

# 1 A repeat pattern of founder events for 2 SARS-CoV-2 variants in Alaska

3  
4 Tracie J. Haan<sup>1</sup>, Lisa K. Smith<sup>2</sup>, Stephanie DeRonde<sup>2</sup>, Elva House<sup>2</sup>, Jacob Zidek<sup>2</sup>, Diana  
5 Puhak<sup>2</sup>, Logan Mullen<sup>1</sup>, Matthew Redlinger<sup>3</sup>, Jayme Parker<sup>2</sup>, Brian M. Barnes<sup>1</sup>, Jason L.  
6 Burkhead<sup>3</sup>, Cindy Knall<sup>4</sup>, Eric Bortz<sup>3,4,\*</sup>, Jack Chen<sup>1,2,5,\*</sup>, and Devin M. Drown<sup>1,5\*</sup>

7  
8 <sup>1</sup> Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, Alaska, USA

9 <sup>2</sup> Alaska Division of Public Health, State of Alaska, Fairbanks, Alaska, USA

10 <sup>3</sup> Department of Biological Sciences, University of Alaska Anchorage, Anchorage, Alaska, USA

11 <sup>4</sup> WWAMI School of Medical Education, University of Alaska Anchorage, Anchorage, Alaska,  
12 USA

13 <sup>5</sup> Department of Biology and Wildlife, University of Alaska Fairbanks, Fairbanks, Alaska, USA

14

15 \* Corresponding authors: Devin M. Drown, [dmdrown@alaska.edu](mailto:dmdrown@alaska.edu); Eric Bortz  
16 [ebortz@alaska.edu](mailto:ebortz@alaska.edu); Jack Chen, [jchen15@alaska.edu](mailto:jchen15@alaska.edu)

17

18 Running title: Repeat founder events in Alaska

19 Key Words: SARS-CoV-2; COVID-19; genomic epidemiology; variant; Variant of Concern  
20 (VOC)

## 21 Abstract

22 Alaska is a unique US state because of its large size, geographically disparate  
23 population density, and physical distance from the contiguous United States. Here, we describe  
24 a pattern of SARS-CoV-2 variant emergence across Alaska reflective of these differences.  
25 Using genomic data, we found that in Alaska the Omicron sublineage BA.2.3 overtook BA.1.1  
26 by the week of 2022-02-27, reaching 48.5% of sequenced cases. On the contrary in the  
27 contiguous United States, BA.1.1 dominated cases for longer, eventually being displaced by  
28 BA.2 sublineages other than BA.2.3. BA.2.3 only reached a prevalence of 10.9% in the  
29 contiguous United States. Using phylogenetics, we found evidence of potential origins of the two  
30 major clades of BA.2.3 in Alaska and with logistic regression estimated how it emerged and  
31 spread throughout the state. The combined evidence is suggestive of founder events in Alaska  
32 and is reflective of how Alaska's unique dynamics influence the emergence of SARS-CoV-2  
33 variants.

## 34 Introduction

35 Throughout the coronavirus disease 2019 (COVID-19) pandemic, variants of severe acute  
36 respiratory syndrome coronavirus 2 (SARS-CoV-2) have repeatedly emerged and spread, often  
37 circulating globally over a relatively short timeframe (Mullen et al., 2020). These variants  
38 frequently underwent mutations affecting viral phenotypes, such as increased transmissibility or  
39 immune escape, which contributed to epidemic waves of cases and hospitalizations occurring  
40 asynchronously across different regions at varying severities (Alcantara et al., 2022). The  
41 sequential wave dynamics of COVID-19 have likely been influenced by a multitude of  
42 epidemiological factors, such as host immunity and vaccination coverage, social measures aimed  
43 at suppressing spread, and the viral characteristics, including transmissibility and the moderately  
44 higher mutation rate for an RNA virus (Saha et al., 2021; Kirby, 2021; Tao, 2021).

