

## **SARS-CoV-2基因组的不寻常特征表明，该基因组是经过复杂的实验室改造而非自然进化的，并确定了其可能的合成途径**

闫丽梦（医学博士，博士） 康舒（博士） 冠杰（博士） 胡善昌（博士）

法治基金及法治社会，美国纽约州纽约市

通讯地址：[team.lmyan@gmail.com](mailto:team.lmyan@gmail.com)

（喜马拉雅国际工作站翻译）

### 摘要

由新型冠状病毒SARS-CoV-2引起的COVID-19大流行，已导致全球91万多人死亡，全球经济受到前所未有的破坏。尽管影响巨大，但SARS-CoV-2的起源仍然是神秘且有争议的。自然起源论虽然被广泛接受，但缺乏实质性的支持。然而，另一种理论认为病毒可能来自研究实验室，但在同行评审的科学期刊上却受到了严格审查。尽管如此，SARS-CoV-2表现出的生物学特征与自然发生的人畜共患病毒不一致。在这份报告中，我们描述了基因组、结构、医学和文献证据，综合这些证据起来看，其强烈地与自然起源理论相矛盾。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/或骨干创建的实验室产物。在这些证据的基础上，我们进一步假设了SARS-CoV-2的合成路径，证明这种冠状病毒在实验室创建的方便性，并且可以在大约六个月内完成。我们的工作强调了对相关研究实验室进行独立调查的必要性。该报告还主张对最近公布的某些数据进行批判性研究，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

何进入人类世界的途径，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。

## 介绍

COVID-19造成了世界性的大流行病，其规模和严重性是前所未有的。尽管国际社会做出了巨大的努力，但对这一流行病的管理和控制仍然困难重重，充满挑战。

作为一种冠状病毒，SARS-CoV-2与其他呼吸道和/或人畜共患的病毒有很大的不同：它能攻击多个器官；它能够进行长时间的无症状感染；它具有高度的传染性，在高危人群中具有明显的致死性；它自出现之初就很好地适应了人类；它能高效地结合人类ACE2受体(hACE2)，其亲和力大于与任何其他潜在宿主的ACE2相关的结合力。

SARS-CoV-2的起源仍有很多争议。一篇被广泛引用的《自然医学》(Nature Medicine)出版物声称，SARS-CoV-2很可能来自自然界。然而，这篇文章及其中心结论现在受到来自世界各地科学家的质疑。此外，《自然医学》这篇文章的作者有利益冲突的迹象，这进一步引起了人们对SARS-CoV-2的关注。

支持自然起源论的现有科学出版物主要依靠一个证据—之前发现的一种名为RaTG13的蝙蝠冠状病毒，它与SARS-CoV-2的核苷酸序列相同度高达96%。

然而，RaTG13在自然界中的存在及其所报道的序列的真实性正受到广泛质疑。值得注意的是，科学杂志明确地审查了任何暗示SARS-CoV-2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，尽管在科学上具有很高的质量，但只能作为预印本或其他非同行

评审的文章存在于各种网络平台上。尽管如此，对这些报道的分析却一再指出，RaTG13的报道存在严重问题，很可能存在欺诈行为。因此，发表捏造的科学数据以误导全世界追查SARS-CoV-2的来源的理论已经具有实质性的说服力，并与SARS-CoV-2是非自然来源的概念相互关联。

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的非自然起源。此外，丰富的文献表明，功能增强研究早已发展到病毒基因组可以被精确工程化和操纵的阶段，从而能够研制出具有独特特性的新型冠状病毒。在本报告中，我们介绍了这样的证据和相关分析。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，而不是简单的病毒序列所能提供的。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，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，展示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研制出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是否经过体外培养（实验室操纵）？

我们提出了三条证据来支持我们的论点，即实验室操作是产生SARS-CoV-2的一部分。

**i、SARS-CoV-2的基因组序列与第三军医大学(中国重庆)和南京指挥部医学研究所(中国南京)军方实验室发现的一种蝙蝠冠状病毒的基因组序列可疑地相似。**

ii、SARS-CoV-2的突刺蛋白内的受体结合基序(RBM)决定了该病毒的宿主特异性，它与2003年流行的SARS-CoV的受体结合基序相似，令人怀疑。基因组证据表明，RBM已被基因改造。

iii、SARS-CoV-2在其突刺蛋白中含有一个独特的福林酶切位点，众所周知，该位点能大大增强病毒的感染力和细胞的趋性。然而，在自然界中发现的这一类特殊的冠状病毒中完全没有这个切位点。此外，与这一额外序列相关的罕见密码子表明，这一福林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多株重组事件以外的技术，在共同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是创建SARS-CoV-2的骨干

长约3万个核苷酸的SARS-CoV-2基因组结构如图1所示。搜索NCBI序列数据库发现，在所有已知的冠状病毒中，有两个相关的蝙蝠冠状病毒ZC45和ZXC21与SARS-CoV-2具有最高的序列同一性（每个蝙蝠冠状病毒与SARS-CoV-2在核苷酸水平上有约89%的相同性）。图1描述了SARS-CoV-2与代表性β冠状病毒基因组的相似性。有97%的相同性和非常相似特征的ZXC21与ZC45没有显示。请注意，鉴于有强有力的证据表明RaTG13病毒的序列可能是捏造的，而且该病毒在自然界中并不存在，因此RaTG13病毒被排除在本分析之外。（不久将提交一份后续报告，总结证明RaTG13虚假性质的最新证据）

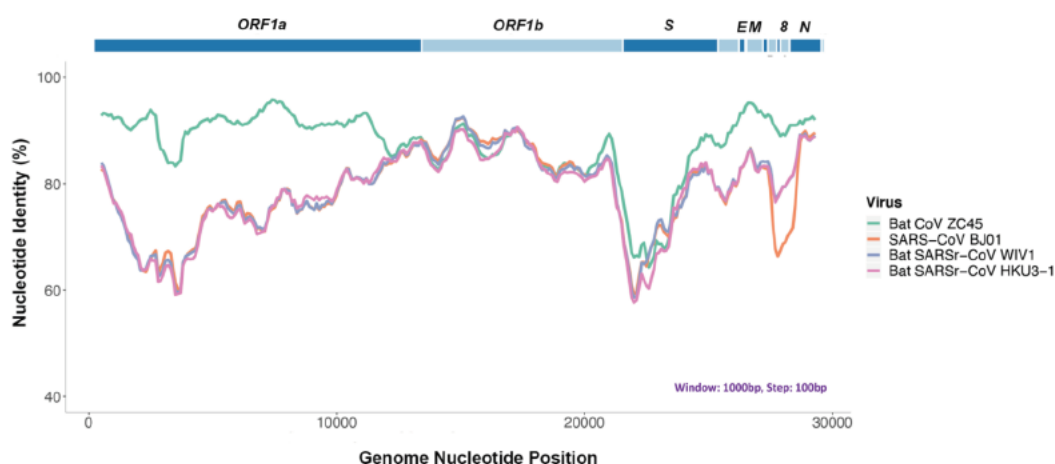


图1、基因组序列分析显示，蝙蝠冠状病毒ZC45与SARS-CoV-2的匹配度最接近。

顶部：SARS-CoV-2 (2019-nCoV WIV04) 的基因组组织。底部：基于2019-nCoV WIV04全长基因组的相似性图。以SARS-CoV BJ01、蝙蝠SARSr-CoV WIV1、蝙蝠SARSr-CoV HKU3-1、蝙蝠冠状病毒ZC45的全长基因组作为参考序列。

当SARS-CoV-2和ZC45/ZXC21在氨基酸水平上进行比较时，发现大多数蛋白质的序列高度一致。核衣壳蛋白的相同度为94%，膜蛋白的相同度为98.6%。突刺蛋白的S2部分（第2半部分）相同度为95%。重要的是，Orf8蛋白的相同度为94.2%，E蛋白的相同度为100%。

Orf8是一种附属蛋白，其功能在大多数冠状病毒中基本上是未知的，尽管最近的数据表明SARS-CoV-2的Orf8通过下调MHC-I来介导宿主适应性免疫逃避。通常情况下，Orf8在冠状病毒中保守性较差。基因序列爆破表明，ZC45/ZXC21的Orf8蛋白与SARS-CoV-2的Orf8有94.2%的相同度，而没有任何其他冠状病毒与SARS-CoV-2在这一特定蛋白上有超过58%的相同度。这里在通常保守性较差的Orf8蛋白上的高度同源性是非常不寻常的

