# Twitter Thread by Billy Bostickson ■■&■ ■

**Billy Bostickson ■■&■ ■**
@BillyBostickson

**1. Thread: RaTG13 Real or Fake?**

**A Norwegian Scientist @Vehuardo recently claimed RaTG13 is an "authentic virus" https://t.co/tAXMNTwdhA**

**This failed to take into account the work of our DRASTIC colleagues & other published analyses.**

**This thread is a polite attempt to educate.**

2. https://t.co/tw606vgHwp
which says RaTG13 doesn't make sense on how it was sequenced, or how the data got so messed up in the lab as to include copy paste mistakes, bat dna, pangolin dna, rice dna, mice dna, & nobody should use it to form conclusions until it's properly vetted

2 (cont)
"sequence is incomplete & provided segment doesn't function. The identified markers suggest partial match basis & actual virus itself as described by the sequence can't functionally exist due to incoherent structures that appear to be the result of sample contamination"

> Basically, modern gene sequencing uses highly automated machines that sequence really fast ("next gen sequencing") but which are less accurate and an only read small snippets of DNA or RNA at a time. The new machines are thousands of times faster, but to get an accurate result, the machines have to read the DNA/RNA over and over many times in small snippets (kind of like reading a randomly chosen a sentence or portion of a sentence out of a book at a time, only the way the sentences are read there are often small errors in the sentences too). Then they use computers to compare all the readouts and try to figure out where the errors are that the machine made in reading the DNA, and to string all the little snippets that the machine reads out into one long sequence. Also, because the machines are so fast and so much cheaper, rather than purifiying the DNA sequences they are trying to analyze, a lot of the times, they just throw a lot or DNA or RNA into the machine at once and let the computer figure out what is what. The DNA or RNA snippets that doesn't belong to the virus (belong to bacteria or whatever other junk was in there) are simply ignored using the computer programs and are not reported in the publication.
>
> Rather than look at the final assembled sequence that the lab in Wuhan reported for this virus called RaTG13, this paper looks at the raw data readings coming off of the machines *before* the computer ran through all the data to remove all the errors, strung the snipped fragments into one single virus sequence, and before all the "junk" sequences were discarded. They are saying that the sequence that was reported in publication for RaTG13 can't be fully reconstructed from the underlying data. And they are saying that there are a lot of other background sequences from stuff that don't make a whole lot of sense based on how the bat specimen was supposedly collected.

3. RaTG13 integrity is further attacked here:
The Validity of critical pieces of evidence for the natural origin of SARS-CoV-2 is Dubious, and needed to be reconsidered
https://t.co/yTNYcsbees

4. and here:
Major Concerns on the Identification of Bat Coronavirus Strain RaTG13 and Quality of Related Nature Paper
https://t.co/o3jWQWwJFQ

5. With reference to the levels of bacteria found and certain other anomalies, we have:
https://t.co/qVPC1HTxcA
by @MonaRahalkar and @BahulikarRahul

6. and another analysis focusing on contamination:
https://t.co/DUvBGawcso
Anomalies in BatCoV/RaTG13 sequencing and provenance
from @flavinkins

7. Then we have the anomalies regarding the amplicon sequences, which also remain to be answered:
https://t.co/dYpqvcebuo
and
https://t.co/QUxsxcmnAo
initially proposed by @interne41914499 #DRASTIC

8. Some more recent work which investigates the authenticity of RaTG13 can be found here:
https://t.co/GGIX4UH06b
by @MonaRahalkar and @BahulikarRahul

9. We should not forget the original attack on RatG13 from
@nerdhaspower which can be found here:
https://t.co/U1F1f8p9A6
much hated by the virology community ;)

10. I believe @Vehuardo addressed some of the concerns in this paper in the recent https://t.co/SpcUDqQuBX note, but in my opinion the challenges were insufficient to overthrow the main claims in the paper.
https://t.co/XmPEVP20f5

11. Finally, another anomaly discovered by @notoriousFIL
Discovery of 2 sites in the raw data that don't match the published RaTG13 genome, also noted by
@shingheizhan who found this mismatch & 2 bases which are identical to their counterparts in SARS-CoV-2,
"kind of bizarre"

12. "The SRA data doesn't appear to agree with the reference genome, there are gaps, different consensus and different bases" They could not explain these discrepancies:
https://t.co/deLDt5kUsu
How about you, can you explain them?

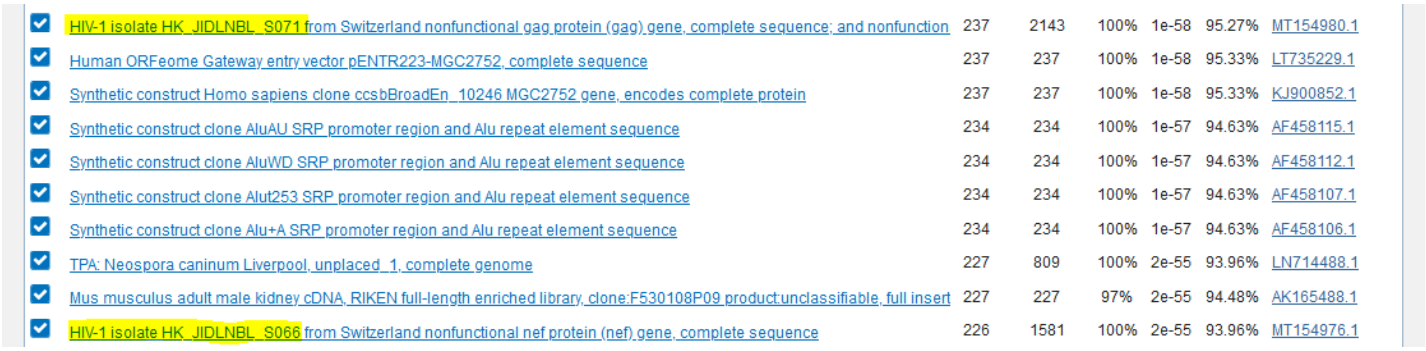
> I mentioned this above because its the most unambiguous example of this. TG also happens to be the SARS-CoV-2 variant at this position. The amino acid selected for at this position (ZN2 domain) in ExoN also seems to follow a pattern for Bat-derived vs common-cold coronaviruses. pic.twitter.com/iMIf8l6uAp
>
> — notoriousFIL (@notoriousFIL) September 17, 2020