45 For many regions of the world, the first notable COVID-19 epidemic wave attributed to a  
46 variant occurred near the end of 2020 into early 2021. During this wave, the Alpha variant (lineage  
47 B.1.1.7; Rambaut et al., 2020), which showed evidence of increased transmissibility, became the  
48 most prevalent variant for most places globally (Volz et al., 2021). Unlike other regions, including  
49 within the contiguous 48 states of the United States (hereafter referred to as Lower 48), Alaska's  
50 dominant lineage was B.1.1.519 throughout early 2021, which was similar to Mexico in late 2020  
51 (Haan et al., 2022; Rodríguez-Maldonado et al., 2021). The timeline in which Alpha and B.1.1.519  
52 emerged in Alaska paired with the striking difference in prevalence between Alaska, which had a  
53 peak B.1.1.519 prevalence of 77.9%, and the Lower 48, which had a peak prevalence of 4.9%,  
54 was indicative of a B.1.1.519 founder event in Alaska (Haan et al., 2022). Since this initial  
55 deviation from the Lower 48, Alaska has displayed similar patterns of variant emergence and  
56 spread both with the sweep of the Delta variant in early August 2021 followed by the sweep of  
57 Omicron beginning in December 2021 (CDC, 2021).

58 Although the recent emergence of Omicron in Alaska was initially similar to that of the  
59 Lower 48, the emergence of sublineages within larger variant classifications has been distinct.  
60 These Omicron sublineages have shown a great degree of divergence that has led to concerns  
61 about antibody evasion and the possibility of repeat infections of Omicron sublineages (Stegger  
62 et al., 2022). Mutations include a spike protein (S) R346K alteration in BA.1.1 (also known as  
63 B.1.1.529.1.1) and 8 unique S mutations in BA.2, which lacks 13 S alterations found in BA.1. The  
64 Omicron sublineage reported in this study, BA.2.3, encodes several key amino acid changes  
65 including A2909V in ORF1a and L140F in ORF3a. These changes may have provided growth  
66 advantages such as antibody evasion and increased reproductive rate that allowed Omicron  
67 lineages to displace Delta initially (Iketani et al., 2022; Smith et al., 2022).

68 Here, we use genomic data readily available via the Global Initiative on Sharing All  
69 Influenza Data (GISAID) to describe the pattern of emergence and spread of the Omicron  
70 sublineage BA.2.3 (also known as B.1.1.529.2.3) in Alaska. We contrast this pattern with  
71 observations in the Lower 48 and several major US states including California, New York, and  
72 Washington. Using data generated through genomic surveillance efforts, we explore the pattern  
73 in Omicron sublineages of Alaska similar to that of the founder event that occurred with B.1.1.519  
74 in early 2021 (Haan et al., 2022). These repetitive patterns of variant emergence are suggestive  
75 of repeat founder events in Alaska.

## 76 Materials and Methods

### 77 *Retrieving and Analyzing SARS-CoV-2 Sequence Data for Alaska*

78 On May 3rd, 2022, we downloaded 11,971 sequences from Alaskan samples available on  
79 GISAID for subsequent analysis (Shu & McCauley, 2017; Elbe & Merret, 2017). This readily  
80 available genomic data was in part generated by the *Alaska SARS-CoV-2 Sequencing*  
81 *Consortium*. The Consortium is a partnership between the University of Alaska and the Alaska  
82 Division of Public Health (AKDPH) with the aim to increase genomic surveillance of SARS-CoV-  
83 2 variants. Genome sequencing in Alaska is from a non-targeted sample of cases, which is the  
84 best available approximation of random samples despite potential disparate coverage across  
85 Alaska's economic regions. We subset these data to the study date of 2021-11-29 through 2022-  
86 04-02 for a total of 3,132 genomes to estimate the prevalence of lineages per week on dates  
87 beginning on Sunday of each week.

88 Lineages were determined by running sequences through PANGO v1.8, Pangolin v4.0.6,  
89 and pangoleARN v1.2.133, and Scorpio v0.3.17 (O'Toole et al., 2021). We estimated the  
90 prevalence of genomes in Alaska from the date of Omicron's (B.1.1.529) first detection in Alaska  
91 on 2021-11-28 through 2022-04-03. All AY sublineages are aggregated into the group B.1.617.2  
92 (Delta). All BA sublineages of Omicron except BA.1.1, BA.2, BA.2.3, and their sublineages are  
93 aggregated into B.1.1.529 (Omicron). Genomes that did not fall into these lineages were grouped  
94 together into the category 'Not Emerging Lineage.'