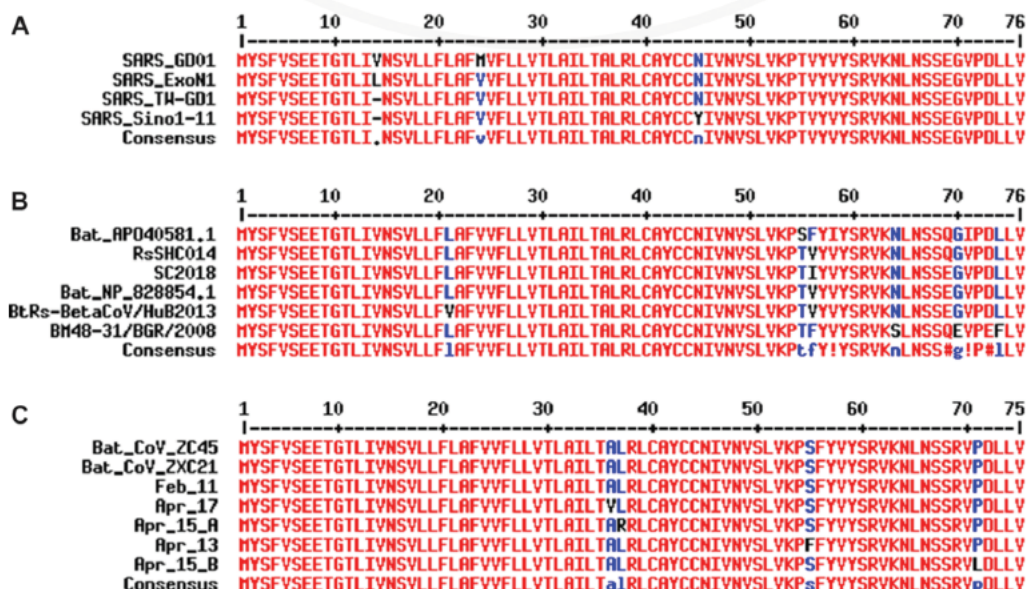


图2、来自不同 $\beta$ 冠状病毒的E蛋白的序列排列，表明了E蛋白对氨基酸突变的宽容性和倾向性。A、在SARS-CoV的不同菌株中观察到了突变。基因库登记号：SARS\_GD01: AY278489.2, SARS\_ExoN1: ACB69908.1, SARS\_TW\_GD1: AY451881.1, SARS\_ino1\_11: AY485277.1。B、来自相关蝙蝠冠状病毒的E蛋白的排列表明其对多个位置的突变的容忍度。基因库登记号：Bat\_AP040581.1: APO40581.1, RsSHC014: KC881005.1, SC2018: MK211374.1, Bat\_NP\_828854.1: NP\_828854.1, BtRs BetaCoV/HuB2013: AIA62312.1, BM48-31/BGR/2008: YP\_003858586.1。C、SARS-CoV-2的早期样本与ZC45和ZXC21在E蛋白上有100%的同一性，但从2020年4月开始的SARS-CoV-2的测序数据表明，在多个位置发生了突变。病毒的登记号。Feb\_11: MN997409, ZC45: MG772933.1, ZXC21: MG772934, Apr\_13: MT326139, Apr\_15\_A: MT263389, Apr\_15\_B: MT293206, Apr\_17: MT350246. 利用MultAlin网络服务器 (<http://multalin.toulouse.inra.fr/multalin/>) 进行了配准。

冠状病毒E蛋白是一种结构性蛋白，它嵌入并内衬在病毒粒子包膜的内部。E蛋白对突变具有耐受性，这在SARS（图2A）和相关蝙蝠冠状病毒（图2B）中都得到了证明。这种对E蛋白氨基酸突变的耐受性在目前的SARS-CoV-2大流行中得到进一步证明。自SARS-CoV-2在人类中爆发以来，经过短短两个月的病毒传播，SARS-CoV-2中的E蛋白已经发生了突变。4月份获得的序列数据显示，不同毒株的四个不同位置已经发生了突变（图2C）。与这一发现相一致的是，序列爆破分析表明，除了SARS-CoV-2之外，没有已知的冠状病毒与ZC45/ZXC21在E蛋白上具有100%的氨基酸序列同一性（排除了当前大流行开始后发表的可疑冠状病毒）。虽然在SARS-CoV和某些与SARS相关的蝙蝠冠状病毒之间观察到了E蛋白的100%同一性，但没有任何一对病毒同时在Orf8蛋白上有超过83%的同一性。因此，SARS-CoV-2和ZC45/ZXC21之间在Orf8蛋白上94.2%的同一性，E蛋白上100%的同一性，以及整体基因组/氨基酸水平的相似性是非常不寻常的。综合考虑这些证据，符合SARS-CoV-2基因组的起源

是以ZC45/ZXC21为骨架和/或模板进行基因功能增强修改的假说。

重要的是，ZC45和ZXC21是由第三军医大学（中国重庆）和南京指挥部医学研究所（中国南京）的军事研究实验室发现的（2015年7月至2017年2月之间）、分离并鉴定的蝙蝠冠状病毒。数据和相关工作已于2018年发表。显然，这种骨干/模板，对于SARS-CoV-2的研制至关重要，并存在于这些实验室和其他相关研究实验室中。

进一步加强我们论点的是已公布的RaTG13病毒，据报道其基因组序列与SARS-CoV-2病毒有96%的同一性。RaTG13病毒虽然表明了SARS-CoV-2的自然起源理论，但也转移了科学界和公众对ZC45/ZXC21的注意力。事实上，中国BSL-3实验室（上海公共卫生临床中心）发表了一篇《自然》杂志的文章，报道SARS-CoV-2与ZC45/ZXC21而不是与RaTG13之间存在矛盾的密切系统发育关系，但该实验室很快就被关闭"整改"。据信，该实验室的研究人员因披露SARS-CoV-2-ZC45/ZXC21之间的联系而受到惩罚。另一方面，我们已经积累了大量的证据，指出与报告的RaTG13序列有关的严重问题，以及质疑这种蝙蝠病毒在自然界的实际存在。最近发表的一篇文章也表明，RaTG13的突刺蛋白的受体结合域(RBD)不能与两种不同类型的马蹄蝠的ACE2结合(它们与RaTG13所谓的天然宿主马蹄蝠*R. affinis*密切相关)，这意味着RaTG13不能感染马蹄蝠。这一发现进一步证实了所报道的RaTG13序列可能是伪造的，因为由该序列编码的突刺蛋白似乎并不具备所声称的功能。捏造病毒以转移人们对ZC45/ZXC21的关注，这一事实说明了ZC45/ZXC21在SARS-CoV-2的研制中的实际作用。

**1.2 SARS-CoV-2突刺蛋白的受体结合基序不能从自然界中诞生，而是应通过基因工程研制出来**

突刺蛋白装饰着冠状病毒粒子的外部。它们在感染人体中起着重要的作用，因为它们介导与宿主细胞受体的相互作用，从而帮助确定病毒的宿主范围和组织趋化性。突刺蛋白分为两半（图3）。前面或N-末端的一半被命名为S1，它完全负责与宿主受体结合。在SARS-CoV和SARS-CoV-2感染中，宿主细胞受体都是hACE2。在S1内，有一段约70个氨基酸的片段与hACE2直接接触，相应地被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hACE2的相互作用。突刺蛋白的C端一半被命名为S2。S2的主要功能包括维持三聚体的形成，并在S1/S2交界处和下游S2'位置的蛋白酶连续裂解后，介导膜融合，使病毒进入细胞。

