JOHNS HOPKINS BLOOMBERG SCHOOL of PUBLIC HEALTH | Center for Health Security

In Response: Yan et al Preprint
Examinations of the Origin of SARS-CoV-2

September 21, 2020

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of novel coronavirus disease 2019 (COVID-19), has caused more than 961,000 known deaths[1] since it was reported to the World Health Organization on December 31, 2019. Determining the origin of the pandemic coronavirus is of great importance, not only to understand the mechanics of how the virus replicates and spreads but also to anticipate and prevent additional viruses from becoming future health security crises. If an origin can be found for SARS-CoV-2, steps can then be taken to prevent a similar pathway for other viruses to lead to a pandemic. For that reason, it is the responsibility of the scientific community to review and analyze data relating to the origin of SARS-CoV-2.

Several analyses of the potential origin of SARS-CoV-2 have been published in scientific journals that provide peer review prior to publication.[2,3,4,5,6,7,8,9] Peer review is central to the scientific process because scrutiny by experts allows for meaningful conclusions to be drawn about available data and reduces inappropriate extrapolation or misinterpretation. It is an imperfect process, often criticized for slowness, but peer review is a necessary part of building reliability in the scientific record. Complex scientific details are best understood and critiqued by others who are also experts in a technical field. When the audience for an article is broadened, even to a technical audience in an adjacent scientific field, data may appear smoother and less conflicting than it is in reality, leading to a blurring or skewing of its real meaning.

In this document, we have undertaken a scientific review of a recent report, released as a preprint put forward by the Rule of Law Society, authored by Li-Meng Yan, Shu Kang, Jie Guan, and Shanchang Hu. The report, *Unusual Features of the SARS-CoV-2 Genome Suggesting Sophisticated Laboratory Modification Rather Than Natural Evolution and Delineation of Its Probable Synthetic Route*,[10] presents a theory about the origin of SARS-CoV-2 but offers contradictory and inaccurate information that does not support their argument. As the report has not been submitted to a scientific peer-reviewed publication, which would provide the expert scrutiny expected by the scientific community and the larger public, we aim to provide an objective analysis of details included in the report, as would be customary in a peer-review process.

Specific Comments on the Report

Page 2

1. **On natural existence of a closely related virus.** Line 17: RaTG13 is a previously discovered bat coronavirus which has about a 96% sequence identity to SARS-CoV-2,[4] indicating that it is a close relative and that bats are likely involved in the evolution of SARS-CoV-2. Yan et al question the existence of RaTG13, which is found in GenBank.[11] The authors cite multiple papers in their reference section that have weaknesses or flaws to make their case. In their paper, reference

This review has been written by Kelsey Lane Warmbrod, MS, MPH; Rachel M. West, PhD; Nancy D. Connell, PhD; and Gigi Kwik Gronvall, PhD; all from the Johns Hopkins Center for Health Security.

7's author is not a scientist or researcher according to his ORCID profile; references 10 and 13 cannot be found online and the links provided are not active; reference 11 is an opinion piece on an anti-GMO interest group website; references 5, 6, 8, 9, and 12 appear to be authored by scientists lacking expertise in coronaviruses and/or viral evolution. Only 2 of these publications (14 and 15) were published in scientific journals with peer review, and none of the authors of these 2 articles specialize in coronaviruses or viral genetics.

2. **On the capacity to predict function from genotype.** Line 28: Yan et al overstate the capabilities of deducing functional changes from genetic manipulation of coronaviruses, referring to an "abundant literature indicat[ing] that gain-of-function research has long advanced to the stage where viral genomes can be precisely engineered and manipulated to enable the creation of novel coronaviruses possessing unique properties."[10] Technologies like CRISPR have enabled precise, directed gene editing, and are major advances for the biological sciences. However, the report overstates current capabilities in designing phenotypes and genetic functions of viruses, which are not already elucidated, including coronaviruses, and vastly overstates the capabilities of genetic manipulation of coronaviruses in 2019, before these viruses were the focus of worldwide interrogation by the scientific community. There were 6 coronaviruses known to infect humans prior to 2020, but their prevalence and pathology in different age groups is incompletely understood, which would hamper any potential design of novel coronavirus functions. Prior to 2020, coronaviruses were not as intensely researched as other viruses that cause human disease, such as HIV, and influenza.

3. **Lack of current evidence countering natural origin theory.** Line 27: Yan et al refer to an extensive scientific literature providing "genomic, structural, and literature evidence"[10] to counter the prevailing theory in the scientific community that the origin of SARS-CoV-2 is a natural zoonosis, emerging from animals, but they do not cite any references to support their claim—a crucial basic practice for any researcher.

