Anomalous biases of reverse mutations in SARS-CoV-2 variants

Hideki Kakeya* and Tatsuya Kanzaki+

*Institute of Systems and Information Engineering, University of Tsukuba, Japan

+Graduate School of Science and Technology, University of Tsukuba, Japan

* Corresponding author: kake@iit.tsukuba.ac.jp

Abstract

Anomalous reverse mutation patterns in the Delta variants and the Omicron variants are investigated. A previous study on the Omicron BA.1, BA.1.1, and BA.2 has found almost all the variations of sequences containing only one reverse mutation and no other point mutations, suggesting lab origin of the Omicron variant. The histograms of mutations in the spike proteins of major variants are compared, where BA.1 is outstanding in its high reverse mutation rate and low variety of mutations. Among the reverse mutations of BA.1, H681P reversion while preserving K679 is least likely to emerge naturally either through homologous recombination, point mutation, or both of them put together when the sequence alignment and mutation spectrum are taken into account. The sequences including both K679 and P681 are registered by multiple submitters, which consolidates their existence. G614D reverse mutations, a reversion of the earliest and the most dominant point mutation in SARS-CoV-2, in the Delta variant are concentrated around the Great Lakes and G614D in the BA.2 lineage are concentrated on each side of the Hudson River. To elucidate the origin of these sequences that are improbable to emerge through community spread, inspections of laboratories are needed, where the geographical distribution of anomalous sequence detections can help to narrow down the potential sources.

Keywords

SARS-CoV-2, variant of concerns, reverse mutation, point mutation, homologous recombination

Introduction

Strong bias toward nonsynonymous (N) mutation over synonymous (S) is observed consistently in the spikes of SARS-CoV-2 variants of concern (VOCs). Among them, the Omicron variant, which includes 30 or more N mutations in the spike protein alone while including only one S mutation [1], is markedly different from the other VOC strains with around 10 spike mutations. Phylogenetic analysis shows that the Omicron variant did not emerge from the other precedent VOCs [2].

The independence of mutations among the VOCs has been discussed from various perspectives. Hassan et al. surveyed mutations in various VOCs across the continents to find that many mutations were specific to each location [3], which could have caused the emergence of independent mutations. Some discuss that the possible origins of the Omicron variant are either unknown human population under strong selective pressure to escape from vaccine-induced immune response, incubation in an immunocompromised patient, or evolution in a non-human host before spilling over back to human [4,5]. However, the highly vaccinated populations in advanced countries are well-monitored by the health administration, which means that evolution accompanying many mutations without being noticed is practically impossible. As for incubation in an immunocompromised patient, the count of mutations observed so far is around 10 or fewer [6-8], which is not comparable to that observed in the Omicron variant.

The dN/dS (Ka/Ks) ratio, which compares N mutations to S mutations, is the most well-known metric to measure the selective pressure [9,10]. Wei et al. argue that dN/dS as high as 6.64, which is observed in the spike protein of the Omicron variant, is extremely unlikely to emerge in an immunocompromised patient [11]. Mutation of virus can have a dN/dS value much higher than unity only when the virus spreads among multiple species [12]. Indeed, even the HIV-1 regulatory gene *tat*, which is known for its high selective pressure, has around 1.5 dN/dS ratio in the human population [13]. In SARS-CoV and SARS-CoV-2, dN/dS is usually smaller than unity [14].

Wei et al. insist that the Omicron variant has evolved in mice [11], which is followed by Zhang et al. [15]. It is known, however, that the original strain of SARS-CoV-2 does not infect mice [16]. Kakeya et al. indicate that a lab origin of the Omicron variant is likely [17], possibly caused by a spill-over from transgenic mice [18]. Arakawa suggests that other variants can also have lab origins considering the consistently high dN/dS ratio [19].

Tanaka et al. have found that the data of the Omicron variants BA.1, BA.1.1, and BA.2 registered in GenBank comprise sequences with only one reverse mutation and no other point mutations in the spike protein, where single reverse mutation can be found in almost all of the spike mutations [20], indicating the trace of reversion experiments to see the effect of each mutation. Kakeya et al. have found reversions in D614G of the spike proteins in the Delta variant and the Omicron variant BA.2 in the late stage of the community spread [21].

It is known that D614 is unstable in humans while G614 is competent in human-to-human transmission [22], which makes extinction of D614 inevitable. Indeed, D614G is known to be the first major mutation observed in the original Wuhan strain [22,23] and the only mutation shared by all the VOCs, which means that the emergence of D614 does not meet up with the expectation of natural mutation process.

Since the lab origin of SARS-CoV-2 and its VOCs can have great impact on life science in general, there have been huge pushbacks against lab leak theories. As for the original Wuhan strain, the US Congress recently revealed the information provided by a whistleblower from the CIA (Central Information Agency) that the agency offered officials who had assessed the COVID-19 origin as a Wuhan lab leak “significant monetary incentive” to change their positions to “unable to determine” the origins [24]. The whistleblower also says that Anthony Fauci, a former head of NIAID (National Institute of Allergy and Infectious Diseases) was escorted to the CIA Headquarters to “influence” its COVID-19 origins investigation without a record of entry [25].

Similar strong pushbacks have been made against the lab origin theory of the SARS-CoV-2 variants. Putting political pressures aside, there have been several counterarguments against it from a scientific point of view. Some insist the odd reverse mutation patterns can be generated through series of point mutations under strong immune pressures. Others insist that the sequence data registered in GenBank is not reliable enough, some of which may include data errors. Others insist that D614, which is considered unstable and unlikely to emerge naturally, can be stabilized with another mutation to survive natural selection.

In the present study, the authors introduce detailed analyses on the major variants of SARS-CoV-2 to answer the above counterarguments. The histograms of mutations in the spike proteins of major variants were generated to see whether the immune pressure could explain the spike mutation pattern. The affiliation and distribution of data submitters were checked to see whether data entry error could explain the reverse mutation pattern. Mutations co-occurring with G614D reverse mutation were checked to see whether emergence of D614 could be explained with a cofactor. Timing and location of D614 sampling were also analyzed to identify the origin of this mutation.

Methods

In the first analysis, the surface glycoprotein (spike protein) data of 13 lineages (B.1.1.7, B,1,351, P.1, B.1.617.2, C.37, B.1.621, BA.1, BA.2, BA.2.12.1, BA.4, BA.5, BQ.1, XBB.1.5) were downloaded from NCBI (National Center for Biotechnology Information) GenBank in June 2023. To save computational cost, protein sequences including deletion and insertion were removed from the analyses. All the point mutations including reverse mutations were counted and the ratio of reverse mutations to all the point mutations were calculated for each lineage. Among these lineages, the histogram of spike mutations for B.1.1.7, P.1, B.1.617.2, BA.1, BA.2, and XBB.1.5 were generated for comparison. The histograms of reverse mutations in the Omicron variants BA.1, BA.2, and XBB.1.5 were also generated.

In the second analysis, the detailed information including the submitter names, dates, and locations of collection of the revertant containing only one reverse mutation and no other mutations in spike protein listed by Tanaka et al. [20] were obtained to see the reliability of data. The registered sequences including reversion of P681H while preserving K679, which is least likely to emerge naturally, were also checked to confirm the existence of this peculiar mutation.

In the third analysis, mutations co-occurring with G614D reverse mutation were checked for B.1.617.2 and BA.2 lineages, where many G614D reversion was found in the previous study [21].

In the fourth analysis, reversions of D614G were counted in 18 lineages (AY.44, AY.103, B.1.1.519, B.1.427, B.1.429, B.1.526, B.1.617.2, B.1.637, BA.1.15, BA.2, BA.5.2.1, BA.5.5 BE.1, BE.3 CH.1.1, D.2, XBB.1, XBB.1.6). The data were collected from NCBI GenBank from the end of September to the beginning of October in 2023. These lineages were selected not to overlap with the previous study [21] except for B.1.617.2 and BA.2, which had been found to include high ratio of G614D revertant. Unlike the previous study [21], which counted only the first emergence of unique sequences that had no insertion or deletion, the whole data were searched to count all the reversions of D614G by picking up "YQDVN" from the amino acid sequences, where the timing and the location of collection were analyzed to identify the epicenter of this mutation.

Results

The ratios of reverse mutations to all the point mutations in the 13 lineages, including Alpha variant (B.1.1.7), Beta variant (B.1.351), Gamma variant (P.1), Delta variant (B.1.617.2), Lambda variant (C.37), Mu variant (B.1.621), and seven Omicron lineages, are shown in Figure 1. As this figure shows, BA.1 is outstanding in its high reverse mutation ratio, which is attained with the probability of 5.9×10^{-3} under a normal distribution based on the data of 13 lineages. The high ratio of BA.1 is outstanding when the comparison is limited within the Omicron lineages, which is attained with the probability of 2.8×10^{-2} under a normal distribution based on the data of seven Omicron lineages.

The histograms of spike mutations for B.1.1.7, P.1, B.1.617.2, BA.1, BA.2, and XBB.1.5 based on the whole data obtained from GenBank are shown in Figure 2. The histogram of BA.1 is apparently smoother than those of the other lineages. The locations of amino acids that have more than 0.1% mutation rate in for B.1.1.7, P.1, B.1.617.2, BA.1, BA.2, and XBB.1.5 is 89, 79, 86, 41, 72, and 80 respectively. The low count of 41 in BA.1 has a statistically significant deviation based on the normal distribution given by the data of these six lineages ($p = 0.027$).

The histograms of reverse mutations in the Omicron variants BA.1, BA.2, and XBB.1.5 are shown in Figure 3. BA.1 is unique in its high ratio of reverse mutations in all locations, while BA.2 and XBB.1.5 includes mutations other than reversions. At the 408th amino acid in BA.2 and XBB.1.5, mutation to Arginine is frequent, while mutation to Lysin is frequent at the 146th amino acid of XBB.1.5.

The detailed information on the revertant containing only one reverse mutation and no other mutations in spike protein found by Tanaka et al. [20] are listed in Supplemental Table 1. The variety of submitters is relatively large in BA.2, while the Center for Disease Control and Prevention (Howard, D., et al.) is the main submitter for all of the three lineages BA.1, BA1.1, and BA.2. What is significant is diverse locations of sampling soon after the first detection of each lineage, which is observed for revertant at K417N, N440K, and N856K in BA.1, revertant at 215ins in BA.1.1, and revertant at D408S in BA.2.

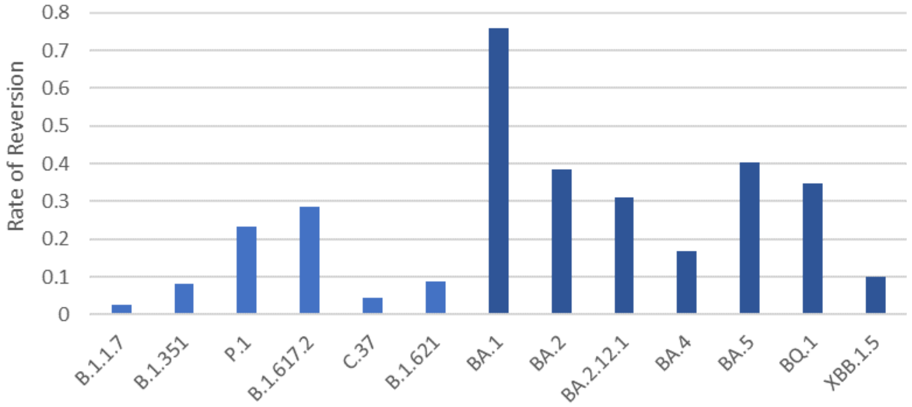


Figure 1. Rate of reversions among point mutations in 13 lineages.

