

# Summary

This document is a summary of the decision making process I used during the debate between Saar/Rootclaim for the Lab Leak (LL) side, and Peter for the Zoonotic Origin (ZO) side.

I have broken it into three sections.

Section 1 is a short description of my process in this debate and my background.

Section 2 summarizes the evidence and arguments of each side through the lens of my first decision making method.

Section 3 summarizes the evidence and arguments of each side through the Bayesian analysis method preferred by the LL side. The vast majority of section 3 is found in the linked spreadsheet. Please see this for extended discussion.

# Section 1

## What did I decide?

**I find that SARS-CoV-2 is substantially more likely to be of zoonotic origin than a product of intentional engineering. I believe that Peter, for the zoonosis side, presented the superior case, and is the winner of the debate.**

I viewed this debate as a truth-seeking exercise with my vote reflecting what I found to be the most plausible origin of SARS-CoV-2, subject to the constraint that I would rely on (to the greatest degree possible) claims made by the debaters. Because of this constraint I was not able to evaluate several lines of inquiry which I thought could have been valuable for both sides. Ultimately, this means my decision in this debate is somewhere between "I think SARS-CoV-2 more likely than not had a natural origin" and "Based on the data presented by each of the sides in this debate, and ~70 hours of time spent evaluating and synthesizing those arguments/evidence, I think SARS-CoV-2 more likely than not had a natural origin".

In my Bayesian analysis, modeled after Michael Weissman's analysis I concluded

$$P_0(LL) / P_0(ZO) = 1.7E-3$$

The product of my updates, I concluded

$$P(LL) / P(ZO) = 2.125$$

The final, then is

$$P_0(LL) / P_0(ZO) * P(LL) / P(ZO) = 3.6E-3$$

I concluded that there is approximately a 1 in 300 chance that SARS-CoV-2 was the result of a lab leak. As you can see from my analysis, the major factors weighing against LL were the probability that WIV could carry out DEFUSE style work, and the epicenter being HSM.

Both my methods (described below) gave a similar result, with a slightly different emphasis on factors, and thus I concluded that zoonosis was the more likely cause of SARS-CoV-2.

## How did I decide?

The debates lasted ~18 hours, and I reviewed over 750 slides provided by the sides. I did not have time to evaluate the vast majority of the truth claims made, and therefore an important part of my process was deciding which claims I was going to subject to higher scrutiny. I tried to do the following four things in the debate:

- Give the benefit of the doubt to each side's claims when the other side did not contest them
- Read/analyze only sources that were listed by the debaters
- Rely on my own analysis only when I felt the debaters misrepresented their evidence or made a claim that wasn't supported by the cited evidence
- Not to be influenced by who I thought did the better "debating" either rhetorically or strategically

I ended up being more skeptical of the evidence that LL provided for two reasons. First, there were multiple instances of internal inconsistency. LL would assume truth claim A in one part of the debate, and then assume not A in another part of the debate, depending on the strategic utility of A/  $\neg$ A. Second, I felt that the LL side misrepresented the strength of evidence by leaving out qualifying information.

I don't believe this was willful dishonesty - I believe both parties have extremely high integrity. I discuss this more in Section 2 - see sub-section "Skepticism of LL evidence").

I made my decision with two "independent" methods. In method one, I synthesized the most plausible scenario I could for each side using a combination of the presented evidence, my background, and reading the sources linked in each slide deck. I compared these synthesized positions dispassionately, and got a sense for what I thought was correct in each key area. I then worked backwards to isolate what pieces I found most convincing and combined them into an overall decision. This method can be criticized for lack of "objectivity" and potential for bias, but I think it's an honest reflection of how most decisions are made. I am skeptical that the Bayesian decision making/evaluation methods are any more "objective" than this method. I think they maximize legibility, not objectivity, and tend to hide the intuitive/heuristic portion in the data inclusion step and values, where it's harder to see.

In the second method, I carried out a Bayesian analysis as advocated by the LL side. I modeled it somewhat after what the LL presented, and the analysis of Michael Weissman (v4 of his argument [here](#)). In brief, I made a calculation of the base rate/prior distribution on the probability that a newly emerged virus like SARS-CoV-2 would have zoonotic or lab origins. This was converted to a likelihood ratio (probability of lab leak/probability of zoonotic origin). For each additional piece of evidence (e.g. "the furin cleavage site (FCS) in SARS-CoV-2 is the only FCS within the subgenus *Sarbecovirus*") I evaluated the probability of that argument under a LL or ZO. The final probability was then the product of the base ratio times all the other ratios. The analysis by Weissman (and another by Stephen Quay) give a good introduction to this method. The spreadsheet I used to justify/carry out each step in the Bayesian analysis is linked [here](#) (and discussed in section 3).

I finished my decision via method 1 first. This obviously introduced bias, but I worked hard to guard against it. During my method 2 process, I did not make any intermediate calculations to make sure I was getting something that agreed with method 1. I only evaluated the decision I would have come to at the end of method 2 after I had completed all the relevant ratio determinations (I did make two updates to numbers in the base rate after I calculated it - both in LL's favor). Had the two decision methods provided different answers I would have refined both until I could reach a consensus.

As a final note, I am not skilled in the Bayesian method, and I am sure I made significant mistakes. More time and practice would improve and refine my estimates. At the fundamental rules of the universe level, Bayesian analysis must be the best way to evaluate evidence. However, I am unsure that it's a good strategy for a human given our cognitive limitations, and doubly unsure it's truly being used (in the dispassionate sense) where the outcome is social desirability/fame/Twitter likes.

## What is my background/bias?

I have a PhD in Microbiology and Immunology from Stanford. I have worked as a bioinformatician and microbiologist for the past 10 years, but I have focused almost exclusively on human commensal bacteria (the microbiome). My background as a scientist, both in academic settings, and running a startup, give

me direct experience with many of the administrative (e.g. grant writing/execution), physical (e.g. fermentation and sterilization), and theoretical (e.g. evolutionary dynamics of pathogen adaptations) portions of this debate.

Conversely, my knowledge of virology and molecular biology is limited; the fact I am extrapolating from work in bacteria and research methods used for bacteria reduces my confidence in my decision.

I also have a background in debate - I competed in and coached debate in high school, college, and have judged hundreds of debates. While I was not evaluating on strategic or rhetorical choices, the heuristics I learned in debate for evaluating the truth of a claim when I couldn't understand/know the source material definitely influenced my decision making.

## Section 2

In this section, I briefly summarize the most plausible scenario for a LL or ZO, and provide some extended discussion of my comparison of these scenarios.

I find that ZO is more likely. The key factors in my decision were as follows (in approximate order of importance)

Factor supported LL	Factor supported ZO
	Apparent emergence of SARS-CoV-2 at the Huanan Seafood Market vs. other places in Wuhan
	Whether Wuhan Institute of Virology had an appropriate backbone strain on which gain of function (GOF) research could have resulted in SARS-CoV-2
	Lack of evidence of circulating virus prior to HSM superspreading event, implying early HSM infections contained (or were close) to the index case
Lack of multiple spillovers or evidence of prior infections more proximal to the likely site of a zoonotic spillover	
Apparent emergence of SARS-CoV-2 in Wuhan vs. other places in China	
FCS appears to be an insert	
Codon usage in FCS (double CGG)	
	FCS insert out of frame
	FCS suboptimal (PRRAR)
Furin cleavage site (FCS) uniqueness in the <i>Sarbecovirus</i> subgenera	

### The best case for LL

Here I give the strongest version of the case I could make from the evidence for Lab Leak.

Multiple betacoronoviruses (e.g. SARS1 and MERS) with pandemic potential emerged in the last 20 years. Early research on the novel biology and potential for public health consequences of these viruses

unlocked academic and public health funding, leading to the establishment of research programs dedicated to understanding coronaviruses. Many of these research programs were started in China due to their unique access to potentially novel coronaviruses and a rapidly improving scientific apparatus. Coinciding with the founding of these research programs, were unprecedented advances in sequencing capabilities (454, Illumina, PacBio, etc.) and synthetic biology tools (e.g. Gibson assembly). Using these new tools, researchers in China and elsewhere started to conduct gain of function (GOF) experiments on both flu and coronaviruses to uncover mechanisms of pathogenesis. The Wuhan Institute of Virology (WIV) emerged as a leader in this work due to its personnel - many of which had experience with SARS1 - and its access to caves with huge reservoirs of coronaviruses. In the early 2010's, WIV (as well as other Chinese labs) begins to accumulate large banks of bat fecal, oral, tissue, and blood samples from both within China and surrounding countries.

The WIV begins to collaborate with Ralph Baric's lab at University of North Carolina (UNC), and publishes multiple papers with GOF research on MERS in the years 2015-2021. Collectively, these papers demonstrate technical expertise in all aspects of coronavirus engineering, including insertion of FCSs, codon optimization, replacement of receptor binding domains (RBDs), and S gene swaps. Note, a nice list of these was provided by Rootclaim in response to a question from Eric (shown below).

