

**The Illusion of Biosafety During SARS-CoV-2 Research:
Multiple Apparent Occult Lab-Acquired Infections Are Identified
Under BSL-3 Conditions at a Premier US-based Coronavirus Laboratory**

Steven E Massey¹ and Steven Quay, MD, PhD²

ABSTRACT

An active debate exists over the use of synthetic biology and other advanced research tools on dangerous pathogens. Virologists doing gain-of-function and related research on dangerous pathogens, including creating synthetic chimeric infectious clones, believe their work is essential to preventing the next pandemic. Many scientists in related fields do not believe the benefit of the research outweighs the risk of a laboratory-acquired infection, leading to community spread. They also believe that current regulations and guidelines for the funding, conduct, and biosafety reporting of research accidents, that is, infections of laboratory personnel, is inadequate.

The consensus of the virologists' position is that, if creating synthetic pathogens is conducted under appropriate Biological Safety Laboratory (BSL) standards, the work can be performed safely. However, abundant evidence indicates that laboratory-acquired infections (LAI) still do occur, even under the highest BSL-3 and even BSL-4 standards.

Here we develop methods and criteria to identify occult LAIs and distinguish them from community-acquired infections. We then apply these tools to a test case.

Using these methods, we identify eight apparent LAI SARS-CoV-2 infections from May 2020 to January 2021, sequenced at the Clinical Molecular Microbiology Laboratory, University of North Carolina (UNC) Hospital, Chapel Hill, NC. While the laboratory from which they were acquired cannot be known with certainty, using the criteria herein, including the response to our inquiry and genome sequence comparison, all of the LAIs have a high probability of being SARS-CoV-2 variants being actively studied at premier coronavirus laboratories on the University of North Carolina Campus (UNC), ostensibly under BSL-3 conditions.

One of the clearest cases of a laboratory-acquired infection involves a 60-year-old woman with COVID-19 whose specimen was collected on May 18, 2020, at the UNC Hospital. Three features establish it as an LAI:

1. Being early in the pandemic, it was the only sample collected on May 18, 2020, in the entire state of North Carolina. Community infections in North Carolina were not clinically relevant at that time.
2. It had the C18060T SNV which was characteristic of the first strains in America (specifically, virus hCoV-19/USA/WA-CDC-WA1/2020, collected in January 2020, in Seattle, WA. All variants studied in UNC laboratories had this SNV.

¹ Biology Dept, University of Puerto Rico - Rio Piedras, San Juan, Puerto Rico, USA; [ORCID](#)

² [ORCID](#); Seattle, WA USA. Corresponding author: Steven@DrQuay.com

The Illusion of Biosafety During SARS-CoV-2 Research
S.E. Massey and S.C Quay – July 23, 2025

3. It had the C23615G SNV, which yields the Spike Protein amino acid substitution, R685G, that disrupts the furin cleavage site. This SNV was introduced in a publication including the Ralph Baric laboratory at UNC for the purpose of creating an attenuated vaccine (Zost, et al., 2020). It does not occur in community infections. Specifically, only five viruses with this SNV are found out of 18,589,989 GISAID sequences and four of these are from the UNC Hospital. The probability of this being a chance observation is near zero, specifically 10^{-33} .

An additional interesting feature of this virus is that it is an intermediate between Lineage A and Lineage B, having the T28144C SNV but not the C8782T SNV that define Lineage A. These rare intermediate genomes have been used previously to challenge the two-spillover hypothesis for SARS-CoV-2 (Massey, et. al., 2023). This case is a new, additional example of an intermediate sequence.

We could find no public records of reported LAIs from the UNC during this period and conclude it is likely these LAIs were unknown to the laboratory itself as knowingly failing to report infections under these circumstances would be a violation of a number of statutes and regulations.

The finding of eight likely occult LAIs in a nine-month period of time from what many consider the premier coronavirus synthetic biology laboratory in the US, or even the world, combined with the apparent failure to identify and report these LAIs by the laboratory or university at large, underscores a failure of current LAI regulations. Although current regulations have mandatory reporting, they do not have a process for finding and reporting occult infections.

Laboratory constructed synthetic viruses from UNC were watermarked with SNV T15102C/A and it was established by examining all GISAID sequence entries from North Carolina for the period of January 1, 2020, to January 31, 2021, a total of 1958 cases, that no sequenced human case had that SNV. This demonstrates the usefulness of such watermarks in synthetic coronavirus biology in determining the attribution, or lack thereof, for a community outbreak in the vicinity of the laboratory. **We suggest such a process be considered as a mandatory step for all research involving significant human pathogens.**

Given the high probability that the COVID-19 pandemic began with one or more scientists at the Wuhan Institute of Virology who became infected during synthetic virus engineering, and this report of undetected laboratory-acquired infections from synthetic clones and laboratory variants at the premier US coronavirus laboratories, it behooves us to pause all such research and develop robust biosafety standards, protocols, and regulations that can meet the heightened infectivity that synthetic viruses can achieve before resuming such research.

Before uploading the first version of this pre-print, the authors contacted the UNC sequencing lab personnel who submitted the suspected LAI cases to GISAID but received no response. Shortly after, the authors were contacted by GISAID's Washington DC office, relaying complaints from unnamed officials at the CDC and UNC about the preprint. GISAID threatened to revoke the authors' access unless metadata and sequence files were removed.

The Illusion of Biosafety During SARS-CoV-2 Research
S.E. Massey and S.C Quay – July 23, 2025

This pressure campaign to suppress inquiry, coupled with a lack of transparency from institutional actors, supports the hypothesis that these infections were potentially acquired in a laboratory setting.

Finally, this work highlights a regulatory conundrum: identifying occult laboratory-acquired infections that arise from laboratories conducting dangerous research requires the willingness for self-governance of the institutions in which the research is being conducted. As such, it is not amenable to an easy third-party or governmental oversight if the closed loop of accountability cannot be pierced.

The Illusion of Biosafety During SARS-CoV-2 Research **S.E. Massey and S.C Quay – July 23, 2025**

INTRODUCTION

A time-tested method of identifying a laboratory-acquired infection (LAI) is to show that a patient has an infection with a virus that is not found in the community at the time of the infection but is under investigation in the laboratory. If both the laboratory and the governmental healthcare officials conducting the investigation do so with openness and transparency, the resolution of a particular case can be done with dispatch.

For example, in November 2021, a young, female lab worker at a high-biosecurity facility in Taipei contracted COVID despite there being no other confirmed local cases at the time, raising suspicions of a lab leak (Silver 2022). The sequence of the virus was then found to match a SARS-CoV-2 Delta variant contained in the lab, rather than the local strains of the virus previously in circulation in the community. This was deemed the first reported lab leak of the COVID-19 virus.

We mapped her travels outside the lab between exposure to the virus, infection, and eventual diagnosis, approximately three weeks later (Figure 1). This is the type of analysis China should have conducted following the likely Wuhan Institute of Virology LAIs.

Date	Event
Mid-Nov	"Exposed to the pathogen" in mid-November while working at the Academia Sinica's Institute of Biomedical Sciences (IBMS), a P3 (Biosafety Level-3) facility located in Taipei's Nangang District.
26-Nov	Slight fever
27-Nov	She got on the Blue Line (Bannan Line) at Nangang Station
27-Nov	She got off at Dongmen Station on the Red Line (Tamsui-Xinyi Line)
28-Nov	She dined at Tokiya restaurant in iFG Farglory Square in New Taipei City's Xizhi District.
28-Nov	She got on the Blue Line at Nangang Station
28-Nov	She got off at Taipei Main Station
28-Nov	She visited the Q Square shopping center in the Taipei Main Station
28-Nov	She shopped in a PUMA shoe store in the Taipei Main Station
28-Nov	Betty's women's clothing boutique in Taipei Main Station
28-Nov	She rode the Blue Line back to Nangang Station.
29-Nov	Work at Academia Sinica
30-Nov	Work at Academia Sinica
1-Dec	Worked at AS. She went to the Sun Tung Pao (孫東寶) steak restaurant on Nanchang Street in New Taipei City's Xizhi District.
2-Dec	Work at AS
3-Dec	Work at AS. She went to the 3 Coins restaurant.
4-Dec	Cough intensified
4-Dec	She took Blue Line from Nangang Station & transferred to Red Line to ride to Dongmen Station.
4-Dec	That evening, she dined in the food court of the Global Mall Nangang Station Store.
5-Dec	She took Blue line from Nangang Station, transferred to the Orange Line (Zhonghe-Xinlu Line) to Guting Station.
5-Dec	She then went to a Watsons near Exit 5 of Guting Station
5-Dec	She dined at the Moonromantic Taipei Curry House.
6-Dec	Work at AS.
7-Dec	Work at AS
7-Dec	She went to the 7-Eleven on Zhongxing Road in Xizhi District.
7-Dec	She took the No. 306 bus from Academic Sinica to the Nangang District Office, from 3:39 p.m. to 3:50 p.m.
7-Dec	She took the Blue 21 bus from Nangang District Office to Zhongxin Road from 4:03 p.m. to 4:19 p.m.
8-Dec	Abnormality with her sense of smell and taste on Wednesday
8-Dec	She went to the Cosmed on Fude 1st Road in Xizhi.
8-Dec	PCR test conducted; positive
Unknown	She had also been studying Japanese at a Tamkang University center in Taipei City; the bustling Yongkang Street area near the university building

Figure 1 A diary of the movements of the infected lab worker in Taipei, November / December 2021

The Illusion of Biosafety During SARS-CoV-2 Research
S.E. Massey and S.C Quay – July 23, 2025

None of the 110 contacts later identified as exposed to the worker tested positive for the virus (Silver 2022).

