Is considering a genetic-manipulation origin for SARS-CoV-2 a conspiracy theory that must be censored?

Preprint · April 2020

CITATIONS
0

READS
12,789

1 author:

Rossana Segreto
University of Innsbruck
12 PUBLICATIONS **345** CITATIONS

SEE PROFILE

DOI: **10.13140/RG.2.2.31358.13129/1**

Is considering a genetic-manipulation origin for SARS-CoV-2 a conspiracy theory that must be censored?

Rossana Segreto[1#] and Yuri Deigin[2]

[1]Department of Microbiology, University of Innsbruck, Austria.

[2]Youthereum Genetics Inc., Toronto, ON Canada.

[#]Correspondence should be addressed to Rossana.Segreto@uibk.ac.at

RS and YD contributed equally to the manuscript.

Abstract

The origin of SARS-CoV-2 is still controversial. Comparative genomic analyses have shown that SARS-CoV-2 is likely to be chimeric, most of its sequence being very close to the CoV detected from a bat, whereas its receptor binding domain is almost identical to that of CoV obtained from pangolins. The furin cleavage site in the spike protein of SARS-CoV-2 was previously not identified in other SARS-like CoVs and might have conferred the ability to cross species and tissue barriers. Chimeric viruses can be the product of natural recombination or genetic manipulation. The latter could have aimed to identify pangolins as possible intermediate hosts for bat-CoV potentially pathogenic for humans. Theories that consider a possible artificial origin for SARS-CoV-2 are censored as they seem to support conspiracy theories. Researchers have the responsibility to carry out a thorough analysis, beyond any personal research interests, of all possible causes for SARS-CoV-2 emergence for preventing this from happening in the future.

Several months have passed since the outbreak of SARS-CoV-2 in Wuhan, China, and its origin is still controversial. The theory that the Wuhan's Huanan Seafood Wholesale Market was the first source for animal–human virus transmission has lost credibility. During the first phase of the epidemic in Wuhan, several hospitalized patients with confirmed SARS-CoV-2 infections had no link with the market.[1] Unfortunately, the market was quickly closed and sanitized before enough animal samples could have been collected; the few market samples that did get collected exhibit only human-adapted SARS-CoV-2 and no traces of zoonotic predecessor strains.[2]

The closest relatives to SARS-CoV-2 are bat and pangolin coronaviruses

Zhou and colleagues[3] from the Wuhan Institute of Virology (WIV) first identified and characterized the new coronavirus (CoV), later named SARS-CoV-2. The genomic sequences obtained from early cases shared 79% sequence identity to the CoVs that caused Severe Acute Respiratory Syndrome (SARS-CoV) in 2002-2003 and 96.2% sequence identity to RaTG13 (MN996532), a total genomic sequence of a CoV detected from a *Rhinolophus affinis* bat. This sample was collected in the Yunnan province (China) by the same group of researchers in 2013. Zhou and colleagues[3] found a short region of RNA-dependent RNA polymerase (RdRp) in their data and then fully sequenced the original sample. This sequence is currently the closest phylogenetic relative for SARS-CoV-2 found[4] and it has not been published before the outbreak of SARS-CoV-2.

The RdRp of RaTG13 has 100% nucleotide identity with the sequence BtCoV/4991 (KP876546) identified by Ge and colleagues[5] in a *Rhinolophus affinis* bat in the Yunnan province in 2013, as RaTG13. The original sample was collected in a mine colonized by bats near Tongguanzhen, Mojiang, Yunnan. The WIV researchers were invited to investigate the mine after 6 miners contracted severe pneumonia in 2012[6], and 3 of the miners had died.[7] The miners were tasked with clearing out bat droppings in the mine, and the severity of their pneumonia correlated with the duration of exposure to the mine.[8] Some of the miners' samples subsequently underwent testing at WIV, where IgG antibodies against SARS were identified in 4 of the samples.[9] Considering that only about 5300

people were infected in mainland China during the SARS outbreak of 2002-2004, most of whom resided in Guandong, the odds of 4 miners in Yunnan retaining antibodies from the 2002-2004 SARS outbreak are quite low. On the other hand, it is possible that the SARS antibody test administered to the miners cross-reacted with a novel SARS-like bat virus that the miners had acquired at the mine. Ge and colleagues[5] had identified a number of CoVs in the mine, but based on the phylogenetic analysis, BtCoV/4991 was the only SARS-related strain, clearly separated from all known alpha- and beta-CoVs at that time. Ge et al. also amplified spike genes of collected CoVs and made them available upon request. BtCoV/4991 differentiates from other bat CoVs also in the phylogenetic analysis carried out by Wang and colleagues in 2019.[10] Chen and colleagues[11] identified BtCoV/4991 as the closest sequence to SARS-CoV-2 because RaTG13 had not yet been published at that time. BtCoV/4991 and RaTG13 have been recently confirmed to be two different coding names of the same strain by their original authors at WIV, as they registered the two strains as one entry in the Database of Bat-associated Viruses (DBatVir).[12]