Page 3

1. **On implausibility of the proposed viral genetic backbone.** Lines 19-20: Scientific evidence suggests that many coronaviruses[12] similar to SARS-CoV-2 have a recent common ancestor or that convergent evolution[13] has occurred. Many coronaviruses infect bats and other animals, most of which have not been analyzed, so the evolutionary record has gaps until more samples are collected. Convergent evolution[14] refers to the evolution of similar traits in independent organisms. Yan et al do not attempt to refute the prevailing scientific evidence on viral evolution, but assert that ZC45, a coronavirus with over 3,000 punctuated, broadly distributed nucleotide differences from SARS-CoV-2 (a significantly large number of differences), could have been used as a "backbone" or template to produce SARS-CoV-2 synthetically. ZC45 is a beta coronavirus[15] isolated from a bat between 2015 and 2017 in Zhoushan city, Zhejiang province, China. ZC45 and ZXC21 were both discovered and characterized in to better understand animal reservoirs of SARS-like coronaviruses. No explanation is given for how the over 3,000 nucleotide differences SARS-CoV-2 and ZC45 could be produced; this process would be highly challenging for deliberate engineering.

2. **Role of Chinese military lab.** Lines 4-6: The United States has a number of high-containment laboratories in which viruses can be studied safely with engineering controls, including negative air pressure. Some of these labs are located at military laboratories, such as the US Army Medical

Research Institute of Infectious Diseases in Frederick, Maryland. China, France, Germany, India, Russia, the United Kingdom, and many other countries similarly have laboratories operated by military researchers that are declared to the Biological Weapons Convention in confidence-building measures. Scientific investigation in military laboratories is not uncommon; coronavirus research performed in a Chinese military research institute is not in itself suspicious, as asserted by Yan et al.

3. **Furin cleavage sites in coronaviruses.** Lines 10-16: The authors assert that a furin cleavage site in its Spike protein is absent in coronaviruses found in nature, which is not the case.* This is fairly common in other coronaviruses[16]; MERS has a furin cleavage site[17] within Spike.

4. **Dissimilarities between SARS-CoV-2 and ZC45.** Figure 1.1: The report features a figure comparing sequences of various coronavirus strains. The figure's data appear accurate and demonstrate a high degree of dissimilarity between ZC45 and SARS-CoV-2, particularly in ORF1a, but the conclusion made by the authors in the text is that the strains are similar. Neither the figure nor the text clarify which genome serves as the reference.

Page 4

1. **Similarity of ORF8 between SARS-CoV-2 and ZC45.** Lines 9-14. The authors' assertion that the similarity between the ORF8 gene in SARS-CoV-2 and ZC45 is unnatural (relative to sequence conservation among coronaviruses) is not supported by evidence presented. While the sequence of ORF8 varies among coronaviruses, its function is not well characterized.[18] In line 10, the authors report that ORF8 may be involved in SARS-CoV-2's ability to evade the host immune response (and thus affect pathogenicity). They then suggest that ORF8 is usually dissimilar among different coronavirus strains, based on a paper by Muth et al[19] that studied deletions in ORF8 during the 2003 SARS-CoV-1 epidemic. Muth et al found that a deletion of 29 nucleotides in ORF8 of SARS-CoV-1 attenuated the virus by decreasing the virus's ability to replicate. A recent paper[20] identified the role of ORF8 in pathogenesis of SARS-CoV-2 as potentially playing a role in viral maturation and assembly. Importantly, this study on ORF8 was published *after* the emergence of SARS-CoV-2, whose mode of action is still not fully understood; this timeline does not align with Yan and colleagues' proposed timeline of events. Furthermore, the authors fail to consider the level of similarity in ORF8 between viral variants of the same strain, which could provide better context for the sequence identity between different strains. It is, therefore, inappropriate to suggest that the similarity of SARS-CoV-2 and ZC45 is unusual.

2. Also, lines 11-13: ZC45 and ZXC21 seem to have an 94% identity with ORF8, which is greater than with other circulating coronaviruses (59%), but this is still quite low. ORF8 has been identified[20] as a protein of interest in aiding in virus assembly/packaging. Yan et al argue that SARS-CoV-2 is suspiciously similar to SARS-CoV-1, yet these 2 viruses contain less than 20% similarity in their ORF8 sequences.

* The original version of this report included an editing error in this paragraph. It was updated on 10/14/20.

Page 5

1. **Mischaracterization of sequence homology data.** Lines 9-10, referring to Figure 2: The authors present a variety of homology data that are superfluous, internally inconsistent, or misinterpreted in the text. For example, the authors state that the E protein, which plays a minimal role in pathogenesis, is highly variable; however, the Figure 2 shows a fairly stable amino acid sequence. In lines 4-5, the authors state that SARS-CoV-2's E gene is highly permissible to mutations because in a 2-month period there have been 4 nonsynonymous mutations. They use this to suggest it is suspicious that early SARS-CoV-2 samples had identical identity to the purported "backbone" viruses, when SARS-CoV-2 is able to tolerate nonsynonymous mutations to the E gene and, therefore, it would be unlikely for SARS-CoV-2 to have evolved naturally to have 100% sequence identity. However, this analysis does not consider the selection bias in the samples' sequences and gaps in the existing phylogenetic trees. It is acknowledged in the field that there are gaps in the phylogenetic trees of the coronavirus family, making it difficult to determine accurately the likelihood of similarity between 2 viral variants. Additionally, Figure 2 shows only 1 sequence from an early time point in the pandemic and 4 samples from April. If other samples from February were to be included, then there might not be 100% amino acid sequence identity between SARS-CoV-2 samples and ZC45 and ZXC21. Finally, 2 strains of coronaviruses showing identical sequences in a particular gene could be an example of convergent evolution.[21]

2. **Binding with ACE2.** Lines 31-34: In a discussion about whether RaTG13 can bind various ACE2 homologs from different types of horseshoe bats, the authors neglect to point out that the ACE2 homolog of the specific species of horseshoe bat from which RaTG13 was isolated was not included in the cited binding studies. This makes conclusions about whether RaTG13 can bind ACE2 homologues incomplete.

