United States Department of State

Washington, D.C. 20520

September 6, 2024

Case No. FL-2022-00062

Mr. Gary Ruskin
U.S. Right to Know
4096 Piedmont Avenue, #963
Oakland, CA 94611

Dear Mr. Ruskin:

As we noted in our letter dated July 26, 2024, we are processing your request for material under the Freedom of Information Act ("FOIA"), 5 U.S.C. § 552. The Department of State ("Department") has identified an additional five responsive records subject to the FOIA. We have determined four records may be released in part, and one must be withheld in full.

An enclosure explains the FOIA exemptions and other grounds for withholding material. Where we have made redactions, the applicable FOIA exemptions are marked on each record. The record withheld in full is exempt from release pursuant to FOIA Exemptions 5, 5 U.S.C. § 552. The document identification number for the record is: A-00000565063. Where applicable, the Department has considered the foreseeable harm standard when reviewing these records and applying FOIA exemptions. All non-exempt material that is reasonably segregable from the exempt material has been released and is enclosed.

We will keep you informed as your case progresses. If you have any questions, your attorney may contact Assistant United States Attorney Stephanie Johnson at stephanie.johnson5@usdoj.gov or (202) 252-7874. Please refer to the case number, FL-2022-00062, and the civil action number, 22-cv-01130, in all correspondence about this case.

Sincerely,

Avery Bullard

Avery D. Bullard
Chief, Litigation and Appeals Branch
Office of Information Programs and Services

Enclosures: As stated.

The Freedom of Information Act (5 USC 552)

FOIA Exemptions

(b)(1) Information specifically authorized by an executive order to be kept secret in the interest of national defense or foreign policy. Executive Order 13526 includes the following classification categories:

1.4(a) Military plans, systems, or operations
1.4(b) Foreign government information
1.4(c) Intelligence activities, sources or methods, or cryptology
1.4(d) Foreign relations or foreign activities of the US, including confidential sources
1.4(e) Scientific, technological, or economic matters relating to national security, including defense against transnational terrorism
1.4(f) U.S. Government programs for safeguarding nuclear materials or facilities
1.4(g) Vulnerabilities or capabilities of systems, installations, infrastructures, projects, plans, or protection services relating to US national security, including defense against transnational terrorism
1.4(h) Weapons of mass destruction

(b)(2) Related solely to the internal personnel rules and practices of an agency

(b)(3) Specifically exempted from disclosure by statute (other than 5 USC 552), for example:

ARMSEXP	Arms Export Control Act, 50a USC 2411(c)
CIA PERS/ORG	Central Intelligence Agency Act of 1949, 50 USC 403(g)
EXPORT CONTROL	Export Administration Act of 1979, 50 USC App. Sec. 2411(c)
FS ACT	Foreign Service Act of 1980, 22 USC 4004
INA	Immigration and Nationality Act, 8 USC 1202(f), Sec. 222(f)
IRAN	Iran Claims Settlement Act, Public Law 99-99, Sec. 505

(b)(4) Trade secrets and confidential commercial or financial information

(b)(5) Interagency or intra-agency communications forming part of the deliberative process, attorney-client privilege, or attorney work product

(b)(6) Personal privacy information

(b)(7) Law enforcement information whose disclosure would:
(A) interfere with enforcement proceedings
(B) deprive a person of a fair trial
(C) constitute an unwarranted invasion of personal privacy
(D) disclose confidential sources
(E) disclose investigation techniques
(F) endanger life or physical safety of an individual

(b)(8) Prepared by or for a government agency regulating or supervising financial institutions

(b)(9) Geological and geophysical information and data, including maps, concerning wells

Other Grounds for Withholding

NR Material not responsive to a FOIA request excised with the agreement of the requester

From:	"Grube, Steven M. EOP/NSC" (b)(6)
To:	Feith, David (b)(6) @state.gov>
Subject:	RE: Urgent HHS COVID statement for review
Date:	Fri, 4 Dec 2020 20:03:32 +0000

We're getting our edits cleared through Anthony but he's in meetings that can't be moved or interrupted.

What did you mean about pinging HHS?

I've attached a fun article about the prevalence of preexisting antibodies that it might be good to take a look at.

Best regards,
Steve

From: Feith, David (b)(6) state.gov>
Sent: Friday, December 4, 2020 2:20 PM
To: Kanapathy, Ivan J. EOP/NSC (b)(6) Grube, Steven M. EOP/NSC (b)(6) Fabina, Lauren C. EOP/NSC (b)(6)
Cc: Ruggiero, Anthony J. EOP/NSC (b)(6)
Subject: RE: Urgent HHS COVID statement for review

With attachment. Also pasting State/EAP edits here:

(b)(5)

(b)(5)

~~SENSITIVE BUT UNCLASSIFIED~~

From: Kanapathy, Ivan J. EOP/NSC (b)(6)
Sent: Friday, December 4, 2020 1:58 PM
To: Feith, David (b)(6) @state.gov>; Grube, Steven M. EOP/NSC (b)(6)
Fabina, Lauren C. EOP/NSC (b)(6)
Cc: Ruggiero, Anthony J. EOP/NSC (b)(6)
Subject: RE: Urgent HHS COVID statement for review

++ Steve and Lauren.

From: Feith, David (b)(6) @state.gov>
Sent: Friday, December 4, 2020 1:56 PM
To: Kanapathy, Ivan J. EOP/NSC (b)(6) Ruggiero, Anthony J. EOP/NSC

(b)(6)

Subject: RE: Urgent HHS COVID statement for review

Gents, see attached my edits. Also suggest you ping HHS etc.

~~SENSITIVE BUT UNCLASSIFIED~~

From: Feith, David
Sent: Friday, December 4, 2020 1:23 PM
To: Kanapathy, Ivan (b)(6)
Subject: FW: Urgent HHS COVID statement for review
Importance: High

Gents –

(b)(5)

Thanks.

--
David Feith
Deputy Assistant Secretary
Bureau of East Asian and Pacific Affairs (EAP)
U.S. Department of State
(b)(6)
(b)(6) @state.gov

~~SENSITIVE BUT UNCLASSIFIED~~

From: Stilwell, David R (b)(6) @state.gov>
Sent: Friday, December 4, 2020 1:06 PM
To: (b)(6) @state.gov>
Cc: Fritz, Jonathan D (b)(6) @state.gov>; Buangan, Richard L (b)(6) @state.gov>; Feith, David (b)(6) @state.gov>; EAP-Press <EAP-Press@state.gov>; (b)(6) @state.gov>
Subject: Re: FOR A/S STILWELL CLEARANCE ASAP: 11:30 AM: Urgent HHS statement for review

Good. There's some great reporting out of Taiwan on this topic. Can we point them to TACRO?

Get Outlook for iOS

From: (b)(6) @state.gov>
Sent: Friday, December 4, 2020 11:30:57 AM

To: Stilwell, David R (b)(6) @state.gov>
Cc: Fritz, Jonathan D <(b)(6) @state.gov>; Buangan, Richard L (b)(6) @state.gov>; Feith, David (b)(6) @state.gov>; EAP-Press <EAP-Press@state.gov>; (b)(6) @state.gov>
Subject: FOR A/S STILWELL CLEARANCE ASAP: 11:30 AM: Urgent HHS statement for review

A/S Stilwell,

(b)(5)

Apologies for the short fuse. Would appreciate your earliest clearance so that we can get our edits in.

Thank you,

(b)(6)

(b)(5)

(b)(5)

From: (b)(6) @state.gov>
Sent: Friday, December 4, 2020 12:13 PM
To: (b)(6) @state.gov>; (b)(6) @state.gov>
Cc: EAP-Press <EAP-Press@state.gov>
Subject: RE: FOR (b)(6) CLEARANCE ASAP: 11:30 AM: Urgent HHS statement for review

Clear for CM. FO should see this one – recommend including Feith when you send it up to Buangan and Fritz.

(b)(6)
Director, Office of Chinese and Mongolian Affairs
Bureau of East Asian and Pacific Affairs
(b)(6) @state.gov
(b)(6)

From: (b)(6) @state.gov>
Sent: Friday, December 4, 2020 12:04 PM
To: (b)(6) @state.gov>; EAP-CM-Global-DL <EAP-CM-Global-DL@state.gov>; EAP-CM-Bilat Unit-DL <EAPCMBilatUnit@state.gov>; EAP-CM-ECON-DL <EAP-CM-ECON-DL@state.gov>
Cc: EAP-Press <EAP-Press@state.gov>
Subject: Re: CLEARANCE ASAP: 11:30 AM: Urgent HHS statement for review

Hi (b)(6

Some edits for CM/Econ below in yellow highlights. Since the CDC study was about cases before the President stopped travel from China, it does not show how that helped as implied in the third paragraph.

(b)(6)

Environment, Science, Technology, and Health Officer

Office of Chinese and Mongolian Affairs

U.S. Department of State

(b)(6)

From: (b)(6) @state.gov>
Sent: Friday, December 4, 2020 10:16 AM
To: EAP-Press <EAP-Press@state.gov>; (b)(6) @state.gov>; IO-Press-DL <IO-Press-DL@state.gov>
Cc: OES-PA-DG <OES-PA-DG@state.gov>
Subject: 11:30 AM: Urgent HHS statement for review

Hello, please see HHS reactive statement below in response to the WSJ article that China is using to claim COVID did not start in Wuhan. I would appreciate your comments/clearance by 11:30.

Thank you,
(b)(6)
OES/PPO

****

The draft reactive statement below is in response to the WSJ article that China is using to claim COVID did not start in Wuhan. Once HHS comments have been received, it will go to State Dept for review.

>>https://www.wsj.com/articles/covid-19-likely-in-u-s-in-mid-december-2019-cdc-scientists-report-11606782449<<;

(b)(5)

(b)(5)

Sender:	"Grube, Steven M. EOP/NSC" (b)(6)
Recipient:	Feith, David (b)(6) @state.gov>

From:	"Feith, David" (b)(6) @state.gov>
To:	davidjfeith(b)(6)
Subject:	COVID, Toy Reid
Date:	Tue, 19 Jan 2021 17:19:48 +0000

From: Kanapathy, Ivan (b)(6)
Sent: Tuesday, November 3, 2020 2:42 PM
To: Feith, David (b)(6) @state.gov>
Subject: FW: links

1 of 2

From: Reid, Toy (Rubio) <Toy_Reid@rubio.senate.gov>
Sent: Thursday, October 1, 2020 1:48 PM
To: Kanapathy, Ivan J. EOP/NSC (b)(6)
Subject: RE: links

Here is Shi Zhengli's belated attempt in July to hush the questions surrounding her work. *Science* does not exactly subject her claims to serious scrutiny. Ebright is quoted here again.

>https://www.sciencemag.org/news/2020/07/trump-owes-us-apology-chinese-scientist-center-covid-19-origin-theories-speaks-out<

>https://www.sciencemag.org/sites/default/files/Shi%20Zhengli%20Q%26A.pdf<

From: Reid, Toy (Rubio)
Sent: Thursday, October 1, 2020 11:41 AM
To: 'Kanapathy, Ivan J. EOP/NSC' (b)(6)
Subject: RE: links

For some strange reason, we can't get the most important paper, the first link below, to print properly. The graphics are totally illegible. It's best viewed online, but perhaps you'll have better luck printing.

From: Reid, Toy (Rubio)
Sent: Thursday, October 1, 2020 11:11 AM

To: 'Kanapathy, Ivan J. EOP/NSC' (b)(6)
Subject: links

Ivan,

I'm planning to bring printed copies of the first three links below, but I wanted to share the links with you as well. I think these articles address most of D/NSA Pottinger's questions, but I can dig up more if necessary. See you at 12:30.

