

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2023 年 5 月 4 日 (04.05.2023)



(10) 国际公布号
WO 2023/070873 A1

(51) 国际专利分类号:
C07K 14/165 (2006.01) *A61K 39/215* (2006.01)
C12N 15/50 (2006.01) *A61P 31/14* (2006.01)
C12N 15/85 (2006.01)

(21) 国际申请号: PCT/CN2021/138037

(22) 国际申请日: 2021 年 12 月 14 日 (14.12.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202111271508.1 2021年10月29日 (29.10.2021) CN

(71) 申请人: 中国科学院深圳先进技术研究院 (SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY CHINESE ACADEMY OF SCIENCES) [CN/CN]; 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。

(72) 发明人: 戴俊彪 (DAI, Junbiao); 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。马英新 (MA, Yingxin); 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。毛国斌 (MAO, Guobin); 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。

(74) 代理人: 北京中巡通大知识产权代理有限公司 (BEIJING ZHONG XUN TONG DA INTELLECTUAL PROPERTY AGENCY CO., LTD.); 中国北京市西城区榆树馆胡同东口 1 层 106, Beijing 100195 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

(54) Title: METHOD FOR PREPARING SARS-COV-2 VIRUS-LIKE PARTICLES AND USE OF SARS-COV-2 VIRUS-LIKE PARTICLES

(54) 发明名称: SARS-CoV-2病毒样颗粒的制备方法及其应用

(57) Abstract: A composition for preparing SARS-CoV-2 virus-like particles. The composition comprises a first plasmid containing a Stru ΔS fragment, and any one or two selected from a second plasmid containing an S fragment and a third plasmid containing a packaging signal fragment of the SARS-CoV-2 virus. The method for preparing the composition comprises: respectively splitting a Stru ΔS fragment, an S fragment and/or an ORF1ab packaging signal fragment into short DNA fragments having homologous sequences between adjacent fragments, and performing homologous recombination on the short DNA fragments and a linear plasmid vector in a yeast cell. The present invention further relates to a method for preparing SARS-CoV-2 virus-like particles using the composition, the obtained SARS-CoV-2 virus-like particles, and the use of the composition in the preparation of a vaccine for preventing or treating SARS-CoV-2 virus infections and in the in-vitro research on cells infected by the SARS-CoV-2 virus.

(57) 摘要: 一种用于制备 SARS-CoV-2 病毒样颗粒的组合物, 包括包含 StruΔS 片段的第一质粒、以及选自包含 S 片段的第二质粒包含 SARS-CoV-2 病毒的包装信号片段的第三质粒的任一种或两种; 该组合物的制备方法包括将 StruΔS 片段、S 片段和/或 ORF1ab 包装信号片段分别拆分为在相邻片段之间具有同源序列的 DNA 短片段, 与线性化质粒载体在酵母细胞中进行同源重组。还涉及用所述组合物制备 SARS-CoV-2 病毒样颗粒的方法、制备获得的 SARS-CoV-2 病毒样颗粒、以及所述组合物在制备用于预防或治疗 SARS-CoV-2 病毒感染的疫苗和在体外用于 SARS-CoV-2 病毒感染细胞的研究中的用途。

本国际公布：

- 包括国际检索报告(条约第21条(3))。
- 包括说明书序列表部分(细则5.2(a))。

SARS-CoV-2 病毒样颗粒的制备方法及其应用

技术领域

本发明涉及分子生物学技术领域，特别是涉及 SARS-CoV-2 病毒样颗粒的制备方法及其应用。

背景技术

病毒样颗粒（VLP）是一种由病毒单一或多个结构蛋白自行装配而成的高度结构化的蛋白颗粒。VLP 在宿主内无法实现复制繁殖，但是其包含部分或全部蛋白，仍然可以很好地引起宿主的强免疫响应。因此，VLP 可以作为一种潜在的高效疫苗制剂。一般而言，所有的 VLP 均可形成独特的结构，其表面布满了重复的蛋白结构，其较强的免疫原性使其可作为疫苗。VLP 在疫苗研究中倍受关注，目前上市或在研的 VLP 疫苗中，人用 VLP 疫苗主要有 HBV 疫苗、HPV 疫苗以及 HEV 疫苗，而兽用 VLP 疫苗为用于预防猪圆环病毒 2 型。在研的人用 VLP 疫苗还涉及了多种病毒，包括流感病毒、艾滋病病毒、埃博拉病毒等，在研兽用 VLP 疫苗则还有口蹄疫病毒、猪细小病毒疫苗、猪瘟病毒等。然而，目前大部分的 VLP 中只含有部分或者全部的结构蛋白，在对非结构蛋白、辅助蛋白等方面的研究受限，且作为疫苗引起的免疫反应有限。

新型冠状病毒(SARS-CoV-2)是一种单链 RNA 病毒，具有较高的病毒变异能力，其变异毒株具有更高的病毒感染力，因此，疫苗的开发尤为重要。尽管基于多种技术的 SARS-CoV-2 疫苗已经进入临床阶段，基于 SARS-CoV-2 VLP 疫苗研究较少。其中一个关键的原因是缺乏一个简单，高效的 SARS-CoV-2 VLP 合成与制备平台。

发明内容

鉴于目前存在的技术问题，本发明提供了一种 SARS-CoV-2 VLP 从头人工合成及制备方法，包括将 SARS-CoV-2 基因组拆分为多个片段，合成这些片段并将其组装在载体上，转染包装细胞获得 VLP。

本发明的第一方面提供一种用于制备 SARS-CoV-2 病毒样颗粒的组合物，包括下述（a），还包括选自下述（b）和（c）中的任一种或任两种：

（a）包含 StruΔS 片段的第一质粒，所述ΔS 片段包含编码 SARS-CoV-2 病毒的至少一种结构蛋白或其突变体和/或至少一种辅助蛋白或其突变体的核酸序列，且所述ΔS 片段不包含编码 SARS-CoV-2 病毒的 S 蛋白或其突变体的核酸序列；

（b）包含 S 片段的第二质粒，所述 S 片段包含编码 SARS-CoV-2 病毒的 S 蛋白或其突变体的核酸序列；

（c）包含 SARS-CoV-2 病毒的包装信号片段的第三质粒，所述包装信号片段包含 SARS-CoV-2 病毒的包装信号序列。

在一些实施方案中，所述的 S 蛋白的突变体相对于 SARS-CoV-2 病毒的 S 蛋白的突变包括 N331Q、N501Y、D614G 和/或 P681H。

在一些实施方案中，所述 StruΔS 片段包含下述任一种：

（1）编码 SARS-CoV-2 病毒的除 S 蛋白之外的全部结构蛋白和全部辅助蛋白的核酸序列；

(2) 与 (1) 相比, 所述结构蛋白和辅助蛋白中的至少一种相对于野生型 SARS-CoV-2 病毒具有突变; 和

(3) 与 (1) 相比, 缺失所述结构蛋白和辅助蛋白中的任一种。

在一些实施方案中, 所述包装信号序列是 NCBI 序列号 NC_045512.2 的第 19900-20000 位核苷酸或 NCBI 序列号 NC_045512.2 的第 19773-20335 位核苷酸。

在一些实施方案中, 所述包装信号片段包含 SARS-CoV-2 病毒基因组的 ORF1ab。

本发明的另一方面提供了上述组合物的制备方法, 包括分别通过下述方法制备第一质粒、第二质粒和/或第三质粒:

(1) 将 StruΔS 片段、S 片段和/或 ORF1ab 包装信号片段分别拆分为在相邻片段之间具有同源序列的 DNA 短片段;

(2) 将由步骤 (1) 获得的 DNA 短片段和线性化质粒载体转入酵母细胞中, 进行同源重组, 分别获得第一质粒、第二质粒和/或第三质粒。

本发明的另一方面提供了 SARS-CoV-2 病毒样颗粒的制备方法, 包括: 用上述任一种的组合物转染包装细胞, 获得 SARS-CoV-2 病毒样颗粒。

在一些实施方案中, 该制备方法, 还包括分别通过下述方法制备第一质粒、第二质粒和/或第三质粒:

(1) 将 StruΔS 片段、S 片段和/或包装信号片段分别拆分为在相邻片段之间具有同源序列的 DNA 短片段;

(2) 将由步骤 (1) 获得的 DNA 短片段和线性化质粒载体转入酵母细胞中, 进行同源重组, 分别获得第一质粒、第二质粒和/或第三质粒。

在一些实施方案中, 所述制备方法, 进一步包括将第一质粒、第二质粒和/或第三质粒分别在大肠杆菌中进行扩增, 然后转染包装细胞。

本发明的另一方面提供由上述制备方法制备获得的 SARS-CoV-2 病毒样颗粒。

本发明的另一方面提供由上述组合物或由上述制备方法制备获得的 SARS-CoV-2 病毒样颗粒在制备用于预防或治疗 SARS-CoV-2 病毒感染的疫苗中的用途。

本发明的另一方面提供由上述组合物或由上述制备方法制备获得的 SARS-CoV-2 病毒样颗粒在体外用于 SARS-CoV-2 病毒感染细胞的研究中的用途。

所述 SARS-CoV-2 病毒感染细胞的研究可以是研究 VLP 与细胞的相互作用、研究病毒对细胞的侵染过程和/或研究病毒成分的作用等。

本发明与现有技术相比, 具有以下优点和效果:

A、本发明利用哺乳细胞真核表达系统表达 SARS-CoV-2 的结构蛋白、非结构蛋白及辅助蛋白等以组装成 VLP, 可用于疫苗, 该疫苗制备方法不涉及活病毒。相比于灭活病毒疫苗和减毒活疫苗, 该方法安全性更好; 相比于多肽或核酸疫苗, 该方法免疫原性更好。

B、本发明采用病毒基因组拆分的方法, 先合成小片段基因, 再通过酵母同源重组系统进行基因组组装, 该方法简便、快速且效率高。

C、本发明利用基因组拆分策略, VLP 中含有结构蛋白与非结构蛋白等完整的病毒粒子形貌, 免疫原性强, 有望发展成为具有交叉保护效力的高效疫苗。

附图说明

通过以下详细的描述并结合附图将更充分地理解本发明, 其中相似的元件以相似的方式编号, 其中:

图 1 显示了 SARS-CoV-2 假病毒的生成。(A) 全长 SARS-CoV-2 基因组和拆分部分的方案, 包括 SARS-CoV-2 S、ORF1ab 和 StruΔS 基因组。(B) 分离的 SARS-CoV-2 基

基因组在酵母中组装以构建质粒 SARS-CoV-2 S、ORF1ab 和 StruΔS。(C) SARS-CoV-2 假病毒是通过共转染质粒 SARS-CoV-2 S、ORF1ab 和 StruΔS 获得的。SARS-CoV-2 假病毒感染 293T/hACE2 细胞后不能产生子代病毒。

图 2 显示了 SARS-CoV-2 假病毒的构建和表征。(A) 转染后 24 小时,共转染质粒 SARS-CoV-2 S、ORF1ab-mCherry 和 StruΔS-EGFP 的 293T 细胞的共聚焦显微镜图像。(B) 通过蛋白质印迹检测细胞裂解液(上)和 SARS-CoV-2 假病毒(下)中的 SARS-CoV-2 S 和 N 蛋白。GAPDH 用作蛋白质印迹的上样对照。(C) TEM 检测到的 SARS-CoV-2 假病毒和 VLPs 的形态和大小。比例尺:20 nm。通过共转染质粒 SARS-CoV-2 S、ORF1ab-mCherry 和 StruΔS-EGFP 获得的 SARS-CoV-2 假病毒用于 293T/hACE2 细胞侵染的共聚焦显微镜图像(D)和实验方案(E)。

图 3 显示了包装信号(PS)的分析和验证。(A)使用 Biopython 软件分析 SARS-CoV-2、SARS-CoV 和类蝙蝠 SARS-CoV 的 ORF1b 区域的序列(nt 19500-20400)。这三种 SARS 相关病毒中相同位置的二级结构由 RNAstructure Web Server 使用默认参数预测,并由 Vienna RNA Web Server 进行分析。两个相同的 RNA 茎环(SL1 和 SL2)用矩形框指示。(B)由质粒 EGFP-PS583 转染的细胞裂解液中 PS583 基因的 RT-PCR 分析。红色箭头指示预期的扩增序列。nRT:未逆转录组。(C)通过蛋白质印迹检测细胞裂解液(左)和 VLP(EGFP-PS583)(右)中的 EGFP、SARS-CoV-2 S 和 N 蛋白。GAPDH 用作蛋白质印迹和 RT-PCR 的上样对照。用 VLP(EGFP-PS583)侵染 293T/hACE2 细胞 48 小时后的共聚焦显微镜成像(D)和实验方案(E)。用 VLP(EGFP-PS583)侵染 293T 细胞和 VLP(EGFP)侵染 293T/hACE2 细胞作为对照。

图 4 显示了不同 SARS-CoV-2 蛋白的功能验证。(A)细胞质中 DiO 标记的 SARS-CoV-2 VLP 的统计分析,以测试 SARS-CoV-2 S 突变对病毒侵染能力的影响。(B)VLP(EGFP-PS583)、VLP(ΔN-EGFP-PS583)或 VLP(EGFP)侵染 293T/hACE2 细胞后的共聚焦显微镜成像图,观察侵染后 48 小时后细胞内的绿色荧光。(C)通过蛋白质印迹检测不同类型 VLP 中的 S 和 N 蛋白。通过 TEM 观察 VLP 的形态和大小。比例尺:20 nm。

图 5 显示了(A)双荧光标记 VLP(QD-DiO)的方案。(B)通过 TEM 表征 VLP(QD-DiO)的形态和大小,白色箭头指示 QD。比例尺:20 nm。(C)VLP(QD-DiO)中 DiO 和 QD 信号的共定位。(D)VLP(QD-DiO)进入 293T/hACE2 细胞的动态过程。(E)293T/hACE2 细胞的微分干涉对比(DIC)图像。黑线指示了病毒的轨迹。(F,G)分析(D)中显示的病毒粒子的平均速度(F)和 MSD 图(G)。(H)VLP(QD-DiO)的 RNA-QD 和 Env-DiO 在 293T/hACE2 细胞中实时解离的动态过程。(I-K)RNA-QD 和 Env-DiO 的轨迹(I)、平均速度(J)和 MSD 图(K)。

图 6 显示了 SARS-CoV-2 S 质粒的构建和表征。(A)细胞裂解液中 SARS-CoV-2 S 蛋白编码基因的 RT-PCR 分析。红色箭头指示预期的扩增序列。nRT:未逆转录组。

(B)通过蛋白质印迹检测细胞裂解液中 SARS-CoV-2 S 蛋白,GAPDH 用作蛋白质印迹和 RT-PCR 的上样对照。

图 7 显示了 SARS-CoV-2 StruΔS-EGFP 质粒的构建和表征。(A)StruΔS-EGFP 构建的示意图。(B)细胞裂解液中 SARS-CoV-2 ORF3a 和 N 蛋白编码基因的 RT-PCR 分析,红色箭头指示预期的扩增序列,GAPDH 用作上样对照,nRT:未逆转录组。(C)SARS-CoV-2 StruΔS-EGFP 质粒转染 293T 细胞 24 小时后的荧光成像。(D)通过蛋白质印迹检测细胞裂解液(左)和 VLP(右)的 N 蛋白,GAPDH 用作上样对照。(E)通过透射电子显微镜检测 SARS-CoV-2 VLP(StruΔS)的形态和大小。比例尺:20nm。

图 8 显示了 SARS-CoV-2 ORF1ab-mCherry 质粒的构建和表征。(A) ORF1ab-mCherry 构建的示意图。(B) 细胞裂解物中 SARS-CoV-2 NSP1 和 NSP16 编码基因的 RT-PCR 分析, 红色箭头指示预期的扩增序列, GAPDH 用作上样对照, nRT: 未逆转录组。(C) SARS-CoV-2 ORF1ab-mCherry 质粒转染 293T 细胞 24 小时后的荧光成像。

图 9 显示了 SARS-CoV-2 VLP (StruΔS-S) 的构建和表征。(A) SARS-CoV-2 StruΔS-EGFP 和 S 质粒共转染 293T 细胞 24 小时后的荧光成像。(B) 通过蛋白质印迹检测细胞裂解液(上)和 VLP (StruΔS-S) (下) 的 N 蛋白, GAPDH 用作上样对照。(C) 通过透射电子显微镜检测 SARS-CoV-2 VLP (StruΔS-S) 的形态和大小。比例尺: 20nm。

图 10 显示了 SARS-CoV-2 VLP (StruΔS-ORF1ab) 的构建和表征。(A) SARS-CoV-2 StruΔS-EGFP 和 ORF1ab-mCherry 质粒共转染 293T 细胞 24 小时后的荧光成像。(B) 通过蛋白质印迹检测细胞裂解液(上)和 VLP (StruΔS-ORF1ab) (下) 的 N 蛋白, GAPDH 用作上样对照。(C) 通过透射电子显微镜检测 SARS-CoV-2 VLP (StruΔS-ORF1ab) 的形态和大小。比例尺: 20nm。

图 11 显示了假 SARS-CoV-2 病毒生物安全性的验证。(A) 通过蛋白质印迹检测构建 293T/hACE2 细胞上 hACE2 受体的表达。(B) 通过蛋白质印迹检测细胞裂解液(左)和上清(右)的 S 和 N 蛋白。(C) 通过 VLP 或质粒 SARS-CoV-2 S, ORF1ab-mCherry 和 Stru ΔS-EGFP 共转染后感染 293T/hACE2 细胞 48 小时, GAPDH 用作上样对照。(C) 感染野生型 SARS-CoV-2 病毒和 SARS-CoV-2 VLP 后 72 小时后的 293T/hACE2 细胞图像。

图 12 显示了基于比对其有序列的 RNA 结构预测。多重比对 LocARNA 使用默认参数预测的蝙蝠 SARS 样冠状病毒 PS (MG772933.1:19773-20355), SARS PS (NC-004718.3:19712-20294) 和 SARS-CoV-2 PS (NC-045512.2:19773-20355) 的共有结构。

图 13 显示了包装信号的分析 and 验证。(A) 通过 RT-PCR 检测细胞裂解液中 PS101 基因, 红色箭头指示预期的扩增序列, GAPDH 用作上样对照, nRT: 未逆转录组。

(B) 通过蛋白质印迹检测细胞裂解液(左)和 VLP (右) 的 S, N 蛋白和 EGFP, GAPDH 用作上样对照。(C) VLP (EGFP-PS101) 和 VLP (EGFP) 感染 293T/hACE2 细胞 48 小时后的荧光成像。

图 14 显示了 SARS-CoV-2 S (N331Q) 构建和表征。(A) SARS-CoV-2 S (N331Q) 构建的示意图。(B) 细胞裂解液中 SARS-CoV-2 S (N331Q) 编码基因的 RT-PCR 分析, 红色箭头指示预期的扩增序列, nRT: 没有逆转录。(C) 通过蛋白质印迹检测细胞裂解液中 SARS-CoV-2 S (N331Q) 蛋白, GAPDH 用作蛋白质印迹和 RT-PCR 的上样对照。

图 15 显示了 SARS-CoV-2 S (N501Y) 构建和表征。(A) SARS-CoV-2 S (N501Y) 构建的示意图。(B) 细胞裂解物中 SARS-CoV-2 S (N501Y) 编码基因的 RT-PCR 分析, 红色箭头表示预期的扩增子, nRT: 未逆转录组。(C) 通过蛋白质印迹检测细胞裂解液中 SARS-CoV-2 S (N501Y) 蛋白, GAPDH 用作蛋白质印迹和 RT-PCR 的上样对照。

图 16 显示了 SARS-CoV-2 S (D614G) 构建和表征。(A) SARS-CoV-2 S (D614G) 构建的示意图。(B) 细胞裂解液中 SARS-CoV-2 S (D614G) 编码基因的 RT-PCR 分析, 红色箭头指示预期的扩增序列, nRT: 未逆转录组。(C) 通过蛋白质印迹检测细胞裂解液中 SARS-CoV-2 S (D614G) 蛋白, GAPDH 用作蛋白质印迹和 RT-PCR 的上样对照。

图 17 显示了 SARS-CoV-2 S (P681H)构建和表征。(A) SARS-CoV-2 S (P681H)构建的示意图。(B) 细胞裂解液中 SARS-CoV-2 S (P681H) 编码基因的 RT-PCR 分析, 红色箭头指示预期的扩增序列, nRT: 未逆转录组。(C) 通过蛋白质印迹检测细胞裂解液中 SARS-CoV-2 S (P681H) 蛋白, GAPDH 用作蛋白质印迹和 RT-PCR 的上样对照。

图 18 显示了具有 S 蛋白突变的 SARS-CoV-2 VLP 的构建和表征。通过蛋白质印迹检测 VLP (N331Q) (A), VLP (N501Y) (B), VLP (D614G) (B) 和 VLP (P681H) (D)。通过透射电子显微镜检测 SARS-CoV-2 VLP 的形态和大小。比例尺: 20nm。

图 19 显示了 SARS-CoV-2 StruΔS-EGFP/ΔN 质粒的构建和表征。(A) SARS-CoV-2 StruΔS-EGFP/ΔN 质粒构建示意图。(B) 细胞裂解液中 SARS-CoV-2 ORF3a 和 N 蛋白编码基因的 RT-PCR 分析, 红色箭头指示预期的扩增序列, nRT: 未逆转录组。(C) 通过蛋白质印迹检测细胞裂解液中 SARS-CoV-2 N 蛋白, GAPDH 用作蛋白质印迹和 RT-PCR 的上样对照。

图 20 显示了 SARS-CoV-2 StruΔS-EGFP/ΔE 质粒的构建和表征。(A) SARS-CoV-2 StruΔS-EGFP/ΔE 质粒构建示意图。(B) 细胞裂解液中 SARS-CoV-2 ORF3a, E 和 N 蛋白编码基因的 RT-PCR 分析, 红色箭头指示预期的扩增序列, nRT: 未逆转录组。(C) 通过蛋白质印迹检测细胞裂解液中 SARS-CoV-2 N 蛋白, GAPDH 用作蛋白质印迹和 RT-PCR 的上样对照。

图 21 显示了 E 蛋白缺失的 SARS-CoV-2 VLP (VLP/ΔE) 的构建和表征。(A) 通过透射电子显微镜检测 VLP/ΔE 的形态和大小。比例尺: 20nm。(B) VLP/ΔE 感染 293T/hACE2 细胞 48 小时后的共聚焦显微镜成像图。

图 22 显示了 SARS-CoV-2 StruΔS-EGFP/ΔM 质粒的构建和表征。(A) SARS-CoV-2 StruΔS-EGFP/ΔM 质粒构建示意图。(B) 细胞裂解液中 SARS-CoV-2 ORF3a, M 和 N 蛋白编码基因的 RT-PCR 分析, 红色箭头指示预期的扩增序列, nRT: 未逆转录组。(C) 通过蛋白质印迹检测细胞裂解液中 SARS-CoV-2 N 蛋白, GAPDH 用作蛋白质印迹和 RT-PCR 的上样对照。

图 23 显示了 M 蛋白缺失的 SARS-CoV-2 VLP (VLP/ΔM) 的构建和表征。(A) 通过透射电子显微镜检测 VLP/ΔM 的形态和大小。比例尺: 20nm。(B) VLP/ΔM 感染 293T/hACE2 细胞 48 小时后的共聚焦显微镜成像图。

图 24 显示了 SARS-CoV-2 StruΔS-EGFP/ΔORF10 质粒的构建和表征。(A) SARS-CoV-2 StruΔS-EGFP/ΔORF10 质粒构建示意图。(B) 细胞裂解液中 SARS-CoV-2 ORF3a, N 和 ORF10 蛋白编码基因的 RT-PCR 分析, 红色箭头指示预期的扩增序列, nRT: 未逆转录组。(C) 通过蛋白质印迹检测细胞裂解液中 SARS-CoV-2 N 蛋白, GAPDH 用作蛋白质印迹和 RT-PCR 的上样对照。

图 25 显示了蛋白缺失的 SARS-CoV-2 VLP 的表征。(A) 蛋白缺失 SARS-CoV-2 VLP 的质粒构建示意图。(B) 蛋白缺失 SARS-CoV-2 VLP 的荧光共定位结果 (DiO 与 QDs)。

具体实施方式

除非另外指出, 本发明中所用的术语具有本领域通常所理解的含义, 可以通过参考本领域技术人员已知的标准教科书, 参考书、文献加以理解。本文所提供的定义和解释用于明确它们在本发明范围内的特定用途。本发明所引用的参考文献均通过引用将其全文并入本文。

不希望受到任何特定理论的束缚，本发明中对于相关理论或机理解释仅仅是用于帮助理解发明，不应视为对本发明所保护的方案的限制。

本发明中所使用的术语“包含”，或与其具有类似含义的其他术语形式“包括”、“含有”或“具有”等，应当理解为包括所列出的要素，但不排除其他要素的存在；这些术语也包括仅由所列出的要素组成的情况。术语“由……组成”是指仅由所列举的要素组成。术语“基本上由……组成”表示不排除对所涉及方案没有显著影响的要素。

除非另外特别说明，本文中的“a”、“an”或“the”包括单数形式和复数形式。术语“至少一个”或“至少一种”或其它类似表述的含义等同于“一个或多个”或“一种或更多种”的含义。

本文中所述的数值范围，例如温度范围，时间范围，组成或浓度范围，或其他数值范围等，包括该范围内的端值、所有中间范围、子范围（例如在某个中间值和某个端值之间的范围）以及所有单个数值，特别是以整数数值为端值的中间范围、子范围和单个的整数数值。并且，所述数值范围中所描述的任何中间范围、子范围以及所有单个数值可以从所述的数值范围中排除。

本文中所述的“约”、“大约”的意思为所述数值的 $\pm 20\%$ 、 $\pm 18\%$ 、 $\pm 15\%$ 、 $\pm 12\%$ 、 $\pm 10\%$ 、 $\pm 9\%$ 、 $\pm 8\%$ 、 $\pm 7\%$ 、 $\pm 6\%$ 、 $\pm 5\%$ 、 $\pm 4\%$ 、 $\pm 3\%$ 、 $\pm 2\%$ 、 $\pm 1\%$ 或 $\pm 0.5\%$ 的范围，所述范围包扩所述范围的端点及所述范围内的任一数值。

本文中所述的术语“和/或”应当理解为由该术语连接的多个要素中的任何一个要素、任意两个要素、任意三个要素或任意的更多个要素组成的组合。

本发明中，术语“核酸序列”和“核苷酸序列”可以互换使用，是指由碱基、糖和磷酸骨架连接构成的序列。本发明的核酸序列可以是脱氧核糖核酸序列（DNA）或核糖核酸序列（RNA），且可包括天然碱基或非天然碱基，其可以是单链或双链，可以是编码序列或非编码序列。

本发明中，当表达基因或基因所编码的蛋白时，“载体”可以是质粒，在一些情况下，术语“载体”和“质粒”可以互换使用。

本发明中，除非另有说明，核酸序列的描述通常是从5'端至3'端，氨基酸序列的描述通常是从N端至C端。

本发明的方法可以是在体外或体内进行。

反向遗传学是病毒拯救（rescuing）或VLP制备的主要手段之一，病毒基因组的获得是其中最为重要的一步。核酸酶连接是获取病毒基因组的主要方法，但是该方法反应过程复杂且效率较低，同时，病毒全长基因组在转化大肠时存在较大的毒性。酵母体内同源重组法是利用酵母细胞内高效的同源重组系统来实现多个相互存在同源系列DNA片段的组装方法。本发明通过独特的基因组拆分设计，以及基因组的合成与酵母体内同源重组组装技术，将SARS-CoV-2的基因组拆分并组装成三个质粒，通过细胞转染，即可获得结构完整性高的SARS-CoV-2 VLP，为进一步的疫苗开发提供了很好的工具平台。

具体而言，本发明提供了SARS-CoV-2的基因组的三个片段，每一个片段被组装到一个质粒中，共获得三个质粒，使用这三个质粒中的任一个、任两个或三个共转染包装细胞，可以获得SARS-CoV-2的VLP。

所述的SARS-CoV-2可以是野生型的SARS-CoV-2或者是其任何突变体。野生型SARS-CoV-2的基因组可由NCBI序列号NC_045512.2确定。

所述三个片段分别是包含 SARS-CoV-2 的 S 蛋白的编码序列的 S 片段、包含除 S 蛋白之外的所有结构蛋白和所有辅助蛋白的编码序列的 StruΔS 片段、以及包含包装信号序列的包装信号片段。

SARS-CoV-2 的全部结构蛋白包括 S 蛋白（刺突糖蛋白）、E 蛋白（小包膜蛋白）、M 蛋白（膜糖蛋白）和 N 蛋白（核衣壳蛋白）。SARS-CoV-2 的全部辅助蛋白包括 ORF3a、ORF3b、ORF6、ORF7a、ORF7b、ORF8 和 ORF10 编码的蛋白。

所述 S 片段编码的 S 蛋白（刺突糖蛋白）可以是野生型 S 蛋白或其突变体。在一些实施方案中，S 蛋白突变体相对于野生型 S 蛋白具有选自 N331Q、N501Y、D614G 及 P681H 的任一个或更多个的突变。野生型 S 蛋白序列可以由 NCBI 序列号 NC_045512.2 的 21563-25384 位核苷酸编码的蛋白质序列，或者由 NCBI 序列号 YP_009724390.1 确定。在一些实施方案中，S 片段中包含的编码 S 蛋白的序列是 NCBI 序列号 NC_045512.2 的第 21563-25384 位核苷酸。

所述 StruΔS 片段编码结构蛋白和辅助蛋白中的任一种或任几种，例如 E 蛋白（小包膜蛋白）、M 蛋白（膜糖蛋白）、N 蛋白（核衣壳蛋白）中的任一种、任两种或三种，和/或辅助蛋白中的任一种或任几种，这些蛋白独立地可以是野生型或其突变体。虽然最优选地，本发明的 StruΔS 片段编码除 S 蛋白之外的所有结构蛋白和辅助蛋白，但在本发明的一些实施例，相比于该最优选的方案而言，所述 StruΔS 片段所编码的蛋白也可以进一步缺少一种或更多种结构蛋白，和/或缺少一种或更多种辅助蛋白，例如，所述 StruΔS 片段所编码的蛋白相比于该最优选的方案，可以不包含 E 蛋白（小包膜蛋白）、M 蛋白（膜糖蛋白）、N 蛋白（核衣壳蛋白）中的任一种、任两种或三种，和/或可以不包含辅助蛋白中的任一种或更多种。野生型的结构蛋白和/或辅助蛋白是指野生型 SARS-CoV-2 的结构蛋白和/或辅助蛋白。在一些实施方案中，StruΔS 片段中包含的编码结构蛋白和辅助蛋白的核酸是 NCBI 序列号 NC_045512.2 的第 25393-29674 位核苷酸。