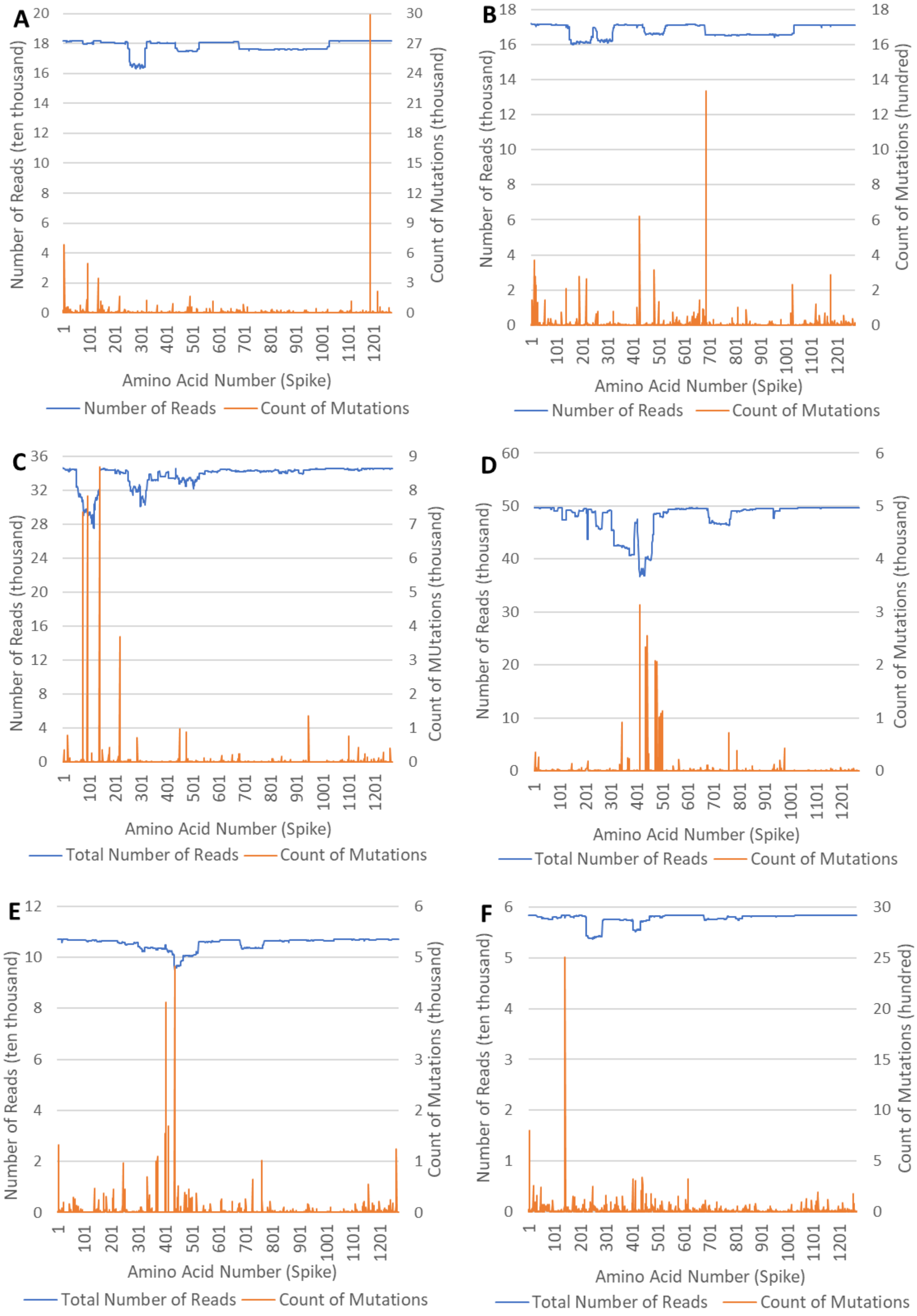


Figure 2. Histograms of spike mutations in six lineages: (A) B.1.1.7; (B) P.1; (C) B.1.617.2; (D) BA.1; (E) BA.2; (F) XBB.1.5.

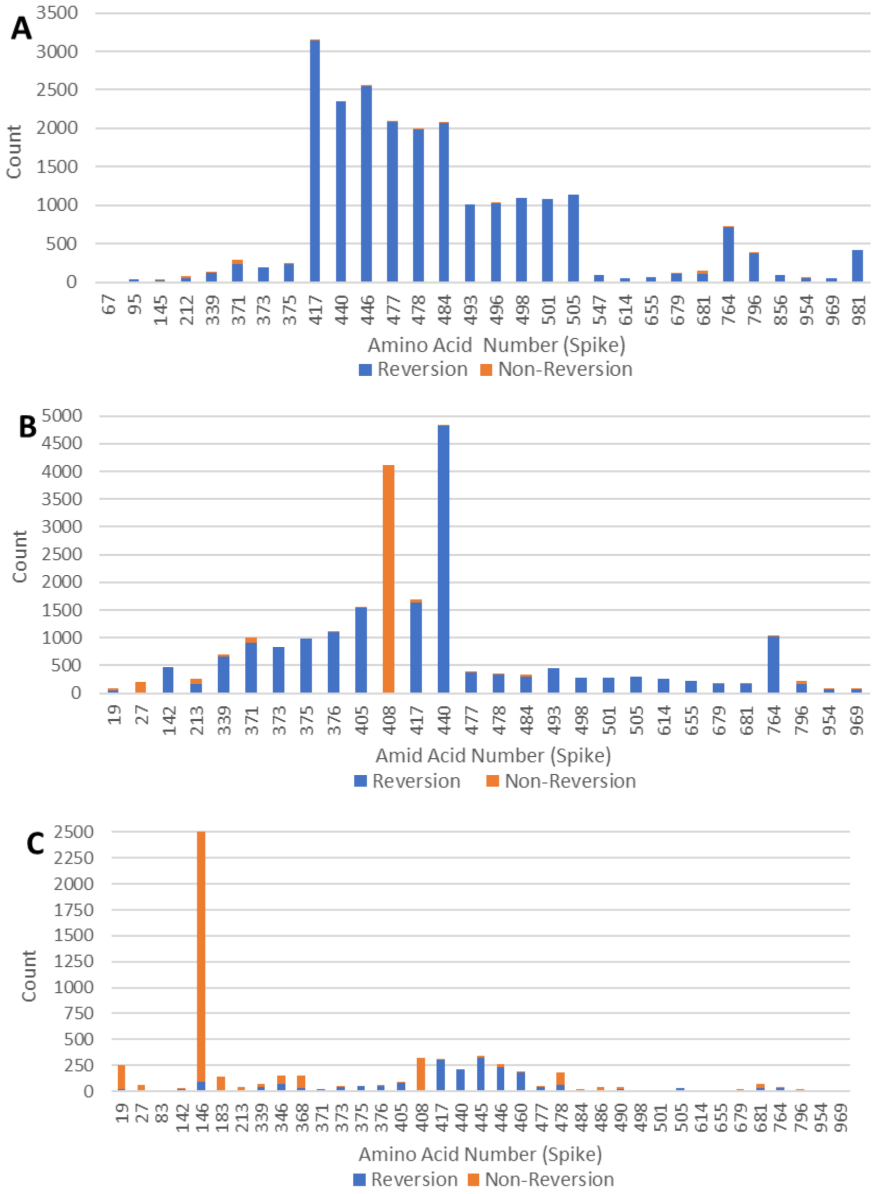


Figure 3. The histograms of reverse mutations in the Omicron variants (A) BA.1, (B) BA.2, and (C) XBB.1.5

Some of the single reversions are impossible to be generated by homologous recombination, for the neighboring mutation is too close to be attained with template switching. One possible way to attain single reversion is to insert point mutation or to combine homologous recombination and point mutation, as shown in Figure 4.

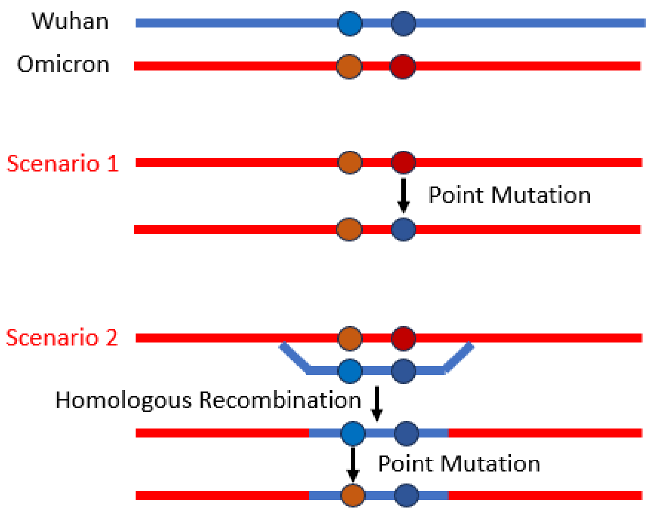
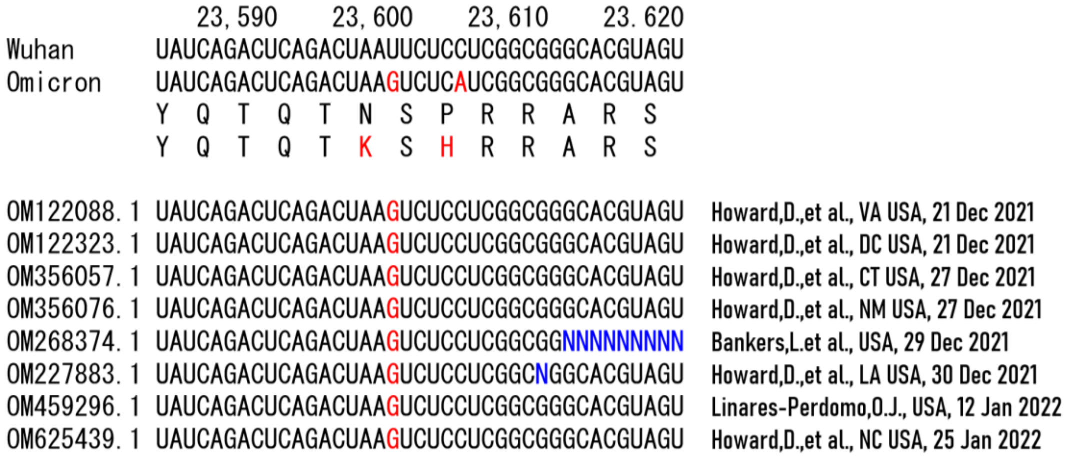


Figure 4. Possible scenarios of reverse mutation in one of the two amino acids close to each other.

BA.1 sequences including reversion at P681 while preserving K679 found in GenBank are listed in Figure 5. Among the neighboring mutations, this pattern is least likely to emerge naturally, for it requires point mutation at H679 from U to G after homologous recombination or point mutation at P681 from A to C, both of which is rare in the mutation spectrum of SARS-CoV-2 [26]. These mutants appeared soon after the emergence of BA.1 lineage. Three submitters registered these mutants independently, which backs up the existence of this unnatural mutation pattern.