3. **Binding of *Rhinolophus affinis* ACE 2.** Lines 34-36: Research[22] has shown that the receptor-binding domain of SARS-CoV-2 binds human, pangolin, and *Rhinolophus macrotis* bat ACE2 receptors optimally, and that the receptor-binding domain of *Rhinolophus affinis*, a type of horseshoe bat, did not bind the ACE2 of orthologous (different) horseshoe bat species' ACE2. *R affinis* ACE2 has not been well characterized, so it could not be tested. This is interesting work in progress but does not provide substantive conclusions about the provenance of SARS-CoV-2.

Page 6

1. **Missing methods section.** The report is missing a methods section, which is typically included in review articles[23] and allows for critical review of the process by which the articles reviewed were chosen. Information should be included about how the alignments were created, sequence quality, and adjustments for sampling bias—all factors that affect the results and conclusions.

Page 8

1. **On variability of Spike sequences.** Lines 1-13: There are various judgments about the similarity of SARS-CoV-2 sequences to other related viruses (ZC45 and ZXC21), but no inclusion of contrasting evidence. For example, S2 is not highly variable among coronaviruses,[24] but S1 is only a 69% match, making the claims that ZC45 was used as a template not credible. Convergent evolution, seen in several other viruses,[25,26] including SARS-CoV-1,[27] often as a virulence factor, should be considered by the authors.

2. **On substitution mutations within the Spike protein.** Lines 16-18: Substitution mutations that are hydrophobic and classified as minor in the report, are structurally significant and not minor; many mutations are lysines to phenylalanines, which alter structure, or phenylalanine to tyrosine which alter the charge of the side group.

3. **Quasiviruses and evolution of RNA viruses.** Lines 23-26: The authors make teleological assumptions in this passage. "As elaborated below, the way that SARS-CoV-2 RBM [receptor-binding motif] resembles SARS-CoV RBM and the overall sequence conservation pattern between SARS-CoV-2 and ZC45/ZXC21 are highly unusual. Collectively, this suggests that portions of the SARS-CoV-2 genome have not been derived from natural quasi-species viral particle evolution."[10] Currently, not enough is understood about SARS quasispecies[28] to argue definitively that a certain population arose from another or to eliminate the possibility of said evolution. Many of the Yan and colleagues' arguments could be explained by a mixture of convergent evolution, quasispecies, sampling bias, methodology issues, and/or a common ancestor.

4. **Viral recombination.** Lines 30, 31, and 43: The description of viral recombination does not accurately describe how this process occurs in viruses.[29,30] Viral recombination is a complex event,[31] which is not a "swapping" of entire genes, as the authors suggest, but a common, important part of viral evolution.[29] Reassortment can occur, but only in segmented, positive-sense RNA viruses. It is likely that ancestors of SARS-CoV-2 underwent viral recombination, though this is not necessarily a complete exchange of entire gene segments.

Page 9

1. **The potential for zoonotic emergence of coronaviruses.** Line 9: There is not enough information available in the scientific literature to know whether strains related to SARS-CoV-2 may infect humans or if infections are possible but limited. Therefore, statements made by the authors about the infectivity of ZC45 are unsupported.

2. **On intermediate hosts in viral evolution.** Lines 21-23: Viruses can have complicated evolutionary origins, sometimes with intermediate hosts,[32] as seen with influenza[33]; influenza viruses[34] are also known to crossover into humans. The human ACE2 (hACE2) receptor may be optimal for SARS-CoV-2, but recent work has found that SARS-CoV-2 can actually use multiple ACE2 receptors,[4] but not mice ACE2. More sampling needs to be done, but assertions about whether the hACE2 is the best receptor to bind SARS-CoV-2 cannot be supported at this time.

3. **Zoonotic emergence of coronaviruses in history.** Lines 36-38: Coronaviruses have caused human disease before, including SARS and MERS, and many have pointed to warning signs that coronaviruses could become a serious problem, which were not heeded prior to SARS-CoV-2. These facts are contradicted by the authors who also describe SARS-CoV-2 as "intelligent," which is teleological and counterfactual.

Page 10

1. **Lack of evidence regarding gain of function research in coronaviruses.** Line 2: Some gain of function research using coronaviruses has been published, but the author's statement of an

"abundant" literature in this area overstates the amount known. The papers referenced do not support the author's claim that such research led to human competent viruses. One paper, Ren et al,[35] inserted the Spike protein gene of all SARS-CoV-like viruses (not SARS) into a viral backbone and did not use the entire SARS virus or infect live animals.

