



SARS-CoV-2基因组的不寻常特征表明其经由实验室复杂改造的特性非源于自然进化及其可能的合成途径的描述

Li-Meng Yan (MD, PhD)¹, Shu Kang (PhD)¹, Jie Guan (PhD)¹, Shanchang Hu (PhD)¹

¹法治社会及法治基金, 纽约, 美国

联络方式: team.lmyan@gmail.com

DOI: 10.5281/zenodo.4028830

【中文翻译】

摘要

由新型冠状病毒 SARS-CoV-2 引起的新冠肺炎 (COVID-19) 大流行已在全球造成 91 万多人死亡, 使全球经济受到前所未有的严重破坏。尽管影响巨大, 但 SARS-CoV-2 的起源仍然神秘且富有争议。自然起源理论虽然被广泛接受, 但缺乏实质性的支持。然而, 另一种认为病毒可能来自研究实验室的理论却受到了同行评审的科学期刊的严格审查。尽管如此, SARS-CoV-2 表现了与自然发生的、人畜共患的病毒不一致的生物学特性。在这份报告中, 我们描述了基因组、结构、医学和文献证据, 将这些证据综合起来看, 与自然起源理论存在强烈矛盾。证据表明, SARS-CoV-2 应该是以蝙蝠冠状病毒 ZC45 和/或 ZXC21 为模板和/或骨干的实验室产品。基于这些证据, 我们进一步推测了 SARS-CoV-2 的合成途径, 证明了在实验室研制该冠状病毒是十分方便的, 可以在大约 6 个月内完成。我们的工作着重强调了对相关研究实验室进行独立调查的必要性。它还主张对最近公布的某些数据进行批判性调查研究, 尽管这些数据存在问题, 但却被用来支持和宣称 SARS-CoV-2 的自然起源理论。从公共卫生的角度来看, 这些行动是必要的, 因为了解 SARS-CoV-2 的来源和病毒如何传播给人类, 对于从根本上控制 COVID-19 大流行以及防止今后类似的大流行至关重要。



名词注释

特殊名词	原文	注释
ACE2 受体 (ACE2)	Angiotensin- Converting Enzyme 2	一种附着在肺，动脉，心脏，肾脏和肠道中细胞表面上的酶；也是 SARS 冠状病毒（SARS-CoV）的受体
人类 ACE2 受体 (hACE2)	human Angiotensin- Converting Enzyme 2	
功能增强研究	gain-of-function	通过改良对象(病毒)，并使其在特定方面增强效果的研究项目
刺突蛋白	Spike protein/Spike	冠状病毒外层的(冠状)突触，用以结合目标细胞并进一步感染目标细胞
受体结合域 (RBD)	Receptor Binding Domain	介导了病毒与靶细胞的相互作用的区域
受体结合基序 (RBM)	Receptor Binding Motif	
Furin 酶切位点	Furin-cleavage site	Furin 酶能够识别特定的氨基酸序列并在多种前体蛋白进行加工。这里专指称该特殊切未点。
武汉病毒学研究所 (WIV)	Wuhan Institute of Virology	位于中国武汉的一间 P4 等级研究院
Gibson 组装	Gibson Assembly	Gibson 组装技术是一种简单、快速、高效的 DNA 定向无缝克隆技术，可将插入片段 (PCR 产物) 定向克隆至任意载体的任意位点
聚合酶连锁反应 (PCR)	Polymerase Chain Reaction	是一项利用 DNA 双链复制的原理，在生物体外复制特定 DNA 片段的核酸合成技术



引言

COVID-19 已经引发了全球范围的大流行，其规模和严重程度都是前所未有的。尽管国际社会作出了巨大的努力，但对这一流行病的管理和控制仍然是困难和具有挑战性的。

作为一种冠状病毒，SARS-CoV-2 与其他呼吸道和/或人畜共患病毒有显著的不同：它攻击多个器官；它能够进行长时间的无症状感染；它在高危人群中具有高度的传染性及致命性；它自出现之初就对人类有良好的适应性¹；它与人类 ACE2 受体（hACE2）的结合效率很高，其亲和力大于任何其他潜在宿主^{2,3}与 ACE2 受体的亲和力。

SARS-CoV-2 的起源至今仍极具争议。一篇被广泛引用的《自然医学杂志》刊登的文章声称，SARS-CoV-2 很可能来源于自然界⁴。然而，这篇文章及其核心结论现在正受到来自世界各地科学家的质疑⁵⁻¹⁵。此外，这篇《自然医学杂志》文章的作者显示出利益冲突的迹象^{16,17}，进一步引发了人们对这篇文章可信度的担忧。

支持自然起源理论的现有科学出版物在很大程度上依赖于一个单一的证据 - 一种先前发现的名为 RaTG13 的蝙蝠冠状病毒，它与 SARS-CoV-2¹⁸ 的核苷酸序列有 96% 的一致性。然而，RaTG13 在自然界中的存在及其所报道序列的真实性正在受到广泛质疑^{6-9,19-21}。值得注意的是，科学期刊已经明确审查了任何暗示 SARS-CoV-2^{8, 22} 的非自然起源的不同意见。由于这种审查制度，质疑 SARS-CoV-2 的自然起源或 RaTG13 的实际存在的文章，尽管就科学角度而言具有很高的质量，但却只能作为预印本^{5-9,19-21} 或其他非同同行评审的文章在各种在线平台上发表^{10-13,23}。尽管如此，对这些报告的分析一再指出，RaTG13^{6,8,9,19-21} 的报告中存在严重的问题和可能的欺诈行为。因此，关于发表伪



造的科学数据以误导全世界追查 SARS-CoV-2 的来源的理论已变得非常令人信服，并与 SARS-CoV-2 是非自然起源的概念交织在一起。

与这一概念相一致的是，基因组、结构和文献证据也表明 SARS-CoV-2 并非源于自然。此外，大量的文献表明，功能增强研究早已发展到病毒基因组可以被精确设计和操纵的阶段，从而创造出独特性的新型冠状病毒。在本报告中，我们介绍了这些证据和相关分析。本报告的第一部分描述了 SARS-CoV-2 的基因组和结构特征，这些特征的存在可能与该病毒是实验室改造的产物这一理论相一致，而不是简单的病毒连续传播所能实现的。本报告的第二部分描述了 SARS-CoV-2 在实验室产生的极有可能的途径，并有着病毒基因组的证据支持其中的关键步骤得到了病毒基因组中的证据支持。重要的是，第二部分应该被看作是一个演示，说明如何利用现有的材料和有据可查的技术，在短时间内方便地在实验室里创造出 SARS-CoV-2。本报告是由经验丰富的科学家团队利用我们在病毒学、分子生物学、结构生物学、计算生物学、疫苗开发和医学方面的综合专业知识制作的。

1. SARS-CoV-2 是否进行过体外操作？

我们提出三条证据链来支持我们的论点，即 SARS-CoV-2 的历史组成部分是包含实验室改造的：

- i. SARS-CoV-2 的基因组序列与第三军医大学（中国重庆）和南京军区医学研究所（中国南京）的军事实验室发现的一种蝙蝠冠状病毒的基因组序列疑似相近。



- ii. SARS-CoV-2 的刺突蛋白内的受体结合基序 (RBM) 决定了病毒的宿主特异性，与 2003 年流行的 SARS-CoV 的受体结合基序有着十分可疑的相似性。基因组证据表明，该 RBM 已被进行了基因改造。
- iii. SARS-CoV-2 在其刺突蛋白中含有一个独特的 Furin 酶切位点，该位点已知可大大增强病毒的感染力和细胞的趋向性。然而，这个切位点在源于自然界的这类的冠状病毒中完全没有。此外，与这一额外序列相关的罕见密码子表明，这一 Furin 酶切位点极有可能不是自然进化的产物，而是通过简单的连续传代或一系列的重组技术，在合并感染的组织培养物或动物体内人工插入 SARS-CoV-2 基因组的。

1.1 基因组序列分析显示，ZC45 或与之密切相关的蝙蝠冠状病毒，应该是用于制造 SARS-CoV-2 的主干病毒。

图 1 显示了约 30,000 个核苷酸长的 SARS-CoV-2 基因组的结构。透过搜索 NCBI 序列数据库发现，在目前已知的冠状病毒中，与 SARS-CoV-2 具有最高序列同一性的两个相关蝙蝠冠状病毒分别为 ZC45 和 ZXC21（每个蝙蝠冠状病毒与 SARS-CoV-2 在核苷酸水平上有约 89% 的一致性）。图 1 描述了 SARS-CoV-2 与典型 β 冠状病毒基因组的相似性。ZXC21 与 ZC45 有 97% 的相同和非常相似的特征，但并未显示出来。请注意，RaTG13 病毒被排除在本分析之外，因为有强有力的证据表明，其序列可能是捏造的，而且该病毒并不存在于自然界^{2,6-9}。（后续报告将会很快提交，届时会总结所有可验证 RaTG13 虚假性质的最新证据）

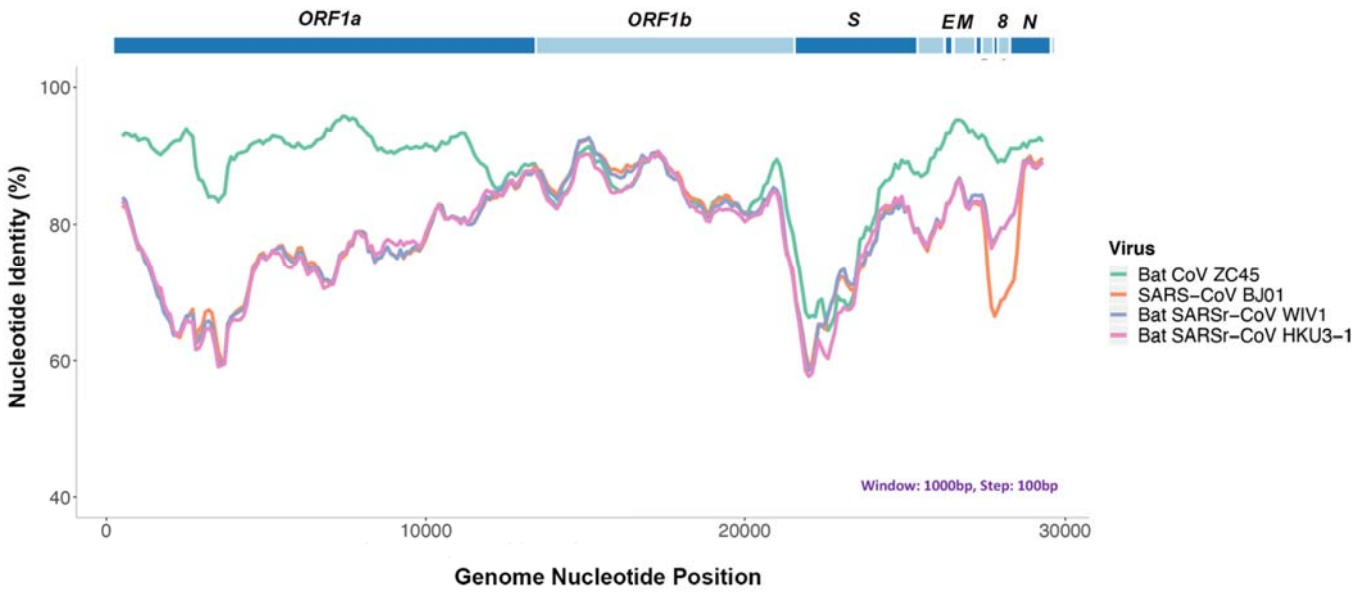


图 1：基因组测序分析显示，蝙蝠冠状病毒 ZC45 与 SARS-CoV-2 病毒最为接近。上图：系 SARS-CoV-2 的组织结构 (2019-nCoV WIV04)。下图：相似度分析系基于 2019-nCoV WIV04 的全段基因组。SARS-CoV BJ01、SARSr-CoV WIV1 蝙蝠病毒、SARSr-CoV HKU3-1 蝙蝠病毒及蝙蝠病毒 ZC45 均作为参比测序 (基因结构分析)。

当对比 SARS-CoV-2 和 ZC45/ZXC21 的氨基酸水平时，可以观察到大部分蛋白质的测序有极高的相似度。核衣壳蛋白（简称“核蛋白”或“N 蛋白”）的相似度是 94%。膜蛋白（或“M 蛋白”）相似度是 98.6%。刺突蛋白的第二部分，即 S2 蛋白的相似度是 95%。极为重要的是，Orf8 蛋白的相似度是 94.2%，E 蛋白相似度达到了 100%。

Orf8 蛋白是一种辅助蛋白，尽管近期表明 SARS-CoV-2 里的 Orf8 蛋白（可能）会通过抑制 MHC-I 类分子来减缓对宿主后天免疫（适应性免疫）系统的入侵²⁴，但 Orf8 蛋白在大部分冠状病毒中的作用仍属未知。一般情况下，Orf8 蛋白在冠状病毒中的保守性极差²⁵。测序 blast 分析显示 ZC45/ZXC21 中的 Orf8 蛋白与 SARS-CoV-2 的 Orf8 蛋白具有 94.2% 的相似度，其他任何的冠状病毒无法在这一特定的蛋白质上与 SARS-