The facility was fined \$150,000 New Taiwan dollars (*ca.* \$5400 US) for the incident, apparently the world's first documented infection with the pandemic coronavirus in a research lab. The purpose of this work is to examine if it is possible to identify a laboratory-acquired infection if neither the laboratory nor governmental healthcare officials conduct an investigation.

Here, we call these cryptic or occult laboratory-acquired infections.

MATERIALS AND METHODS

All genome sequences were found on GISAID (Khare, S., et al., 2021).

In this report Patient 1 is GISAID EPI_ISL_877723, Patient 2 is GISAID EPI_ISL_884293, Patient 3 is GISAID EPI_ISL_877616, Patient 4 is GISAID EPI_ISL_877686, Patient 5 is GISAID EPI_ISL_968173, Patient 6 is GISAID EPI_ISL_1443308, Patient 7 is GISAID EPI_ISL_1163793, and Patient 8 is GISAID EPI_ISL_968142. The GISAID data for these cases can be accessed via a bespoke DOI linkage found at the end of this pre-print.

Table 1 - Below are laboratories at the University of North Carolina that conduct research on the SARS-CoV-2 virus and the type of work they conducted. In many cases the work they conducted was in collaboration with the Baric laboratory, which is the anchor laboratory with respect to coronavirus research at UNC.

Laboratory	Focus Area	Reference
Baric Lab (Ralph Baric) - UNC Gillings School of Global Public Health	Remdesivir, SARS-CoV-2 pathogenesis, vaccine development, D614G variant	https://www.science.org/doi/pdf/10.1126/scitranslmed.abb5883
Garcia Laboratory (UNC School of Medicine)	Acute SARS-CoV-2 infection, innate immune response, therapeutic interventions	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7523135/
UNC-Chapel Hill Institute for Global Health & Infectious Diseases	Clinical trials, epidemiological studies, seroprevalence studies on frontline workers	https://www.cdc.gov/mmwr/volumes/69/wr/mm6935e2.htm
UNC COVID-19 Testing Program & Virology Laboratory	High-throughput SARS-CoV-2 testing, asymptomatic carrier detection	https://www.cdc.gov/mmwr/volumes/69/wr/mm6946e1.htm
UNC Department of Microbiology & Immunology	SARS-CoV-2 mutations, evolution, high-density amplicon sequencing	https://www.cell.com/cell-reports/fulltext/S2211-1247(20)31341-3
UNC-Chapel Hill BSL-3 Laboratory for SARS-CoV-2 Studies	BSL-3 laboratory research, live virus experiments, animal models	https://www.nature.com/articles/s41586-020-2708-8
UNC Department of Pathology & Laboratory Medicine	Serological responses, SARS-CoV-2 diagnostic test evaluation	https://journals.asm.org/doi/pdf/10.1128/mbio.02426-20
North Carolina Policy Collaboratory (at UNC-Chapel Hill)	Policy-driven COVID-19 research, tracking SARS-CoV-2 spread in North Carolina	https://www.pnas.org/doi/pdf/10.1073/pnas.2001046117

The Illusion of Biosafety During SARS-CoV-2 Research
S.E. Massey and S.C Quay – July 23, 2025

Table 2 contains a partial list of SARS-CoV-2 viruses and variants that were in UNC laboratories during the 2020-2021 timeframe and about which publications have been found. The NT substitutions shown in bold were found in the human sequences.

GenBank	Virus Name	NT Substitutions
MN985325.1	hCoV-19/USA/WA1/2020	C8782T, C18060T, T28144C
N.A.	D614G mutant of hCoV-19/USA/WA1/2020	C8782T, C18060T, A23403G, T28144C
MT461669.1	icSARS-CoV-2-WT derived from Severe acute respiratory syndrome coronavirus 2 SARS-CoV-2/human/USA/USA-WA1/2020	C8782T, T15102C, C18060T, T28144C
MT844088.1	Mouse-adapted Mutant Severe acute respiratory syndrome coronavirus 2 clone SARS-CoV-2-MA,	T15102A, C18060T , C23054T, A23056C, C23057A, C23059G, T28144C

Sequence sources

SARS-Cov-2 sequences were obtained from the Global Initiative on Sharing All Influenza Data (GISAID), with the exception of the first North Carolina sequence ('WA1'), which was obtained from the NCBI (Accession MT325591). The sequence for RaTG13 (Zhou et al. 2020), was likewise obtained from the NCBI (Accession MN99653).

According to GISAID, the sequencing of the seven genome sequences was conducted on an Ion Torrent machine, using the Ampliseq protocol. The sequences were submitted by the Dittmer Lab at UNC. The Dittmer group generated the 61 early cases shown in Figure 4, calling positions with a sequencing depth < 3 as an 'N', using CLC Genomics Workbench version 11.0 (McNamara et al. 2020). This indicates that while sequencing depth was not reported on GISAID, the large majority of each of the seven genome sequences had a read depth > 3, indicating that the sequences are unlikely to be artefactual.

Haplotype network analysis

Haplotype network analysis was conducted by aligning sequences of interest using Muscle (Applied Research Press 2015), followed by network construction and visualization in PopArt (Leigh and Bryant 2015) using the TCS method (Templeton, Crandall, and Sing 1992)(Clement, Posada, and Crandall 2000). Sequences with > 5 % ambiguous bases were removed.

Criteria for distinguishing a community-acquired infection from a laboratory-acquired infection

For purposes of distinguishing between a community-acquired infection (CAI) from a laboratory-acquired infection (LAI), the following Table 3 contains the criteria that will be used in this study.

The Illusion of Biosafety During SARS-CoV-2 Research
S.E. Massey and S.C Quay – July 23, 2025

Laboratory-Acquired Infections	Community Acquired Infections
If cooperative, an inquiry of an occult LAI becomes acknowledged as a true LAI. Alternatively, there is no response from the sequencing laboratory. There can even be attempts to silence or censor those making the inquiry. This kind of reaction, called the 'consciousness of guilt' response, is itself independent evidence of an LAI.	An inquiry of the testing laboratory leads to rapid cooperation. If there is an unforeseen explanation for the cases and sequences, it is offered fully and transparently. A suspected LAI can become a confirmed community-acquired infection by this process
Molecular Clock of the suspect virus is significantly slowed or even stopped compared to the molecular clock of community cases collected on the same day	Molecular Clock of suspect virus shows a similar number to SNVs from patients from the community for the date of specimen collection
The phylogeny of the suspect virus is anomalous, arising from a long period into the past	The phylogeny of the suspect virus fits within the community cases around the date of collection
Mutational sweeps that confer strong adaption, like the D614G mutation in SARS2, can be absent in the suspect virus	Mutational sweeps, like D614G, that reach >99% incidence in the community, are typically found in the suspect virus
The SNVs within the suspect virus are related to research work being done in a laboratory or publications near to the patients collection medical facility	Related research is not being done near the suspect virus collection or it is wholly unrelated to the suspect virus

Table 3 Criteria for distinguishing between Laboratory-Acquired Infections and Community-Acquired Infections

RESULTS

Test 1: The molecular clock in the seven suspected LAI-infections has recorded a significantly reduced number of SNVs, given the timing of the sample collection

SARS-CoV-2, like many RNA viruses, accumulate mutations or SNVs during transmission in a pandemic. A suspect community-acquired infection should have a similar number of SNVs as expected, given the date of sample, whether simply calculated based on estimates of the molecular clock from worldwide-evolutionary models or looking at local infections in the community where the suspect infection occurred.

A laboratory-acquired infection is likely to be similar to variants in ongoing experiments, which are often examining scientific questions with early, archetypal strains that can have broad interest within the research community. Consequently, their genome sequences may appear ‘frozen,’ with few differences from early strains, which are used in research or biotechnology (Massey, 2024).