### 95 *SARS-CoV-2 Sequence Data for the Lower 48*

96 On April 12th, 2022, we downloaded metadata, including lineage assignment using the  
97 same version of Pangolin (v4.0.6) for all sequences available on GISAID and filtered for  
98 sequences from the United States of America (USA) collected between 2021-11-28 to 2022-04-

99 12. We removed cases from Alaska, Hawaii, and US territories to limit our comparisons to the  
100 lower 48 contiguous states. We also analyzed data at a state level for New York, California, and  
101 Washington. GISAID metadata was used to calculate the prevalence of variants, which was an  
102 approximation based on percent of sequenced cases each day.

### 103 *Visualizations, statistical analyses, and Nextstrain build*

104 We generated visualizations in RStudio (v 1.4.1106) using packages ggplot2 (v 3.3.5),  
105 ggpubr (v 0.4.0), tidyverse (v 1.3.1), and lubridate (v 1.7.10). A generalized linear model using  
106 the logit link function with the base R stats (v 3.6.2) glm tool was used to generate estimates of  
107 prevalence of BA.2.3 over time for Alaska and two economic regions; the Anchorage-Mat Su and  
108 Gulf Coast. For these models, the daily percent of sequenced cases assigned to BA.2.3 (i.e., daily  
109 prevalence of BA.2.3) was used as the dependent variable with time from 2022-01-01 through  
110 2022-04-02 used as the independent variable. Regressions were plotted using geom\_smooth in  
111 ggplot2. We generated a Nextstrain (cli v-3.2.4) build to examine the phylogenetic relationship of  
112 BA.2.3 in Alaska compared to global sequences. We generated this tree using GISAID's Global  
113 Nextregions for context, all BA.2.3 cases from Alaska, and global cases of BA.2.3 from 2021-12-  
114 02 through 2022-02-06 that included all cases from before 2022-01-01 and then downsampled to  
115 a fifth of the sequences randomly after that date. We colored tree tips by countries of significance  
116 including the Philippines, South Africa, Japan, India, and the USA split by 'USA- Alaska' and  
117 'USA- Other.' Other country's cases were masked from the tree visualization.

## 118 Results and Discussion

### 119 *Higher prevalence of Omicron lineage BA.2.3 in Alaska versus the Lower 48*

120 To examine how the emergence of BA.2.3 differs between Alaska and the Lower 48, we  
121 determined the date of first detection and prevalence of BA.2.3 over time in both locations. The  
122 first Alaska case assigned to Omicron was detected in the Anchorage-Mat Su region, the most  
123 populated region of Alaska, on 2021-11-28. Within four weeks of first detection (by the week of  
124 2021-12-19), Omicron had outcompeted Delta in terms of prevalence both in Alaska and the  
125 Lower 48 (Figure 1). By the week of 2022-01-16, Delta was detected in less than 1% of sequenced  
126 cases for both Alaska and the Lower 48. Omicron cases during this week were dominated by the  
127 sublineage BA.1.1 in both Alaska, at 66.3% prevalence, and the Lower 48, at 67.2% prevalence  
128 (Figure 1). While BA.1.1 was dominant the week of 2022-01-16 in both locations, BA.2 and  
129 sublineages were just starting to be detected in the United States. By the week of 2022-01-16, in  
130 the Lower 48 BA.2 comprised only 0.2% of sequenced cases whereas in Alaska no cases of BA.2

131 had been detected. However, the sublineage BA.2.3 was found in 2.7% of sequenced cases in  
132 Alaska by this week whereas in the Lower 48 BA.2.3 represented only 0.1% of cases. By 2022-  
133 02-27, BA.2.3 comprised the majority of cases in Alaska (45.3%) compared to 6.1% in the Lower  
134 48. At the same time, BA.2 comprised 9.4% of Lower 48 cases and only 2.5% of Alaska's cases.  
135 Although by March BA.2.3 started increasing in prevalence in the Lower 48, BA.2 already  
136 displaced BA.1.1, the previously dominant lineage. By the last week of March, 2022, BA.2.3  
137 comprised 74.3% of cases in Alaska and 19.2% of cases in the Lower 48. These stark differences  
138 in prevalence over time reflect the divergent patterns of emergence of BA.2.3 in Alaska versus  
139 the Lower 48.