2. **Lack of restriction sites in the proposed viral backbone ZC45.** Figure 5: The authors describe a possible pathway for designing viruses that is out of step with current scientific methods for gene editing, casting doubt on both their analysis and their conclusions. While use of restriction sites as presented are theoretically possible in SARS-CoV-2, based on the authors' own analysis, ZC45 does not have the necessary restriction sites (of EcoRI and BstEII). Therefore, ZC45 would have to be genetically modified beyond the sequence presented for a restriction digestion to be possible. This negates the authors' argument that ZC45 is the obvious backbone of SARS-CoV-2. Restriction digests are not favored for manipulation of RNA viruses due to several obstacles: genome sizes, viral proofreading enzymes that can limit the success of restriction enzymes, and the ability to recover viruses after reverse genetic manipulation.

Page 11

1. **On restriction sites present within the Spike protein.** Lines 6-9: Restriction enzyme sites are found in all genomes and naturally occur frequently.[36,37] For instance, in a commonly used adenovirus vector, the BstEII restriction enzyme site occurs 10 times. The frequency of restriction site distribution is due to the fact that they comprise stretches of 6 or 8 consecutive nucleotides, which have high—and measurable—probabilities of occurring by chance within a given genome. With contemporary gene-editing methodologies, restriction sites are rarely used. These arguments aside, Yan and colleagues falsely assert the existence of restriction enzyme sites in the SARS-CoV-2 sequence, but not in the Spike gene sequence of other beta coronaviruses, is evidence of genetic manipulation, or that the presence of restriction sites is rare. A New England BioLabs site search for restriction enzyme sites in the 5' end of the SARS-CoV-2 sequence revealed at least 7 other restriction sites in the RBM, in addition to the *Eco*RI site Yan et al cited as evidence of manipulation.

Page 13

1. **The possibility of convergent evolution in beta coronaviruses.** Lines 10-12: Yan et al state that there is only 1 evolutionary pathway that could explain the appearance of SARS-CoV-2—a homologous recombination event. However, convergent evolution is another pathway for the development of the furin cleavage site, which would result in SARS-CoV-2 having the cleavage site similar to nonbeta coronaviruses. Convergent evolution is a well-established phenomenon in biology.

2. **The evolution of a furin cleavage site.** Lines 14-16: The authors argue that the existence of polybasic furin cleavage sites in other coronaviruses implies that convergent evolution could not have played a role in evolution of the furin cleavage site in SARS-CoV-2. The furin cleavage site refers to a specific position at the S1/S2 junction in SARS-CoV-2. This is a sequence of amino acids where the host (human) enzyme, furin,[38] can cleave. This furin cleavage is essential for the proper maturation[39] of the Spike glycoprotein and subsequent cell-to-cell membrane fusion in

the host. They present the divergent furin cleavage site sequence in SARS-CoV-2 as evidence that homologous recombination between an ancestor beta coronavirus and a furin cleavage site-containing coronavirus is impossible. The argument that homologous recombination is not a likely factor in fact supports a hypothesis of convergent evolution.

3. **Homologous recombination.** Lines 18-19: The report states that the low sequence identity between beta coronavirus and other coronaviruses that contained a furin cleavage site would be too low to allow homologous recombination to occur. If recombination had occurred, it would not have had to have occurred in the immediate area of the sequence coding the furin cleavage site; it could occur in other, more homologous regions.

Page 14

1. **On methods of a literature review.** Typically, the scientific description of the steps to create a transmissible virus (as per the chart on page 15) would require biosecurity review before publication in a reputable scientific journal, as this is a dual-use concern,[40] which has the potential to lower barriers toward biological weapons development. However, it should be noted that the steps described by Yan et al are not individually novel and, in our judgment, do not present a biological weapons risk, particularly as the methods chosen have been supplanted by more accurate genetic engineering tools.

Page 16

1. **On troubleshooting molecular cloning.** Line 16: The authors' statement that there is "almost no risk of [molecular cloning] failing"[10] contradicts experience with the technique, as it can be a finicky method[41] requiring keen problem-solving skills.[42]

2. **Virology protocol inaccuracies.** Lines 25-29: The report inaccurately describes some common laboratory techniques. For example, the report states that sequence information for short segments of coronavirus RNA-dependent RNA polymerase (RdRp) is possible due to the availability of a polymerase chain reaction (PCR) protocol used to identify coronaviruses. However, PCR is not a sequencing method, it only amplifies existent sequences. PCR is a common tool, used to determine if a specific DNA sequence is in a sample and, if so, how many copies of that sequence are in the sample. Using PCR to detect the presence of coronaviruses in a sample is a standard practice in research and clinical laboratories using standard coronavirus-specific primers, as the RdRp is highly conserved between coronaviruses. Approximately 28 current SARS-CoV-2 diagnostics[43] with Emergency Use Authorization use this method and this specific gene target.

Page 17

1. **Serial passaging and virulence.** Lines 19-20: Serial passaging refers to a process wherein a stock viral population is used to infect an animal, then virus from that animal is collected and used to infect another animal for a designated number of "passages." Serial passage of a virus causes the population to adapt to the animal or cell type in which it is being passaged. Passaging would lead to adaptation to another animal (if passaged *in vivo*) or, if *in vitro*, to the specific cell type used. Most human cells used in laboratory culture have significant differences compared to the

commensurate cells in humans. Serial passage, then, would not necessarily make a virus more pathogenic to live humans. Additionally, passage does not necessarily increase fitness of a viral population. The report mischaracterizes the complexity of these processes and projects outcomes from passaging that are not supported by laboratory techniques.

