acccactaat 1320
ggtgttggtt accaaccata cagagtagta gtactttctt ttgaacttct acatgcacca 1380
gcaactgttt gtggacctaa aaagtgataa gcggccgc 1418

<210> 30
<211> 1418
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 针对大肠杆菌(E.coli)表达优化的HEV-GS15-RBD核酸序列
<400> 30

gagctcatga ttgcactgac cctgtttaat ctggcagata ccctgttagg tggtctgccg 60
accgaactga ttagcagtgc cggtggtcag ctgttttata gccgtccggt tgttagcgca 120
aatggtgaac cgaccgttaa actgtatacc agcgttgaaa atgcacagca ggataaaggt 180
attgcaattc cgcatgatat tgatctgggt gaaagccgtg ttgtgattca ggattatgat 240
aatcagcatg aacaggatcg tccgacaccg agtccggcac cgagccgtcc gtttagcgtt 300
ctgcgtgcaa atgatgttct gtggctgagc ctgaccgcag cagaatatga tcagagcacc 360
tatggtagca gcaccggtcc ggtttatgtt agcgatagcg ttaccctggt taatgttgca 420
accggtgcac aggcagttgc acgtagcctg gattggacca aagtgaccct ggatggtcgt 480
ccgctgagca ccattcagca gtatagcaaa acctttttg ttctgccgct gcgtggtaaa 540
ctgagctttt gggaagcagg caccaccaaa gcaggttatc cgtataacta taataccacc 600
gcaagcgatc agctgctggt tgaaaacgca gcaggtcatc gtgttgcaat tagcacctat 660
accaccagtt taggtgcagg tccggttagc attagcgcag ttgcagttct ggcaccgcat 720
tcagccggtg gtggtggtag cggtggtggc ggttcaggtg gcggtggttc acgtgttcag 780
ccgacagaaa gcattgttcg ttttccgaat atcaccaatc tgtgtccgtt tggcgaagtt 840
tttaatgcaa cccgttttgc aagcgtttat gcctggaatc gtaaacgtat tagcaattgc 900
gttgccgatt atagcgtgct gtataatagc gcaagcttta gcacctttaa atgctatggt 960
gttagcccga ccaaactgaa tgatctgtgt tttaccaatg tgtatgccga tagctttgtg 1020
attcgtggtg atgaagttcg tcagattgca ccgggtcaga ccggtaaaat tgcagattat 1080
aactacaaac tgccggatga ttttacgggt tgtgttattg catggaatag caataacctg 1140
gatagcaaag ttggtggcaa ctataactat ctgtatcgcc tgtttcgtaa gagcaatctg 1200
aaaccgtttg aacgtgatat tagcaccgaa atttatcagg caggtagcac cccgtgcaat 1260
ggtgttgaag gttttaattg ttattttccg ctgcagagct atggttttca gcctaccaat 1320
ggtgtgggtt atcagccgta tcgtgttgtt gttctgtcat ttgaactgct gcatgcaccg 1380
gcaaccgttt gtggtccgaa aaaatgataa gcggccgc 1418

<210> 31
<211> 466
<212> PRT
<213> 人工序列(Artificial Sequence)

<220>

<223> HEV-GS15-RBD

<400> 31

Met Ile Ala Leu Thr Leu Phe Asn Leu Ala Asp Thr Leu Leu Gly Gly
1 5 10 15
Leu Pro Thr Glu Leu Ile Ser Ser Ala Gly Gly Gln Leu Phe Tyr Ser
20 25 30
Arg Pro Val Val Ser Ala Asn Gly Glu Pro Thr Val Lys Leu Tyr Thr
35 40 45
Ser Val Glu Asn Ala Gln Gln Asp Lys Gly Ile Ala Ile Pro His Asp
50 55 60
Ile Asp Leu Gly Glu Ser Arg Val Val Ile Gln Asp Tyr Asp Asn Gln
65 70 75 80
His Glu Gln Asp Arg Pro Thr Pro Ser Pro Ala Pro Ser Arg Pro Phe
85 90 95
Ser Val Leu Arg Ala Asn Asp Val Leu Trp Leu Ser Leu Thr Ala Ala
100 105 110
Glu Tyr Asp Gln Ser Thr Tyr Gly Ser Ser Thr Gly Pro Val Tyr Val
115 120 125
Ser Asp Ser Val Thr Leu Val Asn Val Ala Thr Gly Ala Gln Ala Val
130 135 140
Ala Arg Ser Leu Asp Trp Thr Lys Val Thr Leu Asp Gly Arg Pro Leu
145 150 155 160
Ser Thr Ile Gln Gln Tyr Ser Lys Thr Phe Phe Val Leu Pro Leu Arg
165 170 175
Gly Lys Leu Ser Phe Trp Glu Ala Gly Thr Thr Lys Ala Gly Tyr Pro
180 185 190
Tyr Asn Tyr Asn Thr Thr Ala Ser Asp Gln Leu Leu Val Glu Asn Ala
195 200 205
Ala Gly His Arg Val Ala Ile Ser Thr Tyr Thr Thr Ser Leu Gly Ala
210 215 220
Gly Pro Val Ser Ile Ser Ala Val Ala Val Leu Ala Pro His Ser Ala
225 230 235 240
Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Arg
245 250 255
Val Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu
260 265 270
Cys Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr
275 280 285

Ala Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val
290 295 300
Leu Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser
305 310 315 320
Pro Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser
325 330 335
Phe Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr
340 345 350
Gly Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly
355 360 365
Cys Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly
370 375 380
Asn Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro
385 390 395 400
Phe Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro
405 410 415
Cys Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr
420 425 430
Gly Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val
435 440 445
Val Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro
450 455 460
Lys Lys
465