140         Given the Lower 48 is an aggregate of many distinct, yet connected, communities, we  
141 examined the prevalence of sublineages at a finer geographic scale. This finer scale was at a  
142 state level for several populous US states including California, Washington, and New York. Each  
143 of these states reflected a similar pattern in BA sublineages as the overall Lower 48 with a low  
144 prevalence of BA.2.3 compared to Alaska over the study time period. The week of 2022-02-13  
145 BA.2.3 already comprised 25.3% of sequenced cases in Alaska and only 1.8% in the Lower 48,  
146 1.8% in New York, 2.2% in California, and 2.4% in Washington (Figure 2). The week of 2022-03-  
147 13 when BA.2.3 comprised a majority of the cases sequenced in Alaska at 67.9% whereas the  
148 other states had much lower and variable prevalence in the Lower 48 (10.9%), New York (5.5%),  
149 California (16.2%), and in Washington (17.6%) (Figure 2). By the last week of March, the only  
150 other state with a BA.2.3 prevalence greater than 40% was California.

151         Although selective advantages, such as transmission potential, posed by SARS-CoV-2  
152 variants have played a key role in their emergence over the course of the pandemic, changes in  
153 variant prevalence can also be attributed to founder effects (Attwood et al., 2022). In the context  
154 of SARS-CoV-2, and other viral pathogens, founder effects result from a chance colonization  
155 event allowing a new population of viral lineages to emerge. When a chance colonization event  
156 occurs, the growth of a new population of viral lineages can give the impression that one lineage  
157 has a growth advantage over others (Rambaut et al., 2004; Ruan et al., 2021). The apparent  
158 difference in the emergence of BA.2.3 between Alaska, the Lower 48, and other US states  
159 highlights a potential founder effect in which BA.2.3 is acting as the founding sublineage in Alaska.  
160 It may have become dominant here and not in other locations because of the timing of emergence  
161 and social factors that rendered Alaska communities as a naïve population susceptible to infection  
162 by BA.2.3.

163         The composition of SARS-CoV-2 lineages within Alaska and the Lower 48, each  
164 containing distinct sets of mutations that define them, were distinct at the time of BA.2.3's

165 emergence in Alaska. The main difference in community composition was the presence of other  
166 BA.2 lineages in the Lower 48 and their absence in Alaska (Figure 1). The absence of other BA.2  
167 lineages in Alaska could have allowed for the founding of BA.2.3 in the population, similar to how  
168 other lineages with specific mutations emerged and became dominant in other locations  
169 throughout the pandemic (Ozer et al., 2021; Hodcroft et al., 2021). For example, in the summer  
170 of 2020, the 20E lineage of SARS-CoV-2, which had no evidence of increased transmissibility,  
171 became the dominant lineage in Europe, likely driven by its founder event in the population paired  
172 with the increased connectivity across Europe from travel over the summer months (Hodcroft et  
173 al., 2021). It was also suggested the emergence of variant of concern (VOC) Alpha (B.1.1.7) and  
174 the associated mutations, like D614G, in part could have been driven by the founder effect (Ozer  
175 et al., 2021; Tang et al., 2021). This was suggested because of the inconclusive results over the  
176 positive selection of those mutations and coinciding timing with the nexus of dispersal from Asia  
177 to Europe associated with the D614G mutation (Grubaugh et al., 2020). Given the emergence of  
178 other lineages in the Lower 48 even though BA.2.3 was detected around the same time for many  
179 locations, Alaska's emergence of BA.2.3, and B.1.1.519 earlier in the pandemic, implicates repeat  
180 occurrences of variant emergence influenced by the founder effect (Haan et al., 2022).

### 181 *Modeling shows variable emergence of BA.2.3 across Alaska*

182 Spatiotemporal variation in the emergence and spread of SARS-CoV-2 lineages has been  
183 observed at broad geographic levels. The CDC has reported on these regional variations by  
184 dividing the United States into ten regions that show distinct communities based on genomic  
185 surveillance (CDC, 2022). Here, using genomic surveillance data from Alaska, we found within-  
186 state variation in the emergence of BA.2.3 between major economic regions of Alaska (Figure 2;  
187 Alaska Department of Labor and Workforce Development, 2021). In Alaska there are six  
188 economic regions defined by the Department of Labor and Workforce Development: the  
189 Anchorage-Mat Su, Interior, Gulf Coast, Southeast, Southwest, and Northern regions in order  
190 from highest to lowest population. BA.2.3 was first detected from two cases collected on 2022-  
191 01-11 in the Gulf Coast and the Anchorage-Mat Su regions of Alaska. The Gulf Coast is the third  
192 most populous region of Alaska and just south of the most populated region, the Anchorage-Mat  
193 Su. While many economic regions across Alaska are only connected by air or boat transportation,  
194 the Gulf Coast and Anchorage-Mat Su regions are broadly connected via Alaska's road system.  
195 When examining a model estimate of BA.2.3 prevalence over time, we found that the Gulf Coast  
196 region had the earliest emergence. For the state as a whole, prevalence was estimated to be at  
197 greater than 5% the week of 2022-01-19 (Figure 3A), the Anchorage-Mat Su didn't reach 5% until

