after Emerging Viruses Group at the Wuhan Institute of Virology website (archive)

The coronavirus research that may have led to COVID-19

It is plausible that the SARS-CoV-2 genome (the cause of COVID-19) was engineered at the Wuhan Institute of Virology (WIV) through 'gain of function research'. Here I demonstrate why this is a plausible hypothesis through: a review of historical work at the WIV and analysis of the biochemical structure of SARS-CoV-2. I do not speculate on why or how a lab escape may have occurred.

Natural Origin

Based on the documented outbreaks of SARS-CoV and MERS-CoV and the role of bats and intermediate hosts in the SARS-CoV epidemic in particular (Hu et al. 2007, Li et al. 2008), a natural zoonotic origin for a novel coronavirus should in a first instance be given a high likelihood.

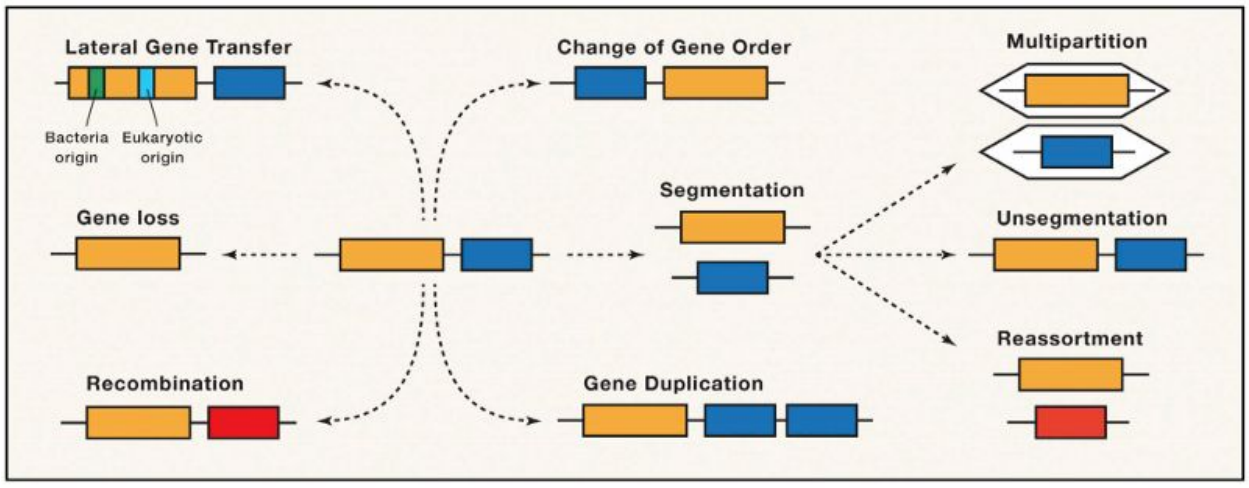


Fig 1. after Zhang et al. (2018) Mechanisms of Genome Evolution in RNA Viruses For each virus, the hypothetical genome is represented by a line, while the encoded genes (open reading frames, ORFs) are represented by boxes. Different colors are used to show genes of differing evolutionary origins.

Further Zhang et al. (2018) note “processes such as gene duplication and loss, genomic rearrangements, and lateral gene transfer occur far more frequently than ever imagined (with some viruses even lacking structural genes), and there are an increasing number of examples of cellular genes integrated into viral genomes (see Fig. 1)”. Given these factors, why even consider the possibility of a synthetic origin? It is this question that I attempt to answer.

First some background on the censorship on the topic of research around lab escape and lab synthesis investigation.

Censorship

In a curious correspondence on the 19th of February to the Lancet Jiang et al. 2020 (including Zhengli Shi) as ‘a group of virologists’ sought to change the name SARS-CoV-2. They write “the name SARS-CoV-2 might have adverse effects on the social stability and economic development in countries where the virus is causing an epidemic, perhaps even around the world. People develop panic at the thought of a re-occurrence of SARS. Travellers and investors might not want to visit a country with an ongoing epidemic or even sporadic cases of SARS. People may also believe that, like SARS-CoV, 2019-nCoV will not re-emerge once the current outbreak ends; therefore, they might not be prepared to prevent 2019-nCoV infection in the near future and could lose a sense of alert”. The argument was considered by the Lancet but rebuffed. While perhaps inconsequential, several other events indicate a persistent strong desire by the Chinese government to limit access to information and control the narrative around the origin of SARS-CoV-2 and the outbreak of COVID-19. However first I will discuss scientific research censorship.

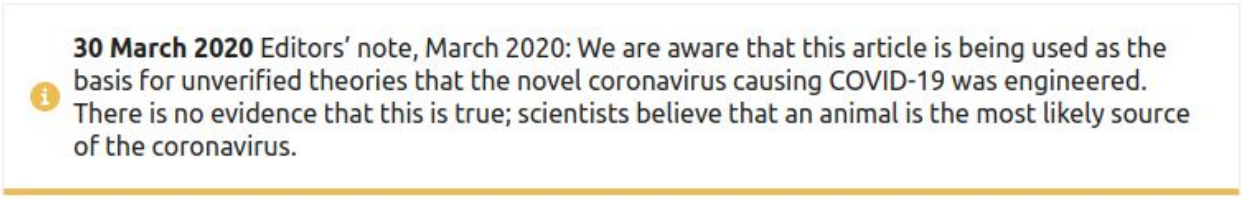
A single paper, a statement and an article in particular (all published within 11 weeks of Chinese officials first reporting a cluster of pneumonia cases ‘with no deaths’ to the WHO) set the scene which led to widespread censorship of research questioning a natural zoonotic origin for SARS-CoV-2.

Zhou et al. authored a Nature paper published on the 3rd of February 2020 describing the discovery of the RaTG13 SARSr-CoV isolated from a Rhinolophus affinis from Yunnan province that shared an overall genome sequence identity of 96.2% to SARS-CoV-2 — close enough that it “provides evidence that 2019-nCoV may have originated in bats”. This paper provided evidence that like two previous coronaviruses SARS-CoV and MERS-CoV which are known to have a natural zoonotic origin, SARS-CoV-2 as well was likely to have a natural origin. However the validity of the evidence in this paper is questioned later in this post.

A statement in the Lancet on February 19th 2020 by 27 prominent public health scientists stated “We stand together to strongly condemn conspiracy theories suggesting that COVID-19 does not have a natural origin”.

Andersen et al. authored a correspondence piece in Nature published on the 17th of March 2020 claiming that “Our analyses clearly show that SARS-CoV-2 is not a laboratory construct or a purposefully manipulated virus.”

An example of resultant censorship can be found in Nature, where several published papers such as Menachery et al. (2015) are tagged with the following note:



(Incidentally, the statement, in regards to SARS-CoV-2 engineering "there is no evidence that this is true" is disingenuous. As no animal species has as yet been identified as the source of zoonotic transfer of SARS-CoV-2 one could also technically state "there is no evidence that natural zoonotic transfer is true".)

In an interesting revelation, in November 2020, the U.S. Right to Know (USRTK) gained access to EcoHealth emails through a lawsuit against the National Institutes of Health after the agency failed to respond to its July 2020 FOIA request for records of gain-of-function experiments relating to the COVID-19 pandemic from the Wuhan Institute of Virology. In the released batch of emails, it appears that the statement mentioned above in the Lancet by 27 public scientists appears have a somewhat political raison d'être. While Daszak may have wanted to support his collaborators in China to do this via a statement refuting any questioning of a natural origin theory for the origin SARS-CoV-2 as a conspiracy theory only 4 weeks after the COVID-19 pandemic had been made known outside China is troubling.

Post the release of the emails, in an interview on skynews Prof. Nikolai Petrovsky discusses the unscientific methods and political nature of the emails.