<210> 32
<211> 4566
<212> DNA
<213> 人工序列(Artificial Sequence)
<220>
<223> 人密码子优化的HBSAg-(EAAAK)3-全长SARS-CoV-2刺突蛋白核酸序列
<400> 32

aagcttgccg ccaccatgga gaacatcaca tcaggattcc taggacccct gctcgtgtta 60
caggcggggt ttttcttgtt gacaagaatc ctcacaatac cacagagtct agactcgtgg 120
tggacttctc tcaattttct agggggatca cccgtgtgtc tgggccaaaa ttcgcagtcc 180
ccaacctcca atcactcacc aacctcttgt cctccaattt gtcctggcta tcgctggatg 240
tgtctgcggc gttttatcat attcctcttc atcctgctgc tatgcctcat cttcttgttg 300
gttcttctgg actaccaggg tatgttgccc gtttgtcctc taattccagg atcaacaact 360
accaacacgg gaccatgcaa gacctgcacg actcctgctc aaggaaactc tatgtttccc 420
tcttgttgct gtacaaaacc taccgacgga aactgcactt gtattcccat cccatcatcc 480

tgggctttcg caaaatacct atgggagtgg gcctcagtcc gtttctcctg gctcagttta 540
ctagtgccat ttgttcagtg gttcgtaggg ctttccccca ctgtttggct ttccgctata 600
tggatgatgt ggtattgggg gccaagtctg tacagcatcg tgagtccctt tatacctcta 660
ttaccaattt tcttttgtct ttgggtatac attgaggctg ccgcaaagga agccgcagct 720
aaagaggcag ctgccaagtt cgtgttcctg gttctgctgc ccctggtgtc tagccagtgc 780
gtgaacctga ccaccagaac acagctgcct ccagcctaca ccaacagctt caccagaggc 840
gtgtactacc ccgacaaggt gttccggtcc tccgtgctgc attctaccca ggacctgttc 900
ctgcctttct tctccaacgt gacctggttc cacgccatcc atgtgtctgg caccaacggc 960
accaagagat tcgacaaccc cgtgctgcct ttcaacgacg gggtgtactt tgcctccacc 1020
gagaagtcca acatcatcag aggctggatc ttcggcacaa ccctggacag caagacccag 1080
agcctgctga tcgtgaacaa cgccaccaac gtggtcatca aagtgtgcga gttccagttc 1140
tgcaacgacc ccttcctggg cgtctactac cacaagaaca acaagtcctg gatggaatcc 1200
gagttccggg tgtactcctc cgccaacaac tgcaccttcg aatacgtgtc ccagcctttc 1260
ctgatggacc tggaaggcaa gcagggcaac ttcaagaacc tgcgcgagtt cgtgttcaag 1320
aacatcgacg gctacttcaa gatctactcc aagcacaccc ctatcaacct cgtgcgggat 1380
ctgcctcagg gcttctctgc tctggaaccc ctggtggatc tgcccatcgg catcaacatc 1440
acccggtttc agaccctgct ggccctgcac cggtcttatt tgacccctgg cgactcctct 1500
tctggctgga ctgctggcgc cgctgcttac tatgtgggct acctgcagcc tcggaccttt 1560
ctgctgaagt acaacgagaa tggcaccatc accgacgccg tggactgtgc tctggatcct 1620
ctgtccgaga caaagtgcac cctgaagtcc ttcaccgtgg aaaagggcat ctaccagacc 1680
tccaacttcc gggtgcagcc caccgagtct atcgtgcggt tccctaacat caccaacctg 1740
tgtcctttcg gcgaggtgtt caatgccacc agattcgcct ctgtgtacgc ctggaaccgg 1800
aagcggatct ctaactgcgt ggccgactac agcgtgctgt acaactccgc ctccttcagc 1860
accttcaagt gctacggcgt gtcccctaca aagctgaacg acctgtgctt cacaaacgtg 1920
tacgccgaca gcttcgtgat ccggggagat gaagtgcggc agatcgctcc tggacagacc 1980
ggcaagatcg ccgattacaa ctacaagctg cccgacgact tcaccggctg tgtgatcgct 2040
tggaactcca acaacctgga ctccaaagtc ggcggcaact acaactacct gtaccggctg 2100
ttccggaagt ctaacctgaa gcctttcgag cgggacatca gcaccgagat ctaccaggct 2160
ggcagcaccc cttgtaacgg cgtggaaggc ttcaactgct acttcccact gcagtcctac 2220
ggctttcagc ctaccaatgg cgtgggctat cagccctaca gagtggtggt gctgtccttc 2280
gagctgctgc atgctcctgc taccgtgtgc ggccctaaga aatctaccaa cctggtcaag 2340
aacaaatgcg tgaacttcaa cttcaacggc ctgaccggca ccggcgtgct gacagagtcc 2400
aacaagaagt tcctgccatt ccagcagttc ggccgggata tcgccgatac cacagatgcc 2460
gtcagggacc ctcagacact ggaaatcctg gacatcaccc cttgctcctt cggcggagtg 2520
tctgtgatca ccccaggcac caacacctct aaccaggtgg ccgtgctgta tcaggacgtg 2580
aactgtaccg aggtgcccgt ggctatccat gccgatcagc tgacccctac atggcgcgtg 2640
tactccaccg gctctaacgt gttccagaca agagctggct gtctgatcgg cgctgagcac 2700
gtgaacaatt cctacgagtg cgacatcccc atcggagccg gaatctgcgc ctcttatcag 2760
acccagacca actctcccag acgggccaga tctgtggcca gccagtctat cattgcttac 2820