CoV-2 具有超出 58% 以上的相似度。在一般情况下保守性很差的 Orf8 蛋白质在这里出现了极高的同源性，这是极不寻常的。

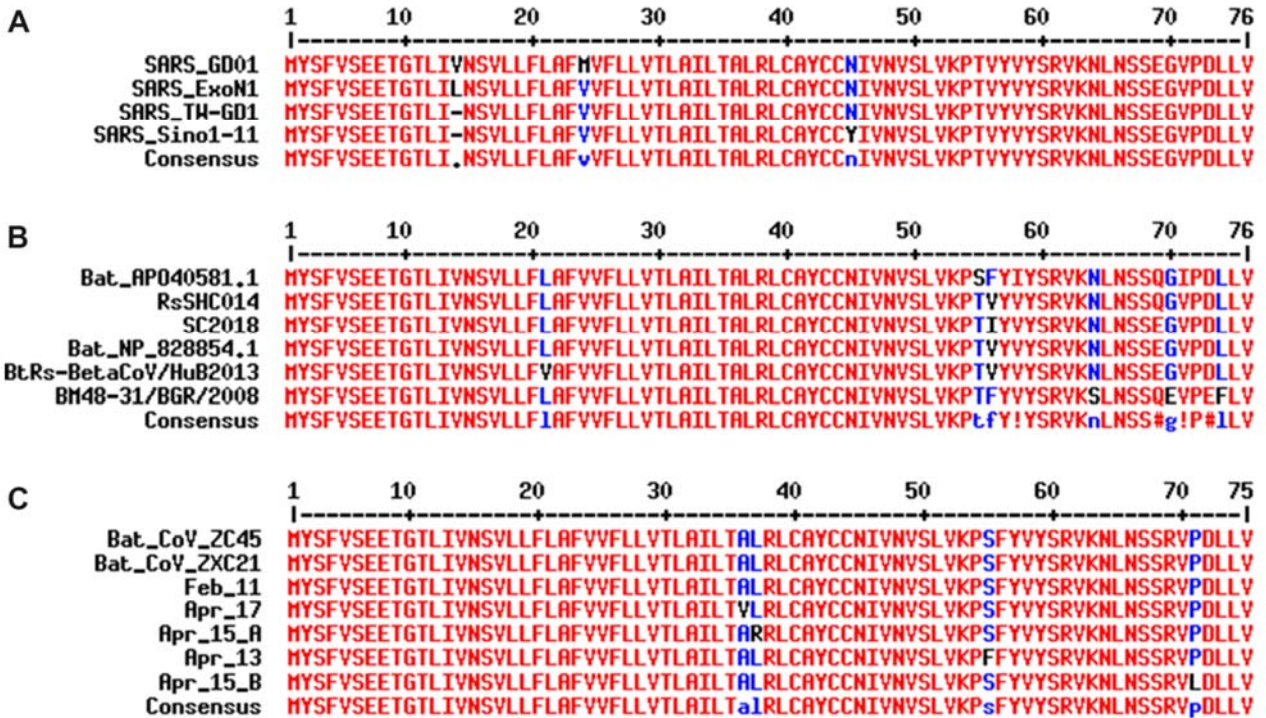


图 2: 比对来自不同 β 属冠状病毒的 E 蛋白测序, 证明了 E 蛋白氨基酸突变的允许性和趋势。A. 观察 SARS-CoV 的不同菌株, 发现了突变。GenBank 登录号: SARS_GD01: AY278489. 2, SARS_ExoN1: ACB69908. 1, SARS_TW_GD1: AY451881. 1, SARS_Sino1_11: AY485277. 1。B. 对来自具有关联性的蝙蝠冠状病毒里的 E 蛋白进行比对, 发现其在不同位置上对突变的容许度。GenBank 登录号: Bat_AP040581. 1: AP040581. 1, RsSHC014: KC881005. 1, SC2018: MK211374. 1, Bat_NP_828854. 1: NP_828854. 1, BtRsBetaCoV/HuB2013: AIA62312. 1, BM 4831/BGR/2008: YP_003858555。C. 虽然 SARS-CoV-2 的早期复制在 E 蛋白上与 ZC45 和 ZXC21 具有 100% 的相似度, 但自 2020 年 4 月起, SARS-CoV-2 的测序数据表明了已在多个位置发生了突变。病毒的登录号: Feb_11: MN997409, ZC45: MG772933. 1, ZXC21: MG772934, Apr_13: MT326139, Apr_15_A: MT263389, Apr_15_B: MT293206, Apr_17: MT350246。使用 MultAlin 网上工具进行比对 (<http://multalin.toulouse.inra.fr/multalin/>)。



冠状病毒 E 蛋白是一种结构性蛋白，其嵌入并排列于病毒体被膜内侧²⁶。E 蛋白对突变具有耐受性，这在 SARS（图 2A）和具有关联性的蝙蝠冠状病毒（图 2B）中都有证明。E 蛋白对氨基酸突变的耐受性在这次 SARS-CoV-2 大流行中得到进一步证明。自病毒在人类中爆发以来，仅经过短短两个月的传播，SARS-CoV-2 中的 E 蛋白就已经经历了多次突变。四月份获得的测序数据显示，不同菌株在四个不同位置发生了突变（图 2C）。与这一发现一致的是，BLAST 测序分析表明，除 SARS-CoV-2 以外，没有任何已知冠状病毒在 E 蛋白氨基酸序列上与 ZC45/ZXC21 具有 100% 相似度（*但那些在本次疫情爆发后所发表的可疑冠状病毒不包含在内*^{18, 27-31}）。虽然在 SARS-CoV 和某些与 SARS 有关联的蝙蝠冠状病毒之间观察到了 E 蛋白具有 100% 的相似度，但这些做对比的配对中并没有任何一对在 Orf8 蛋白上同时具有高于 83% 的相似度³²。因此，在 Orf8 蛋白上具有 94.2% 的相似度、E 蛋白 100% 的相似度、以及在 SARS-CoV-2 和 ZC45/ZXC21 之间的整体基因组/氨基酸水平相似度都是极不寻常的。综合分析，这些证据和以下假设是一致的，即：SARS-CoV-2 基因组的起源是将 ZC45/ZXC21 作为骨干及/或模板并加以基因增强功能改造而得出的。

重要的是，ZC45 和 ZXC21 是（2015 年 7 月至 2017 年 2 月）由第三军医大学军事研究实验室（中国重庆）及南京军区医学研究所（中国南京）发现、分离、并分类的蝙蝠冠状病毒。数据和相关文章发表于 2018 年^{33, 34}。显而易见的是，ZC45 和 ZXC21 作为研制 SARS-CoV-2 病毒至关重要的骨干及/或模板，存在于上述实验室及其他相关的实验室里。

已发表的 RaTG13 病毒可以进一步强化我们的论点，该病毒的公开基因组序列与 SARS-CoV-2 的基因组序列有着 96% 的相似度。虽然（该论文）称 SARS-CoV-2 来自于自然界，但 RaTG13 病毒也转移了科学界和公众对 ZC45/ZXC21 的注意力^{4, 18}。事实上，



中国的 BSL-3 实验室（上海公共卫生临床中心）在《自然》杂志上发表了一篇文章，报道了 SARS-CoV-2 和 ZC45/ZXC21 之间（而非和 RaTG13 之间）具有相互冲突的密切亲缘关系³⁵，该实验室很快被关闭以进行“整改”³⁶。人们相信该实验室的研究人员因披露了 SARS-CoV-2 和 ZC45/ZXC21 之间的关联性而受到了惩处。另一方面，已经有大量证据指向与所报道的 RaTG13 病毒序列有关的严重问题，并对这种蝙蝠病毒是否在自然界实际存在提出了质疑^{6,7,19-21}。近期的一份报道还指出，RaTG13 病毒的刺突蛋白的受体结合域（RBD）无法与两种不同类型的马蹄蝠身上的 ACE2 蛋白相结合（这两种马蹄蝠与 RaTG13 的所谓自然宿主-中华菊头蝠（又称）很接近）²，这表明 RaTG13 病毒并不能感染马蹄蝠。这一发现进一步证实了所公布的 RaTG13 测序可能是伪造的，因为由该测序中的刺突蛋白似乎并不具有所述功能。事实上，这是一个被伪造出来的病毒，用以转移人们对 ZC45/ZXC21 的注意力，这说明了 ZC45/ZXC21 在研制 SARS-CoV-2 病毒中的实际作用。

1.2 SARS-CoV-2 病毒刺突的受体结合基序不可能来自自然界，其应该是通过基因工程改造而来的

刺突蛋白是冠状病毒颗粒外部的装饰。它们在（病毒）感染的过程中起了重要作用，因为它们介导与宿主细胞受体的相互作用，从而有助于确定病毒的宿主范围和组织嗜性。刺突蛋白被分成两个部分（图 3）。前端或 N 端部分被称为 S1，它完全负责与宿主受体的结合。在 SARS-CoV 和 SARS-CoV-2 的感染中，宿主细胞受体是 hACE2。在 S1 里面，大约 70 个氨基酸的片段与 hACE2 有直接接触，因此被相应地命名为受体结合基序（RBM）（图 3C）。在 SARS-CoV 和 SARS-CoV-2 中，RBM 完全决定了与人类 ACE2 受体（hACE2）的相互介导作用。刺突蛋白的 C 端（羧基端）部分被命名为 S2。

S2 的主要功能包括维持三聚体的形成，并在 S1/S2 连接处和下游 S2 的位置的连续蛋白酶裂解时，介导膜突变并融合，使病毒进入细胞。

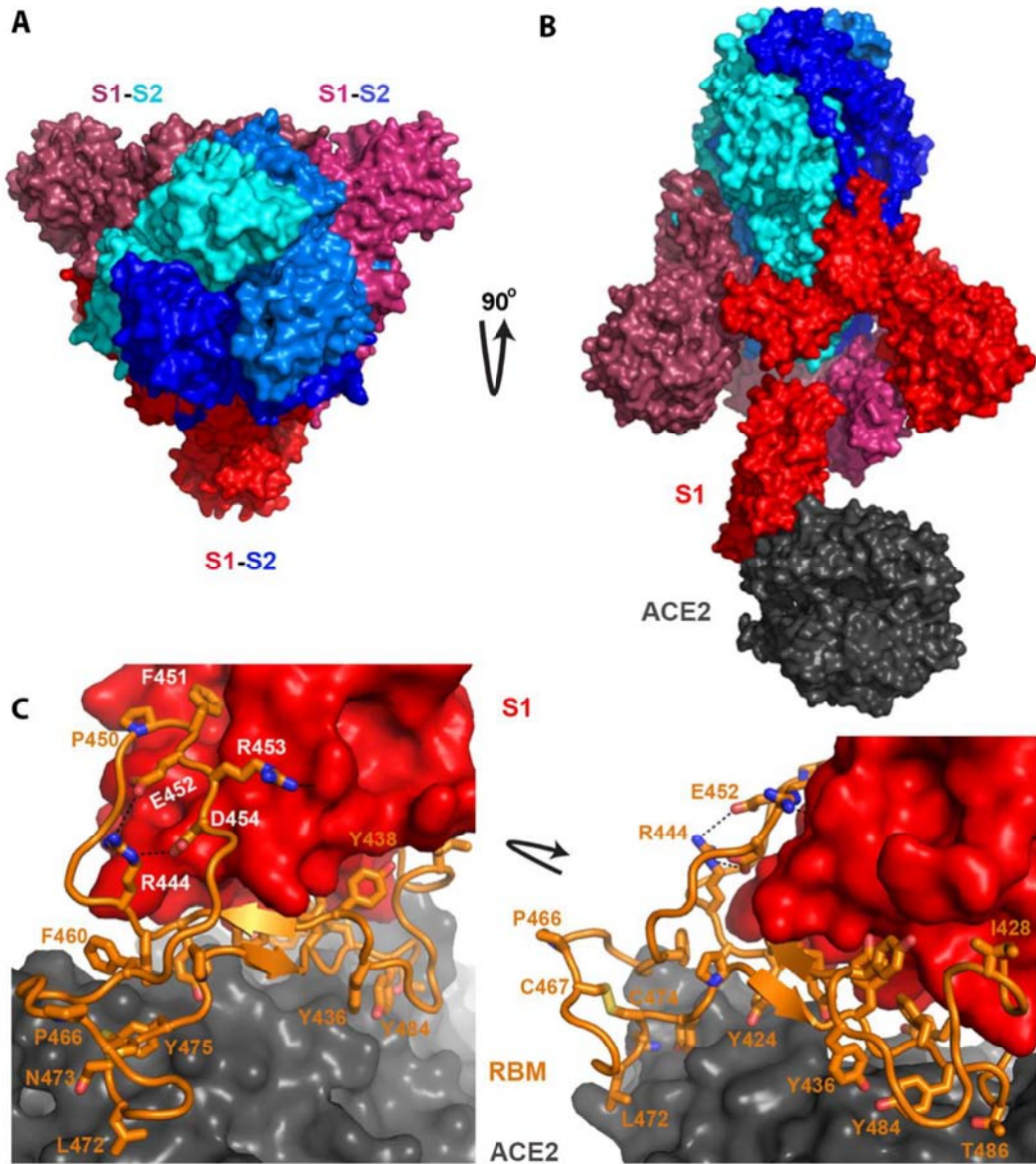


图 3: SARS 刺突蛋白结构图、及其如何与 hACE2 受体所结合。图片来源于 PDB ID: 6acj³⁷。图 A) 三个刺突蛋白，各包含 S1 部分，和 S2 部分，形成三聚体，图 B) S2 部分（蓝色）负责形成三聚体，S1 部分（红色）负责与 hACE2（深灰色）的结合，图 C) S1 和 hACE2 结合的细节。S1 的 RBM（受体结合基序），对于结合作用极为重要也是关键因素（橘黄色）。RBM（受体结合基序）的残基对于与 hACE2 结合，或蛋白折叠程棍状都是至关重要的（残基序号系基于 SARS 刺突序列号而来）。



1 10 20 30 40 50 60 70 80 90 100 110 120 130

Huhan-Hu-1 MFVFLVLLPLVSSQCVN--LITRTQLPPAYTN--SFRGYYYPKVRSSVLSHSTQDLFLPFFSNVTHFHALHVSNGTNGKRFDMNPLPFDNDGYYFAS TEKSNIRGHIFFGT LDSKTSQLLIYNNAIN

2019-nCoV_USA-A21 MFVFLVLLPLVSSQCVN--LITRTQLPPAYTN--SFRGYYYPKVRSSVLSHSTQDLFLPFFSNVTHFHALHVSNGTNGKRFDMNPLPFDNDGYYFAS TEKSNIRGHIFFGT LDSKTSQLLIYNNAIN

Bat_CoV_ZC45 MLFFLLQFALVNSQCVN--LIGRTPLNPNYTN--SSQRGYYYPDIYRSVDTLVSQGLFLPFFSNVSHYSLTTNNA--ATKRTDMNPLDFKDGGLYFARTEKSNIRGHIFFGT LNTSLSQLLIYNNAIN

Bat_CoV_ZXC21 MLFFLLQFALVNSQCVN--LIGRTPLNPNYTN--SSQRGYYYPDIYRSVDTLVSQGLFLPFFSNVSHYSLTTNNA--ATKRTDMNPLDFKDGGLYFARTEKSNIRGHIFFGT LNTSLSQLLIYNNAIN

SARS_G202 HFIFLLFLTLTSGSOLRCTIFDQVQAPNYTQHTSSHRGYYYPDEIFRSOTLYLQDLFLPFFSNVTFGHTINHT-----FDMNVPFIPKDGGLYFARTEKSNIRGHIFFGT LNTSLSQLLIYNNAIN

SARS HFIFLLFLTLTSGSOLRCTIFDQVQAPNYTQHTSSHRGYYYPDEIFRSOTLYLQDLFLPFFSNVTFGHTINHT-----FDMNVPFIPKDGGLYFARTEKSNIRGHIFFGT LNTSLSQLLIYNNAIN

Consensus .#F.f11.1.Lvsqcv#.1T.rt.1.PnYt..Ss.RGVYYPD.IZR5dL.LQdLFLPFFSNVtw.h.i.....tkrfdmP!lpFkOG!YFAATEKSN!rGHIFFGT!#ksQSLII!NNAIN

131 140 150 160 170 180 190 200 210 220 230 240 250 260

Huhan-Hu-1 VVIKVCFFQCNDFPLGVYYHKNNKSHSESEFRVSSANNCTFEYYSQPLHDLGKGNFKNLREFVFKNDIGYFKIYSKHTPIHLVRDLPQGSALPELVDLPVIGINIRFQTLALHRSYLTGDS

2019-nCoV_USA-A21 VVIKVCFFQCNDFPLGVYYHKNNKSHSESEFRVSSANNCTFEYYSQPLHDLGKGNFKNLREFVFKNDIGYFKIYSKHTPIHLVRDLPQGSALPELVDLPVIGINIRFQTLALHRSYLTGDS