Thanks,
Toy

Best Summaries of the Key Virological Issues, the State of the Field, and Shi's Publication Record

>https://medium.com/@yurideigin/lab-made-cov2-genealogy-through-the-lens-of-gain-of-function-research-f96dd7413748<

>https://www.independentsciencenews.org/health/the-case-is-building-that-covid-19-had-a-lab-origin/<

May 2020 Paper on Spike Protein-ACE2 Binding Affinity

>https://arxiv.org/ftp/arxiv/papers/2005/2005.06199.pdf<

Full Text of Shi's Most Famous Study (November 2015)

>https://www.nature.com/articles/nm.3985<

Nov 2015 Article on the Controversy Sparked Over Shi's Study (Gain-of-Function Research)

>https://www.nature.com/news/engineered-bat-virus-stirs-debate-over-risky-research-1.18787#b1<

Richard Ebright in Nature February 2017

>https://www.nature.com/news/inside-the-chinese-lab-poised-to-study-world-s-most-dangerous-pathogens-1.21487<

Xiao Botao (South China University of Technology) Paper References the Work of WCDCP Researcher Tian Junhua

>https://img-prod.tgcom24.mediaset.it/images/2020/02/16/114720192-5eb8307f-017c-4075-a697-348628da0204.pdf<

Mr. Toy I. Reid
Foreign Policy Fellow
U.S. Senator Marco Rubio (FL)
284 Russell Senate Office Building
Washington, DC 20510
Direct: (b)(6)
Toy_Reid@rubio.senate.gov

(b)(6)

Received: 2 September 2020 | Revised: 16 October 2020

DOI: 10.1002/bies.202000240

The genetic structure of SARS-CoV-2 does not rule out a laboratory origin

SARS-COV-2 chimeric structure and furin cleavage site might be the result of genetic manipulation

Rossana Segreto[1] | **Yuri Deigin**[2]

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

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

Correspondence
Rossana Segreto, Department of Microbiology, University of Innsbruck, Technikerstraße 25, 6020 Innsbruck, Austria.
Email: Rossana.Segreto@uibk.ac.at

No external funding was received for this work.

Rossana Segreto and Yuri Deigin contributed equally to this study.

Abstract

Severe acute respiratory syndrome-coronavirus (SARS-CoV)-2′s origin is still controversial. Genomic analyses show SARS-CoV-2 likely to be chimeric, most of its sequence closest to bat CoV RaTG13, whereas its receptor binding domain (RBD) is almost identical to that of a pangolin CoV. Chimeric viruses can arise *via* natural recombination or human intervention. The furin cleavage site in the spike protein of SARS-CoV-2 confers to the virus the ability to cross species and tissue barriers, but was previously unseen in other SARS-like CoVs. Might genetic manipulations have been performed in order to evaluate pangolins as possible intermediate hosts for bat-derived CoVs that were originally unable to bind to human receptors? Both cleavage site and specific RBD could result from site-directed mutagenesis, a procedure that does not leave a trace. Considering the devastating impact of SARS-CoV-2 and importance of preventing future pandemics, researchers have a responsibility to carry out a thorough analysis of all possible SARS-CoV-2 origins.

KEYWORDS

BtCov/4991, furin cleavage site, Gain-of-function studies, pangolin CoV, RaTG13, receptor binding domain, SARS-CoV-2

INTRODUCTION

Nearly a year has passed since the outbreak of severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2) in Wuhan, China, and its origin is still controversial. Despite the international research effort conducted, a natural host, either direct or intermediate, has not yet been identified. The hypothesis that the Wuhan Huanan Seafood Wholesale Market was the first source for animal–human virus transmission has now been conclusively dismissed[i] and the few market samples that were collected showed only human-adapted SARS-CoV-2, with no traces of zoonotic predecessor strains[ii]. Almost all scientific papers published to date purport that SARS-CoV-2 has a natural origin, and the only published paper considering possible a lab origin[1] focuses on serial passage as the technique that could justify SARS-CoV-2 special

[i] Areddy, J. T. (**2020**). China rules out animal market and lab as coronavirus origin. *The Wall Street Journal*. https://www.wsj.com/articles/china-rules-out-animal-market-and-lab-as-coronavirus-origin-11590517508 (last accessed on Oct 15, 2020).

[ii] Zhan, S. H., Deverman, B. E., Chan, Y. A. (**2020**). SARS-CoV-2 is well adapted for humans. What does this mean for re-emergence? *BioRxiv*. https://doi.org/10.1101/2020.05.01.073262 (last accessed on Oct 15, 2020).

adaptation in humans. We identified two main SARS-CoV-2 features, (1) the presence of a furin cleavage site missing in other CoVs of the same group and (2) an receptor binding domain (RBD) optimized to bind to human cells[2] might be the result of lab manipulation techniques such as site-directed mutagenesis. The acquisition of both unique features by SARS-CoV-2 more or less simultaneously is less likely to be natural or caused only by cell/animal serial passage.

SARS-COV-2'S CLOSEST RELATIVES ARE BAT AND PANGOLIN CORONAVIRUSES

Zhou et al.[3] from the Wuhan Institute of Virology (WIV) were the first to identify and characterize a new coronavirus (CoV), 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 CoV sequence detected from a *Rhinolophus affinis* bat. RaTG13 is currently the closest phylogenetic relative for SARS-CoV-2 found,[4] but its complete genomic sequence was not published before the outbreak of SARS-CoV-2 and the original sample was collected in the Yunnan province (China) by the same group of WIV researchers in 2013. Zhou et al.[3] stated to have found a match between SARS-CoV-2 and a short region of RNA-dependent RNA polymerase (RdRp) of a CoV in their database and then fully sequenced the original sample collected in 2013, which they called RaTG13.

We discovered that the RdRp of RaTG13 has 100% nucleotide identity with the sequence BtCoV/4991 (KP876546), which was identified by Ge et al.[5] in a *Rhinolophus affinis* bat in the Yunnan province in 2013, same location and year as RaTG13. BtCoV/4991 was collected in a mine colonized by bats near Tongguanzhen, Mojiang, Yunnan. The WIV researchers were invited to investigate the mine after six miners there had contracted severe pneumonia in 2012[iii], and three of the miners have died.[6] The miners have been tasked with clearing out bat droppings in the mine, and the severity of their pneumonia correlated with the duration of exposure to the mine.[7] Four miners' samples subsequently underwent testing at WIV, where Immunoglobulin G (IgG) antibodies against SARS were identified in all samples.[8] 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 four miners in Yunnan retaining antibodies from the 2002–2004 SARS outbreak are negligible. 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 et al.[5] have 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. BtCoV/4991 was also different from other bat CoVs in the phylogenetic analysis carried out by Wang et al. in 2018[9] and Hou et al.[10] 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 later asserted to be two different coding names of the same strain, as their original authors at WIV registered the two strains as one entry in the Database of Bat-associated Viruses (DBatVir).[iv]

In late July 2020, Zhengli Shi, the leading CoV researcher from WIV, in an email interview [11] asserted the renaming of the RaTG13 sample and unexpectedly declared that the full sequencing of RaTG13 has been carried out as far back as in 2018 and not after the SARS-CoV-2 outbreak, as stated in Zhou et al.[3] The reversal in WIV's stance on when exactly RaTG13 was fully sequenced could have been due to the discovery by independent researchers into the origins of SARS-CoV-2 that the filenames of the raw sequencing reads deposited by WIV on May 19, 2020[v] seem to indicate that sequencing for RaTG13 was done in 2017 and 2018.[vi] However, no formal erratum about year of sequencing and sample renaming from the authors of Zhou et al. [3] has yet appeared, or as far as is currently known, has been submitted.

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 (GD), China.[12] The envelope protein of MP789 shows surprisingly 100% aminoacidic identity with the corresponding protein in RaTG13, in bat-SL-CoVZXC21 (MG772934.1), in bat-SL-CoVZC45 (MG772933.1) and in some early SARS-CoV-2 isolates (e.g. YP_009724392).[13] The envelope protein of CoVs is involved in critical aspects of the viral lifecycle, such as viral entry, replication and pathogenesis.[14]

BAT COVS HAVE BEEN THOROUGHLY STUDIED AND GENETICALLY MANIPULATED

Many studies have reported that bats are natural reservoirs for a broad diversity of potentially pathogenic SARS-like CoVs.[15,16] Some of these viruses can potentially directly infect humans[17], 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.[18] In order to evaluate the emergence potential of novel CoVs, researchers have created a number of chimeric CoVs, consisting of bat CoV backbones, normally unable to infect human cells, whose spike proteins were replaced by those from CoVs compatible with human ACE2. These chimeras were meant to simulate recombination events that might occur in nature.[19,20] Such gain-of-function experiments have raised a number of biosafety concerns and stirred controversy among researchers and the general public. One of the main arguments in favor of gain-of-function studies is the need to be prepared with an arsenal of drugs and vaccines for the next pandemic.[21]

[iii] Qiu, J. (**2020**). How China's 'Bat Woman' hunted down viruses from SARS to the new coronavirus. *Sci. Am.* https://www.scientificamerican.com/article/how-chinas-bat-woman-hunted-down-viruses-from-sars-to-the-new-coronavirus1/ (last accessed on Oct 15, 2020).

[iv] DBatVir – *The Database of Bat-Associated Viruses*. http://www.mgc.ac.cn/cgi-bin/DBatVir/main.cgi?func=accession&acc=MN996532 (last accessed on Oct 15, 2020).

[v] SRX8357956: amplicon sequences of RaTG13. https://www.ncbi.nlm.nih.gov/sra/SRX8357956 (last accessed on Oct 15, 2020).

[vi] Anon. (**2020**). Names of the RaTG13 amplicon sequences. https://web.archive.org/web/20200918174030/https://graph.org/RaTG13-Amplicon-Names-07-03 (last accessed on Oct 15, 2020).

By contrast, the main argument is that the lack of evidence

pandemic itself could be caused by those experiments, due to the risk of lab escape.[22,23]

In recent years, the field of corona-virology had been focused on pan-CoV therapies and vaccines, as evident from research conducted in the past 5 years,[24–27] as well as from media reports.[vii] Synthetically generating diverse panels of potential pre-emergent CoVs was declared a goal of active grants for the EcoHealth Alliance, which funded some of such research at WIV, in collaboration with laboratories in the USA and other international partners.[viii]

CREATING CHIMERIC COVS WITH NOVEL RBDS HAS GONE ON FOR DECADES

Researchers have been generating chimeric CoVs for over two decades, long before the advent of modern sequencing or genetic engineering techniques. For example, in 1999, a group from Utrecht University used targeted RNA recombination to create a "cat-and-mouse" CoV chimera: the RBDs of a feline and murine CoV were swapped, demonstrating that this exchange swapped also species tropism during *in vitro* experiments.[28]

In 2007, the Shi group at WIV created a series of "bat-man" CoV chimeric spike proteins while trying to determine what exactly confers CoVs the ability to jump from one species to another. The researchers used different segments of the spike protein of the human SARS virus to replace corresponding segments in the spike protein of a bat viral backbone. It was concluded that a relatively short region (aa 310 to 518) of the spike protein "was necessary and sufficient to convert Rp3-S into a huACE2-binding molecule,"[29] that is to provide the bat CoV spike protein with a novel ability of binding to a human ACE2 receptor.

In 2008, the Baric group at the University of North Carolina (UNC) took the WIV research one step further: instead of using human immunodeficiency viruses (HIV) pseudo-viruses with bat CoV spike proteins, a live chimeric CoV was created. Following the experiments of their 2007 WIV colleagues, the Baric group used a bat SARS-like CoV as a backbone and replaced its RBD with the RBD from human SARS.[30]

In 2015, the Shi and Baric groups joined forces and published probably the most famous gain-of-function virology paper, which described the creation of another synthetic chimeric virus.[19] This time the RBD of a mouse-adapted SARS backbone (SARS-MA15) was replaced by the RBD of RsSHC014, a bat strain previously isolated from Yunnan bats in 2011 by the Shi group. In 2016, the Baric group repeated their 2015 experiment using the same SARS-MA15 backbone and the RBD from Rs3367,[31] a close relative of RsSHC014 also previously found in Yunnan by WIV and renamed "WIV1" after live culturing.[17]

Probably the largest reported number of novel chimeric viruses created was described in a 2017 paper from the Shi group at WIV,[15] in

[vii] Kahn, J. (**2020**). How scientists could stop the next pandemic before it starts. *NYT* Magazine. https://www.nytimes.com/2020/04/21/magazine/pandemic-vaccine.html (last accessed on Oct 15, 2020).