Several authors have raised the issue of difficulty in publishing papers considering non-natural origin possibilities. For example Segreto and Deigin (2020) in a preprint noted that "Theories that consider a possible artificial origin for SARS-CoV-2 are censored as they seem to support conspiracy theories."

Journalists, presumably guided by the example from prominent signatories to the Lancet statement, the widely cited Andersen et al. article and Zhou et al. paper all linked above have also imposed a blanket ban on investigation into non-natural origin theories, often by using terminology such as 'debunking myths', 'bioweapon' and 'conspiracy theories' [eg a,b,c,d]

Relman D. (2020) in an opinion piece to PNAS writes that there are three potential origin scenarios for SARS-CoV-2:

1. "First, SARS-CoV-2 may have evolved in bats, which are known reservoirs of immense coronavirus diversity, and then spread directly, or indirectly via an intermediate host, to humans through natural mechanisms."
2. "Second, SARS-CoV-2 or a recent ancestor virus may have been collected by humans from a bat or other animal and then brought to a laboratory where it was stored

knowingly or unknowingly, propagated and perhaps manipulated genetically to understand its biological properties, and then released accidentally."

3. "The third scenario, seemingly much less likely, involves laboratory manipulation or release, with the clear intention of causing harm."

The result of censorship is that only the first option is being openly investigated, with scientists researching the second option being forced to publish almost exclusively non-peer reviewed preprints, which then can be easily discounted by other scientists and journalists precisely because the work has not been peer reviewed. It is the second scenario which we consider here.

Even one year after the initial COVID-19 outbreak in Wuhan we still do not know the true source of the SARS-CoV-2 virus. It is with this background that I provide a literature summary into gain of function experiments by the WIV in particular and argue that synthetic engineering of the virus at the WIV is a plausible scenario requiring a thorough, independent investigation with full access to the WIV and its records (Interview with Ralph Baric on collaborative research with the WIV can be found here).

Coronaviruses

SARS-CoV-2 belongs to the β coronavirus (CoV) genera of the *Coronaviridae* family of CoVs, along with two other human pathogenic viruses MERS-CoV and SARS-CoV. The process of SARS-CoV and SARS-CoV-2 infection of a host cell is shown below.

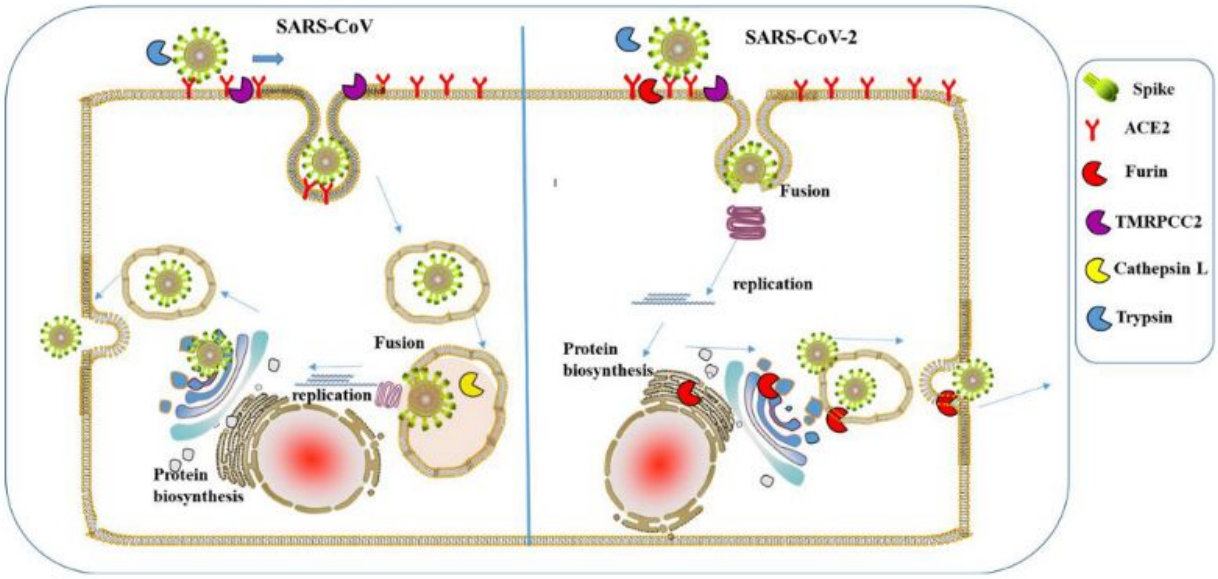


Fig 2. after Wu et al. (2020) A schematic diagram of the process of SARS-CoV and SARS-CoV-2 infecting host cells. ACE2 receptor in red.

The Spike protein (see protrusions around the viruses in figure above) is responsible for binding, tissue tropism and pathogenesis (Millet and Whitaker, 2014). It is particularly important

because without being able to bind to and facilitate entry to a cell the virus could not replicate. It also largely determines the host spectrum. I will revisit spike (S) proteins a bit later in this post.

SARS

The 2002 SARS outbreak was caused by a SARS-related *betacoronavirus* (SARSr-CoV). Palm Civets as well as several other market animals tested positive to SARS-CoV antibodies (Guan et al. 2003) and played a role in transferring the virus to humans. Wang et al. (2006) proposed horseshoe bats (genus Rhinolophus) to be most likely the reservoir host of SARSr-CoV.

After the SARS epidemic interest by researchers in China in sampling and genetically modifying SARSr-CoVs became a significant area of research for the CAS through the Center for Emerging Infectious Diseases (Emerging Viruses Group, headed by Dr. Shi Zhengli) at the WIV. Research on bat hosted SARSr-CoV's by the Emerging Viruses Group appears to have started in 2005 and publications significantly picked up in 2010.

Gain of Function Research by the WIV

International Collaboration

Fig. 3. [Archive of an “International Cooperation” page from the WIV Emerging Viruses Group headed by Zhengli Shi. Page is dated 20150920.]

(selected text translated below)

From 2004 to now, researcher Wang Linfa from Duke-Nus Graduate Medical School in Singapore has collaborated to study human infectious viruses carried by bats and their antiviral immune mechanisms.

From 2004 to now, he has cooperated with researcher Peter Daszak of the American Ecological Health Alliance to conduct research on viral pathogens and their pathogenicity carried by bats and rodents.

Collaboration by the WIV with Lin-Fa Wang at CSIRO, Australia (Wang et al. 2006, Zhou et al. 2009, Hou et al. 2010, Ge et al. 2013) then Duke-NUS Medical School, Singapore (Hu et al. 2017, Yang et al. 2017, Wang et al. 2018, Zhou et al. 2018) and Peter Daszak at the EcoHealth Alliance (ex Consortium for Conservation Medicine, see also Lancet support statement) (Wang et al. 2006, Ge et al. 2013, Hu et al. 2017, Yang et al. 2017, Wang et al. 2018, Zhou et al. 2018, Wang et al. 2019 (Zhu at EcoHealth Alliance)) appears to have been particularly close, while collaboration Ralph Baric's group at the University of North Carolina at Chapel Hill and the WIV (Xing-Yi Ge, Zhengli-Li Shi) is documented in a paper by Menachery et al. (2105).

Although publications by the WIV on gain of function research paused after 2016 perhaps due to widespread concern by scientists after the Menachery et al. (2105) paper (see Butler 2020), it appears research by the WIV may have been ongoing. For example Lin-Fa Wang, Daszak and Baric all attended the 8th International Symposium on Emerging Viral Diseases in 2018 in Wuhan, China (archive).

Further, a grant to the Ecohealth Alliance headed by Peter Daszak commenced on the 01/06/2014 and was scheduled to end on the 24/4/2020. The proposal involved working with the WIV to sample bats and undertake gain of function research on their viruses: "We will use S protein sequence data, infectious clone technology, in vitro and in vivo infection experiments and analysis of receptor binding to test the hypothesis that % divergence thresholds in S protein sequences predict spillover potential."