198 2022-01-25 (Figure 3B), and the Gulf Coast was estimated to reach greater than 5% prevalence  
199 on 2022-01-03, which was before BA.2.3 was even detected (Figure 3C). Having the model  
200 indicate 5% prevalence before first detection suggests BA.2.3 could have been present in the  
201 Gulf Coast region of Alaska before sequencing captured a case of BA.2.3; however, model  
202 uncertainty, indicated by the shading, is also consistent with BA.2.3 being absent until the first  
203 actual case was detected. In the Gulf Coast, BA.2.3 was estimated to comprise the majority of  
204 cases by 2022-02-16, whereas for the state as a whole this didn't occur until weeks later, 2022-  
205 03-07 and for the Anchorage-Mat Su region this didn't occur until 2022-03-11.

### 206 *Phylogenetics of BA.2.3 provides evidence of multiple introduction events*

207 Based on global sequence data of BA.2.3 available on GISAID, we found that within  
208 Alaska there are two clades of BA.2.3 comprising the majority of Alaska's cases. These two clades  
209 both appear to have emerged from cases originally detected in the Philippines, where BA.2.3 was  
210 first detected on 2021-12-02 (Figure 4). When considering cases by economic region of Alaska,  
211 there is no evidence that the two clades were introduced to each economic region independently  
212 given the interspersed nature of the cases (Figure 5). This suggests that there was mixing of  
213 cases between the two regions, rather than two separate introduction events. In other regions of  
214 the United States, BA.2.3 cases appear to have emerged from both the Philippines and a clade  
215 where South Africa and India cases appear to be dominant early in the tree. (Figure 4). By scaling  
216 the branch length by divergence, or number of mutations, we show how the majority of Alaska  
217 cases diverged from the early Philippines cases accumulating mutations with further spread in  
218 Alaska (Figure 4).

## 219 Conclusions

220 Using genomic data available in the GISAID repository, we demonstrated the unique  
221 emergence and spread patterns of the SARS-CoV-2 sublineage of Omicron BA.2.3 in Alaska  
222 compared to the lower 48 contiguous states of the United States. Looking at a finer scale with  
223 several major US states, the same stark difference in prevalence of BA.2.3 was observed  
224 further highlighting the unique occurrence of BA.2.3 in Alaska. Our phylogenetic analysis paired  
225 with logistic regression revealed the potential ancestral origins of BA.2.3 cases in Alaska and  
226 how these clades within BA.2.3 emerged and spread by economic region of the state. These  
227 repetitive patterns of variant emergence with B.1.1.519 followed by BA.2.3 in Alaska are  
228 suggestive of repeat founder events which are reflective of how Alaska's unique location  
229 influences the emergence of distinct SARS-CoV-2 variants

## 230 Data Availability Statement

231 All data used in this study are available online through the SARS-CoV-2 repository,  
232 global initiative on sharing all influenza data (GISAID). These findings are based on analysis of  
233 approximately 4,490 genomes accessible via EPI\_SET\_20220517as and 1,009,539 accessible  
234 via EPI\_SET\_20220517ge. Accession numbers for genomes of Alaska cases, metadata for the  
235 United States of America, and Global Nextregion data retrieved from the GISAID can be found  
236 using the EPI-SET identifiers at <https://www.gisaid.org/> .

## 237 Acknowledgments

238 We gratefully acknowledge all the researchers responsible for obtaining specimens and  
239 laboratories where genetic sequence data were generated and shared via the GISAID Initiative  
240 (<https://www.gisaid.org>), on which this research is based. Work presented here was supported  
241 by Epidemiology and Laboratory Capacity for Prevention and Control of Emerging Infectious  
242 Diseases (ELC) from the Centers for Disease Control and Prevention and a supplement award  
243 to Alaska INBRE, an Institutional Development Award (IDeA) from the National Institute of  
244 General Medical Sciences of the National Institutes of Health under grant number  
245 2P20GM103395.