13. Is RaTG13 a viable virus?
Apart from some ambiguous results from computer models & pseudoviruses, the consensus is that RaTg13 is not a viable virus, supported anecdotally by Beijing Scientists who failed to recreate and infect animals with it.
https://t.co/AwUdiocb6f

> has anyone heard any updates on this RaTG13 research? https://t.co/cM8yLXHLwg
>
> — Billy Bostickson \U0001f3f4\U0001f441&\U0001f441 \U0001f193 (@BillyBostickson) September 23, 2020

14. Not surprising, as @flavinkins pointed out, since last 4 amplicons of RatG13 clearly display signatures of DNA synthesis assembly errors, meaning they were synthetic, in-vitro materials designed to “fill gaps” for a contiguous assembly to join sequences never found together

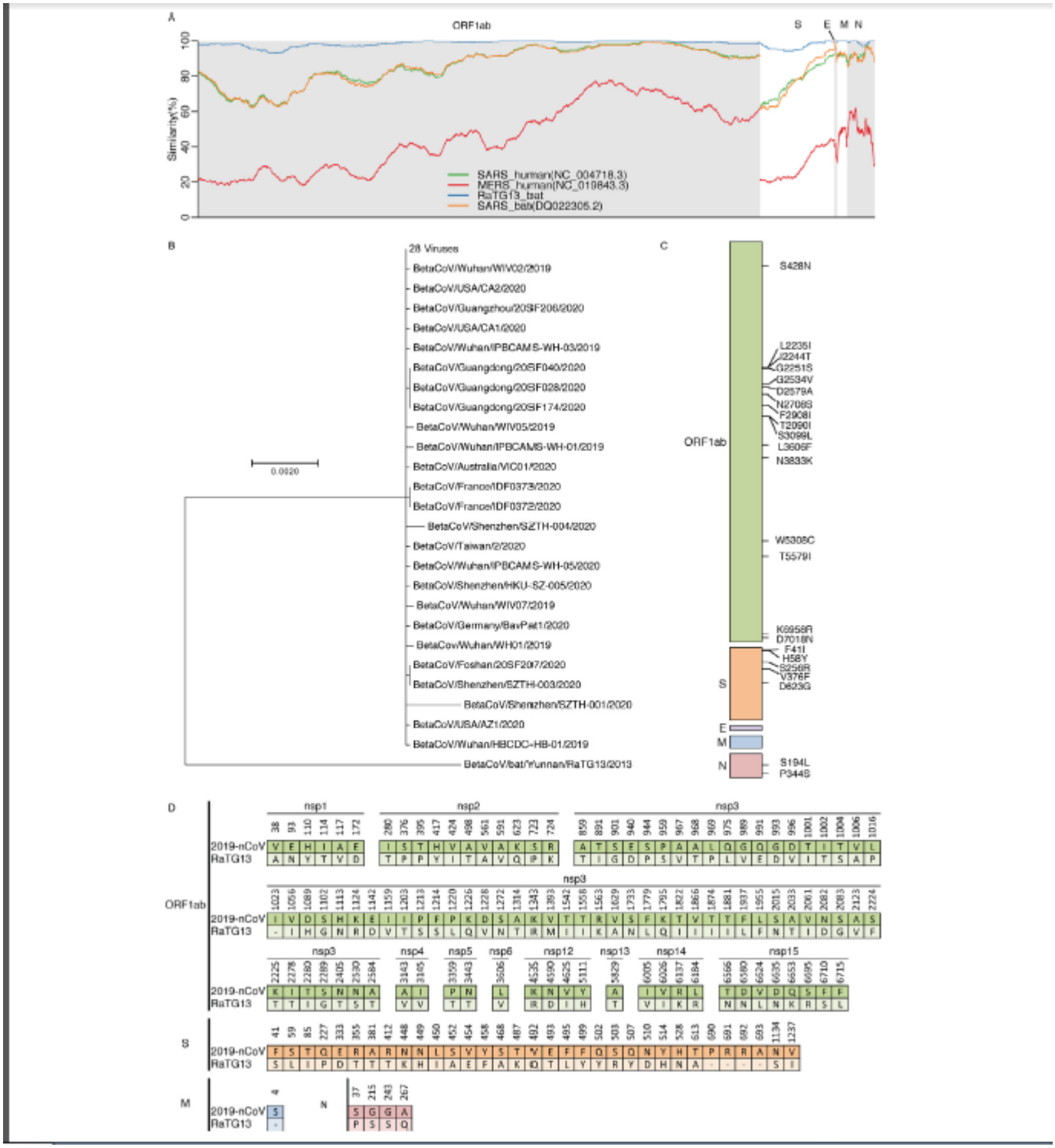
15. Traces of Human ERV and HIV-1 derived material was seen in RaTg13, pointing to it being a classical cell culture product. The HIV was was apparently from the lentiviral plasmids and the nucleotide in question was a part of the Gag gene.