Page 18

1. **Unrealistic timelines.** Lines 25-29: The timeline offered for how an entirely novel protein can be engineered in a little studied virus, circa 2019, is not scientifically realistic.

Page 19

1. **Methods of genetic modification in viruses.** Lines 11-12: The authors incorrectly state that reverse genetics systems are commonly used to assemble coronaviruses. Reverse genetics[44] can be used in other virus synthesis, such as influenza. The paper the authors cite, from Thao and colleagues,[41] did use reverse genetics in a yeast-based system to synthesize full length SARS-CoV-2. However, previous research[45] had identified that coronaviruses can be particularly difficult to engineer using reverse genetics systems, as the large size of Nidovirus genomes, replicase activity, and requirement for large transcript synthesis create obstacles. Certain methods require insertion of mutations elsewhere in the genome to manage the T7 transcription termination signals or require helper viruses to coinfect cells to aid in cloning. Recent work in dengue viruses[46] and MERS[47] has shown the promise of Gibson assembly in synthesizing positive-strand viruses.

2. **Reverse genetics tools (and limitations).** Lines 22-24: Reverse genetics and synthetic biology provide technological tools to synthesize SARS-CoV-2, as demonstrated by the methods section of the Thao paper.[41] The yeast used for this synthesis of SARS-CoV-2 used a specific platform that depended on a mouse hepatitis virus. The description in Yan et al of pooling the DNA fragments together and "transforming" them into yeast will not work,[48,49] as it would require a method known as transformation-associated recombination,[50] calling into question the Yan analysis.

3. **On viral passaging and adaptation.** Lines 34-35: Adaptation for receptors likely improves infectability of a virus, but it does not necessarily make the virus more transmissible, pathogenic, or virulent. Even if the virus adapts for the receptor, it does not mean that the virus will be able to cause viremia or transmit to other hosts. The report falsely asserts that serial passage would "validate the virus' fitness and ensure its receptor-oriented adaptation toward its intended host"[10] and also argue a contradictory theory on page 3 that the virus *was not* serially passaged. While viral passaging can optimize viral fitness, this is never a guarantee and has to be scientifically demonstrated.

4. **SARS-CoV-2 animal models.** Line 39: Finding an animal model for SARS-CoV-2 has been difficult[51] and, before 2020, there was not a good animal model for SARS-CoV, so the idea of "serial passage in laboratory animals"[10] would have been challenging.

Page 20

1. **Serial passaging and virulence.** Lines 2-4: The authors incorrectly assert that serial passage of a virus only leads to increased virulence. The report asserts that 10 to 15 rounds of passage would improve the viral Spike protein's binding affinity and the infectivity and lethality of the virus. However, serial passaging does not always lead to genome stabilization, as some viral populations may die off.[52] Of the strains that do stabilize, infection efficiency is only enhanced for the model species used for passaging, not for all species. Some of these (millions of) virions may be more lethal or infectious, just as many may be less so. Passaging cannot guarantee an outcome of viral evolution. The life cycle of a virus and infection efficiency depend on more than just receptor binding, and adaptation to 1 organ or 1 type of receptor may come at the expense of reduced ability to spread to other organs, cause viremia, shed from 1 host, or cause pathogenicity.[53] Thus, improved receptor binding does not necessarily mean enhanced transmissibility or pathogenicity.[54]

2. **On laboratory adaptation leading to increased virulence.** Lines 16-21: Viral adaptation can include attenuation. That is one reason why viruses and bacteria are sometimes serially passaged for attenuation to be used in vaccines.[55,56] It cannot be stated as Yan and colleagues do that there is a "lack of apparent attenuation"[10] so far in this pandemic, because the global incidence of COVID-19 (especially asymptomatic cases) is unknown, or that viral adaptation, *in vitro* or *in vivo*, led to increase transmissibility or virulence.

3. **Viral mutation rates.** Line 24: The authors state that if serial passage is confined to 1 species, less random mutations occur, but this is incorrect. Mutation rates are a function of the RdRp, as well as the ExoN proofreading enzyme,[57] and so repeated passage will not inherently make a virus more or less likely to mutate. However, passage does affect which mutations[58] become fixed in the viral population. Coronaviruses form a quasispecies, where each variant within the population can be different from the others and have different fitnesses. Together, the population of variants infect a host, disseminate from the initial infection site, and cause pathogenesis. During a passage, the viral variants best suited for infection and pathogenesis within the model organism are selected for, but the rate of mutations occurring does not change. Because mutations can still occur, there is a possibility the virus can adapt, unless the mutations cause so many deleterious mutations that the population collapses.

Page 21 (Conclusion)

1. While the impact of SARS-CoV-2 on global public health is undeniable, the pathogenic effects of SARS-CoV-2 infection at an individual or cellular level are not unprecedented. Many viruses are capable of causing high morbidity and mortality,[59,60] infecting several organs, and/or presymptomatic or asymptomatic transmission.[61] Additionally, other viral infections (eg, chikungunya) also induce long-term sequelae.[62] Humans have contended with many scourges and it is a certainty that COVID-19 will not be the last.