所述包装信号片段至少包含 SARS-CoV-2 的包装信号序列，还可以包含 SARS-CoV-2 的 ORF1ab 序列中的其他序列。在一些实施方案中，所述包装信号序列为 NCBI 序列号 NC_045512.2 的第 19900-20000 位核苷酸或 NCBI 序列号 NC_045512.2 的第 19773-20335 位核苷酸。在一些实施方案中，所述包装信号片段可以包含 ORF1ab 的全长或其片段，即 SARS-CoV-2 的 ORF1ab 全长或其片段。ORF1ab 的全长或片段中所任一种或任几种 ORF 可以是野生型或其突变体。在本发明的一些实施方案中，所述 ORF1ab 的片段可以缺少 ORF1ab 序列中的一种或更多种 ORF，但至少包含包装信号序列。野生型的 ORF1ab 序列中的 ORF 是指野生型 SARS-CoV-2 的 ORF1ab 序列中的 ORF。在一些实施方案中，ORF1ab 是 NCBI 序列号 NC_045512.2 的第 266-21555 位核苷酸。

如本文所使用的，术语“野生型”可用于基因、蛋白质或病毒株，通常是指从天然存在的来源所分离的。野生型通常是在种群中最频繁观察到的基因、蛋白质或病毒株，因而指定为“野生型”。术语“突变体”是指当与野生型基因、蛋白质或病毒株相比，在基因的核苷酸序列、蛋白质的氨基酸序列或病毒株的基因组序列具有一个或更多个核苷酸或氨基酸的插入、取代或缺失。“突变体”通常仍具有野生型的基本功能特性，但其功能特性的相对水平与野生型相比可能发生改变，例如蛋白质或基因编码的蛋白质的活性有所提高或降低。应当理解，“突变体”也可能是天然存在的，例如病毒在会发生自然变异，导致产生天然存在的病毒突变体，“野生型”可以指同一基因、蛋白质或病毒株中最早出现的天然存在的那些。本发明中，野生型 SARS-CoV-2 的基因组可由 NCBI 序列号 NC_045512.2 确定，野生型的各个结构蛋白、辅助蛋白、和/或 ORF1ab 中的各个 ORF 可以由 NCBI 序列号 NC_045512.2 所示的基因组序列编码或确定的。SARS-CoV-

2 在自然界存在和传播过程中通过自然变异而产生病毒以及该病毒基因组中包含的基因及其编码的蛋白质，均可以被称为是“突变体”。本发明中的“突变体”也包括那些对天然存在的野生型 SARS-CoV-2 和自然变异的 SARS-CoV-2 进行人工改造而得到的病毒、其基因组、其任一种基因和其任一种蛋白质的突变体。

5 本发明中，S 片段、StruΔS 片段和 ORF1ab 片段中的每一个被组装到一个质粒中，所述质粒用于表达这些片段中的基因或其编码的蛋白。在一些实施方案中，包含 StruΔS 片段的质粒被称为第一质粒，包含 S 片段的质粒被称为第二质粒，包含包装信号片段的质粒被称为第三质粒。对质粒的编号仅仅是为了进行区分，并非意味着这些质粒中隐含着特定的序列结构。

10 用于制备第一质粒、第二质粒和第三质粒的初始质粒载体可以是相同的或不同的，例如三者可以用相同的一种初始质粒载体，或者可以使用三种不同的初始质粒载体，或者可以是其中二者相同，但与另一者不同。用于制备第一质粒、第二质粒或第三质粒的初始质粒载体可以是本领域常见的任何适当的载体，例如本领域常见的可以在酵母和细菌之间穿梭，并可以转染包装细胞的任何适当的载体，包括但不限于 pEASY-T1、
15 pRS415、pJS356、pEGFP-N1 等。

第一质粒中包含的编码结构蛋白和/或辅助蛋白的序列、第二质粒中编码 S 蛋白的序列和/或第三质粒中的包装信号序列或 ORF1ab 的全长或其片段可以与表达调节序列可操作地连接。所述表达调节序列是质粒中所含的片段的表达和/或翻译具有调控作用的序列，包括但不限于 5' 非翻译区 (5'UTR)、3' 非翻译区 (3'UTR)、启动子、增强
20 子、终止子、选择标记基因、转录后调节元件、内部核糖体进入位点 (IRES)、可切割序列和/或多腺苷酸化信号 (polyA) 等。

可使用的启动子包括但不限于来自病毒或哺乳动物 (包括人) 等的启动子等。启动子可以是组成型的或诱导型的。病毒启动子的示例包括但不限于巨细胞病毒 (CMV) 立即早期启动子、RSV 启动子、鸡 β-肌动蛋白 (CBA) 启动子、CMV 早期增强子/鸡 β-
25 肌动蛋白 (CAG) 启动子、猿猴病毒 40 (SV40) 启动子等。哺乳动物启动子的示例包括但不限于人延伸因子 1a (EF1a) 启动子、人泛素 C (UCB) 启动子、小鼠磷酸甘油酸激酶 (PGK) 启动子、RNA 聚合酶 III 型启动子 (如 U6 和 H1) 等。组成型启动子的示例包括但不限于逆转录病毒劳氏肉瘤病毒 (RSV) LTR 启动子、巨细胞病毒 (CMV) 启动子、SV40 启动子、二氢叶酸还原酶启动子、β-肌动蛋白启动子和磷酸甘
30 油激酶 (PGK) 启动子和 EF1a 启动子。诱导型启动子的示例包括但不限于锌诱导的羊金属硫蛋白 (MT) 启动子、地塞米松 (Dex) 诱导的小鼠乳腺肿瘤病毒 (MMTV) 启动子、T7 聚合酶启动子系统等。

可使用的转录后调节元件例如可以是病毒转录后调节元件如土拨鼠肝炎病毒转录后调节元件 (WPRE)、乙型肝炎病毒转录后调节元件 (HBVPRE) 或 RNA 转运元件
35 (RTE) 等。

可使用的选择标记包括但不限于，例如药物抗性选择标记。这种选择标记基因可编码对于生长在选择培养基中的细胞的存活或生长必要的因子。未用包含选择基因的载体转化的宿主细胞将不会在培养基中存活。可以选择给予对抗生素或其他毒素的抗性、补充营养缺陷型的缺陷、或供给从培养基中扣留的关键营养元素的蛋白作为选择标记，抗
40 生素或其他毒素的示例包括但不限于氨苄青霉素、潮霉素、新霉素、甲氨喋呤、卡那霉素、庆大霉素、Zeocin、或四环素。选择标记还可以是荧光蛋白，例如 mCherry、GFP、BFP、EGFP 等。

可切割序列可以是任何在表达后能自切割或能通过其它方式被切割 (例如被特定的酶切割) 的序列。可切割序列可以是在表达后能够自切割的序列。这样的序列是本领域

技术人员熟知的，例如具有自切割功能的核酶序列，如剪切型核酶序列。剪切型核酶可以催化自身 RNA 在特定的位点自我剪切。剪切型核酶包括但不限于锤头核酶、发夹核酶、HDV 核酶或 RNaseP。可切割序列还可以是能通过其它方式被切割，例如被特定的酶切割的序列。这样的序列是本领域技术人员熟知的，例如 tRNA 序列。tRNA 序列能在 tRNA 5'成熟酶 (RZaseP) 作用下被切掉 tRNA 的 5'端附加序列，或者可以在 3'内切核酸酶 RZase F 的作用下被切掉 tRNA 3'端的序列。因此，所述可切割序列可以是可被 RZaseP 切割 5'端附加序列的 tRNA 序列，或者可以是可被 RZase F 切割 3'端序列的 tRNA 序列。任何在成熟时能够被切割的 tRNA 都可以用于本发明。

多腺苷酸化信号序列可以位于转录终止区，例如 3'非翻译区中，其示例包括但不限于牛生长激素 (bGH) poly(A)、SV40 polyA、胸苷激酶 (TK) poly(A)序列等。

5'非翻译区 (5'UTR)、3'非翻译区 (3'UTR) 可以使用任何适合的病毒 5'UTR 和 3'UTR，例如 SARS-CoV-2 的 5'UTR 和 3'UTR。

在一些实施方案中，第一质粒中编码结构蛋白和/或辅助蛋白的核酸与 5'端的启动子、5'UTR，以及 3'端的 3'UTR、polyA 和任意的可切割序列可操作地连接。在一些实施方案中，第二质粒中编码 S 蛋白的核酸与 5'端的启动子和 3'端的 polyA 可操作地连接。在一些实施方案中，第三质粒中来自 SARS-CoV-2 的 ORF1ab 的序列与 5'端的启动子、5'UTR，以及 3'端的 3'UTR、polyA 和任意的可切割序列可操作地连接。

在本发明中，StruΔS 片段、S 片段或 ORF1ab 片段可以通过转化介导的重组 (Transformation-associated recombination, TAR) 克隆方法被组装到第一质粒、第二质粒或第三质粒中。TAR 克隆的是利用酵母体内重组系统对大片段的目的 DNA 进行拼接合成，具体方法是：将具有同源末端的目的 DNA 的多个片段和线性化的 TAR 载体共同转染到酵母中，利用酵母细胞中高效的重组系统，使载体与目的 DNA 片段的同源序列以及不同目的 DNA 片段的同源序列间进行同源重组，即可产生带有目的 DNA 的载体。本发明中可以有 S 片段、StruΔS 片段或 ORF1ab 片段为目的 DNA 进行 TAR 克隆。所述 TAR 载体可以是质粒载体。具体而言，目的 DNA 被拆分为多个 DNA 片段，相邻 DNA 片段之间具有同源序列 (也可以叫同源臂、重叠序列或重叠片段)，位于目的 DNA 两端的 DNA 片段分别与线性化的载体两端具有同源序列，将这些 DNA 片段和线性化的载体共同转入酵母中，使得它们在酵母体内发生同源重组，以组装成含有目的 DNA 的载体。所述“相邻”是指两个 DNA 片段在目的 DNA 上的位置是相邻的。拆分的每个 DNA 片段的长度通常为 2-5kb，例如拆分的每个 DNA 片段的长度范围的下限和上限可以分别是约 2kb、约 2.1kb、约 2.2kb、约 2.3kb、约 2.4kb、约 2.5kb、约 2.6kb、约 2.7kb、约 2.8kb、约 2.9kb、约 3kb、约 3.1kb、约 3.2kb、约 3.3kb、约 3.4kb、约 3.5kb、约 3.6kb、约 3.7kb、约 3.8kb、约 3.9kb、约 4kb、约 4.1kb、约 4.2kb、约 4.3kb、约 4.4kb、约 4.5kb、约 4.6kb、约 4.7kb、约 4.8kb、约 4.9kb、约 5kb，优选为 3-4kb，更优选为约 3kb。同源序列的长度通常为 50-300bp，优选为 100-250bp，更优选为 150-200bp。

在一些实施方案中，StruΔS 片段可以被拆分为 3 个 DNA 片段。在一些实施方案中，S 片段可以被拆分为 2 个 DNA 片段。在一些实施方案中，ORF1ab 片段可以被拆分为 2 个 DNA 片段。

拆分后的 DNA 片段可以通过本领域已知的方法合成，如体外合成，所合成的 DNA 片段可以是单链或双链片段。

应当理解，本发明所述的 StruΔS 片段、S 片段或 ORF1ab 片段可以仅含有所要表达的 SARS-CoV-2 的基因，不含有其他序列 (例如编码序列或表达调控序列)，也可以包含所要表达的 SARS-CoV-2 的基因和前述的一种、两种或更多种表达调控序列。表达调

控序列可以通过任何方式被添加到载体中，例如，表达调控序列可以被包含在 StruΔS 片段、S 片段和/或 ORF1ab 片段中，与这些片段中所表达的 SARS-CoV-2 的基因一同被拆分为包含同源臂的多个 DNA 片段，并通过酵母同源重组（TAR）被组装到质粒中。前述的表达调控序列也可以被事先添加到质粒载体中，与 StruΔS 片段、S 片段或 ORF1ab 片段拆分的带有同源臂的 DNA 片段在酵母中进行同源重组，此时 StruΔS 片段、S 片段或 ORF1ab 片段仅含有所要表达的 SARS-CoV-2 的基因，而不包含表达调控序列。在另一些实施方案中，一些调控序列被包含在 StruΔS 片段、S 片段和/或 ORF1ab 片段中，与这些片段中所表达的 SARS-CoV-2 的基因一同被拆分为包含同源臂的多个 DNA 片段，并通过酵母同源重组（TAR）被组装到质粒中，另一些表达调控序列则被事先添加到质粒载体中。

DNA 片段和线性化的载体在酵母体内同源重组的方法是本领域技术人员熟知的，例如，一个具体的实例是：首先制备酵母感受态，从平板上挑取酵母菌株并接种到 5 mL 酵母浸出粉胨葡萄糖培养基（YPD），培养过夜；从扩摇的菌液中取出适量再次接种到 5 mL YPD，在 30 °C，200 rpm 条件下培养至 OD₆₀₀=0.4~0.6。清洗菌体，取待连接的 DNA 片段 150 ng、线性化的载体 100 ng，混合均匀，并加入感受态酵母细胞；随后，从同源臂的两端约 100 bp 设计引物扩增，挑取酵母单克隆过夜培养，按照 PCR 程序处理进行 junction PCR；验证正确后，从平板上挑取菌体，接种于 5 mL 液体 SC-LEU 培养基，培养过夜，收集菌体；最后，重悬菌体并提取酵母总 DNA 备用。

在酵母中组装好的含有目的 DNA 的载体，可以通过常规的质粒提取方法从酵母中提取出来，获得第一质粒、第二质粒或第三质粒，然后可选择地转入大肠杆菌扩增；或可以直接提取酵母总 DNA，用酵母总 DNA 转化大肠杆菌并扩增，获得第一质粒、第二质粒或第三质粒。本领域熟知从酵母中获得并扩增重组质粒的方法，例如可以将酵母总 DNA 转入大肠杆菌，挑选单克隆，验证正确后，大量扩增并提取质粒，保存质粒以备用。作为一个具体实例，首先挑取 EPI300 单克隆并接种于 5 mL LB 液体培养基，过夜培养，采用电转的方式转染酵母总 DNA，将转移菌液涂在抗性平板上挑取单克隆，按照 PCR 程序处理进行 junction PCR；随后选择验证正确的菌提取质粒，并进行酶切验证，酶切验证正确后进行测序验证；验证正确后，大量扩增大肠杆菌，并提取质粒，保存质粒以备用。本发明中，使用选自第一质粒、第二质粒和第三质粒中的任意两种或三种转染包装细胞，使他们在包装细胞中表达并组装，即可获得 SARS-CoV-2 的病毒样颗粒。在一些实施方案中，可以用于转染包装细胞的质粒包括第一质粒和选自第二质粒和第三质粒中的任一种或两种。

在进行转染之前，可以先对质粒的表达进行验证，例如可以将任一种质粒单独转染哺乳动物细胞（如 293T 细胞），裂解细胞后提取胞内 RNA 及蛋白并分别进行检测，例如可以通过 RT-PCR 和 Western blot 进行检测，还可以通过选择性标记，例如荧光蛋白成像的方法实现对质粒表达的验证。

本发明中，术语“病毒样颗粒（VLP）”、“假病毒”、“重组病毒”和“重组病毒颗粒（rVP）”可以互换使用，是指通过包含至少一种重组 DNA 技术步骤的方法获得的病毒颗粒，其中病毒基因组中包含的至少一种组分通过基因工程方法进行表达和组装。本发明的“病毒样颗粒”可以不含病毒基因组，或者含有不编码病毒蛋白的基因，因此是非复制性和非感染性的，所述“重组病毒颗粒”在本发明中指使用第一质粒、第二质粒和第三质粒中的任意两种或三种转染包装细胞获得的病毒颗粒，所述病毒颗粒中可以包含或不包含核酸物质。当用于转染包装细胞的质粒中包括具有包装信号的质粒时，所获得的病毒颗粒中包含核酸物质，当用于转染包装细胞的质粒中不包括具有包装

信号的质粒时，所获得的病毒颗粒中不包含核酸物质。例如，本发明中，rSAR2-CoV-2 与 SAR2-CoV-2 VLP 可以互换使用。

本发明所述的“转染”是指使多核苷酸，如核酸分子、质粒等，从细胞外转移到细胞内，使得该多核苷酸在所述细胞内具有功能。转染方法是本领域技术人员熟知的，例如磷酸钙或脂质体介导的转染。如本领域技术人员所知，可使用的脂质体包括 lipofectamine 8000 等。在一些实施方案中，转染方法包括使 293T 长至密度约为 80%，用质粒与 lipofectamine 8000 混合后转染 293T 细胞，培养基为添加了 10%FBS 的 DMEM；6 h 后，更换培养基为 DMEM，48 h 后收集上清溶液。

用于生产重组病毒颗粒的包装细胞是本领域技术人员熟知的，例如，可使用的包装细胞包括，但不限于，哺乳动物（包括人）细胞、昆虫细胞、植物细胞、微生物或酵母，例如 HEK293 系列的细胞（如 HEK293A、HEK293T 或 HEK293FT）、A549 细胞或 Vero 细胞等。

用于生产重组病毒颗粒的培养基和培养方法是本领域技术人员已知的，可以采用商购的或定制的培养基，或者进一步添加补充本领域已知的一个或多个细胞培养组分，包括但不限于葡萄糖、维生素、氨基酸和或生长因子，以增加生产培养物中重组病毒颗粒的滴度。重组病毒载体生产培养物可以在适合于所使用的特定宿主细胞（即包装细胞）的条件下生长。在产生重组病毒载体之后，如果需要，可使用多种常规方法从细胞裂解物或病毒上清液中纯化病毒颗粒，包括超滤法，例如使用硝酸纤维素滤膜；吸附法，例如使用磷酸钙或离子交换熟知吸附病毒或杂质，随后用盐溶液洗脱；层析法，例如葡聚糖凝胶层析、离子交换层析、亲和层析等；离心法，例如差速离心、CsCl 或蔗糖密度梯度离心；沉淀法，例如聚乙二醇沉淀（如使用 PEG2000），等电点沉淀或中性盐沉淀等。上述纯化方法也可以组合使用。在一些实施方案中，作为纯化方法的一个示例，配制浓度为 10% (w/v) 的 PEG20000 溶液，以体积比 1:1 混合 PEG20000 与培养的病毒上清溶液，4°C 放置 16 h 进行浓缩，离心收集沉淀物，然后用 opti-MEM 重悬以备，该方法可用于 VLP 的侵染等实验；作为纯化方法的另一个示例，将收集的病毒培养上清过 20% (w/v) 的蔗糖溶液，以 30000 rpm 超速离心 4 h 对 VLP 进行初步纯化，此 VLP 纯化方法可用于 RNA、蛋白及 VLP 形貌等的表征。

本发明的重组病毒颗粒可以作为活性成分用在药物组合物或疫苗中，以治疗由 SARS-CoV-2 引起的疾病。术语“疫苗”指包含所述重组病毒颗粒的制剂。疫苗中所包含的重组病毒颗粒的剂量可由本领域技术人员例如根据疾病状况、受试者和治疗时间表进行调整。疫苗中通常包含治疗有效量的重组病毒颗粒。本发明所述的“治疗”是指预防或缓解（例如，降低、减轻或治愈）至少一种与疾病状态相关的症状。本发明所述的“治疗有效量”即足以预防或缓解（例如，降低、减轻或治愈）至少一种与疾病状态相关的症状的量。所述药物组合物或疫苗的剂量可方便地由本领域技术人员来确定，例如通过首先鉴定有效引发预防性或治疗性免疫应答的剂量，例如通过测量病毒特异性免疫球蛋白的血清效价或通过测量血清样品或尿样或粘膜分泌物中抗体的抑制比。

本发明的药物组合物或疫苗可以包含本发明的重组病毒颗粒和药学上可接受的载体或赋形剂。药学上可接受的载体包括但不限于盐水、缓冲盐水、右旋糖、水、甘油、无菌等张水性缓冲液及其组合。本发明的药物组合物或疫苗可以包含佐剂，佐剂对于本领域技术人员而言是公知的。示例性的佐剂包括完全弗氏佐剂、不完全弗氏佐剂、氢氧化铝佐剂、BCG 等。

本发明的药物组合物或疫苗应当适合于施用给受试者，例如是是无菌的、非颗粒的和 / 或非致热原性的。所述药物组合物或疫苗可被配制为固体形式，如可用于配制注射

液的冻干粉、液体溶液、悬液、乳液、片剂、丸剂、胶囊剂、持续释放配制物或散剂等。

本发明的药物组合物或疫苗的施用方法包括但不限于肠胃外给药（例如皮内、肌肉内、静脉内和皮下）、硬膜外和粘膜（例如鼻内和经口或肺部途径或通过栓剂）。给药可为系统性的或局部的。

本发明的重组病毒颗粒还可以用在 SARS-CoV-2 病毒感染细胞的体外研究中，例如可用于体外药物筛选，如在体外筛选能够抑制 SARS-CoV-2 病毒感染细胞的药物；开发抗体，例如可以该重组病毒颗粒作为免疫原制备抗体或检测抗体中和活性；还可用于开发病毒疫苗，例如可以针对新出现的病毒突变体，特别是新出现的具有突变 S 蛋白的病毒突变体，方便快捷地开发疫苗。

在一些实施方案中，基因组的拆分、合成及组装原理如图 1 所示。首先，将基因组 ORF1ab 拆分成 6 个大小为 3-5 kb 的片段、S 蛋白基因拆分成 2 个片段、其他结构蛋白或辅助蛋白基因拆分为 3 个片段；然后，体外合成每个片段并通过酵母同源重组技术组装成 ORF1ab、S 及 StruΔS 三个片段；再将三个片段构建在相应载体上，转入大肠杆菌，获得质粒；最后将质粒转染哺乳动物细胞，获得 VLP。各种 VLP 的表征如图 3-4 所示。

在一些实施方案中，本发明提供了一种 SARS-CoV-2 VLP 从头人工合成及制备方法，具体包括：

第一步：SARS-CoV-2 基因组拆分，合成及组装

（1）将 SARS-CoV-2 基因组拆分成 ORF1ab、StruΔS 和 S 三个部分，其中 ORF1ab 序列从 266 到 21555，包含所有非结构蛋白的序列；StruΔS 序列从 25393 到 29674，包含除 S 蛋白外所有的结构蛋白与辅助蛋白的序列；S 序列从 21563 到 25384，包含 S 蛋白的序列（具体序列见附件）；

（2）将 ORF1ab 拆分成 6 个约 3 kb 的片段，将 StruΔS 拆分成 3 个约 3 kb 的片段，将 S 拆分成 2 个 3 kb 的片段，每个片段均采用体外合成的方法，并在相邻片段上设计 overlap 片段。

（3）将上一步骤拆分的片段利用酵母体内平台进行酵母同源重组，利用相邻片段间的 overlap 通过同源重组的方法组装成为 ORF1ab、StruΔS 及 S 三个完整的片段，同时分别组装到质粒载体 pJS356 和 pEASY-T-CAG 上得到三个质粒 SARS-CoV-2 ORF1ab、SARS-CoV-2 StruΔS 和 SARS-CoV-2 S。其中在 ORF1ab、StruΔS 编码序列两端分别加上 5' UTR 及 3' UTR 与 PolyA 尾巴，同时每条序列的 5' 加上 CMV 的增强子；在 S 蛋白编码序列的 3' 端加上 PolyA 尾巴；并在 5' 加上 CAG 的增强子；

酵母同源重组的具体操作步骤如下：首先制备酵母感受态，从平板上挑取酵母菌株并接种到 5 mL 酵母浸出粉胨葡萄糖培养基（YPD），培养过夜；从扩摇的菌液中取出适量再次接种到 5 mL YPD，在 30 °C，200 rpm 条件下培养至 OD₆₀₀=0.4~0.6。清洗菌体，取待连接的 DNA 片段（约 3 kb）150 ng、线性化的载体 100 ng，混合均匀，并加入感受态酵母细胞；随后，从同源臂的两端约 100 bp 设计引物扩增，挑取酵母单克隆过夜培养，按照 PCR 程序处理进行 junction PCR；验证正确后，从平板上挑取菌体，接种于 5 mL 液体 SC-LEU 培养基，培养过夜，收集菌体；最后，重悬菌体并提取酵母总 DNA 备用；

（4）将酵母总 DNA 转入大肠杆菌，挑选单克隆，验证正确后，大量扩增并提取质粒，保存质粒以备用；

具体操作步骤如下：首先挑取 EPI300 单克隆并接种于 5 mL LB 液体培养基，过夜培养，采用电转的方式转染酵母总 DNA，将转移菌液涂在抗性平板上挑取单克隆，按

照 PCR 程序处理进行 junction PCR; 随后选择验证正确的菌提取质粒, 并进行酶切验证, 酶切验证正确后进行测序验证; 验证正确后, 大量扩增大肠杆菌, 并提取质粒, 保存质粒以备用。

第二步: SARS-CoV-2 VLP 表达及制备方法

(1) 对 SARS-CoV-2 基因组拆分成而构建的三个质粒分别进行验证。首先分别将三个质粒单独转染 293T 细胞, 然后裂解细胞, 分别提取细胞内的 RNA 及蛋白, 并分别进行 RT-PCR 及 western blot 等的验证; 同时通过 SARS-CoV-2 StruΔS 中的 EGFP 和 SARS-CoV-2 ORF1ab 中的 mCherry 的荧光成像的方法实现对质粒表达的验证。

(2) 293T 长至密度约为 80%, 分别转染 SARS-CoV-2 StruΔS, SARS-CoV-2 StruΔS 与 SARS-CoV-2 S, SARS-CoV-2 StruΔS 与 SARS-CoV-2 ORF1ab, 及 SARS-CoV-2 StruΔS、SARS-CoV-2 S 及 SARS-CoV-2 ORF1ab, 培养基为 opti-MEM; 6 h 后, 更换培养基为 DMEM, 48 h 后收集上清溶液。

(3) 配置浓度为 10% (w/v) 的 PEG20000 溶液, 以体积比 1:1 混合 PEG20000 与病毒上清溶液, 4 度放置 16 h 进行浓缩, 离心收集沉淀物, 然后用 opti-MEM 重悬以备用, 以上方法可用于 VLP 的侵染等实验。将收集的上清过 20% (w/v) 的蔗糖溶液, 以 30000 rpm 超速离心 4 h 对 VLP 进行初步纯化, 此 VLP 纯化方法主要用于 RNA、蛋白及 VLP 形貌等的表征。

在一些实施方案中, 质粒共转制备 VLP 及其验证, 包括下述任一种步骤:

(1) 将 StruΔS 的质粒单转入 293T 细胞制备 StruΔS 的 VLP

将 293T 铺在 100 mm 的细胞培养皿, 至细胞密度约为 90% 时, 用转染试剂 lipo8000 在每个培养皿中转染 15 μg SARS-CoV-2 StruΔS, 6 h 后更换培养基。48 h 后收集上清溶液, 即为 VLP。将培养皿中细胞裂解进行 RT-PCR 与 western blot 以验证质粒在 293T 细胞中的表达; 采用荧光成像的方法验证质粒的表达, 采用 western blot 分析 VLP 中的 N 蛋白, 采用透射电镜表征的方法验证 VLP 的大小与形貌;

(2) 将 StruΔS 与 S 的质粒共转入 293T 细胞制备 StruΔS + S 的 VLP

将 293T 铺在 100 mm 的细胞培养皿, 至细胞密度约为 90 % 时, 用转染试剂 lipo8000 在每个培养皿中转染 15 μg SARS-CoV-2 StruΔS 与 15 μg SARS-CoV-2 S, 6 h 后更换培养基; 72 h 后收集上清溶液, 即为 StruΔS + S 的 VLP。采用荧光成像的方法验证质粒的表达, 采用 western blot 分析胞内表达及 VLP 中的 S 蛋白和 N 蛋白; 采用透射电镜表征的方法验证 VLP 的大小与形貌;

(3) 将 StruΔS 与 ORF1ab 的质粒共转入 293T 细胞制备 StruΔS + ORF1ab 的 VLP

将 293T 铺在 100 mm 的细胞培养皿, 至细胞密度约为 90 % 时, 用转染试剂 lipo8000 在每个培养皿中转染 15 μg SARS-CoV-2 StruΔS 与 15 μg SARS-CoV-2 ORF1ab, 6 h 后更换培养基。48 h 后收集上清溶液, 即为 StruΔS + ORF1ab 的 VLP。采用荧光成像的方法验证质粒的表达, 采用 western blot 分析胞内表达及 VLP 中的 N 蛋白, 采用透射电镜表征的方法验证 VLP 的大小与形貌;

(4) 将 StruΔS, S 与 ORF1ab 的质粒共转入 293T 细胞制备 StruΔS + S + ORF1ab 的 VLP

将 293T 铺在 100 mm 的细胞培养皿, 至细胞密度约为 90 % 时, 用转染试剂 lipo8000 在每个培养皿中转染 15 μg SARS-CoV-2 StruΔS, 15 μg SARS-CoV-2 S 与 15 μg SARS-CoV-2 ORF1ab, 6 h 后更换培养基; 48 h 后收集上清溶液, 即为 StruΔS + S + ORF1ab 的 VLP; 采用荧光成像的方法验证质粒的表达, 采用 western blot 分析胞内表达及 VLP 中的 S 蛋白和 N 蛋白, 采用透射电镜表征的方法验证 VLP 的大小与形貌;

(5) 将 StruΔS, ORF1ab 与 S^{突变}的质粒共转入 293T 细胞制备 StruΔ S + S^{突变} + ORF1ab 的 VLP

首先, 通过 PCR 的方法对 S 蛋白质粒 (即 SARS-CoV-2 S) 进行单位点的突变, 本案例以 N331Q、N501Y、D614G 及 P681H 为例; 将每个 S 突变质粒转染 293T 细胞, 裂解细胞, 采用 RT-PCR 的方法验证 S 突变质粒的表达。然后, 制备 S 突变 VLP, 对于每种 S 蛋白突变, 将 293T 细胞铺在 100 mm 的细胞培养皿, 至细胞密度约为 90 % 时, 用转染试剂 lipo8000 在每个培养皿中转染 15 μg SARS-CoV-2 StruΔS, 15 μg SARS-CoV-2 S^{突变}与 15 μg SARS-CoV-2 ORF1ab 的质粒, 6 h 后更换培养基; 48 h 后收集上清溶液, 即为 StruΔS + S^{突变} + ORF1ab 的 VLP; 采用荧光成像的方法验证质粒的表达, 采用 western blot 分析胞内表达及 VLP 中的 S 蛋白和 N 蛋白, 采用透射电镜表征的方法验证 VLP 的大小与形貌。

下面通过实施例, 并结合附图, 对本发明的技术方案作进一步详细的说明, 但本发明不限于下面的实施例。除非另有说明, 本发明中以及下述实施例中所述的 SARS-CoV-2 基因组及其部分的核苷酸编号均以 NCBI 序列号 NC_045512.2 为准, S 蛋白及其突变体的氨基酸位置编号均以 S 蛋白的氨基酸序列为准, 即以 S 蛋白的第 1 位氨基酸为编号 1。