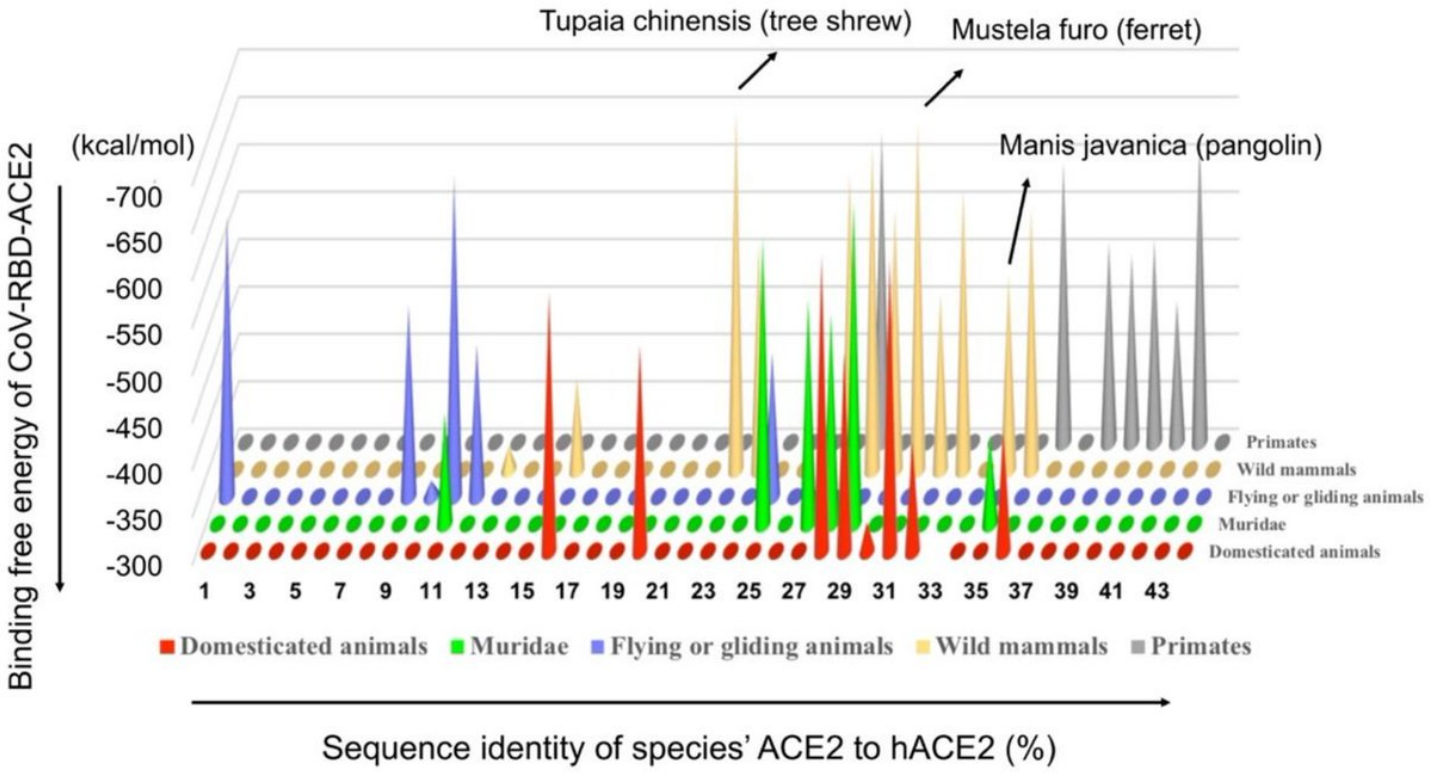
| | | | | | | |
|---|---|---|---|---|---|---|
| HIV-1 isolate HK_JIDLNBL_S071 from Switzerland nonfunctional gag protein (gag) gene, complete sequence; and nonfunction | 237 | 2143 | 100% | 1e-58 | 95.27% | MT154980.1 |
| Human ORFeome Gateway entry vector pENTR223-MGC2752, complete sequence | 237 | 237 | 100% | 1e-58 | 95.33% | LT735229.1 |
| Synthetic construct Homo sapiens clone ccsbBroadEn_10246 MGC2752 gene, encodes complete protein | 237 | 237 | 100% | 1e-58 | 95.33% | KJ900852.1 |
| Synthetic construct clone AluAU SRP promoter region and Alu repeat element sequence | 234 | 234 | 100% | 1e-57 | 94.63% | AF458115.1 |
| Synthetic construct clone AluWD SRP promoter region and Alu repeat element sequence | 234 | 234 | 100% | 1e-57 | 94.63% | AF458112.1 |
| Synthetic construct clone Alut253 SRP promoter region and Alu repeat element sequence | 234 | 234 | 100% | 1e-57 | 94.63% | AF458107.1 |
| Synthetic construct clone Alu+A SRP promoter region and Alu repeat element sequence | 234 | 234 | 100% | 1e-57 | 94.63% | AF458106.1 |
| TPA: Neospora caninum Liverpool, unplaced_1, complete genome | 227 | 809 | 100% | 2e-55 | 93.96% | LN714488.1 |
| Mus musculus adult male kidney cDNA, RIKEN full-length enriched library, clone:F530108P09 product:unclassifiable, full insert | 227 | 227 | 97% | 2e-55 | 94.48% | AK165488.1 |
| HIV-1 isolate HK_JIDLNBL_S066 from Switzerland nonfunctional nef protein (nef) gene, complete sequence | 226 | 1581 | 100% | 2e-55 | 93.96% | MT154976.1 |

16. Back to Viability:
We know that RaTG13 does not even result in infection in bats due to weaker affinity between its S-domain & bat ACE2 receptor compared to SARS-CoV-2 & human ACE2. So, if this is the case, how could it have even emerged from bats?
https://t.co/QbFpgWXYWk

17. Despite similarities between RatG13 & SARS-CoV-2, there is little evidence that RatG13 could infect a bat, let alone a human
https://t.co/R1WDjxzvvf
The question arises, if Scientists are so sure RaTg13 is viable & natural, why haven't more experiments been done to test it?




18. A withdrawn paper used a computer model to analyse potential RaTG13 binding to different animal ACE2, suggested binding would be highest to common lab experimental animals, Ferret & Tree Shrew
https://t.co/z9LNvUzoR4
However, no further in vivo work has been attempted since.




19. In an experiment in July, pseudoviruses bearing spike protein derived from bat raTG13 allegedly could use ACE2 of a diverse range of animal species to "gain entry"
Just "gain entry"?
We have no results about infectivity, so why no in vivo experiments?
https://t.co/3FcszCgk4B

raTG13 S (Fig. 3B&C). Among the five pseudoviruses, bat raTG13 spike appears to be the least efficient one in mediating pseudovirus entry through hACE2. Insertion of PRRA into raTG13 Spike further reduced the pseudovirus entry into all cell types that were tested (Fig. 3A-C). Pseudovirus bearing the Pangolin spike is

20. An earlier paper in April claimed that
"a wide range of ACE2 orthologs support binding to RBD proteins of Bat-CoV RaTG13"
https://t.co/neyH77DnAr
It was later revealed that they used the wrong species of bat for ACE2 sequence (Rs instead of Ra) thus invalidating their claim

**KR** • 6 months ago

I was wondering if you had any thoughts about why the bat coronavirus RBD doesn't bind nor enter the cells expressing the bat ACE2 protein?