Fig.4 Peter Daszak of the Ecohealth Alliance on collaborative research with the WIV

Early Work

In 2008 Ren et al. (Zhengli Shi's team at the WIV) demonstrated that when the receptor-binding domain (RBD) of the S protein of a bat virus (SL-CoV) was replaced with a human virus the hybrid S protein was able to use the huACE2 for cell binding (i.e. infect human cells). While in Wang et al. (2008) Zhengli Shi's team collaborated with CSIRO in Australia to construct a chimeric SARS-CoV by replacing the S protein with a non infective gene.

In 2010 as documented in Hou et al. (2010) Zhengli Shi's team at the WIV analysed the ACE2 to SARS-CoV binding efficiency for multiple bat species and discovered two bat species could be infected with (human) SARS-CoV. Further they discovered that site discrete mutagenesis on ACE2 could dramatically increase receptor activity. WIV collaboration with the CSIRO documented in Yu et al. (2010) again showed that minor sequence changes in bat ACE2 "was all that was required to render it fully functional as a receptor for SARS-CoV".

Significant advances

Menachery et al. (2105) (including Zhengli Shi at the WIV) and Menachery et al. (2016) demonstrated that two closely related Bat SARS-like viruses to SARS-CoV (Bat SARSr-CoV SHC014 and Bat SARSr-CoV WIV1 respectively) could be modified to infect human ACE2 cells. Menachery et al. (2105) synthesized a chimeric CoV "SHC014-MA15" using the SHC014 spike and a mouse-adapted SARS-CoV back-bone. SHC014-MA15 was able to infect hACE2 (ie human ACE2 cells) and (surprisingly to the authors) showed a gain in pathogenesis relative to a wild SARS-CoV Urbani Spike protein with a mouse-adapted SARS-CoV back-bone chimeric virus.

In figure 4 below, d) shows the actual experiment of SHC014 spike–containing viruses (i.e. SARSr-CoV bat coronavirus S protein) recombined with a virulent CoV backbone causing direct human cell infection.

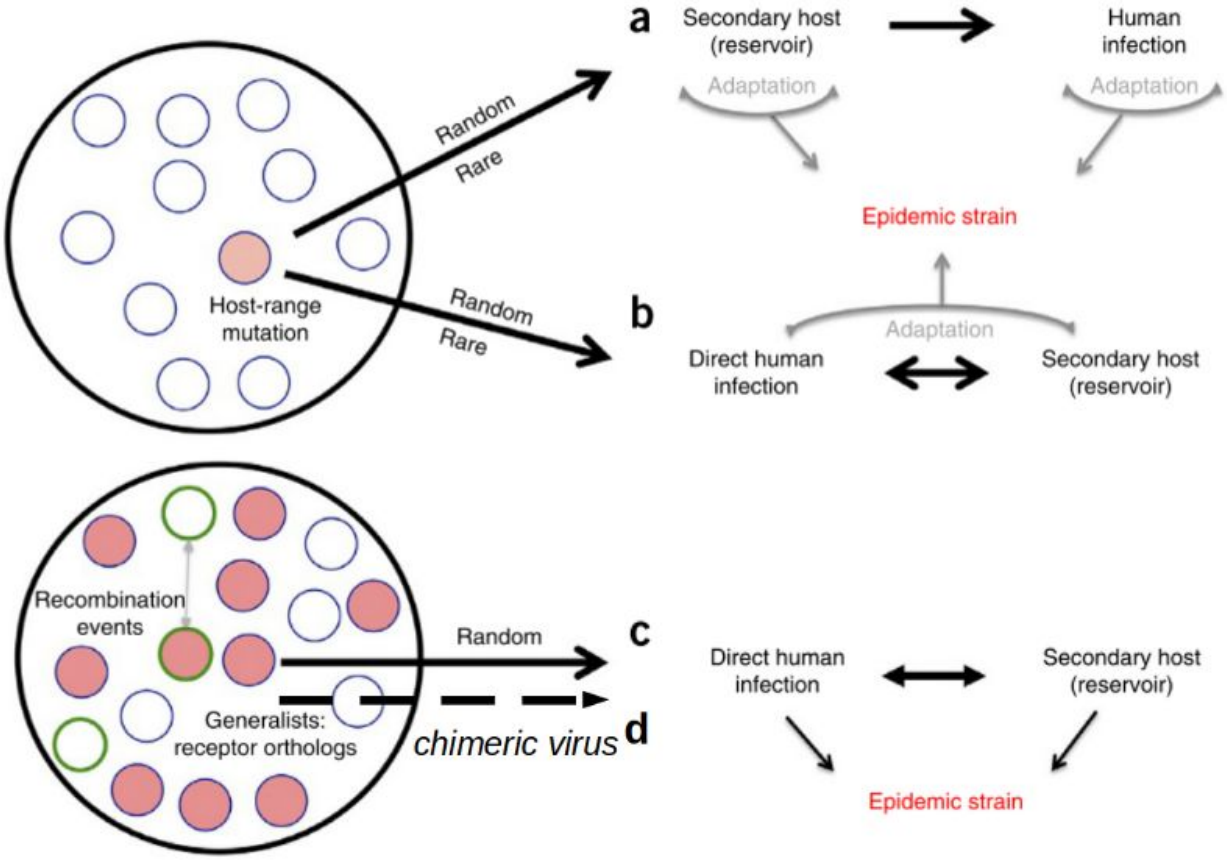


Fig. 5. Modified after Menachery et al. 2015. Emergence paradigms for coronaviruses. Coronavirus strains are maintained in quasi-species pools circulating in bat populations. (a,b) Traditional SARS-CoV emergence theories, d) chimeric lab created virus demonstrated to infect human cells and c) chimeric virus creation demonstrates the possibility that quasi-species pools maintain multiple viruses capable of infecting human cells without the need for mutations (red circles). Although adaptations in secondary or human hosts may be required for epidemic emergence.

In a follow-up to this paper, in a interview published on the 14th of September 2020 published on HuffPost, Ralph Baric, coauthor and head of the group which published the Menachery et al. (2105) work above stated “In the chimera we made in America in 2015 with the SARS virus, together with Professor Zheng-li Shi of the Wuhan Institute of Virology, we had left signature mutations, so it was clear that it was the result of genetic engineering. But, otherwise there is no way to distinguish a natural virus from one made in the laboratory”. He was asked if the possibility of SARS-CoV-2 being a laboratory chimera could be ruled out and responded “Not with the viruses that have been sequenced and reported to date.”

SARS-CoV-2 Furin Cleavage Site

SARS CoV-2 is unique in its betacoronavirus lineage B (Sarbecovirus) in being the only one to contain a furin cleavage site. Indeed of all coronavirus spike proteins in the NCBI database Wu et al. 2020 found that any coronavirus spike even remotely similar (>40% similar) to the SARS-CoV-2 spike protein did not have a furin cleavage site.

The furin cleavage site (FCS) for SARS-CoV-2 is a ‘multibasic’ site due to hosting multiple basic arginine (R) amino acids. The only other betacoronavirus containing a related (dibasic) cleavage site is the otherwise distant MERS- CoV (lineage C) (Lemmin et al. 2020).

The furin cleavage site is located at the S1/S2 boundary and occurs at the end of the “RRAR” (see ‘Site 1’ arrow in figure 5 below). By having an Arginine in the second position (or if it had a basic lysine) in the sequence the efficiency of the FCS is improved tenfold (Henrich et al. 2003), it is also rare, occurring in only 5 out of 132 known furin cleavage sites (Lemmin et al. 2020). Furin is an enzyme that can cleave precursor proteins to produce mature proteins, this facilitates virus cell entry and thus (in the case of SARS-CoV-2) a more efficient spread through the human population.