Model-based clock comparison

Examining four papers that estimate the molecular clock for SARS-CoV-2 indicates a range of SNVs per year of 29.9 to 33.5, a mean of 33.6 and a standard deviation of 1.7 (McLean et al. 2022)(Bar-On et al. 2020)(Abbasian et al. 2023)(Chaurasia and Ghose 2024). Table 4 shows the expected SNVs for the seven suspect infections, given their date of collection.

Patient ID	Collection Date	Expected SNVs	Observed SNVs
1	8/24/2020	21.7	3
2	8/26/2020	21.8	0
3	6/9/2020	14.7	1
4	8/22/2020	21.6	2
5	9/1/2020	22.4	6
6	1/12/2021	34.7	4
7	1/27/2021	36.1	2
8	5/18/2020	12.7	3

Table 4 Expected number of SNVs in the eight suspect infections given their date of collection

These results are a significant deviation from what would be expected if these were community-acquired infections.

Geotemporal clock comparison

To examine for an unusual or unexpected local community molecular clock, all infections from the state of North Carolina on the same day as the suspect patient were examined. The results of that are shown in Table 5.

The Illusion of Biosafety During SARS-CoV-2 Research
S.E. Massey and S.C Quay – July 23, 2025

Patient	SNVs	Community SNVs on Collection Date (mean, SD, n)	Patient Compared to Confidence Interval
1	3	15.8, 8.1, 32	Less than the 99.9999% CI
2	0	12.9, 3.3, 12	Less than the 99.9999% CI
3	1	9, NA, 2	9>1
4	2	15.9, 5.8, 12	Less than the 99.9999% CI
5	6	16, NA, 2	16>6
6	4	20.7, 5.8, 39	Less than the 99.9999% CI
7	2	20.9, 6.1, 83	Less than the 99.9999% CI
8	3	[Patient 8 is the only case]	N.D.

Table 5 Comparison of numbers of SNVs between the seven suspect infections compared to North Carolina genome sequences generated on the same date of collection

As can be seen, five of the patients had their specimen collected on a day when there were between 12 and 83 other cases from North Carolina, making a statistical analysis possible. In each case, the suspect case had a number of SNVs that was below the lower limit of the 99.9999% Confidence Interval. These results are a significant deviation from what would be expected from a series of community-infections.

Patient 8 is unique in another way: being a case in May 2020, it is the only case in the entire state of North Carolina that day.

Both methods of calculating the SNV number for these seven suspect infections indicate these are a significant deviation from what would be expected from community-acquired infections and consistent with what would be expected from a laboratory-acquired infection.

Test 2: Phylogeny of laboratory-acquired infections appears frozen in time

Epidemics can be characterized by progressive lineages that arise, sweep across the globe to achieve near universal prevalence, only to be replaced by the next lineage, in a repeating cycle. Many studies document that these competitive fitness demonstrations are the selective pressure for whatever phenotype provides the greatest evolutionary advantage.

A community-acquired infection should be seen as being part of the wave of evolution that is present at the time of sample collection. Because laboratory strains or variants are not being studied under the high selection pressure of a clinical infection, and also because experiments are usually performed starting with deep frozen aliquots of a variant, to purposely keep genetic consistency during a multi-experiment study, a laboratory-acquired infection would be expected to represent the phylogeny when the experiments were started, even if months or even years before the lab accident that led to an infection.

The following Figures illustrate that principle for these suspected infections, which contains the Pango Variant designation of the 146 cases in the state of North Carolina in July 2020.

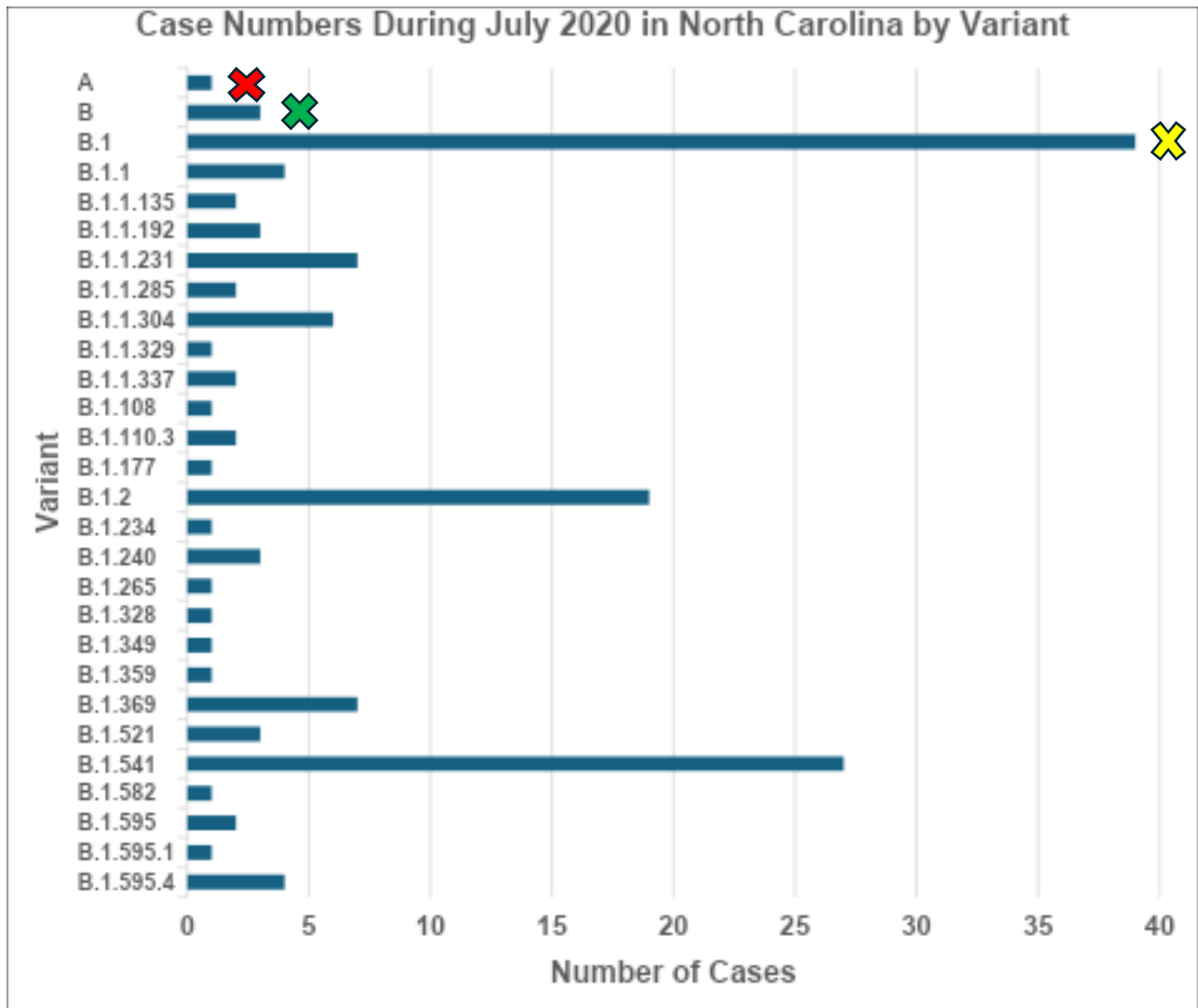


Figure 2 Case numbers during July 2020 in North Carolina by variant

The red X signifies the three suspect infections that are lacking the D614G and have the two ancestral SNVs of what was called Lineage A, the earliest human infections. Please note that these three cases were collected in September 2020, and January 2021, and so are even more out of touch with the community variants than this figure would suggest. The green X signifies the two suspect cases that are Lineage B, having acquired the two mutations away from the ancestral strain but have not acquired the D614G mutation. The yellow X signifies Lineage B variants, with the addition of the D614G mutation.

To show the frozen nature of these cases, the following chart shows the same data for March 2020, the first month when cases in North Carolina could be tabulated. As expected, March cases with the A, B, and B.1 variants would be much more likely to be found in the community. The rapid development of many more B.1 variants between March and July is also noted.