Reply • Share ›

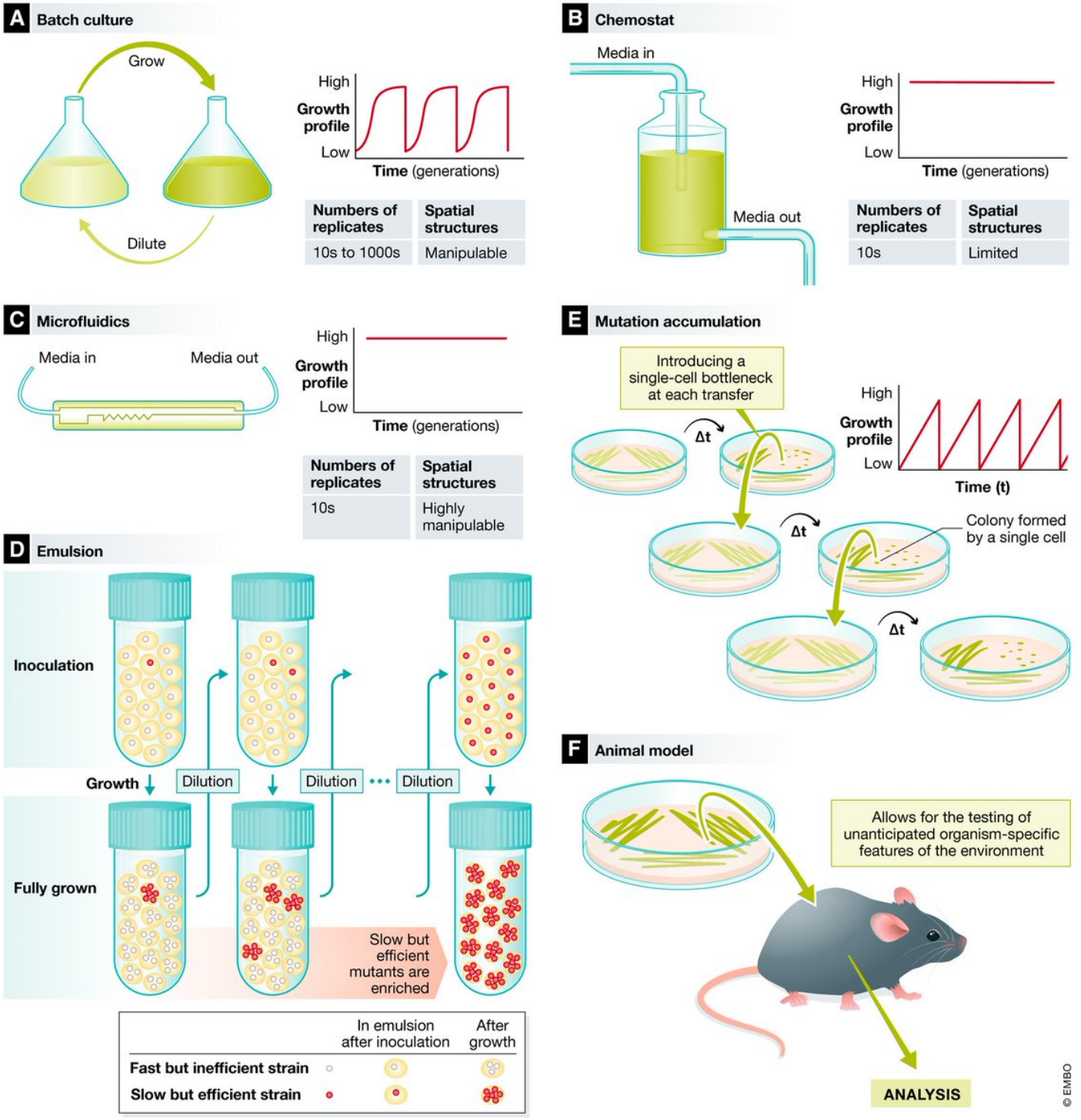
**GUOCAI ZHONG** ➔ KR • 6 months ago

Thank you for asking this. It's an excellent question. Here are what I got when I had the same question before. There are over 1,000 different species of bats around the world. The bat coronavirus RaTG13 was found in one bat species called Rhinolophus affinis (PMID:32015507), while the ACE2 sequence used in this study is from another bat species, Rhinolophus sinicus (GenBank Acc. No. KC881004.1). GenBank doesn't have Rhinolophus affinis's ACE2 sequence. I guess the there are some amino-acid differences between the ACE2 proteins of the two bat species, and some differences make the Rhinolophus sinicus ACE2 nonfunctional for supporting RaTG13 RBD binding and pseudovirus infection. We had provided the species and accession number information of the ACE2 orthologs in the Extended Data Table 1. But it might be helpful to include the species information to a main figure.

Reply • Share ›

21. our DRASTIC team member @Rossana38510044 also finds it "very unusual that the RBM of RaTG13 has no affinity for any other sequence in NCBI. It is indeed a great mystery how it has evolved."

I would suggest that it may have been the result of "directed evolution"




22. To summarise so far, a few flawed in-vitro experiments using RaTG13 pseudoviruses or ACE2 orthologs have revealed that the only class of ACE2 that can bind RaTG13 is ungulate ACE2
@flavinkins points out that despite high in-vitro activity, it does not have in-vivo infectivity

