Twitter Thread by Bloom Lab

Bloom Lab
@jbloom_lab

I'm agnostic on its various hypotheses about mechanisms of origin of furin-cleavage sites, but the part of this paper that suggests furin-cleavage site might be present in two of these SARSr-CoVs as a minor variant is embarrassingly bad science that shouldn't be amplified. (1/n)

> Back to CoVs. In samples from two European bats, the authors found SARSr-CoVs that were just one mutation away from a functional FCS. But minor variants were sequenced from each that *had* a functional FCS already, just as is seen with some low path flu.
> 11/
>
> — Michael Worobey (@MichaelWorobey) December 16, 2021

Here is Table 1 of pre-print (https://t.co/7fQsreN77n) that shows data related to this claim. The mutations in question are present at 0.004% and 0.006%, corresponding to 6 or 7 Illumina reads out of >100,000 total. (2/n)

Table 1. Single nucleotide variants within two European *SrC* at spike position Thr670.

	Consensus sequence	A	C	G
BB99-04	A (%)	146506 (99.801)	8 (0.005)	159 (0.108)
	T (%)	19 (0.013)	187 (0.127)	13 (0.009)
	C (%)	7 (0.005)	147250 (99.858)	10 (0.007)
	G (%)	259 (0.176)	**6* (0.004)**	146817 (99.872)
	Total reads	146798	147459	147005
BB89-98	**Consensus sequence**	**A**	**C**	**A**
	A (%)	117032 (99.593)	1 (0.001)	116893 (99.745)
	T (%)	47 (0.0400)	38 (0.032)	13 (0.011)
	C (%)	10 (0.009)	117540 (99.961)	8 (0.007)
	G (%)	421 (0.358)	**7#* (0.006)**	278 (0.237)
	Total reads	117510	117586	117192

#One of the reads was not paired-end
*This singe nucleotide variant leads to a non-synonymous exchange generating a FCS motif in this virus

There are good deep-sequencing studies of CoVs, eg by @katrina_lythgoe (https://t.co/joX4kCqOEh), @LauringLab (https://t.co/TisAPvkCVp) & @KATarinambraun @tcfriedrich @trvrb @LouiseHMoncla (https://t.co/08EhfVAAL1). These studies find you can call variants to 2-3% frequency (3/n)

> Establishing reliable variant calling thresholds for clinical samples in which true variant frequencies are unknown ideally requires resequencing of multiple samples from RNA to test for concordance. Working within the constraints of small volumes of remnant RNA from laboratory testing, we resequenced 76 high-viral-load samples, of which 27 replicate pairs generated sufficient read numbers (>50,000 unique mapped reads) for reliable minor variant detection. iSNVs with <2% MAF were generally indistinguishable from noise, whereas those with ⩾3% MAF were highly concordant between replicates (Fig. 2A and fig. S2).

Table 1 of this pre-print is reporting "mutations" that introduce a furin-cleavage site at frequencies of 0.004% to 0.006%, which is orders of magnitude lower than what good studies have rigorously defined as a reasonable threshold (2-3%) to call mutations. (4/n)

As anybody who has ever analyzed viral or cancer deep sequencing knows, it's simply impossible to use standard Illumina sequencing to identify mutations at even 0.4% or 0.04%, let alone 0.004%. It's actually surprising when a given sequencing error isn't present at ~0.01%. (5/n)

So what this study should say is: we found some bat SARS-related CoVs that are just one or a few mutations away from having a furin cleavage site, but none of them actually have a furin cleavage site even as a minor variant. (6/n)

Above point also currently caveated, since pre-print only shows 10-codon chunk of BB99-04 & no sequence of BB89-98. This region of spike is subject to substantial alignment uncertainty, so withhold judgment until enough of the sequences released to support the alignments (7/n)

Nucleotide sequences obtained within this study were deposited in GenBank with accession numbers XXX to XXX.