## 246 References.

- 247 Alaska Department of Labor and Workforce Development. (2021). *2020 Population Estimates*  
248 *by Borough, Census Area, and Economic Region*. Nov 5, 2021.  
249 <https://live.laborstats.alaska.gov/pop/index.cfm>
- 250 Alcantara, L. C. J., Nogueira, E., Shuab, G., Tosta, S., Fristch, H., Pimentel, V., ... & Giovanetti,  
251 M. (2022). SARS-CoV-2 epidemic in Brazil: how variants displacement have driven  
252 distinct epidemic waves. *Virus Research*, 198785.
- 253 Attwood, S. W., Hill, S. C., Aanensen, D. M., Connor, T. R., & Pybus, O. G. (2022).  
254 Phylogenetic and phylodynamic approaches to understanding and combating the early  
255 SARS-CoV-2 pandemic. *Nature Reviews Genetics*, 1-16.
- 256 Centers for Disease Control and Prevention. (2021). *Science brief: Omicron (B.1.1.529) variant*,  
257 *CDC COVID-19 science briefs*. CDC, Atlanta, GA.
- 258 Centers for Disease Control and Prevention. (2022). *COVID Data Tracker*. Atlanta, GA: US  
259 Department of Health and Human Services, CDC; 2022, May 11.  
260 <https://covid.cdc.gov/covid-data-tracker>

- 261 Elbe, S., & Buckland-Merrett, G. (2017). Data, disease and diplomacy: GISAID's innovative  
262 contribution to global health. *Global challenges* (Hoboken, NJ), 1(1), 33–46.
- 263 Grubaugh, N. D., Hanage, W. P., & Rasmussen, A. L. (2020). Making sense of mutation: what  
264 D614G means for the COVID-19 pandemic remains unclear. *Cell*, 182(4), 794-795.
- 265 Haan, T. J., Smith, L. K., DeRonde, S., House, E., Zidek, J., Puhak, D., ... & Drown, D. M.  
266 (2022). Founder effect contributes to the unique pattern of SARS-CoV-2 variant B. 1.1.  
267 519 emergence in Alaska. *medRxiv*.
- 268 Iketani, S., Liu, L., Guo, Y., Liu, L., Chan, J.F.W., Huang, Y., Wang, M., Luo, Y., Yu, J., Chu, H.  
269 and Chik, K.K.H. (2022). Antibody evasion properties of SARS-CoV-2 Omicron  
270 sublineages. *Nature*, 1-4.
- 271 Khare, S., et al (2021) GISAID's Role in Pandemic Response. *China CDC Weekly*, 3(49): 1049-  
272 1051. doi: 10.46234/ccdcw2021.255 PMID: 8668406
- 273 Kirby, T. (2021). New variant of SARS-CoV-2 in UK causes surge of COVID-19. *The Lancet*  
274 *Respiratory Medicine*, 9(2), e20-e21.
- 275 Mullen, J.L., Tsueng, G., Latif, A.A., Alkuzweny, M., Cano, M., Haag, E., Zhou, J., Zeller, M.,  
276 Hufbauer, E., Matteson, N., Andersen, K.G., Wu, C., Su, A.I., Gangavarapu, K., Hughes,  
277 L.D. (2020) Center for Viral Systems Biology outbreak.info. Available online:  
278 <https://outbreak.info/>
- 279 O'Toole, Á., Scher, E., Underwood, A., Jackson, B., Hill, V., McCrone, J. T., & Rambaut,  
280 A.(2021). Assignment of epidemiological lineages in an emerging pandemic using the  
281 pangolin tool. *Virus Evolution*, 7(2), veab064.
- 282 Ozer, E.A., Simons, L.M., Adewumi, O.M., Fowotade, A.A., Omoruyi, E.C., Adeniji, J.A., Dean,  
283 T.J., Taiwo, B.O., Hultquist, J.F. and Lorenzo-Redondo, R. (2021). High prevalence of  
284 SARS-CoV-2 B. 1.1. 7 (UK variant) and the novel B. 1.525 lineage in Oyo State, Nigeria.  
285 *Medrxiv*.
- 286 Rambaut, A., Holmes, E. C., O'Toole, Á., Hill, V., McCrone, J. T., Ruis, C., du Plessis, L. &  
287 Pybus, O. G.(2020). A dynamic nomenclature proposal for SARS-CoV-2 lineages to  
288 assist genomic epidemiology. *Nature microbiology*, 5(11), 1403-1407.
- 289 Rambaut, A., Posada, D., Crandall, K. A., & Holmes, E. C. (2004). The causes and  
290 consequences of HIV evolution. *Nature Reviews Genetics*, 5(1), 52-61.
- 291 Rodríguez-Maldonado, A. P., Vázquez-Pérez, J. A., Cedro-Tanda, A., Taboada, B., Boukadida,  
292 C., Wong-Arámbula, C., Nuñez-García, T. E., Cruz-Ortiz, N., Barrera-Badillo, G.,  
293 Hernández-Rivas, L., López-Martínez, I., Mendoza-Vargas, A., Reyes-Grajeda, J. P.,  
294 Alcaraz, N., Peñalosa-Figueroa, F., Gonzalez-Barrera, D., Rangel-DeLeon, D., Herrera-  
295 Montalvo, L. A., Mejía-Nepomuceno, F., Hernández-Terán, A., ... Ramírez-González,  
296 J.E. (2021). Emergence and spread of the potential variant of interest (VOI) B.1.1.519 of