The second non-human RdRp sequence closest to BtCoV/4991 (91.89% nucleotide identity) is the CoV sequence MP789 (MT084071) isolated in 2019 in a Malaysian pangolin (*Manis javanica*) from the Guangdong province, China.[13] The envelope protein of MP789 has 100% aminoacidic identity with the corresponding protein in RaTG13, in bat-SL-CoVZXC21 (MG772934.1), in bat-SL-CoVZC45 (MG772933.1) and in some SARS-CoV-2 isolates (e.g. YP_009724392).[14] The envelope protein of CoVs is involved in critical aspects of the viral life, as viral entry, replication and pathogenesis.[15]

Bat CoVs have been studied intensely and genetically manipulated

Several studies point out that bats are reservoirs for a broad diversity of potentially pathogenic SARS-like CoV.[16, 17] Some of these viruses can directly infect humans[18], whereas others need to mutate their spike protein in order to effectively bind to the human angiotensin 1-converting enzyme 2 (hACE2) receptor and mediate virus entry.[19] In order to evaluate the emergence potential of novel CoVs, chimeric CoVs with Bat CoV backbones not able to infect human cells were fused to spike proteins of CoVs compatible with human ACE2, simulating recombination events that might naturally

occur.[20, 21] These experiments with gain of function have raised biosafety concerns and controversy among researchers and the public. One of the main arguments in favour of gain of function studies is the need to be prepared with an arsenal of drugs and vaccines for the next pandemic. By contrast, one of the main arguments against them is that the next pandemic could be caused by those experiments, due to the risk of lab leakage.[22, 23, 24]

In recent years, the field of corona-virology had been focused on pan-coronavirus therapies and vaccines, as evident from research conducted in the past five years,[25, 26, 27, 28] as well as from media reports.[29] Synthetically generating diverse panels of potential pre-emergent coronaviruses was declared as a goal of active grants for EcoHealth Alliance which funded some of such research at WIV.[30]

Key difference between SARS-CoV-2 and its closest relative RaTG13

SARS-CoV-2 differs from its closest relative RaTG13 by a few key characteristics. The most striking one is the acquisition in the spike protein of SARS-CoV-2 of a cleavage site activated by the host-cell enzyme furin, previously not identified in other beta-CoVs of lineage b[31] and similar to that of Middle East Respiratory Syndrome Coronavirus (MERS-CoV).[32] Host protease processing plays a pivotal role as a species and tissue barrier. Engineering of the cleavage sites of CoV spike proteins modifies virus tropism and virulence.[33] The ubiquitous expression of furin in different organs and tissues may have conferred to SARS-CoV-2 the ability to infect body parts insensitive to other CoVs, leading to systemic infection in the body.[34] Cell-cultured SARS-CoV-2 that was missing the above-mentioned cleavage site caused attenuated symptoms in infected hamsters,[35] and mutagenesis studies have confirmed that the polybasic furin site is essential for SARS-CoV-2's ability to infect human lung cells.[36]

The polybasic furin site in CoV2 was created by a 12-nt insert TCCTCGGCGGGC coding for a PRRA amino acid sequence at the S1/S2 junction (Fig. 1). Interestingly, the two joint arginines are coded by two CGGCGG codons, which are quite rare for these viruses: only 5% of arginines are coded by CGG in CoV2 or RaTG13, and CGGCGG in the new insert is the only doubled instance of this codon in CoV2. The CGGCGG insert includes a *Fau*I restriction site, of which there are six instances in CoV2 and four instances in

RaTG13 (and 2 in MP789). The serendipitous location of the *Fau*I site could allow using restriction fragment length polymorphism (RFLP) techniques[37] for cloning[38] or screening for mutations,[39] as the new furin site is prone to deletions *in vitro*.[40, 41]