References

1. COVID-19 dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. Johns Hopkins University of Medicine Coronavirus Resource Center website. Accessed September 21, 2020. https://coronavirus.jhu.edu/map.html
2. Luan J, Jin X, Lu Y, Zhang L. SARS-CoV-2 Spike protein favors ACE2 from *Bovidae* and *Cricetidae. J Med Virol*. April 1, 2020. https://doi.org/10.1002/jmv.25817
3. Anderson KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARA-CoV-2. *Nat Med.* 2020;26:450-452. https://doi.org/10.1038/s41591-020-0820-9
4. Zhou P, Yang Z-L, Wang X-G, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature.* 2020;579:270-273. https://doi.org/10.1038/s41586-020-2012-7
5. Leitner T, Kumar S. Where did SARS-CoV-2 come from? *Mol Biol Evol.* 2020;37(9):2463-2464. https://doi.org/10.1093/molbev/msaa162
6. Xia X. Extreme genomic CpG deficiency in SARS-CoV-2 and evasion of host antiviral defense. *Mol Biol Evol.* 2020;37(9):2699-2705. https://doi.org/10.1093/molbev/msaa094
7. Zhou H, Chen X, Hu T, et al. A novel bat coronavirus closely related to SARS-CoV-2 contains natural insertions at the S1/S2 cleavage site of the Spike protein. *Curr Biol.* 2020;30(11):2196-2203.e3. https://doi.org/10.1016/j.cub.2020.05.023
8. Zhang Y-Z, Holmes EC. A Genomic Perspective on the Origin and Emergence of SARS-CoV-2. *Cell.* 2020;181(2):223-227. https://doi.org/10.1016/j.cell.2020.03.035
9. Tang Z, Wu C, Li X, et al. On the origin and continuing evolution of SARS-CoV-2. *Natl Sci Rev.* 2020;7(6):1012-1023. https://doi.org/10.1093/nsr/nwaa036
10. Yan L-M, Kang S, Guan J, Hu S. *Unusual Features of the SARS-CoV-2 Genome Suggesting Sophisticated Laboratory Modification Rather Than Natural Evolution and Delineation of Its Probable Synthetic Route.* New York: Rule of Law Society & Rule of Law Foundation; 2020.
11. Bat coronavirus RaTG13, complete genome. GenBank. Accessed September 21, 2020. https://www.ncbi.nlm.nih.gov/nuccore/MN996532.1
12. Oong XY, Ng KT, Takebe Y, et al. Identification and evolutionary dynamics of two novel human coronavirus OC43 genotypes associated with acute respiratory infections: phylogenetic, spatiotemporal and transmission network analyses. *Emerg Microbes Infect.* 2017;6(1):1-13. https://doi.org/10.1038/emi.2016.132
13. Forni D, Cagliani R, Clerici M, Sironi M. Molecular evolution of human coronavirus genomes. *Trends Microbiol.* 2017;25(1):35-48. https://doi.org/10.1016/j.tim.2016.09.001
14. Stayton CT. What does convergent evolution mean? The interpretation of convergence and its implications in the search for limits to evolution. *Interface Focus.* 2015;5(6):20150039. https://dx.doi.org/10.1098%2Frsfs.2015.0039
15. Hu D, Zhu C, Ai L, et al. Genomic characterization and infectivity of a novel SARS-like coronavirus in Chinese bats. *Emerg Microbe Infect.* 2018;7:154. https://dx.doi.org/10.1038%2Fs41426-018-0155-5
16. Cheng J, Zhao Y, Xu G, et al. The S2 subunit of QX-type infectious bronchitis coronavirus Spike protein is an essential determinant of neurotropism. *Viruses.* 2019;11(10):972. https://doi.org/10.3390/v11100972
17. Jaimes JA, Millet JK, Goldstein ME, Whittaker GR, Straus MR. A fluorogenic peptide cleavage assay to screen for proteolytic activity: applications for coronavirus Spike protein activation. *J Vis Exp.* 2019;143:e58892. https://dx.doi.org/10.3791/58892
18. Wu Z, Yang L, Ren X, et al. ORF8-related genetic evidence for Chinese horseshoe bats as the source of human severe acute respiratory syndrome coronavirus. *J Infect Dis.* 2016;213(4):579-583. https://doi.org/10.1093/infdis/jiv476
19. Muth D, Corman VM, Roth H, 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.* 2018;8:15177. https://doi.org/10.1038/s41598-018-33487-8

20. Flower TG, Buffalo CZ, Hooy RM, Allaire M, Ren X, Hurley JH. Structure of SARS-CoV-2 ORF8, a rapidly evolving coronavirus protein implicated in immune evasion. Preprint. *bioRxiv.* Posted August 27, 2020. Accessed September 21, 2020. https://www.biorxiv.org/content/10.1101/2020.08.27.270637v1.full.pdf

21. Lan J, Ge J, Yu J, et al. Structure of the SARS-CoV-2 Spike receptor-binding domain bound to the ACE2 receptor. *Nature.* 2020;581:215-220. https://doi.org/10.1038/s41586-020-2180-5

22. Mou H, Quinlan BD, Peng H, et al. Mutations from bat ACE2 orthologs markedly enhance ACE2-Fc neutralization of SARS-CoV-2. Preprint. bioRxiv. Posted June 30, 2020. Accessed September 21, 2020. https://dx.doi.org/10.1101%2F2020.06.29.178459