Bat_CoV_ZC45 VVIKVCFFQCNDFPLGVYYHKNNKSHSESEFRVSSANNCTFEYYSQPLHDLGKGNFKNLREFVFKNDIGYFKIYSKHTPIHLVRDLPQGSALPELVDLPVIGINIRFQTLALHRSYLTGDS

Bat_CoV_ZXC21 VVIKVCFFQCNDFPLGVYYHKNNKSHSESEFRVSSANNCTFEYYSQPLHDLGKGNFKNLREFVFKNDIGYFKIYSKHTPIHLVRDLPQGSALPELVDLPVIGINIRFQTLALHRSYLTGDS

SARS VVIKVCFFQCNDFPLGVYYHKNNKSHSESEFRVSSANNCTFEYYSQPLHDLGKGNFKNLREFVFKNDIGYFKIYSKHTPIHLVRDLPQGSALPELVDLPVIGINIRFQTLALHRSYLTGDS

SARS VVIKVCFFQCNDFPLGVYYHKNNKSHSESEFRVSSANNCTFEYYSQPLHDLGKGNFKNLREFVFKNDIGYFKIYSKHTPIHLVRDLPQGSALPELVDLPVIGINIRFQTLALHRSYLTGDS

Consensus VVIKVCFFQCNDFPLGVYYHKNNKSHSESEFRVSSANNCTFEYYSQPLHDLGKGNFKNLREFVFKNDIGYFKIYSKHTPIHLVRDLPQGSALPELVDLPVIGINIRFQTLALHRSYLTGDS

261 270 280 290 300 310 320 330 340 350 360 370 380 390

Huhan-Hu-1 SGTAGARAYYVGLQPRITFLKYNENGTIDRVDCLDPLSETKCTLKSFTEKGIYQISNFRVQPTESIVRFPNITNLCPGEVFNATRFASVYBHRKRISNCVADYSVYNSASFSTFKCYGVSPT

2019-nCoV_USA-A21 SGTAGARAYYVGLQPRITFLKYNENGTIDRVDCLDPLSETKCTLKSFTEKGIYQISNFRVQPTESIVRFPNITNLCPGEVFNATRFASVYBHRKRISNCVADYSVYNSASFSTFKCYGVSPT

Bat_CoV_ZC45 NGHTAFSARAYVGLQPRITFLKYNENGTIDRVDCLDPLSETKCTLKSFTEKGIYQISNFRVQPTESIVRFPNITNLCPGEVFNATRFASVYBHRKRISNCVADYSVYNSASFSTFKCYGVSPT

Bat_CoV_ZXC21 NGHTAFSARAYVGLQPRITFLKYNENGTIDRVDCLDPLSETKCTLKSFTEKGIYQISNFRVQPTESIVRFPNITNLCPGEVFNATRFASVYBHRKRISNCVADYSVYNSASFSTFKCYGVSPT

SARS NGHTAFSARAYVGLQPRITFLKYNENGTIDRVDCLDPLSETKCTLKSFTEKGIYQISNFRVQPTESIVRFPNITNLCPGEVFNATRFASVYBHRKRISNCVADYSVYNSASFSTFKCYGVSPT

SARS NGHTAFSARAYVGLQPRITFLKYNENGTIDRVDCLDPLSETKCTLKSFTEKGIYQISNFRVQPTESIVRFPNITNLCPGEVFNATRFASVYBHRKRISNCVADYSVYNSASFSTFKCYGVSPT

Consensus .glt.a.aRAYZVGLkP.rFLKYENGTIDRVDCLDPLSETKCTLKSFTEKGIYQISNFRVQPTESIVRFPNITNLCPGEVFNATRFASVYBHRKRISNCVADYSVYNSASFSTFKCYGVSPT

391 400 410 420 430 440 450 460 470 480 490 500 510 520

Huhan-Hu-1 KLNDLCLTNYRDSFVIRGDEVRQIAPGQGTGADYNYKLPDQFGCVIIRHNSNLDKSYGNGNYLRLFRKSNLKPFERDITSTEIYQAGSTPCNGVEGFCYFPLQSYGFQTLNMGVYQYRIVVYLSF

2019-nCoV_USA-A21 KLNDLCLTNYRDSFVIRGDEVRQIAPGQGTGADYNYKLPDQFGCVIIRHNSNLDKSYGNGNYLRLFRKSNLKPFERDITSTEIYQAGSTPCNGVEGFCYFPLQSYGFQTLNMGVYQYRIVVYLSF

Bat_CoV_ZC45 KLNDLCLTNYRDSFVIRGDEVRQIAPGQGTGADYNYKLPDQFGCVIIRHNSNLDKSYGNGNYLRLFRKSNLKPFERDITSTEIYQAGSTPCNGVEGFCYFPLQSYGFQTLNMGVYQYRIVVYLSF

Bat_CoV_ZXC21 KLNDLCLTNYRDSFVIRGDEVRQIAPGQGTGADYNYKLPDQFGCVIIRHNSNLDKSYGNGNYLRLFRKSNLKPFERDITSTEIYQAGSTPCNGVEGFCYFPLQSYGFQTLNMGVYQYRIVVYLSF

SARS KLNDLCLTNYRDSFVIRGDEVRQIAPGQGTGADYNYKLPDQFGCVIIRHNSNLDKSYGNGNYLRLFRKSNLKPFERDITSTEIYQAGSTPCNGVEGFCYFPLQSYGFQTLNMGVYQYRIVVYLSF

SARS KLNDLCLTNYRDSFVIRGDEVRQIAPGQGTGADYNYKLPDQFGCVIIRHNSNLDKSYGNGNYLRLFRKSNLKPFERDITSTEIYQAGSTPCNGVEGFCYFPLQSYGFQTLNMGVYQYRIVVYLSF

Consensus KLNDLCLTNYRDSFVIRGDEVRQIAPGQGTGADYNYKLPDQFGCVIIRHNSNLDKSYGNGNYLRLFRKSNLKPFERDITSTEIYQAGSTPCNGVEGFCYFPLQSYGFQTLNMGVYQYRIVVYLSF

521 530 540 550 560 570 580 590 600 610 620 630 640 650

Huhan-Hu-1 ELLNAPATVCGPKSTNLVKNQCVNFMNGLTGTGLVTESSKFKLPFOQGRDIAOTDVARDPQTELELDITPCSFGGVSVIIPGTNSQVAVLYQVQNCITVPAIHADQLTPHARYSTGNNVFT

2019-nCoV_USA-A21 ELLNAPATVCGPKSTNLVKNQCVNFMNGLTGTGLVTESSKFKLPFOQGRDIAOTDVARDPQTELELDITPCSFGGVSVIIPGTNSQVAVLYQVQNCITVPAIHADQLTPHARYSTGNNVFT

Bat_CoV_ZC45 ELLNAPATVCGPKSTNLVKNQCVNFMNGLTGTGLVTESSKFKLPFOQGRDIAOTDVARDPQTELELDITPCSFGGVSVIIPGTNSQVAVLYQVQNCITVPAIHADQLTPHARYSTGNNVFT

Bat_CoV_ZXC21 ELLNAPATVCGPKSTNLVKNQCVNFMNGLTGTGLVTESSKFKLPFOQGRDIAOTDVARDPQTELELDITPCSFGGVSVIIPGTNSQVAVLYQVQNCITVPAIHADQLTPHARYSTGNNVFT

SARS ELLNAPATVCGPKSTNLVKNQCVNFMNGLTGTGLVTESSKFKLPFOQGRDIAOTDVARDPQTELELDITPCSFGGVSVIIPGTNSQVAVLYQVQNCITVPAIHADQLTPHARYSTGNNVFT

SARS ELLNAPATVCGPKSTNLVKNQCVNFMNGLTGTGLVTESSKFKLPFOQGRDIAOTDVARDPQTELELDITPCSFGGVSVIIPGTNSQVAVLYQVQNCITVPAIHADQLTPHARYSTGNNVFT

Consensus ELLNAPATVCGPKSTNLVKNQCVNFMNGLTGTGLVTESSKFKLPFOQGRDIAOTDVARDPQTELELDITPCSFGGVSVIIPGTNSQVAVLYQVQNCITVPAIHADQLTPHARYSTGNNVFT

651 660 670 680 690 700 710 720 730 740 750 760 770 780

Huhan-Hu-1 RAGCIGAETHVNSYECDDIPIGAGICASYITQINSRRARSVASQSIAYTHSLGRENVSAYSNNSIAIPTNFISVITTEILPVSHAKTSVDOCTHYICGDSIECSNLLQYGSFCTQLNRRALSGTAEQD

2019-nCoV_USA-A21 RAGCIGAETHVNSYECDDIPIGAGICASYITQINSRRARSVASQSIAYTHSLGRENVSAYSNNSIAIPTNFISVITTEILPVSHAKTSVDOCTHYICGDSIECSNLLQYGSFCTQLNRRALSGTAEQD

Bat_CoV_ZC45 RAGCIGAETHVNSYECDDIPIGAGICASYITASI-----LRSTQKRIAYTHSLGRENVSAYSNNSIAIPTNFISVITTEILPVSHAKTSVDOCTHYICGDSIECSNLLQYGSFCTQLNRRALSGTAEQD

Bat_CoV_ZXC21 RAGCIGAETHVNSYECDDIPIGAGICASYITASI-----LRSTQKRIAYTHSLGRENVSAYSNNSIAIPTNFISVITTEILPVSHAKTSVDOCTHYICGDSIECSNLLQYGSFCTQLNRRALSGTAEQD

SARS RAGCIGAETHVNSYECDDIPIGAGICASYITVSL-----LRSTQKSIAYTHSLGRENVSAYSNNSIAIPTNFISVITTEILPVSHAKTSVDOCTHYICGDSIECSNLLQYGSFCTQLNRRALSGTAEQD

SARS RAGCIGAETHVNSYECDDIPIGAGICASYITVSL-----LRSTQKSIAYTHSLGRENVSAYSNNSIAIPTNFISVITTEILPVSHAKTSVDOCTHYICGDSIECSNLLQYGSFCTQLNRRALSGTAEQD

Consensus RAGCIGAETHVNSYECDDIPIGAGICASYIT.s.....LRSt.qkSIAYTHSLGRENVSAYSNNSIAIPTNFISITTE!#PVSHAKTSVDOCTHYICGDSIECSNLLQYGSFCTQLNRRALSGTAEQD

781 790 800 810 820 830 840 850 860 870 880 890 900 910

Huhan-Hu-1 KNTQEVFAQYKQIYKTPPIKDFGGFNFSQILPDKPKRSFIEDLLFNKVTLDHAGFIKQYGCGLDGINARLDICRAQFNGLTVLPLLDDEHIAQTSLLAGITISGHTFAGARALQIPFAHQHAYR

2019-nCoV_USA-A21 KNTQEVFAQYKQIYKTPPIKDFGGFNFSQILPDKPKRSFIEDLLFNKVTLDHAGFIKQYGCGLDGINARLDICRAQFNGLTVLPLLDDEHIAQTSLLAGITISGHTFAGARALQIPFAHQHAYR

Bat_CoV_ZC45 KNTQEVFAQYKQIYKTPPIKDFGGFNFSQILPDKPKRSFIEDLLFNKVTLDHAGFIKQYGCGLDGINARLDICRAQFNGLTVLPLLDDEHIAQTSLLAGITISGHTFAGARALQIPFAHQHAYR

Bat_CoV_ZXC21 KNTQEVFAQYKQIYKTPPIKDFGGFNFSQILPDKPKRSFIEDLLFNKVTLDHAGFIKQYGCGLDGINARLDICRAQFNGLTVLPLLDDEHIAQTSLLAGITISGHTFAGARALQIPFAHQHAYR

SARS RNTREYFAQYKQIYKTPPIKDFGGFNFSQILPDKPKRSFIEDLLFNKVTLDHAGFIKQYGCGLDGINARLDICRAQFNGLTVLPLLDDEHIAQTSLLAGITISGHTFAGARALQIPFAHQHAYR

SARS RNTREYFAQYKQIYKTPPIKDFGGFNFSQILPDKPKRSFIEDLLFNKVTLDHAGFIKQYGCGLDGINARLDICRAQFNGLTVLPLLDDEHIAQTSLLAGITISGHTFAGARALQIPFAHQHAYR

Consensus KNTQEVFAQYKQIYKTPPIKDFGGFNFSQILPDKPKRSFIEDLLFNKVTLDHAGFIKQYGCGLDGINARLDICRAQFNGLTVLPLLDDEHIAQTSLLAGITISGHTFAGARALQIPFAHQHAYR

911 920 930 940 950 960 970 980 990 1000 1010 1020 1030 1040

Huhan-Hu-1 FNGIGVTQNYLYENQKLIANQFNRAIGKIQDLSSTASRALGKLQDQVYVQNAQALNTLVKQLSSNFGRISVLDNIDLARLDKYEVEVQIDRLITGRLQSLQTYVYVQQLIRRAEIRASALRAATKNSCEVLG

2019-nCoV_USA-A21 FNGIGVTQNYLYENQKLIANQFNRAIGKIQDLSSTASRALGKLQDQVYVQNAQALNTLVKQLSSNFGRISVLDNIDLARLDKYEVEVQIDRLITGRLQSLQTYVYVQQLIRRAEIRASALRAATKNSCEVLG

Bat_CoV_ZC45 FNGIGVTQNYLYENQKLIANQFNRAIGKIQDLSSTASRALGKLQDQVYVQNAQALNTLVKQLSSNFGRISVLDNIDLARLDKYEVEVQIDRLITGRLQSLQTYVYVQQLIRRAEIRASALRAATKNSCEVLG

Bat_CoV_ZXC21 FNGIGVTQNYLYENQKLIANQFNRAIGKIQDLSSTASRALGKLQDQVYVQNAQALNTLVKQLSSNFGRISVLDNIDLARLDKYEVEVQIDRLITGRLQSLQTYVYVQQLIRRAEIRASALRAATKNSCEVLG

SARS FNGIGVTQNYLYENQKLIANQFNRAISQIQDLSSTASRALGKLQDQVYVQNAQALNTLVKQLSSNFGRISVLDNIDLARLDKYEVEVQIDRLITGRLQSLQTYVYVQQLIRRAEIRASALRAATKNSCEVLG

SARS FNGIGVTQNYLYENQKLIANQFNRAISQIQDLSSTASRALGKLQDQVYVQNAQALNTLVKQLSSNFGRISVLDNIDLARLDKYEVEVQIDRLITGRLQSLQTYVYVQQLIRRAEIRASALRAATKNSCEVLG

Consensus FNGIGVTQNYLYENQKLIANQFNRAIGKIQDLSSTASRALGKLQDQVYVQNAQALNTLVKQLSSNFGRISVLDNIDLARLDKYEVEVQIDRLITGRLQSLQTYVYVQQLIRRAEIRASALRAATKNSCEVLG

1041 1050 1060 1070 1080 1090 1100 1110 1120 1130 1140 1150 1160 1170

Huhan-Hu-1 QSKRYDFCGKYHLSHFQSPAGHGVFLHVTYVPAQEKHFTAPRACHEGKHHPFREGVYVNSGTHFVYVQRFVPEQIITONTFVSGNCDVYVIGIINNTVYDLPQELDSFKEELDKYFKNHTSPVD