19. Saar: What specific research project known to have been conducted at the WIV is most akin to the work alleged to create sars-cov-2?

The closest description of the research that could have led to SARS2 is in the DEFUSE proposal, which itself was not funded and clearly any research WIV would have decided to undertake based on DEFUSE would not have been published if a lab leak occurred. But some previous research by WIV and/or collaborators that could have led to SARS2 is as follows:

- WIV creating an FCS in HKU4 together with Baric in 2015:  
<https://journals.asm.org/doi/10.1128/vi.01279-15>
- WIV creating 8 live WIV1 chimeras with various RBDs from other SARS-like viruses in 2017: <https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1006698>
- Shibo Jiang inserting an FCS with a leading CGG codon and then using RFLP to verify the insertion (this is how Faul can be used in the SARS2 PRRA insertion for RFLP to screen the FCS presence/absence):  
<https://dx.doi.org/10.1371%2Fjournal.pone.0080005>
- WIV using RFLP: <https://doi.org/10.1128/JVI.03079-15>
- WIV and EcoHealth constructing MERS chimeras: "We constructed the full-length infectious clone of MERS-CoV, and replaced the RBD of MERS-CoV with the RBDs of various strains of HKU4-related coronaviruses previously identified in bats from different provinces in southern China."  
<https://theintercept.com/2021/10/21/virus-mers-wuhan-experiments>
- "The full-length infectious cDNA clone of MERS-CoV has been successfully constructed. The full-length S gene of 12 different novel bat MERS-related coronaviruses have been amplified and cloned into the T-vectors." "Understanding the Risk of Bat Coronavirus Emergence" (Grant Number 5R01AI110964-05, report submitted by EcoHealth Alliance to the National Institute of Allergy and Infectious Diseases, August 3, 2021)  
[https://s3.documentcloud.org/documents/21089573/priority-grants-for-foia-request-55058-first-look-institute-2\\_redacted.pdf](https://s3.documentcloud.org/documents/21089573/priority-grants-for-foia-request-55058-first-look-institute-2_redacted.pdf)

- 
- WIV tracking bats with satellite trackers in 2019 in Mengla county, Yunnan (similar to what is described in DEFUSE – goes to show that WIV could decide to do research described in DEFUSE by themselves): <https://pubmed.ncbi.nlm.nih.gov/35585799/>
  - Could also have been unpublished ongoing research for the NIH grant 2R01AI110964-06, "Understanding the Risk of Bat Coronavirus Emergence"  
<https://reporter.nih.gov/project-details/9819304>
  - Could also have been unpublished ongoing research for the Chinese grant "Pathogenicity of Two New Bat SARS-Related Coronaviruses to Transgenic Mice Expressing Human ACE2 Receptor."  
[https://www.medsci.cn/sci/nsfc\\_show.do?id=d976823520ca](https://www.medsci.cn/sci/nsfc_show.do?id=d976823520ca)

As part of the larger coronavirus research program, somewhere in ~2015 the WIV initiated different GOF experiments with wild viruses contained within their collection. The DEFUSE proposal, authored by Ecohealth Alliance, Ralph Baric, and using WIV and the Shi lab as a key collaborator, documents some of the work that was occurring at this time. While written as if this work was to occur in the future (e.g. to be funded by the grant), the preliminary work (construction of genetic systems, usable backbones, sample collection, etc.) was already completed. In my PhD training, I was taught to write grants that were "30% completed and 70% to be done", so it seems very plausible that some parts of the DEFUSE work occurred prior to the submission (and rejection) of that grant. Some GOF experiments planned for UNC might even have been transferred to WIV during this period due to the nominal NIH GOF funding ban 2014-2017.

By late 2019, WIV has multiple different research programs ongoing (perhaps each grad student has ~2 research projects). In one of these projects, a grad student is tasked with constructing a coronavirus that will look something like SARS, but with greater potential infectivity. The grad student grabs unsequenced samples from the WIV freezers, sequences them, and finds one or two with a virus with a very interesting RBD/backbone. The grad student manages to retrieve live virus from one of the sequenced samples, and begins to add features (e.g. FCS).

At this point, all that's left is for the virus to escape the lab. In this debate, the LL side argued almost exclusively for a containment breach due to this grad student (or another worker) contracting the virus due to insufficient laboratory protocol, and then infecting someone outside the lab. In short, the LL side argued that the index case was most likely to have been a lab worker. LL suggested that the containment breach could have been due to aerosolization of virus outside a proper tissue culture hood, from improperly contained mouse experiments, etc. This is a plausible scenario, as there were at least 3 cases of SARS1 escape from labs due to improper biosafety procedures.

The lab worker that acquires the index infection does not develop significant symptoms, but spreads the virus outside the lab. Depending on the timing of the lab acquired infection (LAI), multiple different scenarios could have occurred. In this debate, LL argued that the initial LAI was in late September 2019, after which the virus spread slowly through the community with primarily abortive transmission. At some point, the virus infects someone at HSM, and the HSM then becomes the first superspreading event. LL argued that HSM had several properties that would make it the most likely place for the first superspreading event to be ascertained including: the limited ventilation (increased particulates/unit volume), cold temperatures and organic matter coating many surfaces (longer fomite survival times), and unique population (many of the same visitors in close proximity).

Some key features of this scenario are:

- No malicious intent required - a small scale research operation at a national lab (WIV) with normal research incentives
- Only a small number of people - possibly just a single grad student - would need to carry out the project
  - Ecohealth Alliance, Ralph Baric, etc. can all be telling the complete truth that they believe it's natural origin and that their proposed research didn't lead to SARS-CoV-2
- The grad student could have been searching for interesting viruses to work on, starting with sequencing existing WIV samples, construction of a genetic system, and then GOF insertions
  - The sequence information for the project (e.g. the metagenome of the interesting sample, the sequence of the recovered strain, etc.) could have been stored exclusively on the grad student's laptop

- No sequences would have shown up in the internal WIV databases or external (e.g. Genbank). They would not have been submitted until paper submission (sequence submission is always a pain in the ass)
- No coverup is required - the PI might not have even known about the specific research steps the grad student was taking
- WIV could plausibly have had all the necessary samples in house from which to derive the appropriate backbone

I believe the LL side presented a strong case (approximately outlined above). There was one significant area where I thought they presented a weaker case than they could have. LL did not strongly consider that the initial laboratory leak could have occurred via environmental release of SARS-CoV-2 as opposed to LAI. Many of the cases of past lab leaks are due to insufficient sterilization either due to a procedural failure or because of damage to sterilization equipment. I believe this is both a reasonable probability (given all the other parts of the scenario are true) and it resolves some of the questions that ended up hurting the LL side in this debate.

## The best case for ZO

Here I give the strongest version of the case I could make from the evidence for ZO.

The most plausible ZO scenario mimics the emergence of SARS1. Based on sequence divergence rates in betacoronaviruses, the ancestor to SARS-CoV-2 must have diverged from an RATG13-like ancestor 40-70 years ago (see Hu et al, 2017). Based on the bat population densities and the presence of the most similar viruses, this SARS-CoV-2 containing cave was likely located in Yunnan province, northern Laos, or eastern Myanmar.

While there may have been a specific single ancestor to SARS-CoV-2, it is also possible that the pangenome of a population of viruses (e.g. those found in a single cave) evolved towards SARS-CoV-2. This is thought to be the case with SARS1. SARS1 was never identified as a single virus in a single bat, but the viral pangenome of a single cave did contain all the parts necessary to constitute the SARS1 virus (Boni et al. 2017)

At some point, wildlife farmers begin to harvest animals (perhaps bamboo rats or racoon dogs) from, or in close proximity to, the SARS-CoV-2 containing cave in Yunnan. One of these animals is infected with a direct SARS-CoV-2 ancestor with the exact genomic architecture of SARS-CoV-2. It is also possible that the intermediate host was infected with a recombinant SARS-CoV-2 ancestor not containing key features like the FCS and other mutations (e.g. A372). In the scenario ZO defended, if the virus did not have these features during spillover, it quickly gained them.

Thus, shortly after spillover from the bats, the farmed animals are maintaining a reservoir of the fully constituted SARS-CoV-2. Humans interacting with these animals (handlers, workers, etc.) are exposed, and acquire infections as was documented for framed civets with SARS1. Because SARS-CoV-2 transmission patterns require fairly high population density, these early (likely rural) infections of handlers and farmers lead to abortive transmission chains. However, in early December 2019, a shipment of infected animals is transported (most likely alive) to the Huanan Seafood Market (HSM). While other markets are plausible within Wuhan, several vendors at the HSM had previously been cited/fined for trading in wildlife, and the HSM market was known to have Yunnan-farmed species including bamboo

rats. At the HSM, the animals infects the index case (or at least first ascertained case): an HSM seafood vendor named Wei Guixian, on December 10th, 2019.