[viii] Project Number 2R01AI110964-06, *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= (last accessed on Oct 15, 2020).

authors reported viruses using WIV1 as a backbone and transplanting into it various RBDs from bat SARS-like viruses. These viruses were collected over a span of 5 years from the same cave near Kunming, Yunnan Province, where the Shi group originally found Rs3367 and RsSHC014. Only two of the eight live chimeric viruses were successfully rescued, and those two strains were found to possess the ability to bind to the human ACE2 receptor, as confirmed by experiments in hACE2-expressing HeLa cells and RT-PCR quantification of viral RNA.

SARS-COV-2 SHARES ITS RBD WITH A PANGOLIN COV

The possibility that pangolins could be the intermediate host for SARS-CoV-2 has long been under discussion. [32–34] The biggest divergence between SARS-CoV-2 and RaTG13 is observed in the RBD of their spike proteins.[4] Although its overall genome similarity is lower to SARS-CoV-2 than that of RaTG13, the MP789 pangolin strain isolated from GD pangolins has an almost identical RBD 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.[35] 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 and bats.[ix] 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.[12] Recombination events between the RBD of CoV from pangolins and RaTG13-like backbone could have produced SARS-CoV-2 as chimeric strain. For such recombination to occur naturally, the two viruses must have infected the same cell in the same organism simultaneously, a rather improbable event considering the low population density of pangolins and the scarce presence of CoVs in their natural populations.[x] Moreover, receptor binding studies of reconstituted RaTG13 showed that it does not bind to pangolin ACE2.[xi]

THE FURIN CLEAVAGE SITE: THE 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 difference is the acquisition in the

[ix] Piplani, S., Singh, P. K., Winkler, D. A., Petrovsky, N. (**2020**). In silico comparison of spike protein-ACE2 binding affinities across species; significance for the possible origin of the SARS-CoV-2 virus.*arXiv*. http://arxiv.org/abs/2005.06199 (last accessed on Oct 15, 2020).

[x] Lee, J., Hughes, T., Lee, M.-H., Field, H., Rovie-Ryan, J. J., Sitam, F. T., … Daszak, P. (**2020**). No evidence of coronaviruses or other potentially zoonotic viruses in Sunda pangolins (*Manis javanica*) entering the wildlife trade *via* Malaysia. *BioRxiv*. https://doi.org/10.1101/2020.06.19.158717 (last accessed on Oct 15, 2020).

[xi] Mou, H., Quinlan, B. D., Peng, H., Guo, Y., Peng, S., Zhang, L., … Farzan, M. (**2020**). Mutations from bat ACE2 orthologs markedly enhance ACE2-Fc neutralization of SARS-CoV-2. *BioRxiv*. https://doi.org/10.1101/2020.06.29.178459 (last accessed on Oct 15, 2020).

FIGURE 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 *FauI* enzyme restriction site

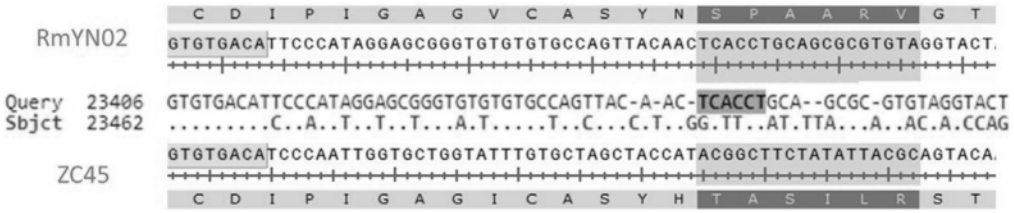


FIGURE 2 Alignment of nucleotide and amino acid sequences of the S protein from bat-SL-CoVZC45 (MG772933.1) and RmYN02 at the S1/S2 junction site. No insertions of nucleotides possibly evolving in a furin cleavage site can be observed (in blue)

spike protein of SARS-CoV-2 of a cleavage site activated by a host-cell enzyme furin, previously not identified in other beta-CoVs of lineage b[36] and similar to that of Middle East respiratory syndrome (MERS) coronavirus.[35] Host protease processing plays a pivotal role as a species and tissue barrier and engineering of the cleavage sites of CoV spike proteins modifies virus tropism and virulence.[37] The ubiquitous expression of furin in different organs and tissues have conferred to SARS-CoV-2 the ability to infect organs usually invulnerable to other CoVs, leading to systemic infection in the body.[38] Cell-cultured SARS-CoV-2 that was missing the above-mentioned cleavage site caused attenuated symptoms in infected hamsters,[39] and mutagenesis studies have confirmed that the polybasic furin site is essential for SARS-CoV-2′s ability to infect human lung cells.[40]

The polybasic furin site in SARS-CoV-2 was created by a 12-nucleotide insert TCCTCGGCGGGC coding for a PRRA amino acid sequence at the S1/S2 junction (Figure 1). Interestingly, the two joint arginines are coded by two CGGCGG codons, which are rare for these viruses: only 5% of arginines are coded by CGG in SARS-CoV-2 or RaTG13, and CGGCGG in the new insert is the only doubled instance of this codon in SARS-CoV-2. The CGGCGG insert includes a *FauI* restriction site, of which there are six instances in SARS-CoV-2 and four instances in RaTG13 (and two in MP789). The serendipitous location of the *FauI* site could allow using restriction fragment length polymorphism (RFLP) techniques[41] for cloning[42] or screening for mutations,[43] as the new furin site is prone to deletions *in vitro*.[39,44]

A study by Zhou et al.[45] 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 bat-SL-CoVZC45 and bat-SL-CoVZXC21, no insertions are apparent, just nucleotide mutations (Figure 2).

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

The insertion of the furin cleavage site in SARS-CoV-2 is not in frame with the rest of the sequence, when compared with the MP789 and the RaTG13 sequences (Figure 3). Therefore, it is possible to exclude that such insertion could have originated by polymerase slippage or by releasing and repriming, because insertion mutations generated by these mechanisms have been postulated to maintain the reading frame of the viral sequence.[46] The possibility that the furin cleavage site could have been acquired by recombination has been recently

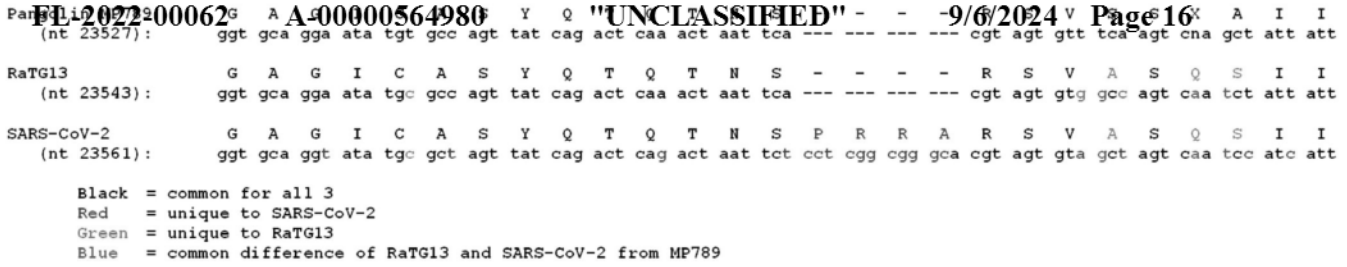


FIGURE 3 Alignment of nucleotide and amino acid sequences of the S protein from RaTG13 (MN996532), MP789 (MT084071) and SARS-CoV-2 (NC045512.2) at the S1/S2 site. The common nucleotides and amino acids are given in black, SARS-CoV-2 unique nucleotides and amino acids in red, RaTG13 unique nucleotides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

questioned by Seyran et al.,[47] because the SARS-CoV-2 spike protein seems to lack any further recombination event in contrast with the recombination model of other CoVs.

CRITIQUE OF "THE PROXIMAL ORIGIN OF SARS-COV-2"

Due to the broad-spectrum of research conducted over almost 20 years on bat SARS-CoVs justified by their potential to spill over from animal to human,[48] a possible synthetic origin by laboratory engineering of SARS-CoV-2 cannot be excluded. The widely cited article of Andersen et al.[2] stated that SARS-CoV-2 has most likely a natural origin. The main argument brought by the authors 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 et al.,[49] 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.[15] All amino acids in the RBD have been extensively analyzed and new models to predict ACE2 affinity are available.[50] In this regard, BatCoV Rs3367 (99.9% identity to WIV1) has been shown to share with SARS-CoV-2 four out of six critical residues in the RBD. Considering that WIV1 was shown to directly bind to hACE2, the same assumption could easily have been made about SARS-CoV-2 RBD.[51]

As described above, creation of chimeric viruses has been carried out over the years with the purpose of studying 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[12], 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. Subsequent serial cell or animal passage, as described by Sirotkin & Sirotkin [1] could have provided the perfect adaptation of the RBD to the hACE2.

Regarding the furin cleavage site, Andersen et al.[2] 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 understand virus transmissibility in different hosts.[50]

Andersen et al.[2] also state, based on the work of Almazan et al.[53] that "the genetic data irrefutably show that SARS-CoV-2 is not derived from any previously used virus backbone." In the last 6 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 et al.[2] affirm 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 et al.[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 with human lung implants and autologous human immune system[57] or in mice expressing the hACE2 receptor.[31] To overcome problems of bat CoV isolation, experiments based on direct inoculation of bat CoV in suckling rats have been carried out.[58] Humanized mice, ferrets, primates and/or other animals with similar ACE2 conformation could have all been used for serial passage experiments, as described in detail by Sirotkin and Sirotkin.[1]

Andersen et al.[2] 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 was published before the publication of the results. It is likely that the authors assumed that the results were never intended to be published.

It is important to mention that RaTG13 and the pangolin CoV sequences from smuggled pangolins confiscated in the GD province in March 2019, and to which most of published papers supporting a natural origin of SARS-CoV-2 refer,[2] have recently been questioned as to the accuracy of their assembly data[xii] and require further analyses to prove their correctness.[xiii,xiv] It should also be noted that *in vitro* receptor binding studies of reconstituted RaTG13 yielded some peculiar results.[xi] The most surprising observation was that RaTG13, unlike SARS-CoV-2, is unable to bind ACE2 in *R. macrotis* bats, a close relative of RaTG13's purported host, *R. affinis*[59] (whose ACE2 receptor has not yet been tested). At the same time, RaTG13 was observed to bind hACE2[60], but not as well as ACE2 of rats and mice, to which SARS-CoV-2 did not bind at all. Is it possible that just as SARS-MA15 was a mouse-adapted strain of SARS, RaTG13 is actually a mouse-adapted version of a CoV extracted from the Mojiang cave, rather than a strain obtained from a bat fecal swab? Unfortunately, the RaTG13 sample has been exhausted and it is no longer available for external examination,[11] which is unfortunate given a number of inconsistencies in its sequencing raw data. Also, the status and availability of the Mojiang miners' samples remain as well an open and highly relevant question. Several samples from the miners have been collected[7,8] and likely stored, and it would be of great value to test them for the presence of SARS-CoV-2-like CoVs.

Another open question is the reason for modification and subsequent deletion of WIV's own viral database. In May 2020, several media outlets have reported that the change tracking system of WIV's internal database showed that the database was renamed from "Wildlife-borne viral pathogen database" to "Bat and rodent-borne viral pathogen database," and its description was edited to replace instances of "wild animal" by "bat and rodent"; in addition, mention of "arthropod vectors" was deleted.[xv] The database description reported that it contained over 60 Mb of data in structured query language (SQL) format, but at as of early May 2020 the download link no longer worked.[xvi] Subsequently, the database page was taken down in its entirety but its snapshot is still available on Web Archive.[xvii] It is possible that other international CoV labs might have downloaded the SQL archive of the WIV database before it was taken down, in which case such groups should make those data publicly available.