2019-nCoV_USA-A21 QSKRYDFCGKYHLSHFQSPAGHGVFLHVTYVPAQEKHFTAPRACHEGKHHPFREGVYVNSGTHFVYVQRFVPEQIITONTFVSGNCDVYVIGIINNTVYDLPQELDSFKEELDKYFKNHTSPVD

Bat_CoV_ZC45 QSKRYDFCGKYHLSHFQSPAGHGVFLHVTYVPAQEKHFTAPRACHEGKHHPFREGVYVNSGTHFVYVQRFVPEQIITONTFVSGNCDVYVIGIINNTVYDLPQELDSFKEELDKYFKNHTSPVD

Bat_CoV_ZXC21 QSKRYDFCGKYHLSHFQSPAGHGVFLHVTYVPAQEKHFTAPRACHEGKHHPFREGVYVNSGTHFVYVQRFVPEQIITONTFVSGNCDVYVIGIINNTVYDLPQELDSFKEELDKYFKNHTSPVD

SARS QSKRYDFCGKYHLSHFQSPAGHGVFLHVTYVPAQEKHFTAPRACHEGKHHPFREGVYVNSGTHFVYVQRFVPEQIITONTFVSGNCDVYVIGIINNTVYDLPQELDSFKEELDKYFKNHTSPVD

SARS QSKRYDFCGKYHLSHFQSPAGHGVFLHVTYVPAQEKHFTAPRACHEGKHHPFREGVYVNSGTHFVYVQRFVPEQIITONTFVSGNCDVYVIGIINNTVYDLPQELDSFKEELDKYFKNHTSPVD

Consensus QSKRYDFCGKYHLSHFQSPAGHGVFLHVTYVPAQEKHFTAPRACHEGKHHPFREGVYVNSGTHFVYVQRFVPEQIITONTFVSGNCDVYVIGIINNTVYDLPQELDSFKEELDKYFKNHTSPVD

1171 1180 1190 1200 1210 1220 1230 1240 1250 1260 1270 1278

Huhan-Hu-1 LGDISGINSVYNIQKEDRLNEVAKNLESIDLQELGKYEYIKAPHYIHLGFIAGLIRIYVYTHLCCHTSCCCLKGCSCGSCCKFDEDDSEPLVKGKVLHYT

2019-nCoV_USA-A21 LGDISGINSVYNIQKEDRLNEVAKNLESIDLQELGKYEYIKAPHYIHLGFIAGLIRIYVYTHLCCHTSCCCLKGCSCGSCCKFDEDDSEPLVKGKVLHYT

Bat_CoV_ZC45 LGDISGINSVYNIQKEDRLNEVAKNLESIDLQELGKYEYIKAPHYIHLGFIAGLIRIYVYTHLCCHTSCCCLKGCSCGSCCKFDEDDSEPLVKGKVLHYT

Bat_CoV_ZXC21 LGDISGINSVYNIQKEDRLNEVAKNLESIDLQELGKYEYIKAPHYIHLGFIAGLIRIYVYTHLCCHTSCCCLKGCSCGSCCKFDEDDSEPLVKGKVLHYT

SARS LGDISGINSVYNIQKEDRLNEVAKNLESIDLQELGKYEYIKAPHYIHLGFIAGLIRIYVYTHLCCHTSCCCLKGCSCGSCCKFDEDDSEPLVKGKVLHYT

SARS LGDISGINSVYNIQKEDRLNEVAKNLESIDLQELGKYEYIKAPHYIHLGFIAGLIRIYVYTHLCCHTSCCCLKGCSCGSCCKFDEDDSEPLVKGKVLHYT

Consensus LGDISGINSVYNIQKEDRLNEVAKNLESIDLQELGKYEYIKAPHYIHLGFIAGLIRIYVYTHLCCHTSCCCLKGCSCGSCCKFDEDDSEPLVKGKVLHYT



图 4：来自不同冠状病毒刺突蛋白的测序排列。进行对比的病毒包括：SARS-CoV-2（Wuhan-Hu-1: NC_045512, 2019-nCoV_USA-AZ1: MN997409），蝙蝠冠状病毒（Bat_CoV_ZC45: MG772933, Bat_CoV_ZXC21: MG772934），及 SARS 冠状病毒（SARS_GZ02: AY390556, SARS: NC_004718.3）。两条橘黄色线条标示的部分是受体结合基序（RBM），其对于和 hACE2 受体的结合至关重要。图顶部用红色线条标示的部分是基础残基。两条绿色线条标示的部分是仅存在于 SARS-CoV-2 病毒里的 Furin 酶切位点，在其他 B 群 β 属冠状病毒中都不存在。

类似在与其他病毒蛋白中观察到的，SARS-CoV-2 的 S2 与 ZC45/ZXC21 的 S2 具有高度的序列同一性（95%）。与之形成鲜明的对比的是，SARS-CoV-2 和 ZC45/ZXC21 两者之间的 S1 蛋白，是表明病毒可以感染哪一个宿主（人或蝙蝠），比氨基酸序列同一性仅为 69% 保守性低得多。

图 4 显示了来自六种 β 冠状病毒的刺突蛋白的序列比对。两个病毒是从目前的大流行中分离出来的（Wuhan-Hu-1, 2019-nCoV_USA-AZ1）；两个是可疑的病毒模板（Bat_CoV_ZC45, Bat_CoV_ZXC21）；两个是 SARS 冠状病毒（SARS_GZ02, SARS）。RBM 在两条橙色线之间被标注出来。显然，尽管整个基因组具有很高的序列同一性，但 SARS-CoV-2 的 RBM 与 ZC45 和 ZXC21 的 RBM 明显不同。有趣的是，SARS-CoV-2 的 RBM 在很大程度上类似于 SARS 刺突的 RBM。尽管这不是精确的“复制粘贴”，但仔细检查刺突-hACE2 结构^{37,38}发现，所有对 hACE2 结合或蛋白质折叠必不可少的残基（图 3C 中的橙色棒和图 4 中用红色短线标出的）被“保留”。这些必需残基中的大多数都被精确保留，包括那些参与二硫键形成的残基（C467, C474）和静电相互作用（R444, E452, R453, D454），这对于 RBM 的结构完整性至关重要（图 3C 和 4）。基本残基组中的几处变化几乎完全是疏水性的“取代”（I428→L, L443→F, F460→Y, L472→F, Y484→Q），这不应影响蛋白质折叠或 hACE2 相互作用。与此同时，大部分非必需氨基酸残基均已“突变”（图 4，RBM 残留物未用短红线标记）。从这个单独的序列分析来判断，我们很早就确信 SARS-CoV-2 刺突蛋白不仅会与



hACE2 结合，而且其结合也恰好类似于原始 SARS 刺突蛋白与 hACE2²³ 之间的结合。最近的架构工作证实了我们的预测³⁹。

如下所述，SARS-CoV-2 RBM 类似于 SARS-CoV RBM 的方式以及 SARS-CoV-2 与 ZC45/ZXC21 之间的整体序列保护模式是非常不寻常的。总的来说，这表明 SARS-CoV-2 基因组的某些部分并非源自天然准物种病毒颗粒的进化。

如果 SARS-CoV-2 确实来自自然进化，则其 RBM 只能通过以下两种可能途径之一获得：1) 久远的重组事件并产生趋同演化；或 2) 相当近期发生的自然重组事件。

在第一种情况下，SARS-CoV-2 的祖先，一个类似 ZC45/ZXC21 的蝙蝠冠状病毒将重组并“交换”其 RBM 和带有相对“完整” RBM 的冠状病毒（参考 SARS）。这种重组将产生一种新型的 ZC45/ZXC21 冠状病毒，其 RBM 中的所有缺口均被“填充”（图 4）。随后，该病毒将不得不广泛适应其新宿主，其中 ACE2 蛋白与 hACE2 高度同源。必须发生基因组上的随机突变，才能最终将 RBM 塑造成当前形式-以高度智能的方式使其类似于 SARS-CoV RBM。但是，这种趋同的进化过程也将导致基因组其他部分中大量突变的积累，从而使整体序列同一性相对较低。SARS-CoV-2 与 ZC45/ZXC21 在各种蛋白质上的高度序列同一性（94-100% 相同性）不支持这种情况，因此，清楚地表明带有此类 RBM 的 SARS-CoV-2 不能是来自类 ZC45/ZXC21 蝙蝠冠状病毒，并通过这种趋同的进化途径。

在第二种情况下，类 ZC45/ZXC21 冠状病毒必须最近重组并和另一种冠状病毒替换其 RBM，该冠状病毒已成功适应与 hACE2 高度同源的动物 ACE2 的结合。此类事件的可能性部分取决于自然重组的一般要求：1) 两种不同的病毒具有明显的序列相似性；2) 它们必须共同感染并存在于同一动物的同一细胞中；3) 重组病毒不会被宿主清



除或使宿主灭绝； 4) 重组病毒最终将必须在宿主物种内变得稳定和可在同物种中传播。

关于最近的重组情况，动物宿主不能是蝙蝠，因为蝙蝠中的 ACE2 蛋白与 hACE2 同源性不高，因此如 SARS-CoV-2 所示，这种适应将无法产生 RBM 序列。该动物宿主也不可能是人类，因为类似 ZC45/ZXC21 的冠状病毒无法感染人类。此外，没有证据表明在 2019 年末之前在人群中传播了任何 SARS-CoV-2 或类 SARS-CoV-2 病毒。有趣的是，根据近期的一项生物信息学研究，自疫情爆发以来，SARS-CoV-2 已非常适应生存于人类¹。

自然进化的可能性只有另一种可能性，那就是类 ZC45/ZXC21 病毒和含有类 SARS 的 RBM 的冠状病毒可以在 ACE2 蛋白与 hACE2 同源的中间宿主中重组。几个实验室报告说，一些从马来西亚走私到中国的 Sunda 穿山甲带有冠状病毒，其受体结合域 (RBD) 与 SARS-CoV-2^{27-29, 31} 几乎相同。然后他们继续暗示穿山甲可能是 SARS-CoV-2^{27-29, 31} 的中间宿主。但是，最近的独立报告发现该数据存在重大缺陷⁴⁰⁻⁴²。此外，与这些报道相反^{27-29, 31}，在 2009 年至 2019⁴³ 年间，这十年里在马来西亚和沙巴收集的 Sunda 穿山甲样品中未检测到冠状病毒。最近的一项研究还表明，SARS-CoV-2 与报道的穿山甲冠状病毒之间共享的 RBD 与 hACE2 的结合强度是与穿山甲 ACE2² 的十倍，进一步排除了穿山甲作为中间宿主的可能性。最后，一项计算机研究表明，回应了穿山甲不可能是中间宿主，但同时也表明，在研究中检测的动物 ACE2 蛋白都没有比 hACE2³ 表现出对 SARS-CoV-2 刺突蛋白更有利的结合潜力。这项最后的研究几乎免除了所有动物作为中间宿主的可疑角色³，这与观察到 SARS-CoV-2 从爆发开始就很好地适应了人类的观察结果一致¹。这是重要的，因为这些发现共同表明，SARS-CoV-2 似乎不存在任何中间宿主，这至少解除了在中间宿主中发生重组事件的可能性。



即使我们忽略了上述证据，即没有适当的宿主可以进行重组，而是假设确实存在这种宿主，但这种重组事件在自然界是不可能发生。

如上所述，如果自然重组事件导致 SARS CoV-2 的出现，那么类 ZC45/ZXC21 病毒和含有类 SARS 的 RBM 冠状病毒将不得不通过交换 S1/RBM 在同一细胞中重组，这是一种罕见的重组形式。此外，由于 SARS 在人类历史上仅发生过一次，因此以这种智能的方式生产类似于 SARS 的病毒在自然界中至少是同样罕见的 - RBM 与 SARS RBM 的区别仅在于少数非必要条件切位点（图 4）。这种独特的类 SARS 冠状病毒与 ZC45/ZXC21 相近的古老病毒驻留在同一细胞中并且两种病毒以“RBM 交换”方式重组的可能性非常低。重要的是，这和下文第 1.3 节所述的其他重组事件（自然界中甚至更不可能发生）都必须发生，才能产生 SARS-CoV-2 中所见的刺突。

尽管以上证据和分析似乎不赞成 SARS-CoV-2 的 RBM 的自然起源，但大量文献表明，透过专门研究冠状病毒的刺突蛋白的增强功能研究，经由多次反复，最终产生了非人类来源冠状病毒能够成功感染人类的案例⁴⁴⁻⁴⁷。

记录还显示，例如武汉病毒学研究所（WIV）已经与美国研究人员合作完成了此类研究⁴⁵，而且还单独开展了研究⁴⁷。此外，WIV 从事了数十年的冠状病毒监测研究，因此拥有世界上最大的冠状病毒收藏库。显然地，WIV 和其他相关实验室开展和成功进行这种 Spike(刺突)/RMB(受体结合基序)工程和增强功能，获得研究的技术障碍并不存在。



这种 EcoRI 和 BstEII 位点在其他 β 冠状病毒中并不存在，这更加证明了它们并不是自然的，是为了方便操纵关键的 RBM 而专门被引入 SARS-CoV-2 的刺突基因中的。虽然 ZC45 的刺突中也没有这两个位点（图 5B），但如本报告第二部分中所述，这两个位点可以很轻易的被引入 ZC45 中。

值得一提的是，在这里把 EcoRI 位点引入后会将相对应的氨基酸从 *WNT* 改变到 *WVS*（图 5AB）。据我们所知，所有的 SARS 和类 SARS 蝙蝠冠状病毒都只在这个位置携带了一个 *T* (threonine, 苏氨酸) 残基。而 SARS-CoV-2 在这方面是唯一例外，它的这个 *T* 已经突变为一个 *S* (serine, 丝氨酸)，因此这完全解除了在病毒大爆发后公布的对 RaTG13 和穿山甲冠状病毒的怀疑⁴⁸。

一旦限制性位点被成功地引入，受体结合基序 (RBM) 片段可以使用常规限制性酶消化和结扎很轻易地被替换。虽然其他的克隆技术也许不会留下任何遗传操作的痕迹 (Gibson 组装就是一个例子)，但是上述这种老旧的 (改造) 方式会被采用是因为这种改造为替换关键的 RBM 提供很大的便利。

鉴于 RBM 能够完全决定与 hACE2 的结合，而且 SARS 的 RBM-hACE2 的结合已在高分辨率的机构 (图 3)^{37,38} 中完全地表现出来了，这种仅仅用 RBM 替换的方式不会比完整的刺突替换方式风险大。事实上，这种 RBM 替换技术的可行性已被证明^{39,47}。在 2008 年，石正丽博士的科研小组已经在将一个限制性位点引入到一个密码子优化的刺突基因 (图 5C) 后⁴⁷，把一种 SARS 的 RBM 替换进了几种类 SARS 蝙蝠冠状病毒的刺突蛋白中。随后，他们验证了所产生的嵌合刺突蛋白与 hACE2 的结合。此外，在最近的发表的论文中，SARS-CoV-2 的受体结合基序 (RBM) 被替换到了 SARS-CoV 的受体结合域 (RBD)，形成了一个完全能和 hACE2 结合的嵌合受体结合区域 (chimeric RBD) (图 5C)³⁹。令人吃惊的是，在之前描述的两种情况下，被操纵了的 RBM 片段都几乎