In this debate, ZO argued that there were two spillover events of lineage A and lineage B SARS-CoV-2 within the HSM market in this timeframe. ZO also argued that HSM contains the index cases (not just first ascertained) and was the first superspreading event of SARS-CoV-2.

Hu et al, 2017. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. Plos Pathogens, Nov 30, 2017,  
<https://doi.org/10.1371/journal.ppat.1006698>

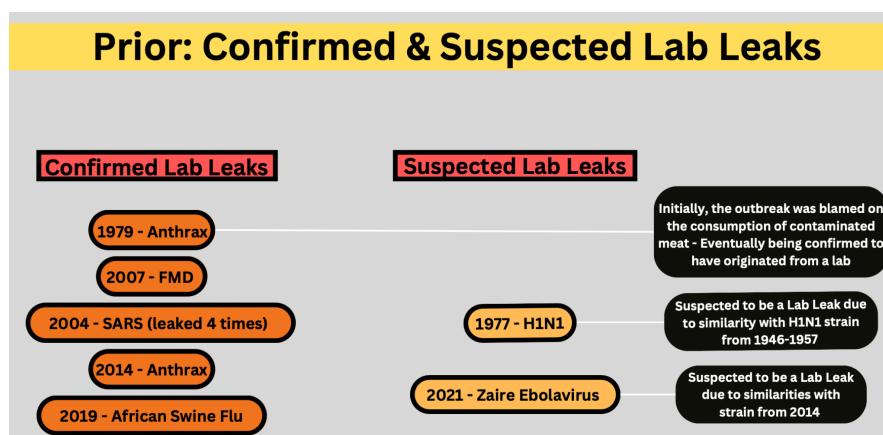
Boni, M.F., Lemey, P., Jiang, X. et al. Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. *Nat Microbiol* 5, 1408–1417 (2020).  
<https://doi.org/10.1038/s41564-020-0771-4>

## Skepticism of LL evidence

As mentioned in Section 1, throughout the course of the debate I became more skeptical of the evidence presented by the LL side. The two main sources of this skepticism were imprecision and motivated reasoning in their use of evidence, and seeming use of mutually inconsistent theories in different parts of the debate.

As an example of the imprecision in the LL side, I have highlighted the way they calculated their lab leak prior. One of the ways LL calculates the prior probability of a lab leak is by considering confirmed and suspected lab leaks. This seems like a good method. However, their choice of examples left me suspicious of motivated reasoning and that they had very limited command of the day-to-day operation of a biological laboratory.

In the first version of Slide 19 of debate 1, LL presents five "confirmed lab leaks" and two "suspected lab leaks".



I asked about these in pre-debate questions. Below are the questions and responses given by the LL side.

### **Question 1**

The slides generally convey the idea that an infected WIV worker is the likely source of the spillover event from lab to general population (e.g slide 56 “more likely to be a location frequented by a lab worker”, slide 66 “meeting between personnel”). In contrast, in slide 7, two of the confirmed lab leaks cited (1979 Anthrax, 2007 Foot and Mouth disease) were environmental releases rather than lab acquired infections of laboratory personnel. Do you believe there is any possibility that a loss of containment from WIV (e.g. incompletely inactivated sample released via scrubber or wastewater) was the spillover event? If so, how would this interact with the Mahjong halls hypothesis or the epicenter data shown on e.g. slide 56?

### **Answer**

We removed the 2007 Foot and Mouth outbreak, as it's not human.

There are indeed additional options for the leak to occur, such as illegal sale of a lab animal, or improper waste disposal. However, given SARS2 is a very infectious respiratory virus handled in BSL-2, and since the lab workers leave the lab every day (i.e. no need for a protocol breach), that seems like the most likely path.

### **Question 2**

Slides 19-22 give a list of confirmed and suspected lab leaks. These slides seem to serve as the basis for an important prior in the calculations. I have multiple questions on these slides:

1. What evidence is there that the 1977 Russian flu was a lab leak vs. the other competing hypotheses - intentional release and challenge vaccination trials?

### **Answer**

We did not look into this. Let me know if and why you find it important, and we'll try to answer.

2. Can you provide a citation for the 2021 - Zaire Ebola outbreak and the 2019 African Swine Flu being lab leaks?

### **Answer**

Zaire is suspected, not confirmed:

<https://twitter.com/ydeigin/status/1642537529554485252>

2019 Swine Fever has now been removed as it is not human.

3. The linked resources for the spillover events on slide 20 don't mention them in any format that I can find (e.g. ctrl+f for “spillover”, “Foshan”, etc.).

### **Answer**

The sources are for lab leaks. We forgot to add the source for SARS zoonotic spillovers. Apologies. Here it is: <https://encyclopedia.pub/entry/29846>

4. All the “major epidemics” listed on slide 22 are viral - why are bacterial epidemics of the 20th century not included?

### **Answer**

We just filtered the Wikipedia epidemic list for the last 100 years and 200k+ deaths.

I was skeptical of the LL side presentation of this data, especially given the level of responses to these questions. I have summarized my concerns with this evidence below:

1979 Anthrax - this lab leak occurred due to the release of 0.5-1 kg of biological material (Scientific American 2017) containing between 0.001 and 1 gram of anthrax spores (Meselson

1994). There are two critical things to realize about this "leak". First, this occurred at an offensive bioweapons facility containing fermentors and dryers capable of growing and drying perhaps thousands of liters per day. The biocontainment standards at this facility were likely much less stringent than a modern BSL2 or 3 lab. The leak from Sverdlovsk occurred from a facility operating at 3-4 orders of magnitude greater scale (in relevant parameters of amount of infectious agent produced), under worse conditions, as part of an offensive program. Whatever was occurring at Wuhan was nothing like this setup. Second, and related, the accidental release occurred due to a failure in the drying system, leading to uncontrolled and unsterilized venting (Alibek 1999; chapter 7, Scientific American 2017). As discussed, the LL side did not seriously consider/defend scenarios in which SARs-CoV-2 leaked via an environmental release. Even if this incident is an example of a lab leak, it is not one that can support the chain of events that LL side argued in this debate.

2007 FMD (foot and mouth disease) - this lab leak (of a cattle virus) was due to broken pipes, allowing the release of unsterilized infectious FMD virus which was then transported via vehicle tires to surrounding farms (Coghlan 2007). As with the previous example, I believe the scale of operations at the Pirbright site, where this leak occurred, was substantially larger than whatever was happening at Wuhan. Immediately prior to the leak, 6000 liters (possibly 12,000) of cell culture media was used for FMD vaccine production (National Emergency Epidemiology Group Food and Farming Group, 2007, page 5). And, as with the previous example, the leak occurred via a failed mechanical system (pipework leading to the sterilizer). I find the lack of attention by the LL side in this incident even more confusing because there are some parallels which would make for an intuitive pro-LL case. Specifically, the pipework at the Pirbright facility was known to be poor, and the facility had advertised seeking a plumber for repairs (Coghlan 2007). This is an interesting parallel to the claims made in Vanity Fair's 'Pishi' article which claimed, among other things, that there were well known problems with the sterilizer systems at the WIV and that the WIV had been purchasing, patenting, and installing new filters to rectify the problem immediately prior to the leak (Eban and Kao, 2022).

2019 African Swine Flu - I could find no evidence of this being a lab leak or even speculation to that effect.

1977 H1N1 - I think that 'suspected' is doing a lot of work here. I believe there is widespread agreement that the introduction of this H1N1 strain was not from a natural source (Rozo and Gronvall, 2015). However, the evidence seems substantially stronger that this was a consequence of a vaccine trial compared to a single point lab release. Yuri cites the Rozo and Gronvall 2015 paper (discussed in next point) but provides no refutation.

2021 Ebola - The evidence that LL side presented for this was a single twitter thread from Yuri Deigin ([here](#)). Yuri's argument stems from the lack of mutations in the 2021 strain compared to the 2014 strain. In short, the 2021 strain shows approximately ~10% of the expected mutations for a strain derived from the 2014 outbreak circulating in the wild. This is interesting evidence and worthy of consideration, but it is the only piece of data cited, and it appears that it is 'suspected' to be a lab leak by only LL proponents. This doesn't exclude it, but it makes its weight much harder to assess.

2014 Anthrax and 2004 SARS - both of these cases cited by the LL side seem like excellent examples of mechanisms of lab leak (incompletely inactivated samples and contaminated

cultures/researchers using inappropriate techniques, respectively). I will discuss the merits of these in the specific issues section below.

Why does this matter?