- 297 SARS-CoV-2 predominantly present in Mexico. *Archives of virology*, 166(11), 3173–  
298 3177.
- 299 Ruan, Y., Luo, Z., Tang, X., Li, G., Wen, H., He, X., Lu, X., Lu, J. & Wu, C.I. (2021). On the  
300 founder effect in COVID-19 outbreaks: how many infected travelers may have started  
301 them all?. *National Science Review*, 8(1), nwaa246.
- 302 Saha, S., Tanmoy, A. M., Tanni, A. A., Goswami, S., Al Sium, S. M., Saha, S., & Saha, S. K.  
303 (2021). New waves, new variants, old inequity: a continuing COVID-19 crisis. *BMJ*  
304 *Global Health*, 6(8), e007031.
- 305 Shu, Y., & McCauley, J. (2017). GISAID: Global initiative on sharing all influenza data—from  
306 vision to reality. *Eurosurveillance*, 22(13), 30494.
- 307 Stegger, M., Edslev, S.M., Sieber, R.N., Ingham, A.C., Ng, K.L., Tang, M.H.E., Alexandersen,  
308 S., Fonager, J., Legarth, R., Utko, M. and Wilkowski, B. (2022). Occurrence and  
309 significance of Omicron BA. 1 infection followed by BA. 2 reinfection. *medRxiv*.
- 310 Tao, K., Tzou, P.L., Nouhin, J., Gupta, R.K., de Oliveira, T., Kosakovsky Pond, S.L., Fera, D.  
311 and Shafer, R.W. (2021). The biological and clinical significance of emerging SARS-  
312 CoV-2 variants. *Nature Reviews Genetics*, 22(12), 757-773.
- 313 Tang, J. W., Tambyah, P. A., & Hui, D. S. (2021). Emergence of a new SARS-CoV-2 variant in  
314 the UK. *Journal of Infection*, 82(4), e27-e28.
- 315 Volz, E., Mishra, S., Chand, M., Barrett, J.C., Johnson, R., Geidelberg, L., Hinsley, W.R.,  
316 Laydon, D.J., Dabrera, G., O'Toole, Á. and Amato, R. (2021). Assessing transmissibility  
317 of SARS-CoV-2 lineage B. 1.1. 7 in England. *Nature*, 593(7858), 266-269.

## 318 Figure Legends

319 **Figure 1.** The percent of sequences by week (estimated prevalence) colored by SARS-CoV-2  
320 lineages detected from 2021-11-28 to 2022-04-03 in (A) Alaska and (B) the Lower 48. BA  
321 lineages of Omicron except BA.1.1, BA.2, and BA.2.3 are aggregated into B.1.1.529. BA.2  
322 includes all sublineages of BA.2 detected except BA.2.3.

323

324 **Figure 2.** The percent of sequences by week (estimated prevalence) belonging to BA.2.3  
325 colored by states including Alaska, California, New York, and Washington detected from 2021-  
326 11-28 to 2022-04-03.