实施例 1: 材料和方法

质粒: 基于酵母中的 TAR 系统构建 SARS-CoV-2 S、ORF1ab-mCherry 和 StruΔS-EGFP 三个质粒, 以制备 SARS-CoV-2 的病毒颗粒。。在密码子优化过的编码 SARS-CoV-2 S 糖蛋白的 cDNA 序列的 5'端添加 CAG 启动子, 在其 3'端添加 bGH poly(A) signal, 将该带有启动子和 poly(A) signal 的序列拆分为两个带有重叠序列的片段 (SEQ ID NO:1, SEQ ID NO:2), 通过酵母中的同源重组将带有启动子和 poly(A) signal 的 SARS-CoV-2 S 糖蛋白的 cDNA 序列组装并克隆到 pEASY-T1 载体上, 形成质粒 pEASY-T-S (在以下实施例中也被称为 SARS-CoV-2 S 质粒或 S 质粒)。在 SARS-CoV-2 的 ORF1ab 的 5'端添加 CMV 启动子、5' UTR, 并在 3'端添加 3' UTR、HDV、SV40 polyA 和 mCherry 编码基因, 并将由此获得的序列拆分为六个带有重叠序列的片段 (SEQ ID NO: 3-8), 在酵母中将这重叠片段组装并克隆到 pJS356 载体 (pJS356 载体是在 pR415 质粒的基础上加入用于诱导低拷贝质粒的高表达的 ori2-oriV 元件和用于在大肠杆菌中高效表达的 SopA、SopB、RepE 及 SopC 功能元件获得的) 中, 形成质粒 pJS356-ORF1ab-mCherry (在以下实施例中又可称 SARS-CoV-2 ORF1ab 质粒、SARS-CoV-2 ORF1ab-mCherry 质粒、ORF1ab-mCherry 质粒或 ORF1ab 质粒)。在编码全部结构蛋白 (刺突蛋白除外) 和全部辅助蛋白 (辅助蛋白包括 ORF3a, ORF3b, ORF6, ORF7a, ORF7b, ORF8, ORF10) 的 DNA 序列的 3'端添加 CMV 启动子、T7 启动子、5' UTR, 并在 3'端添加 3' UTR、HDV、SV40 polyA 和 EGFP 编码基因, 并将由此获得的序列拆分为三个带有重叠序列的片段 (SEQ ID NO: 9-11), 在酵母中将这重叠片段组装, 克隆到 pJS356 载体中, 形成质粒 pJS356-StruΔS-EGFP (在以下实施例中又可称 SARS-CoV-2 StruΔS 质粒、SARS-CoV-2 StruΔS-EGFP 质粒、StruΔS-EGFP 质粒或 StruΔS 质粒)。通过酶切和测序验证获得的质粒。

将通过生信分析得到的 SARS-CoV-2 基因组 RNA (nt 19900-20000 或 19773-20335) 的包装信号 (packaging signal, 简称 PS) 的 DNA 序列插入 pEGFP-N1 载体中 EGFP 基因的 3'非编码区, 构建质粒 pEGFP-N1-PS101 或 pEGFP-N1-PS583。从质粒 pJS356-ORF1ab-mCherry 扩增得到 PS101 (nt 19900-20000)或 PS583 (nt 19773-20335)。用

Not I 限制性核酸内切酶处理质粒 pEGFP-N1，插入含 PS 的片段以产生 pEGFP-N1-PS101 或 pEGFP-N1-PS583。

定点突变：利用 SARS-CoV-2 S 质粒作为模板构建 S 蛋白基因突变体，包括 SARS-CoV-2 S (N331Q)，SARS-CoV-2 S (N501Y)，SARS-CoV-2 S (D614G) 和 SARS-CoV-2 S (P681H) 共四种质粒。选择目标突变位点附近 15-20 碱基作为正向引物，反向引物选择互补序列作为反向引物，其列于表 1。定点突变 PCR 后，用限制性内切酶 DpnI 消化模板链。产物直接转化至大肠杆菌 DH5 α 感受态细胞，选择单克隆并测序。

ORF1b RNA 二级结构的预测：使用 Biopython 从参考基因组中选取 SARS-CoV-2，SARS-CoV 和类蝙蝠 SARS-CoV 的 ORF1b 区域中的序列 (nt 19500-20400)。二级结构由具有默认参数的 RNA 结构 web 服务器预测，并由 Vienna RNA web 服务器进行可视化分析。

相似结构的多序列比对和预测：我们预测并比较了 ORF1b 3'末端的保守 RNA 茎环，这是 SARS 相关病毒中预测的包装信号。使用 LocARNA web 服务器对 SARS-CoV-2 (NC-045512.2:19773-20355)，SARS-CoV (NC-004718.3:19712-20294) 和类蝙蝠 SARS-CoV (MG772933.1:19773-20355) 进行比对。

细胞构建和培养：本研究构建了稳定表达人 ACE2 的 293T 细胞 (293T/hACE2)。用 hACE2 质粒转染 293T 细胞产生受体 hACE2，并在 2 μ g/mL 嘌呤霉素下进行筛选。通过蛋白质印记来检测 hACE2 细胞的表达。293T 细胞在添加 10% FBS、1%青霉素和链霉素的 DMEM 培养基中培养，并将其培养在 37 $^{\circ}$ C，5% CO $_2$ 培养箱中。293T/hACE2 细胞在添加 10% FBS、1%青霉素和链霉素以及 2 μ g/mL 嘌呤霉素的 DMEM 培养基中培养，并将其培养在 37 $^{\circ}$ C，5% CO $_2$ 培养箱中。

VLPs 的构建：为了获得 VLPs，使用 lipofectamine 8000 与 SARS-CoV-2 S (10 μ g)，ORF1ab-mCherry (10 μ g) 和 Stru Δ S-EGFP (10 μ g) 质粒混合后转染到 293T 细胞中，并在含有 10% FBS 的 DMEM 培养基中培养 48 小时。为了形成组装有包装信号的 VLP (EGFP-PS101) 或者 VLP (EGFP-PS583)，将 pEGFP-N1-PS101 (1 μ g) 或 pEGFP-N1-PS583 (1 μ g) 与 SARS-CoV-2 S (10 μ g) 和 Stru Δ S (10 μ g) 质粒一起共转染到 293T 细胞 48 小时。

VLP 纯化：转染后 48 小时后收集培养基上清并用 0.45 μ m 滤膜进行过滤。将培养基上清加到 20%蔗糖垫上，并使用 SW32 转子以 30000rpm 超速离心 3 小时。纯化后的 VLP 用于拍摄透射电子显微镜。对于蛋白质免疫印迹分析，将培养基上清加到 20%至 40%密度梯度蔗糖的顶部，并使用 SW41 转子以 30000rpm 超速离心 4 小时。采用 PEG20000 浓缩的方法收集 VLP 用于荧光成像。

RNA 提取和 RT-PCR：通过 FastPure 细胞/组织总 RNA 分离试剂盒 V2 细胞培养基上清的总 RNA，并根据试剂盒说明书在 48 小时转录成 cDNA。使用 cDNA 作为模板，设计表 1 中列出的特异性引物用以扩增 SARS-CoV-2 S，N，NSP1，NSP16 或 ORF3a 的编码区，通过在 1%琼脂糖凝胶电泳验证 PCR 产物。

蛋白质免疫印迹分析：用含有蛋白酶抑制剂和 PMSF 的 RIPA 缓冲液在冰上裂解细胞或 VLP 五分钟，并以 13000g 离心 5 分钟除去大的细胞碎片。蛋白质裂解物在 100 度下煮沸 10 分钟，在 4-20%丙烯酰胺凝胶上分离，并转移到硝酸纤维素膜上。用含 1% Tween-20 和 5%牛奶的 TBST 缓冲封闭硝酸纤维素膜，4 度过夜，然后与 IRDye 680RD 山羊抗小鼠抗体或 IRDye 800CW 山羊抗兔抗体一起温育。通过 ChemiDoc MP 成像系统来获得分析结果。

透射电子显微镜：将 VLP 粘附到碳涂覆的铜网上 10 分钟，并用 2% (w/v) 磷钨酸 (pH 7.1) 染色 1 分钟，用 200kV 透射电子显微镜观察样品。

量子点纳米信标的制备：将碲粉（20mg）和硼氢化钠（11mg）加入含 1mL 超纯水的圆底烧瓶中，将混合物在无氧和冰浴条件下搅拌 5 小时以获得碲氢化钠。为了制备镉前体，将氯化镉，氯化锌和 N-乙酰基-L-半胱氨酸按照 1:1:4 的混合并通过 NaOH 将 pH 调节至 9.0。将 BHQ2 和硫代磷酸酯共修饰的 DNA，镉前体和碲氢化钠混合并转移到四聚氟乙烯内衬的不锈钢高压反应釜中。将混合物加热至 200 度 36 分钟以获得 BHQ2-DNA 功能化的 CdTe:Zn²⁺ QDs，通过用 Amicon Ultra-4 离心过滤装置（50kDa）以 8000 rpm 离心 10 分钟来纯化量子点。对 BHQ2-DNA 功能化的 CdTe:Zn²⁺ QDs 进行 95 度退火处理保持 10 分钟并在室温下保持 30 分钟以形成茎环结构。

VLP 的荧光标记：为了构建带荧光标记 VLP，在转染 SARS-CoV-2 质粒后 6 小时加入 40 μL 退火的量子点纳米信标（10 μM）到培养基中以标记病毒基因组，并且在病毒自组装期过程中将纳米信标标记在 VLP 中。在 2mL 收集的病毒液中加入 0.5μL DiO（1 mM），在 37℃下孵育 1 小时以标记病毒囊膜。

病毒感染和荧光成像：将 293T/hACE2 细胞接种于共聚焦培养皿中，在含 10% FBS、1%青霉素和链霉素和 2μg/mL 嘌呤霉素的 DMEM 中培养。荧光标记的 VLP 与宿主细胞 293T/hACE2 在 4℃孵育保持 30 分钟，并用新鲜培养基替换培养基以除去未结合的颗粒。共聚焦培养皿用封口膜密封，然后转移到 37 度 5% CO₂ 培养箱中。通过 UltraView Vox 共聚焦激光扫描系统在 488 nm 和 561 nm 激发下进行成像。

实施例 2 结果和讨论

结果

SARS-CoV-2 VLP 的构建和表征

为组装具侵染能力的 SARS-CoV-2 VLP，我们将 SARS-CoV-2 基因组的 cDNA 全长分为三部分：SARS-CoV-2 ORF1ab（核苷酸位置：266 - 21555）、SARS-CoV-2 S（核苷酸位置：21563-25384）和 SARS-CoV-2 StruΔS（核苷酸位置：25393-29674），且都包含组成型启动子，具体如图 1A 所示。使用化学合成的病毒基因组 DNA 片段，分别在酵母中进行组装构建成分别含有 SARS-CoV-2 ORF1ab、SARS-CoV-2 S 和 SARS-CoV-2 StruΔS 的质粒；构建好的质粒共转染到 293T 细胞中以产生 SARS-CoV-2 VLP（图 1B）。在 293T 细胞中，DNA 片段被翻译并表达病毒蛋白；包含包装信号（PS）序列的 SARS-CoV-2 ORF1ab RNA 与结构蛋白组装形成 SARS-CoV-2 VLP；SARS-CoV-2 VLP 的感染性是由刺突蛋白与宿主细胞表面 hACE2 受体的结合所介导。在病毒粒子中仅组装包包装序列基因，不组装所有结构蛋白基因的情况下，已受 SARS-CoV-2 VLP 感染的细胞不会产生子代病毒（图 1C）。因此，我们的 SARS-CoV-2 VLP 构建策略有望为病毒学研究提供一个安全且通用的平台。

SARS-CoV-2 刺突糖蛋白（S）作为受体结合位点，可以介导膜融合和病毒进入。我们组装并优化了 S 蛋白编码序列，并在哺乳动物细胞中表达 S 蛋白。将 SARS-CoV-2 S 质粒转染 293T 细胞，通过 RT-PCR 分析 S 基因的结果如图 6A 所示。表达的 SARS-CoV-2 S 蛋白通过蛋白免疫印迹验证，结果显示有两个主要条带，180 kDa 的全长蛋白和 110 kDa 的裂解蛋白（图 6B）。

SARS-CoV-2 StruΔS 基因序列与 EGFP 融合构建成质粒，命名为 SARS-CoV-2 StruΔS-EGFP（图 7A）；同时，我们构建了结构蛋白（E，包膜；M，膜；N，核衣壳）和辅助蛋白的表达载体。为了验证 SARS-CoV-2 StruΔS-EGFP 质粒的活性，在 293T 细胞中通过 RT-PCR 分析了 ORF3a 和 N 基因（图 7B）；在质粒转染 24 小时后观察到 EGFP 表达（图 7C），并通过蛋白免疫印迹确定 SARS-CoV-2 N 蛋白的表达（图 7D）。结构蛋白可以组装成病毒样颗粒 VLP（StruΔS）；我们在转染后 48 小时收集培

培养基, 转移至超离管中(底部含 20%蔗糖垫)以分离 VLP。通过蛋白免疫印迹法验证 VLP (StruΔS)中确实含有 SARS-CoV-2 N 蛋白(图 7D); VLP (StruΔS)的形态大小则通过 TEM 进行观察, 我们发现不含刺突蛋白的球形颗粒的估计直径为 $75 \pm 14 \text{ nm}$ ($n = 30$, 图 7E)。

我们将 SARS-CoV-2 ORF1ab 基因序列与 mCherry 序列融合构建质粒, 质粒表达的多聚蛋白 1a (pp1a) 和 pp1ab 可以被病毒蛋白酶切割成 16 个非结构蛋白 (NSP) (图 8A)。通过 RT-PCR 评估 SARS-CoV-2 ORF1ab-mCherry 质粒的活性, 同时使用共聚焦显微镜可在转染后 48 小时内观察到 mCherry 信号(图 8B 和 8C)。

VLP (StruΔS-S)的组装和分泌依赖于 SARS-CoV-2 S 质粒和 StruΔS-EGFP 质粒的共表达; 转染后 24 h, 共聚焦显微镜检测到 EGFP 的荧光, 蛋白免疫印迹检测到 SARS-CoV-2 S 和 N 蛋白, 这表明 SARS-CoV-2 S 和 StruΔS-EGFP RNA 可以在转染的细胞中表达(图 9A 和 9B)。为表征其形貌和尺寸, 对 VLP (StruΔS-S)负染色后在 TEM 下进行透射电镜观察(图 9C); 观察结果表明 VLP (StruΔS-S)是带有刺突蛋白的球形颗粒, 平均大小估计为 $74 \pm 16 \text{ nm}$ ($n = 30$, 不包含刺突)。同时我们也对组装 VLP (StruΔS-S)的结构蛋白进行蛋白免疫印迹分析, 也能得到预期分子量的条带(图 9B)。

VLP (StruΔS-ORF1ab)是由 SARS-CoV-2 StruΔS-EGFP 质粒和 ORF1ab-mCherry 质粒共转染 293T 细胞而产生的; 在转染后 24 小时对细胞进行成像, 可以观察到明显的荧光(图 10A); 同时, 通过蛋白免疫印迹检测到 SARS-CoV-2 N 蛋白在转染细胞中的表达(图 10B)。接着, 进行透射电镜观察以分析 VLP (StruΔS-ORF1ab)的形态特征。不含刺突蛋白的球形颗粒的直径约为 $79 \pm 13 \text{ nm}$ ($n = 30$, 图 10C)。VLP (StruΔS-ORF1ab)的 SARS-CoV-2 N 蛋白通过蛋白免疫印迹法测定(图 10B)。

要制备 SARS-CoV-2 VLP, 需要三质粒共转染到 293T 细胞中。转染后 24 小时内, 能观察到 EGFP 和 mCherry 信号, 表明在细胞中质粒进行了复制和转录(图 2A)。为了检测 SARS-CoV-2 S 和 N 蛋白的表达, 在转染后 48 小时, 收集细胞裂解液进行蛋白免疫印迹分析。同时, 为了分析病毒颗粒的形态表征, 进行了透射电镜观察; 透射电镜图像显示圆形病毒颗粒且具有典型尖峰冠状, 且颗粒的直径约为 $76 \pm 12 \text{ nm}$ ($n = 30$, 不包含刺突)。SARS-CoV-2 VLP 和其他 VLPs 似乎具有相似的形状和尺寸(图 2C)。病毒粒子中的 SARS-CoV-2 S 和 N 蛋白, 通过蛋白免疫印迹进行分析可得到预期的条带(图 2B)。

为了避免潜在的生物安全问题, 将拯救的病毒样颗粒在 BSL-3 实验室进行了安全测试。我们构建的 293T/hACE2 细胞株对 SARS-CoV-2 感染具有易感性; 细胞表达的 hACE2 受体, 可以通过蛋白免疫印迹进行验证(图 11A)。用 SARS-CoV-2 VLP 去感染 293T/hACE2 细胞, 以测试受感染的细胞是否产生了子代病毒。病毒结构蛋白是否在细胞中进行了表达并组装成子代病毒, 用蛋白免疫印迹进行检测, 结果显示没能观察到对应的条带。结果表明, SARS-CoV-2 VLP 不能进行复制、表达结构蛋白, 进而不能在宿主细胞中组装产生子代病毒。由于 SARS-CoV-2 VLP 无法复制, 感染 SARS-CoV-2 VLP 和野生型 SARS-CoV-2 病毒的 293T/hACE2 细胞的代表性图像显示病毒存在显著的滴度差异(图 11C)。因此, 我们构建的 SARS-CoV-2 VLP, 可以作为一种安全、适宜的模型来研究 SARS-CoV-2 病毒的生命周期和感染机制, 有望进一步理解病毒的生命周期, 为抗病毒药物及疫苗的开发提供一定理论基础。

为了进一步探究病毒基因组 RNA 的组装, 我们利用 SARS-CoV-2 VLP 感染 293T/hACE2 细胞(图 2E)。在病毒颗粒感染 48 小时后, 通过共聚焦显微镜成像可以观察到 mCherry 信号, 但可能由于 ORF1ab-mCherry mRNA 的低拷贝数, 受感染细胞中的红色荧光非常微弱(图 2D)。VLP (mCherry)由 SARS-CoV-2 S 质粒、StruΔS-EGFP

质粒和 mCherry 质粒（直接将 mCherry 插入 pEGFP-N1 载体获得）三质粒共转染组装而成。在 VLP (mCherry) 感染的细胞中未观察到荧光信号，表明没有包装信号 (PS) 序列的 mCherry mRNA 无法包装成颗粒。这说明位于 ORF1ab 片段中的 SARS-CoV-2 包装信号，对于病毒 RNA 组装形成颗粒具有重要作用。

SARS-CoV-2 中包装信号的预测和验证

根据之前的报道，冠状病毒的核衣壳通过识别特定的包装信号 (PS) 序列与病毒 RNA 相互作用，以促进基因组 RNA 组装成病毒颗粒。用生物信息学分析进行预测 SARS-CoV-2 的 PS 位置，预测结果表明 PS 在靠近 ORF1b 区域的 3' 末端。同时，我们预测了 SARS-CoV-2、SARS-CoV 和类蝙蝠 SARS -CoV 的 RNA 二级结构；预测结果表明，三种类型的病毒的 RNA 二级结构中都含有两个稳定的茎环 (SL1 和 SL2) (图 3A)。PS 共有结构的多重比对进一步证实了 RNA 茎环具有高度保守性，尤其是 SL1 和 SL2 (图 12)。这些结果表明，稳定的茎环结构可能在 SARS-CoV-2 中具备 PS 功能。为了进一步探究，根据包装序列的生信预测结果，我们选择了两个区域来测试包装活性，一个是包含保守茎环结构的短区域 (PS101, nt 19900-20000)，另外一个是有额外侧翼序列的较长区域 (PS583, nt 19773-20335)。

将预测的 PS 序列与 EGFP 报告基因的 3' 非编码区融合表达，构建了质粒 pEGFP-N1-PS101 和质粒 pEGFP-N1-PS583 并分别转染到 293T 细胞中；EGFP-PS101 和 EGFP-PS583 RNA 的表达通过 RT-PCR 来确定 (图 S8A 和 3B)。为了证实预测的 PS RNA 的组装，收集 VLP (EGFP-PS101) 和 VLP (EGFP-PS583) 并用于感染 293T/hACE2 细胞，如图 3E 所示。SARS-CoV-2 S 质粒、StruΔS 质粒和 pEGFP-N1-PS101/pEGFP-N1-PS583 质粒共转染到 293T 细胞中，分别产生 VLP (EGFP-PS101) 或 VLP (EGFP-PS583)。转染细胞中，S 蛋白、N 蛋白和 EGFP 进行表达并组装成 VLP，通过蛋白免疫印迹对几个蛋白进行分析 (图 13B 和 3C)。在 VLP 中检测到结构蛋白，而 EGFP 未能被检测到。为了验证 EGFP-PS101 或 EGFP-PS583 RNA 是否被包装到 VLP 中，用 VLP 感染 293T/hACE2 细胞并检查受感染细胞中的 EGFP 表达 (图 3E)。在 VLP (EGFP-PS101) 或 VLP (EGFP-PS583) 感染细胞 48 小时后，观察到 EGFP 的绿色荧光，并且具有较高的荧光强度，这说明其具有更好的包装活性。如图 13C 和 3D 所示，VLP 感染细胞中的低荧光信号可能是由于包装到 VLP 中的 EGFP-PS RNA 拷贝数较低。VLP (EGFP) 是通过 SARS-CoV-2 S 质粒、StruΔS 质粒和 pEGFP-N1 质粒共转染得到，并作为对照组，表面不含 PS 序列的 EGFP mRNA 不能被包装到 VLP 中 (绿色荧光蛋白)。这些结果证实预测的序列在 SARS-CoV-2 中具有 PS 功能，对于将病毒基因组 RNA 组装成 VLP 非常重要。

结构蛋白在病毒感染和组装中的作用评估

此前的报道证实，SARS-CoV-2 S 蛋白关联受体结合和膜融合，S 蛋白的突变将极大地影响对宿主的感染能力。我们在 SARS-CoV-2 S 质粒的基础上分别构建了 4 个不同区域的突变，包括 RBD 区域 22553 位点 A 到 C 置换和 RBD 区域 22555 位点 C 到 G 置换的 N331Q 突变、23063 位点 A-T 置换的 N501Y；RBD 区附近 23403 位点的 A 到 G 置换的 D614G；以及弗林蛋白酶区附近 23604 位点 C 到 A 置换的 P681H 突变；S 突变体的构建和表征如图 14-17 所示。VLP S (突变) 由 SARS-CoV-2 S (突变) 质粒、StruΔS 质粒和 ORF1ab 质粒的共转染细胞后组装完成。病毒粒子中的 SARS-CoV-2 S 和 N 蛋白通过蛋白质印迹进行鉴定，预期条带如图 18 所示。VLP S (突变) 的形态和大小利用透射电镜进行分析，结果表明突变体与具有野生型 S 蛋白的 VLP 相似 (图 18)。为了检

测 S 突变病毒颗粒的感染性, 用 DiO 标记脂质包膜产生荧光 VLP S (突变) 去感染 293T/hACE2 细胞。将细胞在感染后 1 小时或 2 小时进行固定, 接着随机选择 500 个细胞在共聚焦显微镜下成像, 统计有色颗粒用于分析。与野生型 S 的 VLP 相比, VLP (N501Y)、VLP (D614G) 和 VLP (P681H) 具有更高的传染性, 而 VLP (N331Q) 的传染性则明显降低 (图 4A)。结果与之前的报告一致, 表明我们构建的系统可以用于评估 S 突变对 SARS-CoV-2 组装和感染性影响的能力。

为了研究核衣壳(N)在病毒 RNA 包装中的必要性, 我们在质粒 SARS-CoV-2 StruΔS 上删除了 N 蛋白的编码区, 从而构建了质粒 SARS-CoV-2 StruΔS-EGFP/ΔN (又可称 SARS-CoV-2 StruΔS-EGFP/ΔN 质粒、StruΔS-EGFP/ΔN 质粒或 StruΔS/ΔN 质粒) 图 19A)。SARS-CoV-2 StruΔS-EGFP/ΔN 质粒的活性和表达分别通过 RT-PCR 和蛋白免疫印迹进行检测 (图 19B 和 19C)。VLP (ΔN) 可以通过 SARS-CoV-2 S 质粒、StruΔS-EGFP/ΔN 质粒和 ORF1ab 三质粒共转染细胞后得到。VLP (ΔN) 呈球形、具有冠状结构, 直径为 81 ± 17 nm ($n = 30$, 不包括刺突), 如 TEM 图像所示。产生的 VLP, 我们用蛋白免疫印迹检测进行检测, 检测到 S 蛋白, 未能检测到 N 蛋白, 这与预期的结果一致(图 4C)。病毒 RNA 组装过程中是否依赖核衣壳, 是通过将 VLP (ΔN-EGFP-PS583) 感染 293T/hACE2 细胞确定的; VLP (ΔN-EGFP-PS583) 是通过与质粒 SARS-CoV-2 S 质粒、StruΔS/ΔN 质粒 以及包装信号质粒 pEGFP-N1-PS583 共转染产生。在 VLP (ΔN-EGFP-PS583) 感染的细胞中观察到 EGFP 的绿色荧光很少, 如图 4B 所示; 这与 VLP (EGFP-PS583) 感染的细胞存在显著差异。结果表明, N 蛋白对于 VLP 的形成不是必需的, 但它在病毒 RNA 包装中发挥重要作用。

关于冠状病毒的早期报道中, 膜 (M) 和包膜 (E) 蛋白作为主要的功能成分, 在病毒颗粒的产生中发挥了重要作用; 本研究也探究了 M 和 E 蛋白对 VLP 制备的影响。分别删除 M 或 E 蛋白编码序列, 得到 SARS-CoV-2 StruΔS-EGFP/ΔM 质粒和 SARS-CoV-2 StruΔS-EGFP/ΔE 质粒, 并通过 RT-PCR 和蛋白免疫印迹进行验证 (图 20 和 22)。将 StruΔS 突变体质粒, 即 SARS-CoV-2 StruΔS-EGFP/ΔM 质粒或 SARS-CoV-2 StruΔS-EGFP/ΔE 质粒与质粒 SARS-CoV-2 S 和 ORF1ab-mCherry 共转染到 293T 细胞中, 分别产生 StruΔS 突变体病毒颗粒 rVP (ΔM) 或 rVP (ΔE)。TEM 图像和蛋白质印迹分析表明, 缺乏 M 或 E 蛋白的病毒颗粒在转染细胞中仍然可获得 (图 4C、21A 和 23A)。StruΔS 突变体病毒颗粒进一步用于感染 293T/hACE2 细胞, 在感染 48 小时后没能观察到荧光信号 (图 21B 和 23B), 这与 SARS-CoV-2 VLP 感染后宿主细胞的表现相反, 表明即使缺失 M 或 E 基因, 转染细胞也能产生突变病毒粒子, 但病毒粒子的功能受到了影响。与 N 缺失的突变体不同, M 或 E 蛋白的缺失不影响病毒 RNA 的包装。

SARS-CoV-2 基因组中开放阅读框 10 (ORF10) 区域位于 N 基因下游, 在之前的报道中 ORF10 蛋白似乎没有发挥重大作用。我们构建了敲除 ORF10 编码序列的 SARS-CoV-2 StruΔS-EGFP/ΔORF10 质粒, 并进行表征 (图 24); 该质粒与质粒 SARS-CoV-2 S、ORF1ab 共转染细胞, 从而产生 VLP (ΔORF10)。VLP (ΔORF10) 表现出与我们构建的 SARS-CoV-2 VLP 具有相似的形态和功能 (图 4C), 表明 ORF10 基因的缺失对病毒颗粒的组装没有影响。

为了进一步确认上述结构蛋白/辅助蛋白对 VLP 组装的影响, 我们对 VLP 中的基因组和囊膜分别进行了荧光标记, 并通过荧光共定位效率初步探究了 VLP 的组装效率 (图 25), 结论与上述验证结果基本一致。

rSARS-CoV-2 病毒进入 293T/hACE2 细胞的实时成像。

我们制备了双荧光标记的 SARS-CoV-2 VLP (QD-DiO)，以深入了解病毒进入过程，如图 5A 所示。病毒 RNA 与靶核酸序列杂交形成复合物，用量子点 (QD) 纳米信标标记后带有红色荧光，复合物最终被病毒颗粒包裹。在 QD 标记病毒内部观察到暗电子致密核心 (图 5B)。为了获得双荧光颗粒，用 DiO 标记脂质包膜。RNA-QD 和 Env-DiO 的荧光共定位证实了成功构建双色 rSARS-CoV-2 (QD-DiO) (图 5C)，其中 rSARS-CoV-2 也即 SARS-CoV-2 VLP。

为了显示病毒进入的动态过程，通过共聚焦显微镜成像在 293T/hACE2 细胞中跟踪 SARS-CoV-2 VLP (QD-DiO) 颗粒。在实时成像期间，只有包含 QD 和 DiO 信号共定位，才当作单个病毒粒子。在 293T/hACE2 细胞的膜上观察到双色颗粒，并以主动方式转运到宿主细胞中 (图 5D)。该病毒颗粒的轨迹如图 5E 所示。本研究中，我们对活细胞中的 2000 多个单病毒颗粒的轨迹进行了跟踪和分析。结果表明，47.5% 病毒颗粒通过内吞作用进入 293T/hACE2 细胞中，而其他颗粒仅附着在细胞膜表面，没有转运到细胞质中。

病毒核心在胞吞进入宿主细胞过程中从包膜中的释放也被实时成像捕获。对活细胞中 rSARS-CoV-2 (QD-DiO) 颗粒的跟踪显示，病毒核心释放到细胞质中。具有 QD 和 DiO 共定位信号 (黄色) 的病毒粒子在细胞质中成像，并且在病毒粒子的动态运动期间观察到红点与绿点的分离 (图 5F)。RNA-QD 和 Env-DiO 的轨迹在分离行为后有所不同，如图 5G 所示。结果表明病毒核心从内体中成功逃逸，这是产生感染的必要过程。本研究中，我们在 293T/hACE2 细胞中分析了 1000 个单个粒子，并捕获了 64 个类似事件。较低的比例可能是由于在观察过程中被跟踪的粒子移出焦平面或荧光信号的光漂白。