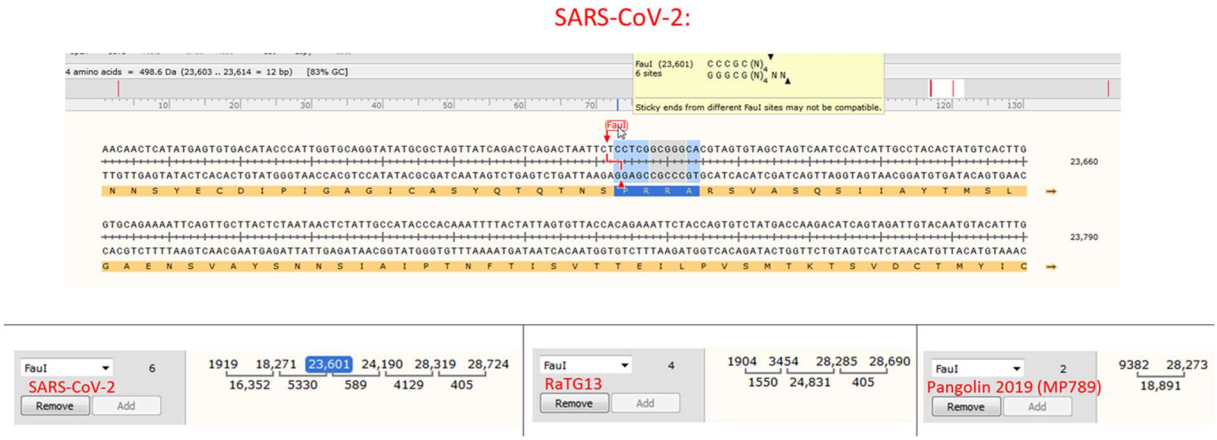

Fig. 1 Nucleotide sequence of the S protein at the S1/S2 junction in SARS-CoV-2 (NC045512.2), showing the furin cleavage site (in blue) that includes a *Fau*I enzyme restriction site.

A study by Zhou et al.[42] recently reported the discovery of a novel CoV strain RmYN02, which the authors claim exhibits natural PAA amino acid insertions at the S1/S2 cleavage site where SARS-CoV-2 has the PRRA insertion. However, upon close examination of the underlying nucleotide sequence of RmYN02 in comparison with its closest ancestors ZC45 and ZXC21, no insertions are apparent, just nucleotide mutations (Fig. 2).

Fig. 2 Alignment of nucleotide and amino acid sequences of the S protein from RaTG13 (MN996532) and RmYN02 at the S1/S2 junction site. No insertions of nucleotides possibly evolving in a furin cleavage site can be observed (in blue).

Therefore, SARS-CoV-2 remains unique among its beta coronavirus relatives not only due to a polybasic furin site at the S1/S2 junction, but also due to the four amino acid insert PRRA which had created it (Fig. 3).

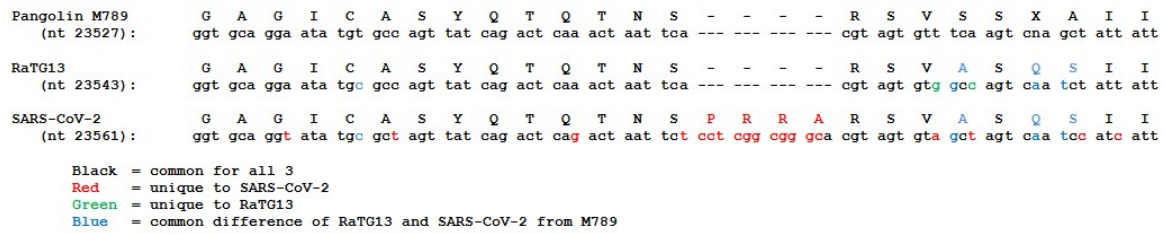

Fig. 3 Alignment of nucleotide and amino acid sequences of the S protein from RaTG13 (MN996532), MP789 (MT084071) and SARS-CoV-2 (NC045512.2) at the S1/S2 site. The common nucleotides and amino acids are given in black, SARS-CoV-2 unique nucleotides and amino acids in red, RaTG13 unique nucleotides and amino acids in green and common nucleotides and amino acids in SARS-CoV-2 and RaTG13 that differ in M789 in blue. The codon for Serine (TCA) in RaTG13 and MP789 is split in SARS-CoV-2 to give part of a new codon for Serine (TCT) and part of the amino acid Alanine (GCA).