Count of mutations co-occurring with G614D reverse mutation in B.1.617.2 and BA.2 lineages are listed in Figure 6. Here the amino acids where co-occurrence of mutation is more than 10% and 20% are shown for B.1.617.2 and BA.2 respectively. In B.1.617.2, almost all major mutations are reversions (the 95th amino acid is the exception) and overall co-occurrence of mutation is infrequent. In BA.2, co-occurrence of mutation is more frequent, while major mutations are again reversions (the 408th amino acid is the exception).



```
            23,590     23,600     23,610     23.620
Wuhan       UAUCAGACUCAGACUAAUUCUCCUCGGCGGGCACGUAGU
Omicron     UAUCAGACUCAGACUAAGUCUCAUCGGCGGGCACGUAGU
            Y  Q  T  Q  T  N  S  P  R  R  A  R  S
            Y  Q  T  Q  T  K  S  H  R  R  A  R  S

OM122088.1  UAUCAGACUCAGACUAAGUCUCCUCGGCGGGCACGUAGU  Howard,D.,et al., VA USA, 21 Dec 2021
OM122323.1  UAUCAGACUCAGACUAAGUCUCCUCGGCGGGCACGUAGU  Howard,D.,et al., DC USA, 21 Dec 2021
OM356057.1  UAUCAGACUCAGACUAAGUCUCCUCGGCGGGCACGUAGU  Howard,D.,et al., CT USA, 27 Dec 2021
OM356076.1  UAUCAGACUCAGACUAAGUCUCCUCGGCGGGCACGUAGU  Howard,D.,et al., NM USA, 27 Dec 2021
OM268374.1  UAUCAGACUCAGACUAAGUCUCCUCGGCGGNNNNNNNNN  Bankers,L.et al., USA, 29 Dec 2021
OM227883.1  UAUCAGACUCAGACUAAGUCUCCUCGGCNGGCACGUAGU  Howard,D.,et al., LA USA, 30 Dec 2021
OM459296.1  UAUCAGACUCAGACUAAGUCUCCUCGGCGGGCACGUAGU  Linares-Perdomo,O.J., USA, 12 Jan 2022
OM625439.1  UAUCAGACUCAGACUAAGUCUCCUCGGCGGGCACGUAGU  Howard,D.,et al., NC USA, 25 Jan 2022
```

Figure 5. BA.1 sequences including reversion at P681 while preserving K679.

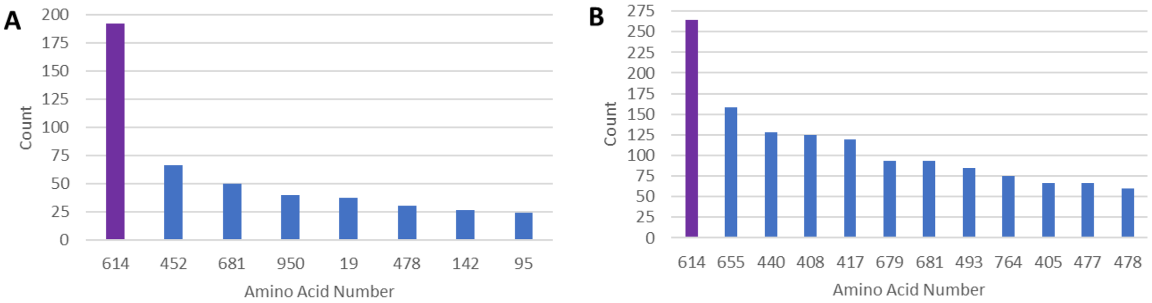


Figure 6. Frequent mutations co-occurring with G614D: (A) B.1.617.2; (B) BA.2.

The counts of the whole sequences and the reversions of D614G among them in 18 lineages (AY.44, AY.103, B.1.1.519, B.1.427, B.1.429, B.1.526, B.1.617.2, B.1.637, BA.1.15, BA.2, BA.5.2.1, BA.5.5 BE.1, BE.3 CH.1.1, D.2, XBB.1, XBB.1.6) are listed in Table 1. The rate of G614D reversion in B.1.617.2 stands out, followed by BA.2, which supports the claim by Kakeya et al. [21]. The counts of G614D reverse mutation and the whole data of B.1.617.2 and BA.2 lineages in each month are shown in Figure 7. The timing of G614D surge is preceded by the surge of the whole data both in B.1.617.2 and BA.2. The delay is significantly long in B.1.617.2.

Table 1. Counts of the whole sequences and the reversions of D614G in 18 lineages.

Lineage	AY.103	AY44	B.1.1.519	B.1.427	B.1.429	B.1.526	B.1.617.2	B.1.637	BA.1.15
# G614D	163	134	5	4	7	10	403	8	57
# All	242444	208550	12534	13062	29260	33588	45699	11185	84850
G614D %	0.067	0.064	0.040	0.031	0.024	0.030	0.882	0.072	0.067
Lineage	BA.2	BA.5.2.1	BA.5.5	BE.1	BE.3	CH.1.1	D.2	XBB.1.16	XBB.1
# G614D	189	5	6	1	0	0	0	0	2
# All	115711	82004	34641	6004	5185	1311	11423	3388	3186
G614D %	0.163	0.006	0.017	0.017	0.000	0.000	0.000	0.000	0.063

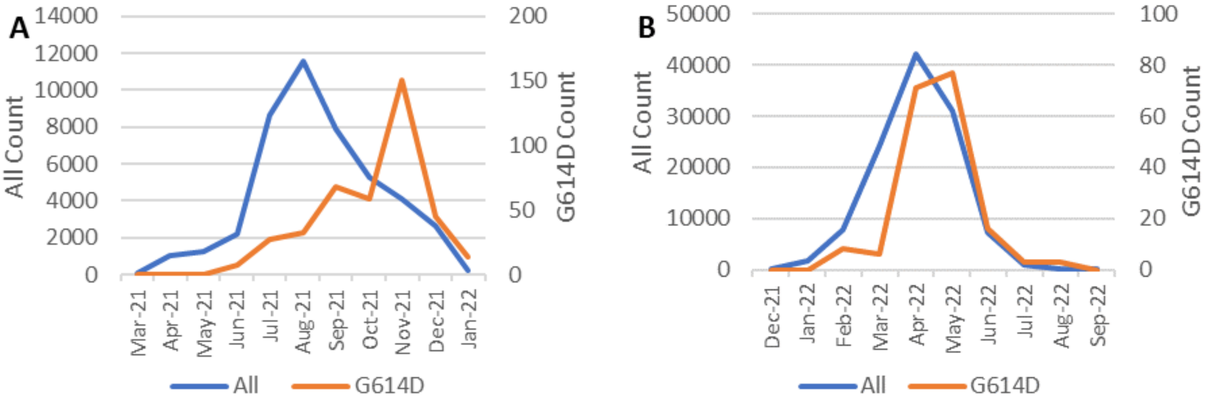


Figure 7. Counts of G614D reversion and the whole data of B.1.617.2 and BA.2 lineages in each month.

The locations in the United States where the whole sequences and the reversions of D614G in B.1.617.2 and BA.2 are expressed in heatmaps in Figure 8. The ratio of G614D reversion is the highest in the States of Michigan and Illinois as for B.1.617.2 and in the States of New York and New Jersy as for BA.2 with statistical significance (Supplemental Table 2).

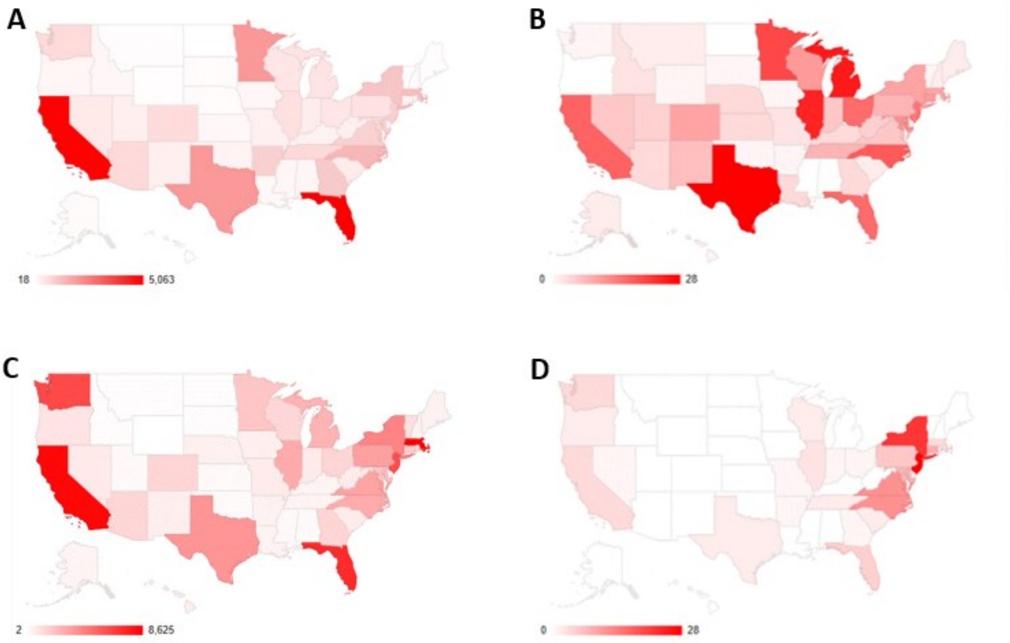


Figure 8. Heatmaps showing the locations of sampling: (A) the whole sequences of B.1.617.2; (B) reversions of D614G in B.1.617.2; (C) the whole sequences of BA.2; (D) reversions of D614G in BA.2.

Discussion

The significantly low rate of mutations other than reversions in BA.1, as shown in Figures 1-3, indicates that these mutations are not the products of immune pressure, for immune pressure leads to a large variety of mutations to escape from the immune response, which is observed in the mutations of other lineages, especially XBB.1.5. Some of the reversions in BA.1 can hardly emerge through natural mutation or recombination considering the sequence alignment and mutations spectrum, as shown in Figures 4 and 5. Since the sequences including these unnatural mutations are registered by multiple submitters independently, the existence of these mutants is quite certain.

Based on the above analyses, it is quite strange that the single reversion mutants are found at almost all the mutation points of the Omicron variant. One possible scenario is that someone in a lab artificially made a variety of mutants with a single reversion to see the effect of each mutation, which escaped from the lab as a result of an unknown incident without being noticed to the public.

As Figure 6 shows, most of the mutations co-occurring with G614D in Delta and Omicron BA.2 are reversions, which shows that these mutants lack any possible factors to stabilize D614. Since D614 is not stable in vivo, not only in humans but also in hamsters [27], emergence of this reverse mutation without any stabilizing cofactors is extremely abnormal.

Table 1 shows that Delta and Omicron BA.2 have notably more numbers of mutants including G614D. As Figure 7 shows, these mutants emerged after the surge of each lineage. Emergence of this unstable reverse mutation in the non-early phase of the prevalence is markedly strange. One possible scenario is that the original strain of the lineage was first sampled from an early patient, which was kept in cell cultures for research, where D614G reverse mutation was obtained possibly through recombination with lab-kept bat sarbecovirus or wild-type SARS-CoV-2, and escaped from the laboratory by accident to spread into human populations.

Figure 8 shows that the epicenter of D614G mutants was Illinois and Michigan as for B.1.617.2 and New York and New Jersy as for BA.2. The State of Illinois has one BSL3+ lab in Chicago. The City of St. Lewis, which is neighboring Illinois State, has a BSL3+ lab and a factory manufacturing mRNA vaccine. The State of Michigan also has a mRNA vaccine factory. New York State has two BSL3+ labs, one of which is located close to the State of New Jersy.

Since the start of the COVID-19 pandemic, quite a large number of laboratories have kept SARS-CoV-2 for experimental purposes. In the end of 2021, a researcher in Taiwan was bitten by a mouse in a biosafety level 3 laboratory and was infected with the Delta variant of SARS-CoV-2, spreading the disease around without noticing [28]. In this case, the incident was confirmed as a lab leak because the virus infection in Taiwan had been subdued due to a strict quarantine policy, which made it easier to identify the researcher as the source of infection. If a lab leak takes place in a city populated with many infected patients, it quite likely remains unnoticed.

Many lab-leak accidents have happened historically and the number of them has been increasing due to the recent spread of genetic engineering [29,30]. Unfortunately, those accidents have been covered up repeatedly in the field of microbiology [31]. A typical example is the Sverdlovsk anthrax leak in 1979 [32], which took 15 years to be accepted officially as a lab-leak event, while it took about 30 years to reach a consensus among virologists that the 1977 Russian influenza H1N1 originated from a frozen virus in a laboratory [33].