讨论

为了研究 SARS-CoV-2，我们构建了病毒基因组拆分系统 (split-virus-genom, SVG)，并进行了对应的表征分析。利用该系统产生了具有单轮感染的 SARS-CoV-2 VLP，有望用于生物学探索和疫苗开发。包含包装信号的 SARS-CoV-2 ORF1ab RNA 被组装到 rSARS-CoV-2 中，而结构基因不在病毒粒子中，从而确保了系统的安全性。我们设计了一系列带有病毒基因的突变或缺失的 rSARS-CoV-2 突变体，以确定病毒成分的功能；探究了四种结构蛋白和 ORF10 蛋白在 SARS-CoV-2 VLP 感染和组装中的作用。同时，我们引入纳米信标和亲脂性染料，通过实时成像显示 SARS-CoV-2 VLP 通过内吞作用进入宿主细胞的动态过程。根据研究需要，可以在系统中设计各种荧光探针，从而可视化病毒生命周期的其他步骤。另外，在 BSL-2 实验室进行病毒学实验的通用平台对于研究 SARS-CoV-2 非常重要。基于我们构建的系统，可运用于目前几种类型的 SARS-CoV-2 研究：(1) SARS-CoV-2 的功能分析可以通过设计 rSARS-CoV-2 突变体来进行每个成分的功能分析。(2) 特定序列 SARS-CoV-2 包装信号的预测和验证，为研究参与病毒 RNA 组装的分子机制提供了可靠的途径。(3) 基于多色标记 SARS-CoV-2 VLP 的单一病毒颗粒追踪，能够对活宿主细胞中病毒颗粒的侵染过程进行实时和精确地成像。(4) 组装的具有不同成分的 VLP 促进了载体和疫苗的开发。结合这些优势，SVG 系统可发展为研究 SARS-CoV-2 和其他冠状病毒的宝贵平台。

我们的实验结果证实 SVG 系统具备安全性。SARS-CoV-2 的全基因组分为三个部分，病毒 RNA 组装所必需的包装信号位于 ORF1ab 片段中。我们的系统缺乏多个基因片段，不会生产野生型病毒。与包含部分功能结构的假病毒不同，我们生产的病毒体几乎具有 SARS-CoV-2 的所有成分。然而，由于病毒包装的 RNA 中缺乏结构基因，SARS-CoV-2 VLP 无法复制和组装子代病毒。这些单轮感染性病毒粒子无法对宿主细胞

进行多轮感染，阻断了病毒的传播。不需要构建特定的稳定细胞系来生产可以在 BSL-2 实验室中安全操作的病毒颗粒，例如反式互补系统。根据病毒研究的需要，可以进一步从三个片段中删除一些 ORF 基因，产生安全性更高的 rSARS-CoV-2 突变体。

在 SVG 系统中,我们惊讶地发现编码 M 蛋白的基因的缺失并不影响病毒颗粒的形成,而在 VLP (Δ M) 感染的细胞中却未观察到病毒 RNA 表达的 mCherry 信号。这可能是由于突变病毒粒子中缺乏 M 蛋白干扰了病毒粒子的功能。因此, SARS-CoV-2 M 蛋白在病毒功能中起主要作用,但对于病毒颗粒的形成并不是必需的。它不同于 SARS-CoV,在 SARS-CoV 中 M 蛋白是病毒组装的关键因素。我们开发的系统也有一些限制。病毒组装效率有待提高,这对系统的进一步提升和疫苗开发具有重要意义。病毒体包装的 RNA 片段的优化可以促进 SARS-CoV-2 感染和复制的研究。本发明的 SVG 系统、组合物和方法可用作促进中和试验和抗病毒测试的强大工具,利用该病毒基因组拆分系统可获得多种假病毒颗粒,可用于病毒抗体的中和抗体试验、抗病毒药物的筛选、病毒疫苗的制备等。总之,我们为 SARS-CoV-2 的研究开发了一个 SVG 系统,可以在 BSL-2 实验室进行操作。产生的病毒粒子 rSARS-CoV-2 很容易进行工程改造,用于分析病毒生命过程和每个组件的功能。

表 1

引物	
Primer S (P681H)-F(CACCGCCGCGCCCCGCAGC)	Tsingke Biotechnology
Primer S (P681H)-R(CGCGGGTGGCTGTGGTCTGGGTCTGG)	Tsingke Biotechnology
Primer S (D614G)-F(CCACCTACGGCGT GGGCTACCA GCC CTACC)	Tsingke Biotechnology
Primer S (D614G)-R(ACGCCGTAGGTGGGCTGGAAGCCGTAGC)	Tsingke Biotechnology
Primer S (N501Y)-F(GCTGTACCAGGGCGTGAA CTGCAC CGA GGT)	Tsingke Biotechnology
Primer S (N501Y)-R(TGCAGTTCACGCCCTGGTA CAGCA CGG CCAC)	Tsingke Biotechnology
Primer S (N331Q)-F(GCGCTTCCCCCAGAT CACCAA CCTGT GCCC)	Tsingke Biotechnology
Primer S (N331Q)-R(GGTTGGTGATCTGGGGGAAGCGCACGATGC)	Tsingke Biotechnology
Primer NSP1-F (ATACCTTCCCAGGTAACAAAC)	Tsingke Biotechnology
Primer NSP1-R (GCATGAAGCTTTACCAGCAC)	Tsingke Biotechnology
Primer NSP16-F (GCTTTGGTGTAAGATGGCC)	Tsingke Biotechnology
Primer NSP16-R (ATTGTTCGCGTGGTTTGCC)	Tsingke Biotechnology
Primer S-F (ATGTTCGTGTTCCCTGGTGCT)	Tsingke Biotechnology
Primer S-R (TAGGTGTAGTGCAGCTTCACG)	Tsingke Biotechnology
Primer ORF3a-F (CCTTCCCAGGTAACAAACCAAC)	Tsingke Biotechnology
Primer ORF3a-R (CACAGTCTTTTACTCCAGATTCCC)	Tsingke Biotechnology
Primer N-F (TGACCAAATTGGCTACTACC)	Tsingke Biotechnology
Primer N-R (GTCTGATTAGTTCCTGGTCC)	Tsingke Biotechnology
Primer GAPDH-F (GTCTCCTCTGACTTCAACAGCG)	Tsingke Biotechnology
Primer GAPDH-R (ACCACCCTGTTGCTGTAGCCAA)	Tsingke Biotechnology
Primer ORF1ab-F (G*G*G*G*G*G*G*G*G*G* AAAAAACAAG GGAC TTCGTCGGCAAT TTTGTTA CCATCAGT ATGA AGT CCCTTGT)	Tsingke Biotechnology

下面的序列中，帶有下列劃線的序列為同源臂

S 蛋白表达序列拆分片段（包含密码子优化的 S 蛋白基因组 3822bp）：

S-片段 1 (SEO ID NO: 1) :

GCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACT
TTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTACACACAGGAAACA
GCTATGACCATGATTACGCCAAGCTTGGTACCGAGCTCGGATCCACTAGTAACGGCCGCCAGT
GTGCTGGAATTCTGCAGATATAACTTCGTATAATGTATGCTATACGAAGTTATccggttacata
5 acttacggtaaataggcccgctggctgaccgcccacgacccccgccattgacgtcaatagt
aacgccaatagggaactttccattgacgtcaatgggtggagtatttacggtaaactgcccaactt
ggcagtacatcaagtgtatcatatgccaaagtacgccccctattgacgtcaatgacggtaaagt
gcccgcctggcattgtgccagtacatgaccttatgggactttcctacttggcagtacatcta
cgtattagtcacgctattaccatggctcgaggtgagccccacgttctgcttcactctccccat
10 cccccccccctccccaccccccaatttttgatattttattttttaattattttgtgcagcgat
gggggcgggggggggggggggggcgggggcgaggggcggggcgaggcgagaggtgcggc
ggcagccaatcagagcggcgcgctccgaaagtttccttttatggcgaggcggcggcggcggcg
gccctataaaaagcgaagcgcgcgggcgggcgagtcgctgcgcgctgccttcgccccgtgcc
ccgctccgcccgcgctcgcgccgccccgcccggctctgactgaccgcgttactcccacaggt
15 gagcgggcgggacggcccttctcctccgggctgtaattagctgagcaagaggtaagggttta
gggatggttggttggtggggtattaatgtttaattacctggagcacctgcctgaaatcacttt
ttttcaggttgaccgggtgccaccATGTTCTGTGTTCTGGTGCTGCTGCCCCCTGGTGAGCAGC
CAGTGCGTGAACCTGACCACCCGCACCCAGCTGCCCCCGCCTACACCAACAGCTTCACCCGC
GGCGTGTACTACCCCGACAAGGTGTTCCGCAGCAGCGTGCTGCACAGCACCCAGGACCTGTTT
20 CTGCCCCCTTCTCAGCAACGTGACCTGGTTCCACGCCATCCACGTGAGCGGCACCAACGGCACC
AAGCGCTTCGACAACCCCGTGCTGCCCTTCAACGACGGCGTGACTTCGCCAGCACCGAGAAG
AGCAACATCATCCGCGGCTGGATCTTCGGCACCCCTGGACAGCAAGACCCAGAGCCTGCTG
ATCGTGAACAACGCCACCAACGTGGTGATCAAGGTGTGCGAGTTCCAGTTCTGCAACGACCCC
TTCCTGGGCGTGACTACCACAAGAACAACAAGAGCTGGATGGAGAGCGAGTTCCGCGTGTA
25 AGCAGCGCCAACAACCTGCACCTTCGAGTACGTGAGCCAGCCCTTCCTGATGGACCTGGAGGGC
AAGCAGGGCAACTTCAAGAACCTGCGCGAGTTCTGTGTTCAAGAACATCGACGGCTACTTCAAG
ATCTACAGCAAGCACACCCCCATCAACCTGGTGCGGACCTGCCCCAGGGCTTCAGCGCCCTG
GAGCCCCCTGGTGACCTGCCCATCGGCATCAACATCACCCGCTTCCAGACCCTGCTGGCCCTG
CACCGCAGCTACCTGACCCCCGGCGACAGCAGCAGCGGTGGACCGCCGGCGCCGCCGCTAC
30 TACGTGGGCTACCTGCAGCCCCGCACCTTCCTGCTGAAGTACAACGAGAACGGCACCATCACC
GACGCCGTGGACTGCGCCCTGGACCCCCTGAGCGAGACCAAGTGCACCCTGAAGAGCTTCACC
GTGGAGAAGGGCATCTACCAGACCAGCAACTTCCGCGTGACGCCACCGAGAGCATCGTGCGC
TTCCCCAACATACCAACCTGTGCCCTTCGGCGAGGTGTTCAACGCCACCCGCTTCGCCAGC
GTGTACGCTGGAACCGCAAGCGCATCAGCAACTGCGTGGCCGACTACAGCGTGCTGTACAAC
35 AGCGCCAGCTTCAGCACCTTCAAGTGCTACGGCGTGAGCCCCACCAAGCTGAACGACCTGTGC
TTCACCAACGTGTACGCCGACAGCTTCGTGATCCGCGGCGACGAGGTGCGCCAGATCGCCCC
GGCCAGACCGGCAAGATCGCCGACTACAACCTACAAGCTGCCCCGACGACTTCACCGGTGCGTG
ATCGCCTGGAACAGCAACAACCTGGACAGCAAGGTGGGCGGCAACTACAACCTGTACCGC
CTGTTCCGCAAGAGCAACCTGAAGCCCTTCGAGCGCGACATCAGCACCGAGATCTACCAGGCC
40 GGCAGCACCCCCCTGCAACGGCGTGAGGGCTTCAACTGCTACTTCCCCCTGCAGAGCTACGGC
TTCCAGCCCCACCAACGGCGTGGGCTACCAGCCCTACCGCGTGGTGGTGCTGAGCTTCGAGCTG
CTGCACGCCCCCGCCACCGTGTGCGGCCCAAGAAGAGCACCAACCTGGTGAAGAACAAGTGC
GTGAACCTCAACTTCAACGGCCTGACCGGCACCGCGTGCTGACCGAGAGCAACAAGAAGTTC
CTGCCCTTCCAGCAGTTTCGGCCGCGACATCGCCGACACCACCGACGCCGTGCGCGACCCCCAG
45 A

S-片段 2 (SEQ ID NO: 2) :

CTTCGAGCTGCTGCACGCCCCCGCCACCGTGTGCGGCCCAAGAAGAGCACCAACCTGGTGAA
GAACAAGTGCGTGAACCTCAACTTCAACGGCCTGACCGGCACCGCGTGCTGACCGAGAGCAA
50 CAAGAAGTTCCTGCCCTTCCAGCAGTTTCGGCCGCGACATCGCCGACACCACCGACGCCGTGCG
CGACCCCCAGACCTGGAGATCCTGGACATCACCCCTGCAGCTTCGGCGGCGTGAGCGTGAT
CACCCCCGGCACCAACACCAGCAACCAGGTGGCCGTGCTGTACCAGGACGTGAACCTGCACCGA

GGTGCCCGTGGCCATCCACGCCGACCAGCTGACCCCCACCTGGCGCGTGTACAGCACCGGCAG
CAACGTGTTCCAGACCCGCGCCGGCTGCCTGATCGGCGCCGAGCACGTGAACAACAGCTACGA
GTGCGACATCCCCATCGGCGCCGGCATCTGCGCCAGCTACCAGACCCAGACCAACAGCCCCCG
CCGCGCCCGCAGCGTGGCCAGCCAGAGCATCATCGCCTACACCATGAGCCTGGGCGCCGAGAA
5 CAGCGTGGCCTACAGCAACAACAGCATCGCCATCCCCACCAACTTCACCATCAGCGTGACCAC
CGAGATCCTGCCCCGTGAGCATGACCAAGACCAGCGTGGACTGCACCATGTACATCTGCGGCGA
CAGCACCGAGTGCAGCAACCTGCTGCTGCAGTACGGCAGCTTCTGCACCCAGCTGAACCGCGC
CCTGACCGGCATCGCCGTGGAGCAGGACAAGAACACCCAGGAGGTGTTTCGCCAGGTGAAGCA
GATCTACAAGACCCCCCCCCATCAAGGACTTCGGCGGCTTCAACTTCAGCCAGATCCTGCCCCGA
10 CCCCAGCAAGCCAGCAAGCGCAGCTTCATCGAGGACCTGCTGTTCAACAAGGTGACCCTGGC
CGACGCCGGCTTCATCAAGCAGTACGGCGACTGCCTGGGCGACATCGCCGCCCGCGACCTGAT
CTGCGCCCAGAAGTTCAACGGCCTGACCGTGCTGCCCCCCTGCTGACCGACGAGATGATCGC
CCAGTACACCAGCGCCCTGCTGGCCGGCACCATCACCAGCGGCTGGACCTTCGGCGCCGGCGC
CGCCCTGCAGATCCCCTTCGCCATGCAGATGGCCTACCGCTTCAACGGCATCGGCGTGACCCA
15 GAACGTGCTGTACGAGAACCAGAAGCTGATCGCCAACCAGTTCAACAGCGCCATCGGCAAGAT
CCAGGACAGCCTGAGCAGCACCGCCAGCGCCCTGGGCAAGCTGCAGGACGTGGTGAACCAGAA
CGCCAGGCCCTGAACACCCTGGTGAAGCAGCTGAGCAGCAACTTCGGCGCCATCAGCAGCGT
GCTGAACGACATCCTGAGCCGCCTGGACAAGGTGGAGGCCGAGGTGCAGATCGACCGCCTGAT
CACCGGCCGCCTGCAGAGCCTGCAGACCTACGTGACCCAGCAGCTGATCCGCGCCCGCCGAGAT
20 CCGCGCCAGCGCCAACCTGGCCGCCACCAAGATGAGCGAGTGCGTGCTGGGCCAGAGCAAGCG
CGTGGACTTCTGCGGCAAGGGCTACCACCTGATGAGCTTCCCCCAGAGCGCCCCCACGGCGT
GGTGTTCCTGCACGTGACCTACGTGCCCCGCCAGGAGAAGAAGTTCAACCACCGCCCCCGCCAT
CTGCCACGACGGCAAGGCCACTTCCCCCGCGAGGGCGTGTTCTGTGAGCAACGGCACCCACTG
GTTCTGTGACCCAGCGCAACTTCTACGAGCCCCAGATCATCACCACCGACAACACCTTCGTGAG
25 CGGCAACTGCGACGTGGTGATCGGCATCGTGAACAACACCGTGTACGACCCCCCTGCAGCCCGA
GCTGGACAGCTTCAAGGAGGAGCTGGACAAGTACTTCAAGAACCACACCAGCCCCGACGTGGA
CCTGGGCGACATCAGCGGCATCAACGCCAGCGTGGTGAACATCCAGAAGGAGATCGACCGCCT
GAACGAGGTGGCCAAGAACCTGAACGAGAGCCTGATCGACCTGCAGGAGCTGGGCAAGTACGA
GCAGTACATCAAGTGGCCCTGGTACATCTGGCTGGGCTTCATCGCCGGCCTGATCGCCATCGT
30 GATGGTGACCATCATGCTGTGCTGCATGACCAGCTGCTGCAGCTGCCTGAAGGGCTGCTGCAG
CTGCGGCAGCTGCTGCAAGTTCGACGAGGACGACAGCGAGCCCGTGCTGAAGGGCGTGAAGCT
GCACTACACCTAAtgtacaaggaatttctaactagagctcgtgatcagcctcgactgtgcctt
ctagttgccagccatctgttgtttgccccctccccgtgccttccttgaccctggaaggtgcc
ctcccactgtccttttctaataaaaatgaggaaattgcatcgcatgtgtctgagtaggtgtcatt
35 ctattctggggggtgggggtggggcaggacagcaagggggaggattgggaagagaatagcaggc
atgctggggaATAACTTCGTATAATGTATGCTATACGAAGTTATATCCATCACACTGGCGGCC
GcCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGG
CCGTCGTTTTTACAACGTCTGACTGGGAAAACCTGGCGTTACCCAACCTTAATCGCCTTGCA
CACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAAC
40 AGTTGCGCAGCCT

ORF1ab 序列拆分片段:

ORF1ab-片段 1 (SEQ ID NO: 3) :

CTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAAGTGAACGC
45 CGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAT
AAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTCTGCCACTCATCGCAGTACTGTTGTAATT
CATTAAgcatgTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATG
GAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGC
CCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGT
50 CAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCA
AGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATG

ACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTG
ATGCGGTTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTTGACTCACGGGGATTTCCAAGT
CTCCACCCCATTTGACGTCAATGGGAGTTTGT TTTGGCACCAAAATCAACGGGACTTTCCAAAA
TGTCGTAACAACCTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTAT
5 ATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCGattaaagggtttataccttcccaggttaaca
aaccaaccaacttttcgatctctttagatctgttctctaaacgaactttaaaatctgtgtggc
tgtcactcggctgcatgcttagtgactcagcaggtataattaataactaattactgtcgttg
acaggacacgagtaactcgtctatcttctgcaggctgcttacggtttcgtccgtgttgagcc
gatcatcagcacatctaggtttcgtccgggtgtgaccgaaaggtaagatggagagccttgtcc
10 ctggtttcaacgagaaaaacacacgtccaactcagtttgctgttttacagggttcgagcgtgc
tcgtacgtggctttggagactccgtggaggaggtcttatcagaggcacgtcaacatcttaaag
atggcacttgtggcttagtagaagttgaaaaaggcggttttgctcaacttgaacagccctatg
tgttcatcaaaggttcggatgctcgaactgcacctcatgggtcatgttatgggtgagctggtag
cagaactcgaaggcattcagtagcgtgtagtggtgagacacttgggtgtccttgtccctcatg
15 tgggcgaaataaccagtggttaccgcaagggttcttcttcgtaagaacggtaataaaggagctg
gtggccatagttacggcgccgatctaaagtcatttgacttaggcgacgagcttggcactgatc
cttatgaagattttcaagaaaactggaacactaaacatagcagtggtgttaccctgtgaactca
tgcgtgagcttaacggaggggcatacactcgtatgtcgataacaacttctgtggccctgatg
gctaccctcttgagtgcattaaagaccttctagcacgtgctggtaaagcttcatgcactttgt
20 ccgaacaactggactttattgacactaagaggggtgtatactgctgccgtgaacatgagcatg
aaattgcttgggtacacggaacgttctgaaaagagctatgaattgcagacaccttttgaaatta
aattggcaaagaaatttgacaccttcaatggggaatgtccaaattttgtatttcccttaaatt
ccataatcaagactattcaaccaagggttgaaaagaaaaagccttgatggctttatgggtagaa
ttcgatctgtctatccagttgcgtcaccaaatgaatgcaaccaaattgtgcctttcaactctca
25 tgaagtgtgatcattgtgtggtgaaacttcatggcagacgggcgattttgttaaagccacttgcg
aattttgtggcactgagaatttgactaaagaagggtgccactacttgtggttacttaccctaaa
atgctgttgttaaaattttattgtccagcatgtcacaattcagaagtaggacctgagcatagtc
ttgccgaataaccataatgaatctggcttgaaaaccatttctcgtagggtgggtgcactattg
cctttggaggctgtgtgttctcttatgttgggtgccataacaagtgtgcctattgggttccac
30 gtgctagcgtacacataggttgtaaccatacaggtgttgggtggagaagggtccgaagggtctta
atgacaaccttcttgaaatactccaaaaagagaaagtcaacatcaatattgttgggtgacttta
aacttaatgaagagatcgccattattttggcatcttttctgcttccacaagtgttttgtgg
aaactgtgaaagggttggattataaagcattcaaacaaattgttgaatcctgtggtaatttta
aagttacaaaaggaaaagctaaaaaagggtgcctggaatattggtgaacagaaatcaatactga
35 gtcctctttatgcatttgcacagaggtgctcgtgttgtagcatcaattttctccgcactc
ttgaaactgctcaaaattctgtgctgttttacagaaggccgctataacaatactagatggaa
tttcacagtattcactgagactcattgatgctatgatgttcacatctgatttggctactaaca
atctagttgtaatggcctacattacaggtgggtgttgggtcagttgacttcgcagtggttaacta
acatctttggcactgtttatgaaaaactcaaaccgctccttgattggcttgaagagaagtta
40 aggaagggtgtagagtttcttagagacgggttgggaaattgttaaatttatctcaacctgtgctt
gtgaaattgtcgggtggacaaattgtcacctgtgcaaaggaaattaaggagagtggtcagacat
tctttaagccttgtaataaaatttttggcttgggtgtgctgactctatcattattgggtggagcta
aacttaaagccttgaatttaggtgaaacatttgtcacgcactcaaagggtgttacagaaagt
gtgttaaatccagagaagaaactggcctactcatgcctctaaaagcccaaaaagaaattatct
45 tcttagagggagaaacacttcccacagaagtgttaacagaggaagtgttcttgaaaactgggtg
atttacaaccattagaacaacctactagtgaagctgttgaaagctccattgggttggtacaccag
tttgtattaaacgggcttatgttgcgtcgaaatcaaagacacagaaaagtactgtgcccttgcac
ctaataatgatggtaacaaacaataccttcacactcaaaggcggtgcaccaacaaagggttactt
ttgggtgatgacactgtgatagaagtgaagggttacaagagtgtgaatatcacttttgaacttg
50 atgaaaggattgataaagtaacttaataagagaagtgtcgtgcctatacagttgaactcgggtacag
aagtaaatgagttcgcctgtgttgtggcagatgctgtcataaaaaactttgcaaccagtatctg
aattacttacaccactgggcattgatttagatgagtggagtatggctacatactacttatttg

atgagtctggtgagtttaaatggcttcacatatgtattgttctttctaccctccagatgagg
atgaagaagaaggtgattgtgaagaagaagagtttgagccatcaactcaatatgagtatggta
ctgaagatgattaccaaggtaaacctttggaatttggtgccacttctgctgctcttcaacctg
aagaagagcaagaagaagattgggttagatgatgatagtcacaaactggttggtcaacaagacg
5 gcagtgaggacaatcagacaactactattcaacaattggtgaggttcaacctcaattagaga
tggaacttacaccagttgttcagactattgaagtgaatagtttttagtggttattttaaactta
ctgacaatgtatacattaaaaatgcagacattgttggaagaagctaaaaaggtaaaaccaacag
tggttggttaatgcagccaatgtttaccttaaacatggaggaggtgttgaggagccttaata
aggctactaacaatgccatgcaagttgaatctgatgattacatagctactaatggaccactta
10 aagtgggtggttagttgtgttttaagcggacacaatcttgctaaacactgtcttcatgttgtcg
gccc aaatg ttaacaaagg tgaagacatt caacttct taagagt gcttat gaaaatt ttaatc
agcacgaagttctacttgcaccattattatcagctgggtat ttttggtgctgaccctatacatt
ctttaagagtt

15 ORF1ab-片段2 (SEQ ID NO: 4) :

AAGTGGGTGGTAGTTGTGTTTTAAGCGGACACAATCTTGCTAAACACTGTCTTCATGTTGTCTG
GCCCAAATGTTAACAAAGGTGAAGACATTCAACTTCTTAAGAGTGCTTATGAAAATTTTAATC
AGCACGAAGTTCTACTTGCACCATTATTATCAGCTGGTATTTTTGGTGCTGACCCTATACATT
CTTTAAGAGTTTGTGTAGATACTGTTTCGCACAAATGTCTACTTAGCTGTCTTTGATAAAAATC
20 TCTATGACAACTTGTTCAGCTTTTTTGAAATGAAGAGTGAAAAGCAAGTTGAACAAAAGA
TCGCTGAGATTCCTAAAGAGGAAGTTAAGCCATTTATAACTGAAAGTAAACCTTCAGTTGAAC
AGAGAAAACAAGATGATAAGAAAATCAAAGCTTGTGTTGAAGAAGTTACAACAACCTCTGGAAG
AAACTAAGTTCCTCACAGAAAACCTTGTTACTTTATATTGACATTAATGGCAATCTTCATCCAG
ATTCTGCCACTCTTGTTAGTGACATTGACATCACTTTCTTAAAGAAAGATGCTCCATATATAG
25 TGGGTGATGTTGTTCAAGAGGGTGTTTTTAACTGCTGTGGTTATACCTACTAAAAGGCTGGTG
GCACTACTGAAATGCTAGCGAAAGCTTTGAGAAAAGTGCCAACAGACAATTATATAACCACTT
ACCCGGGTGAGGGTTTTAAATGGTTACACTGTAGAGGAGGCAAAGACAGTGCTTAAAAGTGTA
AAAGTGCCTTTTACATTCTACCATCTATTATCTCTAATGAGAAGCAAGAAATCTTGGAAGTG
TTTCTTGGAATTTGCGAGAAATGCTTGCACATGCAGAAGAAACACGCAAATTAATGCCTGTCT
30 GTGTGGAACATAAGCCATAGTTTCAACTATACAGCGTAAATATAAGGGTATTTAAATACAAG
AGGGTGTGGTTGATTATGGTGCTAGATTTTACTTTTACACCAGTAAAACAACCTGTAGCGTCAC
TTATCAACACACTTAACGATCTAAATGAACTCTTGTTACAATGCCACTTGGCTATGTAACAC
ATGGCTTAAATTTGGAAGAAGCTGCTCGGTATATGAGATCTCTCAAAGTGCCAGCTACAGTTT
CTGTTTCTTCACCTGATGCTGTTACAGCGTATAATGGTTATCTTACTTCTTCTTAAAACAC
35 CTGAAGAACATTTTATTGAAACCATCTCACTTGCTGGTTCCTATAAAGATTGGTCCCTATTCTG
GACAATCTACACAACCTAGGTATAGAATTTCTTAAGAGAGGTGATAAAGTGATATTACACTA
GTAATCCTACCACATTCCACCTAGATGGTGAAGTTATCACCTTTGACAATCTTAAGACACTTC
TTTCTTTGAGAGAAGTGAGGACTATTAAGGTGTTTACAACAGTAGACAACATTAACCTCCACA
CGCAAGTTGTGGACATGTCAATGACATATGGACAACAGTTTGGTCCAACCTTATTTGGATGGAG
40 CTGATGTTACTAAAATAAAACCTCATAATTCACATGAAGGTAAAACATTTTATGTTTTACCTA
ATGATGACACTCTACGTGTTGAGGCTTTTGAGTACTACCACACAACCTGATCCTAGTTTTCTGG
GTAGGTACATGTCAGCATTAATCACACTAAAAAGTGGAATACCCACAAGTTAATGGTTTTAA
CTTCTATTAAATGGGCAGATAACAACCTGTTATCTTGCCACTGCATTGTTAACTCTCAACAAA
TAGAGTTGAAGTTTAATCCACCTGCTCTACAAGATGCTTATTACAGAGCAAGGGCTGGTGAAG
45 CTGCTAACTTTTGTGCACTTATCTTAGCCTACTGTAATAAGACAGTAGGTGAGTTAGGTGATG
TTAGAGAAACAATGAGTTACTTGTTCACATGCCAATTTAGATTCTTGCAAAAAGAGTCTTGA
ACGTGGTGTGTAACCTTGTGGACAACAGCAGACAACCTTAAAGGGTGTAGAAGCTGTTATGT
ACATGGGCACACTTTCTTATGAACAATTTAAGAAAGGTGTTTACAGATACCTTGTACGTGTGGTA
AACAAGCTACAAAATATCTAGTACAACAGGAGTCACCTTTTGTATGATGTCAGCACCACCTG
50 CTCAGTATGAACTTAAGCATGGTACATTTACTTGTGCTAGTGAGTACACTGGTAATTACCAGT
GTGGTCACATATAACATATAACTTCTAAAGAACTTTGTATTGCATAGACGGTGCTTTACTTA