Interestingly, the insertion of the furin cleavage site in SARS-CoV-2 is not in frame with the rest of sequence, when compared with the MP789 and the RaTG13 sequences. The insertion causes a split in the original codon for Serine (TCA) in MP789 or RaTG13 to give part of a new codon for Serine (TCT) and part of the amino acid Alanine (GCA) in SARS-CoV-2 (Fig. 3).

SARS-CoV-2 seems to merge some exclusive features of SARS-CoV together with those typical for MERS-CoV. A recent study has identified the MERS-CoV transmembrane dipeptidyl peptidase 4 receptor (DDP4) as a candidate binding target or coreceptor of SARS-CoV-2.[43]

Pangolin or not pangolin, that is the question

The possibility that pangolins could be the intermediate host for SARS-CoV-2 is still under discussion.[44, 45, 46] SARS-CoV-2 and RaTG13 mostly diverge because of the RBD of their spike protein.[4] Although the average genome similarity is lower compared to RaTG13, CoV isolated from pangolins has RBDs almost identical to that of SARS-CoV-2. Indeed, pangolin CoVs and SARS-CoV-2 possess identical amino acids at the five critical residues of the RBD, whereas RaTG13 only shares one amino

acid with SARS-CoV-2.[32] ACE2 sequence similarity is higher between humans and pangolins than between humans and bats. Intriguingly, the Spike protein of SARS-CoV-2 has a higher predicted binding affinity to human ACE2 receptor than to that of pangolins.[47] Before the SARS-CoV-2 outbreak, pangolins were the only mammals other than bats documented to carry and be infected by SARS-CoV-2 related CoV.[13, 45] Recombination events between the RBD of CoV from pangolins and RaTG13-like backbone could have produced SARS-CoV-2 as a chimeric strain. For such recombination to occur, the two viruses must have infected the same cell in the same organism simultaneously.[32]

Is a lab origin for SARS-CoV-2 a baseless conspiracy theory?

Due to the broad-spectrum of research conducted over almost 20 years on bat SARS-CoV justified by their potential to spill over from animal to human,[48] a possible synthetic origin by laboratory engineering of SARS-CoV-2 is a reasonable hypothesis. For Andersen and colleagues,[49] strong evidence that SARS-CoV-2 did not result from genetic manipulation is that the high-affinity binding of the SARS-CoV-2 spike protein to hACE2 could not have been predicted by models based on the RBD of SARS-CoV. Based on the structural analysis conducted by Wan and colleagues,[50] SARS-CoV-2 has the potential to recognize hACE2 more efficiently than the SARS-CoV which emerged in 2002. Moreover, generation of CoV chimeric strains has recently demonstrated that bat CoV spikes can bind to the hACE2 receptor with more plasticity than previously predicted.[16] All amino acids in the RBD have been extensively analysed and new models to predict ACE2 affinity are available.[51] As described above, creation of chimeric viruses has been carried out over the years with the purpose to study the potential pathogenicity of bat CoVs for humans. In this context, SARS-CoV-2 could have been synthesized by combining a backbone similar to RaTG13 with the RBD of CoV similar to the one recently isolated from pangolins[13], because the latter is characterized by a higher affinity with the hACE2 receptor. Such research could have aimed to identify pangolins as possible intermediate hosts for bat-CoV potentially pathogenic for humans.

Regarding the furin cleavage site, Andersen and colleagues[49] state that “The functional consequence of the polybasic cleavage site in SARS-CoV-2 is unknown”. New studies from several groups have

lately identified this activation site as possibly enabling the virus to spread efficiently between humans and attack multiple organs.[52] Experiments on proteolytic cleavage of CoV spike proteins have been recently suggested as future key studies to study virus transmissibility in different hosts. [51] The pangolin from which MP789 was isolated was co-infected by several viruses[13], among others the Herpes Virus that is characterized by a furin cleavage site. In a context of an evolutionary study, this observation might have suggested the idea of this insertion into SARS-CoV-2.

Andersen and colleagues[49] also state, based on the work of Almazan and colleagues[53] that "the genetic data irrefutably show that SARS-CoV-2 is not derived from any previously used virus backbone". In the last six years before the outbreak of SARS-CoV-2 the number of potential bat backbones has been undeniably increased by several bat CoV screenings, last but not least bringing RaTG13 to scientific attention in January 2020. Other possible backbones could, as well, still wait for publication.