HOW COULD THE VIRUS HAVE ESCAPED FROM A LAB?

The leak of highly dangerous pathogens from laboratories is not a rare event and occurrences have been documented in several countries. The most notable lab leak known is the 1977 H1N1 lab escape from China that caused a worldwide pandemic.[61] The most recent one is the November 2019 outbreak of brucellosis that occurred in two research centers in Lanzhou, China, infecting over 100 students and staff members.[62] Several lab escapes of the first SARS virus have been reported as well: in the summer of 2003 in Singapore,[63] then in December 2003 in Taiwan,[xviii] and in the spring of 2004 twice in China.[xix]

Concerns about WIV's lab safety were raised in 2018 by U.S. Embassy officials after visiting the Institute and having an interview with Zhengli Shi. The lab auditors summarized their worries in subsequent diplomatic cables to Washington.[xx] Chinese experts have also raised concerns about lab safety in their own country, lamenting that "lab trash can contain man-made viruses, bacteria or microbes" and that "some researchers discharge laboratory materials into the sewer after experiments without a specific biological disposal mechanism."[xxi]

American labs have also had their share of safety issues. Recently, research operations in the Biosafety level (BSL)-4 United States Army Medical Research Institute of Infectious Diseases (USAMRIID) facility in Fort Detrick were interrupted in August 2019 following safety violations, in particular, relating to the disposal of infective materials.[xxii] Other US labs have been cited for safety issues as well.[22]

A number of scenarios causing SARS-CoV-2 to leak from a lab can be hypothesized. For example, an infected animal could have escaped from a lab or it could have scratched or bitten a worker (a concern raised in 2017 about the establishment of a BSL-4 primate vaccine testing facility in Kunming, Yunnan[64]), or a researcher could have accidentally stuck themselves with inoculate (as happened in two cases in Russia[xxiii]). Until 2020, CoVs were not considered particularly deadly or virulent. SARS-like CoVs did not require BSL-4 and could be manipulated under BSL-2 and BSL-3[42] conditions, making an accidental leak more likely. Aerosol experiments with CoVs[65] could result in lab leak as well, because a failure in the equipment used could go unnoticed for a long time before infection of lab workers is detected. Finally, the virus

[xii] Zhang, D. (2020). Anomalies in BatCoV/RaTG13 sequencing and provenance. *Zenodo*. https://zenodo.org/record/3969272 (last accessed on Oct 15, 2020).

[xiii] Singla, M., Ahmad, S., Gupta, C., Sethi, T. (**2020**). De novo assembly of RaTG13 genome reveals inconsistencies further obscuring SARS-CoV-2 origins. *Preprints*. https://doi.org/10.20944/preprints202008.0595.v1 (last accessed on Oct 12, 2020).

[xiv] Chan, Y. A., Zhan, S. H. (**2020**). Single source of pangolin CoVs with a near identical spike RBD to SARS-CoV-2. *BioRxiv*. https://doi.org/10.1101/2020.07.07.184374 (last accessed on Oct 15, 2020).

[xv] Devine, M. (**2020**). What is China covering up about the coronavirus? *NYT* Magazine. https://nypost.com/2020/05/06/what-is-china-covering-up-about-the-coronavirus-devine/ (last accessed on Oct 12, 2020).

[xvi] https://twitter.com/ydeigin/status/1259891518468427776 (last accessed on Oct 15, 2020).

[xvii] Bat and rodent-borne viral pathogen database. https://web.archive.org/web/20200529174243/http://csdata.org/p/308/ (last accessed on Oct 15, 2020).

[xviii] Reuters (**2003**). SARS case confirmed in Taiwan. *Wired*. https://www.wired.com/2003/12/sars-case-confirmed-in-taiwan/ (last accessed on Oct 13, 2020).

[xix] Walgate, R. (**2004**). SARS escaped Beijing lab twice. *The Scientist Magazine*. https://www.the-scientist.com/news-analysis/sars-escaped-beijing-lab-twice-50137 (last accessed on Oct 15, 2020).

[xx] Rogin, J. (**2020**). State Department cables warned of safety issues at Wuhan lab studying bat coronaviruses. *The Washington Post*. https://www.washingtonpost.com/opinions/2020/04/14/state-department-cables-warned-safety-issues-wuhan-lab-studying-bat-coronaviruses/ (last accessed on Oct 15, 2020).

[xxi] Caiyu, L., Shumei, L. (**2020**). Biosafety guideline issued to fix chronic management loopholes at virus labs. *Global Times*. https://www.globaltimes.cn/content/1179747.shtml (last accessed on Oct 15, 2020).

[xxii] Grady, D. (**2020**). Deadly germ research is shut down at army lab over safety concerns. *NYT* Magazine. https://www.nytimes.com/2019/08/05/health/germs-fort-detrick-biohazard.html (last accessed on Oct 15, 2020).

[xxiii] Miller, J. (**2004**). Russian scientist dies in Ebola accident at former weapons Lab. *NYT* Magazine. https://www.nytimes.com/2004/05/25/world/russian-scientist-dies-in-ebola-accident-at-former-weapons-lab.html (last accessed on Oct 15, 2020).

could potentially have leaked through the sewage system if proper waste disposal and/or decontamination procedures were not followed.

CONCLUSIONS AND OUTLOOK

On the basis of our analysis, an artificial origin of SARS-CoV-2 is not a baseless conspiracy theory that is to be condemned[66] and researchers have the responsibility to consider all possible causes for SARS-CoV-2 emergence. The insertion of human-adapted pangolin CoV RBD obtained by cell/animal serial passage and furin cleavage site could arise from site-directed mutagenesis experiments, in a context of evolutionary studies or development of pan-CoV vaccines or drugs. A recent article in Nature[67] affirms that a laboratory origin for SARS-CoV-2 cannot be ruled out, as researchers could have been infected accidentally, and that gain-of-function experiments resulting in SARS-CoV-2 could have been performed at WIV. 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 and it would not leave any trace. Modern technologies based on synthetic genetics platforms allow the reconstruction of viruses based on their genomic sequence, without the need of a natural isolate.[68]

A thorough investigation on strain collections and research records in all laboratories involved in CoV research before SARS-CoV-2 outbreak is urgently needed. Special attention should be paid to strains of CoVs that were generated in virology laboratories but have not yet been published, as those possibly described in the deleted WIV database. Because finding a possible natural host could take years, as with the first SARS,[67] or never succeed, equal priority should be given to investigating natural and laboratory origins of SARS-CoV-2.

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

ACKNOWLEDGMENTS

We are very grateful to Prof. Allan Krill (NTNU) for proof reading the manuscript, all the valuable comments and being open-minded about controversial hypotheses; Prof. Heribert Insam (Head of the Department of Microbiology; University of Innsbruck) for his support and Dr. Lawrence Sellin for all the useful information. A special thanks goes to Dr. Fernando Castro-Chavez (former Post-Doc at the New York Medical College) for his support with Research Gate. We are very thankful to René Bergelt, for having discovered the database that confirmed our finding that BtCoV4991 and RaTG13 refer to the same sample. Finally, we are extremely grateful to members of the D.R.A.S.T.I.C. (Decentralised Radical Autonomous Search Team Investigating COVID-19) Twitter group for all their work in uncovering many previously unpublished facts about SARS-CoV-2 and its relative strains. In particular, we are grateful to Luigi Warren for continuously probing the possible connection of the 2012 Mojiang pneumonia outbreak to WIV and SARS-CoV-2, to @TheSeeker268 for finding Chinese-language 2013 Xu MSc and 2016 Huang PhD theses, which have confirmed the SARS-like viral nature of the 2012 Mojiang pneumonia outbreak and have elucidated WIV's role in investigating that outbreak,[xxiv] including WIV's collection of the 4991/RaTG13 strain from the Mojiang mine, and to Francisco de Asis de Ribera Martin for providing us the English translation of the two theses, and also discovering the RaTG13 amplicon dates.

CONFLICT OF INTEREST

Rossana Segreto and Yuri Deigin do not have any conflicts of interest.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

ORCID

Rossana Segreto https://orcid.org/0000-0002-2566-7042

REFERENCES

1. Sirotkin, K., & Sirotkin, D. (2020). Might SARS-CoV-2 have arisen *via* serial passage through an animal host or cell culture? A potential explanation for much of the novel coronavirus' distinctive genome. *BioEssays*, *42*, 1-7. https://doi.org/10.1002/bies.202000091
2. Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., & Garry, R. F. (2020). The proximal origin of SARS-CoV-2. *Nat. Med.*, *26*, 450-452. https://doi.org/10.1038/s41591-020-0820-9
3. Zhou, P., Yang, X. L., Wang, X. G., Hu, B., Zhang, L., Zhang, W., … Shi, Z. L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, *579*, 270-273. https://doi.org/10.1038/s41586-020-2012-7
4. Cagliani, R., Forni, D., Clerici, M., & Sironi, M. (2020). Computational inference of selection underlying the evolution of the novel coronavirus, severe acute respiratory syndrome coronavirus 2. *J. Virol.*, *94*, 1-11. https://doi.org/10.1128/jvi.00411-20
5. Ge, X. Y., Wang, N., Zhang, W., Hu, B., Li, B., Zhang, Y. Z., … Shi, Z. L. (2016). Coexistence of multiple coronaviruses in several bat colonies in an abandoned mineshaft. *Virol. Sin.*, *31*, 31-40. https://doi.org/10.1007/s12250-016-3713-9
6. Wu, Z., Yang, L., Yang, F., Ren, X., Jiang, J., Dong, J., … Jin, Q. (2014). Novel henipa-like virus, mojiang paramyxovirus, in rats, China, 2012. *Emerg. Infect. Dis.*, *20*, 1064-1066. https://doi.org/10.3201/eid2006.131022
7. Xu, L. (2013). The analysis of 6 patients with severe pneumonia caused by unknown viruses (Master's Thesis). Kunming Medical University, Emergency Medicine (professional degree). http://eng.oversea.cnki.net/Kcms/detail/detail.aspx?filename=1013327523.nh&dbcode=CMFD&dbname=CMFD2014
8. Huang, C. (2016). Novel virus discovery in bat and the exploration of receptor of bat coronavirus HKU9 (PhD Thesis). Chinese Center for Disease Control and Prevention. http://eng.oversea.cnki.net/kcms/detail/detail.aspx?dbcode=CDFD&dbname=CDFDLAST2018&filename=1017118517.nh
9. Wang, N., Luo, C., Liu, H., Yang, X., Hu, B., Zhang, W., … Shi, Z. (2019). Characterization of a new member of alphacoronavirus with unique genomic features in *Rhinolophus* bats. *Viruses*, *11*(4), 379. https://doi.org/10.3390/v11040379
10. Chen, L., Liu, W., Zhang, Q., Xu, K., Ye, G., Wu, W., … Liu, Y. (2020). RNA based mNGS approach identifies a novel human coronavirus from two individual pneumonia cases in 2019 Wuhan outbreak. *Emerg. Microbes Infect.*, *9*, 313-319. https://doi.org/10.1080/22221751.2020.1725399

[xxiv] https://twitter.com/ydeigin/status/1262686286898397189 (last accessed on Oct 15, 2020).