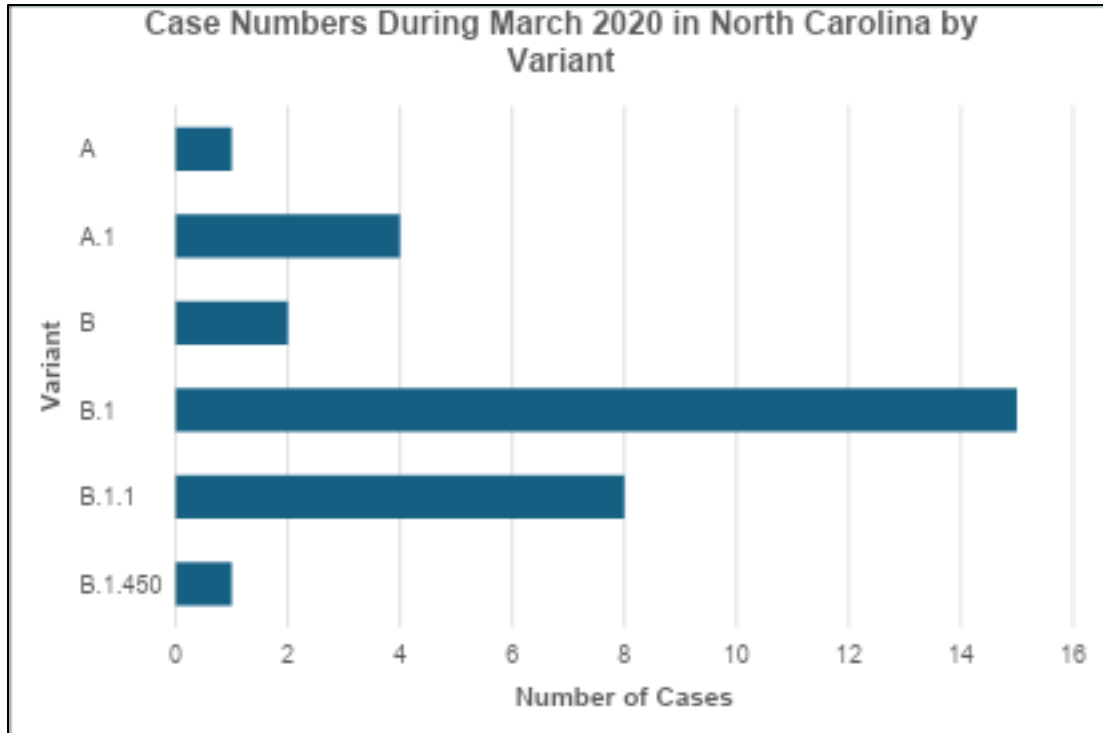


Figure 3 Case numbers during March 2020 in North Carolina by variant

A haplotype network analysis was conducted of the 7 suspect genomes, combined with 61 genome sequences generated by the UNC from early cases in North Carolina (Figure 4). The 61 genome sequences were sampled from 18th March to 14th May 2020. The network shows that the 7 genomes cluster close to the node that connects to the outgroup RaTG13 and so can be considered basal (the thick black line connecting to RaTG13 indicates the large numbers of nucleotide differences with SARS-CoV-2 genomes). This is consistent with their ‘frozen’ nature and indicates unusual genome sequence stasis, given that they were sampled after the 61 early genomes, from 6th June 2020 to 27th January 2021.

The Illusion of Biosafety During SARS-CoV-2 Research S.E. Massey and S.C Quay – July 23, 2025

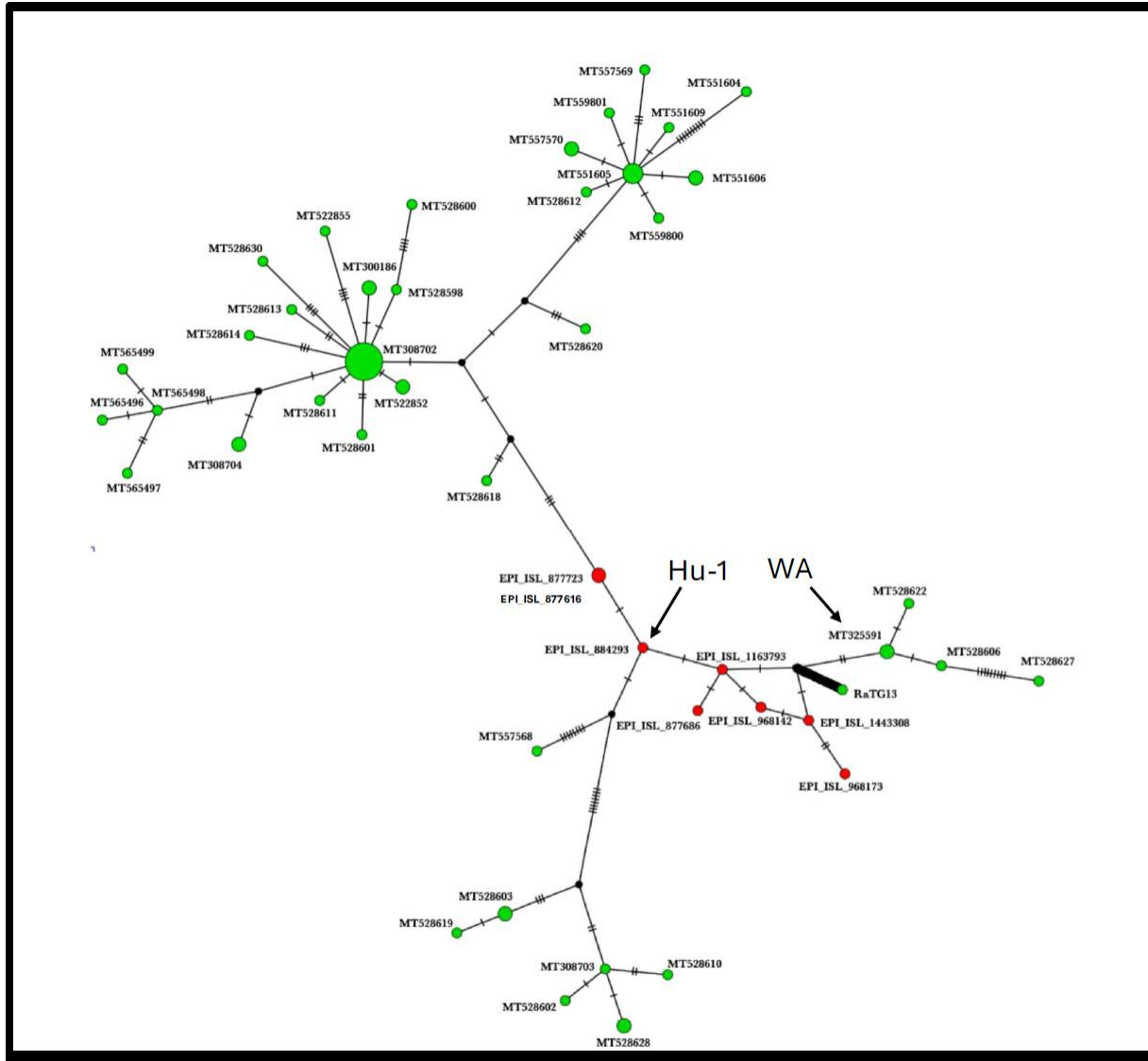


Figure 4 Haplotype network of the eight suspect infections (red) and 61 early sequences generated by UNC (McNamara et al. 2020). ‘Hu-1’ indicates the SARS-Cov-2 reference sequence, which is identical to EPI_ISL_884293 (Patient 2). ‘WA’ represents the first sequenced genome from North Carolina, sampled on the 6th March 2020 from a person who had returned from Washington State. RaTG13 is included as an outgroup, its thick edge is indicative of the large number of SNVs that separate it from SARS-CoV-2.

When the first suspect sequence from Patient 1 is compared to genomes from North Carolina sequenced the same week (Figure 5), it can be seen that the Patient 1 sequence (red) is closest to the node connected to the outgroup RaTG13, indicating its ‘frozen’ nature. Only two SNVs separate the Patient 1 sequence from the RaTG13 connecting node, while 8 SNVs separate the next closest sequence (EPI_ISL_1334533) from the RaTG13 connecting node.