CAAAGTCCTCAGAATACAAAGGTCCTATTACGGATGTTTTCTACAAAGAAAACAGTTACACAA
CAACCATAAAACCAGTTACTTATAAATTGGATGGTGTGTTTGTACAGAAATTGACCCTAAGT
TGGACAATTATTATAAGAAAGACAATTCTTATTTACAGAGCAACCAATTGATCTTGTACCAA
ACCAACCATATCCAAACGCAAGCTTCGATAATTTTAAGTTTGTATGTGATAATATCAAATTTG
5 CTGATGATTTAAACCAGTTAACTGGTTATAAGAAACCTGCTTCAAGAGAGCTTAAAGTTACAT
TTTTCCCTGACTTAAATGGTGATGTGGTGGCTATTGATTATAAACTACACACCCTCTTTTA
AGAAAGGAGCTAAATTTGTTACATAAACCTATTGTTTGGCATGTTAACAATGCAACTAATAAAG
CCACGTATAAAACCAATACCTGGTGTATACGTTGTCTTTGGAGCACAAAACCAGTTGAAACAT
CAAATTCGTTTGATGTACTGAAGTCAGAGGACGCGCAGGGAATGGATAATCTTGCCTGCGAAG
10 ATCTAAAACCAGTCTCTGAAGAAGTAGTGGAATCCTACCATACAGAAAGACGTTCTTGAGT
GTAATGTGAAAACCTACCGAAGTTGTAGGAGACATTATACTTAAACCAGCAAATAATAGTTTAA
AAATTACAGAAGAGGTTGGCCACACAGATCTAATGGCTGCTTATGTAGACAATTCTAGTCTTA
CTATTAAGAAACCTAATGAATTATCTAGAGTATTAGGTTTGAAAACCTTGCTACTCATGGTT
TAGCTGCTGTAAATAGTGTCCCTTGGGATACTATAGCTAATTATGCTAAGCCTTTTCTTAACA
15 AAGTTGTTAGTACAACCTACTAACATAGTTACACGGTGTTTAAACCGTGTGTGTACTAATTATA
TGCCTTATTTCTTTACTTTATTGCTACAATTGTGTACTTTTACTAGAAGTACAAATTCTAGAA
TTAAAGCATCTATGCCGACTACTATAGCAAAGAATACTGTTAAGAGTGTGCGTAAATTTTGTCT
TAGAGGCTTCATTTAATTATTTGAAGTCACCTAATTTTCTAAACTGATAAATATTATAATTT
GGTTTTTACTATTAAAGTGTGTGCTAGGTTCTTTAATCTACTCAACCGCTGCTTTAGGTGTTT
20 TAATGTCTAATTTAGGCATGCCTTCTTACTGTACTGGTTACAGAGAAGGCTATTTGAACTCTA
CTAATGTCACCTATTGCAACCTACTGTACTGGTTCTATACCTTGTAGTGTGTGTCTTAGTGGTT
TAGATTCTTTAGACACCTATCCTTCTTTAGAACTATACAAATTACCATTTTCATCTTTTAAAT
GGGATTTAACTGCTTTTGGCTTAGTTGCAGAGTGGTTTTTGGCATATATTCTTTTCACTAGGT
TTTTCTATGTACTTTGGATTGGCTGCAATCATGCAATTGTTTTTTCAGCTATTTTGCAGTACATT
25 TTATTAGTAATTTCTTGGCTTATGTGGTTAATAATTAATCTTGTACAAATGGCCCCGATTTTTCAG
CTATGGTTAGAATGTACATCTTCTTTGCATCATTTTATTATGTATGGAAAAGTTATGTGCATG
TTGTAGACGGTTGTAATTCATCAACTTGTATGATGTGTGTACAAACGTAATAGAGCAACAAGAG
TCGAATGTACAACCTATTGTTAATGGTGTGTAGAAAGTCCTTTTATGTCTATGCTAATGGAGGTA
AAGGCTTTTGCAACTACACAATTGGAATTGTGTAAATTGTGATACATTCTGTGCTGGTAGTA
30 CATTTATTAGTGTATGAAGTTGCGAGAGACTTGTCACTACAGTTTAAAAGACCAATAAATCCTA
CTGACCAGTCTTCTTACATCGTTGATAGTGTGTACAGTGAAGAATGGTTCCATCCATCTTTACT
TTGATAAAGCTGGTCAAAAGACTTATGAAAGACATTCTCTCTCATTTTGTAACTTAGACA
ACCTGAGAGCTAATAACACTAAAGGTTTATTGCTTATTAATGTTATAGTTTTTGTATGGTAAAT
CAAAATGTGAAGAATCATCTGCAAAATCAGCGTCTGTTTACTACAGTCAGCTTATGTGTCAAC
35 CTATACTGTTACTAGATCAGGCATTAGTGTCTGATGTTGGTGATAGTGCGGAAGTTGCAGTTA
AAATGTTTGATGCTTACGTTAATACGT

ORF1ab-片段3 (SEQ ID NO: 5) :

CACTAAAGGTTTATTGCCTATTAATGTTATAGTTTTTGTATGGTAAATCAAAATGTGAAGAATC
40 ATCTGCAAAATCAGCGTCTGTTTACTACAGTCAGCTTATGTGTCAACCTATACTGTTACTAGA
TCAGGCATTAGTGTCTGATGTTGGTGATAGTGCGGAAGTTGCAGTTAAAATGTTTGTATGCTTA
CGTTAATACGTTTTTCATCAACTTTTAACGTACCAATGGAAAACTCAAAACACTAGTTGCAAC
TGCAGAAGCTGAACTTGCAAGAATGTGTCCTTAGACAATGTCTTATCTACTTTTATTTTTCAGC
AGCTCGGCAAGGGTTTGTGATTGATGATGTAAGAACTAAAGATGTTGTTGAATGTCTTAAATT
45 GTCACATCAATCTGACATAGAAGTTACTGGCGATAGTTGTAATAACTATATGCTCACCTATAA
CAAAGTTGAAAACATGACACCCCGTGACCTTGGTGCTTGTATTGACTGTAGTGCGCGTCATAT
TAATGCGCAGGTAGCAAAAAGTCACAACATTGCTTTGATATGGAACGTTAAAGATTTTCATGTC
ATTGTCTGAACAACCTACGAAAACAAATACGTAGTGTCTGCTAAAAAGAATAACTTACCTTTTAA
GTTGACATGTGCAACTACTAGACAAGTTGTTAATGTTGTAACAACAAAGATAGCACTTAAGGG
50 TGGTAAAATTGTTAATAATTGGTTGAAGCAGTTAATTAAAGTTACACTTGTGTTCCCTTTTTGT
TGCTGCTATTTTCTATTTAATAACACCTGTTTCATGTCTGTCTAAACATACTGACTTTTCAAG

TGAAATCATAGGATACAAGGCTATTGATGGTGGTGTCACTCGTGACATAGCATCTACAGATAC
TTGTTTTGCTAACAAACATGCTGATTTTGACACATGGTTTAGCCAGCGTGGTGGTAGTTATAC
TAATGACAAAGCTTGCCCATTTGATTGCTGCAGTCATAACAAGAGAAGTGGGTTTTGTCGTGCC
TGGTTTTGCCTGGCAGCATATTACGCACAACCTAATGGTGACTTTTTGCATTTCTTACCTAGAGT
5 TTTTAGTGCAAGTTGGTAACATCTGTTACACACCATCAAACTTATAGAGTACACTGACTTTGC
AACATCAGCTTGTGTTTTGGCTGCTGAATGTACAATTTTTAAAGATGCTTCTGGTAAGCCAGT
ACCATATTGTTATGATACCAATGTACTAGAAGGTTCTGTTGCTTATGAAAGTTTACGCCCTGA
CACACGTTATGTGCTCATGGATGGCTCTATTATTCAATTTCCCTAACACCTACCTTGAAGGTTT
TGTTAGAGTGGTAACAACCTTTTGATTCTGAGTACTGTAGGCACGGCACTTGTGAAAGATCAGA
10 AGCTGGTGGTTGTGTATCTACTAGTGGTAGATGGGTACTTAACAATGATTATTACAGATCTTT
ACCAGGAGTTTTCTGTGGTGTAGATGCTGTAAATTTACTTACTAATATGTTTACACCCTAAT
TCAACCTATTGGTGCTTTGGACATATCAGCATCTATAGTAGCTGGTGGTATTGTAGCTATCGT
AGTAACATGCCTTGCCCTACTATTTTATGAGGTTTAGAAGAGCTTTTGGTGAATACAGTCATGT
AGTTGCCTTTAATACTTTACTATTCCTTATGTCATTCACTGTACTCTGTTTAAACACCAGTTTA
15 CTCATTCTTACCTGGTGTATTATTCTGTTATTTACTTGTACTTGACATTTTATCTTACTAATGA
TGTTTTCTTTTTTAGCACATATTCAGTGGATGGTTATGTTACACCTTTAGTACCTTTCTGGAT
AACAATTGCTTATATCATTTGTATTTCCACAAAGCATTTCTATTGGTCTTTTAGTAATTACCT
AAAGAGACGTGTAGTCTTTAATGGTGTTCCTTTAGTACTTTTGAAGAAGCTGCGCTGTGCAC
CTTTTTGTAAATAAAGAAATGTATCTAAAGTTGCGTAGTGATGTGCTATTACCTCTTACGCA
20 ATATAATAGATACTTAGCTCTTTATAATAAGTACAAGTATTTTAGTGAGCAATGGATACAAC
TAGCTACAGAGAAGCTGCTTGTGTCATCTCGCAAAGGCTCTCAATGACTTCAGTAACTCAGG
TTCTGATGTTCTTTACCAACCACCACAAACCTCTATCACCTCAGCTGTTTTGCAGAGTGGTTT
TAGAAAAATGGCATTCCCATCTGGTAAAGTTGAGGGTTGTATGGTACAAGTAACTTGTGGTAC
AACTACACTTAACGGTCTTTGGCTTGATGACGTAGTTTACTGTCCAAGACATGTGATCTGCAC
25 CTCTGAAGACATGCTTAACCCTAATTATGAAGATTTACTCATTCGTAAGTCTAATCATAATTT
CTTGGTACAGGCTGGTAATGTTCAACTCAGGGTTATTGGACATTCTATGCAAAATTGTGTACT
TAAGCTTAAGGTTGATACAGCCAATCCTAAGACACCTAAGTATAAGTTTGTTCGCATTCAACC
AGGACAGACTTTTTTCAGTGTTAGCTTGTTACAATGGTTCACCATCTGGTGTTTACCAATGTGC
TATGAGGCCCAATTTCACTATTAAGGGTTCATTCCTTAATGGTTCATGTGGTAGTGTGGTTT
30 TAACATAGATTATGACTGTGTCTCTTTTTGTTACATGCACCATATGGAATTACCAACTGGAGT
TCATGCTGGCACAGACTTAGAAGGTAACCTTTTATGGACCTTTTGTGACAGGCAAACAGCACA
AGCAGCTGGTACGGACACAACCTATTACAGTTAATGTTTTAGCTTGGTTGTACGCTGCTGTTAT
AAATGGAGACAGGTGGTTTCTCAATCGATTTACCACAACCTCTAATGACTTTAACCTTGTGGC
TATGAAGTACAATTATGAACCTCTAACACAAGACCATGTTGACATACTAGGACCTCTTTCTGC
35 TCAAACCTGGAATTGCCGTTTTAGATATGTGTGCTTCATTAAAAGAATTACTGCAAAATGGTAT
GAATGGACGTACCATATTGGGTAGTGCTTTATTAGAAGATGAATTTACACCTTTTGATGTTGT
TAGACAATGCTCAGGTGTTACTTTCCAAAGTGCAGTGAAAAGAACAATCAAGGGTACACACCA
CTGGTTGTTACTCACAATTTTGACTTCACTTTTAGTTTTAGTCCAGAGTACTCAATGGTCTTT
GTTCTTTTTTTTTGTATGAAAATGCCTTTTTACCTTTTGCTATGGGTATTATTGCTATGTCTGC
40 TTTTGCAATGATGTTTGTCAAACATAAGCATGCATTTCTCTGTTTGTGTTTGTACCTTCTCT
TGCCACTGTAGCTTATTTTAATATGGTCTATATGCCTGCTAGTTGGGTGATGCGTATTATGAC
ATGGTTGGATATGGTTGATACTAGTTTGTCTGGTTTTAAGCTAAAAGACTGTGTTATGTATGC
ATCAGCTGTAGTGTTACTAATCCTTATGACAGCAAGAACTGTGTATGATGATGGTGCTAGGAG
AGTGTGGACACTTATGAATGTCTTGACACTCGTTTATAAAGTTTATTATGGTAATGCTTTAGA
45 TCAAGCCATTTCCATGTGGGCTCTTATAATCTCTGTTACTTCTAACTACTCAGGTGTAGTTAC
AACTGTCAATGTTTTTGGCCAGAGGTATTGTTTTTATGTGTGTTGAGTATTGCCCTATTTTCTT
CATAACTGGTAATACACTTCAGTGATAATGCTAGTTTATTGTTTCTTAGGCTATTTTTGTAC
TTGTTACTTTGGCCTCTTTGTTTACTCAACCGCTACTTTAGACTGACTCTTGGTGTGTTATGA
TTACTTAGTTTCTACACAGGAGTTTAGATATATGAATTCACAGGGACTACTCCACCCCAAGAA
50 TAGCATAGATGCCTTCAAACCTCAACATTAAATTGTTGGGTGTTGGTGGCAAACCTTGTATCAA
AGTAGCCACTGTACAGTCTAAAATGTGAGATGTAAAGTGCACATCAGTAGTCTTACTCTCAGT
TTTGCAACAACCTCAGAGTAGAATCATCATCTAAATTGTGGGCTCAATGTGTCCAGTTACACAA

TGACATTCTCTTAGCTAAAGATACTACTGAAGCCTTTGAAAAAATGGTTTCACTACTTTCTGT
TTTGCTTTCCATGCAGGGTGCTGTAGACATAAACAAGCTTTGTGAAGAAATGCTGGACAACAG
GGCAACCTTACAAGCTATAGCCTCAGAGTTTAGTTCCCTTCCATCATATGCAGCTTTTGCTAC
TGCTCAAGAAGCTTATGAGCAGGCTGTTGCTAATGGTGATTCTGAAGTTGTTCTTAAAAAGTT
5 GAAGAAGTCTTTGAATGTGGCTAAATCTGAATTTGACCGTGATGCAGCCATGCAACGTAAGTT
GGAAAAGATGGCTGATCAAGCTATGACCCAAATGTATAAACAGGCTAGATCTGAGGACAAGAG
GGCAAAAGTTACTAGTGCTATGCAGAC

ORF1ab-片段 4 (SEQ ID NO: 6) :

10 GAGCAGGCTGTTGCTAATGGTGATTCTGAAGTTGTTCTTAAAAAGTTGAAGAAGTCTTTGAAT
GTGGCTAAATCTGAATTTGACCGTGATGCAGCCATGCAACGTAAGTTGGAAAAGATGGCTGAT
CAAGCTATGACCCAAATGTATAAACAGGCTAGATCTGAGGACAAGAGGGCAAAAGTTACTAGT
GCTATGCAGACAATGCTTTTCACTATGCTTAGAAAGTTGGATAATGATGCACTCAACAACATT
15 ATCAACAATGCAAGAGATGGTTGTGTTCCCTTGAACATAATACCTCTTACAACAGCAGCCAAA
CTAATGGTTGTCATACCAGACTATAACACATATAAAAAATACGTGTGATGGTACAACATTTACT
TATGCATCAGCATTGTGGGAAATCCAACAGGTTGTAGATGCAGATAGTAAATTTGTTCAACTT
AGTGAAATTAGTATGGACAATTCACCTAATTTAGCATGGCCTCTTATTGTAACAGCTTTAAGG
GCCAATTCTGCTGTCAAATTACAGAATAATGAGCTTAGTCCTGTTGCACTACGACAGATGTCT
TGTGCTGCCGGTACTACACAACTGCTTGCACTGATGACAATGCGTTAGCTTACTACAACACA
20 ACAAAGGGAGGTAGGTTTGTACTTGCACTGTTATCCGATTTACAGGATTTGAAATGGGCTAGA
TTCCCTAAGAGTGATGGAACCTGGTACTATCTATACAGAAGTGAACCCACCTTGTAGGTTTGT
ACAGACACACCTAAAGGTCCTAAAGTGAAGTATTTATACTTTATTAAAGGATTAACAACCTA
AATAGAGGTATGGTACTTGGTAGTTTAGCTGCCACAGTACGTCTACAAGCTGGTAATGCAACA
GAAGTGCTGCCAATTCAACTGTATTATCTTTCTGTGCTTTTGCTGTAGATGCTGCTAAAGCT
25 TACAAAGATTATCTAGCTAGTGGGGGACAACCAATCACTAATTGTGTTAAGATGTTGTGTACA
CACACTGGTACTGGTCAGGCAATAACAGTTACACCGGAAGCCAATATGGATCAAGAATCCTTT
GGTGGTGCATCGTGTTGTCTGTACTGCCGTTGCCACATAGATCATCCAAATCCTAAAGGATTT
TGTGACTTAAAAGGTAAGTATGTACAAATACCTACAACCTTGTGCTAATGACCCTGTGGGTTTT
ACACTTAAAAACACAGTCTGTACCGTCTGCGGTATGTGGAAAGGTTATGGCTGTAGTTGTGAT
30 CAACTCCGCGAACCCTATGCTTCAGTCAGCTGATGCACAATCGTTTTTAAACGGGTTTGC
GGTGTAAGTGCAGCCCGTCTTACACCGTGCGGCACAGGCACTAGTACTGATGTCGTATACAGGGCTT
TTGACATCTACAATGATAAAGTAGCTGGTTTTGCTAAATTCCTAAAACTAATTGTTGTCGCT
TCCAAGAAAAGGACGAAGATGACAATTTAATTGATTCTTACTTTGTAGTTAAGAGACACACTT
TCTCTAACTACCAACATGAAGAAACAATTTATAATTTACTTAAGGATTGTCCAGCTGTTGCTA
35 AACATGACTTCTTTAAGTTTAGAATAGACGGTGACATGGTACCACATATATCACGTCAACGTC
TTACTAAATACACAATGGCAGACCTCGTCTATGCTTTAAGGCATTTTGATGAAGGTAATTGTG
ACACATTAAGAAGAAATACTTGTACATAACAATTGTTGTGATGATGATTATTTCAATAAAAAGG
ACTGGTATGATTTTGTAGAAAACCCAGATATATTACGCGTATACGCCAAGCTTAGGTGAACGTG
TACGCCAAGCTTTGTTAAAAACAGTACAATTCTGTGATGCCATGCGAAATGCTGGTATTGTTG
40 GTGTACTGACATTAGATAATCAAGATCTCAATGGTAACTGGTATGATTTCCGTGATTTCATAC
AAACCACGCCAGGTAGTGGAGTTCCTGTTGTAGATTCTTATTATTCATTGTTAATGCCTATAT
TAACCTTGACCAGGGCTTTAACTGCAGAGTCACATGTTGACACTGACTTAACAAAGCCTTACA
TTAAGTGGGATTTGTTAAAATATGACTTCACGGAAGAGAGGTTAAAACCTCTTTGACCGTTATT
TTAAATATTGGGATCAGACATAACCCCAAATTGTGTTAACTGTTTGGATGACAGATGCATTC
45 TGCATTGTGCAAACCTTTAATGTTTTATTCTCTACAGTGTTCCACCTACAAGTTTTGGACCAC
TAGTGAGAAAAATATTTGTTGATGGTGTTCATTTGTAGTTTCAACTGGATACCACTTCAGAG
AGCTAGGTGTTGTACATAATCAGGATGTAACTTACATAGCTCTAGACTTAGTTTTAAGGAAT
TACTTGTGTATGCTGCTGACCCTGCTATGCACGCTGCTTCTGGTAATCTATTACTAGATAAAC
GCACTACGTGCTTTTCAGTAGCTGCACTTACTAACAATGTTGCTTTTCAAACGTGCAAAACCCG
50 GTAATTTTAAACAAAGACTTCTATGACTTTGCTGTGTCTAAGGGTTTTCTTTAAGGAAGGAAGTT
CTGTTGAATTAACAACTTCTTCTTTGCTCAGGATGGTAATGCTGCTATCAGCGATTATGACT

ACTATCGTTATAATCTACCAACAATGTGTGATATCAGACAACACTACTATTTGTAGTTGAAGTTG
TTGATAAGTACTTTTGATTGTTACGATGGTGGCTGTATTAATGCTAACCAAGTCATCGTCAACA
ACCTAGACAAATCAGCTGGTTTTCCATTTAATAAATGGGGTAAGGCTAGACTTTATTATGATT
CAATGAGTTATGAGGATCAAGATGCACTTTTCGCATATACAAAACGTAATGTCATCCCTACTA
5 TAACTCAAATGAATCTTAAGTATGCCATTAGTGCAAAGAATAGAGCTCGCACCGTAGCTGGTG
TCTCTATCTGTAGTACTATGACCAATAGACAGTTTCATCAAAAATTATTGAAATCAATAGCCG
CCACTAGAGGAGCTACTGTAGTAATTGGAACAAGCAAATTCTATGGTGGTTGGCACAACATGT
TAAAAACTGTTTATAGTGATGTAGAAAACCCTCACCTTATGGGTGGGATTATCCTAAATGTG
ATAGAGCCATGCCTAACATGCTTAGAATTATGGCCTCACTTGTTCTTGCTCGCAAACATACAA
10 CGTGTTGTAGCTTGTGCACACCGTTTCTATAGATTAGCTAATGAGTGTGCTCAAGTATTGAGTG
AAATGGTCATGTGTGGCGGTTCACTATATGTTAAACCAGGTGGAACCTCATCAGGAGATGCCA
CAACTGCTTATGCTAATAGTGTTTTTAACATTTGTCAAGCTGTCACGGCCAATGTTAATGCAC
TTTTATCTACTGATGGTAACAAAATTGCCGATAAGTATGTCCGCAATTTACAACACAGACTTT
ATGAGTGTCTCTATAGAAAATAGAGATGTTGACACAGACTTTGTGAATGAGTTTTACGCATATT
15 TGCGTAAACATTTCTCAATGATGATACTCTCTGACGATGCTGTTGTGTGTTTCAATAGCACTT
ATGCATCTCAAGGTCTAGTGGCTAGCATAAAGAACTTTAAGTCAGTTCTTTATTATCAAAACA
ATGTTTTTATGTCTGAAGCAAAATGTTGGACTGAGACTGACCTTACTAAAGGACCTCATGAAT
TTTGCTCTCAACATACAATGCTAGTTAAACAGGGTGATGATTATGTGTACCTTCCTTACCCAG
ATCCATCAAGAATCCTAGGGGGCCGGCTGTTTTGTAGATGATATCGTAAAAACAGATGGTACAC
20 TTATGATTGAACGGTTCGTGTCTTTAGCTATAGATGCTTACCCACTTACTAAACATCCTAATC
AGGAGTATGCTGATGTCTTTCATTTGTACTTACAATACATAAGAAAGCTACATGATGAGTTAA
CAGGACACATGTTAGACATGTATTCTGTTATGCTTACTAATGATAACACTTCAAGGTATTGGG
AACCTGAGTTTTATGAGGCTATGTACACACCGCATAACAGTCTTACAGGCTGTTGGGGCTTGTG
TTCTTTGCAATTCACAGACTTCATTAAGATGTGGTGCTTGCATACGTAGACCATTCTTATGTT
25 GTAAATGCTGTTACGACCATGTCATATCAACATCACATAAATTAGTCTTGTCTGTTAATCCGT
ATGTTTGCAATGCTCCAGGTGTGATGTCACAGATGTGACTCAACTTTACTTAGGAGGTATGA
GCTATTATTGTAAATCACATAAACCACCCATTAGTTTTCCATTGTGTGCTAATGGACAAGTTT
TTGGTTTATATAAAAATACATGTGTTGGTAGCGATAATGTTACTGACTTTAATGCAATTGCAA
CATGTGACTGGACAAATGCTGGTGATTACATTTTAGCTAACACCTGTACTGAAAGACTCAAGC
30 TTTTTGCAGCAGAAACGCTCAAAGCTACTGAGGAGACATTTAAACTGTCTTATGGTATTGCTA
CTGTACGTGAAGTGCTGTCTGACAGAGAATTACATCTTTCATGGGAAGTTGGTAAACCTAGAC
CACCCTTAACCGAAATTATGTCTTTACTGGTTATCGTGTAACATAAAAACAGTAAAGTACAAA
TAGGAGAGTACACCTTTGAAAAAGGTGACTATGGTGATGCTGTTGTTTACCGAGGTACAACAA
CTTACAAATTAAATGTTGGTGATTATTTTGTGCTGACATCACATACAGTAATGCCATTAAAGTG
35 CACCTACACTAGTGCCACAAGAGCACTATGTTAGAATTACTGGCTTATACCCAACACTCAATA
TCTCAGATGAGTTTTCTAGCAATGTTGCAAATTATCAAAAGGTTGGTATGCAAAGTATTCTA
CACTC

ORF1ab-片段 5 (SEQ ID NO: 7) :

40 caacaacttacaaattaaatggttggtgattatTTTTgtgctgacatcacatacagtaatgccat
taagtgccactacactagtgccacaagagcactatggttagaattactggcttatacccaacac
tcaatatctcagatgagtttcttagcaatggttgcaaattatcaaaagggttggtatgcaaaagt
attctacactccagggaccacctggtactggttaagagtcattttgctattggcctagctctct
actacccttctgctcgcatagtgatacagcttgctctcatgccgctggtgatgcactatgtg
45 agaaggcattaaaaatatttgcctatagataaatgtagtagaattatacctgcacgtgctcggtg
tagagtgttttgataaattcaaagtgaattcaacattagaacagtatgtcttttgtactgtaa
atgcattgcctgagacgacagcagatatagttgtctttgatgaaatttcaatggccacaaatt
atgatttgagtgttgtaatgccagattacgtgctaagcactatgtgtacattggcgaccctg
ctcaattacctgcaccacgcacattgctaactaagggcacactagaaccagaatatttcaatt
50 cagtggtgtagacttatgaaaactataggtccagacatgttcctcggaacttgctcggcgttgctc
ctgctgaaattggttgacactgtgagtgccttggtttatgataataagcttaaagcacataaag

acaaatcagctcaatgcttttaaagtgttttataaggggtgttatcacgcatgatgtttcatctg
caattaacaggccacaaataggcgtggtaagagaattccttacacgtaaccctgcttgagaa
aagctgtctttatttcaccttataattcacagaatgctgtagcctcaaagattttgggactac
caactcaaactgttgattcatcacaggggtcagaatatgactatgtcatattcactcaaacca
5 ctgaaacagctcactcttgtaatgtaaacagatttaattgttgctattaccagagcaaaagtag
gcatacttttgcataatgtctgatagagacctttatgacaagttgcaatttacaagtcttgaaa
ttccacgtaggaatgtggcaactttacaagctgaaaatgtaacaggactcttttaaagattgta
gtaaggtaatcactgggttacatcctacacaggcacctacacacctcagtgttgacactaaat
tcaaaactgaagggtttatgtgttgacatacctggcatacctaaggacatgacctatagaagac
10 tcatctctatgatgggtttttaaagtgaattatcaagtttaattggttaccctaacaatgtttatca
cccgcgaagaagctataagacatgtacgtgcatggattggcttcgatgtcgaggggtgtcatg
ctactagagaagctgttggtaccaattttacctttacagctaggtttttctacaggtgttaacc
tagttgctgtacctacaggttatgttgatacacctaataatacagattttttccagagtttagt
ctaaaccaccgcctggagatcaattttaaaccacctcataccacttatgtacaaaggacttcctt
15 ggaatgtagtgctgtataaagattgtacaaatgttaagtacacacttaaaaatctctctgaca
gagtcgtattttgtcttatgggcacatggcctttgagttgacatctatgaagtattttgtgaaaa
taggacctgagcgcacctgttgctctatgtgatagacgtgccacatgcttttccactgcttcag
acacttatgcctgttggtcattctatttgatttgattacgtctataatccgttttatgattg
atgttcaacaatgggggttttacaggtaacctacaaagcaacctgatctgtattgtcaagtcc
20 atggtaatgcacatgtagctagttgtgatgcaatcatgactaggtgtctagctgtccacgagt
gctttgttaagcgtgttgactggactattgaatatcctataattgggtgatgaactgaagatta
atgcggctttagaagggttcaacacatgggtgttaagctgcattatttagcagacaaattcc
cagttcttcacgacattggtaaccctaaagctattaagtgtgtacctcaagctgatgtagaat
ggaagtctctatgatgcacagcctttagtgacaaagcttataaaatagaagaattattctatt
25 cttatgccacacattctgacaaattcacagatgggtgtatgcctattttgggaattgcaatgtcg
atagatatcctgctaattccattgtttgtagatttgacactagagtgtctatctaacccttaact
tgcttggtgtgatgggtggcagtttgtatgtaaataaacatgcattccacacaccagcttttg
ataaaagtgttttgttaattttaaacaattaccatttttctattactctgacagttccatgtg
agtctcatggaaaacaagtagtgctcagatatagattatgtaccactaaagtctgctacgtgta
30 taacacgttgcaatttaggtgggtgtctgtctgttagacatcatgctaattgagtacagattgtatc
tcgatgcttataacatgatgatctcagctggccttagcttgtgggtttacaaacaatttgata
cttataacctctggaacacttttacaagacttcagagtttagaaaatgtggccttttaattgttg
taaataagggacactttgatggacaacaggggtgaagtaccagtttctatcattaataacactg
tttacacaaaagttgatgggtgttgatgtagaattgtttgaaaataaaacaacattacctgtta
35 atgtagcatttgagcttttgggctaagcgcaacattaaaccagtaccagaggtgaaaatactca
ataatttgggtgtggacattgctgctaatactgtgatct

ORF1ab-片段6 (SEQ ID NO: 8) :

gggtgaagtaccagtttctatcattaataacactgtttacacaaaagttgatgggtgttgatgt
40 agaattgtttgaaaataaaacaacattacctgttaattgtagcatttgagctttgggctaagcg
caacattaaaccagtaccagaggtgaaaataactcaataatttgggtgtggacattgctgctaa
tactgtgatctgggactacaaaagagatgctccagcacatatatctactattgggtgtttgttc
tatgactgacatagccaagaaaccaactgaaacgatttgtgcaccactcactgtcttttttgga
tggttagagttgatgggtcaagtagacttatttagaaatgcccgtaattgggtgttcttattacaga
45 aggttagtggttaaaggtttacaaccatctgtaggtcccaaacaagctagtcttaattggagtcac
attaattggagaagccgtaaaaacacagttcaattattataagaaagttgatgggtgttgcca
acaattacctgaaacttactttactcagagtagaaatttacaagaatttaaaccaggaggtca
aatggaaattgatttcttagaattagctatggatgaattcattgaacggtataaattagaagg
ctatgccttcgaacatatcgtttatggagatttttagtcatagtcagttaggtgggtttacatct
50 actgattggactagctaaacgttttaaggaatcaccttttgaattagaagattttattcctat
ggacagtacagttaaaaactatttcataacagatgcgcaaacaggttcactaagtgtgtgtg

ttctgttattgatttattacttgatgattttgttgaaataataaaaatcccaagatttatctgt
 agtttctaagggtgtcaaagtgactattgactatacagaaatttcatttatgctttggtgtaa
 agatggccatgtagaaacattttacccaaaattacaatctagtcaagcgtggcaaccgggtgt
 tgctatgcctaattctttacaaaatgcaaagaatgctattagaaaagtgtgaccttcaaaatta
 5 tgggtgatagtgtcaacattacctaagggcataatgatgaatgtcgcaaaatataactcaactgtg
 tcaatattttaaacacattaacattagctgtacctataatatgagagttatacattttggtgc
 tggttctgataaaggagttgcaccaggtacagctgttttaagacagtggttgacctacgggtac
 gctgcttgctgattcagatcttaatgactttgtctctgatgcagattcaactttgattggtga
 ttgtgcaactgtacatacagctaataaatgggatctcattattagtgtatgtacgacctaa
 10 gactaaaaatgttacaaaagaaaatgactctaaagaggggttttttactttacattttgtgggtt
 tatacaacaaaagctagctcttgagggttccgtggctataaagataacagaacattcttgga
 tgctgatctttataagctcatgggacacttcgcatgggtggacagcctttgttactaatgtgaa
 tgcgtcatcatctgaagcatttttaattggatgtaattatcttggcaaacacgcgaacaaat
 agatggttatgtcatgcatgcaaatatacatttttgagggaatacaaatccaattcagttgtc
 15 ttcctattctttattttgacatgagtaaatccccttaaattaaggggtactgctgttatgtc
 tttaaaagaagggtcaaataatgatgatgattttatctcttcttagtaaggttagacttataat
 tagagaaaacaacagagttgttatttctagtgtgttcttgttaacaacGCGGCCGCGAGGG
 CCGCGGCAGCCTGCTGACCTGCGGCGACGTGGAGGAGAACCCCGGCCCGCGGCCCGCCatggt
 gagcaagggcgaggaggacaacatggccatcatcaaggagttcatgcttcaagggtgcacat
 20 ggagggcagcgtgaacggccacgagttcgagatcgagggcgagggcgagggccgcccctacga
 gggcaccagaccgccaagctgaagggtgaccaagggcgggccccctgccccttcgctgggacat
 cctgagccccagttcatgtacggcagcaaggcctacgtgaagcaccgccgacatccccga
 ctacctgaagctgagcttccccgagggcttcaagtgggagcgcgtgatgaacttcgaggacgg
 cggcgtggtgaccgtgaccagagcagcagcctgcaggacggcgagttcatctacaaggtgaa
 25 gctgcgcgccaccaacttccccagcgacggccccgtgatgcagaagaagaccatgggctggga
 ggccagcagcgagcgcgtgtaccccgaggacggcgcccctgaagggcgagatcaagcagcgcct
 gaagctgaaggacggcgccactacgacgcccaggtgaagaccacctacaaggccaagaagcc
 cgtgcagctgcccggcgccctacaacgtgaacatcaagctggacatcaccagccacaacgagga
 ctacaccatcgtggagcagtagcagcgcgcccaggggccacagcaccggcgccatggacga
 30 gctgtacaagGCGGCCGCCtaacaatctttaatcagtggtgtaacattagggaggacttgaaag
 agccaccacatttttaccgagggccacgcggagtagcatcgagtgtacagtgaacaatgctagg
 gagagctgcctatatggaagagccctaattgtgtaaaattaatttttagtagtgctatccccatg
 tgatttttaatagcttcttaggagaatgacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaA
 AAAAAAAAAAagggtcggcatggcatctccacctcctcgcggtccgacctgggcatccgaagg
 35 aggacgcacgtccactcggatggctaaggagGCTAGCagacatgataagatacattgatga
 gtttgacaaaccacaactagaatgcagtgaaaaaatgctttattttgtgaaattttgtgatgc
 tattgctttattttgtaaccattataagctgcaataaacaagttaacaacaacaattgcattca
 ttttatgttttcaggttcagggggaggtgtgggaggttttttaagcaagtaaaacctctacaa
 atgtggtacAGAAAGTCAAAAGCCTCCGACCGGAGGCTTTTGACTATTACTGCCGACATGGAA
 40 GCCATCACAAACGGCATGATGAACCTATAAACGCAGAAAGGCCACCCGAAGGTGAGCCAGTG
 TGATTACCAGCGGCATCAGCACCTTGTCGCCTTGCGTATAATATTTGCAAAAAACCCCTCAAG
 ACCCGTTTAGAGGCCCAA

