

2/9/2022 3:14 PM

**Pre-2019 Antarctic specimens with SARS-CoV-2 sequences found in Hefei, China laboratories are likely contamination from local COVID-19 specimens from Jan 2020 and later**

**SARS-CoV-2 genome sequences identified in specimens collected in the Antarctica before 2019 have a three SNV signature also found uniquely in clinical COVID specimens from early 2020 from Hefei Province, China and are most likely simple clinical lab-research lab cross contamination**

Steven C Quay

The finding of SARS-CoV-2 genomic sequences in soil and related specimens collected in Antarctica from before 2019 as well as extensive human, monkey, and hamster mitochondrial genomes within the same specimens has generated speculation about the role this evidence may play in helping solve what is the origin of the COVID pandemic.<sup>1</sup> The finding in these sequences of three SNVs at C8782T, C18060T, and T28144C which have previously been called “ancestral” mutations because they are closer to RaTG13 and BANAL-52 has fueled speculation this is evidence SARS-CoV-2 was in a laboratory at the University of Technology and Science in Hefei, Anhui Province, China, before the outbreak.<sup>2</sup>

Here I examine a report<sup>3</sup> on early clinical COVID-19 specimens from January 10 to March 10, 2020, from a hospital associated with the same University of Technology and Science in Hefei, Anhui Province, China, that show a high incidence of the same three SNVs. Specifically, of the 263 SNVs found in 174 confirmed early patients with COVID-19, fully 86 (33%) have the same “ancestral” SNVs. The Text-Table below is from the Supplemental information from Reference 3:

Gene region	position	residue in reference	mutated residue	counts of mutated samples	synonymous or not	novel or not
ORF8	28144	T	C	43	Leu Ser	shared
1b	18060	C	T	28	synonymous	shared
1a	8782	C	T	15	synonymous	shared
ORF3a	25499	C	G	1	Pro Arg	novel
M	26894	C	T	1	synonymous	novel

I also note two SNVs in the clinical specimens, 25499 and 26894, that are adjacent to SNVs in the “Antarctica” specimens. These could also be common SNVs from the two specimen sources but with reading errors during sequencing. In speaking about the C18060T SNV in the clinical specimens, the authors state: “Through genomic surveillance, we identified a locally concentrated S-type SARS-CoV-2 Sub-lineage, a1 (C18060T), that was endemic to Hefei and the Yangtze River Delta region, thus providing important insights into the local development of COVID-19.” This paper also provides detailed epidemiology inquiries to note the first patients seen all had visited Wuhan before becoming sick.

In conclusion, given the appearance of the three SNV mutations that define both the Antarctic specimens and the local spread of clinical COVID uniquely in the same university system in the same city in China, the most probable explanation for the appearance of SARS-CoV-2 in specimens collected before 2019 is contamination during processing post January 2020. A definitive answer awaits learning the timing of the sample processing of the Antarctic specimens.

<sup>1</sup> [Unique SARS-CoV-2 variant found in public sequence data of Antarctic soil samples collected in 2018-2019 ; Host genomes for the unique SARS-CoV-2 variant leaked into Antarctic soil metagenomic sequencing data](#)

<sup>2</sup> [Bloom lab Tweet thread Re the papers](#)

<sup>3</sup> [Identification of SARS-CoV-2 Variants and Their Clinical Significance in Hefei, China](#)