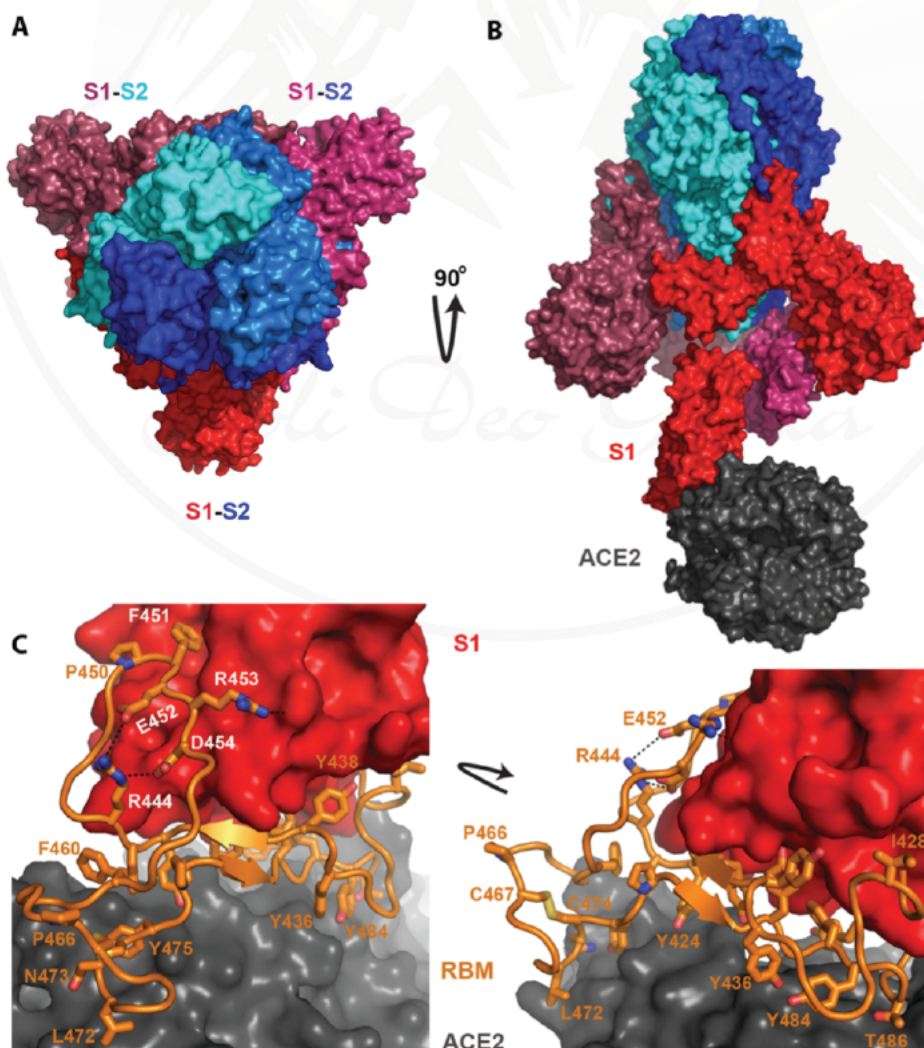


图3、SARS突刺蛋白的结构以及它如何与hACE2受体结合。图片是根据PDB ID: 6acj生成的。A) 三个突刺蛋白，每个由一个S1半和一个S2半组成，形成一个三聚体。B) S2一半（蓝色的阴影）负责三聚体的形成，而S1部分（红色的阴影）负责结合hACE2（深灰色）。C) S1和hACE2之间的结合细节。S1的RBM，是进行充分结合的关键，标为橙色。RBM内对hACE2相互作用或蛋白质折叠都很重要的残基被显示为棒状物（残基数遵循SARS突刺序列）。

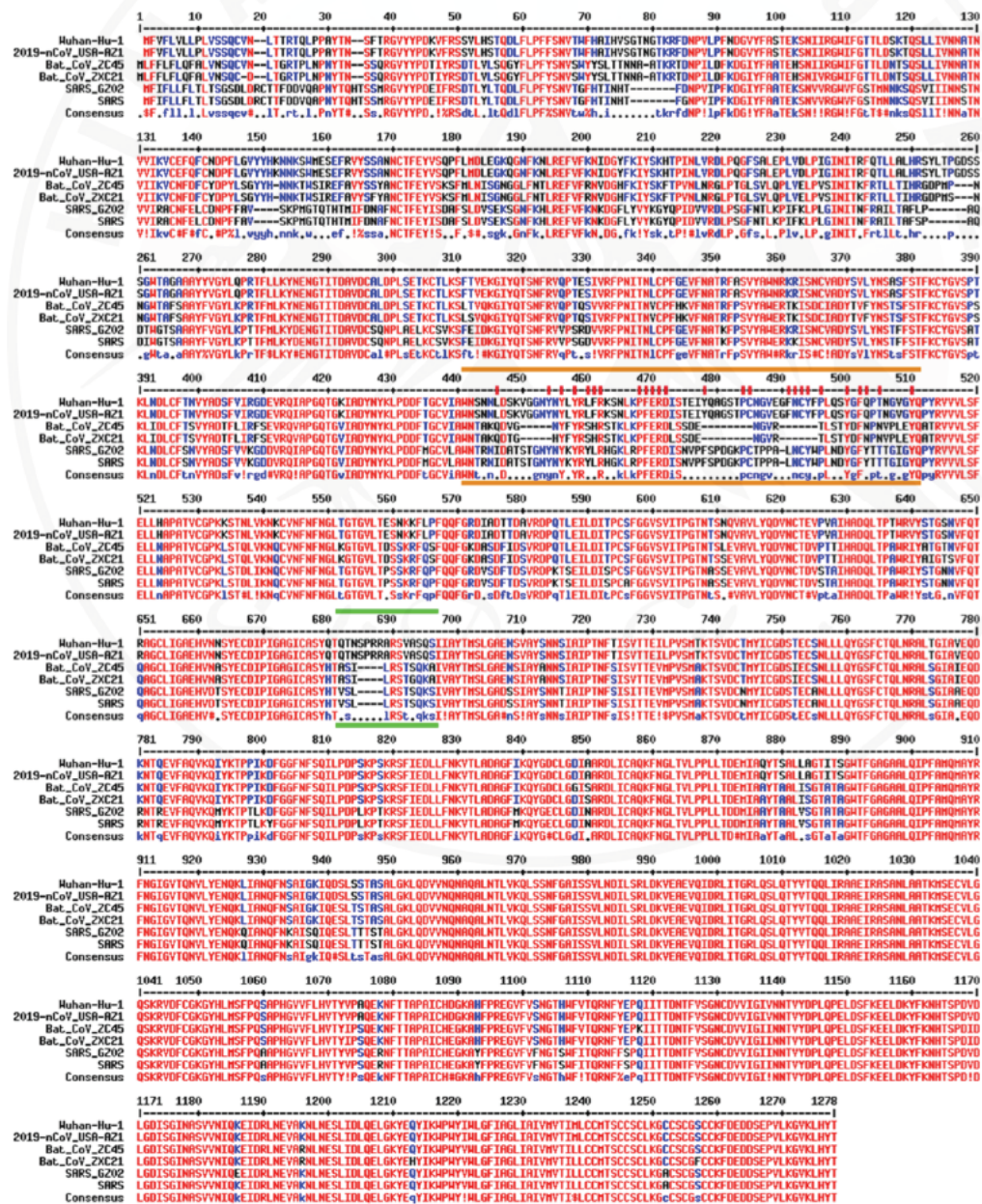


图4、相关冠状病毒的突刺蛋白序列排列。被比较的病毒包括 SARS-CoV-2（武汉-Hu-1: NC\_045512, 2019-nCoV\_USA-AZ1: MN997409）、蝙蝠冠状病毒（Bat\_CoV\_ZC45: MG772933, Bat\_CoV\_ZXC21: MG772934）和SARS冠状病毒（SARS\_GZ02: AY390556, SARS: NC\_004718.3）。由两条橙色线标记的区域是受体结合基序(RBM)，这对与hACE2受体的相互作用很重要。关键的残基在顶部另外用红色小棍突出显示。由两条绿线标记的区域是一个福林酶切位点，该位点只存在于SARS-CoV-2中，而不存在于任何其他B系β冠状病毒中。

与其他观察到的病毒蛋白的情况相似，SARS-CoV-2的S2与ZC45/ZXC21的S2具有很高的序列同一性（95%）。与此形成鲜明对比的是，SARS-CoV-2和ZC45/ZXC21之间，决定病毒可以感染哪种宿主（人类或蝙蝠）的S1蛋白的保守性要低得多，其氨基酸序列同一性只有69%。