StruΔS 蛋白表达序列拆分片段:

45 StruΔS-片段 1 (SEQ ID NO: 9):
 CTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGC
 CGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAAGTCCAGGCATCAAA
 AAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTCCTGCCACTCATCGCAGTACTGTTGTAATT
 CATTAAgcatgTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATG
 50 GAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCG
 CCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACGT

CAATGGGTGGAGTATTTACGGTAAACTGCCCCACTTGGCAGTACATCAAGTGTATCATATGCCA
AGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCCAGTACATG
ACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTG
ATGCGGTTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGA CTACG GGGGATTTC AAGT
5 CTCCACCCCATTTGACGTCAATGGGAGTTTGT TTTGGCACCAAAATCAACGGGACTTTCCAAAA
TGTCGTAACAACCTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTAT
ATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCTAATACGACTCACTATAGGGattaaaggtt
tataccttcccaggtatacaaaccaaccaactttcgatctctttagatctgttctctaaacga
actttaaaatctgtgtggctgtcactcggctgcatgcttagtgactcacgcagtataattaa
10 taactaattactgtcgttgacaggacacgagtaactcgtctatcttctgcaggctgcttacgg
tttcgtccgtggttgacgacgatcatcagcacatctaggttttcgtccgggtgtgaccgaaaggt
aagacgaacttatggatttgtttatgagaatcttcacaattggaactgtaactttgaagcaag
gtgaaatcaaggatgctactccttcagattttgttcgcgctactgcaacgataccgatacaag
cctcactcccttttcggatggcttattgtttggcgttgcaacttcttgctgtttttcagagcgctt
15 ccaaaatcataaccctcaaaaagagatggcaactagcactctccaaggggtgttcactttgttt
gcaacttgctgtttgtttgtttgtaacagtttactcacaccttttgctcgttgctgctggccttg
aagccccctttctctatctttatgcttttagtctacttcttgcaagagtataaactttgtaagaa
taataatgaggcttttggtttgctggaaatgccgttccaaaaaccattactttatgatgcca
actattttctttgctggcactaattgttacgactattgtataccttacaatagtgttaactt
20 cttcaattgtcattacttcagggtgatggcacaacaagtcctatttctgaacatgactaccaga
ttgggtggttatactgaaaaatgggaatctggagtaaaagactgtgttgattacacagttact
tcacttcagactattaccagctgtactcaactcaattgagtacagacactgggtgttgaaatg
ttaccttcttcatctacaataaaattgttgatgagcctgaagaacatgtccaaattcacacaa
tcgacgggttcatccggagttgttaatccagtaatggaaccaatttatgatgaaccgacgacga
25 ctactagcgtgcctttgtaagcacagctgatgagtacgaacttatgtactcattcgttttcgg
aagagacaggtacgttaatagtttaatagcgtacttcttttcttgctttcgtgggtattcttg
tagttacactagccatccttactgcgcttcgattgtgtgcgtactgctgcaatattgttaacg
tgagtcttgtaaaaccttctttttacgtttactctcgtgttataaaatctgaattcttctagag
ttcctgatcttctggtctaaacgaactaaatattatattagtttttctggtttggaactttaat
30 ttttagccatg

StruAS-片段 2 (SEQ ID NO: 10) :

ctagttacactagccatccttactgcgcttcgattgtgtgcgtactgctgcaatattgttaac
gtgagtccttgtaaaaccttctttttacgtttactctcgtgttaaaaatctgaattcttctaga
35 gttcctgatcttctggtctaaacgaactaaatattatattagtttttctggtttggaactttaa
tttttagccatggcagattccaacgggtactattaccggttgaaagacttaaaaagctccttgaac
aatggaacctagtaataggtttcttattccttacatggatttgccttctacaatttgcttatg
ccaacaggaataggtttttgtatataattaagtttaatttctctggtggttatggccagtaa
cttttagcttggtttgtgcttgctgctgtttacagaataaattggatcaccgggtggaattgcta
40 tcgcaatggcttgcttctgtaggcttgatgtggctcagctacttcattgcttcttctcagactgt
ttgcgcgtacgcgttccatgtggtcattcaatccagaaactaacattcttctcaacgtgccac
tccatggcactattctgaccagaccgttctagaaagtgaactcgtaatcggagctgtgatcc
ttcgtggacatcttctgatttgcgtggacaccatctaggacgctgtgacatcaaggacctgccta
aagaaatcactgttgctacatcacgaacgctttcttattacaaattgggagcttcgcagcgtg
45 tagcaggtgactcaggttttgcgtgcatacagtcgctacaggattggcaactataaattaacaa
cagaccattccagtagcagtgacaatattgctttgcttgtagtaagtgacaacagatgttt
catctcgttgactttcaggttactatagcagagatattactaattattatgaggacttttaaa
gtttccatttggaatcttgattacatcataaacctcataattaaaaatttatctaagtcacta
actgagaataaataattctcaattagatgaagagcaaccaatggagattgattaaacgaacatg
50 aaaattattcttttcttggcactgataacactcgctacttgtgagctttatcactaccaagag
tgtgttagagggtacaacagtttttaaaagaaccttgcttcttctggaacatacagagggcaat
tcaccatttcatcctctagctgataacaaatttgcaactgacttgcttttagcactcaatttgc

tttgcttgtcctgacggcgtaaaacacgtctatcagttacgtgccagatcagtttcaccta
ctgttcatcagacaagaggaagttcaagaactttactctccaatttttcttattgttgcggca
atagtgtttataacactttgcttcacactcaaaagaaagacagaatgattgaactttcattaa
ttgacttctatttgtgcttttttagcctttctgctattccttggttttaattatgcttattatct
5 tttggttctcacttgaactgcaagatcataatgaaacttgtcacgcctaacgaacatgaaat
ttcttggttttcttaggaatcatcacaactgtagctgcatttcaccaagaatgtagtttacagt
catgtactcaacatcaaccatattgtagttgatgaccctgtgcctattcacttctattctaaat
ggtatatttagagtaggagctagaaaatcagcacctttaattgaattgtgcgtggatgaggctg
gttctaaatcaccattcagtagacatcgatatcggttaattatacagtttctgtttacctttta
10 caattaattgccaggaacctaaattgggtagtcttgtagtgcttggttcgttctatgaagact
tttttagagtatcatgacgttcgtgttgttttagatttcacataacgaacaaactaaaatgtc
tgataatggaccccaaaatcagcgaaatgcaccccgcatcattggttggtgacccctcagattc
aactggcagtaaccagaatggagaacgcagtggggcgcgatcaaaacaacgtcggccccaagg
tttaccacaataatactgcgtccttggttcaccgctctcactcaacatggcaaggaagaccttaa
15 attccctcgaggacaaggcgttccaattaacaccaatagcagtcacagatgaccaaattggcta
ctaccgaagagctaccagacgaattcgtggtggtgacggtaaaatgaaagatctcagtcacag
atggtattttctactacctaggaactgggcccagaagctggacttccctatggtgctaacaaga
cggcatcatatgggttgcaactgagggagccttgaaatacaccaaaaagatcacattggcaccgc
caatcctgctaacaatgctgcaatcgtgctacaacttctcaaggaacaacattgccaaaagg
20 cttctacgcagaagggagcagagggcggcagtcacagcctcttctcgttctcctcatcacgtagtgc
caacagttcaagaaattcaactccaggcagcagtaggggaacttctcctgctagaatggctgg
caatggcgggtgatgctgctcttgctttgctgctgcttgacagattgaaccagcttgagagcaa
aatgtctggttaaaggccaacaacaagaaggccaaactgtcactaagaaatctgctgctgagggc
ttctaagaagcctcggcaaaaacgtactgccactaaagcatacaatgtaacacaagcctttcgg
25 cagacgtggtccagaacaaacccaaggaaatttttggggaccaggaactaatcagacaaggaac
tgattacaacattggccgcaaattgcacaatttgccccagcgcttcagcgcttcttcggaat
gtcgcgcattggcatggaagtcacaccttcgggaacgtggttgacctacacaggtgccatcaa
attggatgacaaagatccaaatt

StruAS-片段 3 (SEQ ID NO: 11) :

agaacaaacccaaggaaatttttggggaccaggaactaatcagacaaggaactgattacaacaa
ttggccgcaaattgcacaatttgccccagcgcttcagcgcttcttcggaatgtcgcgcattgg
catggaagtcacaccttcgggaacgtggttgacctacacaggtgccatcaaattggatgacaa
agatccaaatttcaagatcaagtcattttgctgaataagcatattgacgcatacaaaaacatt
35 cccaccaacagagcctaaaaaggacaaaaagaagaaggctgatgaaactcaagccttaccgca
gagacagaagaacagcaaaactgtgactcttcttctcctgctgcagatttggatgatttctccaa
acaattgcaacaatccatgagcagtgctgactcaactcaggcctaaactcatgcagaccacac
aaggcagatgggctatataaacgttttctgcttttccgttttacgatataatgtctactcttggtg
cagaatgaattctcgttaactacatagcacaagtagatgtagttaactttaatctcacagAGGG
40 CCGCGGCAGCCTGCTGACCTGCGGCGACGTGGAGGAGAACCCCGGCCCATGGTGAGCAAGGG
CGAGGAGCTGTTACCGGCGTGGTGCCCATCTGGTGAGCTGGACGGCGACGTGAACGGCCA
CAAGTTCAGCGTGAGCGGCGAGGGCGAGGGCGACGCCACCTACGGCAAGCTGACCCTGAAGTT
CATCTGCACCACCGCAAGCTGCCCGTGCCCTGGGCCACCCTGGTGACCACCCTGACCTACGG
CGTGCACTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGAGCGCCAT
45 GCCCCGAGGGCTACGTGCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCG
CGCCGAGGTGAAGTTCGAGGGCGACACCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTT
CAAGGAGGACGGCAACATCCTGGGCCACAAGCTGGAGTACAACATAACAGCCACAACGTGTA
CATCATGGCCGACAAGCAGAAGAAGCGCATCAAGGTGAACCTCAAGATCCGCCACAACATCGA
GGACGGCAGCGTGACGTGGCCGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGT
50 GCTGCTGCCCGACAACCACTACCTGAGCACCCAGAGCGCCCTGAGCAAGGACCCCAACGAGAA
GCGCGACCACATGGTGCTGCTGGAGTTCGTGACCGCCCGCGCATCACCCCTGGGCATGGACGA
GCTGTACAAGGCCACCAACTTCAGCCTGCTGAAGCAGGCCGGCGACGTGGAGGAGAACCCCG

CCCCATGATCGAGCAGGACGGCCTGCACGCCGGCAGCCCCGCCGCCTGGGTGGAGCGCCTGTT
CGGCTACGACTGGGCCCAGCAGACCATCGGCTGCAGCGACGCCGCGTGTTCGCCTGAGCGC
CCAGGGCCGCCCCGTGCTGTTTCGTGAAGACCGACCTGAGCGGCGCCCTGAACGAGCTGCAGGA
CGAGGCCGCCCCGCCTGAGCTGGCTGGCCACCACCGGCGTGCCCTGCGCCGCCGTGCTGGACGT
5 GGTGACCGAGGCGCGCCGCGACTGGCTGCTGCTGGGCGAGGTGCCCCGCCAGGACCTGCTGAG
CAGCCACCTGGCCCCCGCCGAGAAGGTGAGCATCATGGCCGACGCCATGCGCCGCCTGCACAC
CCTGGACCCCGCCACCTGCCCCCTTCGACCACCAGGCCAAGCACCCGCATCGAGCGCGCCCGCAC
CCGCATGGAGGCCGGCCTGGTGGACCAGGACGACCTGGACGAGGAGCACCAGGGCCTGGCCCC
CGCCGAGCTGTTTCGCCCGCCTGAAGGCCAGCATGCCCCGACGGCGAGGACCTGGTGGTGACCCA
10 CGGCGACGCCTGCCTGCCCAACATCATGGTGGAGAACGGCCGCTTCAGCGGCTTCATCGACTG
CGGCAGGCTGGGCGTGGCCGACCGCTACCAGGACATCGCCCTGGCCACCCGCGACATCGCCGA
GGAGCTGGGCGGCGAGTGGGCCGACCGCTTCCTGGTGCTGTACGGCATCGCCGCCCCCGACAG
CCAGCGCATCGCCTTCTACCGCCTGCTGGACGAGTTCTTCTAAcaatctttaatcagtgtgta
acattagggaggacttgaaagagccaccacattttaccgaggccacgcggagtacgatcgag
15 tgtacagtgaacaatgctagggagagctgcctatatggaagagccctaattgtgtaaaattaat
tttagtagtgctatccccatgtgattttaatagcttcttaggagaatgacaaaaaaaaaaaaa
aaaaaaaaaaaaaaaaaaaaaaaaAAAAAAAAAagggtcgccatggcatctccacctctcgcg
tccgacctgggcatccgaaggaggacgcacgtccactcgcatggctaagggagGCTAGCcaga
catgataagatacattgatgagtttggacaaaccacaactagaatgcagtgaaaaaaatgctt
20 tatttgtgaaatttgtgatgctattgctttatttgttaaccattataagctgcaataaacaagt
taacaacaacaattgcattcattttatgtttcaggttcagggggaggtgtgggaggtttttta
aagcaagtaaaacctctacaaatgtggtaCAGAAAGTCAAAAGCCTCCGACCCGAGGCTTTTG
ACTATTACTGCCGACATGGAAGCCATCACAAACGGCATGATGAACCTATAAACGCAGAAAGGC
CCACCCGAAGGTGAGCCAGTGTGATTACCAGCGGCATCAGCACCTTGTCGCCTTGCGTATAAT
25 ATTTGCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAA

实施例 1 中使用的 5'UTR (SEQ ID NO: 12)

ATTAAAGGTTTATACCTTCCCAGGTAACAAACCAACCAACTTTCGATCTCTTGTAGATCTGTT
CTCTAAACGAACCTTTAAATCTGTGTGGCTGTCACTCGGCTGCATGCTTAGTGCACCTACGCA
30 GTATAATTAATACTAATTACTGTCTTTGACAGGACACGAGTAACCTCGTCTATCTTCTGCAGG
CTGCTTACGGTTTCGTCCGTGTTGCAGCCGATCATCAGCACATCTAGGTTTCGTCCGGGTGTG
ACCGAAAGGTAAG

实施例 1 中使用的 3'UTR (包含 polyA 尾巴, SEQ ID NO: 13)

CAATCTTTAATCAGTGTGTAACATTAGGGAGGACTTGAAAGAGCCACCACATTTTCACCGAGG
CCACGCGGAGTACGATCGAGTGTACAGTGAACAATGCTAGGGAGAGCTGCCTATATGGAAGAG
CCCTAATGTGTAAAATTAATTTTAGTAGTGCTATCCCCATGTGATTTTAATAGCTTCTTAGGA
GAATGACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

40 本发明的实施方式并不限于上述实施例所述, 在不偏离本发明的精神和范围的情况
下, 本领域普通技术人员可以在形式和细节上对本发明做出各种改变和改进, 而这些均
被认为落入了本发明的保护范围。

权 利 要 求 书

1、用于制备 SARS-CoV-2 病毒样颗粒的组合物，包括下述（a），还包括选自下述（b）和（c）中的任一种或任两种：

（a）包含 Stru Δ S 片段的第一质粒，所述 Δ S 片段包含编码 SARS-CoV-2 病毒的至少一种结构蛋白或其突变体和/或至少一种辅助蛋白或其突变体的核酸序列，且所述 Δ S 片段不包含编码 SARS-CoV-2 病毒的 S 蛋白或其突变体的核酸序列；

（b）包含 S 片段的第二质粒，所述 S 片段包含编码 SARS-CoV-2 病毒的 S 蛋白或其突变体的核酸序列；

（c）包含 SARS-CoV-2 病毒的包装信号片段的第三质粒，所述包装信号片段包含 SARS-CoV-2 病毒的包装信号序列。

2、根据权利要求 1 所述的组合物，其中所述的 S 蛋白的突变体相对于 SARS-CoV-2 病毒的 S 蛋白的突变包括 N331Q、N501Y、D614G 和/或 P681H。

3、根据权利要求 1 或 2 所述的组合物，其中所述 Stru Δ S 片段包含下述任一种：

（1）编码 SARS-CoV-2 病毒的除 S 蛋白之外的全部结构蛋白和全部辅助蛋白的核酸序列；

（2）与（1）相比，所述结构蛋白和辅助蛋白中的至少一种相对于野生型 SARS-CoV-2 病毒具有突变；和

（3）与（1）相比，缺失所述结构蛋白和辅助蛋白中的任一种。

4、根据权利要求 1-3 任一项所述的组合物，其中所述包装信号序列是 NCBI 序列号 NC_045512.2 的第 19900-20000 位核苷酸或 NCBI 序列号 NC_045512.2 的第 19773-20335 位核苷酸。

5、根据权利要求 1-4 任一项所述的组合物，其中所述包装信号片段包含 SARS-CoV-2 病毒基因组的 ORF1ab。

6、权利要求 1-5 任一项所述的组合物的制备方法，包括分别通过下述方法制备第一质粒、第二质粒和/或第三质粒：

（1）将 Stru Δ S 片段、S 片段和/或 ORF1ab 包装信号片段分别拆分为在相邻片段之间具有同源序列的 DNA 短片段；

（2）将由步骤（1）获得的 DNA 短片段和线性化质粒载体转入酵母细胞中，进行同源重组，分别获得第一质粒、第二质粒和/或第三质粒。

7、SARS-CoV-2 病毒样颗粒的制备方法，包括：

用权利要求 1-5 任一项的组合物转染包装细胞，获得 SARS-CoV-2 病毒样颗粒。

8、根据权利要求 7 所述的制备方法，还包括

分别通过下述方法制备第一质粒、第二质粒和/或第三质粒：

5 (1) 将 Stru Δ S 片段、S 片段和/或包装信号片段分别拆分为在相邻片段之间具有同源序列的 DNA 短片段；

(2) 将由步骤 (1) 获得的 DNA 短片段和线性化质粒载体转入酵母细胞中，进行同源重组，分别获得第一质粒、第二质粒和/或第三质粒。

9、根据权利要求 8 所述的制备方法，进一步包括将第一质粒、第二质粒和/或第三质粒分别在大肠杆菌中进行扩增，然后转染包装细胞。

10、由权利要求 7-9 任一项所述的制备方法制备获得的 SARS-CoV-2 病毒样颗粒。

11、根据权利要求 1-5 任一项的组合物或根据权利要求 10 所述的 SARS-CoV-2 病毒样颗粒在制备用于预防或治疗 SARS-CoV-2 病毒感染的疫苗中的用途。

12、根据权利要求 1-5 任一项的组合物或根据权利要求 10 所述的 SARS-CoV-2 病毒样颗粒在体外用于 SARS-CoV-2 病毒感染细胞的研究中的用途。