Many researchers have suspected that the original strain of SARS-CoV-2 might have might have leaked from a laboratory [34-41]. On June 20 of 2023, the Wall Steet Journal reported the names of three researchers in the Wuhan Institute of Virology who fell sick with typical symptoms of COVID-19 in November 2019 [42]. According to the article, US intelligence had confirmed that these researchers were modifying coronaviruses so that they could bind to human cells.

It is true that all the anomalies shown in this paper are just statistical biases, not the definitive proofs of lab leaks. To pursue what has really happened in laboratories during the COVID-19 pandemic, it is essential that an independent organization with no conflict of interest carries out thorough investigations without limitations of access.

Conclusion

Omicron BA.1 is outstanding in its high reverse mutation rate and low variety of mutations, which cannot be explained to emerge through natural evolution processes including immune escape. The sequences of BA.1 including both K679 and P681, which can hardly emerge naturally, are registered by multiple submitters, which consolidates their existence. G614D reverse mutations, which is also unlikely to emerge through human community spread, have the epicenters in the states of Michigan and Illinois as for the Delta variant and in the states of New York and New Jersy as for the Omicron BA.2. To elucidate the origin of these sequences, inspections of laboratories are needed, where the chronological and geographical distribution of anomalous sequence detections shown in this paper can help to narrow down the potential sources.

Conflicts of interest

The authors declare no conflict of interests exist.

References

[1] Callaway E. Heavily mutated Omicron variant puts scientists on alert. Nature 2021;600(7887):21. DOI: 10.1038/d41586-021-03552-w

[2] Jung C, Kmiec, D, Koepke L, et al. Omicron: what makes the latest SARS-CoV-2 variant of concern so concerning? J Virology 2022;96(6): e02077-21. DOI: 10.1128/jvi.02077-21

[3] Hassan SS, Kodakandla V, Redwan EM, et al. Non-uniform aspects of the SARS-CoV-2 intraspecies evolution reopen question of its origin. International Journal of Biological Macromolecules 2022; 222:972-993.

DOI: 10.1016/j.ijbiomac.2022.09.184

[4] Mallapaty C. The hunt for the origin of Omicron. Nature 2022;602(7898):26-28.

DOI: 10.1038/d41586-022-00215-2

[5] Jung C, Kmiec, D, Koepke L, et al. Omicron: what makes the latest SARS-CoV-2 variant of concern so concerning? J Virology 2022;96(6): e02077-21.

DOI: 10.1128/jvi.02077-21

[6] Choi B, Choudhary MC, Regan J, et al. Persistence and Evolution of SARS-CoV-2 in an Immunocompromised Host. The New England Journal of Medicine 2020;383(23):2291-2293.

DOI: 10.1056/NEJMc2031364

[7] Kemp SA, Collier DA, Datier RP, et al. SARS-CoV-2 evolution during treatment of chronic infection. Nature 2021;592:277-282.

DOI: 10.1038/s41586-021-03291-y

[8] Truong TT, Ryutov A, Pandey U, et al. Increased viral variants in children and young adults with impaired humoral immunity and persistent SARS-CoV-2 infection: A consecutive case series. EBioMedicine 2021 May;67:103355.

DOI: 10.1016/j.ebiom.2021.103355

[9] Miyata T, Yasunaga T. Molecular evolution of mRNA: A method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application. J Molecular Evolution 1980;16(1):23–36.

DOI: 10.1007/BF01732067

[10] Li WH, Wu CI, Luo CC. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. Molecular Biology and Evolution 1985;2(2):150–174.

DOI: 10.1093/oxfordjournals.molbev.a040343

[11] Wei C, Shan KJ, Wang W, et al. Evidence for a mouse origin of the SARS-CoV-2 Omicron variant. J Genet Genomics 2021;48(12):1111-1121.

DOI: 10.1016/j.jgg.2021.12.003

[12] Kryazhimskiy S, Plotkin JB, The population genetics of dN/dS. PLOS Genetics 2008;4(12):e1000304.

DOI: 10.1371/journal.pgen.1000304

[13] Hasan Z, Hasan M, Ashik AI, et al. Prediction of immune pressure on HIV-1 regulatory gene tat by human host through bioinformatics tools. J Adv Biotechnol Exp Ther. 2020 Sep;3(3):233-240.

DOI: 10.5455/jabet.2020.d129

[14] Zhan SH, Deverman BE, Chan YA. SARS-CoV-2 is well adapted for humans. What does this mean for re-emergence? bioRxiv 2020.

DOI: 10.1101/2020.05.01.073262

[15] Zhang W, Shi K, Geng Q, et al. Structural basis for mouse receptor recognition by SARS-CoV-2 omicron variant. PNAS 2022; 119(44): e2206509119.

DOI: 10.1073/pnas.2206509119

[16] Piplani S, Singh PK, Winkler DA, et al. In silico comparison of SARS-CoV-2 spike protein-ACE2 binding affinities across species and implications for virus origin. Science Report 2021;11:13063.

DOI: 10.1038/s41598-021-92388-5

[17] Kakeya H, Matsumoto Y. A probabilistic approach to evaluate the likelihood of artificial genetic modification and its application to SARS-CoV-2 Omicron variant. ISPJ Trans. Bioinformatics 2022;15:22-29.

DOI: 10.2197/ipsjtbio.15.22

[18] Kakeya H, Arakawa H, Matsumoto Y. Multiple probabilistic analyses suggest non-natural origin of SARS-CoV-2 Omicron variant. Zenodo 2023.

DOI: 10.5281/zenodo.7470652

[19] Arakawa H. Mutation signature of SARS-CoV-2 variants raises questions to their natural origins. Zenodo 2022.

DOI: 10.5281/zenodo.6601991

[20] Tanaka A, Miyazawa T. Unnaturalness in the evolution process of the SARS-CoV-2 variants and the possibility of deliberate natural selection. Zenodo 2023.

DOI: 10.5281/zenodo.8361577

[21] Kakeya H, Matsumoto Y. Repeated emergence of probabilistically and chronologically anomalous mutations in SARS-CoV-2 during the COVID-19 pandemic. Zenodo 2023.

DOI: 10.5281/zenodo.8216232

[22] Korber B, Fischer WM, Gnanakaran S, et al. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. Cell 2020; 182(4), 812-827.

DOI: 10.1016/j.cell.2020.06.043

[23] Volz E, Hill V, McCrone JT, et al. Evaluating the effects of SARS-CoV-2 spike mutation D614G on transmissibility and pathogenicity, Cell 2021; 184(1), 64-75.

DOI: 10.1016/j.cell.2020.11.020

[24] Selective Subcommittee on the Coronavirus Pandemic (United States House). Letter to CIA Director William Burns, September 12, 2023.

https://oversight.house.gov/wp-content/uploads/2023/09/2023.09.12-SSCP-HPSCI-Letter-to-CIA-Re.-Origins-of-COVID.pdf [cited Oct 28, 2023]

[25] Selective Subcommittee on the Coronavirus Pandemic (United States House). Letter to OIG, September 26, 2023.

https://oversight.house.gov/wp-content/uploads/2023/09/2023.09.26-SSCP-Letter-to-HHS-OIG-Re.-AF-Movements.pdf [cited Oct 28, 2023]

[26] Shan KJ, Wei C, Wang Y, et al. Host-specific asymmetric accumulation of mutation types reveals that the origin of SARS-CoV-2 is consistent with a natural process. Innovation 2021;2(4), 100159.

DOI: 10.1016/j.xinn.2021.100159

[27] Hou YJ, Chiba S, Halfmann P, et al. SARS-CoV-2 D614G variant exhibits efficient replication ex vivo and transmission in vivo, Science 2020;370(6523):1464-1468

DOI: 10.1126/science.abe8499

[28] Silver A. Taiwan’s science academy fined for biosafety lapses after lab worker contracts COVID-19. Science 2022.

DOI: 10.1126/science.ada0525

[29] Butler D. Fears grow over lab-bred flu. Nature 2011;480(7378):421–422.

DOI: 10.1038/480421a

[30] Biosafety in the balance. Nature 2014;510(7506):443.

DOI: 10.1038/510443a

[31] Young A. Pandora's gamble: lab leaks, pandemics, and a world at risk. Center Street 2023.

ISBN-13: 978-1546002932

[32] Meselson M, Guillemin J, Hugh-Jones M. The Sverdlovsk Anthrax Outbreak of 1979. Science 1994;266(5188):1202-1208.

DOI: 10.1126/science.7973702

[33] Kransnitz M, Levine AJ, Rabadan R. Anomalies in the Influenza Virus Genome Database: New Biology or Laboratory Errors? J Virol. 2008 Sep;82(17):8947-50.

DOI: 10.1128/JVI.00101-08.

[34] Sallard E, Halloy J, Casane D, et al. Tracing the origins of SARS-COV-2 in coronavirus phylogenies: a review. Environ Chem Lett 2021;19(4), 769-785.

DOI: 10.1007/s10311-020-01151-1

[35] Segreto R, Deigin Y, McCairn K, et al. Should we discount the laboratory origin of COVID-19? Environ Chem Lett 2021;Mar 15, 1-15

DOI: 10.1007/s10311-021-01211-0

[36] Quay SC. A Bayesian analysis concludes beyond a reasonable doubt that SARS-CoV-2 is not a natural zoonosis but instead is laboratory derived, Zenodo 2021.

DOI: 10.5281/zenodo.4477081

[37] Wiesendanger R. Studie zum Ursprung der Coronavirus-Pandemie. ResearchGate 2021.

DOI:10.13140/RG.2.2.31754.80323

[38] Markson S. What really happened in Wuhan. HarperCollins 2021.

ISBN-13: 978-1460761083

[39] Chan AJ and Ridley M. Viral: the search for the origin of Covid-19. Fourth Estate 2021.

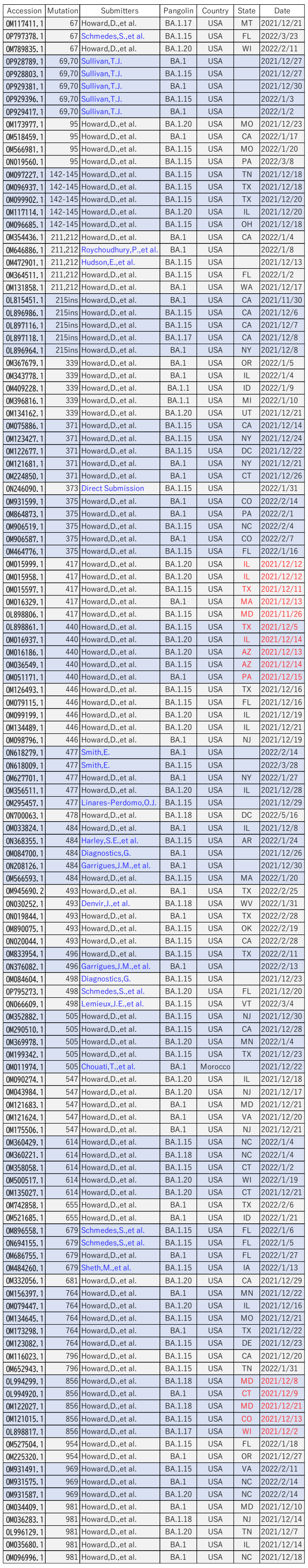
ISBN-13: 978-0008487492

[40] Harrison NL, Sachs JD. A call for an independent inquiry into the origin of the SARS-CoV-2 virus, PNAS 2022;119(21):e2202769119
DOI: 10.1073/pnas.2202769119