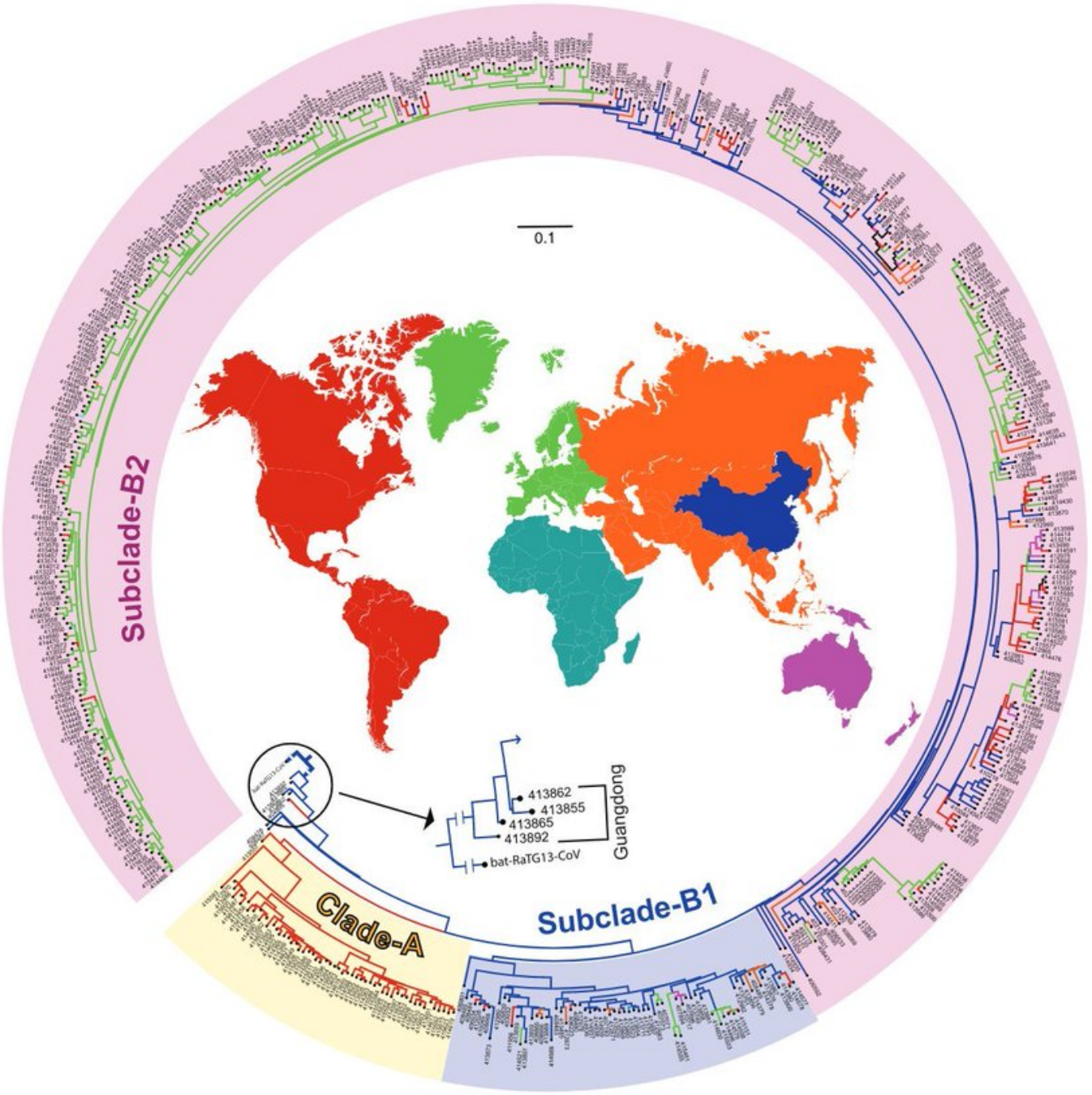
23. This is likely due to ungulate factors that block entry of Sarbecoviruses:
1 https://t.co/vELGSA8qO1
2 https://t.co/boEKPr2gKt
3 https://t.co/S8glmOJafa
This raises questions:
Where are the in-vivo experiments with RaTG13?
Can it even be recreated?
Why has nobody even tried?

24. Finally, there are the phylogenetic tree inconsistencies when RatG13 is used, clearly elucidated by #DRASTIC team member @AntGDuarte here:
https://t.co/OjJLMpZNZa

> RaTG13 & RmYN02 have also produced weird results in phylogenetic analyses of SARS-2, at odds with epidemiological data.
>
> With so many variables at play, it's easy to explain away one odd result. But anomaly seems the rule here, rather than the exception.
> https://t.co/GhAutTIr6t
>
> — Ant\xf3nio Duarte (@AntGDuarte) October 10, 2020

25. An example here that claimed to everyone's great surprise, that when the phylogenetic tree is rooted in RatG13, Guandong was the proximal origin in September 2019!
https://t.co/jWniyY9WRI

*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 June 2020 doi:10.20944/preprints202005.0332.v2



26. More Evidence of errors distorting Phylogenetic Trees
https://t.co/Vazefx6lz6
The PDF has more complete data and is essential reading:
https://t.co/qghoWdlMDV

27. There is ample evidence that RatG13 is a "construct" rather than a viable virus & little evidence to the contrary
@flavinkins suggests that the amplicons could have been generated by simulators such as MetaSim & short fragments synthesised & sequenced
https://t.co/zKT2Mmm4HW

28. @CZilcho concludes:
TG doesn't add credit to the RaTG13 story
Sanger amplicon only released in May to fill a gap in seqs
Complete SRR full of eukaryotic contamination that didn't even match its alleged host R. affinis
BALF or feces?
Smells like a mineshaft full of guano!

29. RaTG13 has been investigated and almost invariably the researchers are either befuddled, confused, shocked or surprised at the anomalies they found.
For example:
https://t.co/iofBxkQUXg
See Table S1. The whole genome nucleotide similarity https://t.co/DBINNuE35h