图4显示了6种β冠状病毒的突刺蛋白的序列排列。两种是本次大流行中分离出来的病毒（武汉-Hu-1, 2019-nCoV\_USA-AZ1）；两种是疑似模板病毒（Bat\_CoV\_ZC45, Bat\_CoV\_ZXC21）；两种是SARS冠状病毒（SARS\_GZ02, SARS）。RBM在两条橙色线条之间突出显示。显然，尽管整体基因组的序列同一性很高，但SARS-CoV-2的RBM与ZC45和ZXC21的RBM有显著差异。耐人寻味的是，SARS-CoV-2的RBM与SARS突刺蛋白的RBM非常相似。虽然这不是一个精确的“复制和粘贴”，但仔细检查突刺蛋白-hACE2结构显示，所有对hACE2结合或蛋白质折叠（图3C中的橙色小棍和图4中的红色短线所强调的内容）至关重要的残基都被“保留”下来了。这些必要的残基大部分被精确地保留下来，包括那些参与二硫键形成（C467, C474）和静电相互作用（R444, E452, R453, D454），这是RBM的结构完整性的关键（图3C和4）。关键残基组内的少数变化几乎都是疏水性的“替代”（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。同时，大部分非关键的氨基酸残基都发生了"突变"(图4，RBM残基未用短红线标注)。仅从这个序列分析来看，我们很早就确信SARS-CoV-2突刺蛋白不仅会与hACE2结合，而且其结合方式也恰恰类似于原来SARS突刺蛋白与hACE2之间的结合方式。最近的结构性分析工作证实了我们的预测。

如下文所述，SARS-CoV-2 RBM与SARS-CoV RBM相似的方式以及SARS-CoV-2与ZC45/ZXC21之间的整体序列保守模式是非常不寻常的。综上所述，这表明SARS-CoV-2基因组的部分内容并非来自自然的准种病毒颗粒进化。

如果SARS-CoV-2确实来自自然进化，那么它的RBM只能通过两种可能的途径之一获得。1) 古老的重组事件，然后是趋同进化；2) 最近发生的自然重组事件。

在第一种情况下，SARS-CoV-2的祖先，一种类似ZC45/ZXC21的蝙蝠冠状病毒会与一种携带相对"完整"RBM的冠状病毒（参照SARS）进行重组和"交换"。这种重组将导致一种新型的ZC45/ZXC21样冠状病毒，其RBM中的所有空白都被"填补"了（图4）。随后，该病毒将不得不在其新的宿主中进行广泛的适应，其中ACE2蛋白与hACE2高度同源。整个基因组的随机突变将不得不发生，最终将RBM塑造为其当前的形式—以一种高度智能的方式类似于SARS-CoV的RBM。然而，这种趋同的进化过程也会导致基因组其他部分大量突变的积累，使得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序列同一性(94-100%同一性)并不支持这种情况，因此，能清楚地表明SARS-CoV-2携带的这种RBM不可能通过这种趋同进化途径来自于类似ZC45/ZXC21的蝙蝠冠状病毒。

在第二种情况下，ZC45/ZXC21类冠状病毒必须在最近与另一种冠状病毒进行重组和交换RBM，这种冠状病毒已经成功地适应

了与hACE2高度同源的动物ACE2的结合。这种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显著的序列相似性；(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) 重组病毒最终必须变得稳定并在宿主物种内可传播。

关于最近的这种重组情况，动物源不可能是蝙蝠，因为蝙蝠的ACE2蛋白与hACE2的同源性不够，因此适应性不能产生SARS-CoV-2中所看到的RBM序列。这个动物源也不可能是人类，因为ZC45/ZXC21类的冠状病毒将无法感染人类。此外，在2019年底之前，还没有证据表明任何SARS-CoV-2或SARS-CoV-2类病毒在人群中传播。耐人寻味的是，根据最近的一项生物信息学研究，SARS-CoV-2自疫情开始以来就对人类有良好的适应性。

自然进化只剩下另外一种可能，即ZC45/ZXC21类病毒和含有SARS类RBM的冠状病毒可能在一个中间宿主中重组，其中ACE2蛋白与hACE2同源。多个实验室报道，部分从马来西亚走私到中国的巽他穿山甲携带了冠状病毒，其受体结合域

(RBD) 与SARS-CoV-2几乎完全相同。然后，他们继续提出，穿山甲很可能是SARS-CoV-2的中间宿主。然而，最近的独立报告发现这一数据存在重大缺陷。此外，与这些报告相反，2009年至2019年间，在马来西亚和沙巴州收集了十多年的巽他穿山甲样本中没有检测到冠状病毒。最近的一项研究还表明，SARS-CoV-2和报道的穿山甲冠状病毒之间共享的RBD与hACE2的结合力是穿山甲ACE2的十倍，这进一步否定了穿山甲作为可能的中间宿主。最后，一项模拟研究虽然回应了穿山甲不可能是中间宿主的观点，但也表明，在他们的研究中，没有一种动物ACE2蛋白比hACE2与SARS-CoV-2突刺蛋白表现出更有利的结合潜力。最后一项研究几乎排除了所有动物作为中间宿主的嫌疑，这与SARS-CoV-2从爆发之初就对人类有良好适应性的观察结果一致。这一点很重要，因为这些研究结果共同表明，

**SARS-CoV-2**似乎不存在中间宿主，这至少降低了中间宿主发生重组事件的可能性。

即使我们忽略上述证据，即不存在合适的宿主来进行重组，而假设确实存在这样的宿主，那么自然界中发生这种重组事件的可能性也几乎没有。

如上所述，如果自然重组事件是SARS-CoV-2出现的原因，那么ZC45/ZXC21类病毒和含有SARS类RBM的冠状病毒必须在同一细胞中通过交换S1/RBM进行重组，这是一种罕见的重组形式。此外，由于SARS在人类历史上只发生过一次，因此，自然界以如此智能的方式产生一种类似SARS的病毒—只在几个非关键的位点上拥有与SARS RBM不同的RBM，至少也是同样罕见的

（图4）。这种独特的SARS类冠状病毒与ZC45/ZXC21类的祖先病毒驻留在同一细胞中，两种病毒以“RBM交换”的方式进行重组的可能性极低。重要的是，这一点，以及下面1.3节中描述的其他重组事件（在自然界中更不可能发生），且都必须发生在SARS-CoV-2中所看到的突刺蛋白。

上述证据和分析综合起来似乎不赞成SARS-CoV-2的RBM有自然来源，且有大量文献表明，功能增强研究，即对冠状病毒的突刺蛋白进行特异性工程改造，已经多次成功地从非人类来源的冠状病毒中产生了感染人类的冠状病毒。

记录还显示，武汉病毒研究所（WIV）等研究实验室与美国研究人员合作（也有单独的研究工作）已经成功开展了此类研究。此外，武汉病毒所还从事了几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，对于WIV和其他相关实验室来说，开展并成功进行这种突刺蛋白/RBM工程和功能增强实验，技术上的障碍是不存在的。



图5、SARS-CoV-2的RBM两端都有两个限制性位点，为替换突刺蛋白基因内的RBM提供了方便。A) SARS-CoV-2 (武汉-Hu-1) 的RBM的核苷酸序列。在RBM的5'-端发现一个EcoRI位点，在3'-端发现一个BstEII位点。B) 虽然这两个限制性位点不存在于ZC45的原始突刺蛋白基因中，但考虑到这两个位点的序列差异较小(2个核苷酸)，它们可以被方便地引入。C) 氨基酸序列排列与RBM区域突出显示 (标色和下划线)。橙色 (上) 的RBM是由SARS-CoV-2 (武汉-Hu-1) 突刺蛋白中的EcoRI和BstEII位点定义的。品红色(中间)的RBM是李芳博士及其同事换成SARS突刺蛋白骨架的区域。蓝色 (下) 高亮的RBM来自SARS-BJ01 (AY278488.2) 的突刺蛋白 (RBM: 424-494)，该区域被石正丽实验室置换成不同蝙蝠冠状病毒的突刺蛋白，从而替换了相应的片段。