11. Cohen, J. (2020). Wuhan coronavirus hunter Shi Zhengli speaks out. *Science*, *369*, 487-488. https://doi.org/10.1126/science.369.6503.487
12. Liu, P., Chen, W., & Chen, J. P. (2019). Viral metagenomics revealed sendai virus and coronavirus infection of malayan pangolins (*Manis javanica*). *Viruses*, *11*(11), 979. https://doi.org/10.3390/v11110979
13. Bianchi, M., Benvenuto, D., Giovanetti, M., Angeletti, S., Ciccozzi, M., & Pascarella, S. (2020). Sars-CoV-2 envelope and membrane proteins: Structural differences linked to virus characteristics? *Biomed. Res. Int.*, *2020*. https://doi.org/10.1155/2020/4389089
14. Schoeman, D., & Fielding, B. C. (2019). Coronavirus envelope protein: Current knowledge. *Virol. J.*, *16*, 1-22. https://doi.org/10.1186/s12985-019-1182-0
15. Hu, B., Zeng, L. P., Yang, X. Lou, Ge, X. Y., Zhang, W., Li, B., … Shi, Z. L. (2017). Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLoS Pathog.*, *13*, 1-27. https://doi.org/10.1371/journal.ppat.1006698
16. Fan, Y., Zhao, K., Shi, Z. L., & Zhou, P. (2019). Bat coronaviruses in China. *Viruses*, *11*(3), 210-. https://doi.org/10.3390/v11030210
17. Ge, X. Y., Li, J. L., Yang, X. Lou, Chmura, A. A., Zhu, G., Epstein, J. H., … Shi, Z. L. (2013). Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature*, *503*, 535-538. https://doi.org/10.1038/nature12711
18. Graham, R. L., & Baric, R. S. (2010). Recombination, reservoirs, and the modular spike: Mechanisms of coronavirus cross-species transmission. *J. Virol.*, *84*, 3134-3146. https://doi.org/10.1128/jvi.01394-09
19. Menachery, V. D., Yount, B. L., Debbink, K., Agnihothram, S., Gralinski, L. E., Plante, J. A., … Baric, R. S. (2015). A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nat. Med.*, *21*, 1508-1513. https://doi.org/10.1038/nm.3985
20. Johnson, B. A., Graham, R. L., & Menachery, V. D. (2018). Viral metagenomics, protein structure, and reverse genetics: Key strategies for investigating coronaviruses. *Virology*, *517*, 30-37. https://doi.org/10.1016/j.virol.2017.12.009
21. Racaniello, V. (2016). Moving beyond metagenomics to find the next pandemic virus. *PNAS*, *113*, 2812-2814. https://doi.org/10.1073/pnas.1601512113.
22. Weiss, S., Yitzhaki, S., & Shapira, S. C. (2015). Lessons to be learned from recent biosafety incidents in the United States. *Isr. Med. Assoc. J.*, *17*, 269-273. https://doi.org/10.1073/pnas.1601512113
23. Casadevall, A., & Imperiale, M. J. (2014). 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*, *5*, 1-5. https://doi.org/10.1128/mBio.01730-14
24. Agostini, M. L., Andres, E. L., Sims, A. C., Graham, R. L., Sheahan, T. P., Lu, X., … Denison, M. R. (2018). Coronavirus susceptibility to the antiviral remdesivir (GS-5734) is mediated by the viral polymerase and the proofreading exoribonuclease. *MBio*, *9*, 1-15. https://doi.org/10.1128/mBio.00221-18
25. Xia, S., Liu, M., Wang, C., Xu, W., Lan, Q., Feng, S., … Lu, L. (2020). 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.*, *30*, 343-355. https://doi.org/10.1038/s41422-020-0305-x
26. Totura, A. L., & Bavari, S. (2019). Broad-spectrum coronavirus antiviral drug discovery. *Expert Opin. Drug Discov.*, *14*, 397-412. https://doi.org/10.1080/17460441.2019.1581171
27. Wang, Y., Sun, Y., Wu, A., Xu, S., Pan, R., Zeng, C., … Guo, D. (2015). Coronavirus nsp10/nsp16 Methyltransferase can be targeted by nsp10-derived peptide in vitro and in vivo to reduce replication and pathogenesis. *J. Virol.*, *89*, 8416-8427. https://doi.org/10.1128/jvi.00948-15
28. Kuo, L., Godeke, G. J., Raamsman, M. J. B., Masters, P. S., & Rottier, P. J. M. (2000). Retargeting of coronavirus by substitution of the spike glycoprotein ectodomain: Crossing the host cell species barrier. *J. Virol.*, *74*, 1393-1406. https://doi.org/10.1128/jvi.74.3.1393-1406.2000
29. Maier, H. J., Bickerton, E., & Britton, P. (2015). *Coronaviruses – Methods and protocols*. London: Humana Press.
30. Becker, M. M., Graham, R. L., Donaldson, E. F., Rockx, B., Sims, A. C., Sheahan, T., … Denison, M. R. (2008). Synthetic recombinant bat SARS-like coronavirus is infectious in cultured cells and in mice. *PNAS*, *105*, 19944-19949. https://doi.org/10.1073/pnas.0808116105
31. Menachery, V. D., Yount, B. L., Sims, A. C., Debbink, K., Agnihothram, S. S., Gralinski, L. E., … Baric, R. S. (2016). SARS-like WIV1-CoV poised for human emergence. *PNAS*, *113*, 3048-3053. https://doi.org/10.1073/pnas.1517719113
32. Li, X., Zai, J., Zhao, Q., Nie, Q., Li, Y., Foley, B. T., & Chaillon, A. (2020). Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. *J. Med. Virol.*, *92*, 602-611. https://doi.org/10.1002/jmv.25731
33. Lam, T. T. Y., Jia, N., Zhang, Y. W., Shum, M. H. H., Jiang, J. F., Zhu, H. C., … Cao, W. C. (2020). Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature*, *583*, 282-285. https://doi.org/10.1038/s41586-020-2169-0
34. Xiao, K., Zhai, J., Feng, Y., Zhou, N., Zhang, X., Zou, J. J., … Shen, Y. (2020). Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature*, *583*, 286-289. https://doi.org/10.1038/s41586-020-2313-x
35. Zhang, T., Wu, Q., & Zhang, Z. (2020). Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. *Curr. Biol.*, *30*, 1346-1351.E2. https://doi.org/10.1016/j.cub.2020.03.022
36. Coutard, B., Valle, C., de Lamballerie, X., Canard, B., Seidah, N. G., & Decroly, E. (2020). The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin- like cleavage site absent in CoV of the same clade. *Antivir. Res.*, *176*, 104742. https://doi.org/10.1016/j.antiviral.2020.104742
37. Letko, M., Marzi, A., & Munster, V. (2020). Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat. Microbiol.*, *5*, 562-569. https://doi.org/10.1038/s41564-020-0688-y
38. Wang, Q., Qiu, Y., Li, J. Y., Zhou, Z. J., Liao, C. H., & Ge, X. Y. (2020). A unique protease cleavage site predicted in the spike protein of the novel pneumonia coronavirus (2019-nCoV) potentially related to viral transmissibility. *Virol. Sin.*, *35*, 337-339. https://doi.org/10.1007/s12250-020-00212-7
39. Lau, S., Wang, P., Mok, B. W., Zhang, A. J., Chu, H., Lee, A. C., … Chen, H. (2020). Attenuated SARS-CoV-2 variants with deletions at the S1 / S2 junction. *Emerg. Microbes Infect.*, *9*, 837-842. https://doi.org/10.1080/22221751.2020.1756700
40. Hoffmann, M., & Kleine-Weber, H. (2020). A multibasic cleavage site in the spike protein of SARS-CoV-2 is essential for infection of human lung cells. *Mol. Cell.*, *78*, 779-784.E5. https://doi.org/10.1016/j.molcel.2020.04.022
41. Kaundun, S. S., Marchegiani, E., Hutchings, S. J., & Baker, K. (2019). Derived polymorphic amplified cleaved sequence (dPACS): A novel PCR-RFLP procedure for detecting known single nucleotide and deletion – insertion polymorphisms. *Int. J. Mol. Sci.*, *20*(13), 3193. https://doi.org/10.3390/ijms20133193
42. Zeng, L. P., Gao, Y. T., Ge, X. Y., Zhang, Q., Peng, C., Yang, X. L., … Shi, Z. L. (2016). Bat severe acute respiratory syndrome-like coronavirus WIV1 encodes an extra accessory protein, ORFX, involved in modulation. *J. Virol.*, *90*, 6573-6582. https://doi.org/10.1128/JVI.03079-15
43. Khan, S. G., Muniz-Medina, V., Shahlavi, T., Baker, C. C., Inui, H., Ueda, T., … Kraemer, K. H. (2002). 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.*, *30*, 3624-3631. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC134237/.
44. Liu, Z., Zheng, H., Lin, H., Li, M., Yuan, R., Peng, J., … Lu, J. (2020). Identification of common deletions in the spike protein of severe acute

//doi.org/10.1128/JVI.00790-20

45. Zhou, H., Chen, X., Hu, T., Li, J., Song, H., Liu, Y., … Shi, W. (2020). 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.*, *30*, 2196-2203.E3. https://doi.org/10.1016/j.cub.2020.05.023
46. Steinhauer, D. A. (1999). Role of hemagglutinin cleavage for the pathogenicity of influenza virus. *Virology*, *258*, 1-20. https://doi.org/10.1006/viro.1999.9716
47. Seyran, M., Pizzol, D., Adadi, P., El-Aziz, T. M. A., Hassan, S. S., Soares, A., … Brufsky, A. M. (2020). Questions concerning the proximal origin of SARS-CoV-2. *J. Med. Virol.*, 03. https://doi.org/10.1002/jmv.26478
48. Wang, L. F., & Anderson, D. E. (2019). Viruses in bats and potential spillover to animals and humans. *Curr. Opin. Virol.*, *34*, 79-89. https://doi.org/10.1016/j.coviro.2018.12.007
49. Wan, Y., Shang, J., Graham, R., Baric, R. S., & Li, F. (2020). Receptor recognition by the novel coronavirus from Wuhan: An analysis based on decade-long structural studies of SARS coronavirus. *J. Virol.*, *94*(7), 1-9. https://doi.org/10.1128/jvi.00127-20
50. Cui, J., Li, F., & Shi, Z. L. (2019). Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.*, *17*, 181-192. https://doi.org/10.1038/s41579-018-0118-9
51. Fraguas Bringas, C., & Booth, D. (2020). Identification of a SARS-like bat coronavirus that shares structural features with the spike glycoprotein receptor-binding domain of SARS-CoV-2. *Access Microbiol.*, 10-17. https://doi.org/10.1099/acmi.0.000166.
52. Mallapati, S. (2020). Why does the coronavirus spread so easily between people? *Nature*, *579*, 183. https://www.nature.com/articles/d41586-020-00660-x.
53. Almazán, F., Sola, I., Zuñiga, S., Marquez-Jurado, S., Morales, L., Becares, M., & Enjuanes, L. (2014). Coronavirus reverse genetic systems: Infectious clones and replicons. *Virus Res.*, *189*, 262-270. https://doi.org/10.1016/j.virusres.2014.05.026
54. Cheng, J., Zhao, Y., Xu, G., Zhang, K., Jia, W., Sun, Y., … Zhang, G. (2019). The S2 subunit of QX-type infectious bronchitis coronavirus spike protein is an essential determinant of neurotropism. *Viruses*, *11*(10), 972. https://doi.org/10.3390/v11100972
55. Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C. L., Abiona, O., … McLellan, J. S. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, *367*, 1260-1263. https://doi.org/10.1126/science.abb2507
56. Du, L., Tai, W., Yang, Y., Zhao, G., Zhu, Q., Sun, S., … Li, F. (2016). Introduction of neutralizing immunogenicity index to the rational design of MERS coronavirus subunit vaccines. *Nat. Commun.*, *7*, 1-9. https://doi.org/10.1038/ncomms13473
57. Wahl, A., De, C., Abad Fernandez, M., Lenarcic, E. M., Xu, Y., Cockrell, A. S., … Garcia, J. V. (2019). Precision mouse models with expanded tropism for human pathogens. *Nat. Biotechnol.*, *37*, 1163-1173. https://doi.org/10.1038/s41587-019-0225-9
58. Hu, B., Zeng, L., Yang, X., Ge, X., Zhang, W., Li, B., … Shi, Z. (2017). Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLoS Pathog.*, *13*(11), e1006698. https://doi.org/10.1371/journal.ppat.1006698
59. Hron, T., Farkašová, H., Gifford, R. J., Benda, P., Hulva, P., Görföl, T., … Elleder, D. (2018). Remnants of an ancient deltaretrovirus in the genomes of horseshoe bats (Rhinolophidae). *Viruses*, *10*(4), 185-. https://doi.org/10.3390/v10040185
60. Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., … Li, F. (2020). Structural basis of receptor recognition by SARS-CoV-2. *Nature*, *581*, 221-224. https://doi.org/10.1038/s41586-020-2179-y
61. Wertheim, J. O. (2010). The re-emergence of H1N1 influenza virus in 1977: A cautionary tale for estimating divergence times using biologically unrealistic sampling dates. *PLoS ONE*, *5*, 2-5. https://doi.org/10.1371/journal.pone.0011184
62. Cyranoski, D. (2019). Chinese institutes investigate pathogen outbreaks in lab workers. *Nature*, https://www.nature.com/articles/d41586-019-03863-z
63. Lim, P. L., Kurup, A., Gopalakrishna, G., Chan, K. P., Wong, C. W., & Leo, Y. S. (2004). Laboratory-acquired severe acute respiratory syndrome. *N. Engl. J. Med.*, *350*, 1740-1745. https://doi.org/10.1056/NEJMoa032565
64. Cyranoski D. (2017). Inside the Chinese lab poised to study world's most dangerous pathogens. *Nature*, *542*, 399-401 https://doi.org/10.1038/nature.2017.21487
65. Totura, A., Livingston, V., Frick, O., Dyer, D., Nichols, D., & Nalca, A. (2020). Small particle aerosol exposure of African Green Monkeys to MERS-CoV as a model for highly pathogenic coronavirus infection. *Emerg. Infect. Dis.*, *26*. https://doi.org/10.3201/eid2612.201664.
66. Calisher, C., Carroll, D., Colwell, R., Corley, R. B., Daszak, P., Drosten, C., … Turner, M. (2019). Statement in support of the scientists, public health professionals, and medical professionals of China combatting COVID-19. *The Lancet*, *395*, E42-E43. https://doi.org/10.1016/S0140-6736(20)30418-9
67. Cyranoski, D. (2020). The biggest mystery: What it will take to trace the coronavirus source. *Nature*, https://doi.org/10.1038/d41586-020-01541-z.
68. Thao, T. T. N., Labroussaa, F., Ebert, N., V'kovski, P., Stalder, H., Portmann, J., … Thiel, V. (2020). Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform. *Nature*, *582*, 561-565. https://doi.org/10.1038/s41586-020-2294-9