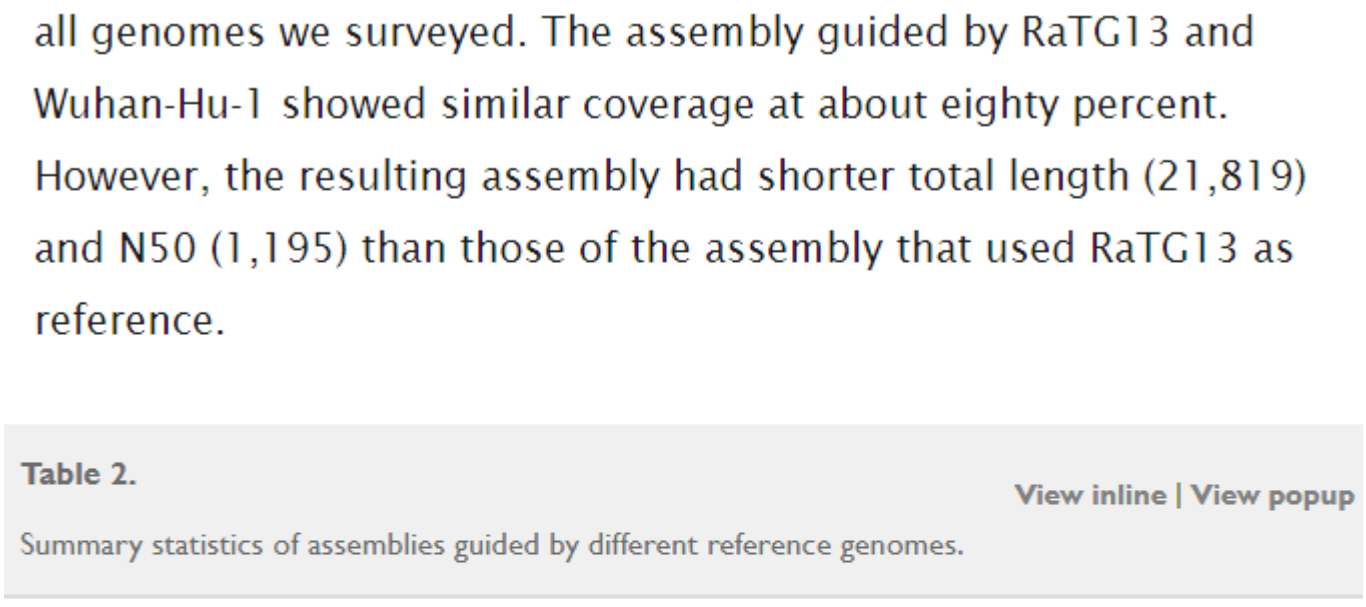
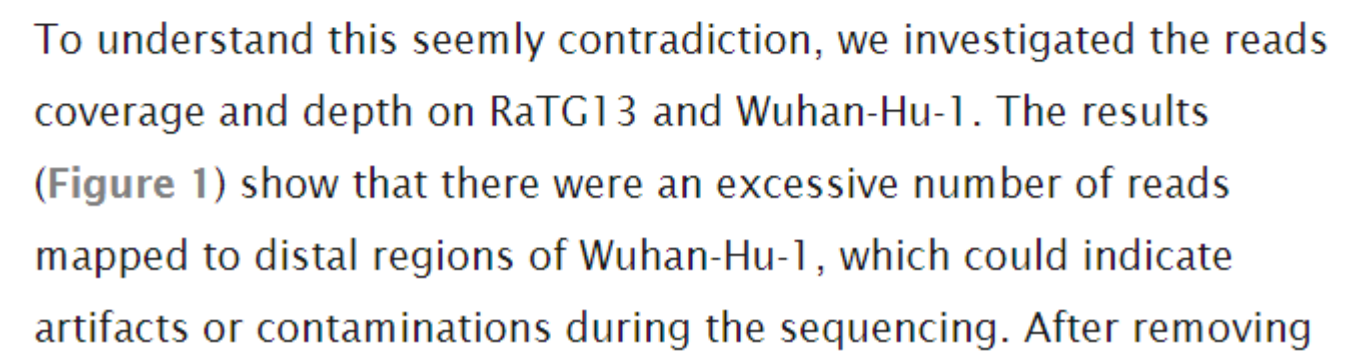

all genomes we surveyed. The assembly guided by RaTG13 and Wuhan-Hu-1 showed similar coverage at about eighty percent. However, the resulting assembly had shorter total length (21,819) and N50 (1,195) than those of the assembly that used RaTG13 as reference.

**Table 2.**

View inline | View popup

Summary statistics of assemblies guided by different reference genomes.

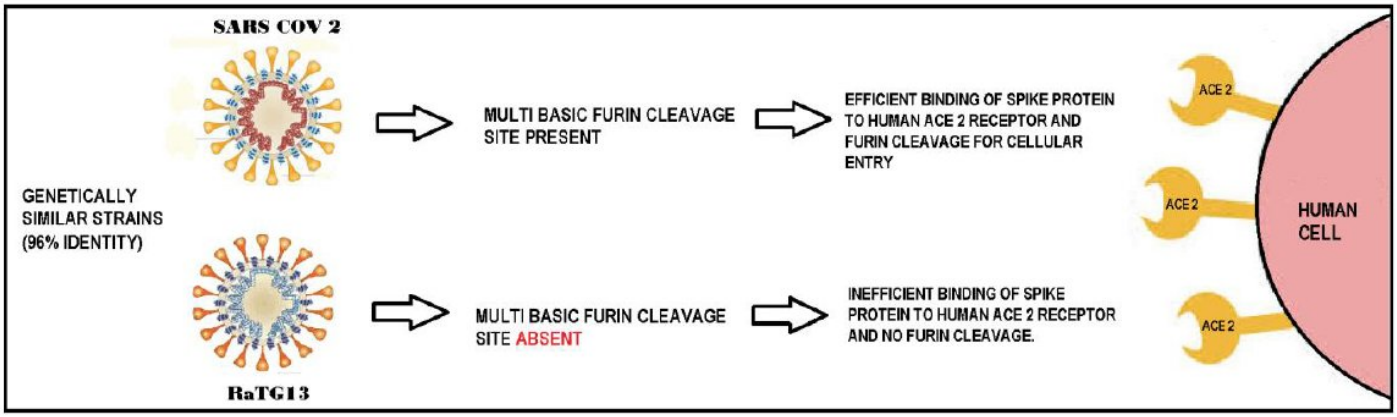

To understand this seemly contradiction, we investigated the reads coverage and depth on RaTG13 and Wuhan-Hu-1. The results (Figure 1) show that there were an excessive number of reads mapped to distal regions of Wuhan-Hu-1, which could indicate artifacts or contaminations during the sequencing. After removing

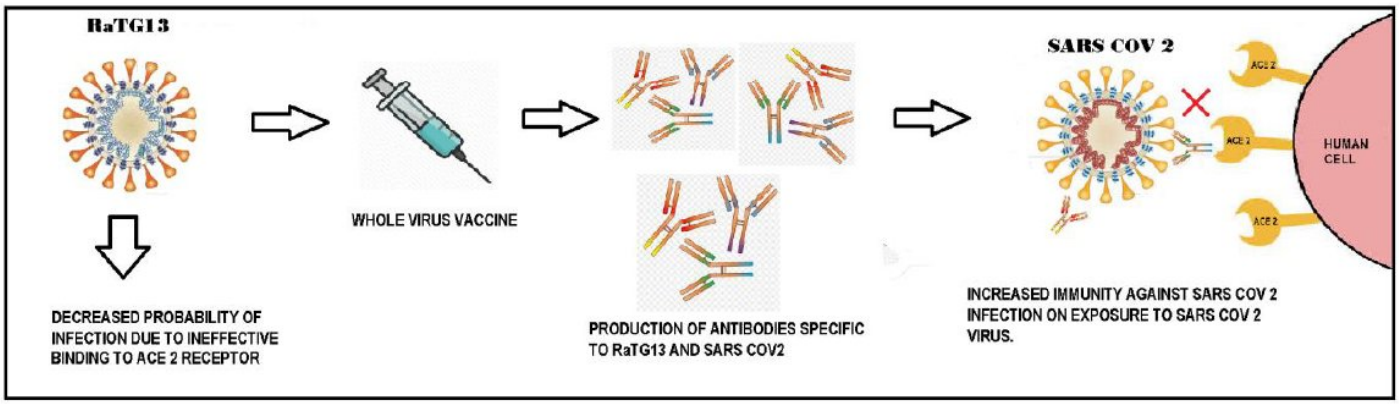
30. A further example can be found here:
SARS-CoV-2 & bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects
https://t.co/ilMmjCLjzu
They were indeed forced to take extreme measures to obtain viable RatG13 for their experiments!