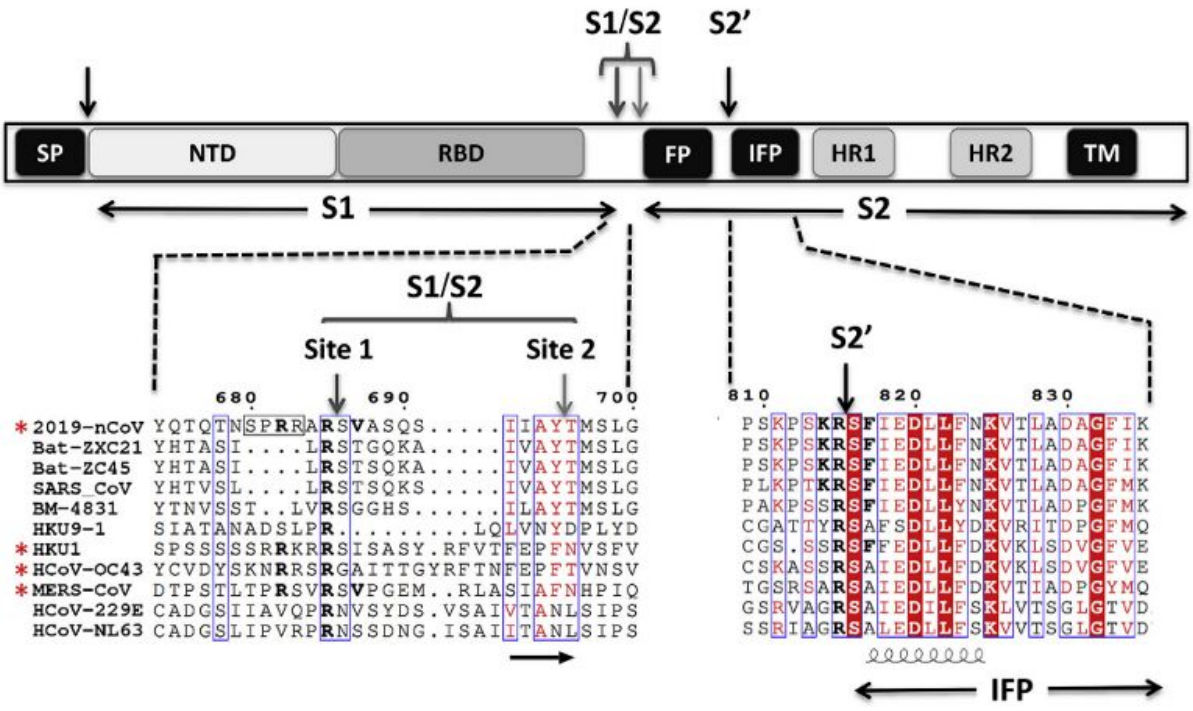


Fig. 6. after Coutard et al. 2020. Schematic representation of the human 2019-nCoV S-protein with a focus on the putative maturation sites. SARS-CoV related sequences in the top 5 rows. Signal peptide (SP), N-terminal domain (NTD), receptor-binding domain (RBD), fusion peptide (FP), internal fusion peptide (IFP), heptad repeat 1/2 (HR1/2), and the transmembrane domain (TM). The SP, S1↓S2 and S2' cleavage sites are indicated by arrows.Insertion of furin like cleavage site is surrounded by a black frame. Red asterisks indicate the presence of a canonical furin-like cleavage motif at the S1/S2 site.

In addition to furin, “the acquisition of the four amino acid insert (‘SPRR’) distinctively broadens the activating protease repertoire of the SARS-CoV-2 S1/S2 cleavage site to all major classes of proteolytic enzymes known to potentially activate coronavirus S proteins.”(Jaimes et al. 2020). The proline residue (the P in ‘SPRR’) in this position is “rare and appears in only 5 out of 132” known furin cleavage sites (Lemmin et al. 2020). It “imposes strong conformational restraints on the peptide chain.” (Lemmin et al. 2020) and often occurs in structural bends which can be seen in the figure below. It “separates the cleavage site from other structural elements, which might better expose it to the proteases.” (Lemmin et al. 2020).

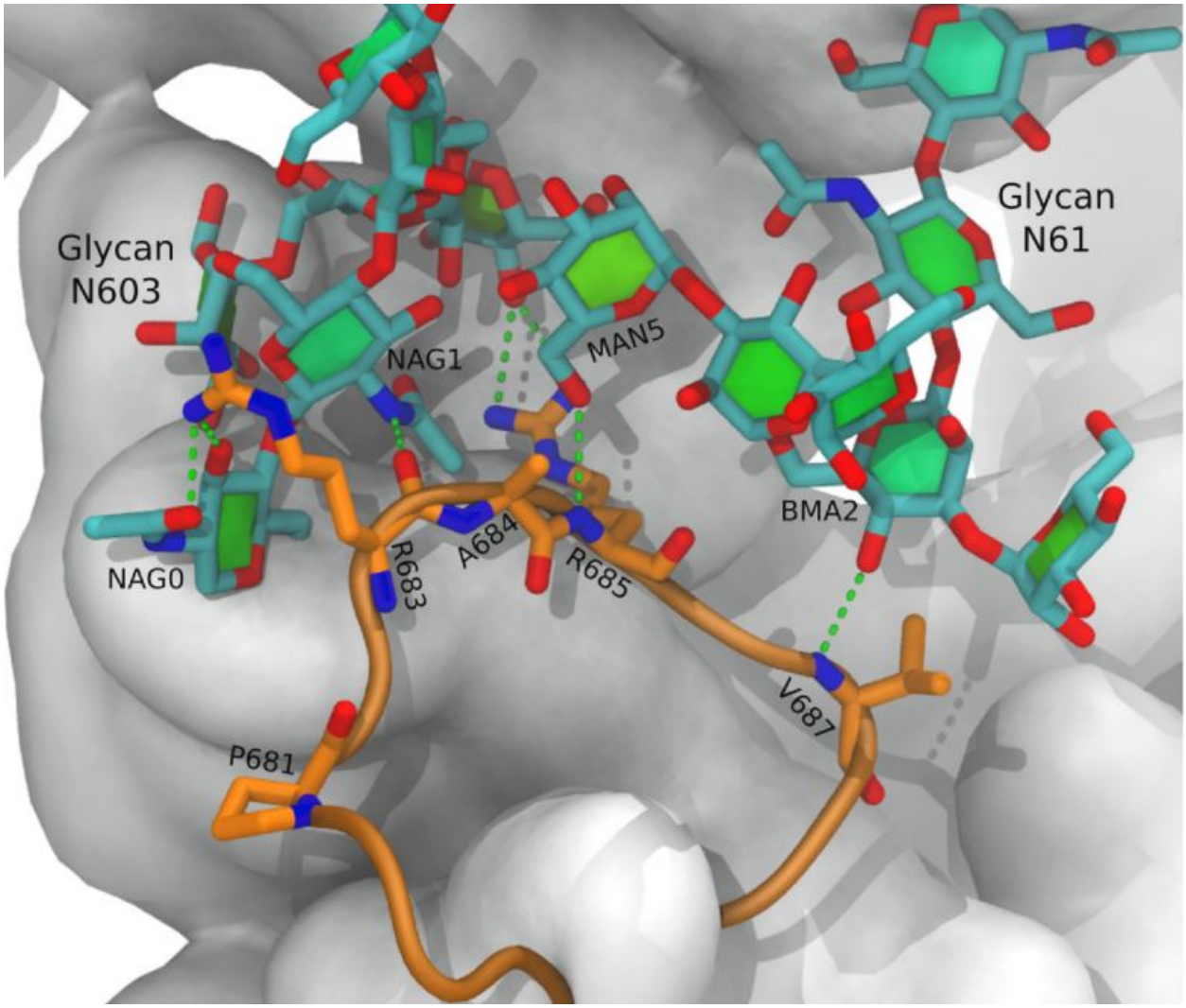


Fig. 7. after Lemmin et al. S protein (white), furin cleavage site (orange) and neighboring glycans. The proline ‘P’ in SPRR insert is denoted P681.