值得注意的是，与RBM基因工程理论一致，我们在SARS-CoV-2基因组RBM的两端分别发现了两个独特的限制性位点EcoRI和BstEII (图5A)。这两个位点，是日常分子克隆的热门选择，但在这个突刺蛋白基因的其他部位并不存在。这种特殊的设置使得在突刺蛋白内交换RBM极为方便，为测试不同的RBM和相应的突刺蛋白提供了一种快速的方法。

这样的EcoRI和BstEII位点在其他β冠状病毒的突刺蛋白基因中并不存在，这强烈地表明它们不是自然的，是为了方便操作关键的RBM而专门引入SARS-CoV-2的这个突刺蛋白基因中的。虽然ZC45突刺蛋白也没有这两个位点(图5B)，但如本报告第2部分所述，它们可以非常容易被引入。

值得注意的是，这里引入EcoRI位点会将相应的氨基酸从-WNT-变为-WNS-（图5AB）。据我们所知，所有SARS和类似SARS的蝙蝠冠状病毒专门在这个位置携带一个T（苏氨酸）残基。SARS-CoV-2是唯一的例外，因为这个T已经突变为S（丝氨酸），除了疫情爆发后公布的可疑的RaTG13和穿山甲冠状病毒外。

一旦限制性位点被成功地引入，RBM片段可以使用常规限制性内切酶和连接反应进行交换。虽然替代性克隆技术可能不会留下任何遗传操作的痕迹（吉布森组装作为一个例子），这种老式的方法可以被选择，因为它在交换这个关键的RBM中提供了极大的方便。

鉴于RBM完全决定hACE2的结合度，而且SARS RBM-hACE2的结合完全由高分辨率的结构来表征（图3），这种单纯的RBM交换不会比完全的突刺蛋白交换更有风险。事实上，这种RBM交换策略的可行性已经得到了证明。2008年，石正丽博士的研究小组在将一个限制性位点引入密码子优化的突刺蛋白基因后，将SARS RBM交换到几种SARS类蝙蝠冠状病毒的突刺蛋白中（图5C）。然后，他们验证了所得到的嵌合突刺蛋白与hACE2的结合。此外，在最近的一份出版物中，SARS-CoV-2的RBM被交换到SARS-CoV的受体结合域（RBD）中，导致嵌合

RBD在结合hACE2方面完全发挥了作用（图5C）。令人注意的是，在这两种情况下，人为操纵的RBM片段几乎完全类似于由EcoRI和BstEII位点的位置定义的RBM（图5C）。虽然这两份出版物都缺乏克隆细节的描述，但可以想象，实际的限制性位点可能会有所不同，这取决于接受RBM插入的突刺蛋白基因，以及在感兴趣的区域引入独特的限制性位点的便利性。值得一提的是，最近发表的这篇论文的对应作者李芳博士从2010年开始就一直是石正丽博士的积极合作者。李博士是世界上第一个从结构上阐明SARS-CoV RBD与hACE238结合的人，也是在结构上理解突刺蛋白-ACE2相互作用的主要专家。在SARS-CoV-2 RBM的两端关于EcoRI和BstEII限制性位点的惊人发现，以及同一RBM区域分别被石博士和她的长期合作者用限制性酶切方法进行了交换的事实，都不太可能是巧合。相反，它是证明SARS-CoV-2的RBM/突刺蛋白是基因操纵产物的实锤证据。

虽然复制SARS RBM的确切序列可能很方便，但这太明显是人为设计和操纵的迹象。更具欺骗性的方法是改变一些非关键的残基，而保留那些关键的结合。高分辨率结构可以很好地指导这种设计（图3）。这样一来，当RBM的整体序列看起来与SARS RBM的序列更加不同时，hACE2的结合能力将得到很好的保留。我们认为，所有的关键残基(图4中用红色棍子标记的残基，与图3C中小棍所示的残基相同)都应该已被"保留"。如前所述，虽然有些应该是直接保留，但有些应该被替换成具有类似性质的残基，这就不会破坏hACE2的结合力，甚至可能进一步加强了结合功能。重要的是，可能在非关键的位点上有意做了改变，使之不像SARS RBM的“复制粘贴”那样明显。

### 1.3 SARS-CoV-2的突刺蛋白中存在一个不寻常的福林酶切位点，与病毒的毒性增强有关

SARS-CoV-2的突刺蛋白中另一个独特的基序是位于S1/S2交界处的多碱基福林酶切位点（图4，两条绿线之间的片段）。这样的位点可以被福林蛋白酶识别和切割。在β冠状病毒的B系内，除了SARS-CoV-2外，没有病毒在S1/S2交界处含有福林酶切位点（图6）。与此相反，在其他组冠状病毒中也观察到这个位置的福林酶切位点。似乎存在一定的选择性压力，使β冠状病毒的B系无法在自然界中获得或维持这样的位点。

Human SARS-CoV BJ01	655 - GICASYHTVSL---RSTS - 670
Human SARS-CoV CUHK-W1	655 - GICASYHTVSL---RSTS - 670
Human SARS-CoV Tor2	655 - GICASYHTVSL---RSTS - 670
Human SARS-CoV Frankfurt-1	655 - GICASYHTVSL---RSTS - 670
Human SARS-CoV Urbani	655 - GICASYHTVSL---RSTS - 670
Civet SARS-CoV civet020	655 - GICASYHTVSSL---RSTS - 670
Civet SARS-CoV sz16	655 - GICASYHTVSSL---RSTS - 670
Raccoon dog SARS-CoV A030	655 - GICASYHTVSSL---RSTS - 670
<b>SARS-CoV-2</b>	669 - GICASYQTQNS <b>PRR</b> ARVA - 688
Pangolin CoV MP789	n/a - GICASYQTQNS---RSVS - n/a
Bat SARSr-CoV RaTG13	669 - GICASYQTQNS---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 JTM15	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、在突刺蛋白的S1/S2交界处发现的福林酶切位点是SARS-CoV-2所独有的，在其他B系β冠状病毒中不存在。图片转载自Hoffmann等。

如前所述，在进入细胞过程中，突刺蛋白首先在S1/S2交界处被切割。这一步骤以及随后在下游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，并假定也影响了病毒感染的致病性。与其他蛋白酶不同，福林蛋白酶在许多类型的细胞中广泛表达，并存在于多个细胞和细胞外位置。重要的是，在S1/S2交界处引入一个福林蛋白酶切位点，可以显著增强病毒的感染力，以及大大扩展其细胞趋性—这一现象在流感病毒和其他冠状病毒中都有很好的记录。

如果我们撇开自然界中任何B系 $\beta$ 冠状病毒中都没有发现福林酶切位点的事实，而假设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，即福林酶切位点必须来自于同源重组事件。具体来说，一个不含福林酶切位点的祖先 $\beta$ 冠状病毒必须与一个含有福林酶切位点的近亲冠状病毒重组。

然而，有两个事实不利于这种可能性。首先，尽管一些来自其他组或系的冠状病毒确实含有多基福林酶切位点，但没有一种冠状病毒含有SARS-CoV-2确切的多基序列（-PRRAR/SVA-）。其次，在SARS-CoV-2和任何含有合理福林蛋白酶切位点的冠状病毒之间，突刺蛋白上的序列同一性不超过40%。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** 突刺蛋白中的福林酶切位点不可能源自自然，而应该是实验室改造的结果。

与这一说法相一致的是，对SARS-CoV-2突刺蛋白中福林酶切位点的核苷酸序列进行仔细检查后发现，插入序列内的两个连续的Arg残基(-PRRA-)均由罕见的密码子CGG（SARS-CoV-2中Arg使用最少的密码子）编码(图7)。事实上，这种CGGCGG排列是在SARS-CoV-2基因组中发现的唯一实例，这种罕见的密码子被串联使用。这一观察结果强烈地表明，这个福林酶切位点

应该是基因工程的结果。更加让人怀疑的是，这里的密码子选择制定了一个FauI限制性位点，这说明限制性片段长度多态性（WIV实验室精通的技术）有可能涉及其中。在那里，FauI切位所产生的片段模式可以用来监测突刺蛋白中福林酶切位点的保存情况，因为这个福林酶切位点在体外容易发生缺失。具体来说，可以对从细胞培养物或实验室动物中回收的病毒的突刺蛋白基因进行RT-PCR检测，其产物要经过FauI切位。保留或失去福林酶切位点的病毒将产生不同的模式，从而可以方便地跟踪要研究的病毒。