How to cite this article: Segreto, R., & Deigin, Y. (2020). The genetic structure of SARS-CoV-2 does not rule out a laboratory origin. *BioEssays*, e2000240. https://doi.org/10.1002/bies.202000240.

From:	"Stilwell, David R" (b)(6) @state.gov>
To:	SES_FO Paper <SES_FOPaper@state.gov>
CC:	Feith, David (b)(6) @state.gov>
Subject:	Alina Chan Must Read
Date:	Tue, 5 Jan 2021 01:43:12 +0000

Mr Secretary
Among scientists, suspicions about the WIV have been ongoing since January—this article lays out the sequence well. It also paints a disturbing picture of supposedly objective scientists actively supporting a theory they know to be false:

When word spread in January that a novel coronavirus had caused an outbreak in Wuhan—which is a thousand miles from where the bats that carry this lineage of viruses are naturally found—many experts were quietly alarmed. There was no proof that the lab was the source of the virus, but the pieces fit.

Despite the evidence, the scientific community quickly dismissed the idea. Peter Daszak, president of EcoHealth Alliance, which has funded the work of the Wuhan Institute of Virology and other labs searching for new viruses, called the notion "preposterous," and many other experts echoed that sentiment.

That wasn't necessarily what every scientist thought in private, though. "They can't speak directly," one scientist told me confidentially, referring to the virology community's fear of having their comments sensationalized in today's politically charged environment. "Many virologists don't want to be hated by everyone in the field."

This researcher, Alina Chan, was too young and too idealistic to be cowed by the scientific apparatus—her persistence helped change the narrative in the scientific world. Dr Peter Dazsak (that name keeps coming up) tried to undermine her research, but she held her ground and he eventually had to concede.

Our task isn't to sit in judgment of Fauci or Dazsak or the Virology world. It's to get them to admit that the WIV was the most likely cause of the pandemic. This article suggests the following outcome:

Antonio Regalado, biomedicine editor of *MIT Technology Review*, put it more bluntly. If it turned out COVID-19 came from a lab, he tweeted, "it would shatter the scientific edifice top to bottom." That's a pretty good incentive to simply dismiss the whole hypothesis, but it quickly amounted to a global gaslighting of the media—and, by proxy, the public. An unhealthy absolutism set in: Either you insisted that any questions about lab involvement were absurd, or you were a tool of the Trump administration and its desperation to

blame China for the virus. I was used to social media pundits ignoring inconvenient or politically toxic facts, but I'd never expected to see that from some of our best scientists.

V/R
Dave

_____

Could COVID-19 Have Escaped from a Lab?

The world's preeminent scientists say a theory from the Broad Institute's Alina Chan is too wild to be believed. But when the theory is about the possibility of COVID being man-made, is this science or censorship?

by **ROWAN JACOBSEN**. *9/9/2020, 9:29 a.m.*

PRINT

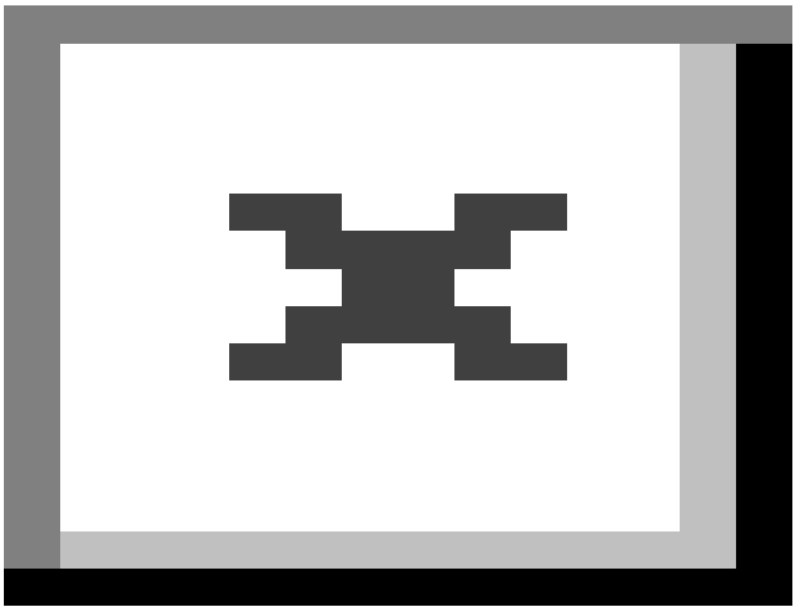


Illustration by Benjamen Purvis

In January, as she watched the news about a novel virus spreading out of control in China, Alina Chan braced for a shutdown. The molecular biologist at the Broad Institute of Harvard and MIT started stockpiling medicine and supplies. By the time March rolled around and a quarantine seemed imminent, she'd bought hundreds of dollars' worth of fillets from her favorite fishmonger in Cambridge and packed them into her freezer. Then she began to

ramp down her projects in the lab, isolating her experimental cells from their cultures and freezing them in small tubes.

As prepared as she was for the shutdown, though, she found herself unprepared for the frustration of being frozen out of work. She paced the walls of her tiny apartment feeling bored and useless. Chan has been a puzzle demon since childhood, which was precisely what she loved about her work—the chance to solve fiendishly difficult problems about how viruses operate and how, through gene therapy, they could be repurposed to help cure devastating genetic diseases. Staring out her window at the eerily quiet streets of her Inman Square neighborhood, she groaned at the thought that it could be months before she was at it again. Her mind wandered back to 2003, when she was a teenager growing up in Singapore and the first SARS virus, a close relative of this coronavirus, appeared in Asia. It hadn't been anything like this. That one had been relatively easy to corral. *How had this virus come out of nowhere and shut down the planet? Why was it so different?* she asked herself.

Then it hit her: The world's greatest puzzle was staring her in the face. Stuck at home, all she had to work with was her brain and her laptop. Maybe they were enough. Chan fired up the kettle for the first of what would become hundreds of cups of tea, stacked four boxes on her kitchen counter to raise her laptop to the proper height, pulled back her long dark hair, and began reading all of the scientific literature she could find on the coronavirus.

It wasn't long before she came across an article about the remarkable stability of the virus, whose genome had barely changed from the earliest human cases, despite trillions of replications. This perplexed Chan. Like many emerging infectious diseases, COVID-19 was thought to be zoonotic—it originated in animals, then somehow found its way into people. At the time, the Chinese government and most scientists insisted the jump had happened at Wuhan's seafood market, but that didn't make sense to Chan. If the virus had leapt from animals to humans in the market, it should have immediately started evolving to life inside its new human hosts. But it hadn't.

On a hunch, she decided to look at the literature on the 2003 SARS virus, which had jumped from civets to people. *Bingo*. A few papers mentioned its rapid evolution in its first months of existence. Chan felt the familiar surge of puzzle endorphins. The new virus really wasn't behaving like it should. Chan knew that delving further into this puzzle would require some deep genetic analysis, and she knew just the person for the task. She opened Google Chat and fired off a message to Shing Hei Zhan. He was an old friend from her days

at the University of British Columbia and, more important, he was a computational god.

“Do you want to partner on a very unusual paper?” she wrote.

Sure, he replied.

One thing Chan noticed about the original SARS was that the virus in the first human cases was subtly different—a few dozen letters of genetic code—from the one in the civets. That meant it had immediately morphed. She asked Zhan to pull up the genomes for the coronaviruses that had been found on surfaces in the Wuhan seafood market. Were they at all different from the earliest documented cases in humans?

Zhan ran the analysis. Nope, they were 100 percent the same. Definitely from humans, not animals. The seafood-market theory, which Chinese health officials and the World Health Organization espoused in the early days of the pandemic, was wrong. Chan’s puzzle detectors pulsed again. “Shing,” she messaged Zhan, “this paper is going to be insane.”

In the coming weeks, as the spring sun chased shadows across her kitchen floor, Chan stood at her counter and pounded out her paper, barely pausing to eat or sleep. It was clear that the first SARS evolved rapidly during its first three months of existence, constantly fine-tuning its ability to infect humans, and settling down only during the later stages of the epidemic. In contrast, the new virus looked a lot more like late-stage SARS. “It’s almost as if we’re missing the early phase,” Chan marveled to Zhan. Or, as she put it in their paper, as if “it was already well adapted for human transmission.”

That was a profoundly provocative line. Chan was implying that the virus was already familiar with human physiology when it had its coming-out party in Wuhan in late 2019. If so, there were three possible explanations.

Perhaps it was just staggeringly bad luck: The mutations had all occurred in an earlier host species, and just happened to be the perfect genetic arrangement for an invasion of humanity. But that made no sense. Those mutations would have been disadvantageous in the old host.

Maybe the virus had been circulating undetected in humans for months, working out the kinks, and nobody had noticed. Also unlikely. China’s health officials would not have missed it, and even if they had, they’d be able to go back now through stored samples to find the trail of earlier versions. And they weren’t coming up with anything.