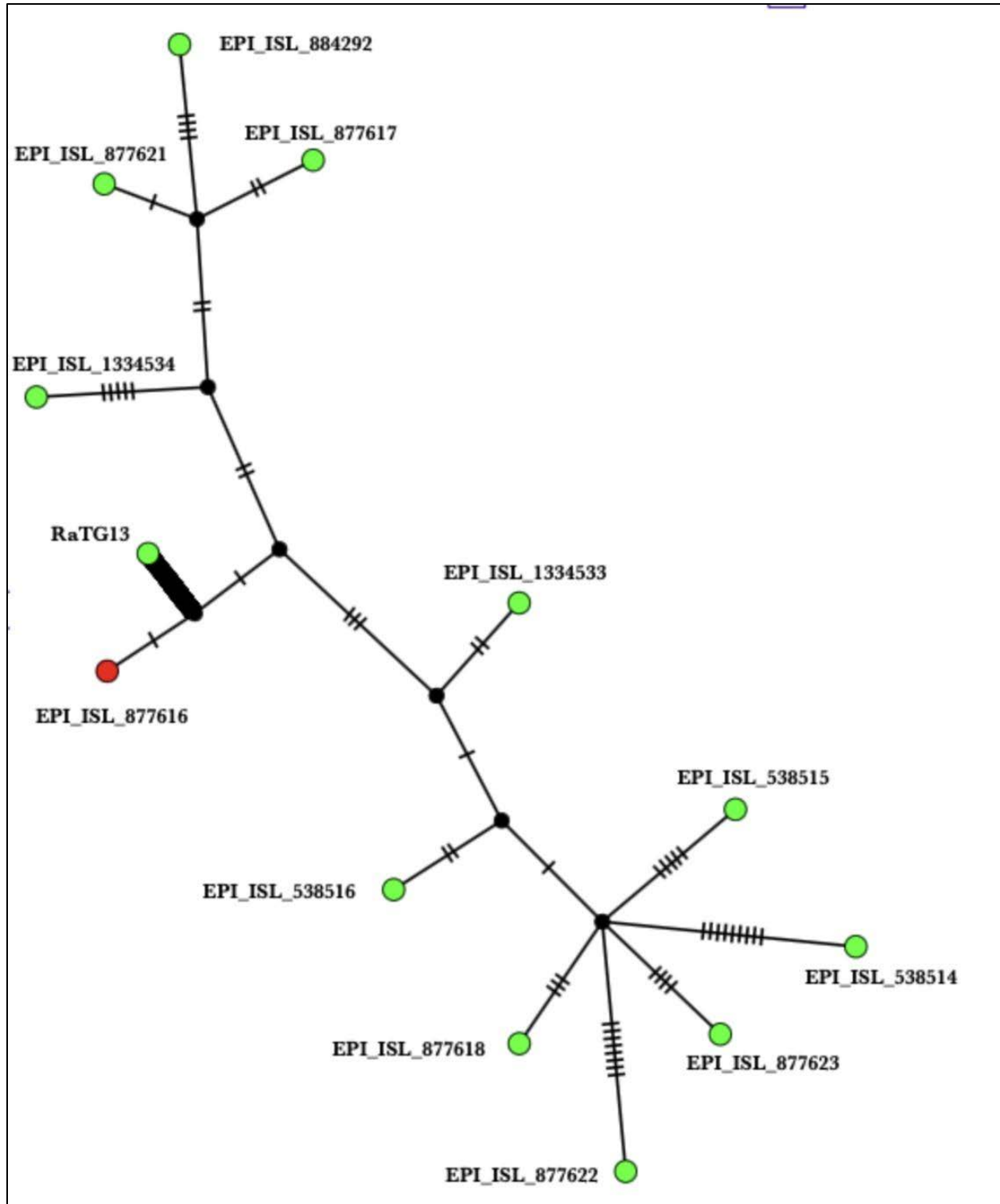


Figure 5 Haplotype network of sequence from Patient 1 (red) compared to sequences generated the following week (24 -31 Aug 2020) from North Carolina. RaTG13 is provided as an outgroup.

The Illusion of Biosafety During SARS-CoV-2 Research **S.E. Massey and S.C Quay – July 23, 2025**

Likewise, when the sequence from Patient 7 is compared to sequences from North Carolina generated the same day (27th Jan 2021), the Patient 7 sequence (red) is close to the RaTG13 connecting node, separated by 3. In contrast, the next closest sequence to the RaTG13 connecting node (EPI_ISL_1031889) is separated by 20 SNVs. This emphasizes the ‘frozen’ nature of the Patient 7 sequence.

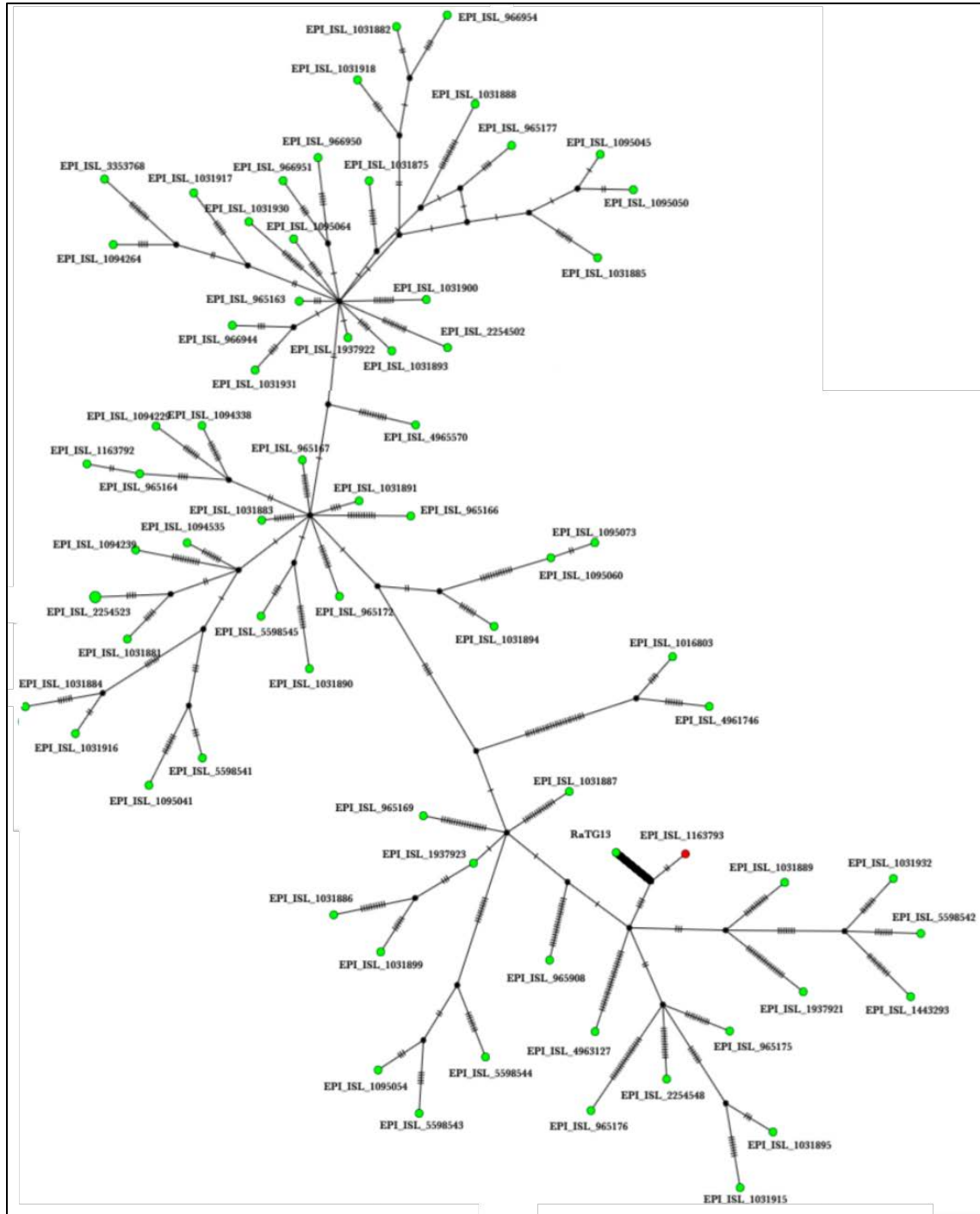


Figure 6 Haplotype analysis of sequence from Patient 7 (red) compared to sequences generated on the same day (27 Jan 2021) from North Carolina. RaTG13 is provided as an outgroup.

Test 3: The presence or absence of significant gene sweeps in the suspect infections

Gene sweeps are typically seen in a new virus outbreak, like SARS-CoV-2, when a particular SNV provides such a substantial selection advantage that, within a short period of time, all new cases include the SNV.

The Spike Protein mutation D614G, deriving from SNV A23403G, is a classic example of such a gene sweep. Prior to March 1, 2020, it was found in 10% of global sequences; between March 1 and March 31, 2020, it represented 67% of sequences; and between April 1 and May 18, 2020 it represented 78% of sequences (McNamara et al. 2020; Korber et al. 2020). The frequency of the D614G mutation and the frequency of the original D614 mutation within the GISAID database can be observed in Table 6. Depending on the collection date, between only 0.5% and 2.3% of all viruses sequenced worldwide have retained the original D614 nt. The probability of finding these five, rare events in the same hospital, if they are independent of each other, is approximately 1 in 3×10^{11} . If, on the other hand, they are all from the same laboratory then they are not independent and might even be expected, if that laboratory was studying the effect of adding the D614G mutation to a background strain that did not contain the SNV.