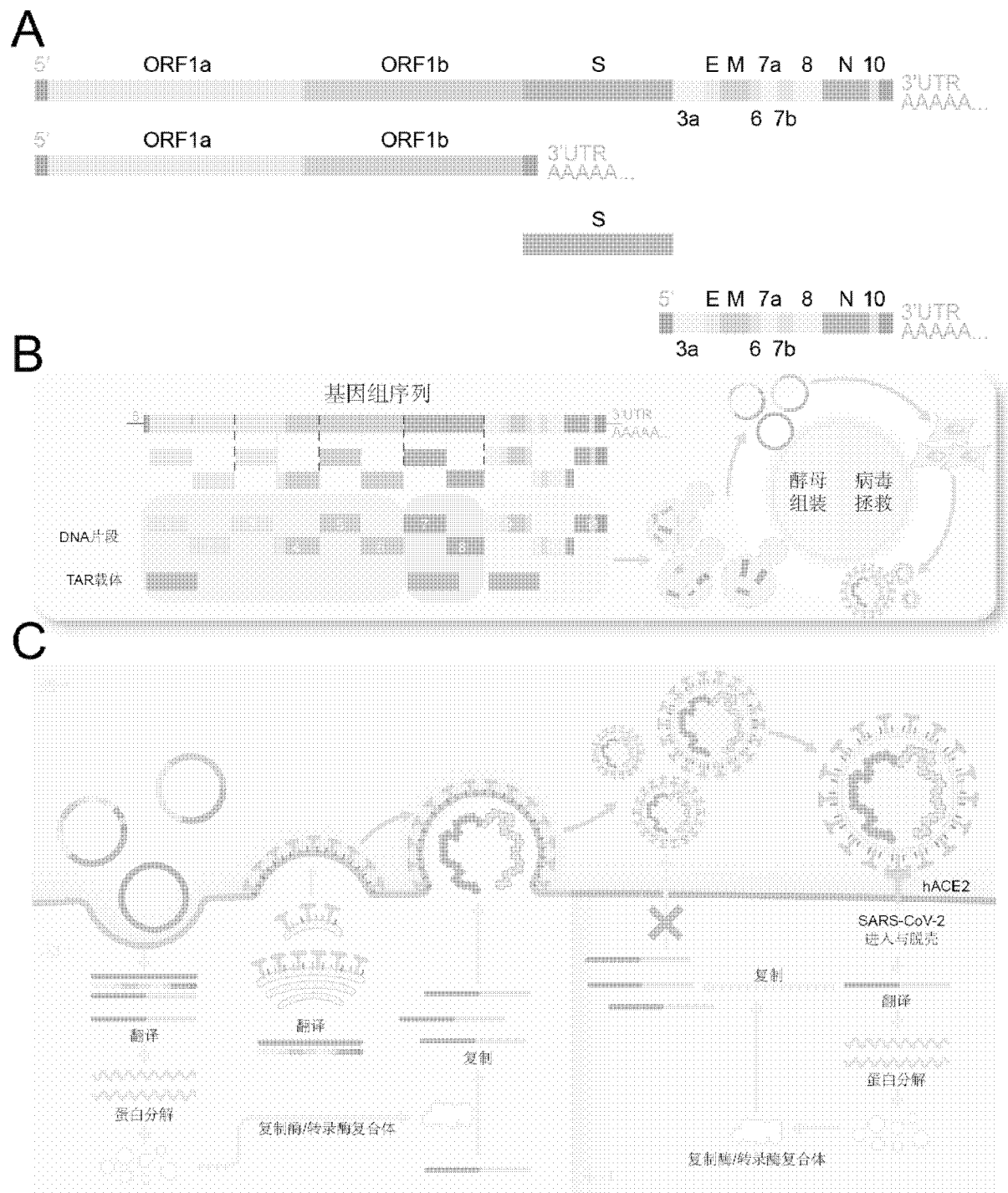


图 1

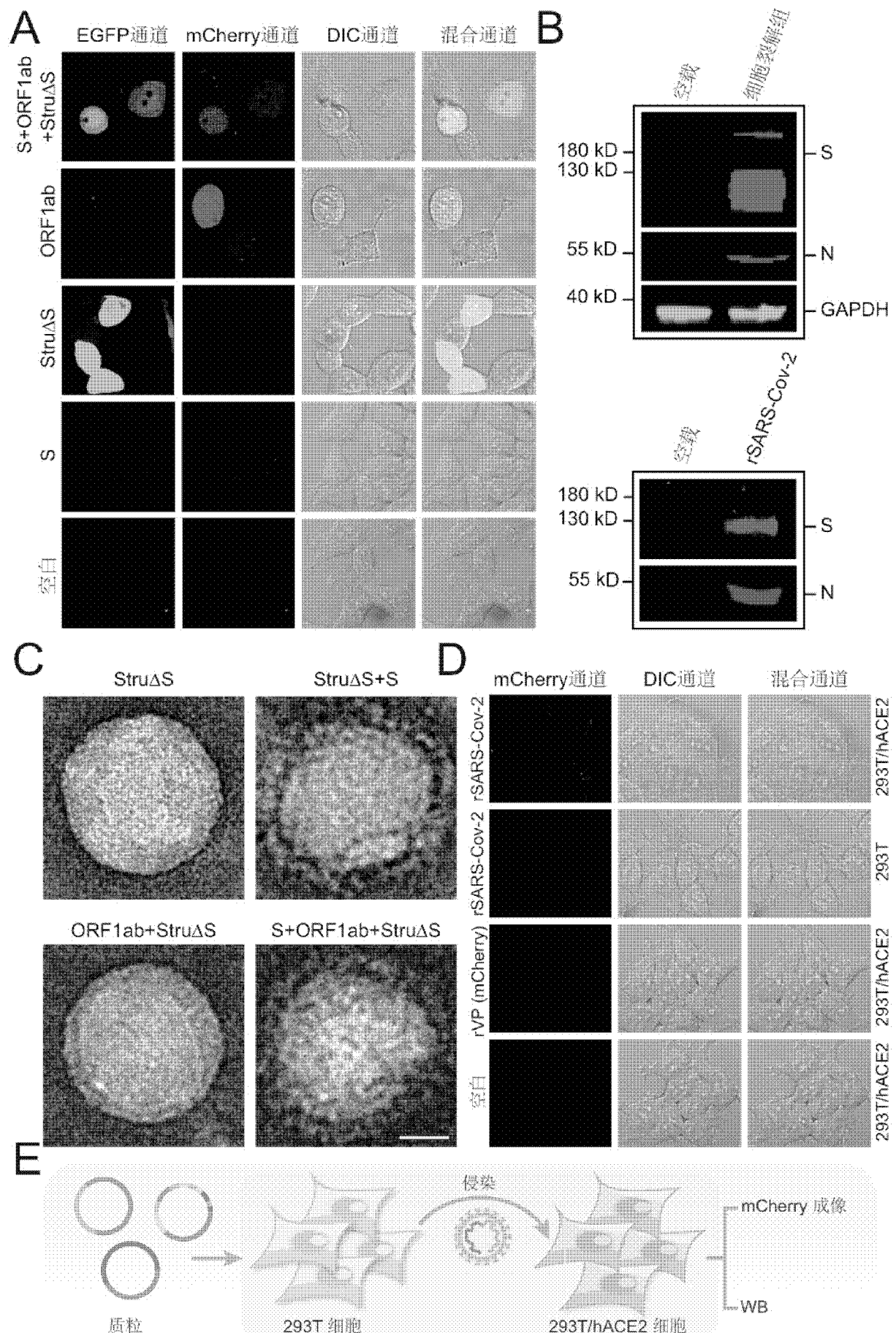


图 2

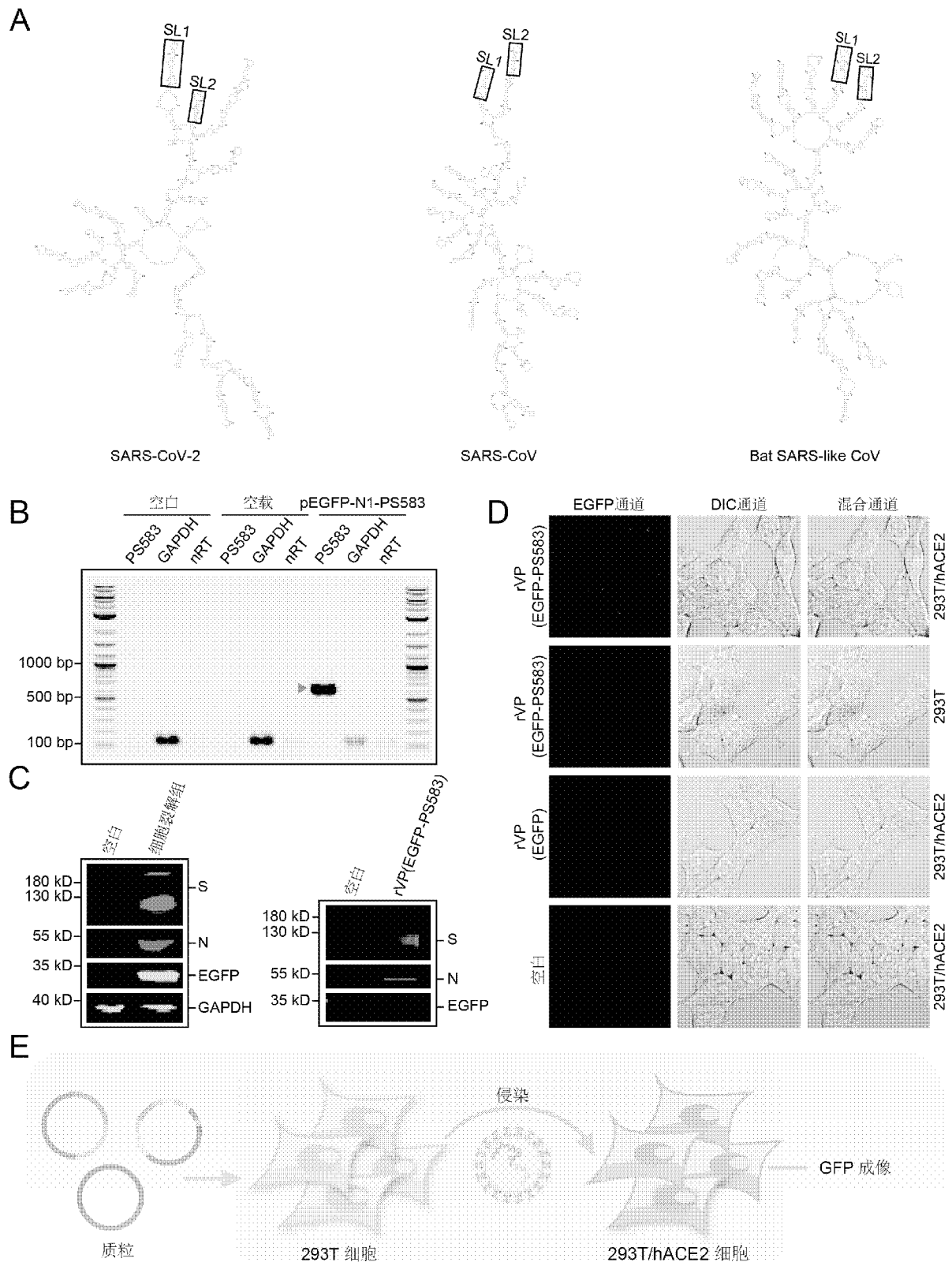


图 3

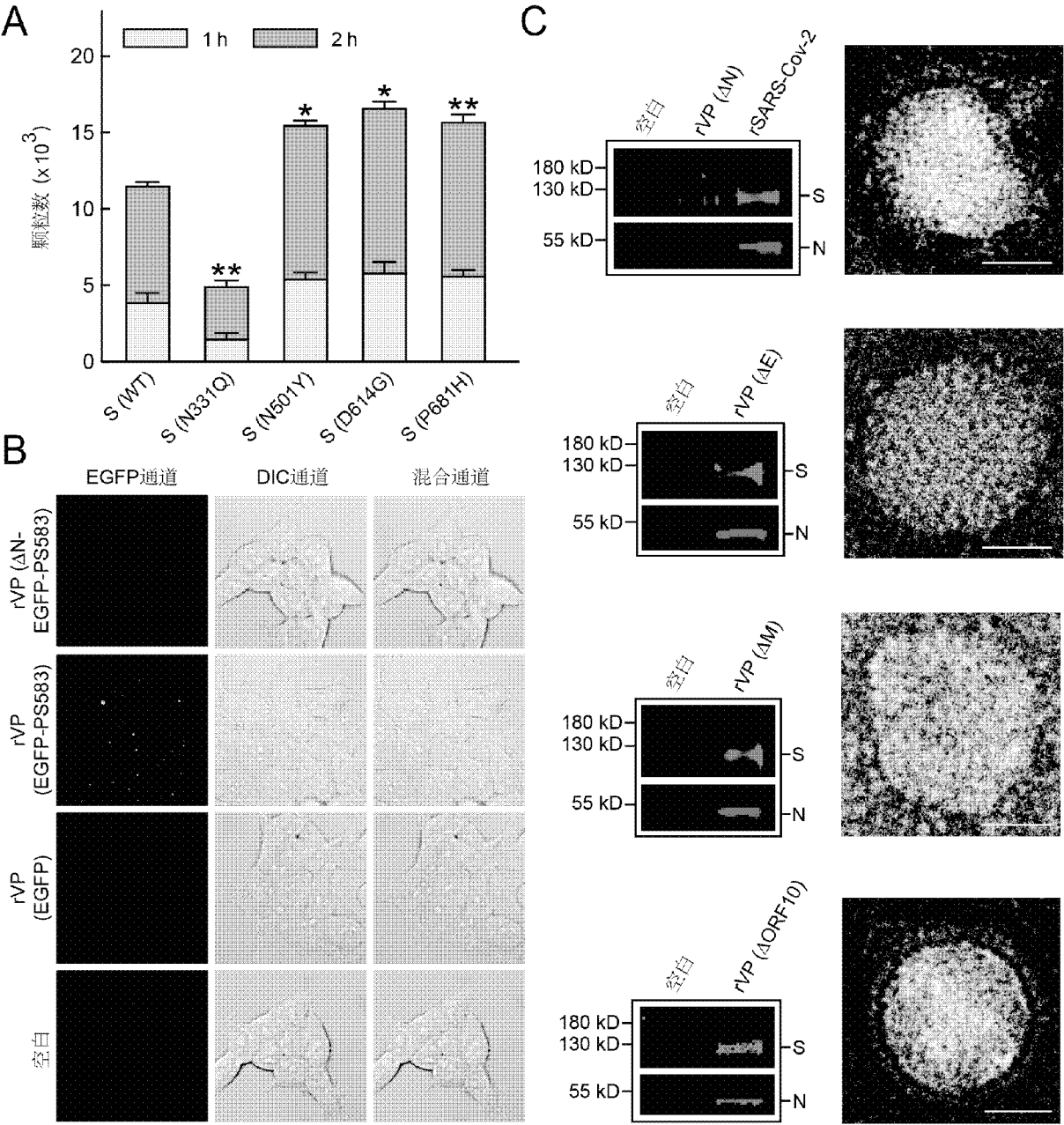


图 4

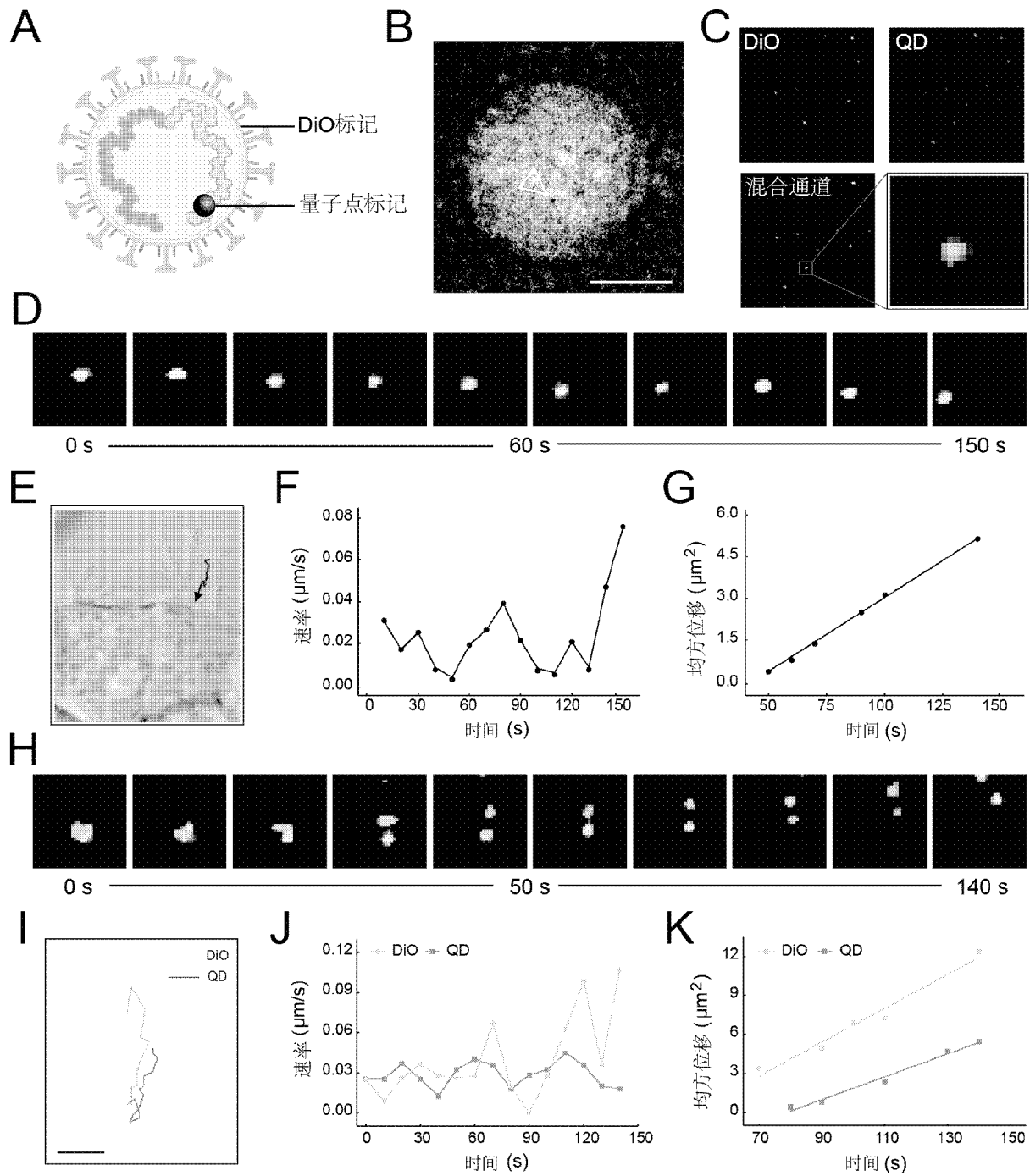


图 5

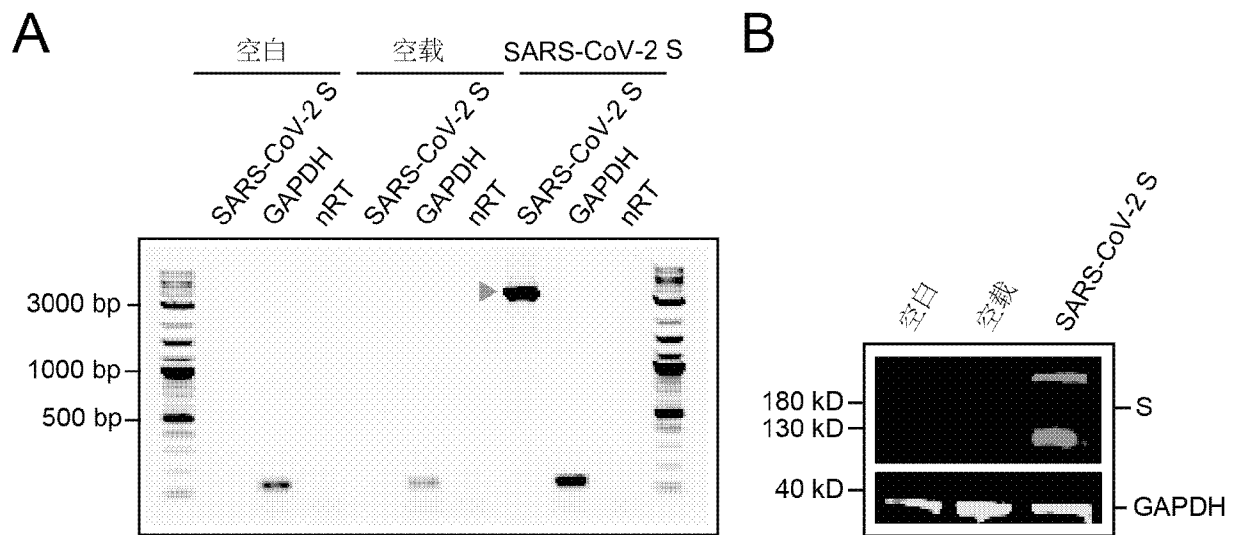


图 6

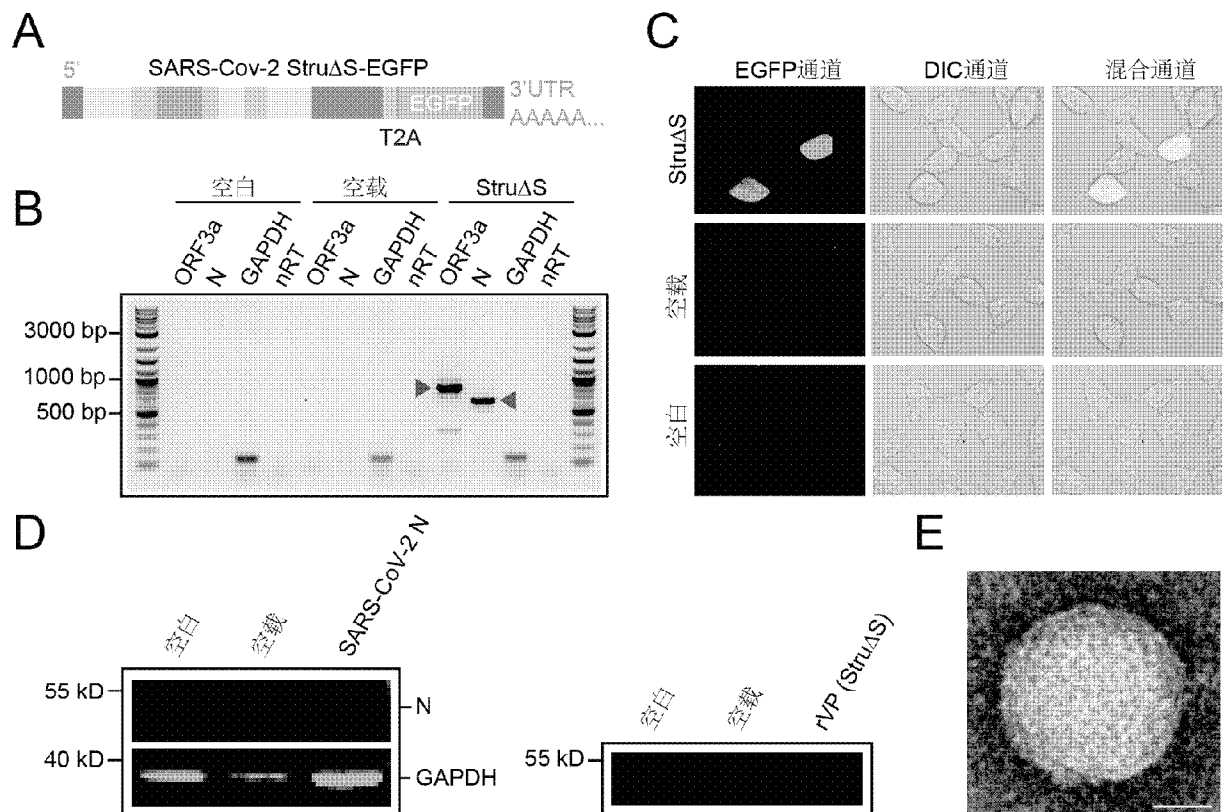


图 7

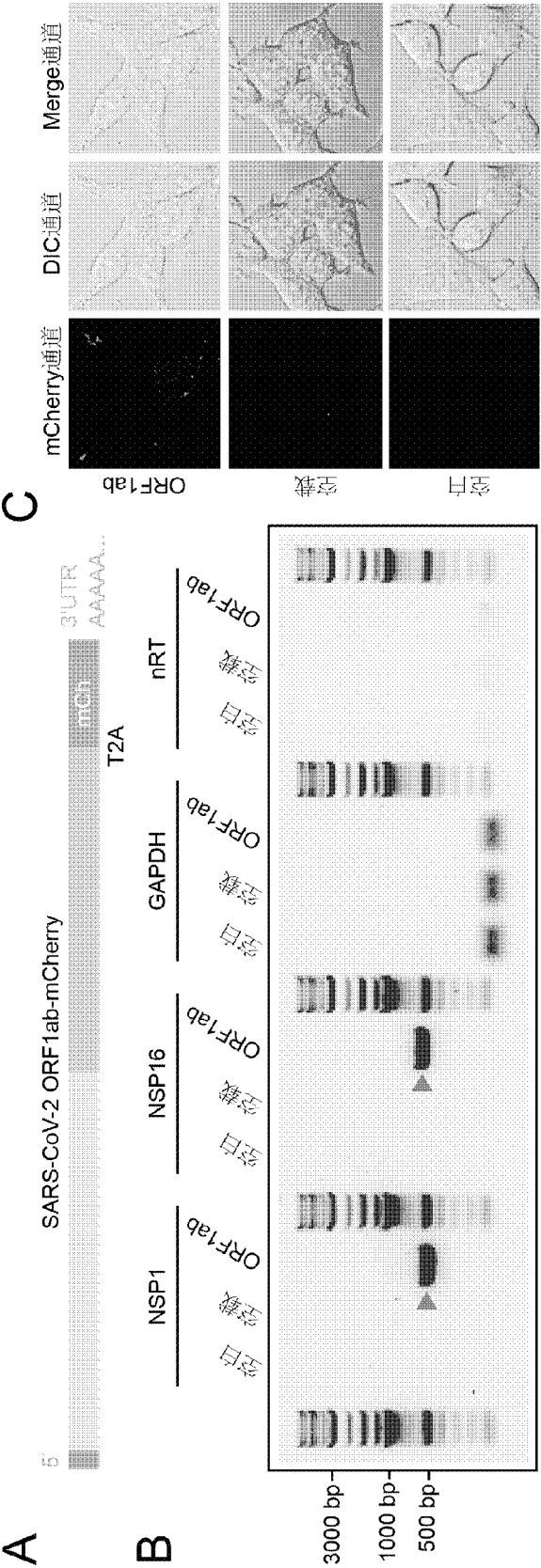


图 8

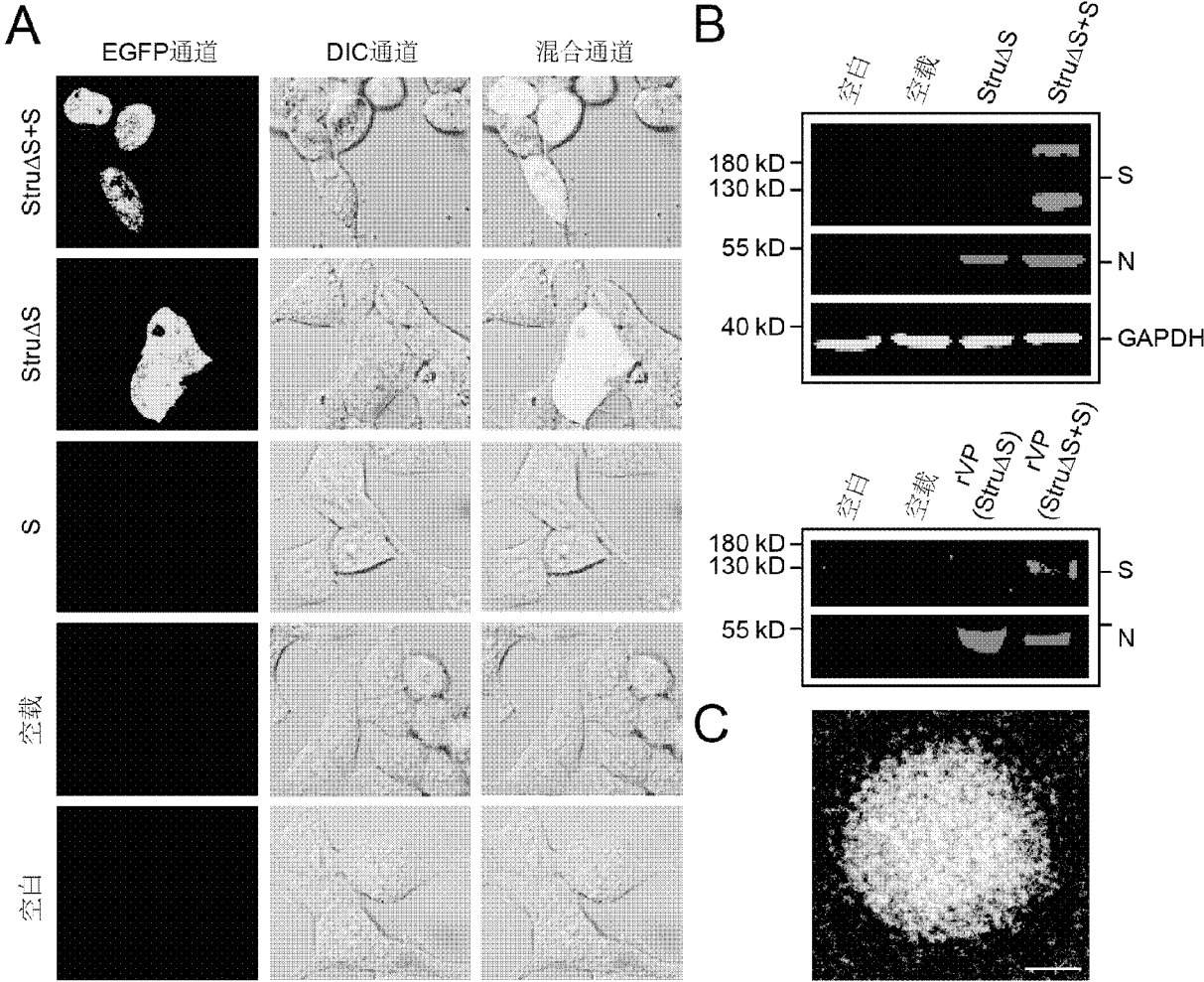


图 9

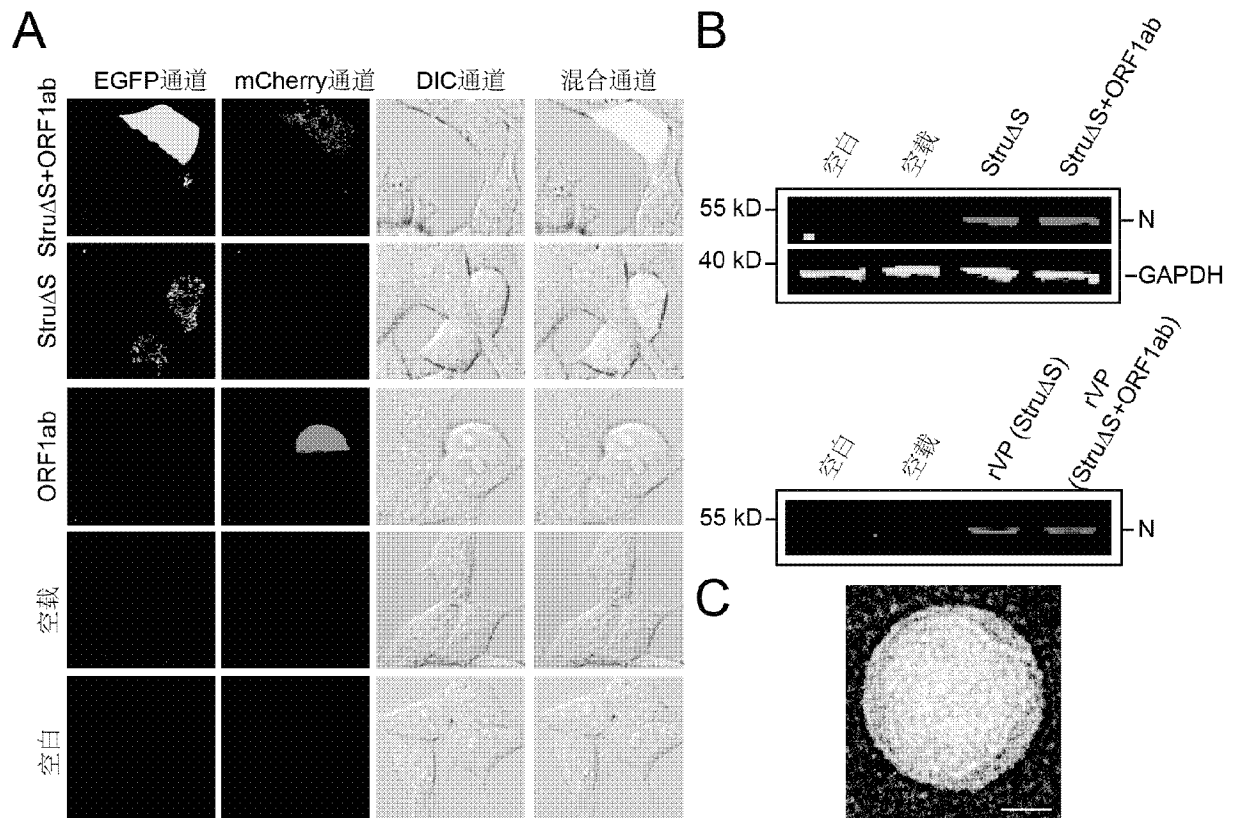


图 10

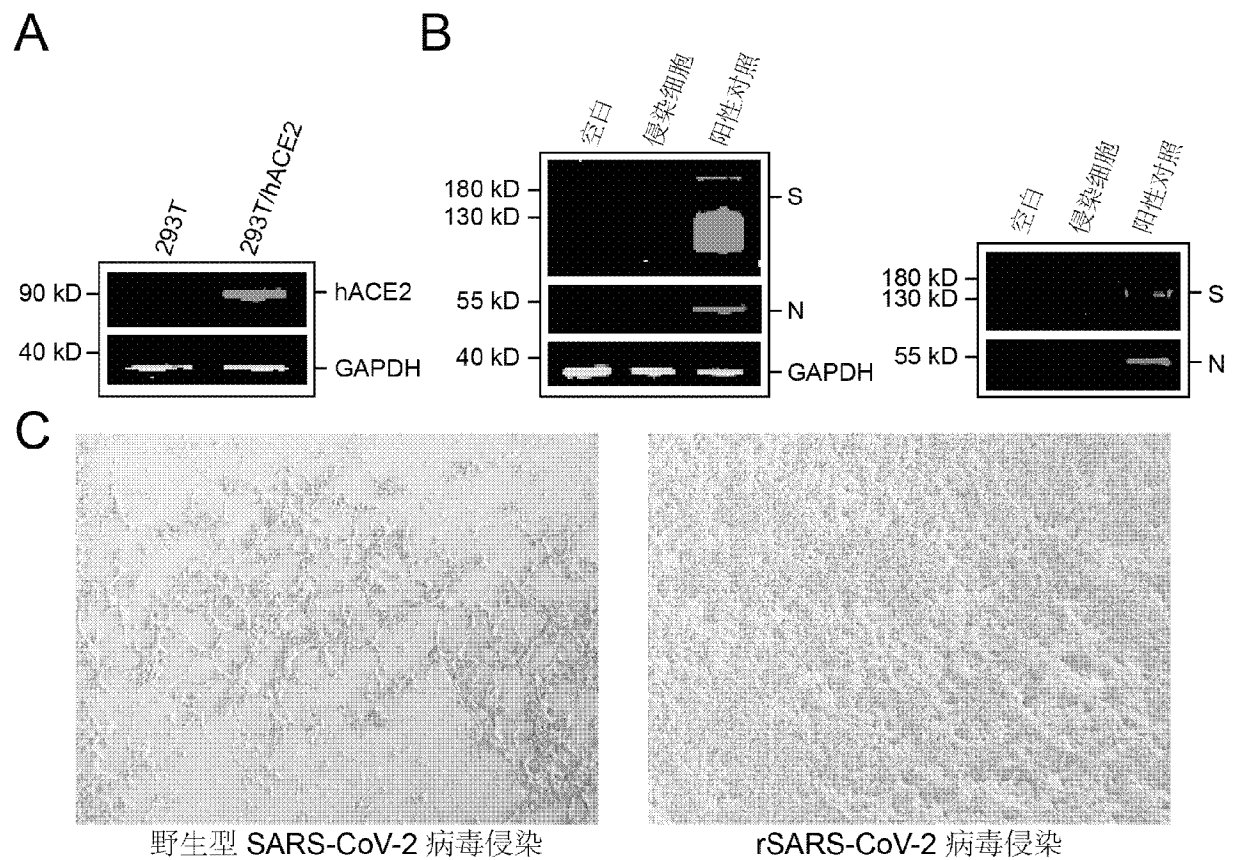


图 11

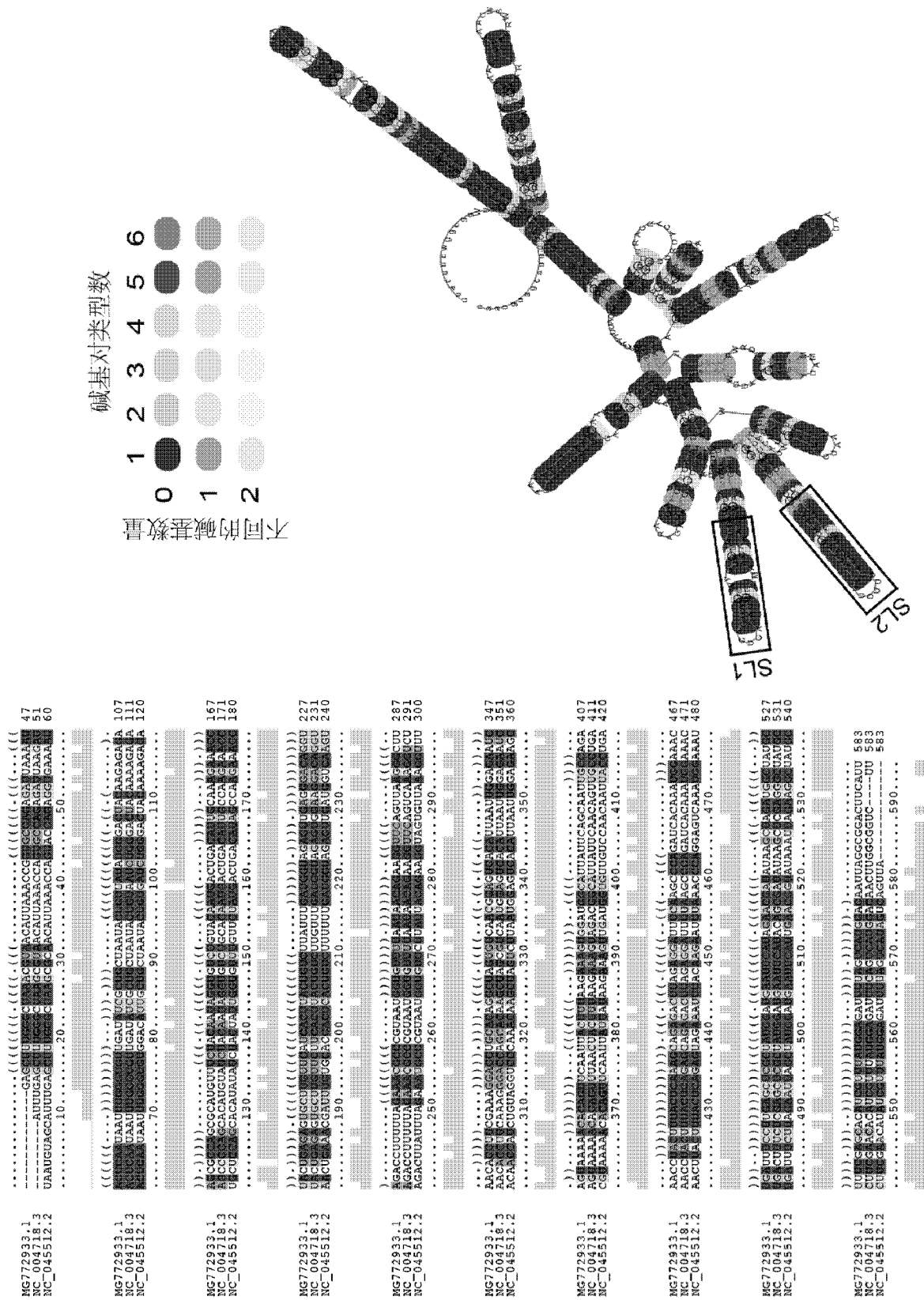


图 12

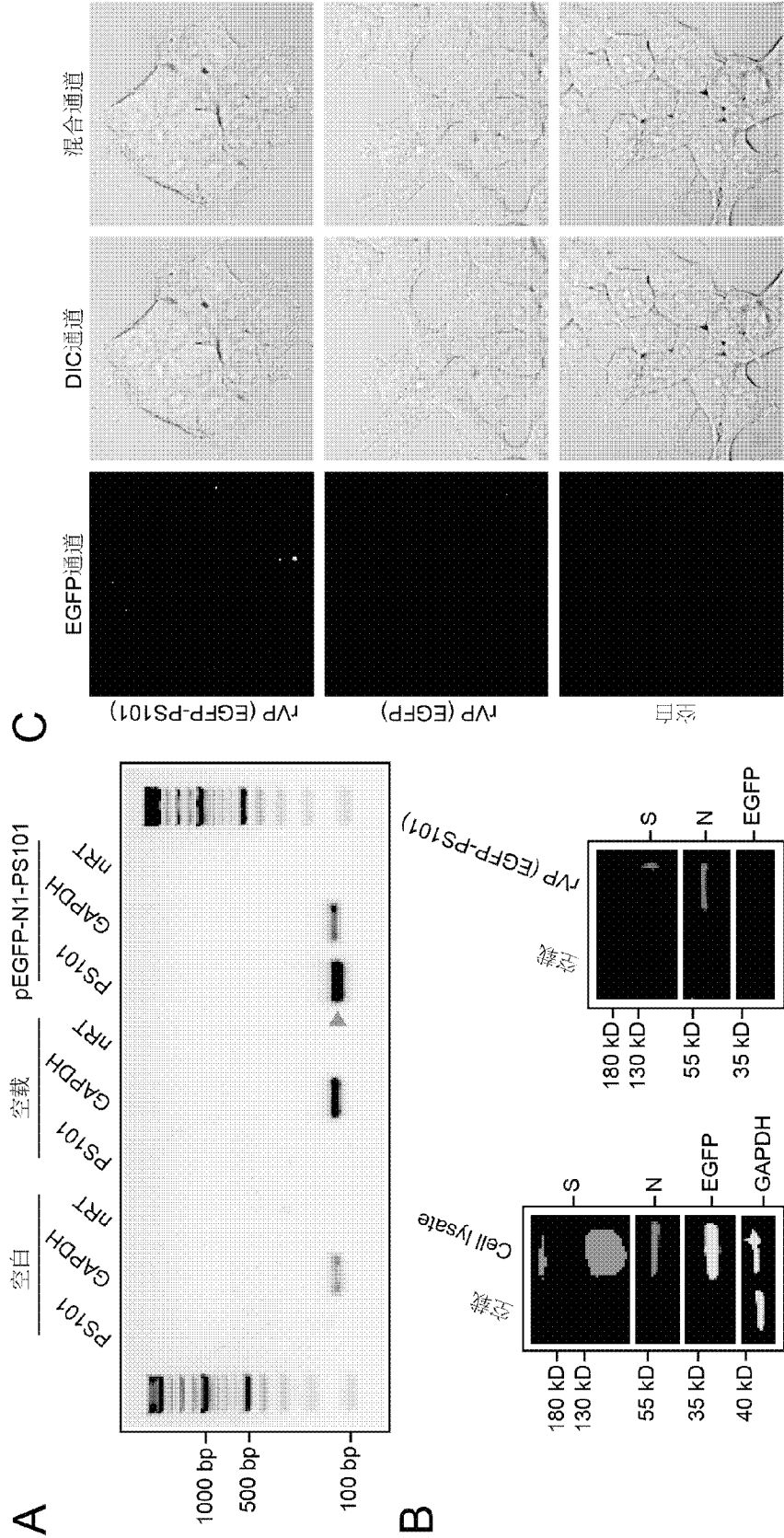


图 13

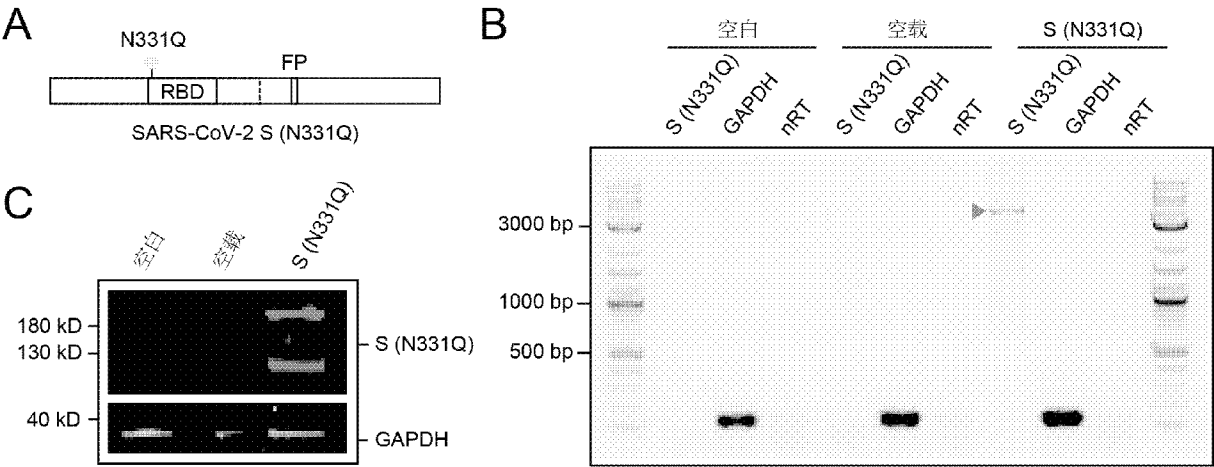


图 14

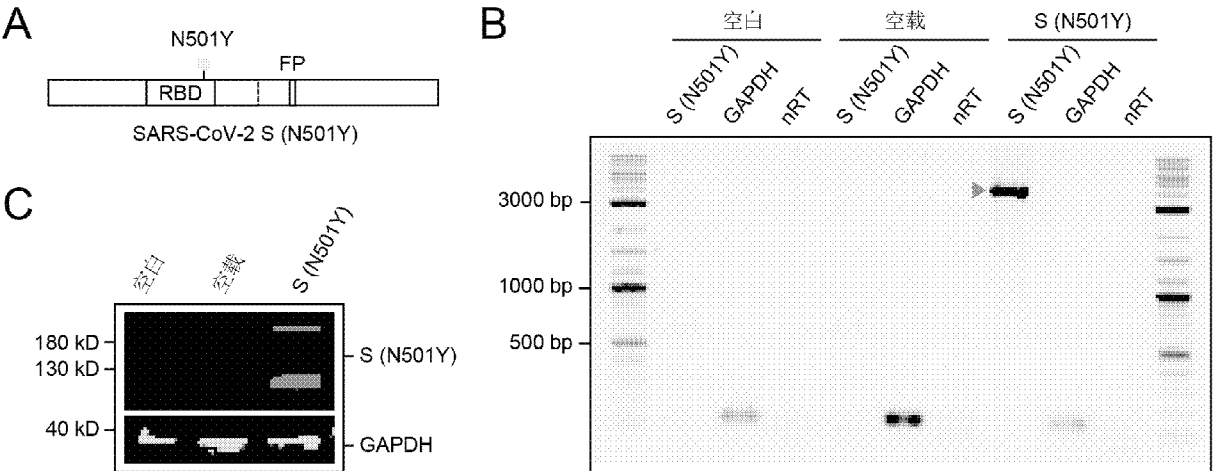


图 15

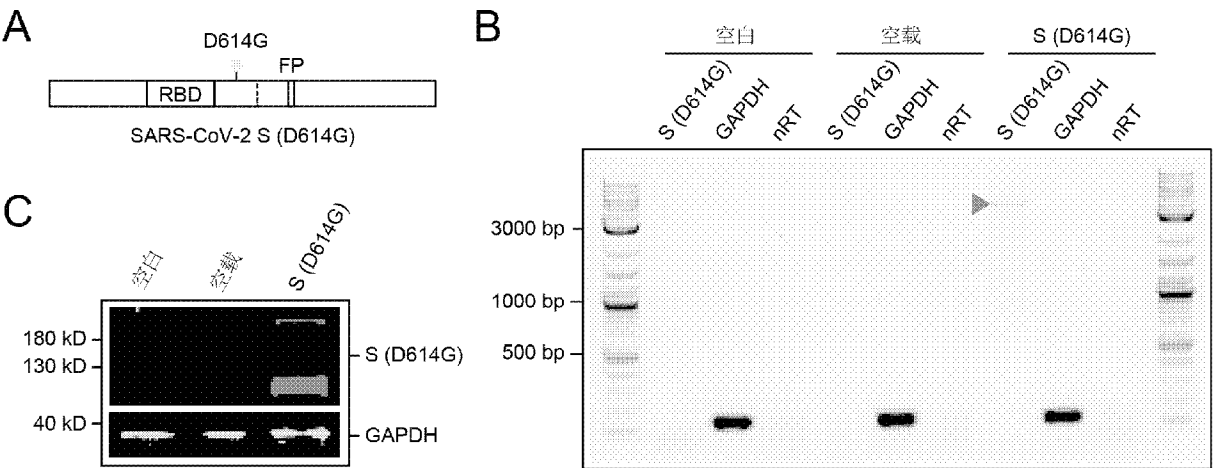


图 16

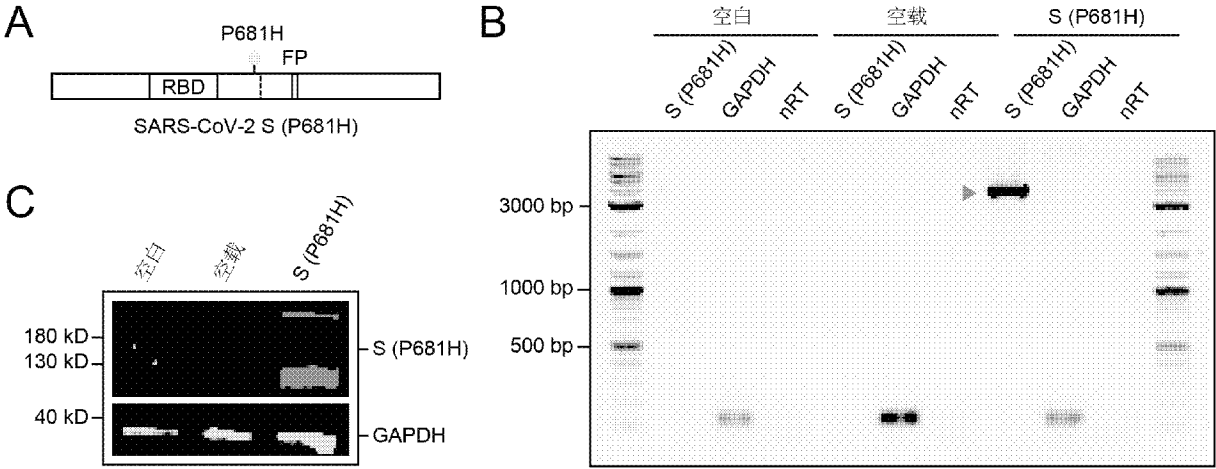


图 17

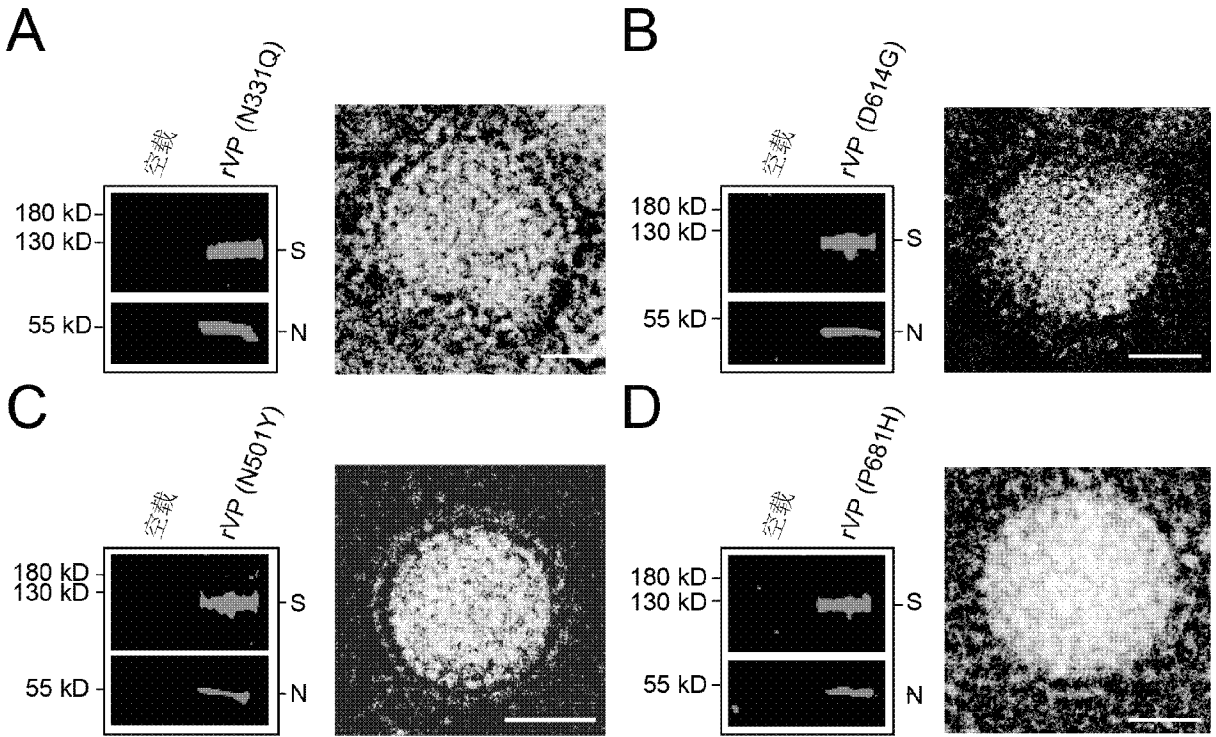


图 18

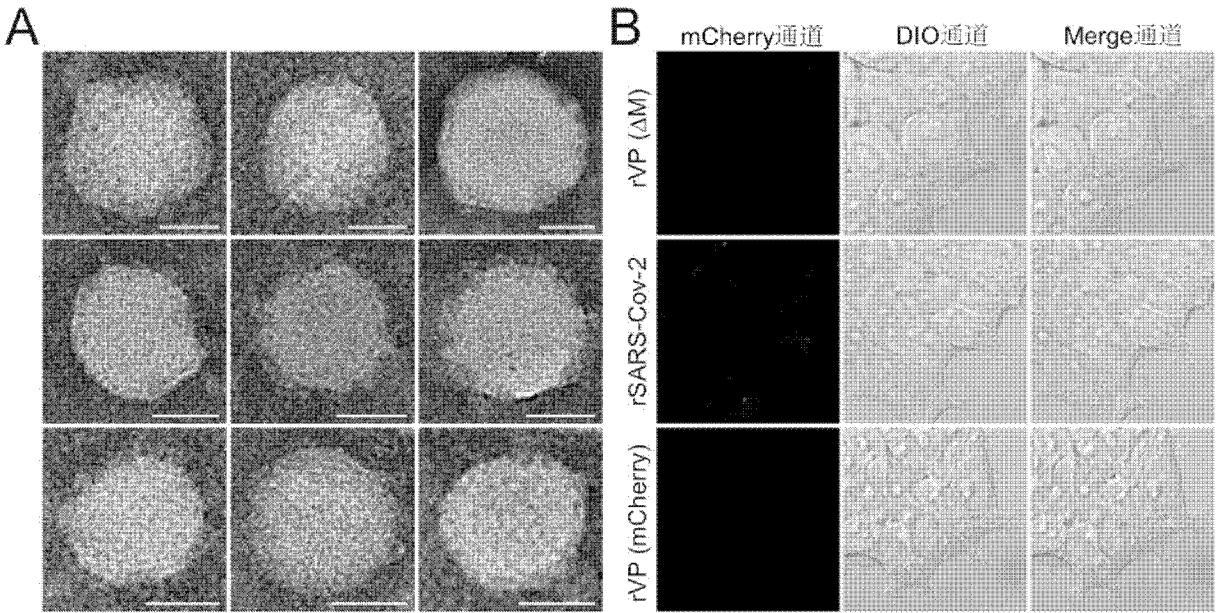


图 23

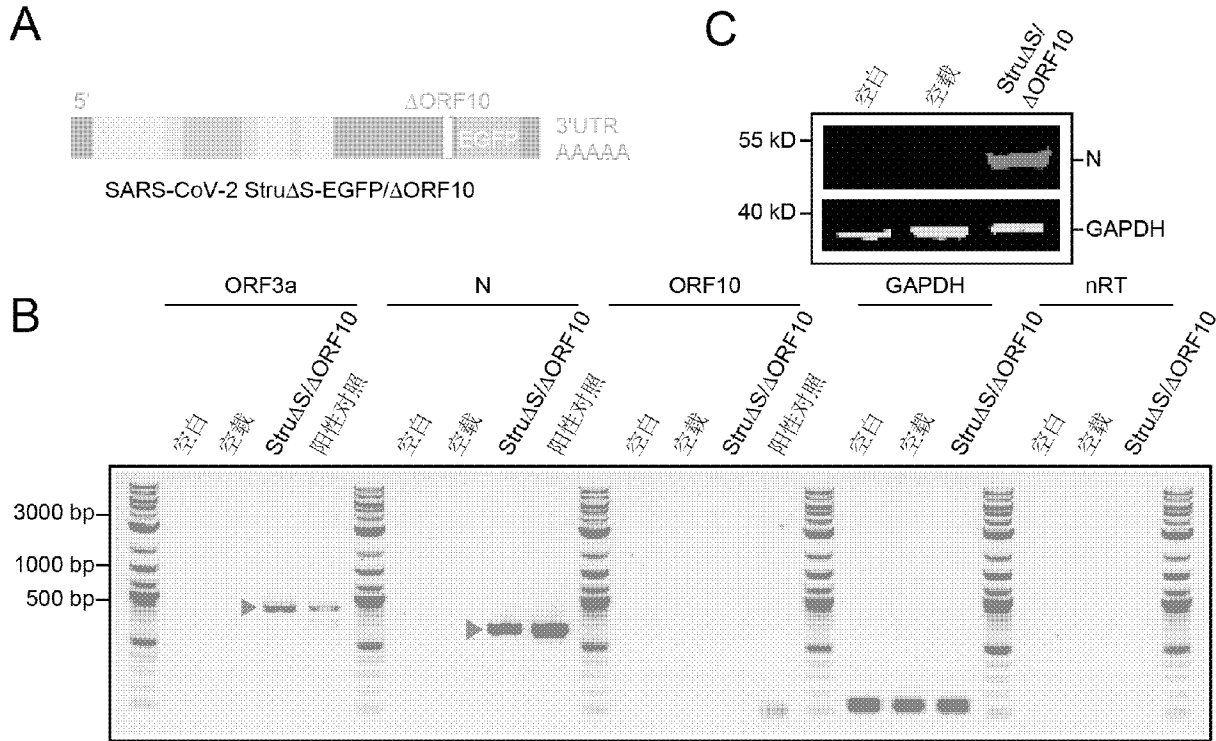


图 24

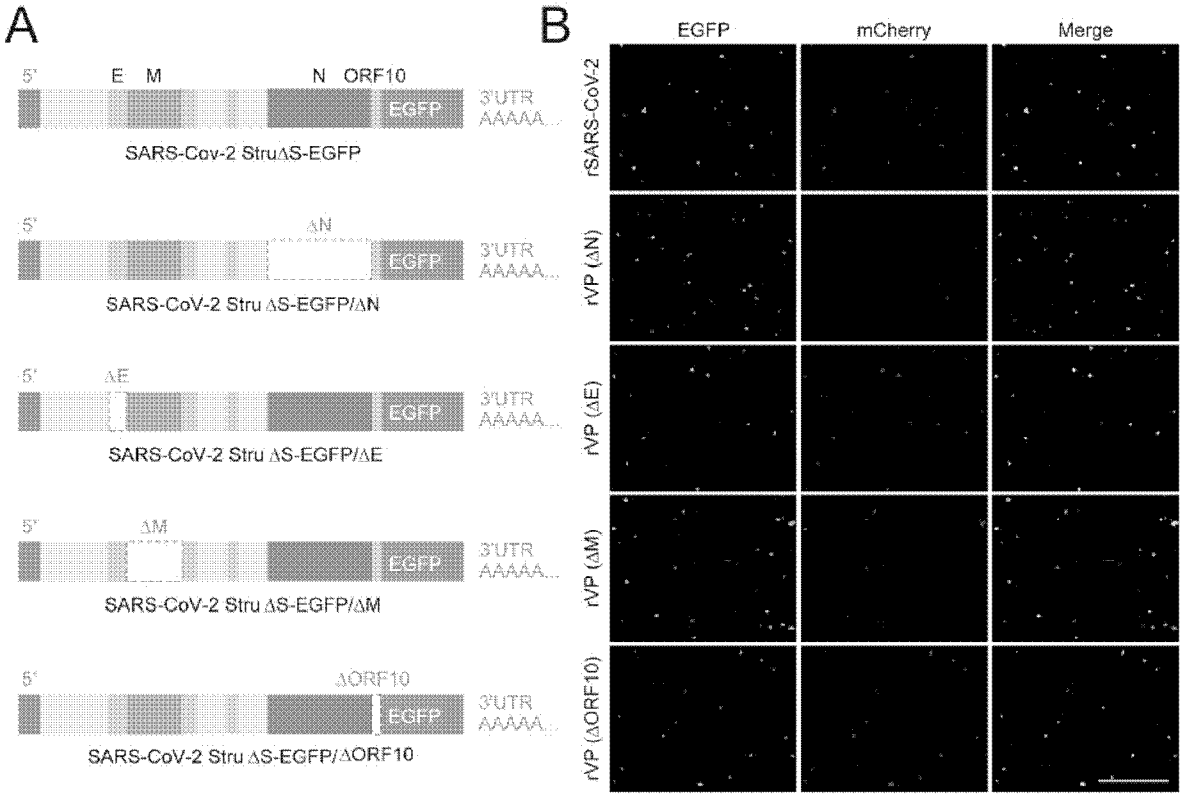


图 25

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/138037

A. CLASSIFICATION OF SUBJECT MATTER

C07K 14/165(2006.01)i; C12N 15/50(2006.01)i; C12N 15/85(2006.01)i; A61K 39/215(2006.01)i; A61P 31/14(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C07K; C12N; A61K; A61P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPTXT; VEN; USTXT; WOTXT; CNABS; GOOGLE SCHOLAR; PUBMED; ISI WEB OF SCIENCE; CNKI; 万方, WANFANG; GENBANK; STN; SEQ ID NOs: 1-3, SARS-CoV-2, 刺突蛋白, S蛋白, StruΔS, VLP, 假病毒, 病毒样颗粒, 刺突糖蛋白, ORF1ab, S片段, Spike, E蛋白, 小包膜蛋白, Envelope, M蛋白, 膜糖蛋白, Membrane, N蛋白, 核衣壳蛋白, nucleocapsid, 同源重组, 酵母, 表达载体

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 114395017 A (SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY, CHINESE ACADEMY OF SCIENCES) 26 April 2022 (2022-04-26) claims 1-12	1-12
X	WO 2021176397 A2 (PREMAS BIOTECH PVT. LTD.) 10 September 2021 (2021-09-10) description, paragraphs 77-95, embodiment 4.2	1, 7, 10
Y	WO 2021176397 A2 (PREMAS BIOTECH PVT. LTD.) 10 September 2021 (2021-09-10) description, paragraphs 77-95, embodiment 4.2	1-12
Y	WO 2021211749 A1 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 21 October 2021 (2021-10-21) claims 1-21, and description, paragraphs 156, 232-236, and 378	1-12
Y	US 2021299245 A1 (SEMENTIS LIMITED) 30 September 2021 (2021-09-30) abstract, description, paragraphs 20-32, 98-115, 209, 215, and 244, embodiment 1, and figure 3	1-12



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

13 July 2022

Date of mailing of the international search report

28 July 2022

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/CN)

No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088, China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/138037

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 113248579 A (CHONGQING MEDICAL UNIVERSITY et al.) 13 August 2021 (2021-08-13) description, paragraphs 19-37	1-12
A	CN 111647053 A (ACADEMY OF MILITARY SCIENCES, PHYSIOME INSTITUTE) 11 September 2020 (2020-09-11) entire document	1-12
A	刘彬 等 (LIU, Bin et al.). "新型冠状病毒基因组结构与蛋白功能 (Genomic Structure and Protein Profiles of 2019 Novel Coronavirus)" <i>微生物与感染 (Journal of Microbes and Infections)</i> , Vol. 15, No. 1, 25 February 2020 (2020-02-25), pages 52-57	1-12
A	THOMAS, S. "Mapping the Nonstructural Transmembrane Proteins of Severe Acute Respiratory Syndrome Coronavirus 2" <i>JOURNAL OF COMPUTATIONAL BIOLOGY</i> , Vol. 28, No. 9, 28 June 2021 (2021-06-28), pages 909-921	1-12

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2021/138037