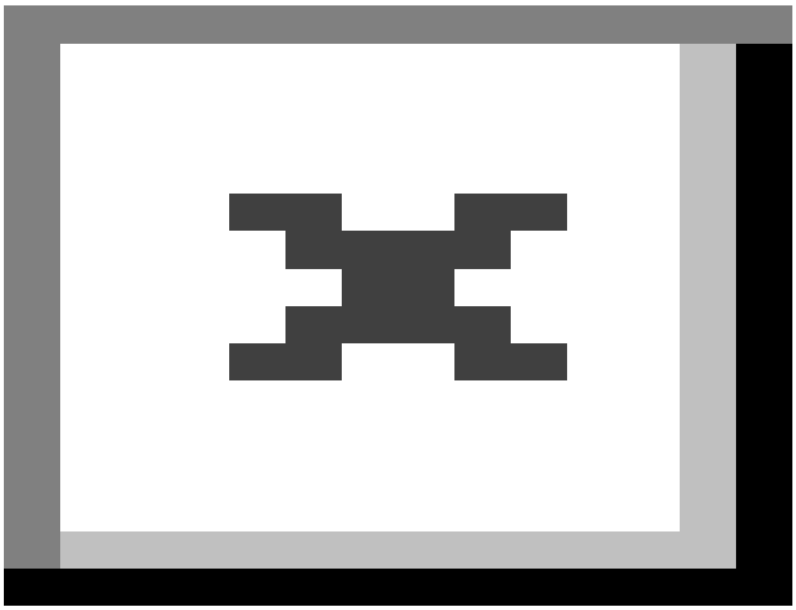
That left a third possibility: The missing phase had happened in a lab, where the virus had been trained on human cells. Chan knew this was the third rail of potential explanations. At the time, conspiracy theorists were spinning bioweapon fantasies, and Chan was loath to give them any ammunition. But she also didn't want to play politics by withholding her findings. Chan is in her early thirties, still at the start of her career, and an absolute idealist about the purity of the scientific process. Facts were facts.

Or at least they used to be. Since the start of the pandemic, the Trump administration has been criticized for playing fast and loose with facts—denying, exaggerating, or spinning them to suit the president's political needs. As a result, many scientists have learned to censor themselves for fear that their words will be misrepresented. Still, Chan thought, if she were to sit on scientific research just to avoid providing ammunition to conspiracy theorists or Trump, would she be any better than them?

Chan knew she had to move forward and make her findings public. In the final draft of her paper, she torpedoed the seafood-market theory, then laid out a case that the virus seemed curiously well adapted to humans. She mentioned all three possible explanations, carefully wording the third to emphasize that if the novel coronavirus did come from a lab, it would have been the result of an accident in the course of legitimate research.

On May 2, Chan uploaded the paper to a site where as-yet-unpublished biology papers known as "preprints" are shared for open peer review. She tweeted out the news and waited. On May 16, the *Daily Mail*, a British tabloid, picked up her research. The very next day, *Newsweek* ran a story with the headline "Scientists Shouldn't Rule Out Lab as Source of Coronavirus, New Study Says."

And that, Chan says, is when "shit exploded everywhere."



Alina Chan, a molecular biologist at the Broad Institute, says we can't rule out the possibility that the novel coronavirus originated in a lab—even though she knows it's a politically radioactive thing to say. / Photo by Mona Miri

Chan had come to my attention a week before the *Newsweek* story was published through her smart and straightforward tweets, which I found refreshing at a time when most scientists were avoiding any serious discussion about the possibility that COVID-19 had escaped from a biolab. I'd written a lot about genetic engineering and so-called gain-of-function research—the fascinating, if scary, line of science in which scientists alter viruses to make them more transmissible or lethal as a way of assessing how close those viruses are to causing pandemics. I also knew that deadly pathogens escape from biolabs with surprising frequency. Most of these accidents end up being

harmless, but many researchers have been infected, and people have died as a result.

For years, concerned scientists have warned that this type of pathogen research was going to trigger a pandemic. Foremost among them was Harvard epidemiologist Marc Lipsitch, who founded the Cambridge Working Group in 2014 to lobby against these experiments. In a series of policy papers, op-eds, and scientific forums, he pointed out that accidents involving deadly pathogens occurred more than twice a week in U.S. labs, and estimated that just 10 labs performing gain-of-function research over a 10-year period would run a nearly 20 percent risk of an accidental release. In 2018, he argued that such a release could "lead to global spread of a virulent virus, a biosafety incident on a scale never before seen."

Thanks in part to the Cambridge Working Group, the federal government briefly instituted a moratorium on such research. By 2017, however, the ban was lifted and U.S. labs were at it again. Today, in the United States and across the globe, there are dozens of labs conducting experiments on a daily basis with the deadliest known pathogens. One of them is the Wuhan Institute of Virology. For more than a decade, its scientists have been discovering coronaviruses in bats in southern China and bringing them back to their lab in Wuhan. There, they mix genes from different strains of these novel viruses to test their infectivity in human cells and lab animals.

When word spread in January that a novel coronavirus had caused an outbreak in Wuhan—which is a thousand miles from where the bats that carry this lineage of viruses are naturally found—many experts were quietly alarmed. There was no proof that the lab was the source of the virus, but the pieces fit.

Despite the evidence, the scientific community quickly dismissed the idea. Peter Daszak, president of EcoHealth Alliance, which has funded the work of the Wuhan Institute of Virology and other labs searching for new viruses, called the notion "preposterous," and many other experts echoed that sentiment.

That wasn't necessarily what every scientist thought in private, though. "They can't speak directly," one scientist told me confidentially, referring to the virology community's fear of having their comments sensationalized in today's politically charged environment. "Many virologists don't want to be hated by everyone in the field."

There are other potential reasons for the pushback. There's long been a sense that if the public and politicians really knew about the dangerous pathogen research being conducted in many laboratories, they'd be outraged. Denying the possibility of a catastrophic incident like this, then, could be seen as a form of career preservation. "For the substantial subset of virologists who perform gain-of-function research," Richard Ebright, a Rutgers microbiologist and another founding member of the Cambridge Working Group, told me, "avoiding restrictions on research funding, avoiding implementation of appropriate biosafety standards, and avoiding implementation of appropriate research oversight are powerful motivators." Antonio Regalado, biomedicine editor of *MIT Technology Review*, put it more bluntly. If it turned out COVID-19 came from a lab, he tweeted, "it would shatter the scientific edifice top to bottom."

That's a pretty good incentive to simply dismiss the whole hypothesis, but it quickly amounted to a global gaslighting of the media—and, by proxy, the public. An unhealthy absolutism set in: Either you insisted that any questions about lab involvement were absurd, or you were a tool of the Trump administration and its desperation to blame China for the virus. I was used to social media pundits ignoring inconvenient or politically toxic facts, but I'd never expected to see that from some of our best scientists.

Which is why Chan stood out on Twitter, daring to speak truth to power. "It is very difficult to do research when one hypothesis has been negatively cast as a conspiracy theory," she wrote. Then she offered some earnest advice to researchers, suggesting that most viral research should be done with neutered viruses that have had their replicating machinery removed in advance, so that even if they escaped confinement, they would be incapable of making copies of themselves. "When these precautions are not followed, risk of lab escape is exponentially higher," she explained, adding, "I hope the pandemic motivates local ethics and biosafety committees to think carefully about how they can reduce risk." She elaborated on this in another tweet several days later: "I'd also—personally—prefer if high biosafety level labs were not located in the most populous cities on earth."

How Safe Are Boston's Biolabs?

As one of the world centers of biotech, the Hub is peppered with academic and corporate labs doing research on pathogens. Foremost among them is Boston University's National Emerging Infectious Diseases Laboratories (NEIDL), the only lab in the city designated as BSL-4 (the highest level of biosafety and the same level as the Wuhan Institute of Virology). It is one of just a dozen or so in the United States equipped to work with live versions of the world's most dangerous viruses, including Ebola and Marburg. Researchers there began doing so in 2018 after a decade of controversy: Many locals objected to the risks of siting such a facility in the center of a major metropolitan area.

The good news? Before opening, NEIDL undertook one of the most thorough risk assessments in history, learning from the mistakes of other facilities. Even Lynn Klotz, a senior science fellow at the Washington, DC–based Center for Arms Control and Non-Proliferation, who advised local groups that opposed NEIDL, told the medical website Contagion that the lab likely has the best possible security protocols and measures in place.

But the reality, Klotz added, is that most lab accidents are caused by human error, and there is only so much that can be done through good design and protocols to proactively prevent such mistakes. (Or to guard against an intentional release by a disgruntled researcher, as allegedly happened in the anthrax attacks of 2001.) Rutgers molecular biologist Richard Ebright, a longtime critic of potentially dangerous pathogen research, says the risks introduced by NEIDL are not low enough and “definitely not” worth the negligible benefits.

Still, risk is relative. Klotz has estimated the chance of a pathogen escape from a BSL-4 lab at 0.3 percent per year, and NEIDL is probably significantly safer than the typical BSL-4 lab. And if catching a deadly pathogen is your fear, well, currently you run a good risk of finding one in your own neighborhood. Until that gets cleared up, the city’s biolabs are probably among the safer spaces in town.

Chan had started using her Twitter account this intensely only a few days earlier, as a form of outreach for her paper. The social platform has become the way many scientists find out about one another’s work, and studies have shown that attention on Twitter translates to increased citations for a paper in scientific literature. But it’s a famously raw forum. Many scientists are not prepared for the digital storms that roil the Twitterverse, and they don’t handle it well. Chan dreaded it at first, but quickly took to Twitter like a digital native. “Having Twitter elevates your work,” she says. “And I think it’s really fun to talk to nonscientists about that work.”

After reading her tweets, I reviewed her preprint, which I found mind-blowing, and wrote her to say so. She thanked me and joked that she worried it might be “career suicide.”

It wasn’t long before it began to look like she might be right.

Speaking her mind, it turns out—even in the face of censure—was nothing new for Chan, who is Canadian but was raised in Singapore, one of the more repressive regimes on earth. Her parents, both computer science professionals, encouraged free thinking and earnest inquiry in their daughter, but the local school system did not. Instead, it was a pressure-cooker of a system that rewarded students for falling in line, and moved quickly to silence rebels.

That was a bad fit for Chan. “You have to bow to teachers,” she says. “Sometimes teachers from other classes would show up and ask me to bow to them. And I would say, ‘No, you’re not my teacher.’ Back then they believed in corporal punishment. A teacher could just take a big stick and beat you in front of the class. I got whacked so many times.”

Still, Chan rebelled in small ways, skipping school and hanging out at the arcade. She also lost interest in her studies. “I just really didn’t like school.

And I didn't like all the extracurriculars they pack you with in Singapore," she says. That changed when a teacher recruited her for math Olympiads, in which teams of students compete to solve devilishly hard arithmetic puzzles. "I really loved it," she says. "You just sit in a room and think about problems."

Chan might well have pursued a career in math, but then she came up against teams from China in Olympiad competitions. "They would just wipe everyone else off the board," she says. "They were machines. They'd been trained in math since they could walk. They'd hit the buzzer before you could even comprehend the question. I thought, *I'm not going to survive in this field.*"

Chan decided to pursue biology instead, studying at the University of British Columbia. "I liked viruses from the time I was a teen," she says. "I remember the first time I learned about HIV. I thought it was a puzzle and a challenge." That instinct took her to Harvard Medical School as a postdoc, where the puzzle became how to build virus-like biomolecules to accomplish tasks inside cells, and then to Ben Deverman's lab at the Broad Institute. "When I see an interesting question, I want to spend 100 percent of my time working on it," she says. "I get really fixated on answering scientific questions."

Deverman, for his part, says he wasn't actively looking to expand his team when Chan came along, but when "opportunities to hire extraordinary people fall in my lap," he takes them. "Alina brings a ton of value to the lab," he explains, adding that she has an ability to pivot between different topics and cut to the chase. Nowhere was that more on display than with her coronavirus work, which Deverman was able to closely observe. In fact, Chan ran so many ideas past him that he eventually became a coauthor. "She is insightful, determined, and has the rare ability to explain complex scientific findings to other scientists and to the public," he says.

Those skills would prove highly useful when word got out about her coronavirus paper.

If Chan had spent a lifetime learning how to pursue scientific questions, she spent most of the shutdown learning what happens when the answers you come up with are politically radioactive. After the *Newsweek* story ran, conservative-leaning publications seized on her paper as conclusive evidence that the virus had come from a lab. "Everyone focused on the one line," Chan laments. "The tabloids just zoomed in on it." Meanwhile, conspiracists took it as hard evidence of their wild theories that there had been an intentional leak.

Chan spent several exhausting days putting out online fires with the many people who had misconstrued her findings. "I was so naive," she tells me with

a quick, self-deprecating laugh. “I just thought, *Shouldn’t the world be thinking about this fairly?* I really have to kick myself now.”