An additional cleavage site S2' is also shown in figure 5 above. SARS-CoV-2, its close relatives Bat-ZXC21 and Bat-ZC45 as well as SARS-CoV contain a dibasic cleavage site (KR↓SF) which is a form more pathogenic to humans than a monobasic site seen in other CoV’s (Coutard et al. 2020 and Lemmin et al. 2020).

Another unique feature of SARS-CoV-2 when compared to related CoV’s is the length of the loop containing the FCS, which being longer than other related CoV’s makes it more efficient (Lemmin et al. 2020).

SARS-CoV-2 is remarkably well adapted to humans

Based on genetic mutation rate studies Zhan et al. (2020) were “surprised to find that SARS-CoV-2 resembles SARS-CoV in the late phase of the 2003 epidemic after SARS-CoV had developed several advantageous adaptations for human transmission” and that by the time

it was “first detected in late 2019, it was already pre-adapted to human transmission to an extent similar to late epidemic SARS-CoV”.

Phylogenetic relationships for SARS-CoV documenting spike evolution towards efficient infection of human cells in Sheahan et al. (2008), and discussion of phylogenetic grouping and pathogenicity by Enjuanes et al. (2008) provide a useful background to findings by Zhan et al. (2020).

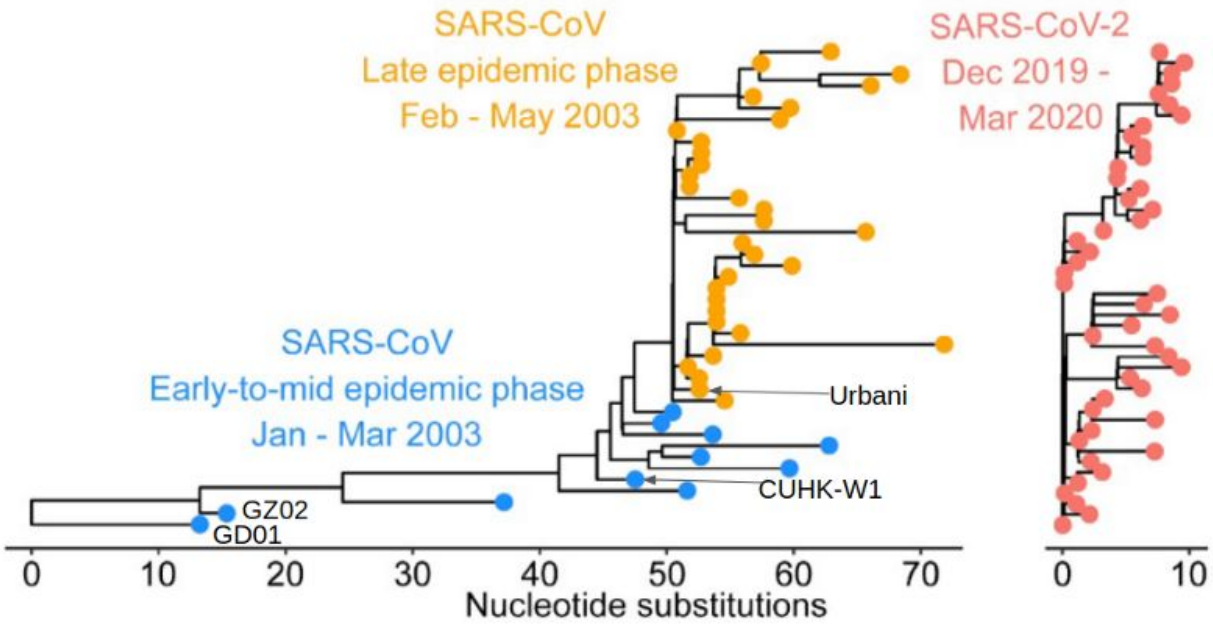


Fig. 8. after Zhan et al. Comparison of the genetic divergence of SARS-CoV and SARS-CoV-2. Max likelihood trees, using curated 11 early-to-mid epidemic SARS-CoV genomes, 32 late epidemic SARS-CoV genomes, and 46 SARS-CoV-2 genomes consisting of a December, 2019 Wuhan-Hu-1 isolate and 15 isolates from each month of January, February, and March, 2020

Similarly Jia et al. (2020) find that SARS-CoV-2 has a much lower mutation rate than SARS-CoV, and has a highly conserved Spike gene.

Damas et al. 2020 undertook a study of ACE2 (protein entry point) from 410 vertebrate species to “study the conservation of ACE2 and its potential to be used as a receptor by SARS-CoV-2”. The study showed that SARS-CoV-2 is well adapted to humans with only rare variants in binding sites, and that “Positive selection was found at multiple ACE2/SARS -CoV-2 S-binding residues in the bat-specific alignment”. As noted by @__ice9 on a twitter thread this is unexpected for a naturally evolved virus from bats, where we would expect some positive section from bat ACE2 towards hACE2 (i.e. adapting to a new (human) host).

Using computer modelling, Piplani et al. (2020) in an e-print find that the SARS-CoV-2 spike protein had the highest overall binding energy for human ACE2, greater than the other 13 tested species including bat, the postulated source of the virus. Piplani et al. (2020) note that “the

binding strength of SARS-CoV-2 for bat ACE2 is considerably lower than for human ACE2, suggesting that even if SARS-CoV-2 did originally arise from a bat precursor it must later have adapted its spike protein to optimise its binding to human ACE2. There is no current explanation for how, when or where this might have happened."

A massive heptapeptide (7 long amino acid sequence) sharing exists between SARS- CoV-2 spike glycoprotein and human (and mouse) proteins. "Such a peptide commonality is unexpected and highly improbable from a mathematical point of view" (Kanduc and Shoenfeld, 2020).