The cases restricted to North Carolina on the day of collection were also assessed. For four collection days, the suspect case here was the only ancestral D614 variant collected that day. We note that there are two additional cases with the D614 variant. While these cases are suspicious for a laboratory-acquired infection we have chosen not to include them because it is less clear they are a laboratory-acquired infection

Patient ID	Collection Date	AA Substitutions	NT Substitutions	D614 or G614 on Collection day in GISAID		North Carolina Cases Only	
				Total Cases	D614 Variants; %	Total Cases	D614 Variants
2	8/26/2020	None	None	1838	13; 0.7%	13	1
4	8/22/2020	NS8 L84S	C18060T, C22993T	825	6; 0.7%	12	2
5	9/1/2020	Spike R685G, NS6 E59stop, NS8 L84S	C8782T, A11332T, C18060T, C23615G, G27376T, T28144C	1768	40; 2.3%	2	1
6	1/12/2021	Spike R685G, NS8 L84S	C8782T, C18060T, C23615G, T28144C	9317	52; 0.6%	39	2
7	1/27/2021	NS8 L84S	C18060T, T28144C	10,034	49; 0.5%	83	1
8	5/18/2020	Spike R685G, NS8 L84S	C18060T, C23615G, T28144C	1,456	158; 10.9%	1	1

Table 6 Frequency of D614 or G614 on the collection dates of five of the suspect infections, all of which possess D614.

This analysis of the absence of the universal gene sweep variant, D614G, in five of these cases is a highly significant deviation from the expected outcome if these were a community acquired infection. It is consistent with a laboratory-acquired infection, especially if that laboratory was studying the effects on infectivity of adding the D614G mutation.

Test 4: Coronavirus research focus at the University of North Carolina (UNC)

Coronavirus research at UNC has been a focus of effort for more than two decades. The leading scientist there, Ralph Baric, has over 500 publications on the NCBI website.³ With the SARS-

³ <https://www.ncbi.nlm.nih.gov/myncbi/ralph.baric.1/bibliography/public/>

The Illusion of Biosafety During SARS-CoV-2 Research
S.E. Massey and S.C Quay – July 23, 2025

CoV-2 pandemic, the research output has increased significantly. For purposes of this work, we will focus on the Baric labs work on studying the role of the D614G mutation on SARS-CoV-2 transmission. Two papers contain extensive work of G614 and D614 variants (Meganck et al. 2024)(Hou et al. 2020).

During this research the D614G was put into the WA1 variant of SARS-CoV-2, the first case in North America with the virus genome sequenced.

In addition, Baric has published on the utility of the ablation of the furin cleavage site via a R685G SNV in the design of sarbecovirus vaccines (Lee J., et al., 2024). This mutation greatly reduces SARS-CoV-2 fitness and therefore would be considered a strong signal of a laboratory construct if found in a patient specimen. In fact, filtering the 15,839,989 virus sequences in GISAID in which low coverage is excluded, there are only five viruses with the R685G mutation. Four of these were from patients who presented to the University of North Carolina Hospital (Patient 5, 6, 8, and one additional patient).

Test 5: The response from the sequencing laboratory and others to an inquiry concerning laboratory-acquired infections

On March 23, 2025, the following email was sent to the GISAID submitters, Drs. Dirk Dittmer and Mellisa Miller, at the Clinical Molecular Microbiology Lab, UNC Hospital.

Comments requested
1 message

Steven Quay, MD, PhD <steven@drquay.com>
To:
Cc:
Bcc:

Sun, Mar 23, 2025 at 1:49 PM

Dear Drs. Dittmer & Miller-

I hope this email finds you well.

Dr. Massey and I have prepared an early draft of a paper in which we believe we have identified up to eight laboratory acquired SARS-CoV-2 infections among the cases that your lab sequenced in 2020 and early 2021 and submitted to GISAID. I am writing this to the two of you as you are both identified as corresponding authors for the excellent paper you wrote about early SARS-CoV-2 cases from North Carolina and so I believe you both are in a position to opine on our manuscript.

Can you please comment on this manuscript and especially, if you were aware of these as laboratory acquired infections at the time or subsequently. If you were aware, were these reported to the NIH, UNC Biosafety Committees, or the CDC? If not, can you comment on their occult nature?

We will be placing this manuscript, when completed, on a preprint server in the very near future and so if we can have a prompt response, hopefully by Wednesday, March 26, it would be appreciated. Can you also give us permission to include your response to this inquiry in the final manuscript.

--
Regards, Steve

Steven Quay, MD, PhD, FCAP

The authors have received no direct response to this email.

On March 25, 2025, the corresponding author had a 12-minute telephone call initiated by the staff of the Washington, DC, office of GISAID, the German database from which the authors retrieved the patient data for this study. The purpose of the call was to inform the corresponding author of

The Illusion of Biosafety During SARS-CoV-2 Research
S.E. Massey and S.C Quay – July 23, 2025

complaints to GISAID from the Centers for Disease Control (CDC), in Atlanta, GA and individuals of the University of North Carolina about this pre-print. According to GISAID, failure to respond fully would be grounds for terminating the GISAID accounts of these authors.

The names of the individuals from these institutions that requested this pre-print be withdrawn were not provided.

During the call the corresponding author agreed to make changes to the preprint, including removing GISAID metadata and sequence FASTA files and providing the same information in a bespoke hyperlink.

The nexus of the CDC and a research laboratory-acquired infection is noted in the following Table, which contains a list of the entities to whom a report of a laboratory-acquired infection must be made, the timing, and other circumstances.

	Entity	When to Report	Why
Federal	NIH Office of Science Policy (OSP), if NIH-funded	Within 30 days	Required for NIH-funded work
	CDC Division of Select Agents and Toxins	Gain of function research leading to a LAI	HHS P3CO (Potential Pandemic Pathogen Care and Oversight) policy.
University of North Carolina	Institutional Biosafety Committee (IBC)	Immediately	Required by NIH Guidelines
	Institutional Environmental Health & Safety (EHS) Office / Biosafety Officer	Immediately	Manages lab safety and compliance
	Occupational Health Services	Immediately	Medical management of exposure
Non-Fed Gov	Local/State Public Health Dept.	As required	Legal disease reporting obligations

On March 26, 2025, the corresponding author wrote the following email to Dr. Associate Vice Chancellor Kemp concerning the apparent occult LAIs at UNC.

The Illusion of Biosafety During SARS-CoV-2 Research
S.E. Massey and S.C Quay – July 23, 2025

Our preprint on apparent occult laboratory acquired infections (LAI) at UNC
1 message

Steven Quay, MD, PhD <steven@drquay.com>

Wed, Mar 26, 2025 at 10:32 PM

Cc: Steven E Massey <stevenemassey@gmail.com>, "Steven Quay, MD, PhD" <steven@drquay.com>

Dear Associate Vice Chancellor Kemp-

I hope this email finds you well.

I have become aware of your email to Edward Hammond of March 26 concerning his inquiry about reports of laboratory-acquired infections at UNC via the X social media site. His inquiry may be related to our pre-print claiming seven apparent LAIs during 2020 and early 2021 at the UNC hospital system.

Before we uploaded our pre-print we did send the manuscript to Drs. Dirk Dittmer and Melissa Miller to get their comments. Frankly we hoped they would have an explanation that explained what looks like very anomalous sequences. We would not have uploaded the pre-print if we were in error.

For whatever reason, we never heard back from them.

Here is the link to our pre-print: <https://zenodo.org/records/15083349>

I hope we can receive a response from UNC investigators soon and that there is an understandable explanation. Our analysis is pretty simple and so we recommend an investigation of the sequences identified in the pre-print and look forward to a satisfactory resolution of this matter.

--

Regards, Steve

Steven Quay, MD, PhD, FCAP

The authors have received no response to this email.

Given the lack of cooperation by UNC and the affirmative efforts of UNC and CDC to contact GISAID instead of the authors of this preprint, the conclusion that these are laboratory-acquired infections cannot be dismissed.

Description of each of the seven-suspect laboratory-acquired infections and their likely providence

Patients 1, 2, and 3

Patient 1 was a 24-year-old male with a sample collection date of August 24, 2020. He was infected with a virus that could be described as the SARS-CoV-2 Reference Sequence, with the addition of the D614G mutation, that conferred a notable transmission advantage. The finding of a B Lineage virus with just a single SNV, the D614G change, is rare. For example, in the entire world, there are only 911 genomes out of 15,743,406 complete sequences in the GISAID database that have this genotype, a frequency of 0.006%.

It should also be noted that Patient 3 has a virus with the exact same sequence, collected on June 9, 2020, over two months earlier, allowing one to calculate the probability of these two rare sequences being in the same community and then showing up in the same hospital. The probability of two genomes with this rare sequence in the community is approximately one in 296 million.

Noting further, Patient 2 is infected with the National Center for Biotechnology Information (NCBI) Reference Sequence Hu-1 (NC_045512.2). because it does not contain the D614G change.

The Illusion of Biosafety During SARS-CoV-2 Research
S.E. Massey and S.C Quay – July 23, 2025

This is also a rare sequence in the GISAID database, primarily because it is so close to the start of the pandemic, having been the first genome uploaded to GenBank on January 5, 2020. There are only 3515 genomes in GISAID or 0.02% that have this sequence.