accatgagcc tgggcgccga gaactctgtg gcctacagca acaactctat cgctatcccc 2880
accaacttca ccatctccgt gaccacagag atcctgccag tgtccatgac caagaccagc 2940
gtggactgca ccatgtacat ctgcggcgac tctaccgagt gctccaacct gctgctccag 3000
tacggctcct tctgcaccca gctgaataga gccctgaccg gaatcgccgt ggaacaggac 3060
aagaacaccc aagaggtgtt cgcccaagtg aagcagatct acaagacccc tcctatcaag 3120
gacttcggcg gcttcaattt ctcccagatt ctgcccgatc ctagcaagcc ctccaagcgg 3180
tctttcatcg aggacctgct gttcaacaaa gtgacactgg ccgacgccgg cttcatcaag 3240
cagtacggcg actgtctggg cgacattgcc gctagggatc tgatctgcgc ccagaagttt 3300
aacggactga cagtgctgcc tcctctgctg accgatgaga tgatcgccca gtacacctcc 3360
gcactgctgg ctggcacaat cacctctgga tggacatttg gcgctggcgc tgctctgcaa 3420
atcccattcg ctatgcaaat ggcctaccgg ttcaacggca tcggcgtgac ccagaatgtg 3480
ctgtacgaga accagaagct gatcgccaac cagttcaaca gcgccatcgg aaagatccag 3540
gacagcctgt ccagcaccgc ttctgccctg ggaaagctgc aggatgtggt caaccagaac 3600
gctcaggccc tgaacaccct cgtgaagcag ctgtctagca acttcggcgc catctcctct 3660
gtgctgaacg atatcctgag ccggctggac aaggtggaag ccgaggtgca gatcgacaga 3720
ctgatcaccg gacggctgca gtccctgcag acctatgtta cccagcagct gatccgggct 3780
gccgagatta gagcctctgc caatctggcc gcaaccaaga tgtctgagtg tgtgctggga 3840
cagtccaaga gagtggactt ctgcggcaag ggctaccacc tgatgagctt ccctcagtct 3900
gctcctcacg gcgtggtgtt tctgcacgtg acctacgtgc ccgctcaaga gaagaacttt 3960
accaccgctc ctgccatctg ccacgacggc aaggctcact ttcctagaga aggcgtgttc 4020
gtgtctaacg gcacccattg gttcgtgaca cagcggaact tctacgagcc ccagatcatc 4080
accaccgaca acaccttcgt gtccggcaac tgcgacgtcg tgatcggaat tgtgaacaat 4140
accgtgtacg accctctgca gcccgagctg gactccttca aagaggaact ggacaagtac 4200
tttaagaacc acacaagccc cgacgtggac ctgggagaca tctctggcat caacgcctcc 4260
gtggtcaaca tccagaaaga gatcgaccgg ctgaacgagg tggccaagaa tctgaacgag 4320
tccctgatcg acctgcaaga actggggaag tacgagcagt acatcaagtg gccctggtac 4380
atctggctgg gctttatcgc tggcctgatc gctatcgtga tggtcacaat catgctgtgc 4440
tgtatgacct cctgttgctc ctgcctgaag ggctgctgct cttgcggctc ttgctgcaag 4500
ttcgacgagg acgactctga gcccgtgctg aaaggcgtga agctgcacta tacctgatga 4560
ctcgag 4566

<210> 33
<211> 1513
<212> PRT
<213> 人工序列(Artificial Sequence)
<220>
<223> HBSAg-(EAAAK)3-全长SARS-CoV-2刺突蛋白
<400> 33

Met Glu Asn Ile Thr Ser Gly Phe Leu Gly Pro Leu Leu Val Leu Gln
1 5 10 15

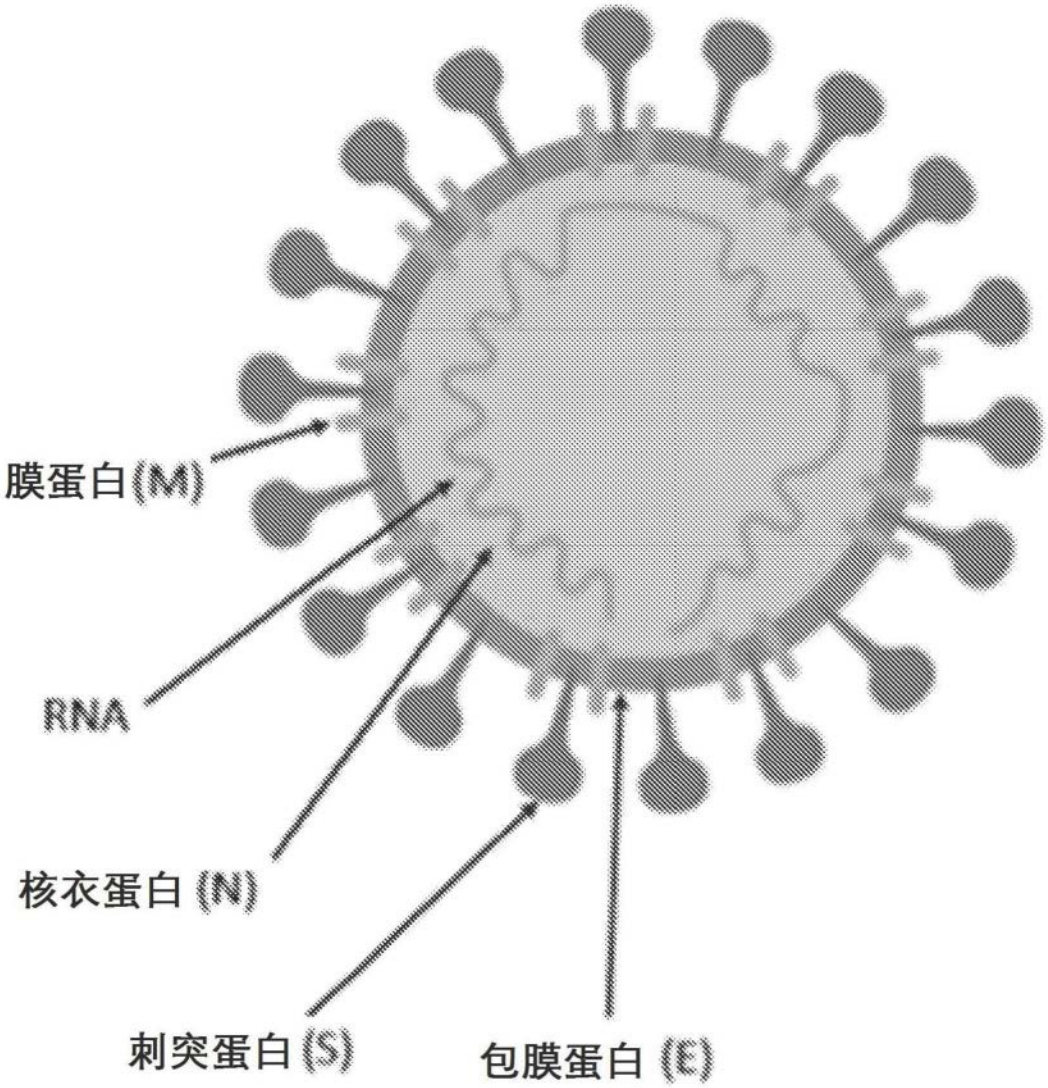
Ala Gly Phe Phe Leu Leu Thr Arg Ile Leu Thr Ile Pro Gln Ser Leu
20 25 30
Asp Ser Trp Trp Thr Ser Leu Asn Phe Leu Gly Gly Ser Pro Val Cys
35 40 45
Leu Gly Gln Asn Ser Gln Ser Pro Thr Ser Asn His Ser Pro Thr Ser
50 55 60
Cys Pro Pro Ile Cys Pro Gly Tyr Arg Trp Met Cys Leu Arg Arg Phe
65 70 75 80
Ile Ile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