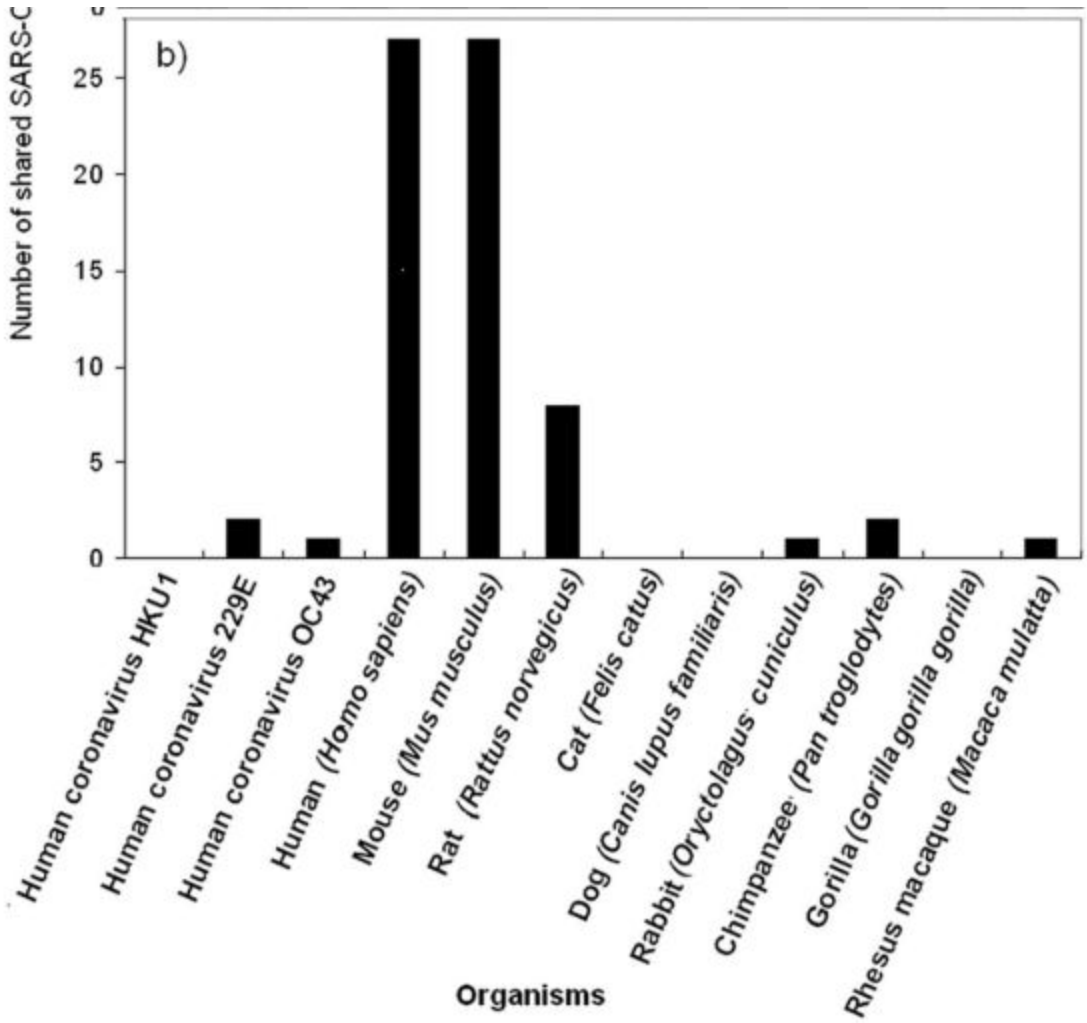


Fig. 9. after Kanduc and Shoenfeld (2020). Peptide sharing between SARS-CoV-2 spike glycoprotein and mammalian and coronavirus proteomes at the 7-mer level.

Sørensen et al. (2020) find that the SARS-CoV-2 spike protein shows 78.4% similarity with human epitopes. Further they find that the spike carries additional charge over SARS-CoV. The additional charge will strongly improve the interactions with other non ACE2 receptors (such as CLEC4M/DC-SIGNR) (Sørensen et al. (2020), see also Zhou et al. (2018)) helping its

pathogenicity (and incidentally potentially explaining the loss of smell experienced by COVID-19 patients). Such dual action capability is another unique feature of SARS-CoV-2.

In another article by Sørensen et al. (2020b)(which was rejected from publication) the authors expand on five features of the spike protein that are more consistent with "purposive manipulation" in a laboratory than a natural origin. Validly (which applies to this article as well) they write that "The longer the chain of causation of individual findings that is shown, especially converging from different disciplines, the greater the confidence in the whole."

Laporte et al. (2020) created S-bearing pseudoviruses in human cells (HEK293T) at 33°C and 37°C to measure S content. In parallel, they were used to transduce HEK293T target cells expressing ACE2/, DPP4/, ANPEP receptors and TMPRSS2. The S proteins of SARS-CoV and MERS-CoV favor 37°C (lower airway replication), while the spikes of SARS-CoV-2 (and HCoV-229E) prefer 33°C (upper airway and thus potentially more easily spread) for pseudoparticle production (but still compatible with 37°C). The question thus arises how did a virus so well adapted to replication at 33°C degrees arise from zoonotic transfer — pangolins which have a body temperature of 33°C and have been proposed but significant questions as to the validity of the pangolin datasets have been raised by independent researchers which I hope to follow-up with in a separate article.

That the SARS-CoV-2 genome was highly adapted to humans from the earliest strain, exhibits an improbably high peptide sharing of the Spike protein with human and mouse peptides as well as additional charge carried by the spike potentially proffering it with additional receptors not used by SARS-CoV are more consistent with a chimeric virus developed using humanised mice, than a natural zoonotic origin. (Not to mention no animal has been identified as the direct source at 12 months post initial outbreak in Wuhan).

RaTG13

The RaTG13 genome was introduced by Zhou et al. (2020) (including Zhengli Shi) in a Nature paper on the 3rd of February 2020, 4 weeks after the Chinese government announced a pneumonia outbreak to the WHO. Zhou et al. (2020) write that after sampling a patient and isolating 2019-nCoV (ie SARS-CoV-2) "We then found that a short region of RNA-dependent RNA polymerase (RdRp) from a bat coronavirus (BatCoV RaTG13) — which was previously detected in Rhinolophus affinis from Yunnan province — showed high sequence identity to 2019-nCoV. We carried out full-length sequencing on this RNA sample". Clearly they indicate that RaTG13 was fully sequenced after they identified high similarity of a section of it to SARS-CoV-2.

Multiple authors in pre-prints have raised concerns regarding the documented origin history and or data validity of RaTG13. Segreto and Deigin (2020) note the registering of BtCoV/4991 and RaTG13 as the same strain; Rahalkar and Bahulikar (2020) review the limited information around collection of the RaTG13 sample and raise questions requiring further information about RaTG13 data collection and history; Lin and Chen (2020) document concerns about the characterization of this strain; Mou et al. (2020) find that RaTG13 does not bind to horseshoe

bat ACE2 orthogs; Bengston (2020) (copy here) concludes RaTG13 is invalid for research unless the genome can be independently verified; Zhang (2020) identifies anomalies in RaTG13 sequencing; Rahalkar and Bahulikar (2020) note anomalous sequence reads in the fecal swab used for RaTG13 sequencing; Singla et al. (2020) reveal inconsistencies with regards to the RaTG13 genome; Yan et al. (2020) discusses the potential fabrication of RaTG13; Petric (2020) writes “Unavailability of the samples of RaTG13 virus, so that the experiment can be independently repeated in another laboratory, together with contradiction in Shi Zhengli´s statements about when the virus was sequenced casts a shadow of doubt whether are these data accurate”; Yan et al. (2020) note the abnormal features of RaTG13 and interpret that the sequence exhibits signs on fabrication; Signus (2020) reveals a potential fabrication pipeline used in RaTG13 sequencing.

On the 20th of October, 2020, Rahalkar and Bahulikar (2020) published a perspective article on a review of the Mojiang miners infection and raise several unanswered questions regarding RdRp 4991/RaTG13.

On the 17th of November 2020, Zhou et al. (2020) published an addendum to their Nature paper stating that unlike they had stated previously they actually sampled RaTG13 in 2018, and provide information regarding four miners, who were mining copper from a “mine cave” and contracted an unknown virus. Incidentally it is odd that individual miners would attempt to mine copper from an abandoned mine shaft, given the price of copper in 2012 was c. $1.50 USD per kg, and the amount of ore required for extraction to extract even one kg of refined copper. The revised story has raised more questions and inconsistencies (see thread by @AntGDuarte here, and criticism of addendum by Rhalkar, M. here).

The inconsistencies apparent in the documentation of the discovery of RaTG13 alone raises questions as to whether the sequence can be relied upon for research use.