1. LL presented seven examples of lab leaks, of which five are one or more of: inapplicable to their theory of lab leak, technically so different as to be hard to compare, or wilfully imprecise. I have a longstanding interest in pandemics and bioweapons research (as well as spending the last 10 years as a microbiologist) which made it easier for me to easily identify how these cases don't support the theory that LL is advancing. In the cases where I have less background, I do not have the time to evaluate as much background evidence. If LL is being this imprecise (in a seemingly biased way) I should be skeptical of all their other evidentiary claims.
2. The LL side seems not to read this literature carefully. If you read the 2014 CDC report on the Anthrax leak, it provides at least six other cases which would be much more applicable to the LL theory than the ones they present (CDC, 2014). It could be that LL is assuming these and presenting other evidence, but I suspect this is not the case. I think there is a level of epistemic closure here. Collectively, this seems to somewhat implicate the RootClaim method that "avoids fundamental human reasoning errors". It appears to me that they concluded that a lab leak was most likely, and then post hoc prepared this presentation of the evidence. I would have expected a more in depth, and more unbiased discussion of a critical prior parameter in their calculations. After multiple revisions to this slide, and the pandemics they are claiming as lab leak evidence, they did not modify their overall probability assessments.

Rozo and Gronvall, 2015. Michelle Rozo and Gigi Kwik Gronvall, The Reemergent 1977 H1N1 Strain and the Gain-of-Function Debate, mBio. 2015 Jul-Aug; 6(4): e01013-15,  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4542197/>.

Scientific American 2017. Time to Worry about Anthrax Again, Scientific American Magazine Vol. 316 No. 4 (April 2017), p. 70. Accessed via  
doi:10.1038/scientificamerican0417-70

Ken Alibek 1999. Biohazard: The Chilling True Story of the Largest Biological Weapons Program in the World, - Told from the Inside by the Man Who Ran It. Random House, ISBN: 0-375-50231-9.

Meselson 1994. Meselson, M., Guillemin, J., Hugh-Jones, M., Langmuir, A., Popova, I., Shelokov, A., & Yampolskaya, O. (1994). The Sverdlovsk anthrax outbreak of 1979. Science, 266(5188), 1202–1208.  
doi:10.1126/science.7973702

Coglan 2007. Andy Coghlan, Faulty pipe blamed for UK foot and mouth outbreak, New Scientist, 2007. Accessed via  
<https://www.newscientist.com/article/dn12615-faulty-pipe-blamed-for-uk-foot-and-mouth-outbreak/>

Eban and Kao, 2022. Katherine Eban and Jeff Kao, COVID-19 Origins: Investigating a “Complex and Grave Situation” Inside a Wuhan Lab. Vanity Fair, October 28, 2022, accessed via  
<https://www.vanityfair.com/news/2022/10/covid-origins-investigation-wuhan-lab>.  
October 28, 2022

For an example of the “mutually inconsistent arguments” I provide the example of the negative binomial distributed spread of SARS-CoV-2. The ZO side presented evidence/a claim that SARS-CoV-2 shows a negative binomial distributed transmission distribution. Most people spread it to no one, and some superspreaders spread it to 20+. The LL side contested this, saying that we should be very skeptical of complex models. Below is a screenshot of the LL side urging caution (their response to my question is in green) at accepting a modeling result based on the negative binomial.

3. Covid spreading events are modeled by an overdispersed negative binomial.  
 Most infected people spread to no one, and about an equal number spread to 3 people as to 20+ people (Debate 3, slide 21). In a sentence, "Covid either grows exponentially or it goes extinct".  
 Copying my response to the claim of overdispersed infections:  
 This is a complex modeling claim newly introduced. As we learned the hard way during the debate, it needs to be reviewed for mistakes and wrong assumptions (e.g. could this be relevant only during social distancing, where most people are careful and a few are not?), before concluding anything from it.  
 Even if true, it just means that case growth will stutter (1 infect 2 infect 7, dies down back to 3) rather than grow smoothly.

The idea that I should be skeptical of the models is a point well taken, and I was sympathetic. However, at a different point in the debate (slide 48 and 49, debate 3), the LL side presents the following study (slide reproduced below) to bolster their Mahjong halls hypothesis. This study relies on...a negative binomial distribution of SARS-CoV-2 infections, and in fact provides very strong evidence for this.

## Specific conditions in HSM : Mahjong halls

In subgroup analysis, we found that cases in mahjong houses had substantial potential of superspreading, with 5% of cases seeding 80% of the total transmission (Table 2 and Fig. S5 in Appendix). Moreover, infectors aged more than 65 had relatively higher transmissibility, with a higher estimated R value compared to other age groups; infectors aged between 0 and 17 had sizable superspreading potential, with 8% of cases generating 80% of all transmission (Table 2 and Fig. S5 in Appendix).

Table 2. Estimated reproductive number (R) and dispersion parameter (k) of the negative binomial distributions, and inferred superspreading potential of Delta variants stratified by different contact settings, and age groups of seed cases.

	R (95% CI)	k (95% CI)	* Prop80% (95% CI)	** Prop0% (95% CI)
<b>Contact Settings</b>				
<b>Household (n = 108)</b>	0.58 (0.47–0.70)	13.42 (1.88-Inf)	32% (25%–36%)	57% (50%–66%)
<b>Community (n = 266)</b>	0.71 (0.59–0.84)	0.53 (0.37–0.76)	22% (18%–26%)	64% (57%–70%)
<b>Mahjong house (n = 91)</b>	0.82 (0.29–1.82)	0.05 (0.03–0.10)	5% (3%–9%)	86% (74%–94%)
<b>Market (n = 38)</b>	0.59 (0.39–0.81)	1.92 (0.58-Inf)	29% (18%–39%)	60% (44%–74%)

While this particular example didn't figure centrally in the debate, I found this same pattern in other uses of evidence by the LL side. In short, LL was either not precise enough in their use of evidence, or

provided internally inconsistent evidence at various points, raising my general skepticism level towards their claims.

## Evaluation of specific arguments

In this section, I add some details for arguments that I had to evaluate, learn about more in depth. Many of the specific conclusions I came to are detailed more extensively in the Bayesian analysis spreadsheet linked in Section 3.

### Apparent emergence of SARS-CoV-2 at the Huanan Seafood Market vs. other places in Wuhan

See the extended discussion in the cells for the update factor labeled “P (HSM is first superspread event in Wuhan| LL)”.

In summary, I think Wuhan was a reasonably unlikely place for a single SARS-CoV-2 zoonotic emergence. But, conditioned on occurring in Wuhan, I think the HSM is by far the most likely place for an outbreak. I was not convinced by the LL arguments that HSM was unique in its physical properties (e.g. Mahjong halls, ventilation, fomite survival, etc.). I think under the LL hypothesis, HSM was not a likely place for emergence.

### Whether Wuhan Institute of Virology had an appropriate backbone strain on which gain of function (GOF) research could have resulted in SARS-CoV-2

See the extended discussion in the cells for the prior factor labeled “P(WIV had an undisclosed backbone close enough to make SARS-CoV-2)”.

Both sides of the debate agree that the viruses with the highest similarity to SARs-CoV-2 are BANAL-52, BANAL-236 (Yuri: ~3% NT, 1% AA difference, BANAL-52 has 77/78 and 175/180 AA identity in RBM and RBD, respectively), RmYN02, and RaTG13.

Both sides of this debate agree that none of the viruses WIV was known to have could have been the backbone for SARS-CoV-2. LL believes that WIV had sequences that were not shared publicly, and that one or more of these was used in the GOF experiments that produced SARs-CoV-2.

LL couldn't point to any data that showed sequences of SARS-like viruses in WIV's possession. As detailed in the Bayesian spreadsheet, I believe WIV had many unsequenced samples, but I don't think they had working backbones. The embargoed Genbank files released later provided the most convincing anecdote in this area for me. Essentially, if WIV had a large number of sequences they were using, they should have shown up in publications, in research presentations, etc.

### Lack of evidence of circulating virus prior to HSM superspreading event, implying early HSM infections contained (or were close) to the index case

I think ZO won this part of the debate cleanly. Below is a summary of the data presented by ZO showing that SARS-CoV-2 couldn't have been circulating prior to the HSM. The two most convincing pieces of evidence for me were the throat swabs of influenza-like-illness (ILI) patients and routine blood draws. Neither of these would be susceptible to the ascertainment bias that both sides agreed existed for a brief late-December to late-January period. They were retrospective analyses that found no evidence of seropositive patients.

ZO debate 1, slide 116 - Wuhan, Baidu data shows very little signal for symptoms of a circulating coronavirus prior to December

ZO debate 1, slide 114 - Wuhan, 3850 blood bank samples from Sep-Dec 2019 show no positives for antibodies

ZO debate 1, slide 113 - Wuhan, 640 throat swabs Oct 6th - Jan 21st 2020. No positives in 2019, 9 in 2020.

ZO debate 1, slide 112 - WHO report data with hundreds of influenza like illness samples and no positives in 2019 from Hubei and Wuhan.

ZO debate 1, slide 110/111 - Wuhan excess mortality in early 2020, no Hubei excess mortality signal until 2020 (and it's very small)

ZO debate 1, slide 109 - Wuhan, no uptick in hospitalizations until December 2019.