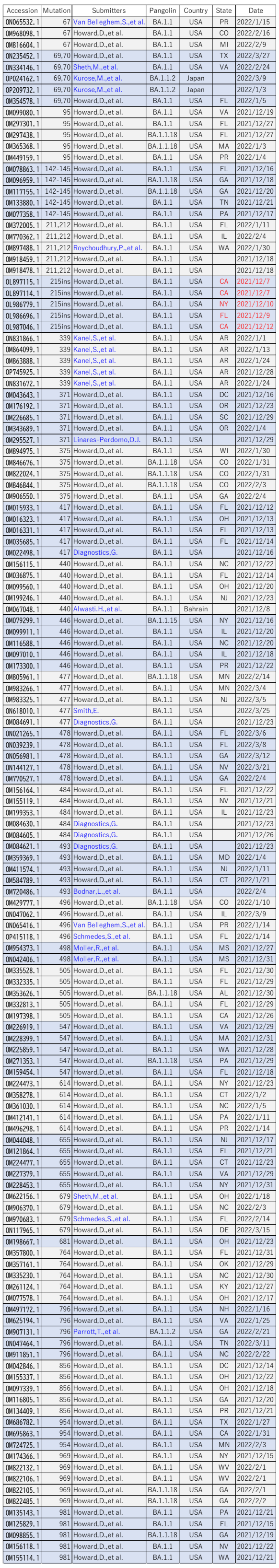
[41] Coccia M. Meta-analysis to explain unknown causes of the origins of SARS-CoV-2, Environmental Research 2022; 211:113062
DOI: 10.1016/j.envres.2022.113062

[42] Gordon MR. U.S.-funded scientist among three Chinese researchers who fell ill amid early Covid-19 outbreak. Wall Street Journal June 20, 2023.
https://www.wsj.com/articles/u-s-funded-scientist-among-three-chinese-researchers-who-fell-ill-amid-early-covid-19-outbreak-3f919567 [cited Oct 28, 2023]

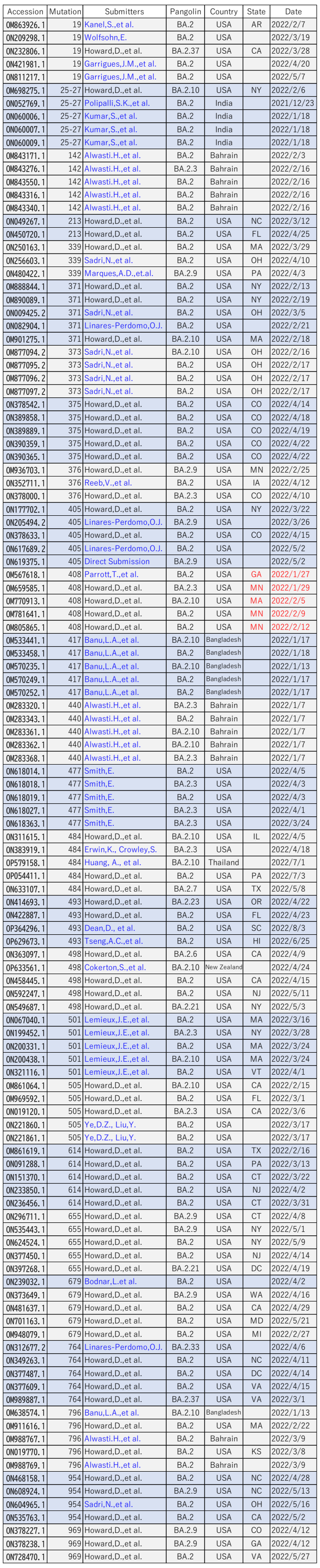
Supplemental Table 1. List of revertant found by Tanaka et al. containing only one reverse mutation and no other mutations in spike protein [20]



Accession	Mutation	Submitters	Pangolin	Country	State	Date
OM117411.1	67	Howard,D.,et al.	BA.1.17	USA	MT	2021/12/21
OP797378.1	67	Schmedes,S.,et al.	BA.1.15	USA	FL	2022/3/23
OM789835.1	67	Howard,D.,et al.	BA.1.20	USA	WI	2022/2/11
OP928789.1	69,70	Sullivan,T.J.	BA.1	USA		2021/12/27
OP928803.1	69,70	Sullivan,T.J.	BA.1.15	USA		2021/12/27
OP929381.1	69,70	Sullivan,T.J.	BA.1	USA		2021/12/30
OP929396.1	69,70	Sullivan,T.J.	BA.1.15	USA		2022/1/3
OP929417.1	69,70	Sullivan,T.J.	BA.1	USA		2022/1/2
OM173977.1	95	Howard,D.,et al.	BA.1.20	USA	MO	2021/12/23
OM518459.1	95	Howard,D.,et al.	BA.1	USA	CA	2022/1/17
OM566981.1	95	Howard,D.,et al.	BA.1.15	USA	MO	2022/1/20
ON019560.1	95	Howard,D.,et al.	BA.1.15	USA	PA	2022/3/8
OM097227.1	142-145	Howard,D.,et al.	BA.1.15	USA	TN	2021/12/18
OM096937.1	142-145	Howard,D.,et al.	BA.1.15	USA	TX	2021/12/18
OM099902.1	142-145	Howard,D.,et al.	BA.1.15	USA	TX	2021/12/20
OM117114.1	142-145	Howard,D.,et al.	BA.1.20	USA	IL	2021/12/20
OM096685.1	142-145	Howard,D.,et al.	BA.1.15	USA	OH	2021/12/18
OM354436.1	211,212	Howard,D.,et al.	BA.1	USA	CA	2022/1/4
OM646886.1	211,212	Roychoudhury,P.,et al.	BA.1	USA		2022/1/8
OM472901.1	211,212	Hudson,E.,et al.	BA.1.15	USA		2021/12/13
OM364511.1	211,212	Howard,D.,et al.	BA.1.15	USA	FL	2022/1/2
OM131858.1	211,212	Howard,D.,et al.	BA.1	USA	WA	2021/12/17
OL815451.1	215ins	Howard,D.,et al.	BA.1	USA	CA	2021/11/30
OL896986.1	215ins	Howard,D.,et al.	BA.1.15	USA	CA	2021/12/6
OL897116.1	215ins	Howard,D.,et al.	BA.1.15	USA	CA	2021/12/7
OL897118.1	215ins	Howard,D.,et al.	BA.1.17	USA	CA	2021/12/8
OL896964.1	215ins	Howard,D.,et al.	BA.1	USA	NY	2021/12/8
OM367679.1	339	Howard,D.,et al.	BA.1	USA	OR	2022/1/5
OM343778.1	339	Howard,D.,et al.	BA.1	USA	IL	2022/1/4
OM409228.1	339	Howard,D.,et al.	BA.1.1	USA	ID	2022/1/9
OM396816.1	339	Howard,D.,et al.	BA.1.1	USA	MI	2022/1/10
OM134162.1	339	Howard,D.,et al.	BA.1.20	USA	UT	2021/12/21
OM075886.1	371	Howard,D.,et al.	BA.1.15	USA	CA	2021/12/14
OM123427.1	371	Howard,D.,et al.	BA.1.15	USA	NY	2021/12/24
OM122677.1	371	Howard,D.,et al.	BA.1.15	USA	DC	2021/12/22
OM121681.1	371	Howard,D.,et al.	BA.1	USA	NY	2021/12/21
OM224850.1	371	Howard,D.,et al.	BA.1	USA	CT	2021/12/26
ON246090.1	373	Direct Submission	BA.1.15	USA		2022/1/31
OM931599.1	375	Howard,D.,et al.	BA.1	USA	CO	2022/2/14
OM864873.1	375	Howard,D.,et al.	BA.1	USA	PA	2022/2/1
OM906519.1	375	Howard,D.,et al.	BA.1.15	USA	NC	2022/2/4
OM906587.1	375	Howard,D.,et al.	BA.1	USA	CO	2022/2/7
OM464776.1	375	Howard,D.,et al.	BA.1.15	USA	FL	2022/1/16
OM015999.1	417	Howard,D.,et al.	BA.1.20	USA	IL	2021/12/12
OM015958.1	417	Howard,D.,et al.	BA.1.20	USA	IL	2021/12/12
OM015597.1	417	Howard,D.,et al.	BA.1.15	USA	TX	2021/12/11
OM016329.1	417	Howard,D.,et al.	BA.1	USA	MA	2021/12/13
OL898806.1	417	Howard,D.,et al.	BA.1.15	USA	MD	2021/11/26
OL898861.1	440	Howard,D.,et al.	BA.1.15	USA	TX	2021/12/5
OM016937.1	440	Howard,D.,et al.	BA.1.20	USA	IL	2021/12/14
OM016186.1	440	Howard,D.,et al.	BA.1.20	USA	AZ	2021/12/13
OM036549.1	440	Howard,D.,et al.	BA.1.15	USA	AZ	2021/12/14
OM051171.1	440	Howard,D.,et al.	BA.1	USA	PA	2021/12/15
OM126493.1	446	Howard,D.,et al.	BA.1.15	USA	TX	2021/12/16
OM079115.1	446	Howard,D.,et al.	BA.1.15	USA	FL	2021/12/16
OM099199.1	446	Howard,D.,et al.	BA.1.20	USA	IL	2021/12/19
OM134489.1	446	Howard,D.,et al.	BA.1.20	USA	IL	2021/12/21
OM098796.1	446	Howard,D.,et al.	BA.1	USA	NJ	2021/12/19
ON618279.1	477	Smith,E.	BA.1	USA		2022/2/14
ON618009.1	477	Smith,E.	BA.1.15	USA		2022/3/28
OM627701.1	477	Howard,D.,et al.	BA.1	USA	NY	2022/1/27
OM356511.1	477	Howard,D.,et al.	BA.1.20	USA	IL	2021/12/28
OM295457.1	477	Linares-Perdomo,O.J.	BA.1.15	USA		2021/12/29
ON700063.1	478	Howard,D.,et al.	BA.1.18	USA	DC	2022/5/16
OM033824.1	484	Howard,D.,et al.	BA.1	USA	IL	2021/12/8
ON368355.1	484	Harley,S.E.,et al.	BA.1.15	USA	AR	2022/1/24
OM084700.1	484	Diagnostics,G.	BA.1	USA		2021/12/26
ON208126.1	484	Garrigues,J.M.,et al.	BA.1	USA		2021/12/30
OM566593.1	484	Howard,D.,et al.	BA.1.15	USA	MA	2022/1/20
OM945690.2	493	Howard,D.,et al.	BA.1	USA	TX	2022/2/25
ON030252.1	493	Denvir,J.,et al.	BA.1.18	USA	WV	2022/1/31
ON019844.1	493	Howard,D.,et al.	BA.1	USA	TX	2022/2/28
OM890075.1	493	Howard,D.,et al.	BA.1.15	USA	OK	2022/2/19
ON020044.1	493	Howard,D.,et al.	BA.1.15	USA	CA	2022/2/28
OM833954.1	496	Howard,D.,et al.	BA.1.15	USA	TX	2022/2/11
ON376082.1	496	Garrigues,J.M.,et al.	BA.1	USA		2022/2/13
OM084604.1	498	Diagnostics,G.	BA.1.15	USA		2021/12/23
OP795273.1	498	Schmedes,S.,et al.	BA.1.20	USA	FL	2021/12/20
ON066609.1	498	Lemieux,J.E.,et al.	BA.1.15	USA	VT	2022/3/4
OM352882.1	505	Howard,D.,et al.	BA.1.15	USA	NJ	2021/12/30
OM290510.1	505	Howard,D.,et al.	BA.1.15	USA	CA	2021/12/28
OM369978.1	505	Howard,D.,et al.	BA.1.20	USA	MN	2022/1/4
OM199342.1	505	Howard,D.,et al.	BA.1.15	USA	TX	2021/12/23
OM011974.1	505	Chouati,T.,et al.	BA.1	Morocco		2021/12/22
OM090274.1	547	Howard,D.,et al.	BA.1.20	USA	IL	2021/12/18
OM043984.1	547	Howard,D.,et al.	BA.1.20	USA	NJ	2021/12/17
OM121683.1	547	Howard,D.,et al.	BA.1	USA	MD	2021/12/21
OM121624.1	547	Howard,D.,et al.	BA.1	USA	VA	2021/12/20
OM175506.1	547	Howard,D.,et al.	BA.1	USA	NJ	2021/12/21
OM360429.1	614	Howard,D.,et al.	BA.1.15	USA	NC	2022/1/4
OM360221.1	614	Howard,D.,et al.	BA.1.18	USA	NC	2022/1/4
OM358058.1	614	Howard,D.,et al.	BA.1.15	USA	CT	2022/1/2
OM500517.1	614	Howard,D.,et al.	BA.1.20	USA	WI	2022/1/19
OM135027.1	614	Howard,D.,et al.	BA.1.20	USA	CT	2021/12/21
OM742858.1	655	Howard,D.,et al.	BA.1	USA	TX	2022/2/6
OM521685.1	655	Howard,D.,et al.	BA.1	USA	ID	2022/1/21
OM896558.1	679	Schmedes,S.,et al.	BA.1.15	USA	FL	2022/1/6
ON694155.1	679	Schmedes,S.,et al.	BA.1.15	USA	FL	2022/1/5
OM686755.1	679	Howard,D.,et al.	BA.1	USA	FL	2022/1/27
OM484260.1	679	Sheth,M.,et al.	BA.1.15	USA	IA	2022/1/13
OM332056.1	681	Howard,D.,et al.	BA.1.20	USA	CA	2021/12/29
OM156397.1	764	Howard,D.,et al.	BA.1	USA	MN	2021/12/22
OM079447.1	764	Howard,D.,et al.	BA.1.20	USA	IL	2021/12/16
OM134645.1	764	Howard,D.,et al.	BA.1.15	USA	MO	2021/12/21
OM173298.1	764	Howard,D.,et al.	BA.1	USA	TX	2021/12/22
OM123082.1	764	Howard,D.,et al.	BA.1.15	USA	DE	2021/12/23
OM116023.1	796	Howard,D.,et al.	BA.1.15	USA	CA	2021/12/20
OM652943.1	796	Howard,D.,et al.	BA.1.15	USA	TN	2022/1/31
OL994299.1	856	Howard,D.,et al.	BA.1.18	USA	MD	2021/12/8
OL994920.1	856	Howard,D.,et al.	BA.1	USA	CT	2021/12/9
OM122027.1	856	Howard,D.,et al.	BA.1.18	USA	MD	2021/12/21
OM121015.1	856	Howard,D.,et al.	BA.1.15	USA	CO	2021/12/13
OL898817.1	856	Howard,D.,et al.	BA.1.17	USA	WI	2021/12/2
OM527504.1	954	Howard,D.,et al.	BA.1.15	USA	FL	2022/1/18
OM225320.1	954	Howard,D.,et al.	BA.1	USA	OR	2021/12/27
OM931491.1	969	Howard,D.,et al.	BA.1.15	USA	VA	2022/2/11
OM931575.1	969	Howard,D.,et al.	BA.1	USA	NC	2022/2/14
OM931587.1	969	Howard,D.,et al.	BA.1.20	USA	NC	2022/2/14
OM034409.1	981	Howard,D.,et al.	BA.1	USA	MD	2021/12/10
OM036283.1	981	Howard,D.,et al.	BA.1.18	USA	NJ	2021/12/14
OL996129.1	981	Howard,D.,et al.	BA.1.20	USA	TN	2021/12/7
OM035680.1	981	Howard,D.,et al.	BA.1	USA	IL	2021/12/14
OM096996.1	981	Howard,D.,et al.	BA.1	USA	NC	2021/12/18