FauI

```

tat cag act cag act aat tct cct cgg cgg gca cgt agt gta gct agt caa tcc atc att
Y  Q  T  Q  T  N  S  P  R  R  A  R  S  V  A  S  Q  S  I  I
  
```

图7、在SARS-CoV-2 突刺蛋白的S1/S2交界处的-PRRA-插入的两个连续的Arg残基都由一个罕见的密码子CGG编码。一个FauI限制性位点，5'-(N)6GCGGG-3'，嵌入在"插入的"PRRA段的编码序列中，它可作为监测引入的福林酶切位点保存情况的标记。

此外，虽然目前还没有已知的冠状病毒含有SARS-CoV-2 突刺蛋白中确切的-PRRAR/SVA-序列，但在2017年石正丽博士发表的啮齿类动物冠状病毒AcCoV-JC34的突刺蛋白的S1/S2交界处观察到了类似的-RRAR/AR-序列。可见，-RRAR-作为功能性福林酶切位点的合理性自2017年起就已被WIV专家所认知。

相关证据共同表明，SARS-CoV-2突刺蛋白中的福林酶切位点可能不是来自自然界，且可能是基因操纵的结果。这种操作的目的是可能是评估实验室研制的冠状病毒的传染性和致病性的任何潜在增强作用。事实上，最近的研究已经证实，福林酶切位点确实赋予了SARS-CoV-2显著的致病性优势

## 1.4 小结

这一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/或模板产生的。突刺蛋白，特别是其中的RBM，应该是经过人工操作的，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切位点也支持了这一点。一个不寻常的福林酶切位点可能已经被引入并插入突刺蛋白的S1/S2交界处，这有助于增加病毒的毒性和致病性。这些转化就使SARS-CoV-2病毒分期最终成为一种高传染性、发病隐匿性、致死性、后遗症不明确、破坏性大的病原体。

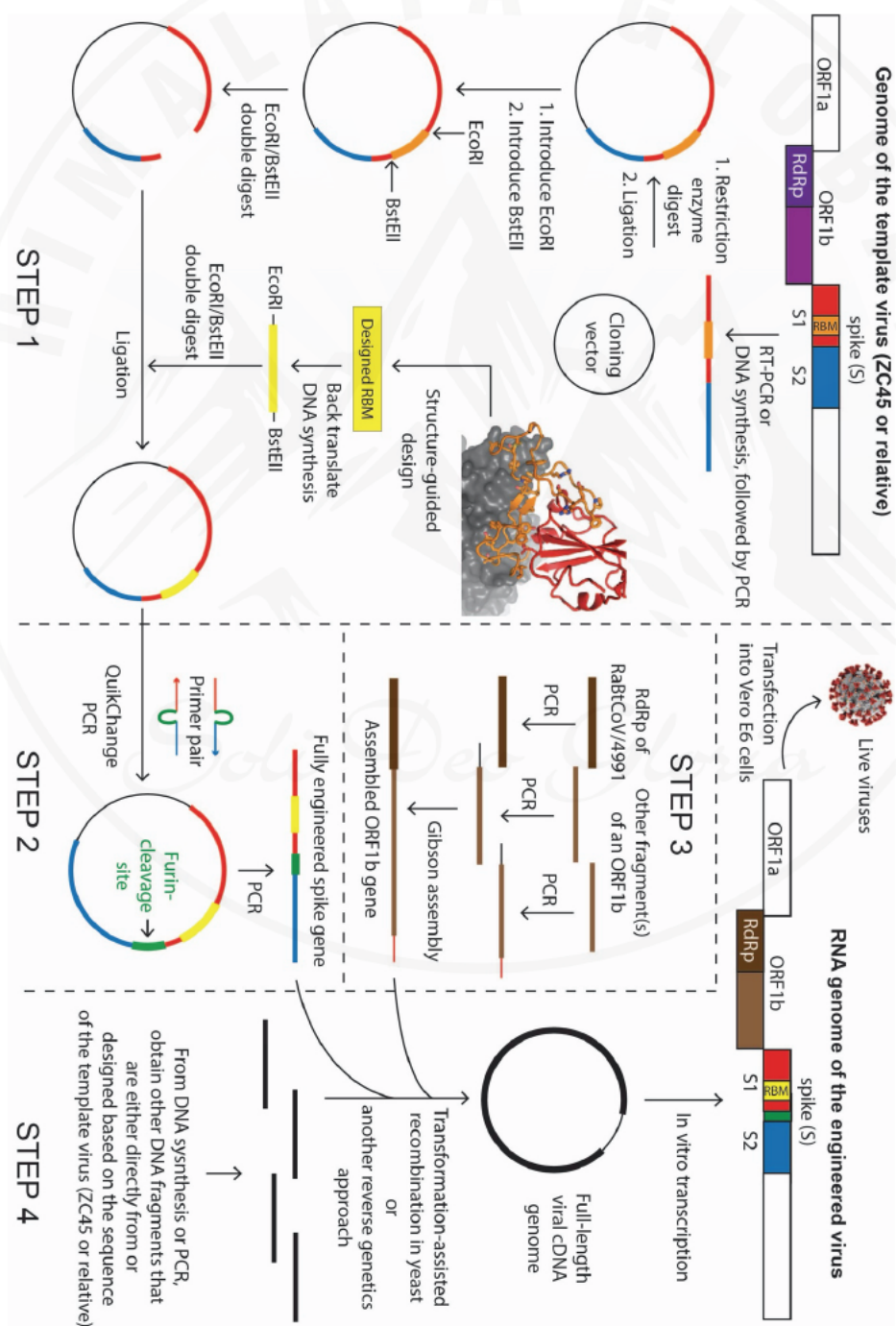
显然，SARS-CoV-2可能是通过WIV的功能增强操作产生的，这种可能性很大，应该进行彻底和独立的调查。

## 2. 确定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研制SARS-CoV-2的合成路径。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假定的。虽然这里提出的步骤不应该被视为需完全采取的步骤，我们相信，关键过程应没有太大的不同。重要的是，我们在这里的工作应该作为一个示范，证明如何遵循公认的概念并使用公认的技术，可在研究实验室中方便地设计和研制SARS-CoV-2。

无论是在香港还是在中国大陆，相关实验室在冠状病毒研究方面，无论是在资源还是研究产出，都处于世界领先地位。后者不仅体现在过去20年里他们发表的大量论文上，还体现在他们在这一领域取得的里程碑式的成就上。他们是第一个发现麝猫是SARS-CoV的中间宿主，并分离出第一株病毒；他们是第一个发现SARS-CoV来源于蝙蝠；他们首次揭示了SARS-CoV感染

的抗体依赖性增强（ADE）；他们在了解MERS的各个领域（人畜共患病、病毒学和临床研究）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。最后但并非最不重要的是，他们拥有世界上最大的冠状病毒收藏库（基因组序列和活体病毒）。香港和大陆的研究实验室都有现成的知识、专业技术和资源（他们广泛合作）来开展和完成下面描述的工作。



## 2.1 新型冠状病毒实验室研制设计的可能方案

在本小节中，我们概述了项目设计阶段可能制定的整体策略和主要考虑因素。

要想设计和研制一个针对人类的冠状病毒，他们就必须选择一种蝙蝠冠状病毒作为模板/骨干。这可以很方便地完成，因为在过去的20年里，许多研究实验室一直在积极收集蝙蝠冠状病毒。然而，考虑到石正丽博士一直从事冠状病毒的功能增强研究，这个模板病毒最好不是石正丽博士收藏的病毒。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将适合作为模板/骨干病毒。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，这些新型蝙蝠冠状病毒是由军事实验室发现并拥有的。因此，实际的模板可能是ZC45、或ZXC21、或它们的近亲病毒。无论三者中哪一个实际的模板，下面描述的假定途径都是一样的。

一旦他们选择了模板病毒，他们首先需要通过分子克隆，对突刺蛋白进行工程设计，使其能够与hACE2结合。这种操作所涉及的概念和克隆技术已经在文献中得到了充分的记载。在几乎没有失败风险的情况下，模板蝙蝠病毒就可以转化为可以结合hACE2并感染人类的冠状病毒。