ZO debate 1, slide 106 - WHO report, followed up on 92 ILI cases later, could find only 67, none with antibodies.

The LL side presented evidence here about Connor Reed being a November case (proving community circulation prior to HSM superspreading) which I found nearly impossible to evaluate. I think on balance, the fact that ILI samples and blood draws found no seropositivity is much stronger than self-reported claims of SARS-CoV-2 positive test results.

I do think there was a more interesting debate to be had here about seropositivity rates in the rest of the world. A cursory search suggested some (more credible) reports of early seropositivity in retrospective studies of US blood banks that would have been hard to explain under the ZO theory.

#### Lack of multiple spillovers or evidence of prior infections more proximal to the likely site of a zoonotic spillover

I think the observed lack of multiple spillovers or prior infections supports the lab leak hypothesis. As I summarized in a set of questions to Peter (shown below) it seems like it would be more likely that SARS-CoV-2 would have multiple spillovers under ZO, as SARS1 did.

SARS-CoV-2 doesn't have a slow/smoldering infection pattern like that seen in SARS-CoV (e.g. Yuri, debate 2, slide 55). It needs to spread fast or it dies out. This means that wherever fully functional SARS-CoV-2 was constituted (e.g. by recombining/evolving the functional FCS), it needed to quickly get into humans or it needed to retain human infectivity in a reservoir/intermediate species. Since a human didn't transport the infection to HSM, we conclude that there exists an intermediate host which doesn't establish a sterilizing immune reaction. It could also be the case that the intermediate host does build immunity, but due to logistics of animal husbandry (e.g. constant turnover of naive hosts in a farmed population), an infectious animal could be transported to HSM. Once at HSM, this animal (or animals) transmits two distinct viruses (A and B) to humans in rapid succession.

Under this logic, it seems far likelier that the first human outbreak location would be near the physical origin of the infectious virus. This is because the populations of intermediate hosts maintains high infectivity (due to immune or logistic factors) and would presumably transmit in locations with humans and these animals in close proximity (e.g. you conclude "There is nothing unique about the market, other than the wild animals" - debate 3, slide 15). The second most likely outbreak location seems to me when the infected animals were shipped to HSM. Trains have particularly low air turnover and (I think) have proven to be good at spreading the virus.

As we discussed somewhat at debate 1, Pekar et al. claims that the number of social contacts/linkages is important to establishing an outbreak. Under this thesis, I can understand how a human infection that occurs on a rural farm doesn't trigger a wider pandemic; not enough other human contacts. However, it's harder to understand how a train (bad ventilation, co-location between human and animals in transit or loading/unloading) doesn't seed multiple infections that would show up outside the HSM.

Collectively, it seems like a fairly narrow range of circumstances have to hold to: have an infectious intermediate host that maintains a reservoir but doesn't cause a chain of transmission at that reservoir, but is infectious enough to transmit two different strains in a short time frame at an unremarkable (re ventilation) market, but is not infectious enough to have transmitted during a train/truck ride or loading/unloading.

#### Apparent emergence of SARS-CoV-2 in Wuhan vs. other places in China

Wuhan is not close to high bat density. Both sides agree that the probability of zoonotic emergence of a pandemic in Wuhan (conditioned on it being a coronavirus in China) is ~1%. ZO argues for up to 5% based on some claims about rail lines and population density that are hard to evaluate. In short, it seems relatively unlikely that Wuhan would be the first place that a bat-derived coronavirus pandemic would emerge.

#### FCS appears to be an insert

See extended discussion in See the extended discussion in the cells for the update factor labeled "P(12 nt insert | ZO)".

I have reproduced below an extended discussion between Yuri and I, in questions given before debate 3. Yuri's responses are highlighted in green.

In summary, this length of insert appears rare and is more likely under a LL. More convincing to me was the idea that the sequences on either side of the insert look like they are from the same virus, it doesn't appear to be multiple recombination events. Mitigating this is what appears to be sequences of coronaviruses across the other *Betacoronavirus* subgenera that have various S1/S2 sites that look like they are on the way to becoming functional FCSs. It seems that perhaps the viral populations have some stable equilibrium where:

- The majority don't have an FCSs in bats because they are reasonably detrimental, but
- mutants with them that find a susceptible host (e.g. a human) have such a huge fitness advantage that overall the ability to develop an FCS is not selected away from

Similar equilibria are found in bacteria - often around resistance to antibiotics.

#### **Response to Judge Will**

As I alluded to last time, I have been trying to summarize and compare some of the key claims discussed so far. The three issues highlighted below are places where I would like to understand each side's position better. The first two issues don't necessarily need a written response. The third issue is important for the debate, but does require you to fill out the table. If this is too much work (understandable), I will fill this out to the best of my abilities from the debates and slides.

**The first issue is the FCS. I was expecting there to be more discussion of the molecular mechanisms of FCS formation, perhaps contrasting Yuri's claims in his 2021 Bioessays publication vs. suggestions of virologists like Zhou, Anderson, or**

Gallaher. I don't need this discussion, but I want to raise some points and understand what each side thinks of them. I would like Yuri/Saar to weigh in on 1-5, and Peter to weigh in on 1 and 2.

1. I want to understand whether both sides agree that 'intermediate' forms of FCS have been identified in the subgenera *Sarbecovirus*. Peter shows a bat *Sarbecovirus* with an S1/S2 insertion one nucleotide away from RXXR, and a recently discovered bat *Hibecovirus* (subgenus closest to *Sarbecovirus*) with an RXXR (Peter, Debate 2, slide 41). Coupled with the sequences shown on slide 58 (reproduced below) there seem to be a range of *Sarbecovirus* taxa with inserts at the S1/S2 junction ranging from 1-6 NT's away from a functional FCS.

[Yuri:] Not "inserts". Neither the Hibecovirus (i.e. not a SARS-like virus) with an existing FCS, nor SARS-like viruses like Peter's virus found in UK bats with RAKQ ("one nucleotide change from RAKR") have had those sequences inserted.

Same goes for the viruses with PAAR/PVAR at the S1/S2 junction – when only RmYN02 was found with PAAR, there was an argument (albeit a weak one) that PAA was inserted, but once other similar viruses were discovered with PAA or PVA at the very same spot (i.e. RacCS203, RacCS271, RacCS264, BANAL-116, BANAL-247), then claiming that all those viruses had insertions coupled with QTQT deletions became simply ridiculous. No alignment algorithm treats those as insertions anymore, see my post here:

<https://twitter.com/ydeigin/status/1694516196303892829>

Moreover, even if those PAA/PVA fragments did arise through an insertion (and a concurrent QTQT deletion) in some ancestor of all those viruses (RmYN02, RacCS203, RacCS271, RacCS264, BANAL-116, BANAL-247), then that insertion must have happened hundreds of years ago based on how distant those viruses are (and were sampled in different geographies: Yunnan, Cambodia, Laos).

That said, all SARS-like viruses already have a proto-furin cleavage site, as all SARS-like viruses have an R at the S1/S2 junction (e.g. TNSR in RaTG13), so if an amino acid that is 3 positions before it or after it mutates into another R, we will get RxxR which is a minimal FCS (e.g. RNSR in RaTG13). The fact that we don't see a single SARS-like virus with an FCS out of 1500+ known ones despite how easy it would be for an FCS to arise is strong evidence of negative selective pressure against an FCS in bats.

Finally, even if PAAR in RmYN02/BANAL-116/etc. mutates into PRRAR in those viruses to then have only 12 nucleotides from that novel mutation get inserted into a SARS2 progenitor at precisely the S1/S2 junction in a SARS2 progenitor is astronomically unlikely.

Not to mention that for PAA to mutate into PRRA in nature is highly unlikely in itself because PAA in RmYN02 or BANAL-116/246 is coded by CCT GCA GCG but PRRA in

SARS2 is coded by **CCT CGG CGG GCA**, so you need the second alanine to disappear but the first one to be preserved, and the two CGG arginine codons to get inserted between the P and the first A. (This is something a genetic engineer is much more likely to have done – taken the PA coded by CCT GCA and inserted two CGG arginines between them to arrive at PRRA, and then just inserted the resulting sequence into a SARS2 progenitor.)

As it stands, SARS2 does not have any signs of natural recombination with PAA-containing viruses that could have resulted in the PRRA insertion — say, one side of the S1/S2 region (with QTQTNs) coming from a BANAL-52-like virus and the other side with the PAARVGT cleavage site fragment coming from a RmYN02-like or BANAL-116-like virus. Instead, both sides of the S1/S2 junction in SARS2 are from a BANAL-52-like virus, but then there's an insertion of just 12 nucleotides creating the PRRAR cleavage site.