Andersen and colleagues[49] also state that "The acquisition of both the polybasic cleavage site and predicted O-linked glycans also argues against culture-based scenarios". Methods for insertion of a polybasic cleavage site in infectious bronchitis CoV are given in Cheng and colleagues[54] and resulted in increased pathogenicity. Concerning the predicted O-linked glycans around the newly inserted polybasic site, it should be noted that this prediction was not confirmed by Cryo-EM inquiry into the SARS-CoV-2 spike glycoprotein.[55] Nevertheless, while it is true that O-linked glycans are much more likely to arise under immune selection, they could be added in the lab through site-directed mutagenesis[56] or arise in the course of *in vivo* experiments, for example, in BLT-L mice that have human lung implants and autologous human immune system[57] or in mice expressing human ACE2 receptor.[58] To overcome problems of bat CoV isolation, experiments based on direct inoculation of bat CoV in suckling rats have been carried out.[59] Pangolins or other animals with similar ACE2 conformation could have been used as experimental animals as well.

The authors also state that "Subsequent generation of a polybasic cleavage site would have then required repeated passage in cell culture or animals with ACE2 receptors similar to those of humans,

but such work has also not previously been described." It should not be excluded that such experiments could have been aborted due to the SARS-CoV-2 outbreak, before a possible publication of the results or that the results were never intended to be published.

Conclusion

Due to the gravity of SARS-CoV-2 impact on humanity, researchers have the responsibility to carry out a thorough analysis, beyond any personal research interests, of all possible causes for SARS-CoV-2 emergence. Unfortunately, theories that consider a possible artificial origin for SARS-CoV-2 are censored by international scientific journals as they seem to support conspiracy theories. Genetic manipulation of SARS-CoV-2 may have been carried out in any laboratory in the world with access to the backbone sequence and the necessary equipment. New technologies based on synthetic genetics platforms even allow the reconstruction of viruses based on their genomic sequence, without the need of a natural isolate.[60]

Xiao Qiang, a research scientist at the School of Information at the University of California at Berkeley, recently stated: "To understand exactly how this virus has originated is critical knowledge for preventing this from happening in the future".[61]

Acknowledgements

We thank Prof. Allan Krill (NTNU) for proof reading the manuscript and the valuable comments; Prof. Heribert Insam (Head of the Department of Microbiology; University of Innsbruck) for his support and Dr. Lawrence Sellin for all the valuable information. A special thanks goes to Dr. Fernando Castro-Chavez (New York Medical College) and to René Bergelt for their support.

Conflicts of Interest Statement

RS and YD do not have any conflicts of interest.

References

[1] Huang C, Wang Y, Li X et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020; **395**: 497–506. https://doi.org/10.1016/S0140-6736(20)30183-5

[2] Zhan SH, Deverman BE, Chan YA. SARS-CoV-2 is well adapted for humans. What does this mean for re-emergence? *bioRxiv* 2020.05.01.073262. DOI: https://doi.org/10.1101/2020.05.01.073262

[3] Zhou P, Yang X, Wang X et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 2020; **579:** 270–273. DOI: https://doi.org/10.1038/s41586-020-2012-7

[4] Cagliani R, Forni D, Clerici M, Sironi M. Computational inference of selection underlying the evolution of the novel coronavirus, SARS-CoV-2. *J Virol.* 2020. DOI: 10.1128/JVI.00411-20

[5] Ge XY, Wang N, Zhang W et al. Coexistence of multiple coronaviruses in several bat colonies in an abandoned mineshaft. *Virol Sin.* 2016; **31**: 31–40. DOI: 10.1007/s12250-016-3713-9

[6] Qiu J. How China's 'Bat Woman' Hunted Down Viruses from SARS to the New Coronavirus. *Sci Am*. June 2020. https://www.scientificamerican.com/article/how-chinas-bat-woman-hunted-down-viruses-from-sars-to-the-new-coronavirus1/

[7] Wu Z, Yang L, Yang F, et al. Novel Henipa-like Virus, Mojiang Paramyxovirus, in Rats, China, 2012. *Emerg Infect Dis*. 2014; **20**(6): 1064–1066. DOI: 10.3201/eid2006.131022.