However, in the context of doing research on the effects of the D614G mutation on viral fitness and transmission, doing paired experiments with the Reference Standard sequence, plus or minus the single change, D614G, is the proper experiment to perform.

Based on the evidence presented herein, Patient 1, 2, and 3, had laboratory-acquired infections.

Patient 4

Patient 4 was a 21-year-old male with a sample collection date of August 22, 2020. He was infected with a virus with only two mutations from the Reference Sequence, C18060T and C22993T. It is thus an L Clade, Pango B Lineage. It is one of only two genomes in the 15,743,406 GISAID sequence database to be an L Clade, Pango B virus with the C18060T mutation. It is the only genome in the 15,743,406 GISAID sequence database to be an L Clade, Pango B virus with the C22993T mutation.

A possible reason for the rarity of the 18060 mutation in the Pango B lineage is that the first 18060 virus was in the ancestor to Pango B, the Pango A lineage. It was from the earliest sequence of a case in the United States, a 35-year-old man returning from Wuhan to Seattle, WA, and who got sick and had a specimen collected January 19, 2020. The virus was originally named USA-WA1/2020 by the University of Washington but both the FDA and the CDC would quickly sequence it and send their own named version of the virus, with the same sequence, to GISAID as virus hCoV-19/USA/WA-FDA-001/2020 and virus hCoV-19/USA/ WA-CDC-02982586-001/2020, respectively. Compared to the Reference Sequence, it had three SNVs: C8782T, C18060T, T28144C. This was one of the natural virus sequences the Baric lab would use to create a synthetic clone of SARS-CoV-2, which they named icSARS-CoV-2-WT.

Based on these considerations, Patient 4 appears to have a laboratory-acquired infection.

Patient 5, 6, 7, and 8

These four patients have the Pango Lineage A SNV of T28144C in common, and the other “required” Lineage A SNV, C8782T. Interestingly, Patient 7 has a reversion at 8782, making it one of the rare intermediate sequences. Because a widely cited paper suggested all intermediate genome sequences between Pango A and B were sequencing artefacts, (Massey et al. 2023) identified 14 intermediate sequences that appear to be inappropriately excluded. Patient 7 could be a 15th intermediate sequence.

Cases 6 and 7 appear to be directly connected, both because they are collected in January 2021 but they seem to be derived from the canonical WA1 US case, which has three SNVs relative to the Reference Sequence, C8782T, C18060T, T28144C. Patient 6 has one additional SNV and Patient 7 has lost C8782T but has gained nothing. The SNV in Patient 6 is rare, appearing in only 112 (0.0007%) genomes in the entire 15,743,406 total sequences in GISAID. This patient 6 SNV was not acquired in the community but may be private to this patient.

The Illusion of Biosafety During SARS-CoV-2 Research
S.E. Massey and S.C Quay – July 23, 2025

All four sequences share the WA1-related SNV, C18060T and are the closest of the seven sequences to ‘WA,’ the first North Carolina case, in Figure 4. A GISAID examination of all cases in North Carolina from May 1, 2020, to January 31, 2021, the timing of collection of these cases, identifies 11 cases with the C18060T SNV, out of a total of 1856 cases with the C18060 SNV. Thus, less than 1% of community infections during this time contained the C18060T mutation. On the other hand, the WA1 strain and derivatives thereof, were the workhorse of laboratory experiments being conducted during this time at the University of North Carolina.

Patients 5, 6, and 8 have the R685G mutation in the furin cleavage site of the Spike Protein that prevents furin cleavage. This was part of the vaccine design features proposed by Baric. The importance of the furin cleavage site for community infection is highlighted by the observation that only 5 sequences out of 15,838,989 had this mutation and four of these are from UNC Hospital. Because all four patients have collection dates that are at least four months apart, this is not a simple human-to-human cluster but instead are distinct infections.

Patient 5 has a sequence with a nonsense mutation, E59stop, in nonstructural protein 6 (nsp6) which truncates the protein at about 20% of its full length. While it has been shown that this protein is a transmembrane protein, it may not be essential for replication, and deletions may, in fact, be an adaptation to humans, there are no examples which we could find of such a severe deletion (Feng et al., 2023). This therefore may be a sequencing error. In support of this, while there are over 7,000 sequences in GISAID with the A11332T SNV, the vast majority are associated with notes indicating significant sequencing anomalies.

Based on these considerations, patients 6-7 appear to have laboratory-acquired infections. The genomic sequence in Patient 5 may have an error at position 11332.

Patient 8 is a unique case for two reasons: it has the rare R685G mutation that the Baric lab was studying. This mutation is one of only 5 out of 15,838,989 genomes in GISAID. This patient is also the only case that day in the entire state of North Carolina, since it was in May 2020, before the pandemic had taken off. This case appears to be an LAI.

Evidence of the absence of water-marked laboratory-created viruses in community infections in North Carolina

Both UNC laboratory created viruses, MT461669.1 and MT844088.1, have an inserted synonymous SNV, T15102C, as a “watermark” of laboratory creation. An examination of the GISAID sequences from January 1, 2020, to January 31, 2021, in North Carolina identified a total of 1958 sequences, none of which had a T15102C SNV. This is strong evidence that neither of these strains resulted in a laboratory-acquired infection.

This also suggests a method for any laboratory doing legitimate synthetic biology research on coronaviruses or any other virus, for that matter, to provide a ‘watermark’ on their synthetic strains and thus avoid being wrongly accused of being responsible for an outbreak in the vicinity of the laboratory.

Pango A Lineage cases were rare in the world in January 2021

The Illusion of Biosafety During SARS-CoV-2 Research
S.E. Massey and S.C Quay – July 23, 2025

In this study, two of the suspected LAIs from UNC were from patients with specimen collection dates in January 2021. They were Pango Lineage A and had two (C18060T, T28144C) and four (C8782T, C18060T, C23615G, T28144C) SNVs, respectively.

To address the hypothesis that sporadic ancestral cases were occurring throughout the world in early 2021, GISAID was interrogated for all Pango Lineage A cases from January 1, 2021, to January 31, 2021. Excluding low coverage cases, there were only 62 cases reported in the entire world, of a total of 241,464 sequenced cases that month or 0.03%.

When limiting the search to the United States there were only eight cases. Of the eight, three of them had over 3% uncalled bases and two of them were Pango A but had an additional 20 SNVs, as would be expected for a community virus. Of the remaining three cases, two were reported here from the UNC.

The third case was an 18-year-old male from a hospital in Bozeman, MT with an infection with four SNVs (C1812T, A6604G, C22311A, T28144C). Given the proximity of this individual to the Rocky Mountain Laboratory in Hamilton, MT, this might be a worthy potential case for a future LAI analysis.

DISCUSSION

Recent reports of potential laboratory-acquired infections (LAIs) of SARS-CoV-2 at the University of North Carolina (UNC) may seem shocking, but they are far from unprecedented. A deeper look into the biosafety history of UNC's high-containment laboratories reveals a concerning pattern of safety breaches involving lab-engineered coronaviruses, including strains closely related to both SARS and MERS. These earlier incidents provide critical context for understanding how repeated failures in biosafety management, some involving the same laboratories implicated in recent suspected LAIs, have gone largely unaddressed despite posing significant public health risks.

From January 2015 to June 2020, UNC-Chapel Hill reported 28 laboratory incidents involving genetically modified organisms to the NIH Office of Science Policy. Of those, at least six incidents involved engineered coronaviruses, specifically SARS- and MERS-associated strains, used in mouse infection models in their biosafety level 3 (BSL-3) laboratories. Notably, each of these incidents involved some level of potential exposure to lab personnel, who were subsequently placed under medical surveillance. In one of the most alarming cases, a researcher was bitten in February 2016 by a mouse infected with a synthetic SARS coronavirus. Despite the bite puncturing two layers of gloves and breaching skin, the scientist was not isolated, merely asked to report her temperature and symptoms while continuing to work and move through public spaces (Young and Blake, 2020).

This pattern of lenient post-incident procedures persisted. For instance, in April 2020, a researcher was bitten by a mouse carrying a mouse-adapted strain of SARS-CoV-2, triggering a 14-day home quarantine. While the response seemed more cautious, it also underscored how similar exposure events had previously led to far more relaxed responses. UNC officials consistently declined to provide key details about the viruses involved, the extent of the exposure, and the risk posed to the surrounding community, despite federal guidance recommending full disclosure for all incidents involving genetically modified pathogens.