23. Gülpinar Ö, Güçlü AG. How to write an review article? *Turk J Urol.* 2013;39(suppl 1):44-48. https://dx.doi.org/10.5152%2Ftud.2013.054

24. Lokman SM, Rasheduzzaman M, Salauddin A, et al. Exploring the genomic and proteomic variations of SARS-CoV-2 spike glycoprotein: a computational biology approach. *Infect Genet Evol.* 2020;84:104389. https://dx.doi.org/10.1016%2Fj.meegid.2020.104389

25. Higginbottom A, Quinn ER, Kuo C-C, et al. Identification of amino acid residues in CD81 critical for interaction with hepatitis C virus envelope glycoprotein E2. *J Virol.* 2000;74(8):3642-3649. https://doi.org/10.1128/jvi.74.8.3642-3649.2000

26. Kimalov B, Gal=on A, Stav R, Belausov E, Arazi T. Maintenance of coat protein N-terminal net charge and not primary sequence is essential for zucchini yellow mosaic virus systemic infectivity. *J Gen Virol.* 2004;85(110):3421-3430. https://doi.org/10.1099/vir.0.80417-0

27. Qu X-X, Hao P, Song X-J, et al. Identification of two critical amino acid residues of the severe acute respiratory syndrome coronavirus Spike protein for its variation in zoonotic tropism transition via a double substitution strategy. *J Biol Chem.* 2005;280:29588-29595. https://doi.org/10.1074/jbc.M500662200

28. Xu D, Zhang Z, Wang F-S. SARS-associated coronavirus quasispecies in individual patients. *N Engl J Med.* 2004;350:1366-1367. https://doi.org/10.1056/NEJMc032421

29. Bentley K, Evans DJ. Mechanisms and consequences of positive-strand RNA virus recombination. *J Gen Virol.* 2018;99(10):1345-1356. https://doi.org/10.1099/jgv.0.001142

30. Worobey M, Holmes EC. Evolutionary aspects of recombination in RNA viruses. *J Gen Virol.* 1999;80(10):2535-2543. https://doi.org/10.1099/0022-1317-80-10-2535

31. Simon-Loriere, EC Holmes. Why do RNA viruses recombine? *Nat Rev Microbiol.* 2011;9:617-626. https://doi.org/10.1038/nrmicro2614

32. Yuan S, Jiang S-C, Li Z-L. Analysis of possible intermediate hosts of the new coronavirus SARS-CoV-2. *Front Vet Sci.* 2020;7:379. https://dx.doi.org/10.3389%2Ffvets.2020.00379

33. Parrish CR, Murcia PR, Holmes EC. Influenza virus reservoirs and intermediate hosts: dogs, horses, and new possibilities for influenza virus exposure of humans. *J Virol.* 2015;89(6):2990-2994. https://doi.org/10.1128/JVI.03146-14

34. Naffakh N, van der Werf S. April 2009: an outbreak of swine-origin influenza A(H1N1) virus with evidence for human-to-human transmission. *Microbes Infect*. 2009;11(8-9):725-728. https://doi.org/10.1016/j.micinf.2009.05.002

35. Ren W, Li W, Yu M, et al. Full-length genome sequences of two SARS-like coronaviruses in horseshoe bats and genetic variation analysis. *J Gen Virol.* 2006;87(Pt 11):3355-3359. https://doi.org/10.1099/vir.0.82220-0

36. Herrara S, Reyes-Herrara PH, Shank TM. Predicting RAD-seq marker numbers across the Eukaryotic Tree of Life. *Genome Biol Evol.* 2015;7(12):3207-3225. https://doi.org/10.1093/gbe/evv210

37. New England BioLabs Inc. Frequencies of restriction sites. Accessed September 21, 2020. https://www.neb.com/tools-and-resources/selection-charts/frequencies-of-restriction-sites

38. New England BioLabs Inc. Furin. Accessed September 21, 2020. https://www.neb.com/products/p8077-furin#Product%20Information

39. Hoffman M, Kleine-Weber H, Pöhlmann S. A multibasic cleavage site in the Spike protein of SARS-CoV-2 is essential for infection of human lung cells. *Mol Cell.* 2020;78(4):779-784.e5. https://doi.org/10.1016/j.molcel.2020.04.022

40. National Institutes of Health. Dual use research of concern. Accessed September 21, 2020. https://osp.od.nih.gov/biotechnology/dual-use-research-of-concern/

41. Thao TTN, Labroussaa F, Ebert N, et al. Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform. *Nature.* 2020;582;561-565. https://doi.org/10.1038/s41586-020-2294-9

42. ThermoFisher Scientific. Cloning troubleshooting guide. Accessed September 21, 2020. https://www.thermofisher.com/us/en/home/life-science/cloning/cloning-learning-center/invitrogen-school-of-molecular-biology/molecular-cloning/cloning/cloning-troubleshooting-guide.html

43. Vela Diagnostics. ViroKey SARS-CoV-2 RT-PCR test v1.0 (EUA). Accessed September 21, 2020. https://www.veladx.com/product/qpcr-respiratory-viruses/virokey-sars-cov-2-rt-pcr-test.html