完全类似于由 EcoRI 和 BstEII 位点位置定义的 RBM (图 5C)。虽然两篇已发表的论文^{39,47}中都缺少克隆细节,可以想象的是,实际的限制性位点可能会根据接收 RBM 插入的刺突基因和在重点区域引入独特的限制性位点(单个或多个)的便利程度而有所不同。值得一提的是,最新发表的论文³⁹的通讯作者 Fang Li 博士,自 2010 以来⁴⁹⁻⁵³,一直是石正丽博士的积极合作者,而 Fang Li 博士是世界上第一个从结构上阐明 SARS-CoV 的 RBD 和 hACE2³⁸之间的结合机理的人,同时也是对刺突-ACE2 间的相互反应的结构理解上的顶级专家^{38,39,53-56}。在 SARS-CoV-2 的 RBM 的两端分别发现 EcoRI 和 BstEII 限制性位点这个惊人发现,和石博士和她的长期合作者(Fang Li 博士)分别在同一个 RBM 位置用限制性酶消化的方法进行了替换的事实不太可能只是一个巧合。**相反的,这恰恰是证明 SARS-CoV-2 的受体结合基序(RBM)/或刺突蛋白(刺突)是基因工程改造的产物的证据。**

虽然简单复制 SARS 的受体结合基序(RBM)一模一样的序列可能很方便,但这样做人为设计和操纵的迹象就太明显了。更具欺骗性的方式是在保留那些在结合中的关键基因的同时改变一些非必要的残基。而高分辨率结构(图 3)^{37,38}可以很好的引导这种设计。如此一来,即便 RBM 的整体排序看上去会与 SARS 的 RBM 很不同,但其与 hACE2 结合的能力却能得到很好的保存。我们认为,所有的关键残基(残基在图 4 中用红色棍状标出,与在图 3C 中显示的棍状残基相同)应该被“保留”。如前所述,有些残基应该被直接保留,而有些应该被具有类似性质的残基取代,这也不会破坏与 hACE2 的结合,这甚至还能进一步加强结合的牢固性。重要的是,这里对非必要位点的改变是可能是刻意而为,目的是为了让他看上去不像 SARS 的受体结合基序(SARS RBM)的翻版。



1.3 SARS-CoV-2 的刺突蛋白中存在一个不寻常的 Furin 酶切位点，而这很可能与该病毒的毒性增强有关

SARS-CoV-2 的刺突蛋白中的另一个独特的基因团是一个位于 S1/S2 交界处的多元 Furin 酶切位点（图 4，两条绿线之间的片段）。这样的位点可以被 Furin 蛋白酶识别和切割。在 B 系 β 属冠状病毒中，除了 SARS-CoV-2 外，没有病毒在 S1/S2 的交界处含有 Furin 酶切位点（图 6）⁵⁷。相反，在其他种类的冠状病毒中已经在这个位置观察到了的 Furin 酶切位点^{57,58}。这里似乎存在着某种刻意的选择，防止这种自然生成的位点在 B 系 β 属冠状病毒中内获得或保存。



Human SARS-CoV BJ01	655 - GICASYHTVSLI-----RSTS - 670
Human SARS-CoV CUHK-W1	655 - GICASYHTVSLI-----RSTS - 670
Human SARS-CoV Tor2	655 - GICASYHTVSLI-----RSTS - 670
Human SARS-CoV Frankfurt-1	655 - GICASYHTVSLI-----RSTS - 670
Human SARS-CoV Urbani	655 - GICASYHTVSLI-----RSTS - 670
Civet SARS-CoV civet020	655 - GICASYHTVSSL-----RSTS - 670
Civet SARS-CoV sz16	655 - GICASYHTVSSL-----RSTS - 670
Racoon dog SARS-CoV A030	655 - GICASYHTVSSL-----RSTS - 670
SARS-CoV-2	669 - GICASYQTQTNSPRRARSVA - 688
Pangolin CoV MP789	n/a - GICASYQTQTNS-----RSVS - n/a
Bat SARSr-CoV RaTG13	669 - GICASYQTQTNS-----RSVA - 684
Bat SARSr-CoV LYRa11	659 - GICASYHTASLL-----RNTD - 674
Bat SARSr-CoV LYRa3	659 - GICASYHTASLL-----RNTG - 674
Bat SARSr-CoV RsSHC014	656 - GICASYHTVSSL-----RSTS - 671
Bat SARSr-CoV Rs4084	656 - GICASYHTVSSL-----RSTS - 671
Bat SARSr-CoV WIV1	656 - GICASYHTVSSL-----RSTS - 671
Bat SARSr-CoV Rs3367	656 - GICASYHTVSSL-----RSTS - 671
Bat SARSr-CoV Rs7327	656 - GICASYHTVSSL-----RSTS - 671
Bat SARSr-CoV Rs9401	656 - GICASYHTVSSL-----RSTS - 671
Bat SARSr-CoV Rs4231	655 - GICASYHTVSSL-----RSTS - 670
Bat SARSr-CoV WIV16	655 - GICASYHTVSSL-----RSTS - 670
Bat SARSr-CoV Rs4874	655 - GICASYHTVSSL-----RSTS - 670
Bat SARSr-CoV ZC45	646 - GICASYHTASIL-----RSTS - 661
Bat SARSr-CoV ZXC21	645 - GICASYHTASIL-----RSTG - 660
Bat SARSr-CoV Rf4092	634 - GICASYHTASTL-----RGVG - 649
Bat SARSr-CoV Rf/JL2012	636 - GICASYHTASLL-----RSTG - 651
Bat SARSr-CoV JTMC15	636 - GICASYHTASLL-----RSTG - 651
Bat SARSr-CoV 16B0133	636 - GICASYHTASLL-----RSTG - 651
Bat SARSr-CoV B15-21	636 - GICASYHTASLL-----RSTG - 651
Bat SARSr-CoV YN2013	633 - GICASYHTASTL-----RSIG - 648
Bat SARSr-CoV Anlong-103	633 - GICASYHTASTL-----RSVG - 648
Bat SARSr-CoV Rp/Shaanxi2011	640 - GICASYHTASVL-----RSTG - 655
Bat SARSr-CoV Rs/HuB2013	641 - GICASYHTASVL-----RSTG - 656
Bat SARSr-CoV YNLF/34C	641 - GICASYHTASVL-----RSTG - 656
Bat SARSr-CoV YNLF/31C	641 - GICASYHTASVL-----RSTG - 656
Bat SARSr-CoV Rf1	641 - GICASYHTASHL-----RSTG - 656
Bat SARSr-CoV 273	641 - GICASYHTASHL-----RSTG - 656
Bat SARSr-CoV Rf/SX2013	639 - GICASYHTASLL-----RSTG - 654
Bat SARSr-CoV Rf/HeB2013	641 - GICASYHTASLL-----RSTG - 656
Bat SARSr-CoV Cp/Yunnan2011	641 - GICASYHTASLL-----RNTG - 656
Bat SARSr-CoV Rs672	641 - GICASYHTASTL-----RSVG - 656
Bat SARSr-CoV Rs4255	641 - GICASYHTASTL-----RSVG - 656
Bat SARSr-CoV 4081	641 - GICASYHTASTL-----RSVG - 656
Bat SARSr-CoV Rm1	641 - GICASYHTASVL-----RSTG - 656
Bat SARSr-CoV 279	641 - GICASYHTASVL-----RSTG - 656
Bat SARSr-CoV Rs/GX2013	642 - GICASYHTASVL-----RSTG - 657
Bat SARSr-CoV Rs806	641 - GICASYHTASLL-----RSTG - 656
Bat SARSr-CoV HKU3-1	642 - GICASYHTASVL-----RSTG - 657
Bat SARSr-CoV Longquan-140	642 - GICASYHTASVL-----RSTG - 657
Bat SARSr-CoV Rp3	641 - GICASYHTASTL-----RSVG - 656
Bat SARSr-CoV Rs4247	642 - GICASYHTASTL-----RSVG - 657
Bat SARSr-CoV Rs4237	641 - GICASYHTASTL-----RSVG - 656
Bat SARSr-CoV As6526	641 - GICASYHTASTL-----RSVG - 656

图 6: 在刺突的 S1/S2 的交界处发现的 Furin 酶切位点是 SARS-CoV-2 独有的, 在其他的 B 系 β 属冠状病毒内都不存在。本图转自 Hoffmann 等人⁵⁷。



如前所述，在病毒进入细胞的过程中，刺突蛋白首先在 S1/S2 的交界处被裂解。这一步以及随后进行的暴露融合肽的裂解都是由宿主蛋白酶介导的。这些蛋白酶在不同细胞类型中的存在与否，极大地影响了细胞的趋化性，也很有可能影响了病毒感染的致病性。与其他蛋白酶不同，Furin 蛋白酶在许多类型的细胞中都能看到，并且它在多个细胞和细胞外位置存在。重要的是，在 S1/S2 交界处引入一个 Furin 酶切位点，可以显著增强病毒的感染力并极大地扩大其细胞滋养性 - 这是一个在流感病毒和其他冠状病毒中都被充分记载的现象⁵⁹⁻⁶⁵。

如果我们暂且撇开在自然界任何 β 冠状病毒的 B 系内都没有发现 Furin 酶切位点的事实，相反的，我们假设在 SARS-CoV-2 中的这个位点是自然进化的结果，那么，只有一种进化途径是可能的，那就是 Furin 酶切位点必须来自同源重组。具体地说，一个不含 Furin 酶切位点的 β 属冠状病毒前代必须与一个含有 Furin 酶切位点的近亲冠状病毒重组。

但是，有两个事实削弱了这种可能性。第一，虽然一些其他组或系的冠状病毒确实含有多元 Furin 酶切位点，但是没有一种冠状病毒含有与 SARS-CoV-2 中所见的多元序列 (-PRRAR/SVA-) 一模一样的多元 Furin 酶切位点。第二，在 SARS-CoV-2 和任何含有合理的 Furin 酶切位点的冠状病毒之间，刺突上的序列同一性不超过 40%⁶⁶。如此低的序列同一性排除了这些病毒祖先之间发生成功的同源重组的可能性。因此，SARS-CoV-2 的刺突蛋白中含有的 Furin 酶切位点不可能是来自于自然，相反的，它应该是实验室改造的结果。

无独有偶，透过对 SARS-CoV-2 刺突的核苷酸仔细检查发现了该两段连续的 Arg 残基皆由稀有的密码子 CGG 所构成（此为 SARS-CoV-2 病毒中最少使用的密码子）(图 7)⁸。而实际上，这个 CGGCGG 排序是在 SARS-CoV-2 基因组中发现的、该稀有密码子串联排列的唯一实例。该观察结果强烈表明，该 Furin 酶切位点应该是基因改造的结果。令人怀疑的是，此处通过密码子选择形成了一个 FauI 限制性酶切位点，这表明可能引



入了限制性片段长度多态性技术，这一技术是武汉病毒研究生实验室所精通的⁶⁷。在那里，由 FauI 消化产生的片段化模式可用于监测刺突中的 Furin 酶切位点保存，因为该 Furin 酶切位点在体外很容易就丧失了^{68,69}。尤其是，可以对从细胞培养物或实验动物中回收的病毒的刺突基因进行 RT-PCR，其产物可进行 FauI 消化。保留或丧失 Furin 酶切位点的病毒将产生不同的模式，从而可以方便地跟踪目标病毒。



图7：SARS-CoV-2刺突的S1/S2连接处的PRRA插入中的两个连续Arg残基，均带有稀有密码子CGG编码。FauI限制性酶切位点5'-(N) 6GCGGG-3'，被嵌入进“插入的”PRRA片段的编码序列中，可以用作标记来监测引入的Furin酶切位点的保存。

此外，虽然没有在已知的冠状病毒中含有与 SARS-CoV-2 刺突蛋白中一模一样的-PRRAR/SVA-序列，但在一种但在啮齿类冠状病毒 AcCoV-JC34 的刺突蛋白（刺突 protein）的 S1/S2 交界处观察到了一种类似的序列-RRAR/AR-，而这个结果是由石正丽博士在 2017 年发表的⁷⁰。可见，-RRAR-作为合理的功能性 Furin 酶切位点自 2017 年起就已经被中国科学院武汉病毒研究所的专家所熟知。

所有的这些证据共同表明，SARS-CoV-2 刺突蛋白可能并非源自自然而且可能是基因改造的结果。这种人为改造的目的可能是为了评估实验室制造的冠状病毒在感染性和致病性上任何潜在的增强⁵⁹⁻⁶⁴。事实上，最近的研究已经证实了 Furin 酶切位点的确赋予了 SARS-CoV-2 在致病性上显著的优势^{57,68}。



1.4 总结

本部分提出的证据显示，SARS-CoV-2 基因组的某些方面极难被解释为自然进化的结果。我们提出的另一种理论是，该病毒可能是以 ZC45/ZXC21 蝙蝠冠状病毒为骨干和/或模板而产生的。刺突蛋白，特别是其中的受体结合基序（RBM），应该是经过人为操纵的，在此基础上，病毒获得了与 hACE2 结合并感染人类的能力。在 RBM 两端发现一个独特的限制性酶消化位点也支持了这一点。一个不寻常的 Furin 酶切位点可能被引入并插入刺突蛋白的 S1/S2 连接处，这有助于提高病毒的毒性和致病性。这些转变最终使得 SARS CoV-2 病毒成为一种具有高传播性、发病隐蔽性、致死性、后遗症不明的大规模破坏性病原体。

显然，SARS-CoV-2 可能是通过在武汉病毒学研究所（WIV）的功能增强操作产生的，而且这种可能性极大，需要进行彻底和独立的调查。

2. 描述 SARS-CoV-2 的合成途径

在本报告的第二部分，我们描述了在实验室环境下创造 SARS-CoV-2 的合成途径。这是基于大量文献支持以及 SARS-CoV-2 基因组中存在的遗传证据而推测的。虽然在此提及的流程并非完全和那些实验室所采用的流程一模一样，但我们相信其中几个关键步骤是大同小异的。更重要的是，我们在本文的报导应该被看作是一个示范，展示如何通过遵循成熟的概念和熟练的技术，在研究实验室中简便地设计和创造 SARS-CoV-2 病毒。

重要的是，香港和中国大陆的研究实验室在冠状病毒研究方面，无论在资源还是研究成果方面，都处于世界领先地位。后者不仅体现在过去 20 年的大量出版物上，还体现在他们在这一领域取得的里程碑式的成就上。他们率先发现果子狸是 SARS-CoV 的中间宿主，并分离出第一株病毒⁷¹；他们第一个发现 SARS-CoV 来源于蝙蝠^{72, 73}；

他们首次揭示了 SARS-CoV 感染的抗体依赖性增强 (ADE)⁷⁴；他们在了解 MERS 的各个领域 (人畜共患病、病毒学和临床研究)⁷⁵⁻⁷⁹ 方面做出了重大贡献；他们在 SARS-CoV-2 研究方面取得了多项突破^{18, 35, 80}。最后，同样重要的是，他们拥有世界上最多的冠状病毒 (基因组序列和活病毒)。香港和大陆的研究实验室 (二者广泛合作) 都有现成的知识、专业技术和资源来开展和完成以下所描述的工作。