These incidents also highlight systemic failures in risk assessment. Despite conducting experiments with high-consequence respiratory viruses, including synthetic variants designed to enhance infectivity, UNC relied heavily on the personal protective equipment of researchers and assumed containment within the lab. However, as demonstrated herein, viral sequence analysis of eight suspected SARS-CoV-2 LAIs originating from UNC shows that such containment may be illusory. This data suggests that laboratory strains, "frozen in time" compared to the community phylogeny, found their way into hospital patients without detection or official recognition, paralleling the earlier, documented breaches from 2015–2020.

The historical record undermines the often-repeated assurance that high-level biocontainment is sufficient to guarantee biosafety. If prior incidents involving SARS and MERS coronaviruses could result in direct exposure to UNC lab staff, without illness, but also without rigorous quarantine, then it is reasonable to infer that under slightly different circumstances, such exposures could have caused community outbreaks. The notion that biosafety is assured simply by virtue of

The Illusion of Biosafety During SARS-CoV-2 Research

S.E. Massey and S.C Quay – July 23, 2025

operating under BSL-3 protocols collapses when faced with the repeated real-world failures at even the most prestigious research institutions.

The question is no longer whether LAIs can occur in top-tier labs: they do happen and have happened. The real issue is whether the regulatory system is prepared to recognize, monitor, and mitigate them. The answer, judging from this pre-print, is no.

REFERENCES

- Abbasian, Mohammad Hadi, Mohammadamin Mahmanzar, Karim Rahimian, Bahar Mahdavi, Samaneh Tokhanbigli, Bahman Moradi, Mahsa Mollapour Sisakht, and Youping Deng. 2023. "Global Landscape of SARS-CoV-2 Mutations and Conserved Regions." *Journal of Translational Medicine* 21 (1): 1–15.
- Applied Research Press. 2015. *MUSCLE: A Multiple Sequence Alignment Method with Reduced Time and Space Complexity*. CreateSpace Independent Publishing Platform.
- Bar-On, Yinon M., Avi Flamholz, Rob Phillips, and Ron Milo. 2020. "Science Forum: SARS-CoV-2 (COVID-19) by the Numbers," March. <https://doi.org/10.7554/eLife.57309>.
- Chaurasia, Rajashree, and Udayan Ghose. 2024. "Genome-Wide Mutation Frequency Variation among SARS-CoV-2 Variants and Its Effects on the Untranslated Regions." *The Nucleus*, February, 1–18.
- Clement, M., D. Posada, and K. A. Crandall. 2000. "TCS: A Computer Program to Estimate Gene Genealogies." *Molecular Ecology* 9 (10): 1657–59.
- Feng S, O'Brien A, Chen D, Saeed M, Baker SC.2023. SARS-CoV-2 nonstructural protein 6 from Alpha to Omicron: evolution of a transmembrane protein. *mBio*14:e00688-23. <https://doi.org/10.1128/mbio.00688-23>
- Hou, Yixuan J., Kenichi Okuda, Caitlin E. Edwards, David R. Martinez, Takanori Asakura, Kenneth H. Dinno 3rd, Takafumi Kato, et al. 2020. "SARS-CoV-2 Reverse Genetics Reveals a Variable Infection Gradient in the Respiratory Tract." *Cell* 182 (2): 429–46.e14.
- Khare, S., et al (2021) GISAID's Role in Pandemic Response. *China CDC Weekly*, 3(49): 1049-1051. doi: 10.46234/ccdcw2021.255
- Korber, Bette, Will M. Fischer, Sandrasegaram Gnanakaran, Hyejin Yoon, James Theiler, Werner Abfalterer, Nick Hengartner, et al. 2020. "Tracking Changes in SARS-CoV-2 Spike: Evidence That D614G Increases Infectivity of the COVID-19 Virus." *Cell* 182 (4): 812–27.e19.
- Lee, J., Stewart, C., Schäfer, A. *et al.* A broadly generalizable stabilization strategy for sarbecovirus fusion machinery vaccines. *Nat Commun* **15**, 5496 (2024). <https://doi.org/10.1038/s41467-024-49656-5>
- Leigh, Jessica W., and David Bryant. 2015. "Popart: Full-Feature Software for Haplotype Network Construction." *Methods in Ecology and Evolution* 6 (9): 1110–16.
- McLean, Gary, Jeremy Kamil, Benhur Lee, Penny Moore, Thomas F. Schulz, Alexander Muik, Ugur Sahin, Özlem Türeci, and Shanti Pather. 2022. "The Impact of Evolving SARS-CoV-2 Mutations and Variants on COVID-19 Vaccines." *mBio* 13 (2): e0297921.
- Massey SE. The Growing Phenomenon of "Frozen" Virus Genome Sequences and Their Likely Origin in Research Facility Escapes. *Microorganisms*. 2024;12: 2412.
- Massey, Steven E., Adrian Jones, Daoyu Zhang, Yuri Deigin, and Steven C. Quay. 2023. "Unwarranted Exclusion of Intermediate Lineage A-B SARS-CoV-2 Genomes Is Inconsistent with the Two-Spillover Hypothesis of the Origin of COVID-19" *Microbiology Research* 14, no. 1: 448-453. <https://doi.org/10.3390/microbiolres14010033>
- McNamara, Ryan P., Carolina Caro-Vegas, Justin T. Landis, Razia Moorad, Linda J. Pluta, Anthony B. Eason, Cecilia Thompson, et al. 2020. "High-Density Amplicon Sequencing Identifies Community Spread and Ongoing Evolution of SARS-CoV-2 in the Southern United States." *Cell Reports* 33 (5): 108352.
- Meganck, Rita M., Caitlin E. Edwards, Michael L. Mallory, Rhianna E. Lee, Hong Dang, Alexis B. Bailey, Jason A. Wykoff, et al. 2024. "SARS-CoV-2 Variant of Concern Fitness and Adaptation in Primary Human Airway Epithelia." *Cell Reports* 43 (4): 114076.
- Silver, A. 2022. "Taiwan's Science Academy Fined for Biosafety Lapses after Lab Worker Contracts COVID-19." *Science*.

The Illusion of Biosafety During SARS-CoV-2 Research
S.E. Massey and S.C Quay – July 23, 2025

- Templeton, A. R., K. A. Crandall, and C. F. Sing. 1992. "A Cladistic Analysis of Phenotypic Associations with Haplotypes Inferred from Restriction Endonuclease Mapping and DNA Sequence Data. III. Cladogram Estimation." *Genetics* 132 (2): 619–33.
- Young, A. and Jessica Blake, "Here Are Six Accidents UNC Researchers Had With Lab-Created Coronaviruses," *ProPublica*, August 17, 2020, <https://www.propublica.org/article/here-are-six-accidents-unc-researchers-had-with-lab-created-coronaviruses>.
- Zhou, Peng, Xing-Lou Yang, Xian-Guang Wang, Ben Hu, Lei Zhang, Wei Zhang, Hao-Rui Si, et al. 2020. "Discovery of a Novel Coronavirus Associated with the Recent Pneumonia Outbreak in Humans and Its Potential Bat Origin." *bioRxiv*. <https://doi.org/10.1101/2020.01.22.914952>.
- Zost SJ, Gilchuk P, Case JB, Binshtein E, Chen RE, Nkolola JP, Schäfer A, Reidy JX, Trivette A, Nargi RS, Sutton RE, Suryadevara N, Martinez DR, Williamson LE, Chen EC, Jones T, Day S, Myers L, Hassan AO, Kafai NM, Winkler ES, Fox JM, Shrihari S, Mueller BK, Meiler J, Chandrashekar A, Mercado NB, Steinhardt JJ, Ren K, Loo YM, Kallewaard NL, McCune BT, Keeler SP, Holtzman MJ, Barouch DH, Gralinski LE, Baric RS, Thackray LB, Diamond MS, Carnahan RH, Crowe JE Jr. Potently neutralizing and protective human antibodies against SARS-CoV-2. *Nature*. 2020 Aug;584(7821):443-449. doi: 10.1038/s41586-020-2548-6. Epub 2020 Jul 15. PMID: 32668443; PMCID: PMC7584396.

Acknowledgements

We gratefully acknowledge all data contributors, i.e., the Authors and their Originating laboratories responsible for obtaining the specimens, and their Submitting laboratories for generating the genetic sequence and metadata and sharing via the GISAID Initiative, on which this research is based.

SUPPLEMENTAL TABLE

Data Availability:

GISAID Identifier: EPI_SET_250327so

DOI: <https://doi.org/10.55876/gis8.250327so>

All genome sequences and associated metadata in this dataset are published in GISAID's EpiCoV database. To view the contributors of each individual sequence with details such as accession number, Virus name, Collection date, Originating Lab and Submitting Lab and the list of Authors, visit [10.55876/gis8.250107tk](https://gis8.250107tk)

Data Snapshot

EPI_SET_250327so is composed of 132 individual genome sequences.

The collection dates range from 2013-07-24 to 2021-03-20;

Data were collected in 4 countries and territories.