[8] Li Xu. [The analysis of 6 patients with severe pneumonia caused by unknown virus] (MSc thesis in Chinese). Kunming Medical University, Emergency Medicine (professional degree), 2013. http://eng.oversea.cnki.net/Kcms/detail/detail.aspx?filename=1013327523.nh&dbcode=CMFD&dbname=CMFD2014

[9] Canping Huang. [Novel Virus Discovery in Bat and the Exploration of Receptor of Bat Coronavirus HKU9] (PhD thesis in Chinese). National Institute for Viral Disease Control and Prevention, Chinese Center for Disease Control and Prevention, June 2016. http://eng.oversea.cnki.net/kcms/detail/detail.aspx?dbcode=CDFD&dbname=CDFDLAST2018&filename=1017118517.nh

[10] Wang N, Luo C, Liu H et al. Characterization of a new member of alphacoronavirus with unique genomic features in *Rhinolophus* bats. *Viruses* 2019; **11**: 379. DOI: 10.3390/v11040379

[11] Chen L, Liu W, Zhang Q et al. RNA based mNGS approach identifies a novel human coronavirus from two individual pneumonia cases in 2019 Wuhan outbreak. *Emerg Microbes Infect.* 2020; **9**: 313–9. DOI: 10.1080/22221751.2020.1725399

[12] DBatVir – The Database of Bat-Associated Viruses. http://www.mgc.ac.cn/cgi-bin/DBatVir/main.cgi?func=accession&acc=MN996532

[13] Liu P, Chen W, Chen J-P. Viral metagenomics revealed *Sendai virus* and coronavirus infection of Malayan Pangolins (*Manis javanica*). *Viruses* 2019; **11**: 979. DOI: 10.3390/v11110979

[14] Bianchi M, Benvenuto D, Giovanetti M, Angeletti S, Ciccozzi M, Pascarella S. Sars-CoV-2 Envelope and Membrane proteins: differences from closely related proteins linked to cross-species transmission? http://doi.org/10.20944/preprints202004.0089.v1

[15] Schoeman D, Fielding BC. Coronavirus envelope protein: current knowledge. *Virol J*. 2019; **16**(1): 69. https://doi.org/10.1186/s12985-019-1182-0

[16] Hu B, Zeng L-P, Yang X-L et al. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLoS Pathog.* 2017; **13**: e1006698. DOI: 10.1371/journal.ppat.1006698

[17] Fan Y, Zhao K, Shi Z-L, Zhou P. Bat coronaviruses in China. *Viruses* 2019; **11**: 210. DOI: 10.3390/v11030210

[18] Ge XY, Li JL, Yang XL, Chmura AA et al. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* 2013; **503**: 535. https://doi.org/10.1038/nature12711

[19] Graham RL, Baric RS. Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. *J Virol*. 2010; **84**: 3134–3146. DOI: 10.1128/JVI.01394-09.

[20] Agnihothram S, Yount BL Jr., Donaldson EF et al. A mouse model for *Betacoronavirus* subgroup 2c using a bat coronavirus strain HKU5 variant. *mBio*. 2014; **5**: e00047-14. DOI: 10.1128/mBio.00047-14

[21] Johnson BA, Graham RL, Menachery VD. Viral metagenomics, protein structure, and reverse genetics: key strategies for investigating coronaviruses. *Virology* 2018; **517**: 30–37. https://doi.org/10.1016/j.virol.2017.12.009

[22] Weiss S, Yitzhaki S, Shapira SC. Lessons to be learned from recent biosafety incidents in the United States. *Isr Med Assoc J*. 2015; **17**: 269–273. PMID: 26137650

[23] Racaniello V. Moving beyond metagenomics to find the next pandemic virus. *Proc Natl Acad Sci U S A*. 2016; **113**: 2812–2814. DOI: 10.1073/pnas.1601512113

[24] Casadevall A, Imperiale MJ. Risks and benefits of gain-of-function experiments with pathogens of pandemic potential, such as influenza virus: a call for a science-based discussion. *MBio*. 2014; **201**: 5–e1730–e1714. DOI: 10.1128/mBio.01730-14

[25] Agostini ML, Andres EL, Sims AC, et al. Coronavirus Susceptibility to the Antiviral Remdesivir (GS-5734) Is Mediated by the Viral Polymerase and the Proofreading Exoribonuclease. *mBio*. 2018; **9**(2): e00221-18. DOI: 10.1128/mBio.00221-18