其次，他们利用分子克隆在突刺蛋白的S1/S2交界处引入一个福林酶切位点。根据已知的知识，这种操作很可能产生一株传染性和致病性更强的冠状病毒。

第三，他们制造出一个ORF1b基因构建。ORF1b基因编码多聚蛋白Orf1b，它经过后转译加工产生单个病毒蛋白：RNA依赖的RNA聚合酶（RdRp）、解旋酶、胍基-N7甲基转移酶、尿苷酸特异的核糖核酸内切酶和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蛋白是最关键的一个，在冠状病毒中高度保守。重要的是，石正丽博士的实验室采用PCR方案，将RdRp基因的特定片段进行扩增，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。由于这种做法，石博士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这个短段的序列信息。

在这里，由于Orf1b是保守性的，而且很可能来自任何 $\beta$ 冠状病毒的Orf1b都足以胜任这项工作，所以基因操作的要求不高，也不复杂。然而，我们相信，他们会希望将特定的Orf1b引入病毒中，原因可能有两种。

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，因此，在基因组中使用不同的RdRp可以确保在系统发育研究中把SARS-CoV-2和ZC45/ZXC21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，唯一公开的信息是其短RdRp片段的序列，而从未报道过其完整的基因组序列或病毒分离史。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。RdRp序列的微小变化可以在一开始就引入（通过DNA合成），也可以在以后通过传代产生。在另一条研究主线上，当他们从事RaTG13序列的研制时，他们可以从RaBtCoV/4991的RdRp短段入手，而不对其序列引入任何改变，从而使这两种病毒在这个RdRp短段上的核苷酸序列100%一致。这样这个RaTG13病毒就可以宣称早在2013年就被发现了。

2. 来自RaBatCoV/4991的RdRp蛋白的独特之处在于，它比任何其他 $\beta$ 冠状病毒的RdRp在开发抗病毒药物方面更有优势。RdRp在人体细胞中没有同源，这使得这种重要的病毒酶成为抗病毒开发的非常理想的靶点。以目前正在进行临床试验的瑞德西韦（Remdesivir）为例，其靶点就是RdRp。在研制一种新型的、针对人类的病毒时，他们也会对开发解药感兴趣。即使发现这样的药物可能不容易实现，但他们有意加入更适合抗病毒药物开发的RdRp也是合理的。

第四，他们利用反向遗传学将突刺蛋白的基因片段、ORF1b和模板ZC45的其余部分组装成病毒基因组的cDNA版本。然后他们会进行体外转录，得到病毒RNA基因组。将RNA基因组转染到细胞中，就可以回收具有所需人工基因组的活病毒和传染性病毒。

第五，采用体内连续传代，他们对病毒株进行特征分析和优化，以提高病毒株的适配性、感染力和整体适应性。然后获得一个或几个符合一定标准的病毒株作为最终产品。

## 2.2 制成SARS-CoV-2的假定合成途径

在本小节中，我们更详细地描述了如何在实验室环境中利用现有材料和常规的分子、细胞和病毒学技术进行每一个步骤。这个过程的示意图如图8所示。我们估计，整个过程可以在大约6个月内完成。

步骤1：为hACE2结合的突刺蛋白设计RBM（1.5个月）。

蝙蝠冠状病毒的突刺蛋白由于其RBM内重要残基的缺失而不能或不能有效结合hACE2。这可以由模板病毒ZC45的RBM（图4）为例。在SARS-CoV-2的研制中，第一个也是最关键的步骤

是对突刺蛋白进行工程设计，使其获得结合hACE2的能力。有文献证明，自2008年以来，研究实验室已经反复进行了这样的操作，成功地产生了具有感染人类细胞能力的工程化冠状病毒。虽然有许多可能的方法可以对突刺蛋白进行工程设计，但我们相信，实际上他们所做的是以SARS的RBM为指导，用设计好的、可能经过优化的RBM替换了原来的RBM。如第一部分所述，我们观察到SARS-CoV-2基因组中RBM的两端存在两个独特的限制性位点EcoRI和BstEII（图5A），而且这种RBM交换已经由石正丽博士和她的长期合作者、结构生物学专家李芳博士成功地进行了。

虽然ZC45突刺蛋白不包含这两个限制性位点(图5B)，但它们可以非常容易地被引入。原始的突刺蛋白基因将用RT-PCR扩增或通过DNA合成获得（一些变化可以安全地引入序列的某些可变区域），然后进行PCR。之后使用EcoRI和BstEII以外的限制性位点将该基因克隆到质粒中。一旦进入质粒，可以很容易地修改突刺基因。首先，可以通过将高亮的 "gaacac "序列（图5B）转换为所需的 "gaattc "来引入EcoRI位点（图5A）。它们之间的区别是两个连续的核苷酸。使用市售的QuikChange位点定向诱变试剂盒，这样的二核苷酸突变可以在不超过一周的时间内产生。随后，BstEII位点可以类似地引入在RBM的另一端。具体来说，"gaatacc "序列(图5B)将被转换为所需的 "ggttacc"(图5A)，这同样需要一周的时间。

一旦这些限制性位点（这是在SARS-CoV-2的突刺蛋白基因内独有的）被成功地引入，不同的RBM段可以方便地进行交换，由此产生的突刺蛋白随后使用既定的检测方法进行评估。

如第一部分所述，RBM段的设计可以通过高分辨率的结构（图3）得到很好的指导，产生一个类似于SARS RBM的智能序列。当进行结构指导下的RBM设计时，他们会按照常规生成几个（例如十几个）这样的RBM，希望一些特定的变体（s）在结合

hACE2方面可能优于其他变体。一旦设计完成，他们可以让每个设计好的RBM基因进行商业化合成（快速且非常经济实惠），在5'-端有一个EcoRI位点，在3'-端有一个BstEII位点。然后，这些新型RBM基因可以分别克隆到突刺蛋白基因中。基因合成和随后的克隆，可以以批量模式完成设计的小型RBM库，这大约需要一个月的时间。

然后，这些工程化的突刺蛋白可能会使用既定的伪型病毒感染试验来测试hACE2的结合力。具有良好至异常优越结合力的工程化的突刺蛋白将被选择。（虽然不是必要的，但可以在这里进行定向进化(在RBM基因上进行容易出错的PCR)，再加上体外结合试验或伪型病毒感染试验，以获得与hACE2结合的具有特殊亲和力的RBM。）

鉴于关于突刺蛋白工程的丰富文献和现有的突刺蛋白-hACE2复合物的高分辨率结构，这一步的成功将得到很大的保证。在这一步骤结束时，可如愿以偿获得一个新的突刺蛋白基因，它编码了一个新的突刺蛋白能够以高亲和力结合hACE2。

步骤2：在S1/S2交界处设计一个福林酶切位点(0.5个月)。

步骤1的产物，含有工程突刺蛋白的质粒，将被进一步修改，以包括在S1/S2交界处的福林酶切位点（在图4中由绿线表示的段）。这段短的基因序列可以使用几个常规克隆技术方便地插入，包括QuikChange位点直接PCR，重叠PCR，然后限制性酶切和连接，或吉布森组装。这些技术都不会在序列中留下任何痕迹。无论选择哪种克隆方法，插入的基因片都将包括在引物中，引物将被设计、合成并用于克隆。这一步，导致在S1/S2交界处添加了福林酶切位点的突刺蛋白得到进一步修改，可以在不超过两周的时间内完成。

步骤3：获得包含RaBtCoV/4991短RdRp段序列的ORF1b基因（1个月，但可与步骤1和2同时进行）。

与突刺蛋白的工程不同，这里不需要复杂的设计，除了需要包括RaBtCoV/4991的RdRp基因段。这里可以使用吉布森组装。在这种技术中，几个片段，每个相邻的对共享20-40 bp重叠，在一个简单的反应中组合在一起，以组装成一个长的DNA产物。两个或三个片段，每个覆盖ORF1b基因的重要部分，将根据已知的蝙蝠冠状病毒序列进行选择。其中一个片段将是RaBtCoV/4991的RdRp片段。每个片段将与引物中引入的适当的重叠区域进行PCR扩增。最后，所有纯化的片段将以等摩尔浓度汇集，并加入到吉布森反应混合物中，经过短暂的孵育，将产生所需的ORF1b基因整体。