325 330 335
Lys Ser Asn Ile Ile Arg Gly Trp Ile Phe Gly Thr Thr Leu Asp Ser
340 345 350
Lys Thr Gln Ser Leu Leu Ile Val Asn Asn Ala Thr Asn Val Val Ile
355 360 365
Lys Val Cys Glu Phe Gln Phe Cys Asn Asp Pro Phe Leu Gly Val Tyr
370 375 380
Tyr His Lys Asn Asn Lys Ser Trp Met Glu Ser Glu Phe Arg Val Tyr
385 390 395 400
Ser Ser Ala Asn Asn Cys Thr Phe Glu Tyr Val Ser Gln Pro Phe Leu
405 410 415
Met Asp Leu Glu Gly Lys Gln Gly Asn Phe Lys Asn Leu Arg Glu Phe
420 425 430
Val Phe Lys Asn Ile Asp Gly Tyr Phe Lys Ile Tyr Ser Lys His Thr
435 440 445
Pro Ile Asn Leu Val Arg Asp Leu Pro Gln Gly Phe Ser Ala Leu Glu
450 455 460
Pro Leu Val Asp Leu Pro Ile Gly Ile Asn Ile Thr Arg Phe Gln Thr
465 470 475 480
Leu Leu Ala Leu His Arg Ser Tyr Leu Thr Pro Gly Asp Ser Ser Ser
485 490 495
Gly Trp Thr Ala Gly Ala Ala Ala Tyr Tyr Val Gly Tyr Leu Gln Pro
500 505 510
Arg Thr Phe Leu Leu Lys Tyr Asn Glu Asn Gly Thr Ile Thr Asp Ala
515 520 525
Val Asp Cys Ala Leu Asp Pro Leu Ser Glu Thr Lys Cys Thr Leu Lys
530 535 540
Ser Phe Thr Val Glu Lys Gly Ile Tyr Gln Thr Ser Asn Phe Arg Val
545 550 555 560
Gln Pro Thr Glu Ser Ile Val Arg Phe Pro Asn Ile Thr Asn Leu Cys
565 570 575
Pro Phe Gly Glu Val Phe Asn Ala Thr Arg Phe Ala Ser Val Tyr Ala
580 585 590
Trp Asn Arg Lys Arg Ile Ser Asn Cys Val Ala Asp Tyr Ser Val Leu
595 600 605
Tyr Asn Ser Ala Ser Phe Ser Thr Phe Lys Cys Tyr Gly Val Ser Pro
610 615 620
Thr Lys Leu Asn Asp Leu Cys Phe Thr Asn Val Tyr Ala Asp Ser Phe
625 630 635 640