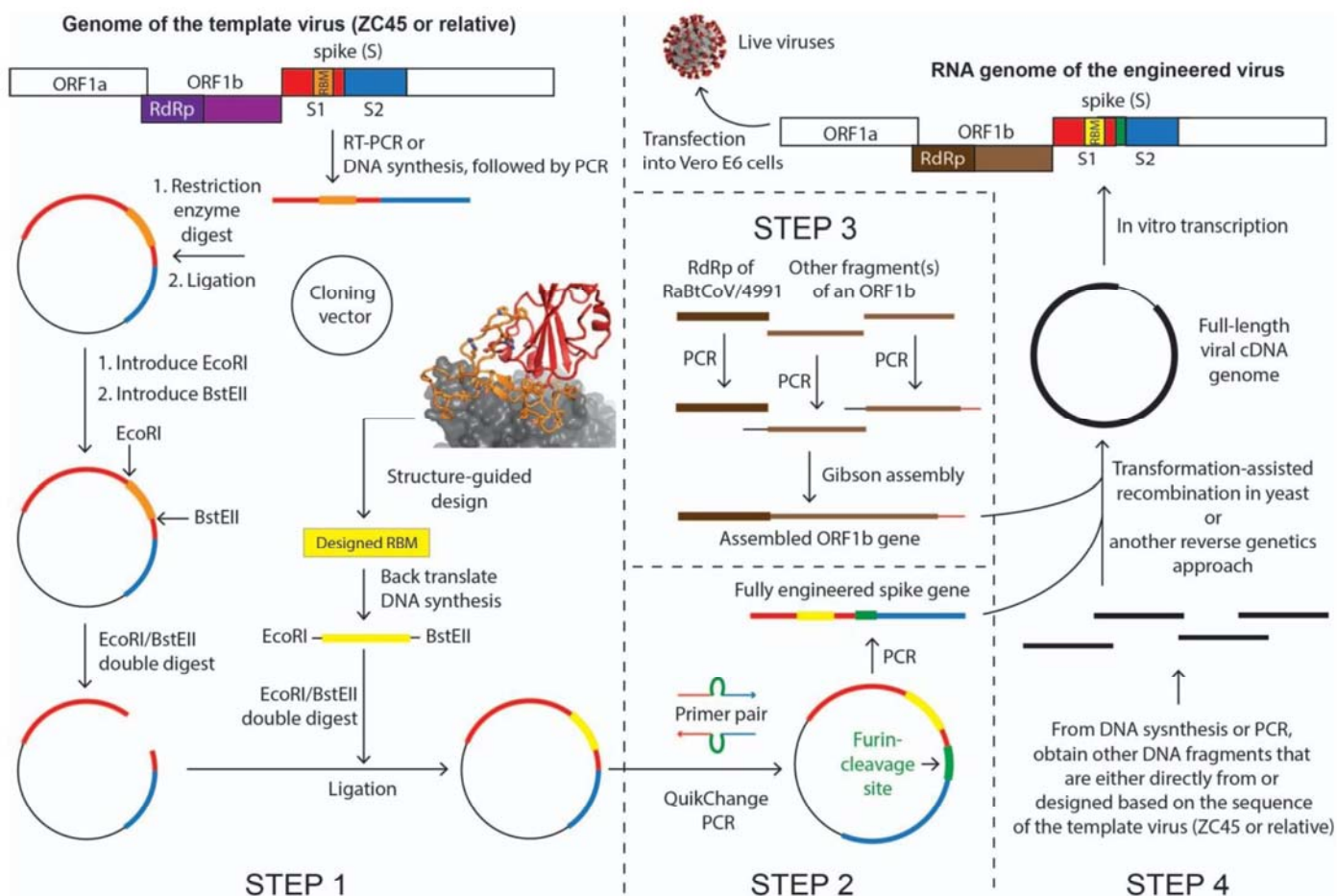


图 8：实验室设计制造的 SARS-CoV-2 病毒可能的流程图

2.1 实验室设计制造的新型冠状病毒的可能方案

在本小节中，我们将概述在项目设计阶段可能制定的总体策略和主要考虑因素。

为了设计和制造一种以人类为靶点的冠状病毒，他们必须要选择一个蝙蝠冠状病毒作为模板/骨干。这可以很简便地完成，因为在过去的二十年里，许多研究实验室



一直在积极收集蝙蝠冠状病毒^{32, 33, 70, 72, 81-85}。然而，这个理想的模板病毒应该不是源于石正丽博士的收藏，毕竟她是众所周知，一直从事冠状病毒的功能增强研究的。因此，ZC45 和/或 ZXC21，这些由军方实验室发现并拥有的新型蝙蝠冠状病毒³³，更适合作为模板/骨干。当然也有可能是这些军事实验室从同一地点发现了其他密切相关的病毒，并保留了一些未公布的病毒，因此，实际的模板可能是 ZC45 和/或 ZXC21，或它们的近亲。无论三者中哪一个是实际的模板，下面描述的推测途径都是一样的。

一旦他们选择了一个模板病毒，他们将首先需要通过分子克隆刺突蛋白，使其能够与 hACE2 相结合。这个操作所涉及的概念和克隆技术已经在文献^{44-46, 84, 86}中进行了充分记录。模板蝙蝠病毒可以之后转换为冠状病毒，以便结合 hACE2 并被感染人类⁴⁴⁻⁴⁶，而这些操作几乎没有失败的风险。

其次，他们将使用分子克隆技术在刺突蛋白的 S1/S2 连接处引入一个 Furin 酶切位点。这种操作，基于已知的知识^{60, 61, 65}，很可能会产生一种更具感染力和致病性的冠状病毒毒株。

第三，他们将产生一个 ORF1b 基因结构。ORF1b 基因编码多聚蛋白 Orf1b，它经过转化后生成单个病毒蛋白：依赖 RNA 的 RNA 聚合酶 (RdRp)、螺旋酶、胍基-N7 甲基转移酶、尿苷酸特异性核糖核酸内切酶和 2'-O-甲基转移酶。这些蛋白都是病毒复制机制的一部分。其中，RdRp 蛋白最为关键，在冠状病毒中是高度固有的部分。重要的是，石正丽博士的实验室采用 PCR 协议，将 RdRp 基因的特定片段进行放大扩增，作为他们检测原始样本（蝙蝠粪便交换、粪便等）中是否存在冠状病毒的主要方法。采用这种方法，石的课题组已经记录了他们成功检测和/或收集到的所有冠状病毒的 RdRp 这个片段的序列信息。

在这里，基因操作要求较低且并不复杂，因为 Orf1b 是固有的，很可能来自任何 β 属冠状病毒的 Orf1b 都将足以胜任这项任务。然而，我们相信，他们会希望将一种特定的 Orf1b 引入病毒中是出于以下两个可能的原因之一：



1. 由于许多系统发育分析仅根据 RdRp 基因的序列相似性对冠状病毒进行分类^{18, 31, 35, 83, 87}，因此在基因组中具有不同的 RdRp 可以确保 SARS-CoV-2 和 ZC45/ZXC21 在系统发育研究中被分成不同的组/亚系。然而，选择 RdRp 基因是很方便的，因为 RdRp 短段序列已经被记录在所有收集/检测到的冠状病毒资料中。他们最终选择的是 2013 年发现的蝙蝠冠状病毒 RaBtCoV/4991 的 RdRp 序列。对于 RaBtCoV/4991 而言，曾经公布的唯一信息是其短 RdRp 段的序列⁸³，而其全基因组序列和病毒分离都没有被报道过。在扩增了 RaBatCoV/4991 的 RdRp 片段（或整个 ORF1b 基因）后，他们就会将其用于 SARS-CoV-2 基因组的后续组装和创建上。RdRp 序列的微小变化可能是在开始时（通过 DNA 合成）引入的，也可以在随后通过传代产生。在另一条不同的轨道上，当他们在制造 RaTG13 序列的时候，他们可以从 RaBtCoV/4991 的 RdRp 短段开始，而不改变其序列，从而使这段短 RdRp 片段⁸³上有 100%的核苷酸序列一致性。这样就可以宣称早在 2013 年这个所谓的 RaTG13 就被发现了。
2. 来自 RaBatCoV/4991 的 RdRp 蛋白是独一无二的，因为它在开发抗病毒药物方面优于来自其他任何 β 冠状病毒的 RdRp。RdRp 在人类细胞中没有同源，这使得这种重要的病毒酶成为非常理想的抗病毒开发靶点。例如，目前正在正在进行临床试验的雷梅西韦（Remdesivir）就是以 RdRp 为靶点。在创造一种新型的、攻击人类的病毒时，他们也会有兴趣开发解药。尽管发明这样的药物可能不太容易，合理的解释是他们故意加入一种更适合开发抗病毒药物的 RdRp 蛋白。

第四，他们会利用反向遗传学的方法将刺突的基因碎片、ORF1b 和 ZC45 的其余的模板组装成一个 cDNA 的版本病毒基因组。然后，他们通过在体外进行嫁接，以此获得病毒 RNA 基因组。RNA 基因的传染性进入细胞，可以使带有所需的人造基因组合病毒恢复活性和传染性。



第五，他们会对病毒株进行特性化和优化，利用在体内的连续传代提高病毒株的适应性和感染性以及整体适应性。一个或多个符合特定标准的病毒株将成为最终产品。

2.2 制造 SARS-CoV-2 的假设合成途径

在这一分类章节部分，我们将更详细地介绍如何通过使用现有的材料和常规分子、细胞和病毒学技术，在实验室中进行每一个步骤。此过程的示意图如图 8 所示。我们估计，整个过程可以在 6 个月左右完成。

第 1 步：为与 hACE2 结合，制造刺突上的受体结合基序（RBM）（1.5 个月）

蝙蝠冠状病毒的刺突蛋白在与 hACE2 结合方面要么不可能，要么效率低下，因为它的 RBM 中缺少重要的残留物。模板病毒 ZC45 的 RBM 可以说明这一点（图 4）。创造 SARS-CoV-2 的首要的也是最关键的步骤是设计出一种刺突，使它具有结合 hACE2 的能力。文献中证明，自 2008 年以来⁴⁴，这种人工干预已经多次在研究实验室中进行，并成功地产生了能够感染人类细胞的冠状病毒^{44-46, 88, 89}。虽然有许多可能的方法可以设计刺突蛋白，但是我们相信，实际上，他们用以 SARS 的 RBM 作为基准，用经过设计并可能优化过的 RBM 取代了原来的 RBM。正如在第 1 部分所描述的，我们的观察可以支持这一理论，即两个独特的限制位点 EcoRI 和 BstEII 存在于 SARS-CoV-2 基因组（图 5A）中 RBM 的两端，事实上，是由石正丽博士和她的长期合作者，结构生物学专家 Fang Li 博士已经成功地进行了这种 RBM 交换技术^{39, 47}。

尽管 ZC45 的刺突不包含这两个限制位点（图 5B），但他们可以非常容易被引入。原始的刺突基因要么和 RT-PCR 一起被放大，要么通过 DNA 合成获得（有些改变可以安全地被引入序列的某些可变区域），然后是 PCR。之后，该基因将通过使用除 EcoRI 和 BstEII 以外的限制位点被克隆成质粒。



一旦进入质粒，刺突的基因可以很容易地被修改。首先，通过将突出显示的“gaacac”序列（图 5B）转换为所需的“gaattc”（图 5A）就可以引入一个 EcoRI 位点。它们之间的区别是个连续的核苷酸。使用市场上有售的 QuikChange 位点定向突变试剂盒，这种二核苷酸突变可以在不超过一周内产生。随后，BstEII 位点也可以以相似方法引入到 RBM 的另一端。具体来说，就是将“gaatacc”序列（图 5B）转换为所需的“ggtacc”（图 5A），同样需要一周的时间。

一旦这些限制位点（SARS-CoV-2 的刺突基因中独一无二的）被成功的引入，就可以方便地交换不同的 RBM 片段，随后就可以使用既定的测定方法对由此产生的刺突蛋白进行评估。

如第 1 部分所述，一个 RBM 片段的设计可以很好地遵循高分辨率结构（图 3）^{37,38}，从而以智能方式生成类似于 SARS 的 RBM 的序列。在对 RBM 进行结构指导设计时，他们会按照常规，并生成几个（例如十几个）此类的 RBM，希望某些特定的变体在绑定 hACE2 方面可能优于其他变体。一旦设计完成，他们可以用每个设计的 RBM 基因与一个 EcoRI 位点在 5' 端和一个 BstEII 位点在 3' 端进行商业化合成（快速和低廉）。然后，这些新的 *RBM 基因* 可以分别被克隆到刺突基因中。基因合成和随后的克隆，可以为特定设计的 RBM 小型基因库以批量的模式完成，大约需要一个月的时间。

然后，这些改进的刺突蛋白可能通过已建立的伪型病毒感染检测^{45,49,50}方式来检测与 hACE2 的结合性。改进后具有良好至特殊结合亲和力的刺突将被选用。（虽然不是必需的，但这里有可能涉及定向进化（RBM 基因上的 PCR 比较容易出错），加上体外结合检测^{39,90}或伪型病毒感染检测^{45,49,50}，使绑定 hACE2 的 RBM 获得特殊的亲和力。

鉴于关于刺突改进^{44-46,84,86}的大量文献和可供参考的刺突-hACE2 复合物^{37,38}高分辨率结构，这一步骤的成功是非常有保证的。在此步骤结束后，如需要，就可以获得一种新的刺突基因，该刺突蛋白的基因编码与 hACE2 结合时具有高亲和力。



第 2 步：在 S1/S2 交界处设计 Furin 酶切位点（0.5 个月）

步骤 1 的产品（包含改进后刺突的质粒）将进一步修改，包括改进在 S1/S2 交汇处包括 Furin 酶切位点（图 4 中的绿线所示的片段）。这种短段的基因序列可以通过使用几种常规克隆技术方便地插入，包括 QuikChange 的位点定向 PCR⁶⁰，重叠 PCR，然后是限制性酶消化和结扎⁹¹，或者是 Gibson 组装。这些技术不会在序列中留下任何痕迹。无论选择哪种克隆方法，插入的基因片段将包含在底像中，该底像将被重新设计、合成，再用于克隆。此步骤，在 S1/S2 交界处添加了经改造 Furin 酶切位点的刺突，可以在不超过两周内完成。

第 3 步：从 RaBtCoV/4991 获取包含短 RdRp 片段序列的 ORF1b 基因（1 个月，但可与步骤 1 和步骤 2 同时进行）

与刺突的（基因）编辑不同，这里不需要复杂的设计，除了需要加入来自 RaBtCoV/4991 的 RdRp 基因片段。可以在这里使用 Gibson 组装。每个相邻的片段共享 20–40 bp（碱基对）的重叠部分，那么这个技术则可以通过在一个简单的反应把几个片段组合在一起，从而组装一个长的 DNA 产品。将根据已知的蝙蝠冠状病毒序列，选择两个或三个片段，每个片段都覆盖 ORF1b 基因的重要部分。这些片段之一是 RaBtCoV/4991 的 RdRp 片段⁸³。每个片段都将被聚合酶链式反应（PCR）扩增，并在引物（primers）中引入适当的重叠区域。最后，将所有提纯后的片段以等摩尔浓度的浓缩物（equimolar concentrations）合并，然后添加至 Gibson 反应混合物中，这将在短暂孵育（incubation）后完整地产出所需的 ORF1b 基因。