步骤4：利用反向遗传学制作设计病毒基因组，并回收活病毒（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。最近的例子是利用酵母中的转化辅助重组法重建SARS-CoV-2基因组。利用这种方法，瑞士科研小组组装了整个病毒基因组，并在短短一周内产生了活病毒。这种高效的技术，不会在创建的病毒基因组中留下任何人工操作的痕迹，自2017年起就可以使用。除了工程化的突刺基因（来自步骤1和2）和ORF1b基因（来自步骤3），覆盖基因组其余部分的其他片段将通过从模板病毒的RT-PCR扩增或通过遵循从模板病毒的序列略微改变的DNA合成获得。我们认为，后一种方法更有可能，因为它可以将序列变化引入到保守性较差的蛋白质的可变区域，其过程可以很容易地被多序列排列所引导。更保守功能的氨基酸序列，如E蛋白的氨基酸序列，可能会被保持不变。然后，所有的DNA片段将被汇集在一起，并转化到酵母中，在那里，SARS-CoV-2基因组的cDNA版本将通过转化辅助重组进行组装。当然，也可以采用另一种反向遗传学的方法，WIV过去已

经成功地使用了这种方法。虽然一些早期的反向遗传学方法可能会在不同片段连接的地方留下限制性位点，但这些痕迹将很难被检测到，因为连接的确切位置可以在约30kb基因组的任何地方。无论使用哪种方式，病毒基因组的cDNA版本将从反向遗传学实验中获得。随后，以该cDNA为模板进行体外转录，将产生病毒RNA基因组，转染到Vero E6细胞中后，可以生产具有所有设计特性的活病毒。

#### 步骤5：优化病毒的适应性，提高病毒在体内的hACE2结合亲和力(2.5-3个月)

从第4步中回收的病毒需要经过经典的实验—实验室动物的连续传代—进行进一步调整。这最后一步将验证病毒的适应性，并确保其面向受体的适应性，以适应其预定的宿主，根据上述分析，宿主应该是人类。重要的是，RBM和福林酶切位点，被分别引入到突刺蛋白中，现在将作为一个功能单元一起优化。在各种可用的冠状病毒动物模型(如小鼠、仓鼠、雪貂和猴子)中，hACE2转基因小鼠(hACE2-mice)应该是这里最合适、最方便的选择。这种动物模型在SARS-CoV的研究过程中已经建立起来了，并且在杰克逊实验室已经有多年的使用经验。

连续传代的程序是简单直接的。简而言之，从步骤4中选择的病毒株，SARS-CoV-2的前体，将被鼻内接种到一组麻醉的hACE2-小鼠中。感染后2-3天左右，肺部的病毒通常会放大到一个峰值滴度。然后，小鼠将死去，其肺部最后均质化。通常，携带病毒量最高的鼠肺上清液将被用于提取候选病毒，以备下一轮传代。经过大约10~15轮的传代，病毒株的hACE2结合亲和力、感染效率和致死率都会得到充分的提高，病毒基因组也会趋于稳定。最后，经过一系列的表征实验（如病毒动力学实验、抗体反应实验、症状观察和病理检查等），得到最终产品SARS-CoV-2，从而结束了整个创制过程。从此，这种病毒病原体可以被扩增(很可能使用Vero E6细胞)，并进行常规生产。

值得注意的是，根据对SARS-CoV的研究，hACE2小鼠虽然适合SARS-CoV-2的适应性，但并不是反映病毒在人类中的传播性和相关临床症状的良好模型。我们认为，在COVID-19爆发前，那些科学家可能没有使用合适的动物模型（如叙利亚仓鼠）来测试SARS-CoV-2的传播性。如果他们使用适当的动物模型进行实验，SARS-CoV-2的高传染性就会非常明显，因此SARS-CoV-2不会在疫情爆发之初被描述为“不会造成人与人之间的传播”。

我们还假定，针对提高传播能力和杀伤力的广泛实验室适应措施可能会将病毒逼得更加强大。因此，SARS-CoV-2在目前对人类的适应过程中，可能已经失去了在传播性和致死性上弱化的能力。这一假设与SARS-CoV-2尽管在大流行，但到目前为止还没有明显的减弱迹象，也与最近出现的主要变异体只表现出更好的传播性的观察相一致。

连续传代是一个快速而密集的过程，在这个过程中，病毒的适应性被加速。虽然意在模仿自然进化，但序列传代在时间和规模上都受到更多限制。因此，与自然进化相比，在序列传代过程中随机突变的情况会更少。这对于保守的病毒蛋白来说尤其如此，例如E蛋白。E蛋白在病毒复制中至关重要，是毒性的决定因素，对它进行工程化处理可能会使SARS-CoV-2减弱。因此，在最初的组装阶段，这些科学家可能决定保持E蛋白的氨基酸序列与ZC45/ZXC21的序列不变。由于E蛋白的保守性和序列传代的限制，实际上没有发生氨基酸突变，导致SARS-CoV-2和ZC45/ZXC21之间E蛋白的序列100%相同。同样的情况也可能发生在分子克隆的标记（RBM两侧的限制性位点）上。连续传代，应该使SARS-CoV-2基因组部分归化，但可能没有消除所有人工操作的痕迹。

### 3. 结束语

关于SARS-CoV-2的起源，许多问题仍然没有答案。著名病毒学家在《自然医学》的一封信中暗示，实验室逃逸虽然没有完全排除，但不太可能，而SARS-CoV-2基因组中没有基因操作的迹象。然而，我们在这里表明，SARS-CoV-2基因组的突刺蛋白基因内的遗传证据（限制性位点侧翼的RBM；在插入的福林酶切位点使用的串联稀有密码子）确实存在，并表明SARS-CoV-2基因组应该是基因操纵的产物。而且，成熟的概念、成熟的技术、知识和专业技术都已具备，可以在短时间内方便地研制出这种新型冠状病毒。

抛开动机不谈，以下关于SARS-CoV-2的事实是有充分依据的。

- 1、如果它是实验室产品，那么它产生的最关键因素—骨干/模板病毒（ZC45/ZXC21），是属于军事研究实验室的。
2. SARS-CoV-2的基因组序列很可能经过了基因工程改造，通过这个，病毒获得了以人类为目标的能力，并增强了毒性和感染力。
3. SARS-CoV-2的特点和致病作用是前所未有的。该病毒具有高度传播性、发病隐匿性、多器官靶向性、后遗症不明确、致死性强、伴有各种症状和并发症等特点。
4. SARS-CoV-2引起了世界性的大流行，夺去了数十万人的生命，使全球经济瘫痪。它具有前所未有的破坏力。

根据我们和其他人收集到的证据，我们认为，要找到SARS-CoV-2的来源，就应该对WIV的P4实验室及其密切合作者的实验室进行独立审查。这样的调查早就应该进行了，不应该再拖延下去。

我们还注意到，在2015年发表的嵌合病毒SHC015-MA15中，最初遗漏了美国国家过敏症与传染病研究所（NIAID）对石正丽的经费资助声明。或许是在2016年1月的会议上恢复了美国国立卫生研究院（NIH）对病毒的增强功能研究的经费资助，因此在2016年发表的论文勘误中恢复了此声明。这是一种不寻常的科研行为，需要一个解释。

本报告中没有彻底描述的是，各种证据表明，最近公布的几种冠状病毒（RaTG13、RmYN02和几种穿山甲冠状病毒）非常可疑，很可能是造假行为。这些造假除了欺骗科学界和公众，使SARS-CoV-2的真实身份被隐藏外，没有其他目的。虽然排除这些证据的细节并没有改变本报告的结论，但我们认为这些细节将为我们的论点提供更多支持，即**SARS-CoV-2是一种实验室强化的病毒，是功能增强研究的产物**。目前，我们正在编写一份后续报告，重点阐述这些补充证据，不久后将提交。

鸣谢

我们要感谢Zhang Daoyu与我们分享了 $\beta$ 冠状病毒不同亚群中E蛋白突变的发现。我们也感谢所有匿名的科学家和其他个人，他们为揭示SARS-CoV-2的起源做出了贡献。