2. I want to understand how each side thinks these forms came about. Specifically, Peter presents evidence that these forms are insertions in the same location as SARS-CoV-2. Aligned against the SARS-CoV-2 genome (insert and flanking R; PRRAR) these insertions are PA-A-R or P-V-A-R. A single additional R insertion would result in a functional cleavage site (e.g. PRAAR utilizing the flanking R). In his [Bioessays paper](#), Yuri claimed at least the RmYN02 strain shows evidence of mutation rather than insertion. [Zhou et al](#) claim this is an insertion, and Peter presented it that way, but I want to understand who's explanation is stronger.

Wuhan-Hu-1	MN908947.3	Y Q T Q T N S P R R A R S V A S Q S
RmYN02	EPI_ISL_412977	Y - - - N S P A - A R - V G T N S
RacCS203	MW251308.1	Y - - - N S P V - A R - V G T N S
BANAL-20-116	MZ937002.1	Y - - - N S P A - A R - V G T N S
BANAL-20-246	MZ937004.1	Y - - - N S P A - A R - V G T N S
RaTG13	MN996532.2	Y Q T Q T N S - - - R S V A S Q S
RShSTT182	EPI_ISL_852604	Y Q T Q T N S - - - R S V T S Q S
RShSTT200	EPI_ISL_852605	Y Q T Q T N S - - - R S V T S Q S
BANAL-20-52	MZ937000.1	Y Q T Q T N S - - - R S V A S Q S
BANAL-20-103	MZ937001.1	Y Q T Q T N S - - - R S V A S Q S
BANAL-20-236	MZ937003.1	Y Q T Q T N S - - - R S V A S Q S

Peter slide 58.

[Yuri: see my reply to question 1]

3. What is the probability that the PRRA insert could be created naturally in a two step process? For example, RmYN02 gets the PA-A insertion (so total sequence is PA-A-R) and then gets another single R insertion resulting in a functional PRAAR? This is the theory of Bill Gallaher (post referenced by Peter, debate 2, slide 79).

SARS-like CoVs, however, have never been observed with an FCS despite being just 1 or 2 nucleotide mutations away from one. So the SARS2 FCS is very unlikely to have come from bats, as bats have strong selective pressure against an FCS. This means it needs to happen in the intermediate host, so there is a short time frame for it to happen. Combining this with the observation of a single spillover in the entire world (or supposedly two at the same location and time), which indicates very few intermediate hosts were infected, it means this rare recombination with a non-viral genome has very few chances to occur.

#### Codon usage in FCS (double CGG)

I agree with LL that this is unusual. However, it feels somewhat cherry picked (as Yuri and I discussed in questions reproduced below). I think the fact that the virus has maintained this codon usage suggests that it is evolutionarily robust, which supports the ZO theory. It could be that the designers knew ahead of time, but then we shouldn't expect it to be that odd(?)

5. The double CGG found in the PRRA insert is not found anywhere in the *Sarbecovirus* subgenus (Yuri, debate 2, slide 38). I am having trouble assessing how unusual this is. Both sides agree it's evidence for the lab leak, but differ significantly in how much. One thing I find very difficult to understand is: the realization of how unusual the CGG codon is seems like textbook retrospective analysis. Peter points out that upon a first perusal, Yuri didn't notice this as odd (Peter, debate 2, slide 88). Furthermore, there is no a priori reason presented by the LL side why they should focus on the codon frequency. If DEFUSE had mentioned codons (or restriction sites) it would seem non-retrospective, but I haven't seen that evidence. Peter's point on slides 99-101 in debate 3 materials highlight how if you just look at statistical distributions of genome features, you are going to find a tremendous number of things that could subsequently look like a smoking gun. I think some sort of multiple hypothesis testing penalty is appropriate here, or at least a background distribution on how many "oddities" I should expect to see when doing this kind of genome gazing to assess my surprise at this fact.

- There is not much retrospective cherry-picking you can do with 6 out of 12 nt.
- To avoid this issue altogether, it may be easier not to think about the CGGCGG specifically, but rather view it as strong evidence that the insert is not from a viral source (viruses don't like CGG, so a duplet is very unlikely), meaning the insert comes from host RNA or from engineering. Insertions are almost always a result of replication failures (polymerase slippage and template switching)  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8077628/>  
Having an insert from a non viral source is very rare in nature, but expected for engineering.
- [Yuri:] Yes, in isolation, CGG-CGG coding for 2 arginines in a SARS-like virus is very rare but rare things could happen in nature. But seeing CGG-CGG in a 12-nt insertion that created a novel FCS, especially in light of DEFUSE, is extremely suspicious.  
Personally, my suspicion is amplified by the Faul site that the CGG-CGG fragment creates, in light of WIV's work on using RFLP before, and Shibo Jiang's work of inserting a novel FCS with a leading FCS and also using RFLP to verify his insertion. Jiang was a close collaborator of WIV, including on a pan-coronavirus therapeutic.

#### FCS insert out of frame and FCS suboptimal (PRRAR)

Neither of these factored strongly. Both seem to suggest ZO to some degree, but the LL side mitigates this by showing that the MERS sequences could have been a reasonable precursor.

#### Furin cleavage site (FCS) uniqueness in the *Sarbecovirus* subgenera

I spent more time on this question than was warranted by the importance in the debate. This is partially because the "uniqueness of the FCS" was one of the first things that got me interested in the origin debate. Ultimately, I didn't favor either side strongly on this question.

SARS-CoV-2 contains a furin cleavage site (FCS). This FCS is the only one found in the subgenus (*Sarbecovirus*) containing both SARS-CoV and SARS-CoV-2. The lack of FCSs in other taxa in *Sarbecovirus* is often cited as evidence of genetic engineering in SARS-CoV-2. Based on brief scans of some references (PMID32123347, PMID34778878), the classification of SARS-CoV-2 is as follows. The number in parenthesis is how many other groups are at the given level (e.g. there are 5 subgenera of which *Sarbecovirus* is one).

Kingdom: *Orthornavirae*

Phylum: *Pisuviricota*

Class: *Pisoniviricetes*

Order: *Nidovirales*

Suborder: *Cornidovirineae* (8)

Family: *Coronaviridae* (1)

Subfamily: *Orthocoronavirinae* (2)

Genus: *Betacoronavirus* (5)

Subgenus: *Sarbecovirus* (5)

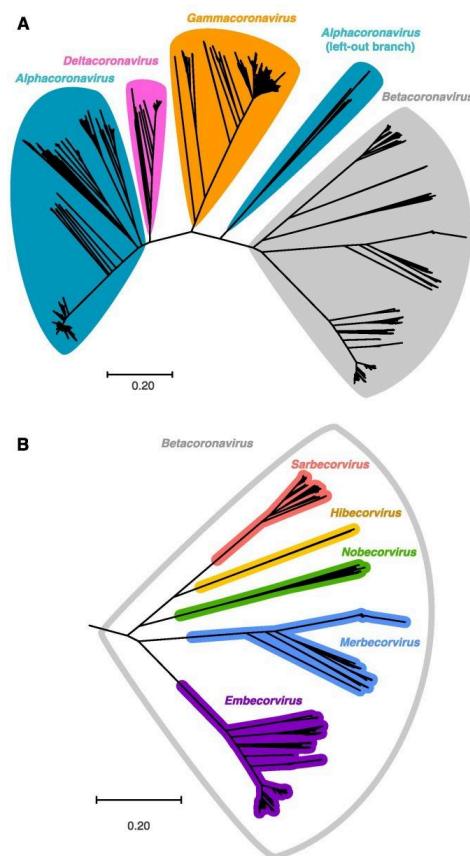


Image from Wu and Zhao (PMID33340798). Shown are the phylogenetic tree of (A) genera in *Coronaviridae*, and (B) the subgenera within the *Betacoronavirus*.

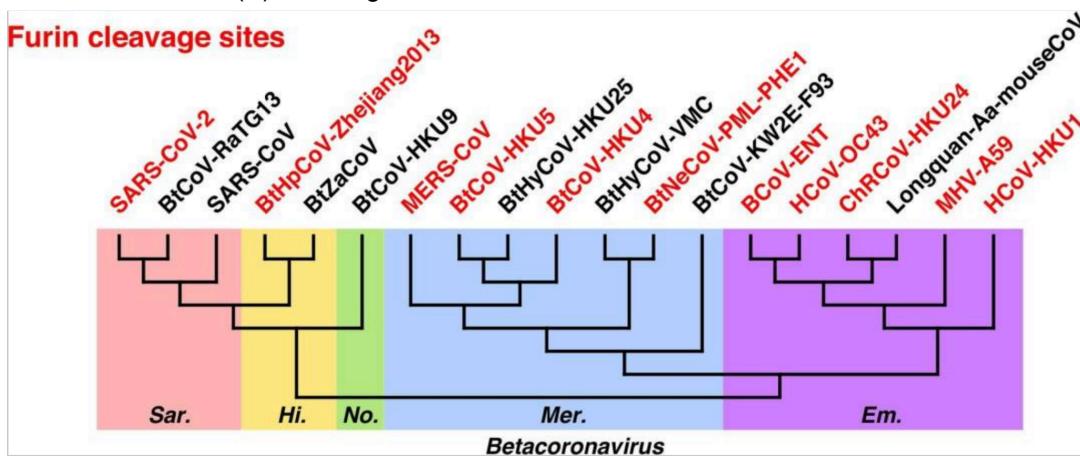


Image from graphical abstract Wu and Zhao (PMID33340798). Note the FCSs highlighted in this plot are RXXR and this is a highly pruned tree.