Accession	Mutation	Submitters	Pangolin	Country	State	Date
ON065532.1	67	Van Belleghem,S.,et al.	BA.1.1	USA	PR	2022/1/15
OM968098.1	67	Howard,D.,et al.	BA.1.1	USA	CO	2022/2/16
OM816604.1	67	Howard,D.,et al.	BA.1.1	USA	MI	2022/2/9
ON235452.1	69,70	Howard,D.,et al.	BA.1.1	USA	TX	2022/3/27
ON334146.1	69,70	Sheth,M.,et al.	BA.1.1	USA	VA	2022/2/24
OP024162.1	69,70	Kurose,M.,et al.	BA.1.1.2	Japan		2022/3/9
OP209732.1	69,70	Kurose,M.,et al.	BA.1.1.2	Japan		2022/1/3
OM354578.1	69,70	Howard,D.,et al.	BA.1.1	USA	FL	2022/1/5
OM099080.1	95	Howard,D.,et al.	BA.1.1	USA	VA	2021/12/19
OM297301.1	95	Howard,D.,et al.	BA.1.1	USA	FL	2021/12/27
OM297438.1	95	Howard,D.,et al.	BA.1.1.18	USA	FL	2021/12/27
OM365368.1	95	Howard,D.,et al.	BA.1.1.18	USA	MA	2022/1/3
OM449159.1	95	Howard,D.,et al.	BA.1.1	USA	PR	2022/1/4
OM078863.1	142-145	Howard,D.,et al.	BA.1.1	USA	FL	2021/12/16
OM096959.1	142-145	Howard,D.,et al.	BA.1.1.18	USA	GA	2021/12/18
OM117155.1	142-145	Howard,D.,et al.	BA.1.1.18	USA	GA	2021/12/20
OM133880.1	142-145	Howard,D.,et al.	BA.1.1	USA	TN	2021/12/21
OM077358.1	142-145	Howard,D.,et al.	BA.1.1	USA	PA	2021/12/17
OM372005.1	211,212	Howard,D.,et al.	BA.1.1	USA	FL	2022/1/11
OM770362.1	211,212	Howard,D.,et al.	BA.1.1	USA	IL	2022/2/4
OM897488.1	211,212	Roychoudhury,P.,et al.	BA.1.1	USA	WA	2022/1/30
OM918459.1	211,212	Howard,D.,et al.	BA.1.1	USA		2021/12/18
OM918478.1	211,212	Howard,D.,et al.	BA.1.1	USA		2021/12/18
OL897115.1	215ins	Howard,D.,et al.	BA.1.1	USA	CA	2021/12/7
OL897114.1	215ins	Howard,D.,et al.	BA.1.1	USA	CA	2021/12/7
OL986779.1	215ins	Howard,D.,et al.	BA.1.1	USA	NY	2021/12/10
OL986696.1	215ins	Howard,D.,et al.	BA.1.1	USA	FL	2021/12/9
OL987046.1	215ins	Howard,D.,et al.	BA.1.1	USA	CA	2021/12/12
ON831866.1	339	Kanel,S.,et al.	BA.1.1	USA	AR	2022/1/1
OM864099.1	339	Kanel,S.,et al.	BA.1.1	USA	AR	2022/1/13
OM863888.1	339	Kanel,S.,et al.	BA.1.1	USA	AR	2022/1/24
OP745925.1	339	Kanel,S.,et al.	BA.1.1	USA	AR	2021/12/28
ON831672.1	339	Kanel,S.,et al.	BA.1.1	USA	AR	2022/1/24
OM043643.1	371	Howard,D.,et al.	BA.1.1	USA	DC	2021/12/16
OM176192.1	371	Howard,D.,et al.	BA.1.1	USA	OR	2021/12/23
OM226685.1	371	Howard,D.,et al.	BA.1.1	USA	SC	2021/12/29
OM343689.1	371	Howard,D.,et al.	BA.1.1	USA	OR	2022/1/4
OM295527.1	371	Linares-Perdomo,O.J.	BA.1.1	USA		2021/12/29
OM894975.1	375	Howard,D.,et al.	BA.1.1	USA	WI	2022/1/30
OM846676.1	375	Howard,D.,et al.	BA.1.1.18	USA	CO	2022/1/31
OM822024.1	375	Howard,D.,et al.	BA.1.1.18	USA	CO	2022/1/31
OM846844.1	375	Howard,D.,et al.	BA.1.1.18	USA	CO	2022/2/3
OM906550.1	375	Howard,D.,et al.	BA.1.1	USA	GA	2022/2/4
OM015933.1	417	Howard,D.,et al.	BA.1.1	USA	FL	2021/12/12
OM016323.1	417	Howard,D.,et al.	BA.1.1	USA	OH	2021/12/13
OM016331.1	417	Howard,D.,et al.	BA.1.1	USA	FL	2021/12/13
OM035685.1	417	Howard,D.,et al.	BA.1.1	USA	FL	2021/12/14
OM022498.1	417	Diagnostics,G.	BA.1.1	USA		2021/12/16
OM156115.1	440	Howard,D.,et al.	BA.1.1	USA	NC	2021/12/22
OM036875.1	440	Howard,D.,et al.	BA.1.1	USA	FL	2021/12/14
OM099560.1	440	Howard,D.,et al.	BA.1.1	USA	OH	2021/12/20
OM199246.1	440	Howard,D.,et al.	BA.1.1	USA	NJ	2021/12/23
OM067048.1	440	Alwasti.H.,et al.	BA.1.1	Bahrain		2021/12/8
OM079299.1	446	Howard,D.,et al.	BA.1.1.15	USA	NY	2021/12/16
OM099911.1	446	Howard,D.,et al.	BA.1.1	USA	IL	2021/12/20
OM116588.1	446	Howard,D.,et al.	BA.1.1	USA	NC	2021/12/20
OM097010.1	446	Howard,D.,et al.	BA.1.1	USA	IL	2021/12/18
OM173300.1	446	Howard,D.,et al.	BA.1.1	USA	PR	2021/12/22
OM805961.1	477	Howard,D.,et al.	BA.1.1.18	USA	MN	2022/2/14
OM983266.1	477	Howard,D.,et al.	BA.1.1	USA	MN	2022/3/4
OM983325.1	477	Howard,D.,et al.	BA.1.1	USA	NJ	2022/3/5
ON618010.1	477	Smith,E.	BA.1.1	USA		2022/3/25
OM084691.1	477	Diagnostics,G.	BA.1.1	USA		2021/12/23
ON021265.1	478	Howard,D.,et al.	BA.1.1	USA	FL	2022/3/6
ON039239.1	478	Howard,D.,et al.	BA.1.1	USA	FL	2022/3/8
ON056981.1	478	Howard,D.,et al.	BA.1.1	USA	GA	2022/3/12
ON144127.1	478	Howard,D.,et al.	BA.1.1	USA	NV	2022/3/21
OM770527.1	478	Howard,D.,et al.	BA.1.1	USA	GA	2022/2/4
OM156164.1	484	Howard,D.,et al.	BA.1.1	USA	FL	2021/12/22
OM155119.1	484	Howard,D.,et al.	BA.1.1	USA	NV	2021/12/21
OM199353.1	484	Howard,D.,et al.	BA.1.1	USA	IL	2021/12/23
OM084630.1	484	Diagnostics,G.	BA.1.1	USA		2021/12/23
OM084605.1	484	Diagnostics,G.	BA.1.1	USA		2021/12/26
OM084621.1	493	Diagnostics,G.	BA.1.1	USA		2021/12/23
OM359369.1	493	Howard,D.,et al.	BA.1.1	USA	MD	2022/1/4
OM411574.1	493	Howard,D.,et al.	BA.1.1	USA	NJ	2022/1/11
OM584789.1	493	Howard,D.,et al.	BA.1.1	USA	CT	2022/1/21
OM720486.1	493	Bodnar,L.,et al.	BA.1.1	USA		2022/2/4
OM429777.1	496	Howard,D.,et al.	BA.1.1.18	USA	CO	2022/1/10
ON047062.1	496	Howard,D.,et al.	BA.1.1	USA	IL	2022/3/9
ON065416.1	496	Van Belleghem,S.,et al.	BA.1.1	USA	PR	2022/1/14
OP415118.1	496	Schmedes,S.,et al.	BA.1.1	USA	FL	2022/1/14
OM954373.1	498	Moller,R.,et al.	BA.1.1	USA	MS	2021/12/27
ON042406.1	498	Moller,R.,et al.	BA.1.1	USA	MS	2021/12/31
OM335528.1	505	Howard,D.,et al.	BA.1.1	USA	FL	2021/12/30
OM332335.1	505	Howard,D.,et al.	BA.1.1	USA	FL	2021/12/29
OM353626.1	505	Howard,D.,et al.	BA.1.1.18	USA	AL	2021/12/30
OM332813.1	505	Howard,D.,et al.	BA.1.1	USA	FL	2021/12/29
OM197398.1	505	Howard,D.,et al.	BA.1.1	USA	CA	2021/12/26
OM226919.1	547	Howard,D.,et al.	BA.1.1	USA	VA	2021/12/29
OM228399.1	547	Howard,D.,et al.	BA.1.1	USA	MA	2021/12/31
OM225859.1	547	Howard,D.,et al.	BA.1.1	USA	WA	2021/12/28
OM271353.1	547	Howard,D.,et al.	BA.1.1.18	USA	PA	2021/12/29
OM159454.1	547	Howard,D.,et al.	BA.1.1	USA	FL	2021/12/18
OM224473.1	614	Howard,D.,et al.	BA.1.1	USA	NY	2021/12/23
OM358278.1	614	Howard,D.,et al.	BA.1.1	USA	CT	2022/1/2
OM361030.1	614	Howard,D.,et al.	BA.1.1	USA	NC	2022/1/5
OM412141.1	614	Howard,D.,et al.	BA.1.1	USA	PA	2022/1/11
OM496298.1	614	Howard,D.,et al.	BA.1.1	USA	PR	2022/1/14
OM044048.1	655	Howard,D.,et al.	BA.1.1	USA	NJ	2021/12/17
OM121864.1	655	Howard,D.,et al.	BA.1.1	USA	FL	2021/12/21
OM224477.1	655	Howard,D.,et al.	BA.1.1	USA	CT	2021/12/23
OM227379.1	655	Howard,D.,et al.	BA.1.1	USA	VA	2021/12/29
OM228453.1	655	Howard,D.,et al.	BA.1.1	USA	NY	2021/12/31
OM622156.1	679	Sheth,M.,et al.	BA.1.1	USA	OH	2022/1/18
OM906370.1	679	Howard,D.,et al.	BA.1.1	USA	NC	2022/2/3
OM970683.1	679	Schmedes,S.,et al.	BA.1.1	USA	FL	2022/2/14
ON117965.1	679	Howard,D.,et al.	BA.1.1	USA	DE	2022/3/15
OM198667.1	681	Howard,D.,et al.	BA.1.1	USA	OH	2021/12/23
OM357800.1	764	Howard,D.,et al.	BA.1.1	USA	FL	2021/12/31
OM357161.1	764	Howard,D.,et al.	BA.1.1	USA	OK	2021/12/29
OM335230.1	764	Howard,D.,et al.	BA.1.1	USA	NC	2021/12/30
OM261124.1	764	Howard,D.,et al.	BA.1.1	USA	KY	2021/12/27
OM077578.1	764	Howard,D.,et al.	BA.1.1	USA	OH	2021/12/17
OM497172.1	796	Howard,D.,et al.	BA.1.1	USA	NH	2022/1/16
OM625194.1	796	Howard,D.,et al.	BA.1.1	USA	VA	2022/1/25
OM907131.1	796	Parrott,T.,et al.	BA.1.1.2	USA	GA	2022/2/21
ON047464.1	796	Howard,D.,et al.	BA.1.1	USA	TN	2022/3/11
OM911851.1	796	Howard,D.,et al.	BA.1.1	USA	NC	2022/2/22
OM042846.1	856	Howard,D.,et al.	BA.1.1	USA	DC	2021/12/14
OM155337.1	856	Howard,D.,et al.	BA.1.1	USA	OH	2021/12/22
OM097339.1	856	Howard,D.,et al.	BA.1.1	USA	OH	2021/12/18
OM116805.1	856	Howard,D.,et al.	BA.1.1	USA	GA	2021/12/20
OM134409.1	856	Howard,D.,et al.	BA.1.1	USA	PR	2021/12/21
OM686782.1	954	Howard,D.,et al.	BA.1.1	USA	TX	2022/1/27
OM695863.1	954	Howard,D.,et al.	BA.1.1	USA	CA	2022/1/31
OM724725.1	954	Howard,D.,et al.	BA.1.1	USA	MN	2022/2/3
OM174366.1	969	Howard,D.,et al.	BA.1.1	USA	NY	2021/12/15
OM822132.1	969	Howard,D.,et al.	BA.1.1	USA	WV	2022/2/1
OM822106.1	969	Howard,D.,et al.	BA.1.1	USA	WV	2022/2/1
OM822105.1	969	Howard,D.,et al.	BA.1.1.18	USA	GA	2022/2/1
OM822485.1	969	Howard,D.,et al.	BA.1.1.18	USA	GA	2022/2/2
OM135143.1	981	Howard,D.,et al.	BA.1.1	USA	PA	2021/12/21
OM125829.1	981	Howard,D.,et al.	BA.1.1	USA	FL	2021/12/15
OM098855.1	981	Howard,D.,et al.	BA.1.1.18	USA	GA	2021/12/19
OM156118.1	981	Howard,D.,et al.	BA.1.1	USA	NV	2021/12/22
OM155114.1	981	Howard,D.,et al.	BA.1.1	USA	WA	2021/12/21