步骤 4：使用反向遗传学生产所设计的病毒基因组并恢复活病毒（0.5 个月）

反向遗传学经常被用于组装整个病毒基因组，包括冠状病毒基因组^{67, 92–96}。最新的例子是使用酵母菌中的转化 - 辅助重组法来重建 SARS-CoV-2 的基因组⁹⁷。瑞士的研究团队使用这种方法在短短一周内就组装了整个病毒基因组并产生了活病毒⁹⁷。这种



有效的技术自从2017年就有了^{98,99}，且不会在制备的病毒基因组中留下任何人工操作的痕迹。除了编辑过的刺突基因（来自步骤1和2）和ORF1b基因（来自步骤3）之外，覆盖基因组其余部分的其他片段将通过模板病毒的逆转录聚合酶链式反应（RT-PCR）的扩增或者通过DNA合成与模板病毒的略有不同的序列。我们相信，后一种方法更有可能，因为它将允许把序列改变引入到保守性较低的蛋白质的可变区段，该过程可以很容易地通过多重序列比对来引导。具有更保守功能的氨基酸序列，例如E蛋白的氨基酸序列，可能没有被改变。然后将所有DNA片段将被集中在一起并转化到酵母中，在其中通过转化-辅助重组组装SARS-CoV-2基因组的cDNA版本。当然，也可以采用另一种反向遗传学方法^{67,92-96,100}，武汉病毒研究所在过去成功地使用了该方法⁶⁷。尽管一些较早的反向遗传学方法可能会在连接不同片段的位置留下限制性位点，但由于连接的确切位点可能在大约30 kb（千碱基）的基因组中的任何位置，因此很难检测到这些痕迹。无论哪种方式，都可以从反向遗传学实验中获得病毒基因组的cDNA版本。随后，以cDNA为模板的体外转录将产生病毒RNA基因组，转染到Vero E6细胞后将产生带有一起所设计特性的活病毒。

步骤5：优化病毒的适应性并提高其与hACE2的体内结合亲和力（2.5-3个月）

从第4步中回收的病毒需要进行经典实验 - 实验动物的连续传代以进一步改进¹⁰¹。最后一步将验证病毒的适应性，并确保其针对预期宿主的面向受体的适应性，根据上述分析，宿主应为人。重要的是，分别被引入到刺突蛋白中的RBM和Furin酶切位点现在将作为一个功能单元而一同进行优化。在各种已有的冠状病毒的动物模型中（例如小鼠，仓鼠，雪貂和猴子），人类ACE2受体(hACE2)转基因小鼠(hACE2-小鼠)应该是此处最合适、最便捷的选择。该动物模型是在研究SARS-CoV的过程中建立的，并已在Jackson实验室中使用了多年¹⁰²⁻¹⁰⁴。

连续传代的过程很直接明了。简而言之，将从步骤4中选出的毒株，即SARS-CoV-2的前身，将通过鼻内接种到一组麻醉的hACE2小鼠中。感染后约2-3天，肺部病毒



通常会扩增至峰值滴度。然后处死小鼠并将其肺匀浆。通常，携带最高病毒载量的小鼠肺上清液将用于提取备选病毒用于下一轮传代。经过大约 10~15 轮后，hACE2 的结合亲和力，感染效率和毒株的致死性将得到充分提高而且病毒基因组会稳定下来¹⁰¹。最后，经过一系列表征实验（例如病毒动力学测定，抗体反应测定，症状观察和病理学检查），将获得 SARS-CoV-2 这一最终产品，整个制备过程结束。从此刻开始，这种病毒病原体可以被扩增（最有可能使用 Vero E6 细胞）并按常规生产。

值得注意的是，基于在 SARS-CoV 上所做的工作，虽然 hACE2-小鼠很适合 SARS-CoV-2 的适应，但并不是反映病毒在人体内的传染能力和相关临床症状的最佳模型。我们相信，这些科学家可能并未在 COVID-19 爆发之前使用适当的动物模型（例如叙利亚金仓鼠）来测试 SARS-CoV-2 的传染能力。如果他们使用了适当的动物模型来进行此实验，那么 SARS-CoV-2 的高传染性将显得极其明显，因此在疫情爆发的开始，SARS-CoV-2 则不会被说成“不引起人对人的传播”。

我们还推测，以增强传播性和致命性为导向的大规模实验室改造行为已经使该病毒彻底失控。结果造成 SARS-CoV-2 可能在当前适应人类的过程中无法减弱传播能力和致死性。该假设与 SARS-CoV-2 迄今仍未出现明显的衰减一致，尽管它具有很高的流行性，而且观察到最近出现的一种主要变异显示出其传播能力有所增强¹⁰⁵⁻¹⁰⁸。

连续传代是一个快速而剧烈的过程，病毒的适应性在此过程中得以加速。尽管旨在模仿自然进化，但连续传代在时间和规模上都受到更大的限制。所以，连续传代中预期的随机突变比自然进化的更少。对于诸如 E 蛋白的保守病毒蛋白，尤为如此。E 蛋白是病毒复制的关键因素，它是毒性决定因素而且编辑 E 蛋白可能会导致 SARS-CoV-2 毒性减弱¹⁰⁹⁻¹¹¹。因此，在最初的组装阶段，这些科学家可能决定保持 E 蛋白的氨基酸序列不变，使其与 ZC45/ZXC21 的 E 蛋白相同。由于 E 蛋白的保守性和连续传代的限制，实际上并没有发生氨基酸的突变，从而导致 SARS-CoV-2 与 ZC45/ZXC21 之间的 E 蛋白的序列有 100% 同一性。分子克隆的标记（在受体结合基序两侧的限制性



酶切位点)也可能发生同样的情况。连续传代本应部分地让 SARS-CoV-2 基因组变得自然,但也许并未消除所有人工操作的痕迹。

3. 结语

关于 SARS-CoV-2 的来源,还有很多问题亟需解答。著名的病毒学家们在发表于《自然医学杂志》的一封信中表示,虽然无法完全排除实验室泄露,但这是几乎不可能的,而且在 SARS-CoV-2 的基因组⁴中不存在基因改造的迹象。然而,我们在本文中展示了 SARS-CoV-2 基因组的刺突(蛋白)基因内部的确实存在基因证据(在受体结合基序两侧的限制性酶切位点;在 Furin 酶切位点使用的一串稀有密码子),这些证据显示 SARS-CoV-2 的基因组应该是基因改造的产物。再者,经过验证的概念、成熟的技术、知识和专长都已经到位,足以在短期内便捷地制备出这个新型冠状病毒。

撇开动机不谈,以下关于 SARS-CoV-2 的事实都得到了很好的支持:

1. 如果它是实验室的产物,在它的制备过程中最关键的成分即病毒骨干/模板(ZC45/ZXC21)为军方实验室所拥有。
2. SARS-CoV-2 的基因组序列可能被基因修饰过,病毒由此已获得了通过增强的毒性和传染性来攻击人类的能力。
3. SARS-CoV-2 的特点和致病效果(pathogenic effects)是前所未有的。该病毒有高传染性、发病隐蔽性(onset-hidden)、攻击多个器官、后遗症不明确、致命以及和多种症状及并发症有关。
4. SARS-CoV-2 导致了全球范围内的大流行,造成数十万人丧失且让全球经济停摆。它有着无以伦比的破坏力。

从我们和其他人收集的证据来判断,我们认为找到 SARS-CoV-2 的来源应当要对武汉病毒研究所 P4 实验室和与之有紧密联系的合作者进行独立的调查。这种调查应该在很久以前就进行了而且不应当再拖延下去。



我们也注意到在 2015 年发表的嵌合病毒 SHC015-MA15 (chimeric virus) 的论文中，最开始漏掉了美国国家过敏和传染病研究所 (NIAID) 为石正丽所提供的经费的事实。或许是在 2016 年 1 月召开的美国国立卫生研究院恢复对病毒功能增强研究的提供资金支持的会议之后，这件事重新出现在发表于 2016 年的勘误之中。

在这份报告中并未全面讲述的是，有各种各样的证据显示最近发表的好几个冠状病毒 (RaTG13¹⁸、RmYN02³⁰、以及多个穿山甲病毒^{27-29,3}) 都高度存疑而且很可能是伪造的。这些伪造的病毒除了欺骗科学界和普通民众以确保 SARS-CoV-2 的真实身份被掩盖以外没有任何用处。尽管把这些证据的细节排除在外并不会改变本报告的结论，我们确实坚信这些细节将为我们的观点提供额外的支持，我们认为 SARS-CoV-2 是实验室增强的病毒和功能增强研究的产物。一份着眼于这些额外证据的后续报告正在准备之中并将被很快提交。

致谢

感谢 Daoyu Zhang 分享给我们 E 蛋白在 β 类冠状病毒不同亚群内变异的发现。我们也感谢所有匿名的科学家和其他人士为发现与 SARS-CoV-2 起源有关的事实所做出的贡献。



参考文献：

1. Zhan, S.H., Deverman, B.E. & Chan, Y.A. SARS-CoV-2 is well adapted for humans. What does this mean for re-emergence? *bioRxiv*, <https://doi.org/10.1101/2020.05.01.073262> (2020).
2. Mou, H. et al. Mutations from bat ACE2 orthologs markedly enhance ACE2-Fc neutralization of SARS-CoV-2. *bioRxiv*, <https://doi.org/10.1101/2020.06.29.178459> (2020).
3. Piplani, S., Singh, P.K., Winkler, D.A. & Petrovsky, N. In silico comparison of 刺突 protein-ACE2 binding affinities across species; significance for the possible origin of the SARS-CoV-2 virus. *arXiv*, arXiv:2005.06199 (2020).
4. Andersen, K.G., Rambaut, A., Lipkin, W.I., Holmes, E.C. & Garry, R.F. The proximal origin of SARS-CoV-2. *Nat Med* **26**, 450-452 (2020).
5. Maiti, A.K. On The Origin of SARS-CoV-2 Virus. *Preprint (authorea.com)*, DOI: 10.22541/au.159355977.76503625 (2020).
6. Lin, X. & Chen, S. Major Concerns on the Identification of Bat Coronavirus Strain RaTG13 and Quality of Related Nature Paper. *Preprints*, 2020060044 (2020).
7. Bengston, D. All journal articles evaluating the origin or epidemiology of SARS-CoV-2 that utilize the RaTG13 bat strain genomics are potentially flawed and should be retracted. *OSFPreprints*, DOI: 10.31219/osf.io/wy89d (2020).
8. Segreto, R. & Deigin, Y. Is considering a genetic-manipulation origin for SARS-CoV-2 a conspiracy theory that must be censored? *Preprint (Researchgate)* DOI: 10.13140/RG.2.2.31358.13129/1 (2020).
9. Rahalkar, M.C. & Bahulikar, R.A. Understanding the Origin of ‘BatCoV-RaTG13’, a Virus Closest to SARS-CoV-2. *Preprints*, 2020050322 (2020).
10. Robinson, C. Was the COVID-19 virus genetically engineered? (<https://gmwatch.org/en/news/latest-news/19383>, 2020).
11. Robinson, C. Another expert challenges assertions that SARS-CoV-2 was not genetically engineered. (<https://gmwatch.org/en/news/latest-news/19383>, 2020).
12. Sørensen, B., Dalgleish, A. & Susrud, A. The Evidence which Suggests that This Is No Naturally Evolved Virus. *Preprint*, <https://www.minervanett.no/files/2020/07/13/TheEvidenceNoNaturalEvol.pdf> (2020).
13. Zhang, B. SARS-CoV-2 Could Come from a Lab - A Critique of “The Proximal Origin of SARS-CoV-2” Published in Nature Medicine. (https://www.linkedin.com/pulse/sars-cov-2-could-come-from-lab-critique-proximal-origin-billy-zhang/?articleId=6651628681431175168#comments-6651628681431175168&trk=public_profile_article_view, 2020).
14. Sirotkin, K. & Sirotkin, D. Might SARS - CoV - 2 Have Arisen via Serial Passage through an Animal Host or Cell Culture? *BioEssays*, <https://doi.org/10.1002/bies.202000091> (2020).
15. Seyran, M. et al. Questions concerning the proximal origin of SARS-CoV-2. *J Med Virol* (2020).
16. China Honors Ian Lipkin. (<https://www.publichealth.columbia.edu/public-health-now/news/china-honors-ian-lipkin>, 2020).
17. Holmes, E. Academic CV. (<https://www.sydney.edu.au/AcademicProfiles/profile/resource?urlid=edward.holmes&type=cv>, 2020).



18. Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*(2020).
19. Rahalkar, M. &Bahulikar, R. The Abnormal Nature of the Fecal Swab Sample used for NGS Analysis of RaTG13 Genome Sequence Imposes a Question on the Correctness of the RaTG13 Sequence. *Preprints.org*, 2020080205 (2020).
20. Singla, M., Ahmad, S., Gupta, C. &Sethi, T. De-novo Assembly of RaTG13 Genome Reveals Inconsistencies Further Obscuring SARS-CoV-2 Origins. *Preprints*, 2020080595 (doi: 10.20944/preprints202008.0595.v1) (2020).
21. Zhang, D. Anomalies in BatCoV/RaTG13 sequencing and provenance. *Preprint(zenodo.org)*, <https://zenodo.org/record/3987503#.Xz9GzC-z3GI> (2020).
22. Robinson, C. Journals censor lab origin theory for SARS-CoV-2. (<https://www.gmwatch.org/en/news/latest-news/19475-journals-censor-lab-origin-theory-for-sars-cov-2>, 2020).
23. Scientific evidence and logic behind the claim that the Wuhan coronavirus is man-made. <https://nerdhaspower.weebly.com> (2020).
24. Zhang, Y. et al. The ORF8 Protein of SARS-CoV-2 Mediates Immune Evasion through Potently Downregulating MHC-I. *bioRxiv*, <https://doi.org/10.1101/2020.05.24.111823> (2020).
25. Muth, D. et al. Attenuation of replication by a 29 nucleotide deletion in SARS-coronavirus acquired during the early stages of human-to-human transmission. *Sci Rep***8**, 15177 (2018).
26. Schoeman, D. & Fielding, B.C. Coronavirus envelope protein: current knowledge. *Virol J***16**, 69 (2019).
27. Lam, T.T. et al. Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature* (2020).
28. Liu, P. et al. Are pangolins the intermediate host of the 2019 novel coronavirus (SARS-CoV-2)? *PLoS Pathog***16**, e1008421 (2020).
29. Xiao, K. et al. Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature* (2020).
30. Zhou, H. et al. A Novel Bat Coronavirus Closely Related to SARS-CoV-2 Contains Natural Insertions at the S1/S2 Cleavage Site of the 刺突 Protein. *Curr Biol***30**, 2196-2203 e3 (2020).
31. Zhang, T., Wu, Q. & Zhang, Z. Probable Pangolin Origin of SARS-CoV-2 Associated with the COVID- 19 Outbreak. *Curr Biol***30**, 1578 (2020).
32. Yang, X.L. et al. Isolation and Characterization of a Novel Bat Coronavirus Closely Related to the Direct Progenitor of Severe Acute Respiratory Syndrome Coronavirus. *J Virol***90**, 3253-6 (2015).
33. Hu, D. et al. Genomic characterization and infectivity of a novel SARS-like coronavirus in Chinese bats. *Emerg Microbes Infect***7**, 154 (2018).
34. Wang, Y. Preliminary investigation of viruses carried by bats on the southeast coastal area (东南沿海地区蝙蝠携带病毒的初步调查研究). Master Thesis (2017).
35. Wu, F. et al. A new coronavirus associated with human respiratory disease in China. *Nature***579**, 265-269 (2020).
36. Lab That First Shared Novel Coronavirus Genome Still Shut Down by Chinese Government. *Global Biodefense*, <https://globalbiodefense.com/headlines/chinese-lab-that-first-shared-novel-coronavirus-genome-shut-down/>(2020).