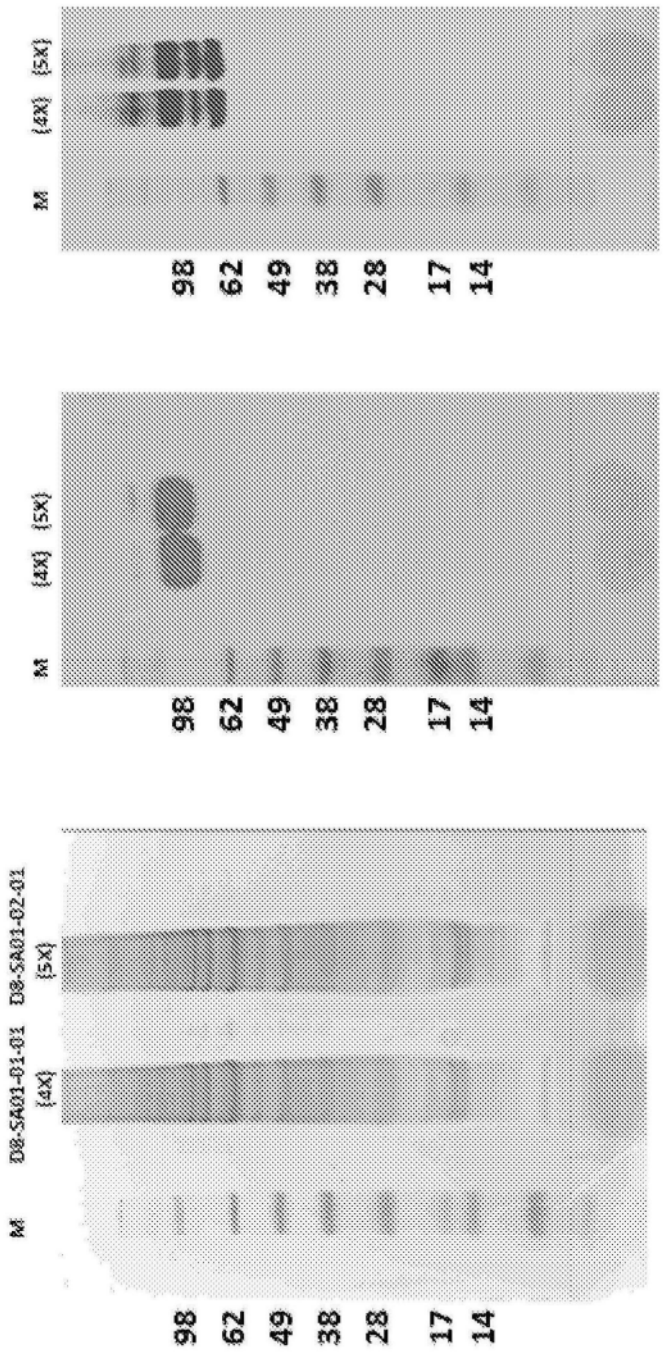
Val Ile Arg Gly Asp Glu Val Arg Gln Ile Ala Pro Gly Gln Thr Gly
645 650 655
Lys Ile Ala Asp Tyr Asn Tyr Lys Leu Pro Asp Asp Phe Thr Gly Cys
660 665 670
Val Ile Ala Trp Asn Ser Asn Asn Leu Asp Ser Lys Val Gly Gly Asn
675 680 685
Tyr Asn Tyr Leu Tyr Arg Leu Phe Arg Lys Ser Asn Leu Lys Pro Phe
690 695 700
Glu Arg Asp Ile Ser Thr Glu Ile Tyr Gln Ala Gly Ser Thr Pro Cys
705 710 715 720
Asn Gly Val Glu Gly Phe Asn Cys Tyr Phe Pro Leu Gln Ser Tyr Gly
725 730 735
Phe Gln Pro Thr Asn Gly Val Gly Tyr Gln Pro Tyr Arg Val Val Val
740 745 750
Leu Ser Phe Glu Leu Leu His Ala Pro Ala Thr Val Cys Gly Pro Lys
755 760 765
Lys Ser Thr Asn Leu Val Lys Asn Lys Cys Val Asn Phe Asn Phe Asn
770 775 780
Gly Leu Thr Gly Thr Gly Val Leu Thr Glu Ser Asn Lys Lys Phe Leu
785 790 795 800
Pro Phe Gln Gln Phe Gly Arg Asp Ile Ala Asp Thr Thr Asp Ala Val
805 810 815
Arg Asp Pro Gln Thr Leu Glu Ile Leu Asp Ile Thr Pro Cys Ser Phe
820 825 830
Gly Gly Val Ser Val Ile Thr Pro Gly Thr Asn Thr Ser Asn Gln Val
835 840 845
Ala Val Leu Tyr Gln Asp Val Asn Cys Thr Glu Val Pro Val Ala Ile
850 855 860
His Ala Asp Gln Leu Thr Pro Thr Trp Arg Val Tyr Ser Thr Gly Ser
865 870 875 880
Asn Val Phe Gln Thr Arg Ala Gly Cys Leu Ile Gly Ala Glu His Val
885 890 895
Asn Asn Ser Tyr Glu Cys Asp Ile Pro Ile Gly Ala Gly Ile Cys Ala
900 905 910
Ser Tyr Gln Thr Gln Thr Asn Ser Pro Arg Arg Ala Arg Ser Val Ala
915 920 925
Ser Gln Ser Ile Ile Ala Tyr Thr Met Ser Leu Gly Ala Glu Asn Ser
930 935 940
Val Ala Tyr Ser Asn Asn Ser Ile Ala Ile Pro Thr Asn Phe Thr Ile

945 950 955 960
Ser Val Thr Thr Glu Ile Leu Pro Val Ser Met Thr Lys Thr Ser Val
965 970 975
Asp Cys Thr Met Tyr Ile Cys Gly Asp Ser Thr Glu Cys Ser Asn Leu
980 985 990
Leu Leu Gln Tyr Gly Ser Phe Cys Thr Gln Leu Asn Arg Ala Leu Thr
995 1000 1005
Gly Ile Ala Val Glu Gln Asp Lys Asn Thr Gln Glu Val Phe Ala
1010 1015 1020
Gln Val Lys Gln Ile Tyr Lys Thr Pro Pro Ile Lys Asp Phe Gly
1025 1030 1035
Gly Phe Asn Phe Ser Gln Ile Leu Pro Asp Pro Ser Lys Pro Ser
1040 1045 1050
Lys Arg Ser Phe Ile Glu Asp Leu Leu Phe Asn Lys Val Thr Leu
1055 1060 1065
Ala Asp Ala Gly Phe Ile Lys Gln Tyr Gly Asp Cys Leu Gly Asp
1070 1075 1080
Ile Ala Ala Arg Asp Leu Ile Cys Ala Gln Lys Phe Asn Gly Leu
1085 1090 1095
Thr Val Leu Pro Pro Leu Leu Thr Asp Glu Met Ile Ala Gln Tyr
1100 1105 1110
Thr Ser Ala Leu Leu Ala Gly Thr Ile Thr Ser Gly Trp Thr Phe
1115 1120 1125
Gly Ala Gly Ala Ala Leu Gln Ile Pro Phe Ala Met Gln Met Ala
1130 1135 1140
Tyr Arg Phe Asn Gly Ile Gly Val Thr Gln Asn Val Leu Tyr Glu
1145 1150 1155
Asn Gln Lys Leu Ile Ala Asn Gln Phe Asn Ser Ala Ile Gly Lys
1160 1165 1170
Ile Gln Asp Ser Leu Ser Ser Thr Ala Ser Ala Leu Gly Lys Leu
1175 1180 1185
Gln Asp Val Val Asn Gln Asn Ala Gln Ala Leu Asn Thr Leu Val
1190 1195 1200
Lys Gln Leu Ser Ser Asn Phe Gly Ala Ile Ser Ser Val Leu Asn
1205 1210 1215
Asp Ile Leu Ser Arg Leu Asp Lys Val Glu Ala Glu Val Gln Ile
1220 1225 1230
Asp Arg Leu Ile Thr Gly Arg Leu Gln Ser Leu Gln Thr Tyr Val
1235 1240 1245