[26] Xia S, Liu M, Wang C, et al. Inhibition of SARS-CoV-2 (previously 2019-nCoV) infection by a highly potent pan-coronavirus fusion inhibitor targeting its spike protein that harbors a high capacity to mediate membrane fusion. *Cell Res*. 2020; **30**(4): 343–355. DOI: 10.1038/s41422-020-0305-x

[27] Totura AL, Bavari S. Broad-spectrum coronavirus antiviral drug discovery. *Expert Opin Drug Discov*. 2019; **14**(4): 397–412. DOI: 10.1080/17460441.2019.1581171

[28] Wang Y, Sun Y, Wu A, et al. Coronavirus nsp10/nsp16 Methyltransferase Can Be Targeted by nsp10-Derived Peptide In Vitro and In Vivo To Reduce Replication and Pathogenesis. J Virol. 2015;89(16):8416-8427. doi:10.1128/JVI.00948-15

[29] Kahn J. How Scientists Could Stop the Next Pandemic Before It Starts. *NYT Magazine*. March 2020. https://www.nytimes.com/2020/04/21/magazine/pandemic-vaccine.html

[30] Project Number 2R01AI110964-06. UNDERSTANDING THE RISK OF BAT CORONAVIRUS EMERGENCE. Awardee Organization: ECOHEALTH ALLIANCE, INC. https://projectreporter.nih.gov/project_info_description.cfm?aid=9819304&icde=49645421&ddparam=&ddvalue=&ddsub=&cr=1&csb=default&cs=ASC&pball=

[31] Coutard B, Valle C, de Lamballerie X, Canard B, Seidah NG, Decroly E. The spike glycoprotein of the new coronavirus 2019- nCoV contains a furin-like cleavage site absent in CoV of the same clade. *Antiviral Res*. 2020; **176**: 104742. https://doi.org/10.1016/j.antiviral.2020.104742

[32] Zhang T, Wu Q, Zhang Z. Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. *Curr Biol*. 2020; **30**: 1346–1351. https://doi.org/10.1016/j.cub.2020.03.022

[33] Letko M, Marzi A, Munster V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat Microbiol.* 2020; **5**: 562–569. https://doi.org/10.1038/s41564-020-0688-y

[34] Wang Q, Qiu Y, Li JY, Zhou ZJ, Liao CH, Ge XY. A unique protease cleavage site predicted in the spike protein of the novel pneumonia Coronavirus (2019-nCoV) potentially related to viral transmissibility. *Virol Sin*. 2020. https://doi.org/10.1007/s12250-020-00212-7

[35] Lau SY, Wang P, Mok B W-Y et al. Attenuated SARS-CoV-2 variants with deletions at the S1/S2 junction. *Emerg Microbes Infect*. 2020. https://doi.org/10.1080/22221751.2020.1756700

[36] Hoffmann 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 [published online ahead of print, 2020 Apr 28]. Mol Cell. 2020;S1097-2765(20)30264-1. doi:10.1016/j.molcel.2020.04.022

[37] Kaundun SS, Marchegiani E, Hutchings SJ, Baker K. Derived Polymorphic Amplified Cleaved Sequence (dPACS): A novel PCR-RFLP procedure for detecting known single nucleotide and deletion-insertion polymorphisms. *Int J Mol Sci.* 2019; **20**(13): 3193. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6651057/

[38] Zeng LP, Gao YT, Ge XY, 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*. 2016; **90**(14): 6573–6582. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4936131/

[39] Khan SG, Muniz-Medina V, Shahlavi T, et al. The human XPC DNA repair gene: arrangement, splice site information content and influence of a single nucleotide polymorphism in a splice acceptor site on alternative splicing and function. *Nucleic Acids Res*. 2002; **30**(16): 3624–3631. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC134237/

[40] Liu Z, Zheng H, Yuan R, et al. Identification of a common deletion in the spike protein of SARS-CoV-2. *bioRxiv* 2020.03.31.015941. https://doi.org/10.1101/2020.03.31.015941

[41] Lau SY, Wang P, Mok B W-Y et al. Attenuated SARS-CoV-2 variants with deletions at the S1/S2 junction. *Emerg Microbes Infect*. 2020. https://doi.org/10.1080/22221751.2020.1756700

[42] 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 [published online ahead of print, 2020 May 11]. *Curr Biol*. 2020. DOI: 10.1016/j.cub.2020.05.023