37. Song, W., Gui, M., Wang, X. & Xiang, Y. Cryo-EM structure of the SARS coronavirus 刺突 glycoprotein in complex with its host cell receptor ACE2. *PLoS Pathog***14**, e1007236 (2018).
38. Li, F., Li, W., Farzan, M. & Harrison, S.C. Structure of SARS coronavirus 刺突 receptor-binding domain complexed with receptor. *Science***309**, 1864-8 (2005).
39. Shang, J. et al. Structural basis of receptor recognition by SARS-CoV-2. *Nature* (2020).
40. Hassanin, A. The SARS-CoV-2-like virus found in captive pangolins from Guangdong should be better sequenced. *bioRxiv*, <https://doi.org/10.1101/2020.05.07.077016> (2020).
41. Zhang, D. The Pan-SL-CoV/GD sequences may be from contamination. *Preprint (zenodo.org)*, DOI: 10.5281/zenodo.3885333 (2020).
42. Chan, Y.A. & Zhan, S.H. Single source of pangolin CoVs with a near identical 刺突 RBD to SARS- CoV-2. *bioRxiv*, <https://doi.org/10.1101/2020.07.07.184374> (2020).
43. Lee, J. et al. No evidence of coronaviruses or other potentially zoonotic viruses in Sunda pangolins (*Manis javanica*) entering the wildlife trade via Malaysia. *bioRxiv*, <https://doi.org/10.1101/2020.06.19.158717> (2020).
44. Becker, M.M. et al. Synthetic recombinant bat SARS-like coronavirus is infectious in cultured cells and in mice. *Proc Natl Acad Sci US A***105**, 19944-9 (2008).
45. Menachery, V.D. et al. A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nat Med***21**, 1508-13 (2015).
46. Menachery, V.D. et al. SARS-like WIV1-CoV poised for human emergence. *Proc Natl Acad Sci US A* **113**, 3048-53 (2016).
47. Ren, W. et al. Difference in receptor usage between severe acute respiratory syndrome (SARS) coronavirus and SARS-like coronavirus of bat origin. *J Virol***82**, 1899-907 (2008).
48. Li, X. et al. Emergence of SARS-CoV-2 through Recombination and Strong Purifying Selection. *bioRxiv* (2020).
49. Hou, Y. et al. Angiotensin-converting enzyme 2 (ACE2) proteins of different bat species confer variable susceptibility to SARS-CoV entry. *Arch Virol***155**, 1563-9 (2010).
50. Yang, Y. et al. Two Mutations Were Critical for Bat-to-Human Transmission of Middle East Respiratory Syndrome Coronavirus. *J Virol***89**, 9119-23 (2015).
51. Luo, C.M. et al. Discovery of Novel Bat Coronaviruses in South China That Use the Same Receptor as Middle East Respiratory Syndrome Coronavirus. *J Virol***92**(2018).
52. Cui, J., Li, F. & Shi, Z.L. Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol***17**, 181- 192 (2019).
53. Wan, Y. et al. Molecular Mechanism for Antibody-Dependent Enhancement of Coronavirus Entry. *J Virol***94**(2020).
54. Li, F. Receptor recognition mechanisms of coronaviruses: a decade of structural studies. *J Virol***89**, 1954- 64 (2015).
55. Li, F. Structure, Function, and Evolution of Coronavirus 刺突 Proteins. *Annu Rev Virol***3**, 237-261 (2016).
56. Shang, J. et al. Cell entry mechanisms of SARS-CoV-2. *Proc Natl Acad Sci U S A***117**, 11727-11734 (2020).
57. Hoffmann, M., Kleine-Weber, H. & Pohlmann, S. A Multibasic Cleavage Site in the 刺突 Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells. *Mol Cell***78**, 779-784 e5 (2020).



58. Coutard, B. et al. The 刺突 glycoprotein of the new coronavirus 2019-nCoV contains a Furin-like cleavage site absent in CoV of the same clade. *Antiviral Res***176**, 104742 (2020).
59. Claas, E.C. et al. Human influenza A H5N1 virus related to a highly pathogenic avian influenza virus. *Lancet***351**, 472-7 (1998).
60. Watanabe, R. et al. Entry from the cell surface of severe acute respiratory syndrome coronavirus with cleaved S protein as revealed by pseudotype virus bearing cleaved S protein. *J Virol***82**, 11985-91 (2008).
61. Belouzard, S., Chu, V.C. & Whittaker, G.R. Activation of the SARS coronavirus 刺突 protein via sequential proteolytic cleavage at two distinct sites. *Proc Natl Acad Sci U S A***106**, 5871-6 (2009).
62. Kido, H. et al. Role of host cellular proteases in the pathogenesis of influenza and influenza-induced multiple organ failure. *BiochimBiophys Acta***1824**, 186-94 (2012).
63. Sun, X., Tse, L.V., Ferguson, A.D. & Whittaker, G.R. Modifications to the hemagglutinin cleavage site control the virulence of a neurotropic H1N1 influenza virus. *J Virol***84**, 8683-90 (2010).
64. Cheng, J. et al. The S2 Subunit of QX-type Infectious Bronchitis Coronavirus 刺突 Protein Is an Essential Determinant of Neurotropism. *Viruses***11**(2019).
65. Ito, T. et al. Generation of a highly pathogenic avian influenza A virus from an avirulent field isolate by passaging in chickens. *J Virol***75**, 4439-43 (2001).
66. Canrong Wu, Y.Y., Yang Liu, Peng Zhang, Yali Wang, Hua Li, Qiqi Wang, Yang Xu, Mingxue Li, Mengzhu Zheng, Lixia Chen. Furin, a potential therapeutic target for COVID-19. *Preprint (chinaXiv)*, <http://www.chinaxiv.org/abs/202002.00062> (2020).
67. Zeng, L.P. et al. Bat Severe Acute Respiratory Syndrome-Like Coronavirus WIV1 Encodes an Extra Accessory Protein, ORFX, Involved in Modulation of the Host Immune Response. *J Virol***90**, 6573-6582 (2016).
68. Lau, S.Y. et al. Attenuated SARS-CoV-2 variants with deletions at the S1/S2 junction. *Emerg Microbes Infect***9**, 837-842 (2020).
69. Liu, Z. et al. Identification of common deletions in the 刺突 protein of SARS-CoV-2. *J Virol* (2020).
70. Ge, X.Y. et al. Detection of alpha- and betacoronaviruses in rodents from Yunnan, China. *Virol J***14**, 98 (2017).
71. Guan, Y. et al. Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science***302**, 276-8 (2003).
72. Ge, X.Y. et al. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature***503**, 535-8 (2013).
73. Lau, S.K. et al. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proc Natl Acad Sci U S A***102**, 14040-5 (2005).
74. Kam, Y.W. et al. Antibodies against trimeric S glycoprotein protect hamsters against SARS-CoV challenge despite their capacity to mediate FcγR2-dependent entry into B cells in vitro. *Vaccine***25**, 729-40 (2007).
75. Chan, J.F. et al. Middle East respiratory syndrome coronavirus: another zoonotic betacoronavirus causing SARS-like disease. *Clin Microbiol Rev***28**, 465-522 (2015).



76. Zhou, J., Chu, H., Chan, J.F. & Yuen, K.Y. Middle East respiratory syndrome coronavirus infection: virus-host cell interactions and implications on pathogenesis. *Viol* **J12**, 218 (2015).
77. Yeung, M.L. et al. MERS coronavirus induces apoptosis in kidney and lung by upregulating Smad7 and FGF2. *Nat Microbiol* **1**, 16004 (2016).
78. Chu, D.K.W. et al. MERS coronaviruses from camels in Africa exhibit region-dependent genetic diversity. *Proc Natl Acad Sci U S A* **115**, 3144-3149 (2018).
79. Ommeh, S. et al. Genetic Evidence of Middle East Respiratory Syndrome Coronavirus (MERS-Cov) and Widespread Seroprevalence among Camels in Kenya. *Viol Sin* **33**, 484-492 (2018).
80. Sia, S.F. et al. Pathogenesis and transmission of SARS-CoV-2 in golden hamsters. *Nature* (2020).
81. Ren, W. et al. Full-length genome sequences of two SARS-like coronaviruses in horseshoe bats and genetic variation analysis. *J Gen Virol* **87**, 3355-9 (2006).
82. Yuan, J. et al. Intraspecies diversity of SARS-like coronaviruses in *Rhinolophus sinicus* and its implications for the origin of SARS coronaviruses in humans. *J Gen Virol* **91**, 1058-62 (2010).
83. Ge, X.Y. et al. Coexistence of multiple coronaviruses in several bat colonies in an abandoned mineshaft. *Viol Sin* **31**, 31-40 (2016).
84. Hu, B. et al. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLoS Pathog* **13**, e1006698 (2017).
85. Luo, Y. et al. Longitudinal Surveillance of Betacoronaviruses in Fruit Bats in Yunnan Province, China During 2009-2016. *Viol Sin* **33**, 87-95 (2018).
86. Kuo, L., Godeke, G.J., Raamsman, M.J., Masters, P.S. & Rottier, P.J. Retargeting of coronavirus by substitution of the 刺突 glycoprotein ectodomain: crossing the host cell species barrier. *J Virol* **74**, 1393- 406 (2000).
87. Drexler, J.F. et al. Genomic characterization of severe acute respiratory syndrome-related coronavirus in European bats and classification of coronaviruses based on partial RNA-dependent RNA polymerase gene sequences. *J Virol* **84**, 11336-49 (2010).
88. Agnihothram, S. et al. A mouse model for Betacoronavirus subgroup 2c using a bat coronavirus strain HKU5 variant. *mBio* **5**, e00047-14 (2014).
89. Johnson, B.A., Graham, R.L. & Menachery, V.D. Viral metagenomics, protein structure, and reverse genetics: Key strategies for investigating coronaviruses. *Virology* **517**, 30-37 (2018).
90. Wu, K., Peng, G., Wilken, M., Geraghty, R.J. & Li, F. Mechanisms of host receptor adaptation by severe acute respiratory syndrome coronavirus. *J Biol Chem* **287**, 8904-11 (2012).
91. Follis, K.E., York, J. & Nunberg, J.H. Furin cleavage of the SARS coronavirus 刺突 glycoprotein enhances cell-cell fusion but does not affect virion entry. *Virology* **350**, 358-69 (2006).
92. Yount, B., Denison, M.R., Weiss, S.R. & Baric, R.S. Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. *J Virol* **76**, 11065-78 (2002).
93. Yount, B. et al. Reverse genetics with a full-length infectious cDNA of severe acute respiratory syndrome coronavirus. *Proc Natl Acad Sci U S A* **100**, 12995-3000 (2003).
94. Almazan, F. et al. Construction of a severe acute respiratory syndrome coronavirus infectious cDNA clone and a replicon to study coronavirus RNA synthesis. *J Virol* **80**, 10900-6 (2006).
95. Scobey, T. et al. Reverse genetics with a full-length infectious cDNA of the Middle East respiratory syndrome coronavirus. *Proc Natl Acad Sci U S A* **110**, 16157-62 (2013).
96. Almazan, F., Marquez-Jurado, S., Nogales, A. & Enjuanes, L. Engineering infectious cDNAs of coronavirus as bacterial artificial chromosomes. *Methods Mol Biol* **1282**, 135-52 (2015).



97. Thao, T.T.N. et al. Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform. *Nature* (2020).
98. Oldfield, L.M. et al. Genome-wide engineering of an infectious clone of herpes simplex virus type 1 using synthetic genomics assembly methods. *Proc Natl Acad Sci U S A***114**, E8885-E8894 (2017).
99. Vashee, S. et al. Cloning, Assembly, and Modification of the Primary Human Cytomegalovirus Isolate Toledo by Yeast-Based Transformation-Associated Recombination. *mSphere***2**(2017).
100. Xie, X. et al. An Infectious cDNA Clone of SARS-CoV-2. *Cell Host Microbe***27**, 841-848 e3 (2020).
101. Roberts, A. et al. A mouse-adapted SARS-coronavirus causes disease and mortality in BALB/c mice. *PLoS Pathog***3**, e5 (2007).
102. Roberts, A. et al. Animal models and vaccines for SARS-CoV infection. *Virus Res***133**, 20-32 (2008).
103. Takayama, K. In Vitro and Animal Models for SARS-CoV-2 research. *Trends Pharmacol Sci***41**, 513- 517 (2020).
104. Wang, Q. hACE2 Transgenic Mouse Model For Coronavirus (COVID-19) Research. The Jackson Laboratory Research Highlight, <https://www.jax.org/news-and-insights/2020/february/introducing-mouse-model-for-corona-virus#>(2020).
105. Zhang, L. et al. The D614G mutation in the SARS-CoV-2 刺突 protein reduces S1 shedding and increases infectivity. *bioRxiv*, <https://doi.org/10.1101/2020.06.12.148726>(2020).
106. Yurkovetskiy, L. et al. Structural and Functional Analysis of the D614G SARS-CoV-2 刺突 Protein Variant. *bioRxiv*, <https://doi.org/10.1101/2020.07.04.187757> (2020).
107. Korber, B. et al. Tracking Changes in SARS-CoV-2 刺突: Evidence that D614G Increases Infectivity of the COVID-19 *Virus*. *Cell***182**, 812-827 e19 (2020).
108. Plante, J.A. et al. 刺突 mutation D614G alters SARS-CoV-2 fitness and neutralization susceptibility. *bioRxiv*, <https://doi.org/10.1101/2020.09.01.278689> (2020).
109. Poon, L.L. et al. Recurrent mutations associated with isolation and passage of SARS coronavirus in cells from non-human primates. *J Med Virol***76**, 435-40 (2005).
110. Pervushin, K. et al. Structure and inhibition of the SARS coronavirus envelope protein ion channel. *PLoS Pathog***5**, e1000511 (2009).
111. Nieto-Torres, J.L. et al. Severe acute respiratory syndrome coronavirus envelope protein ion channel activity promotes virus fitness and pathogenesis. *PLoS Pathog***10**, e1004077 (2014).