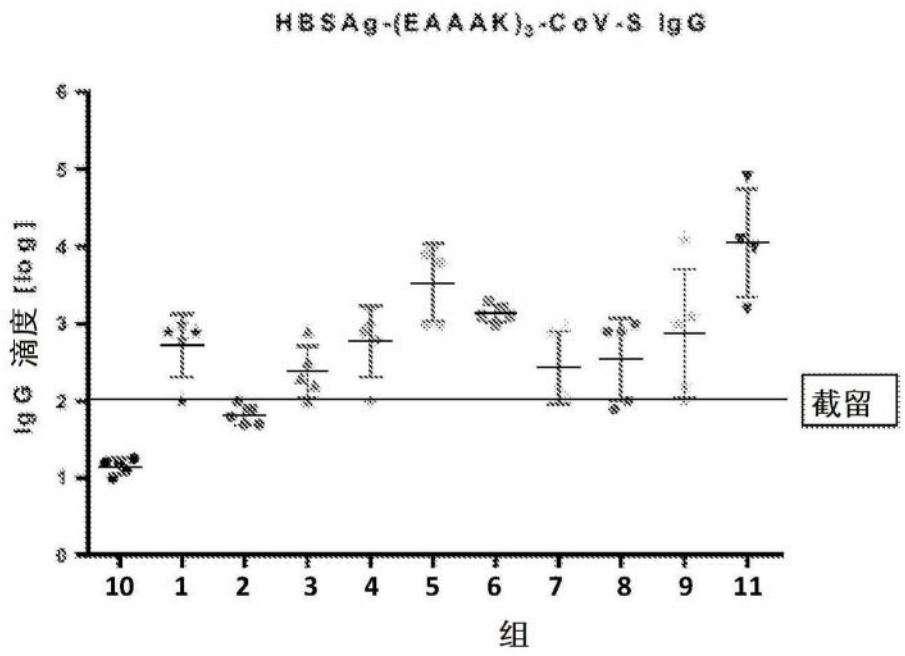
Thr Gln Gln Leu Ile Arg Ala Ala Glu Ile Arg Ala Ser Ala Asn
1250 1255 1260
Leu Ala Ala Thr Lys Met Ser Glu Cys Val Leu Gly Gln Ser Lys
1265 1270 1275
Arg Val Asp Phe Cys Gly Lys Gly Tyr His Leu Met Ser Phe Pro
1280 1285 1290
Gln Ser Ala Pro His Gly Val Val Phe Leu His Val Thr Tyr Val
1295 1300 1305
Pro Ala Gln Glu Lys Asn Phe Thr Thr Ala Pro Ala Ile Cys His
1310 1315 1320
Asp Gly Lys Ala His Phe Pro Arg Glu Gly Val Phe Val Ser Asn
1325 1330 1335
Gly Thr His Trp Phe Val Thr Gln Arg Asn Phe Tyr Glu Pro Gln
1340 1345 1350
Ile Ile Thr Thr Asp Asn Thr Phe Val Ser Gly Asn Cys Asp Val
1355 1360 1365
Val Ile Gly Ile Val Asn Asn Thr Val Tyr Asp Pro Leu Gln Pro
1370 1375 1380
Glu Leu Asp Ser Phe Lys Glu Glu Leu Asp Lys Tyr Phe Lys Asn
1385 1390 1395
His Thr Ser Pro Asp Val Asp Leu Gly Asp Ile Ser Gly Ile Asn
1400 1405 1410
Ala Ser Val Val Asn Ile Gln Lys Glu Ile Asp Arg Leu Asn Glu
1415 1420 1425
Val Ala Lys Asn Leu Asn Glu Ser Leu Ile Asp Leu Gln Glu Leu
1430 1435 1440
Gly Lys Tyr Glu Gln Tyr Ile Lys Trp Pro Trp Tyr Ile Trp Leu
1445 1450 1455
Gly Phe Ile Ala Gly Leu Ile Ala Ile Val Met Val Thr Ile Met
1460 1465 1470
Leu Cys Cys Met Thr Ser Cys Cys Ser Cys Leu Lys Gly Cys Cys
1475 1480 1485
Ser Cys Gly Ser Cys Cys Lys Phe Asp Glu Asp Asp Ser Glu Pro
1490 1495 1500
Val Leu Lys Gly Val Lys Leu His Tyr Thr
1505 1510