[43] Li Y, Zhang Z, Yang L, et al. The MERS-CoV receptor DPP4 as a candidate binding target of the SARS-CoV-2 spike. *iScience* 2020. https://doi.org/10.1016/j.isci.2020.101160

[44] Li X, Zai J, Zhao Q, Nie Q, Li Y, Foley BT, Chaillon A. Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. *J Med Virol*. 2020. https://doi.org/10.1002/jmv.25731

[45] Lam TT, Shum MH, Zhu H et al. Identifying SARS-CoV-2 related coronaviruses in Malayan pangolins. *Nature* 2020. https://doi.org/10.1038/s41586-020-2169-0

[46] Xiao K, Zhai J, Feng Y. et al. Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature* 2020. https://doi.org/10.1038/s41586-020-2313-x

[47] Piplani S, Singh PK, Winkler DA, Petrovsky N. In silico comparison of spike protein-ACE2 binding affinities across species; significance for the possible origin of the SARS-CoV-2 virus. arXiv:2005.06199 [q-bio.BM]. 2020. https://arxiv.org/abs/2005.06199

[48] Wang LF, Anderson DE. Viruses in bats and potential spillover to animals and humans. *Curr Opin Virol*. 2019; **34**: 79–89. https://doi.org/10.1016/j.coviro.2018.12.007

[49] Andersen KG, Rambaut A, Lipkin WI, Holmes, Garry RF. The proximal origin of SARS-CoV-2. *Nat Med*. 2020; **26**: 450–452. https://doi.org/10.1038/s41591-020-0820-9

[50] Wan Y, Shang J, Graham R, Baric RS, Li F. Receptor recognition by novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS. *J Virol*. 2020. https://doi.org/10.1128/jvi.00127-20.

[51] Cui J, Li F, Shi Z. Origin and evolution of pathogenic coronaviruses. *Nat Rev Microbiol.* 2019; **17**: 181–192. https://doi.org/10.1038/s41579-018-0118-9

[52] Mallapati S. Why does the coronavirus spread so easily between people? *Nature* 2020; **579**: 183. DOI: 10.1038/d41586-020-00660-x

[53] Almazan F, Gonzalez JM, Penzes Z. Coronavirus reverse genetic systems: Infectious clones and replicons. *Virus Res*. 2014; **189**: 262–270. DOI: 10.1016/j.virusres.2014.05.026

[54] Cheng J, Zhao Y, Xu G. The S2 Subunit of QX-type infectious bronchitis coronavirus spike protein is an essential determinant of neurotropism. *Viruses* 2019; **11**: 10. DOI: 10.3390/v11100972

[55] Daniel Wrapp NW, Kizzmekia S. Corbett, Jory A. Goldsmith, Ching-Lin Hsieh, Olubukola Abiona, Barney S. Graham, Jason S. McLellan. Cryo-EM Structure of the 2019-nCoV Spike in the Prefusion Conformation. *Science* 2020. https://doi.org/10.1016/j.cell.2020.02.058

[56] Du, L., Tai, W., Yang, Y. et al. Introduction of neutralizing immunogenicity index to the rational design of MERS coronavirus subunit vaccines. Nat Commun 7, 13473 (2016). https://doi.org/10.1038/ncomms13473

[57] Wahl A, De C, Abad Fernandez M, et al. Precision mouse models with expanded tropism for human pathogens. *Nat Biotechnol*. 2019; **37**(10): 1163–1173. DOI: 10.1038/s41587-019-0225-9

[58] Menachery VD, Yount BL Jr, Sims AC, et al. SARS-like WIV1-CoV poised for human emergence. *Proc Natl Acad Sci U S A*. 2016; **113**(11): 3048–3053. DOI: 10.1073/pnas.1517719113

[59] Hu D, Zhu C, Ai L et al. Genomic characterization and infectivity of a novel SARS-like coronavirus in Chinese bats. *Emerg Microbes Infect.* 2018; **7**: 154. DOI: 10.1038/s41426-018-0155-5

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

[61] Josh Rogin. State Department cables warned of safety issues at Wuhan lab studying bat coronaviruses. *The Washington Post*. April 14, 2020. https://www.washingtonpost.com/opinions/2020/04/14/state-department-cables-warned-safety-issues-wuhan-lab-studying-bat-coronaviruses/