Li et al. 2020 show very poor binding efficiency for Bat-CoV RaTG13 RBD to Rhinolophus sinicus ACE2, and Chu et al. (2020) found human ACE2 (also Shang et al. (2020)) followed by Chinese tree shrew and European domestic ferret (both lab animals in China) had the strongest ACE2 binding efficiency to RaTG13 with the two bat species studied Myotis lucifugus (Little brown bat) and Pipistrellus abramus (Japanese pipistrelle) having poor binding efficiency. Although bat species ACE2-RBD binding efficiency can vary, RaTG13 is still purportedly a bat virus (Zhou et al. 2020)

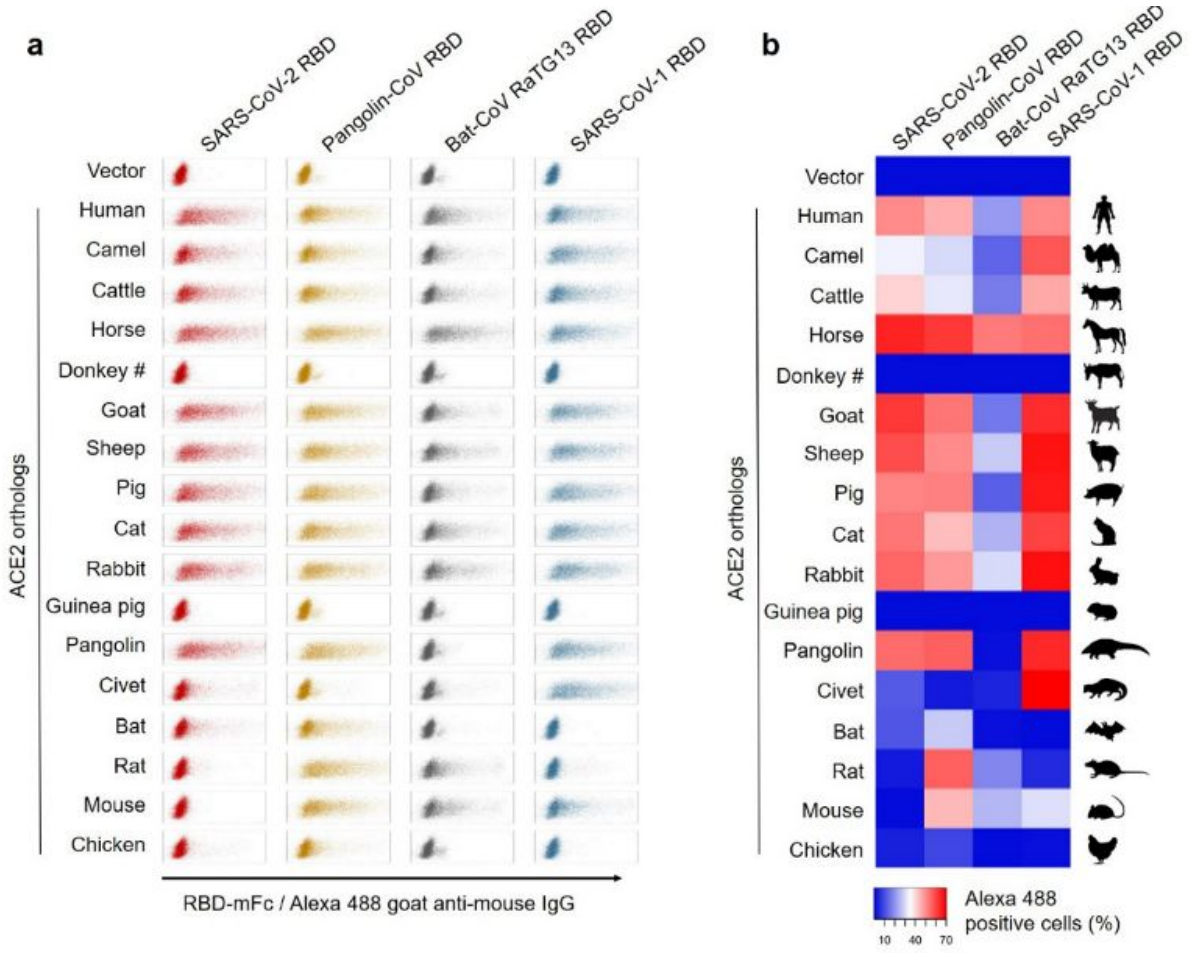


Fig. 10. after Li et al. ACE2 orthologs support binding to RBD proteins of SARS-CoV-2 and three related coronaviruses.

Wrobel et al. (2020) find that “thermal-stability data show that the uncleaved SARS-CoV-2 S trimer has a markedly higher stability than the bat virus protein does, whereas the cleaved SARS-CoV-2 has a similar stability to the (uncleaved) bat virus protein”. The core temperature of *Rhinolophus ferrumequinum* (Rhinolophidae) during flight has been measured at 41 degrees Celsius (Shea et al. (2014) and refs. therein). Thermal stability measurements are shown in Figure 9 indicating extensive protein unfolding by 41 degrees indicated by dF/dT measurements. This indicates that RaTG13 is unlikely to viably propagate in bats in the wild.

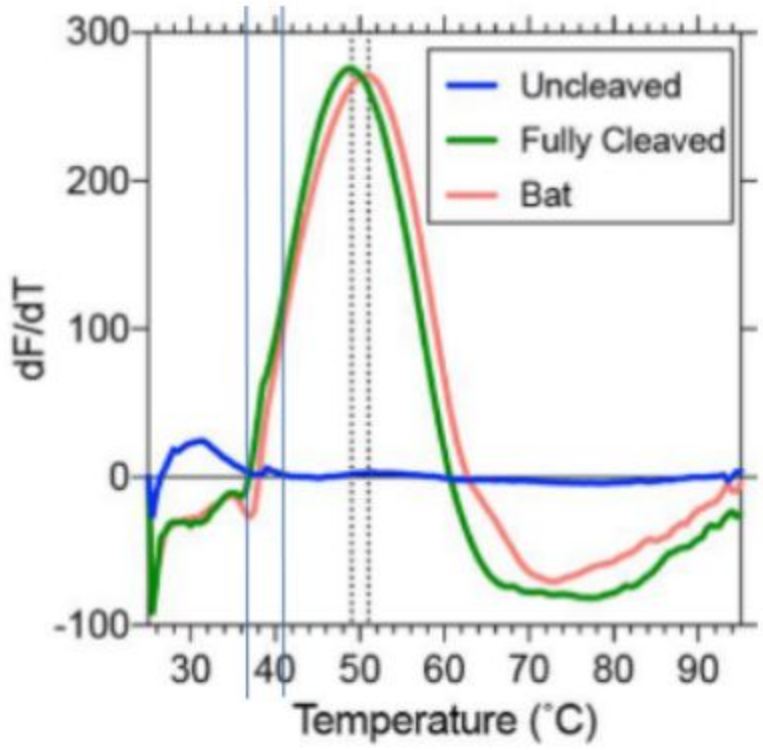


Fig. 11. After Extended Data Fig 1. Wrobel et al. 2020. Biochemical analyses of spike proteins. Differential Scanning Fluorimetry measurement of melting temperature for uncleavable RaTG13 and SARS-CoV-2 S, and fully-furin-cleaved SARS-CoV-2 S proteins. Solid lines at ~37 and 41 degrees, melting indicated by dashed lines.

Flavinkins (pers. comm. 2020) citing Wrobel et al. 2020 notes that “the amino acid position 1–29 of “R.Affinis ACE2” is in fact completely unsupported by all SRA data that is available (see Zhang (2020) and Signus (2020)), and therefore the critical “R24” was most likely fabricated using yeast display. And even that generated binding affinity that is actually lower than that of human ACE2, which is nearly 1000 times less than that of the binding affinity between hACE2 and SARS-CoV or SARS-CoV-2. This makes it impossible to infect the alleged bat host in-vivo.”

Flavinkins (pers. comm. 2020) further notes that “the SARS-CoV-2 Spike protein has a melting point of ~39C, judging by Wrobel et al. (2020) and Peng et al (2020) studies. This agrees well with the observation that despite high in-vitro binding affinity of SARS-CoV-2 to Ungulate and dog ACE2 (Shi et al. 2020) the animals themselves were found to be not susceptible to productive infection in-vivo, with the daily maximum body temperature of the dog at 39.0C being the clear cut-off point where no replication is possible.