Even more troubling, though, were the reactions from other scientists. As soon as her paper got picked up by the media, luminaries in the field sought to censure her. Jonathan Eisen, a well-known professor at UC Davis, criticized the study in *Newsweek* and on his influential Twitter account, writing, “Personally, I do not find the analysis in this new paper remotely convincing.” In a long thread, he argued that comparing the new virus to SARS was not enough to show that it was preadapted to humans. He wanted to see comparisons to the initial leap of other viruses from animals to humans.

Moments later, Daszak piled on. The NIH had recently cut its grant to his organization, EcoHealth Alliance, after the Trump administration learned that some of it had gone to fund the Wuhan Institute of Virology’s work. Daszak was working hard to get it restored and trying to stamp out any suggestion of a lab connection. He didn’t hold back on Chan. “This is sloppy research,” he tweeted, calling it “a poorly designed phylogenetic study with too many inferences and not enough data, riding on a wave of conspiracy to drive a higher impact.” Peppering his tweets with exclamation points, he attacked the wording of the paper, arguing that one experiment it cited was impossible, and told Chan she didn’t understand her own data. Afterward, a Daszak supporter followed up his thread with a GIF of a mike drop.

It was an old and familiar dynamic: threatened silverback male attempts to bully a junior female member of the tribe. As a postdoc, Chan was in a vulnerable position. The world of science is still a bit medieval in its power structure, with a handful of institutions and individuals deciding who gets published, who gets positions, who gets grants. There’s little room for rebels.

What happened next was neither old nor familiar: Chan didn’t back down. “Sorry to disrupt mike drop,” she tweeted, providing a link to a paper in the prestigious journal *Nature* that “does that exact experiment you thought was impossible.” Politely but firmly, she justified each point Daszak had attacked, showing him his mistakes. In the end, Daszak was reduced to arguing that she had used the word “isolate” incorrectly. In a coup de grâce, Chan pointed out that actually the word had come from online data provided by GenBank, the NIH’s genetic sequence database. She offered to change it to whatever made sense. At that point, Daszak stopped replying. He insists, however, that Chan is overinterpreting her findings.

With Eisen, Chan readily agreed to test her hypothesis by finding other examples of viruses infecting new hosts. Within days, a perfect opportunity

came along when news broke that the coronavirus had jumped from humans to minks at European fur farms. Sure enough, the mink version began to rapidly mutate. “You actually see the rapid evolution happening,” Chan said. “Just in the first few weeks, the changes are quite drastic.”

Chan also pointed out to Eisen that the whole goal of a website such as bioRxiv (pronounced “bioarchive”)—where she posted the paper—is to elicit feedback that will make papers better before publication. Good point, he replied. Eventually he conceded that there was “a lot of interesting analysis in the paper” and agreed to work with Chan on the next draft.

The Twitter duels with her powerful colleagues didn’t rattle Chan. “I thought Jonathan was very reasonable,” she says. “I really appreciated his expertise, even if he disagreed with me. I like that kind of feedback. It helped to make our paper better.”

With Daszak, Chan is more circumspect. “Some people have trouble keeping their emotions in check,” she says. “Whenever I saw his comments, I’d just think, *Is there something I can learn here? Is there something he’s right about that I should be fixing?*” Ultimately, she decided, there was not.

By late May, both journalists and armchair detectives interested in the mystery of the coronavirus were discovering Chan as a kind of Holmes to our Watson. She crunched information at twice our speed, zeroing in on small details we’d overlooked, and became a go-to for anyone looking for spin-free explications of the latest science on COVID-19. It was thrilling to see her reasoning in real time, a reminder of why I’ve always loved science, with its pursuit of patterns that sometimes leads to exciting revelations. The website CNET featured her in a story about “a league of scientists-turned-detectives” who were using genetic sequencing technologies to uncover COVID-19’s origins. After it came out, Chan added “scientist-turned-detective” to her Twitter bio.

She’s lived up to her new nom de tweet. As the search for the source of the virus continued, several scientific teams published papers identifying a closely related coronavirus in pangolins—anteater-like animals that are heavily trafficked in Asia for their meat and scales. The number of different studies made it seem as though this virus was ubiquitous in pangolins. Many scientists eagerly embraced the notion that the animals might have been the intermediate hosts that had passed the novel coronavirus to humans. It fit their preexisting theories about wet markets, and it would have meant no lab had been involved.

As Chan read the pangolin papers, she grew suspicious. The first one was by a team that had analyzed a group of the animals intercepted by anti-smuggling authorities in southern China. They found the closely related virus in a few of them, and published the genomes for that virus. Some of the other papers, though, were strangely ambiguous about where their data was coming from, or how their genomes had been constructed. Had they really taken samples from actual pangolins?

Once again, Chan messaged Shing Hei Zhan. “Shing, something’s weird here,” she wrote. Zhan pulled up the raw data from the papers and compared the genomes they had published. Individual copies of a virus coming from different animals should have small differences, just as individuals of a species have genetic differences. Yet the genomes in all of the pangolin papers were perfect matches—the authors were all simply using the first group’s data set. Far from being ubiquitous, the virus had been found only in a few pangolins who were held together, and it was unclear where they had caught it. The animals might have even caught it from their own smuggler.

Remarkably, one group of authors in *Nature* even appeared to use the same genetic sequences from the other paper as if it were confirmation of their own discovery. “These sequences appear to be from the same virus (Pangolin-CoV) that we identified in the present study.”

Chan called them out on Twitter: “Of course it’s the same Pangolin-CoV, you used the same dataset!” For context, she later added, “Imagine if clinical trials were playing fast and loose with their patient data; renaming patients, throwing them into different datasets without clarification, possibly even describing the same patient multiple times across different studies unintentionally.”

She and Zhan posted a new preprint on bioRxiv dismantling the pangolin papers. Confirmation came in June when the results of a study of hundreds of pangolins in the wildlife trade were announced: Not a single pangolin had any sign of a coronavirus. Chan took a victory lap on Twitter: “Supports our hypothesis all this time.” The pangolin theory collapsed.

Chan then turned her Holmesian powers on bigger game: Daszak and the Wuhan Institute of Virology. Daszak had been pleading his case everywhere from *60 Minutes* to the *New York Times* and has been successful in rallying sympathy to his cause, even getting 77 Nobel laureates to sign a letter calling for the NIH to restore EcoHealth Alliance’s funding.

In several long and detailed “tweetorials,” Chan began to cast a cloud of suspicion on the WIV’s work. She pointed out that scientists there had discovered a virus that is more than 96 percent identical to the COVID-19 coronavirus in 2013 in a mineshaft soon after three miners working there had died from a COVID-like illness. The WIV didn’t share these findings until 2020, even though the goal of such work, Chan pointed out, was supposedly to identify viruses with the potential to cause human illnesses and warn the world about them.

Even though that virus had killed three miners, Daszak said it wasn’t considered a priority to study at the time. “We were looking for SARS-related virus, and this one was 20 percent different. We thought it was interesting, but not high risk. So we didn’t do anything about it and put it in the freezer,” he told a reporter from *Wired*. It was only in 2020, he maintained, that they started looking into it once they realized its similarity to COVID-19. But Chan pointed to an online database showing that the WIV had been genetically sequencing the mine virus in 2017 and 2018, analyzing it in a way they had done in the past with other viruses in preparation for running experiments with them. Diplomatic yet deadpan, she wrote, “I think Daszak was misinformed.”

For good measure, almost in passing, Chan pointed out a detail no one else had noticed: COVID-19 contains an uncommon genetic sequence that has been used by genetic engineers in the past to insert genes into coronaviruses without leaving a trace, and it falls at the exact point that would allow experimenters to swap out different genetic parts to change the infectivity. That same sequence can occur naturally in a coronavirus, so this was not irrefutable proof of an unnatural origin, Chan explained, “only an observation.” Still, it was enough for one Twitter user to muse, “If capital punishment were as painful as what Alina Chan is doing to Daszak/WIV regarding their story, it would be illegal.”

Daszak says that indeed he had been misinformed and was unaware that that virus found in the mine shaft had been sequenced before 2020. He also says that a great lab, with great scientists, is now being picked apart to search for suspicious behavior to support a preconceived theory. “If you believe, deep down, something fishy went on, then what you do is you go through all the evidence and you try to look for things that support that belief,” he says, adding, “That is not how you find the truth.”

Many of the points in Chan’s tweetorials had also been made by others, but she was the first reputable scientist to put it all together. That same week, London’s *Sunday Times* and the BBC ran stories following the same trail of

breadcrumbs that Chan had laid out to suggest that there had been a coverup at the WIV. The story soon circulated around the world. In the meantime, the WIV has steadfastly denied any viral leak. Lab director Yanyi Wang went on Chinese television and described such charges as “pure fabrication,” and went on to explain that the bat coronavirus from 2013 was so different than COVID that it could not have evolved into it this quickly and that the lab only sequenced it and didn’t obtain a live virus from it.

To this day, there is no definitive evidence as to whether the virus occurred naturally or had its origins in a lab, but the hypothesis that the Wuhan facility was the source is increasingly mainstream and the science behind it can no longer be ignored. And Chan is largely to thank for that.

In late spring, Chan walked through the tall glass doors of the Broad Institute for the first time in months. As she made her way across the gleaming marble foyer, her sneaker squeaks echoed in the silence. It was like the zombie apocalypse version of the Broad; all the bright lights but none of the people. It felt all the weirder that she was wearing her gym clothes to work.

A few days earlier, the Broad had begun letting researchers back into their labs to restart their projects. All computer work still needed to be done remotely, but bench scientists such as Chan could pop in just long enough to move along their cell cultures, provided they got tested for the virus every four days.

In her lab, Chan donned her white lab coat and took inventory, throwing out months of expired reagents and ordering new materials. Then she rescued a few samples from the freezer, took her seat at one of the tissue-culture hoods—stainless steel, air-controlled cabinets in which cell engineers do their work—and began reviving some of her old experiments.

She had mixed emotions about being back. It felt good to free her gene-therapy projects from their stasis, and she was even more excited about the new project she and Deverman were working on: an online tool that allows vaccine developers to track changes in the virus’s genome by time, location, and other characteristics. “It came out of my personal frustration at not being able to get answers fast,” she says.

On the other hand, she missed being all-consumed by her detective work. “I wanted to stop after the pangolin preprint,” she says, “but this mystery keeps drawing me back in.” So while she waits for her cell cultures to grow, she’s been sleuthing on the side—only this time she has more company: Increasingly, scientists have been quietly contacting her to share their own theories and papers about COVID-19’s origins, forming something of a

growing underground resistance. “There’s a lot of curiosity,” she says. “People are starting to think more deeply about it.” And they have to, she says, if we are going to prevent future outbreaks: “It’s really important to find out where this came from so it doesn’t happen again.”

That is what keeps Chan up at night—the possibility of new outbreaks in humans from the same source. If the virus emerged naturally from a bat cave, there could well be other strains in existence ready to spill over. If they are closely related, whatever vaccines we develop might work on them, too. But that might not be the case with manipulated viruses from a laboratory. “Someone could have been sampling viruses from different caves for a decade and just playing mix-and-match in the lab, and those viruses could be so different from one another that none of our vaccines will work on them,” she says. Either way, “We need to find where this came from, and close it down.”

Whatever important information she finds, we can be sure Chan will share it with the world. Far from being shaken by the controversy her paper stirred, she is more committed than ever to holding a line that could all too easily be overrun. “Scientists shouldn’t be censoring themselves,” she says. “We’re obliged to put all the data out there. We shouldn’t be deciding that it’s better if the public doesn’t know about this or that. If we start doing that, we lose credibility, and eventually we lose the public’s trust. And that’s not good for science.” In fact, it would cause an epidemic of doubt, and that wouldn’t be good for any of us.

DAVID R. STILWELL
Assistant Secretary, East Asia Pacific
(b)(6)

(b)(6)