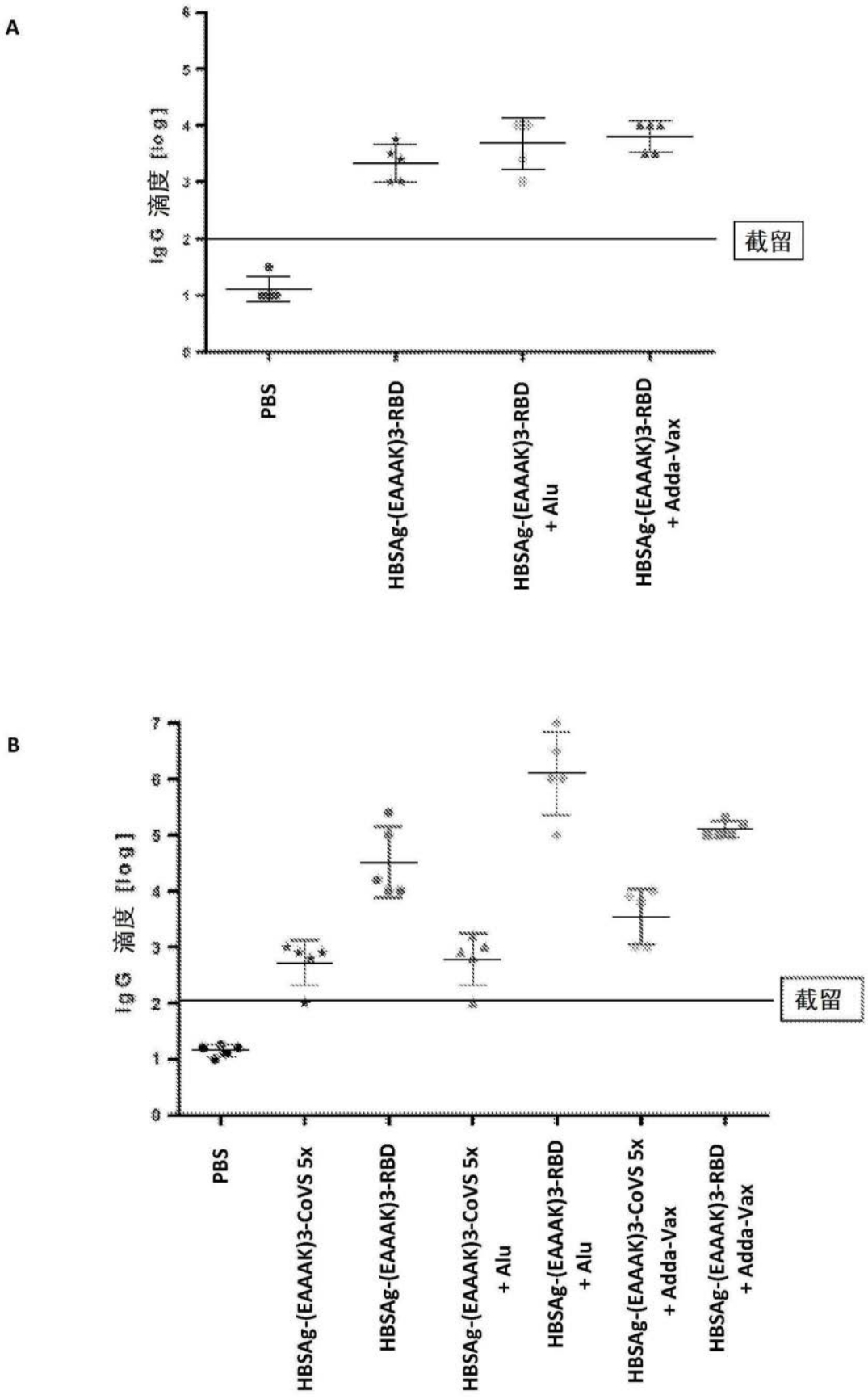
附图1


M
{4X}
{5X}
98
62
49
38
28
17
14
M
{4X}
{5X}
98
62
49
38
28
17
14
M
D8-SA01-01-01
{4X}
D8-SA01-02-01
{5X}
98
62
49
38
28
17
14