Box No. I **Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)**

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☒ forming part of the international application as filed:
 - ☒ in the form of an Annex C/ST.25 text file.
 - ☐ on paper or in the form of an image file.
 - b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
 - c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
 - ☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).
 - ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2021/138037

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	114395017	A	26 April 2022	None			
WO	2021176397	A2	10 September 2021	None			
WO	2021211749	A1	21 October 2021	WO	2021211760	A1	21 October 2021
				WO	2021211748	A1	21 October 2021
US	2021299245	A1	30 September 2021	WO	2021195694	A1	07 October 2021
				TW	202203966	A	01 February 2022
CN	113248579	A	13 August 2021	None			
CN	111647053	A	11 September 2020	None			

国际检索报告

国际申请号

PCT/CN2021/138037

A. 主题的分类

C07K 14/165(2006.01)i; C12N 15/50(2006.01)i; C12N 15/85(2006.01)i; A61K 39/215(2006.01)i; A61P 31/14(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

C07K; C12N; A61K; A61P

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

EPTXT;VEN;USTXT;WOTXT;CNABS;GOOGLE SCHOLAR;PUBMED;ISI WEB OF SCIENCE;CNKI;万方;GENBANK;STN:SEQ ID N0s:1-3, SARS-CoV-2, 刺突蛋白, S蛋白, Stru Δ S, VLP, 假病毒, 病毒样颗粒, 刺突糖蛋白, ORF1ab, S片段, Spike, E蛋白, 小包膜蛋白, Envelope, M蛋白, 膜糖蛋白, Membrane, N蛋白, 核衣壳蛋白, nucleocapsid, 同源重组, 酵母, 表达载体

C. 相关文件

类 型*

引用文件, 必要时, 指明相关段落

相关的权利要求

PX

CN 114395017 A (中国科学院深圳先进技术研究院) 2022年4月26日 (2022 - 04 - 26)
权利要求1-12

1-12

X

WO 2021176397 A2 (PREMAS BIOTECH PVT. LTD.) 2021年9月10日 (2021 - 09 - 10)
说明书第77-95段, 实施例4.2

1, 7, 10

Y

WO 2021176397 A2 (PREMAS BIOTECH PVT. LTD.) 2021年9月10日 (2021 - 09 - 10)
说明书第77-95段, 实施例4.2

1-12

Y

WO 2021211749 A1 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 2021年10月21日 (2021 - 10 - 21)
权利要求1-21, 说明书第156、232-236、378段

1-12

Y

US 2021299245 A1 (SEMENTIS LIMITED) 2021年9月30日 (2021 - 09 - 30)
摘要, 说明书第20-32、98-115、209、215、244、实施例1, 图3

1-12

Y

CN 113248579 A (重庆医科大学 等) 2021年8月13日 (2021 - 08 - 13)
说明书第19-37段

1-12



其余文件在C栏的续页中列出。



见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2022年7月13日

国际检索报告邮寄日期

2022年7月28日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

授权官员

李恩

电话号码 86-(10)-53961874

C. 相关文件		
类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	CN 111647053 A (军事科学院军事医学研究院生命组学研究所) 2020年9月11日 (2020 - 09 - 11) 全文	1-12
A	刘彬 等. “新型冠状病毒基因组结构与蛋白功能” 微生物与感染, 第15卷, 第1期, 2020年2月25日 (2020 - 02 - 25), 第52-57页	1-12
A	THOMAS, S. “Mapping the Nonstructural Transmembrane Proteins of Severe Acute Respiratory Syndrome Coronavirus 2” JOURNAL OF COMPUTATIONAL BIOLOGY, 第28卷, 第9期, 2021年6月28日 (2021 - 06 - 28), 第909-921页	1-12

第I栏

核苷酸和/或氨基酸序列(续第1页第1.c项)

1. 关于国际申请中所公开的任何核苷酸和/或氨基酸序列, 国际检索是基于下列序列表进行的:

a. ☒ 作为国际申请的一部分提交的:

☒ 附件C/ST. 25文本文件形式

☐ 纸件或图形文件形式

b. ☐ 根据细则13之三. 1(a) 仅为国际检索目的以附件C/ST. 25文本文件形式与国际申请同时提交的:

c. ☐ 仅为国际检索目的在国际申请日之后提交的:

☐ 附件C/ST. 25文本文件形式(细则13之三. 1(a))

☐ 纸件或图形文件形式(细则13之三. 1(b)和行政规程第713段)

2. ☐ 另外, 在提交/提供了多个版本或副本的序列表的情况下, 提供了关于随后提交的或附加的副本中的信息与申请时提交的作为申请一部分的序列表的信息相同或未超出申请时提交的申请中的信息范围(如适用)的所需声明。

3. 补充意见:

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2021/138037

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	114395017	A	2022年4月26日	无	
WO	2021176397	A2	2021年9月10日	无	
WO	2021211749	A1	2021年10月21日	WO 2021211760 A1	2021年10月21日
				WO 2021211748 A1	2021年10月21日
US	2021299245	A1	2021年9月30日	WO 2021195694 A1	2021年10月7日
				TW 202203966 A	2022年2月1日
CN	113248579	A	2021年8月13日	无	
CN	111647053	A	2020年9月11日	无	