## Comparison with the S glycoprotein from the bat virus RaTG13

Next, we determined the cryo-EM structures of S from the closest known bat virus (RaTG13) and of uncleaved SARS-CoV-2 S (Extended Data Fig. 1b). The bat virus protein was expressed in mammalian cells, but was found to be unstable during preparation of EM grids and required chemical cross-linking to produce particles for data collection and analysis. The

31. In fact @Rossana38510044 discovered RaTG13 has the dubious honour of being proposed as the basis for a vaccine against SARS-COV-2!
Due to its inability to infect anything more than Nature Papers crafted by Zhengli Shi and her desperate lab members!
https://t.co/30aqZqDb7i




32. "The mutation between SARS-CoV-2 & RaTG13 is unique across coronavirus species"
Comparative genomic analysis revealed specific mutation pattern between SARS-CoV-2 & Bat-SARSr-CoV RaTG13
https://t.co/F2fHoGWMfN
Note the bizarre mention of "5-FU" in the conclusion.
Anyone?

Compared with Bat-SARSr-CoV RaTG13, there were 293 nucleotideswere altered while only 29 amino acids were altered in SARS-CoV-2; their ratio was as high as 9.07. However, this ratio detected in the comparisons of other human coronaviruses with their similar animal coronaviruses was less than 5.0. For example, the ratio of human SARS-CoV Tor2 to bat SARS-like CoVs LYRa11 was 4.81; that ratio of human MERS-CoV EMC2012 to bat SARS-like CoVs RSA2011 was 2.80; that ratio of human coronavirus 229E to bat coronavirus HKU2 was 4.55. These results indicate that the relative level of synonymous substitutions between human SARS-CoV-2 and its possible animal origin (RaTG13) is much higher than that between other human coronaviruses and their potential animal sources.

33. Above reference to 5-FU is "5-fluorouracil":
a drug used for lethal mutagenesis, a broad-spectrum antiviral strategy that increases viral mutation rates & burdens viral populations with mutations reducing infectious progeny
https://t.co/hWedAMiMH2
and
https://t.co/I8iqsFIAsE

**TABLE 2** Mutation types found in USUV populations treated with different drugs

| Drug[a] | No. of mutations in each population | No. of nucleotides sequenced in each population | No. of mutations by type[b] | | | | | Transition frequency ($10 \times ^{-4}$)[c] | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | A to G | U to C | G to A | C to U | Tv | A to G | U to C | G to A | C to U |
| No drug | 8 | 45,337 | 2 | 1 | 2 | 0 | 3 | **1.8** | 1.0 | 1.6 | <0.9 |
| DEC | 8 | 33,947 | 1 | 2 | 0 | 2 | 3 | 1.2 | **2.6** | <1.1 | 2.4 |
| FAV | 20 | 34,218 | 3 | 5 | 5 | 6 | 1 | 3.5 | 6.4 | 5.2 | **7.2** |
| RBV | 11 | 25,897 | 0 | 2 | 4 | 4 | 1 | <1.5 | 3.4 | 5.5 | **6.4** |
| FU | 37 | 34,453 | 10 | 15 | 1 | 3 | 8 | 11.5 | **19.2** | 1.0 | 3.6 |

[a]USUV populations were serially passaged five times in the absence of drug (no drug) or in the presence of either decitabine (DEC), favipiravir (FAV), ribavirin (RBV), or 5-fluorouracil (FU) at a concentration of 800 $\mu$M.
[b]Different type of mutations found in the analysis. Given are the number of times that each transition type is found in the analysis. Tv indicates the number of transversions found in the analysis.
[c]Transition frequencies found in populations treated with mutagenic drugs. These numbers have been normalized to the nucleotide composition in the sequenced amplicon. Highlighted in bold is the most frequent transition in each sample.

34. Assymetric Mutations leave Researchers Puzzled ■
Excessive G–U transversions in novel allele variants in SARS-CoV-2 genomes (RaTG13)
https://t.co/hQgsE7aDVs

The second remarkable feature is that this excess of mutations is asymmetric: there is no similar effect for C–A mutations. SARS-CoV-2 is a positive (+) RNA strand virus. The copying of positive and negative strands of coronavirus RNA is executed by the same enzymes (Sola et al., 2015). If RNA copying was prone to G–U errors when creating the positive strand, the same mechanism would be expected to introduce G–U errors when copying the negative strand, resulting in additional C–A errors on the positive strand. Note that in SARS-CoV-2 the G (19.6%) and C (18.4%) content are similar, as are A (29.9%) and U (32.1%) content. Recently it was suggested that SARS-CoV-2 mutation rates could be affected by variations in its RNA-dependent RNA polymerase (Pachetti et al., 2020), but it is unclear how this could explain the asymmetric increase of G–U transversions.

35. The Puzzling Mutation Bias

(underlying divergence between RaTG13 & SARS-CoV-2)

Mutation Patterns of Human SARS-CoV-2 and Bat RaTG13 Coronavirus Genomes Are Strongly Biased Towards C>U Transitions

https://t.co/AelHag0Z12

**4. Discussion**

It is well established that mutation biases for C>U, U>C, A>G, and G>A transitions are the most frequent SNVs in eukaryotic and prokaryotic genomes [16,17,19]. The bias towards these types of mutations is, however, exceptionally high between the SARS-CoV-2 and related bat RATG13 genomes since more than 80% of the mutation spectrum can be attributed to these four substitutions despite them representing just one-third of all possible substitutions. The remaining eight types of substitutions (all transversions) represented about 17% of all SNVs. The skewed mutation rates towards transitions are consistent with highly conserved pyrimidine to purine ratios despite the genome divergence. Together, the RNA genomes of bat RATG13 and SARS-CoV-2 exhibit a strong transition–transversion bias typical for recently diverging virus strains [21].