Accession	Mutation	Submitters	Pangolin	Country	State	Date
OM863926.1	19	Kanel,S.,et al.	BA.2	USA	AR	2022/2/7
ON209298.1	19	Wolfsohn,E.	BA.2	USA		2022/3/19
ON232806.1	19	Howard,D.,et al.	BA.2.37	USA	CA	2022/3/28
ON421981.1	19	Garrigues,J.M.,et al.	BA.2	USA		2022/4/20
ON811217.1	19	Garrigues,J.M.,et al.	BA.2	USA		2022/5/7
OM698275.1	25-27	Howard,D.,et al.	BA.2.10	USA	NY	2022/2/6
ON052769.1	25-27	Polipalli,S.K.,et al.	BA.2	India		2021/12/23
ON060006.1	25-27	Kumar,S.,et al.	BA.2	India		2022/1/18
ON060007.1	25-27	Kumar,S.,et al.	BA.2	India		2022/1/18
ON060009.1	25-27	Kumar,S.,et al.	BA.2	India		2022/1/18
OM843171.1	142	Alwasti.H.,et al.	BA.2	Bahrain		2022/2/3
OM843276.1	142	Alwasti.H.,et al.	BA.2.3	Bahrain		2022/2/16
OM843550.1	142	Alwasti.H.,et al.	BA.2	Bahrain		2022/2/16
OM843316.1	142	Alwasti.H.,et al.	BA.2	Bahrain		2022/2/16
OM843340.1	142	Alwasti.H.,et al.	BA.2	Bahrain		2022/2/16
ON049267.1	213	Howard,D.,et al.	BA.2	USA	NC	2022/3/12
ON450720.1	213	Howard,D.,et al.	BA.2	USA	FL	2022/4/25
ON250163.1	339	Howard,D.,et al.	BA.2	USA	MA	2022/3/29
ON256603.1	339	Sadri,N.,et al.	BA.2	USA	OH	2022/4/10
ON480422.1	339	Marques,A.D.,et.al.	BA.2.9	USA	PA	2022/4/3
OM888844.1	371	Howard,D.,et al.	BA.2	USA	NY	2022/2/13
OM890089.1	371	Howard,D.,et al.	BA.2	USA	NY	2022/2/19
ON009425.2	371	Sadri,N.,et al.	BA.2	USA	OH	2022/3/5
ON082904.1	371	Linares-Perdomo,O.J.	BA.2	USA		2022/2/21
OM901275.1	371	Howard,D.,et al.	BA.2.10	USA	MA	2022/2/18
OM877094.2	373	Sadri,N.,et al.	BA.2.10	USA	OH	2022/2/16
OM877095.2	373	Sadri,N.,et al.	BA.2	USA	OH	2022/2/17
OM877096.2	373	Sadri,N.,et al.	BA.2	USA	OH	2022/2/17
OM877097.2	373	Sadri,N.,et al.	BA.2	USA	OH	2022/2/17
ON378542.1	375	Howard,D.,et al.	BA.2	USA	CO	2022/4/14
ON389858.1	375	Howard,D.,et al.	BA.2	USA	CO	2022/4/18
ON389889.1	375	Howard,D.,et al.	BA.2	USA	CO	2022/4/19
ON390359.1	375	Howard,D.,et al.	BA.2	USA	CO	2022/4/22
ON390365.1	375	Howard,D.,et al.	BA.2	USA	CO	2022/4/22
OM936703.1	376	Howard,D.,et al.	BA.2.9	USA	MN	2022/2/25
ON352711.1	376	Reeb,V.,et al.	BA.2	USA	IA	2022/4/12
ON378000.1	376	Howard,D.,et al.	BA.2.3	USA	CO	2022/4/10
ON177702.1	405	Howard,D.,et al.	BA.2	USA	NY	2022/3/22
ON205494.2	405	Linares-Perdomo,O.J.	BA.2.9	USA		2022/3/26
ON378633.1	405	Howard,D.,et al.	BA.2	USA	CO	2022/4/15
ON617689.2	405	Linares-Perdomo,O.J.	BA.2	USA		2022/5/2
ON619375.1	405	Direct Submission	BA.2.9	USA		2022/5/2
OM567618.1	408	Parrott,T.,et al.	BA.2	USA	GA	2022/1/27
OM659585.1	408	Howard,D.,et al.	BA.2.3	USA	MN	2022/1/29
OM770913.1	408	Howard,D.,et al.	BA.2.10	USA	MA	2022/2/5
OM781641.1	408	Howard,D.,et al.	BA.2	USA	MN	2022/2/9
OM805865.1	408	Howard,D.,et al.	BA.2	USA	MN	2022/2/12
OM533441.1	417	Banu,L.A.,et al.	BA.2.10	Bangladesh		2022/1/17
OM533458.1	417	Banu,L.A.,et al.	BA.2	Bangladesh		2022/1/18
OM570235.1	417	Banu,L.A.,et al.	BA.2.10	Bangladesh		2022/1/13
OM570249.1	417	Banu,L.A.,et al.	BA.2	Bangladesh		2022/1/17
OM570252.1	417	Banu,L.A.,et al.	BA.2	Bangladesh		2022/1/17
OM283320.1	440	Alwasti.H.,et al.	BA.2.3	Bahrain		2022/1/7
OM283343.1	440	Alwasti.H.,et al.	BA.2	Bahrain		2022/1/7
OM283361.1	440	Alwasti.H.,et al.	BA.2.10	Bahrain		2022/1/7
OM283362.1	440	Alwasti.H.,et al.	BA.2.10	Bahrain		2022/1/7
OM283368.1	440	Alwasti.H.,et al.	BA.2.3	Bahrain		2022/1/7
ON618014.1	477	Smith,E.	BA.2	USA		2022/4/5
ON618018.1	477	Smith,E.	BA.2.3	USA		2022/4/3
ON618019.1	477	Smith,E.	BA.2	USA		2022/4/3
ON618027.1	477	Smith,E.	BA.2.3	USA		2022/4/1
ON618363.1	477	Smith,E.	BA.2.3	USA		2022/3/24
ON311615.1	484	Howard,D.,et al.	BA.2.10	USA	IL	2022/4/5
ON383919.1	484	Erwin,K., Crowley,S.	BA.2.3	USA		2022/4/18
OP579158.1	484	Huang, A., et al.	BA.2.10	Thailand		2022/7/1
OP054411.1	484	Howard,D.,et al.	BA.2	USA	PA	2022/7/3
ON633107.1	484	Howard,D.,et al.	BA.2.7	USA	TX	2022/5/8
ON414693.1	493	Howard,D.,et al.	BA.2.23	USA	OR	2022/4/22
ON422887.1	493	Howard,D.,et al.	BA.2	USA	FL	2022/4/23
OP364296.1	493	Dean,D., et al.	BA.2	USA	SC	2022/8/3
OP629673.1	493	Tseng,A.C.,et al.	BA.2	USA	HI	2022/6/25
ON363097.1	498	Howard,D.,et al.	BA.2.6	USA	CA	2022/4/9
OP633561.1	498	Cokerton,S.,et al.	BA.2.10	New Zealand		2022/4/24
ON458445.1	498	Howard,D.,et al.	BA.2	USA	CA	2022/4/15
ON592247.1	498	Howard,D.,et al.	BA.2	USA	NJ	2022/5/11
ON549687.1	498	Howard,D.,et al.	BA.2.21	USA	NY	2022/5/3
ON067040.1	501	Lemieux,J.E.,et al.	BA.2	USA	MA	2022/3/16
ON199452.1	501	Lemieux,J.E.,et al.	BA.2.3	USA	NY	2022/3/28
ON200331.1	501	Lemieux,J.E.,et al.	BA.2	USA	MA	2022/3/24
ON200438.1	501	Lemieux,J.E.,et al.	BA.2.10	USA	MA	2022/3/24
ON321116.1	501	Lemieux,J.E.,et al.	BA.2	USA	VT	2022/4/1
OM861064.1	505	Howard,D.,et al.	BA.2.10	USA	CA	2022/2/15
OM969592.1	505	Howard,D.,et al.	BA.2	USA	FL	2022/3/1
ON019120.1	505	Howard,D.,et al.	BA.2.3	USA	CA	2022/3/6
ON221860.1	505	Ye,D.Z., Liu,Y.	BA.2	USA		2022/3/17
ON221861.1	505	Ye,D.Z., Liu,Y.	BA.2	USA		2022/3/17
OM861619.1	614	Howard,D.,et al.	BA.2	USA	TX	2022/2/16
ON091288.1	614	Howard,D.,et al.	BA.2	USA	PA	2022/3/13
ON151370.1	614	Howard,D.,et al.	BA.2	USA	CT	2022/3/22
ON233850.1	614	Howard,D.,et al.	BA.2	USA	NJ	2022/4/2
ON236456.1	614	Howard,D.,et al.	BA.2	USA	CT	2022/3/31
ON296711.1	655	Howard,D.,et al.	BA.2.9	USA	CT	2022/4/8
ON535443.1	655	Howard,D.,et al.	BA.2.9	USA	NY	2022/5/1
ON624524.1	655	Howard,D.,et al.	BA.2	USA	NY	2022/5/9
ON377450.1	655	Howard,D.,et al.	BA.2	USA	NJ	2022/4/14
ON397268.1	655	Howard,D.,et al.	BA.2.21	USA	DC	2022/4/19
ON239032.1	679	Bodnar,L.et al.	BA.2	USA		2022/4/2
ON373649.1	679	Howard,D.,et al.	BA.2.9	USA	WA	2022/4/16
ON481637.1	679	Howard,D.,et al.	BA.2	USA	CA	2022/4/29
ON701163.1	679	Howard,D.,et al.	BA.2	USA	MD	2022/5/21
OM948079.1	679	Howard,D.,et al.	BA.2	USA	MI	2022/2/27
ON312677.2	764	Linares-Perdomo,O.J.	BA.2.33	USA		2022/4/6
ON349263.1	764	Howard,D.,et al.	BA.2	USA	NC	2022/4/11
ON377487.1	764	Howard,D.,et al.	BA.2	USA	DC	2022/4/14
ON377609.1	764	Howard,D.,et al.	BA.2	USA	VA	2022/4/15
OM989887.1	764	Howard,D.,et al.	BA.2.37	USA	VA	2022/3/1
OM638574.1	796	Banu,L.A.,et al.	BA.2.10	Bangladesh		2022/1/13
OM911616.1	796	Howard,D.,et al.	BA.2	USA	MA	2022/2/22
OM988767.1	796	Alwasti.H.,et al.	BA.2	Bahrain		2022/3/9
ON019770.1	796	Howard,D.,et al.	BA.2	USA	KS	2022/3/8
OM988769.1	796	Alwasti.H.,et al.	BA.2	Bahrain		2022/3/9
ON468158.1	954	Howard,D.,et al.	BA.2	USA	NC	2022/4/28
ON608924.1	954	Howard,D.,et al.	BA.2.9	USA	NC	2022/5/13
ON604965.1	954	Sadri,N.,et al.	BA.2	USA	OH	2022/5/16
ON535763.1	954	Howard,D.,et al.	BA.2	USA	CA	2022/5/2
ON378227.1	969	Howard,D.,et al.	BA.2.9	USA	CO	2022/4/12
ON378238.1	969	Howard,D.,et al.	BA.2.9	USA	GA	2022/4/12
ON728470.1	969	Howard,D.,et al.	BA.2	USA	VA	2022/5/27