44. Neumann G, Fujii K, Kino Y, Kawaoka Y. An improved reverse genetics system for influenza A virus generation and its implications for vaccine production. *Proc Natl Acad Sci U S A.* 2005;102(46):16825-16829. https://doi.org/10.1073/pnas.0505587102

45. Almazán F, Sola I, Zuñiga S, et al. Coronavirus reverse genetic systems: infectious clones and replicons. *Virus Res.* 2014;189:262-270. https://dx.doi.org/10.1016%2Fj.virusres.2014.05.026

46. Siridechadilok B, Gomutsukhavadee M, Sawaengpol T, et al. A simplified positive-sense-RNA virus construction approach that enhances analysis throughput. *J Virol.* 2013;87(23):12887-12674. https://dx.doi.org/10.1128%2FJVI.02261-13

47. Lee JY, Bae S, Myoung J. Generation of full-length infectious cDNA clones of Middle East respiratory syndrome coronavirus. *J Microbiol Biotechnol.* 2019;29(6):999-1007. https://doi.org/10.4014/jmb.1908.08004

48. Vashee S, Stockwell TB, Alperovich N, et al. Cloning, assembly, and modification of the primary human cytomegalovirus isolate toledo by yeast-based transformation-associated recombination. *mSphere.* 2017;2(5):e00331-17. https://dx.doi.org/10.1128%2FmSphereDirect.00331-17

49. Liao C-L, Lai MMC. RNA recombination in a coronavirus: recombination between viral genomic RNA and transfected RNA fragments. *J Virol.* 1992;66(10):6117-6124. Accessed September 21, 2020. https://jvi.asm.org/content/66/10/6117.short

50. Kouprina N, Larionov V. TAR cloning: perspectives for functional genomics, biomedicine, and biotechnology. *Mol Ther Methods Clin Dev.* 2019;14:16-26. https://doi.org/10.1016/j.omtm.2019.05.006

51. Takayama K. *In vitro* and animal models for SARS-CoV-2 research. *Trends Pharmacol Sci.* 2020;41(8):513-517. https://doi.org/10.1016/j.tips.2020.05.005

52. Warmbrod KL, Patterson EI, Kautz TF, et al. Viral RNA-dependent RNA polymerase mutants display an altered mutation spectrum resulting in attenuation in both mosquito and vertebrate hosts. *PLoS Pathog.* 2019;15(4):e1007610. https://dx.doi.org/10.1371%2Fjournal.ppat.1007610

53. Mandary MB, Masomian M, Poh CL. Impact of RNA virus evolution on quasispecies formation and virulence. *Int J Mol Sci.* 2019;20(18):4657. https://doi.org/10.3390/ijms20184657

54. Mair CM, Ludwig K, Herrmann A, Sieben C. Receptor binding and pH stability — how influenza A virus hemagglutinin affects host-specific virus infection. *Biochim Biophys Acta Biomembr.* 2014;1838(4):1153-1168. https://doi.org/10.1016/j.bbamem.2013.10.004

55. Hanley KA. The double-edged sword: how evolution can make or break a live-attenuated virus vaccine. *Evolution (N Y).* 2011;4(4):635-643. https://dx.doi.org/10.1007%2Fs12052-011-0365-y

56. Eckels KH, Scott RM, Bancroft WH, et al. Selection of attenuated dengue 4 viruses by serial passage in primary kidney cells. V. Human response to immunization with a candidate vaccine prepared in fetal rhesus lung cells. *Am J Trop Med Hyg.* 1984;33(4):684-689. https://doi.org/10.4269/ajtmh.1984.33.684

57. Ogando NS, Ferron F, Decroly E, Canard B, Posthuma CC, Snijder EJ. The curious case of the nidovirus exoribonuclease: its role in RNA synthesis and replication fidelity. *Front Microbiol.* 2019;10:1813. https://dx.doi.org/10.3389%2Ffmicb.2019.01813

58. Peck KM, Lauring AS. The complexities of viral mutation rates. *J Virol.* 2018;92(14):e01031-17. https://doi.org/10.1128/JVI.01031-17

59. Lee N, Lui GCY, Wong KT, et al. High morbidity and mortality in adults hospitalized for respiratory syncytial virus infections. *Clin Infect Dis.* 2013;57(8):1069-1077. https://doi.org/10.1093/cid/cit471

60. World Health Organization. Ebola virus disease: latest numbers as of 21 June 2020. Accessed September 21, 2020. https://www.afro.who.int/health-topics/ebola-virus-disease

61. Rieg G, Lewis RJ, Miller LG, Witt MD, Guerrero M, Daar ES. Asymptomatic sexually transmitted infections in HIV-infected men who have sex with men: prevalence, incidence, predictors, and screening strategies. *AIDS Patient Care STDS.* 2008:22(12):947-954. https://dx.doi.org/10.1089%2Fapc.2007.0240

62. Marimoutou C, Ferraro J, Javelle E, Deparis X, Simon F. Chikungunya infection: self-reported rheumatic morbidity and impaired quality of life persist 6 years later. *Clin Microbiol Infect.* 2015;21(7):688-693. https://doi.org/10.1016/j.cmi.2015.02.024