When comparing bat RaTG13 and SARS-CoV-2, the C>U transitions outnumber those of reverse transitions almost two-fold. This unexpected nonequilibrium may point to a higher probability of C mutation into U in SARS-CoV-2 than U into C. This bias may potentially explain the relatively lower C nucleotide content of SARS-CoV-2 compared to the bat RaTG13 coronavirus. The reference genome (MN908947) to which the rest of the SARS-CoV-2 genomes were compared is assumed to originate from one of the first cases detected in Wuhan province in China, and may be closest to the SARS-CoV-2 ancestor. Most SARS-CoV-2 isolates from different geographical localities exhibited C>U substitutions when compared to the Wuhan-Hu-1. Moreover, phylogenetically divergent S and L' lineages seem to harbor a higher proportion of C>U and U>C substitutions than the L lineage which contained the Wuhan-Hu-1 reference genome. The observations suggest that a trend towards the loss of Cs may be ongoing during the pandemic.

36. Inconsistencies in RaTG13 Data noted by many!
An Update on the Origin of SARS-CoV-2: Though Closest Identity, Bat (RaTG13) and Pangolin Derived Coronaviruses Varied in the Critical Binding Site and O-Linked Glycan Residues
https://t.co/SEZIrPJjM4

Wuhan seafood market, Huang C, et al. reported a total of 41 patients, and 14 cases are not related to the seafood market and no trace of bats has been found, so exact place of origin need to be studied in detail [27]. Subsequently, Zhou P. et al. from Wuhan institute of virology (Zheng Li Shi lab) showed that SARS-CoV-2 was highly similar throughout the genome to RaTG13 with an overall genome sequence identity of 96.2% and 93.1% nucleotide identity to S protein. Also, the author did not mention when it has been sequenced and RNA dependent RNA polymerase (RdRp) data not shown to compare SARS-CoV-2 [9]. RaTG13 was isolated from the bat (*Rhinolophus affinis*) on 24 July 2013 by Zheng Li Shi group and the reason unclear why they did not submit the sequence before instead on 27 Jan 2020, although it is proximal to bat-SL-CoV (accession number: AVP78042.1, AVP78031.1 and ACU31051.1) (**Supplement-3**). SARS-CoV (Rs806/2006) (accession number: ACU31051.1) already has proven for Intraspecies diversity and its implications for the origin of SARS coronaviruses in humans [28]. Hence, the detailed investigation needed for RaTG13 isolate and origin. Scientists report genetic sequences of viruses

37. Confusion leads to a bizarre hypothesis
Synonymous mutations & molecular evolution of SARS-Cov-2 origins
https://t.co/JrNOJzeob7
"Another possibility is that RaTG13 has been maintained under conditions which have allowed it to continue to evolve after its initial sampling."

three ingroup strains. The largest value is observed for RaTG13 (1.0366), despite this sequence being the most early sampled sequence, perhaps caused by additional undetected recombination into RaTG13. Another possibility is that RaTG13 has been maintained under conditions which have allowed it to continue to evolve after its initial sampling.

38. Reactions?
Recently, Dr. Yan published 2 "reports" on SARS-COV-2, alleging it is a CCP bioweapon.
Whether or not you agree, her discussion of RaTG13 was accurate and clear, but was attacked in a "review":
https://t.co/XN9Q1BmUfo
See the image for their pathetic "arguments"

## Specific Comments on the Report

### Page 2

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

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



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

39. Conclusion?
RatG13 is not what it seems
unroll @threadreaderapp


40. Update regarding specific claims made by @K_G_Andersen and @Vehuardo here:
https://t.co/tAXMNTwdhA
answered by @flavinkins

1. The so-called “amplicons” were published long after the problems with the ILLUMINA were exposed, so no assembly can be made for RaTG13 using the SRA

41. These amplicons have strange things hanging off the ends and are artefacts from gene synthesis assembly errors
2.
They have a publishing date almost a month after the problems with the raw data were found in April.

42.
3. They were made up in May, synthesised on 2 plates, 1 for forward reads, 1 for reverse reads. They then had a problem with S1 & had to quickly make the last 4 amplicons
The result was these amplicons had random, non-vector & non-Biological sequences hanging off the ends

43.
4. Also, the real problem is that the “fecal swab” is anything but a “fecal swab”. You can look with your own eyes and perform a search on “bat feces” on NCBI. Whenever bacteria drops below 2.5% in total cellular organisms, all RNA viral reads are gone.

44. via @flavinkins
5. Bacterial 16S rRNA is a lot tougher than Eukaryotic and viral RNA. When 16S is destroyed, viral RNA will already be obliterated and no assemblable sequences will be possible.

45.
6. In fact, it took them over 3 months to publish the amplicons—“gap filling PCR data”—the timeline does not match up at all for a real sequencing project—those “amplicons” appeared just the same time as you would expect for a gene synthesis company from order to shipment.

46. via @flavinkins
7. The RaTg13 Pipeline is completely inverted by date and time:
Assembly comes first and ILLUMINA comes 2 weeks after assembly, which will be easy as a point-to-point fabrication is easy.
Most rRNA was degraded beyond repair, no mRNA properly matches anything

47.
8. “mRNA” matches match the clone more than
Predicted RNA, & the “virus” had far higher quality than anything else. They could easily get fake data, for example, CoV2 data. Manually alter the base calls by alignment & change the basecall without changing the quality score.

48.
9. They could have merged this into a fake “sample”.
For example, the WGS of the bat, filtered coarsely to generate an exome.
Then after 3 months they published the “amplicons” which is suspicious as 3 months is the time for a gene synthesis company to deliver their order.

49. via @flavinkins
10. Here you can see the Bat and Viral Metagenome:

every dataset is mostly bacteria.
(lowest bacteria in all cellular organisms = 30%)
https://t.co/nADnfIUPzZ

50. Unroll @threadreaderapp

51. Update
https://t.co/uORg2OzuPf

> 103. John Signus on the Significance of RaTG13
> 21-10-2020 - 37p
> Anomalous Datasets Reveal Metagenomic Fabrication Pipeline That Further Questions Legitimacy of Ratg13 Genome & Associated Nature Paper
> Print https://t.co/RXC1mj5Uvb
> PDF https://t.co/Lekg0rlPN4
> also mentions RmYN02 pic.twitter.com/fOI3eXaFIr
>
> — Billy Bostickson \U0001f3f4\U0001f441&\U0001f441 \U0001f193 (@BillyBostickson) November 1, 2020