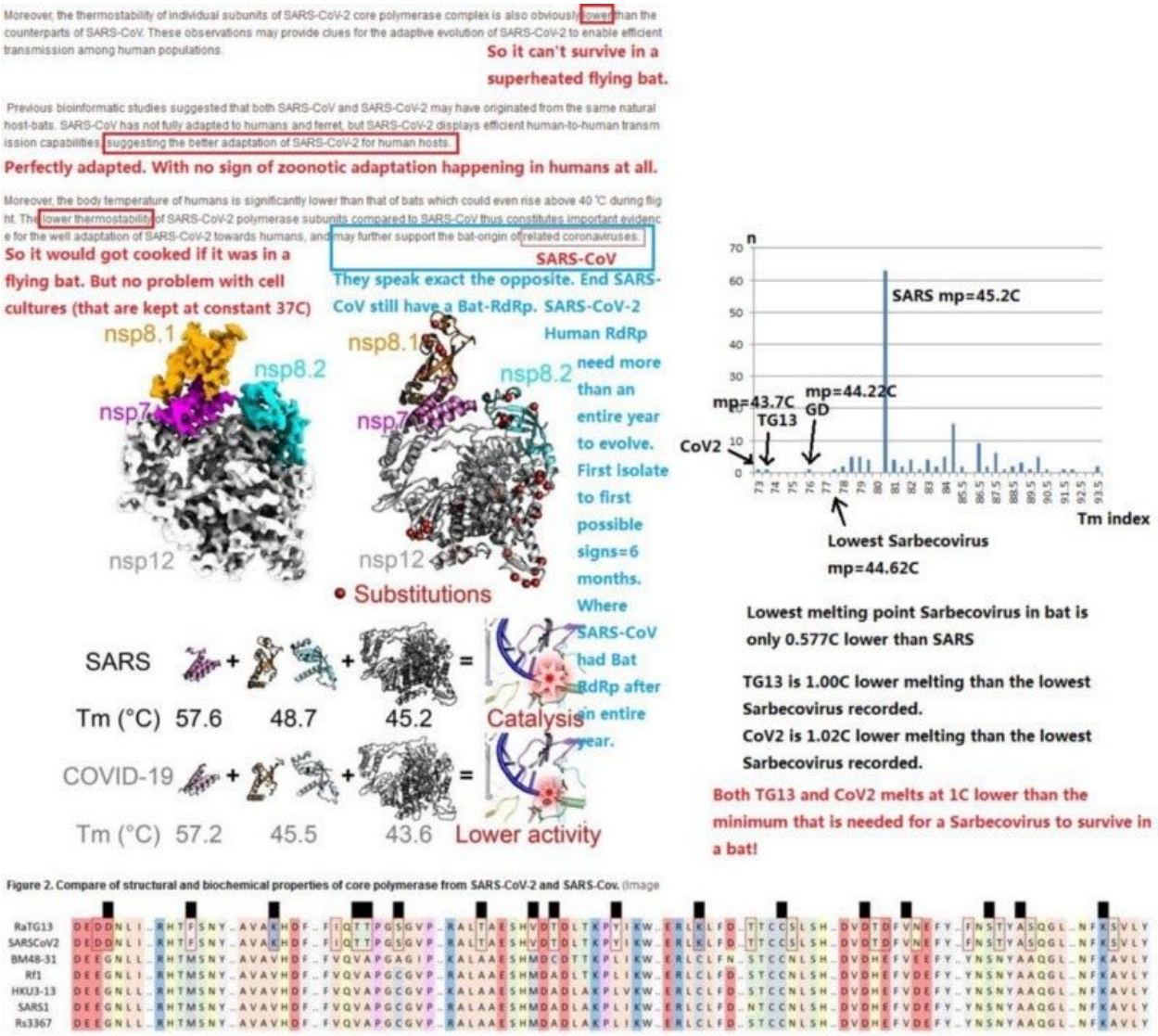


Figure 12. after Flavinkins (pers comm 2020), on thermostability of SARS-CoV-2 and RaTG13.

WIV secrecy

The Chinese government enacted strict censorship of any publications related to SARS-CoV-2 origins post the COVID-19 outbreak.

Early in 2020 historical work by the Emerging Viruses Group was removed from the WIV website. Unfortunately only limited internet archives exist and there is scant archived material from mid 2019 prior to the COVID-19 outbreak. You can access some of the archives here, and newsletters here.

Even well prior to the COVID-19 pandemic it appears that WIV research was censored as indicated by a 2018 interview with the the WIV director "Guangzhou Daily: *What powerful viruses are currently stored in the P4 laboratory?*

Song Donglin: *We currently have the ability to do research on viruses like Ebola, but what typeof virus has been preserved, how many, where it has been preserved, how the virus has been preserved, and what research has been done. The disclosure must be controlled.*"

More obfuscation

Scientists in China have been publishing prodigiously on research on SARS-CoV-2 and the COVID-19 pandemic. A recent work by Yu et al. (2020) on the evolution and transmission of SARS-CoV-2 using genomic data, document extensive tracking of 93 genomes of SARS-CoV-2 within China. They write that the theory that the origin of SARS-CoV-2 in Wuhan "has not been fully validated because the Hua Nan market has not been confirmed as the single source of SARS-CoV-2 transmission to humans and other possible original sources of SARS-CoV-2 have not been identified in Wuhan yet".

It is astounding that the WIV, the very laboratory where extensive gain of function research on bat coronaviruses has been documented, where Zhou et al. led by Zhengli Shi researched and published the discovery of RaTG13, the closest known relative of SARS-CoV-2, has not been openly investigated as a possible original source of SARS-CoV-2. Even though the first identified cases were in Wuhan, and that COVID-19 was more severe in Wuhan in its early stage than when it later spread nationwide in China (Jin et al. 2020), a push by the Chinese government to focus on an origin outside China is becoming evident.

Kristian Andersen and Edward Holmes (Shi et al. 2013, Huang et al. 2016, Lin et al. 2017, Zhang et al. 2018, Wu et al. 2020, Zhou et al. 2020) both with a history of collaborative research with PRC Universities, Hospitals and the Chinese Center for Disease Control and Prevention, have stated they are in support of, and not opposed to investigating labs (presumably WIV in particular).

I urge Andersen, Holmes and Daszak to use their connections to facilitate an unbiased thorough inspection of samples and reports at the WIV.

Until a clear natural zoonotic origin for the virus has been proved, or the escape of a natural virus from the WIV is established, we need to accept the unfortunate possibility that SARS-CoV-2 was a product of gain of function research at the WIV.

Peer reviewed articles questioning a natural origin

As discussed above, the topic of a synthetic origin for SARS-CoV-2 is for the most part censored by relevant journals and such research is mostly available only as pre-prints. However several researchers have published peer reviewed papers including: Sirotkin and Sirotkin (2020) who discuss the potential laboratory origin through an animal or cell culture host; Seyran et al. (2020) who discuss anomalies in the S protein for SARS-CoV-2; and Segreto and Deigin (2020)

who discuss the SARS-CoV-2 FCS in particular and raise the possibility of a chimeric virus creation at the Wuhan Institute of Virology.

Acknowledgements

This article benefited from review by Ersa Flavinkins (@flavinkins) and @angoffinet. Proofreading by @AntGDuarte, @BillyBostickson and @rebelfd.

I would also like to thank the following (in alphabetical order) for inspirational ideas which contributed to my research.

@Ayjchan, @MonaRahalkar, @nerdhaspower, @Rossana38510044, @TheSeeker268, @ydeigin and the #DRASTIC team.

pdf format available here.

Last updated 1/12/2020