The genus containing the SARS viruses (*Betacoronavirus*) has five subgenera. Four of five of these subgenera have one or more taxa with an RXXR FCSs, including several with an RRXR FCS (e.g. HCoV-HKU1 - for data see PMID37141989).

In this debate the LL side mentions that if the FCS were the result of natural processes, it would “Not be the first virus in its family to have an FCS” (Yuri, Debate 2, slide 35). Overall, the LL side does not place much weight on the uniqueness of the FCS, but does mention it several times (e.g. see Yuri, Debate 2, slides 19, 20 and 35).

The Z side makes three arguments that respond to the LL claim. First, the Z side claims that the FCS is not that unique either within *Betacoronavirus* or in the greater viral realm (Peter, Debate 2, slides 40, 42, 49, 50). Second, the Z side suggests more bat coronaviruses with FCSs will be found. As evidence, Peter shows a bat *Sarbecovirus* with an S1/S2 insertion one nucleotide away from RXXR, and a recently discovered bat *Hibecovirus* (subgenus closest to *Sarbecovirus*) with an RXXR (Peter, Debate 2, slide 41). Third, the Z side points out that at least four closely related coronaviruses have insertions at the S1/S2 junction (Peter, Debate 2, slides 57, 58). Aligned against the SARS-CoV-2 genome (insert and flanking R; PRRAR) these insertions are PA-AR or PV-AR. A single additional R insertion would result in a functional cleavage site (e.g. PRAAR utilizing the flanking R). This suggests that insertions at the S1/S2 junction can occur naturally.

#### How did I evaluate this question?

For the purpose of this question, I am treating all FCSs as the same (e.g. ignoring differences between RXXR and RRXR, insert length, in/out of frame, etc).

Both sides agree that, conditioned on observing a pandemic, it is likely that the causal agent would have some unique genetic traits differentiating it from its non-pandemic relatives (Yuri, Debate 2, slide 20; Peter, Debate 2, slides 40-50). The question is should I be more surprised that SARS-CoV-2 has the only FCS in the *Sarbecovirus* if it evolved naturally or was constructed in a lab? I think I need to consider two pieces of information here:

1. What does it mean to be “SARS-like”?
2. How common are FCSs in natural origin SARS-like viruses?

In regards to (1) - I am relying on my experience in bacterial systematics and phylogeny. In short, taxonomy is not synonymous with phylogeny. Scientists demarcate taxonomic groups for a host of reasons and multiple different systematics groupings can reveal valuable information. They are also somewhat arbitrary. It is good to ask the same question using different taxonomic levels and see how the answer changes.

In regards to (2), the evidence presented in the debate shows that four of five of the subgenera of *Betacoronavirus* contain at least one taxon with an FCS (Peter, Debate 2, slides 40, 41 and PMID33340798). Similarly, a bat-infecting *Hibecovirus* (closest subgenus to *Sarbecovirus*) with an FCS and a bat-infecting *Sarbecovirus* one nucleotide away from an FCS have been identified (Peter, Debate 2, slide 41). Finally, there are several closely related viruses (including RmYN02 - with 93.3% overall nucleotide identity) with insertions at the S1/S2 junction that align with the PRRA insertion in SARS-CoV-2. These viruses look somewhere between FCS- and FCS+.

Wuhan-Hu-1	MN908947.3	Y Q T Q T N S P R R A R S V A S Q S I
RmYN02	EPI ISL 412977	Y - - - N S P A A R - - V G T N S
RacCS203	MW251308.1	Y - - - N S P V A R - - V G T N S
BANAL-20-116	MZ937002.1	Y - - - N S P A A R - - V G T N S
BANAL-20-246	MZ937004.1	Y - - - N S P A A R - - V G T N S
RaTG13	MN996532.2	Y Q T Q T N S - - - R S V A S Q S I
RShSTT182	EPI ISL 852604	Y Q T Q T N S - - - R S V T S Q S I
RShSTT200	EPI ISL 852605	Y Q T Q T N S - - - R S V T S Q S I
BANAL-20-52	MZ937000.1	Y Q T Q T N S - - - R S V A S Q S I
BANAL-20-103	MZ937001.1	Y Q T Q T N S - - - R S V A S Q S I
BANAL-20-236	MZ937003.1	Y Q T Q T N S - - - R S V A S Q S I

Hearing that “SARS-CoV-2 has the only FCS in the *Sarbecovirus*” is initially very surprising under natural origin. However, when I consider this claim, it’s not clear why *Sarbecovirus* is the right reference category compared to *Betacoronavirus*. Both groupings spawned a near-pandemic virus (*Sarbecovirus* - SARS-CoV, *Merbecovirus* - MERS), and there are many *Betacoronavirus* taxa with an FCS. If I do restrict myself to *Sarbecovirus*, the data shows intermediate FCS-like forms within the *Sarbecovirus* - the FCS is not ‘never seen’ in *Sarbecovirus*.

I have summarized the interplay of the evidence in the following table.

SARS-like means	How unique is the FCS at this level?	Drawing a pandemic-ready virus with an FCS from the natural reservoir would be...	But...
Compare at subgenera level - i.e. <i>Sarbecovirus</i>	SARS-CoV-2 is the only virus in its group which has an FCS.	Unprecedented - repeated draws from the <i>Sarbecovirus</i> natural reserve will not produce an FCS containing taxon	There are <i>Sarbecovirus</i> taxa that appear in various stages of developing an FCS including inserts at the same location and RAKQ amino acid sequence.  There are also not millions of distinct taxa in the <i>Sarbecovirus</i> - maybe 10's-100's
Compare at genera level - i.e. <i>Betacoronavirus</i>	SARS-CoV-2 is one of ~10 viruses within its genus ( <i>Betacoronavirus</i> ) which have an FCS, and <i>Sarbecovirus</i> is among the 4 of 5 subgenera of <i>Betacoronavirus</i> which	Likely - repeated draws from the <i>Betacoronavirus</i> natural reserve will produce an FCS containing taxon	

	have an FCS containing taxon		
Compare at highest level - i.e. <i>Virus</i>	SARS-CoV-2 is one of many viruses across many viral phyla which has an FCS	Likely - repeated draws from the natural reserve of viruses will produce an FCS containing taxon	

### Is SARS-CoV-2 unusually well adapted to humans?

This part of the debate was not a clear win for either side. The questions at issue were

- Did SARS-CoV-2 show lower whole-genome evolution rates than would be expected for a new zoonosis?
- Did specific parts of the genome show unusual pre-adaptation (e.g. the RBD)?
- Are there features that would be expected not to be in early versions of the virus that spilled over to humans?

Ultimately, I couldn't make a firm conclusion on the RBD's level of pre-adaptation to humans. Both sides made contested fact claims which I couldn't resolve.

Supporting the ZO side are

- Conditioned on a pandemic level human virus, we ought to expect a good RBD
- The RBD looks good, but not wildly better than others
- You wouldn't choose this RBD if you were basing your initial research off SARS. You'd have to know something else first.

Supporting the LL side are

- The RBD appears unusually good

The question about levels of pre-adaptation at the whole genome level and the early mutation rate were easier to resolve and are described below. In summary, I find this doesn't give strong evidence to either side. The virus appears to have a slightly faster mutation rate in deer than it did in early humans (~35%). The paper LL mainly relies on suggests this rate can be species specific, and it certainly doesn't appear to strongly suggest pre-adaptation.

The questions about genomic features that were good in bats/maladaptive in humans were never compared in the debate. My explanation for a stable population without an FCS, but close to having one (see FCS appears to be an insert above) resolves this for me.

*Features that are good in humans/bad in bats:* Conditioned on observing a pandemic, we assume that the responsible virus must be at least reasonably well adapted to humans immediately prior to the first superspreading event, regardless of its origin. If the particular genomic features that make the virus well adapted to humans ALSO appear to be seriously maladaptive in the presumed viral origin species (e.g. bats), it makes a zoonotic origin less likely. This is because the virus would have to acquire and sustain the maladaptive traits in the origin species while waiting to spillover and competing with more fit variants. In contrast, a lab origin would remove this competitive pressure.

In this debate, both sides agree that the FCS is of critical importance to human infectivity. They also agree

that the FCS is likely detrimental to SARS-CoV-2 spread/maintenance in bats. *Prima facie*, this fact pattern supports a lab origin.