Supplemental Table 2. The counts of sequences in each state. The probability for each state is based on the cumulative value of binomial distribution using the probability given by the total data. One, two, and three stars mean probabilities less than 10^{-2}, 10^{-4}, and 10^{-6} respectively.

B.1.617.2

State	All	G614D	Ratio	Probability	
AL	175	0	0.00E+00	1.00E+00	
AK	88	2	2.27E-02	2.34E-01	
AZ	733	4	5.46E-03	9.47E-01	
AR	936	1	1.07E-03	1.00E+00	
CA	5009	17	3.39E-03	1.00E+00	
CO	715	10	1.40E-02	2.18E-01	
CT	228	11	4.82E-02	3.42E-05	**
DE	85	1	1.18E-02	5.90E-01	
DC	347	5	1.44E-02	2.97E-01	
FL	5063	16	3.16E-03	1.00E+00	
GA	1041	4	3.84E-03	9.95E-01	
HI	48	1	2.08E-02	3.96E-01	
ID	184	3	1.63E-02	3.02E-01	
IL	606	24	3.96E-02	5.00E-08	***
IN	374	6	1.60E-02	1.99E-01	
IA	183	1	5.46E-03	8.53E-01	
KS	161	4	2.48E-02	8.92E-02	
KY	260	4	1.54E-02	2.88E-01	
LA	133	4	3.01E-02	5.14E-02	
ME	38	2	5.26E-02	5.98E-02	
MD	784	12	1.53E-02	1.24E-01	
MA	1291	10	7.75E-03	8.64E-01	
MI	535	25	4.67E-02	1.02E-09	***
MN	2040	20	9.80E-03	6.40E-01	
MS	181	0	0.00E+00	1.00E+00	
MO	293	3	1.02E-02	5.91E-01	
MT	51	2	3.92E-02	9.93E-02	
NE	82	4	4.88E-02	1.09E-02	
NV	434	6	1.38E-02	3.02E-01	
NH	106	2	1.89E-02	3.03E-01	
NJ	901	15	1.66E-02	5.50E-02	
NM	335	7	2.09E-02	6.39E-02	
NY	1141	10	8.76E-03	7.51E-01	
NC	1419	18	1.27E-02	2.34E-01	
ND	18	0	0.00E+00	1.00E+00	
OH	491	16	3.26E-02	8.10E-05	**
OK	237	1	4.22E-03	9.17E-01	
OR	188	0	0.00E+00	1.00E+00	
PA	563	8	1.42E-02	2.38E-01	
RI	198	0	0.00E+00	1.00E+00	
SC	402	2	4.98E-03	9.23E-01	
SD	75	1	1.33E-02	5.45E-01	
TN	721	8	1.11E-02	4.79E-01	
TX	2035	28	1.38E-02	8.98E-02	
UT	317	5	1.58E-02	2.38E-01	
VT	64	1	1.56E-02	4.89E-01	
VA	880	6	6.82E-03	8.96E-01	
WA	708	1	1.41E-03	9.99E-01	
WV	341	5	1.47E-02	2.85E-01	
WI	501	11	2.20E-02	1.78E-02	
WY	34	1	2.94E-02	3.00E-01	
Total	33773	348	1.03E-02		

BA.2

State	All	G614D	Ratio	Probability	
AL	175	0	0.00E+00	1.00E+00	
AK	328	0	0.00E+00	1.00E+00	
AZ	1345	0	0.00E+00	1.00E+00	
AR	406	1	2.46E-03	4.99E-01	
CA	8368	4	4.78E-04	1.00E+00	
CO	1393	0	0.00E+00	1.00E+00	
CT	1701	9	5.29E-03	2.96E-03	*
DE	198	2	1.01E-02	4.52E-02	
DC	801	4	4.99E-03	4.92E-02	
FL	7127	5	7.02E-04	9.93E-01	
GA	1398	1	7.15E-04	9.07E-01	
HI	464	0	0.00E+00	1.00E+00	
ID	67	0	0.00E+00	1.00E+00	
IL	2795	3	1.07E-03	8.53E-01	
IN	636	1	1.57E-03	6.61E-01	
IA	497	0	0.00E+00	1.00E+00	
KS	184	0	0.00E+00	1.00E+00	
KY	337	0	0.00E+00	1.00E+00	
LA	379	0	0.00E+00	1.00E+00	
ME	398	0	0.00E+00	1.00E+00	
MD	1933	8	4.14E-03	1.92E-02	
MA	8625	4	4.64E-04	1.00E+00	
MI	2506	0	0.00E+00	1.00E+00	
MN	1889	0	0.00E+00	1.00E+00	
MS	170	0	0.00E+00	1.00E+00	
MO	628	1	1.59E-03	6.56E-01	
MT	39	0	0.00E+00	1.00E+00	
NE	482	0	0.00E+00	1.00E+00	
NV	702	1	1.42E-03	6.97E-01	
NH	427	0	0.00E+00	1.00E+00	
NJ	6021	28	4.65E-03	3.32E-06	**
NM	537	0	0.00E+00	1.00E+00	
NY	4112	22	5.35E-03	4.26E-06	**
NC	2784	12	4.31E-03	3.52E-03	*
ND	38	0	0.00E+00	1.00E+00	
OH	1385	1	7.22E-04	9.05E-01	
OK	130	0	0.00E+00	1.00E+00	
OR	871	2	2.30E-03	4.35E-01	
PA	3166	6	1.90E-03	4.50E-01	
RI	1253	3	2.39E-03	3.58E-01	
SC	627	2	3.19E-03	2.88E-01	
SD	82	0	0.00E+00	1.00E+00	
TN	325	2	6.15E-03	1.06E-01	
TX	3542	2	5.65E-04	9.83E-01	
UT	156	0	0.00E+00	1.00E+00	
VT	987	0	0.00E+00	1.00E+00	
VA	3085	12	3.89E-03	7.69E-03	*
WA	6111	4	6.55E-04	9.92E-01	
WV	652	0	0.00E+00	1.00E+00	
WI	1333	2	1.50E-03	6.61E-01	
WY	2	0	0.00E+00	1.00E+00	
Total	83595	142	1.70E-03		