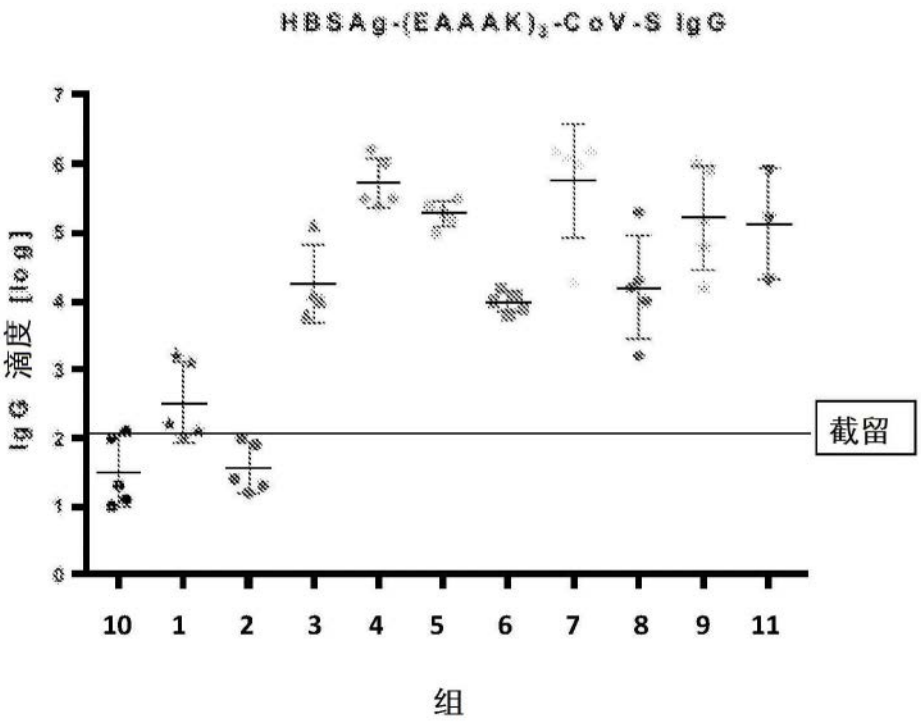
附图2



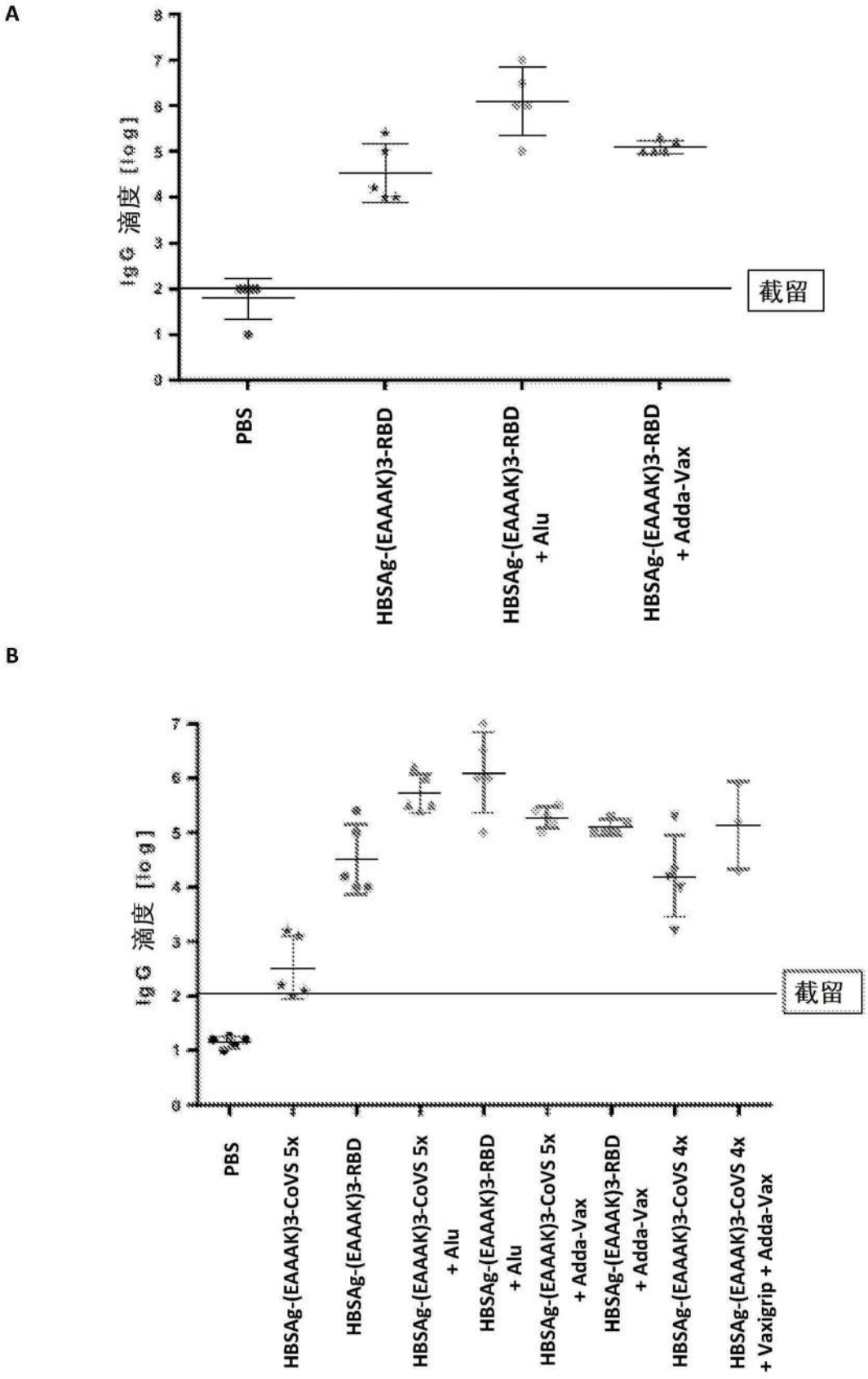
附图3


A
IgG 滴度 [log]
截留
PBS
HBSAg-(EAAAK)3-RBD
HBSAg-(EAAAK)3-RBD + Alu
HBSAg-(EAAAK)3-RBD + Adda-Vax
B
IgG 滴度 [log]
截留
PBS
HBSAg-(EAAAK)3-CoVS 5x
HBSAg-(EAAAK)3-RBD
HBSAg-(EAAAK)3-CoVS 5x + Alu
HBSAg-(EAAAK)3-RBD + Alu
HBSAg-(EAAAK)3-CoVS 5x + Adda-Vax
HBSAg-(EAAAK)3-RBD + Adda-Vax

附图4



附图5


A
IgG 滴度 [log]
截留
PBS
HBSAg-(EAAAK)3-RBD
HBSAg-(EAAAK)3-RBD + Alu
HBSAg-(EAAAK)3-RBD + Adda-Vax
B
IgG 滴度 [log]
截留
PBS
HBSAg-(EAAAK)3-CoVS 5x
HBSAg-(EAAAK)3-RBD
HBSAg-(EAAAK)3-CoVS 5x + Alu
HBSAg-(EAAAK)3-RBD + Alu
HBSAg-(EAAAK)3-CoVS 5x + Adda-Vax
HBSAg-(EAAAK)3-RBD + Adda-Vax
HBSAg-(EAAAK)3-CoVS 4x
HBSAg-(EAAAK)3-CoVS 4x + Vaxigrip + Adda-Vax

附图6