LL also argues that a specific single nucleotide difference between SARS-CoV-2 and all other known Sarbecoviruses demonstrates the same pattern. Specifically, the T372 variant found in all bat Sarbecoviruses allows fecal-oral transmission which is critical to fitness in bats. In SARS-CoV-2, this residue is an A (A372) which is detrimental to bat transmission, but substantially improves human transmission. ZO contests the importance of this claim.

There are three obvious ways the ZO side can resolve this. First, if an intermediate host is involved where the human-critical mutations are at least fitness neutral, the mutations have a larger chance of surviving and emerging as a pandemic capable virus. Second, if the virus without the adaptations can transmit in humans to some degree, it can acquire the human-specific adaptations over a long time before emerging (apparently) all of a sudden in pandemic-capable form. Third, if the frequency and number of bat/human contacts is large enough, even an unfit bat variant may survive long enough to make it to a human where the mutations will be selected for.

In this debate, the ZO side favors a rapid spillover with a negative binomial distributed infection pattern. In short, ZO argues that the index case is close in time to the first superspreading event at HSM, which makes them more reliant on intermediate hosts or high frequency bat/human contact. The ZO side also contests the overall impact of the good in human/bad in bat mutations.

*Features that look unnatural:* If SARS-CoV-2 showed specific features which were unusually well adapted to humans, and those features were unusual/absent in closely related viruses, we might suspect an active design process. We should be even more suspicious that the virus was designed, if close to the point of origin there were researchers known to be interested in designing pandemic capable viruses.

In this debate, LL claims that the high binding efficiency of the SARS-CoV-2 spike protein to human ACE2, the existence of the FCS, the local genetic context (cleanliness of insert) of the FCS, the nucleotides encoding the FCS, the rate of SARS-CoV-2 evolution, and the N-glycan pattern (T372A mutation) are all unusual features that would not be expected (or have never been observed) from a randomly selected Sarbecovirus.

ZO heavily contests all but the N-glycan pattern. In addition, the ZO side makes some counterclaims that the ‘unnatural’ features found in SARS-CoV-2 would be unexpected under a laboratory origin.

#### *Deer/mink adaptation rates*

L claims (debate 2, slides 72-75) that there is a lower early mutation rate in SARS-CoV-2 in humans than would be expected under a zoonotic origin. To compare this, they point to evidence of higher early mutation rates in other immunologically naive species (deer and mink) and compare mutation rates. Yuri notes evidence in McBride, Garushyants, Franks et al. 2023, where the authors conclude that early variants evolved at ~76% the rate in humans compared to white tailed deer when averaged across 5 genomic regions. The authors note (and LL does not mention) that dN/dS is higher in the human spike protein, and that different species can have inherently different mutation rates. LL also cites mink early evolution rates being 9x human, but ZO contests that the source LL chose is substantially higher (~1 order of magnitude by eye) than 5 other papers (ZO debate 2, slides 100-102).

I think LL wins that the deer evolutionary rate is slightly faster, and I am uncomfortable concluding for either side on the mink evolutionary rate. However, I am not sure how this helps LL. Consider LL debate

2, slide 75 “Under Lab Leak: The low early mutation rate is expected under DEFUSE style research, which screened for RBDs that match human ACE2.”

DEFUSE style research optimizing the RBD would only lead to pre-adaptation of the RBD. The rest of the genome should not be under selection. The data cited by LL is consistent with pre-adaptation of the entire genome, but they cite no mechanism that would generate whole genome pre-adaptation under DEFUSE style research.

One plausible response would be that during construction of GOF mutants, the viral passaging that occurred allowed pre-adaptation. Two objections to this. First, I don't think in vitro passaging would lead to the same pan-genome pre-adaptation as in vivo passage (e.g. infection). Second, the ZO side makes strong arguments suggesting that passage results in tell-tale genetic signatures rapidly accumulating (ZO, debate 2, slides 122-133).

Ultimately, I conclude that the deer/mink data does not support pre-adaptation of SARS-CoV-2 to humans.

## Section 3

In this section, I describe some of the factors that were important in my Bayesian analysis. Please see the full spreadsheet as it contains >10,000 words of explanation.

### Base rate

For both my decision making methods I had to assess the base probability of lab leak or a zoonotic event - e.g. the probability of each hypothesis prior to taking the updated information into account. In my opinion, from a strategic perspective, neither side spent enough time in this debate establishing what my prior odds should have been.

In this debate, I decided that I had to estimate the following quantities:

P(WIV made SARS-CoV-2 by 2019)

P(WIV leaked SARS-CoV-2 in 2019)

P(Natural processes made SARS-CoV-2 by 2019)

P(Zoonotic processes leaked SARS-CoV-2 in 2019)

To get my starting odds ratio I compared:

$$P_0(LL) / P_0(ZO) = [P(WIV \text{ made SARS-CoV-2 by 2019}) * P(WIV \text{ leaked SARS-CoV-2 in 2019})] / [P(\text{Natural processes made SARS-CoV-2 by 2019}) * P(\text{Zoonotic processes leaked SARS-CoV-2 in 2019})]$$

The reason I decided on this calculation was because I couldn't figure out how to incorporate arguments about whether WIV could have made, or would have leaked, SARS-CoV-2 in my Bayesian update process. Specifically, updates took the form of  $P(\text{new evidence} | LL) / P(\text{new evidence} | ZO)$ . If I wanted to discount the probability of a lab leak based on evidence that WIV could not have made SARS-CoV-2, or didn't leak if they did, I couldn't figure out how to do that with a condition of LL has occurred. I suspect that if I were more comfortable with Bayesian analytical methods and mathematically more gifted, I could have done this a better way.

### Would WIV have leaked SARS-CoV-2 if it made it?

LL advanced three main arguments.

1. There have been many previous lab leaks across the world, including four leaks of SARS1. The conclusion was that the base rate of leakage of pathogens from lab is much higher than popularly believed.
2. WIV practices for working with infectious viruses were particularly bad and substantially increased the likelihood of release. The LL side focused on statements by the Zhi (check spelling) lab that the DEFUSE work (as well as other coronavirus work) were done in BSL-2 or BSL-3. Most researchers, including some of the DEFUSE authors, believed that BSL-2 was woefully insufficient for the work they proposed.
3. Biotech trends over the last twenty years all increase the baseline probability of lab leaks. The LL side noted decreased sequencing costs, decreased cost/difficulty in using synthetic biology tools, and increased construction of BSL-3 and BSL-4 labs in China. In short: there is a (small) probability of leak from any BSL-2/3/4 lab, and more labs doing more work (because of cheaper tools) linearly increases this risk.

### Previous lab leaks

Please see my section in "Skepticism of LL evidence" for a more in depth discussion of this. As a brief summary of that section, I have synthesized my main concern below.

The examples of previous lab leaks that LL cites include a mixture of LAIs, environmental release due to sterilization failures, accidental release of bioweapons, and insufficient decontamination in shipment. In this debate, LL really only defended an LAI scenario, even when I prompted them to consider an environmental release. LL claimed that they were agnostic to how the pathogen would be released (and thus could capture the probability of leaks associated with e.g. sterilization failure) but I found this unpersuasive; they never discussed these scenarios. Thus, I felt that they were really only arguing for one potential scenario (LAI), but inflating the odds of that scenario with every other type of lab leak, even ones that were so distinct as to be incomparable (e.g. the 1979 anthrax outbreak from a dedicated offensive bioweapons facility; again, see my "meta considerations" section for in depth discussion). There are in depth assessments of per-year lab leak probabilities from a variety of sources, but those were not included in this debate, and I didn't look into them extensively because neither side brought them up.

#### WIV practices

I think the LL side convinced me that WIV practices for doing GOF work were substantially inadequate. LL cites instances of BSL-2 work with Coronaviruses in previous research papers by the Shi group, Chinese scientific body guidelines, and the (proposed) DEFUSE grant work. The cited data modify the probability of LAI, but don't speak much to environmental release, etc.

#### Biotech trends

It was hard to evaluate these claims. Collectively, it seems like I should assign a probability of leak per unit time weighted by a huge number of factors (e.g. work done in BSL2, number of workers, amount of infectious agent produced per person per hour, etc.). Cheaper sequencing and more BSL3+ labs seem like they increase the number of labs over which this calculation is done, but I am not sure how to evaluate that in the context of a leak from WIV. More relevant evidence would have spoken to the activities at WIV - did they have

Ultimately, I think a much more fruitful discussion could have been had on the prior probability. I believe the relevant factors influencing the probability of a lab leak are: the goal of the research, the scale of the operation (numbers of infectious particles created/year), the number of types of experiments - especially those at multiple different biosafety levels or where samples are transported out of labs, the laboratory equipment (e.g. hoods, respirators, aerosol minimizing lab equipment), the sterilization infrastructure (on-site vs. offsite, autoclave/EO/H<sub>2</sub>O<sub>2</sub>, etc.), and the training and experience of the lab workers.