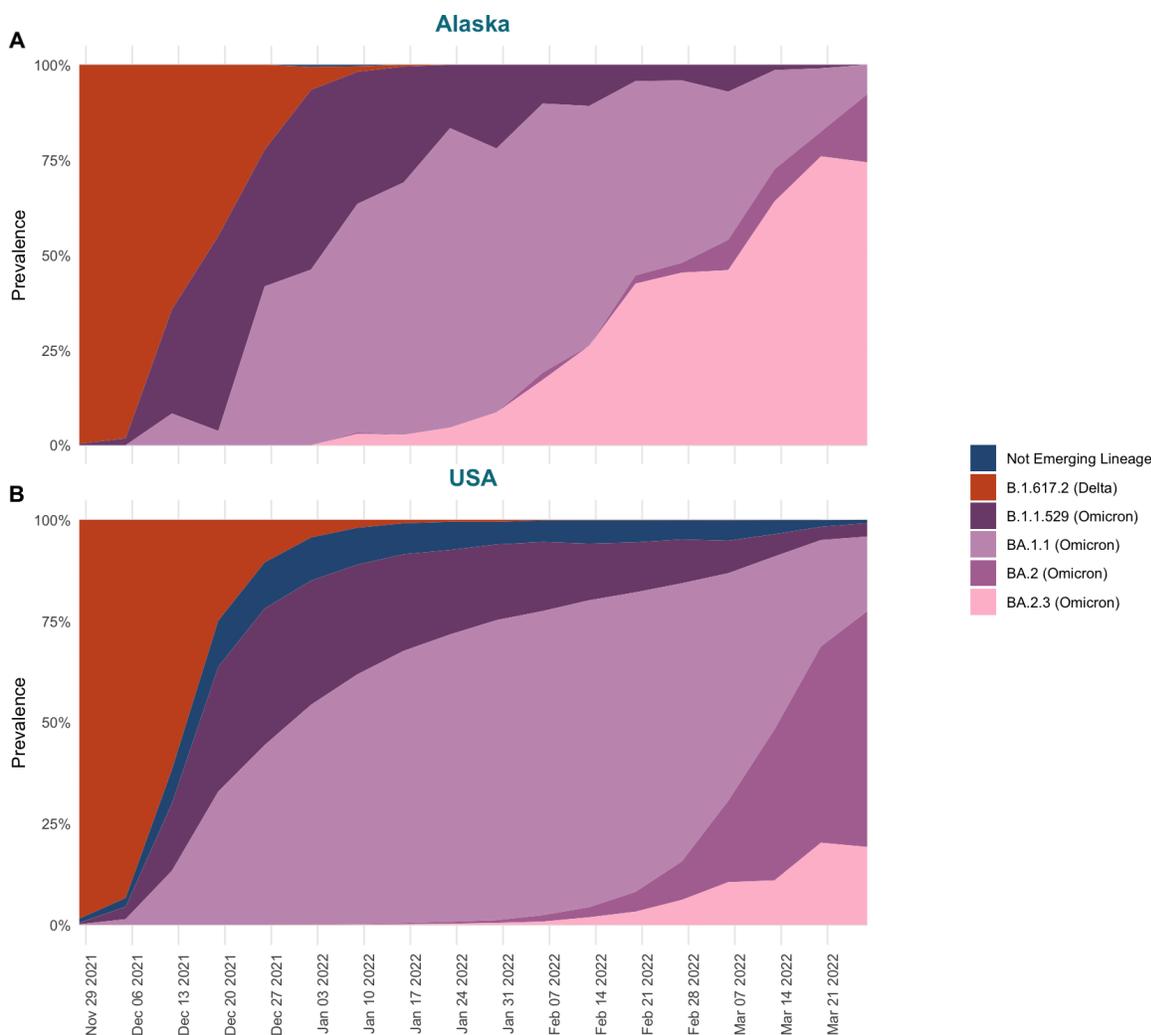
327 **Figure 3.** Logistic regression (line = regression; shaded region = standard error) estimating the  
328 prevalence of BA.2.3 over time in A) Alaska and the two economic regions of Alaska with a deep  
329 enough coverage of cases including (B) the Anchorage-Mat Su and (C) Gulf Coast. Points  
330 represent the daily percent of cases assigned to BA.2.3 used to calculate the regression. The red  
331 arrow highlights when the regression estimated BA.2.3 was at greater than 5% prevalence.

332

333 **Figure 4.** Phylogenetic tree of BA.2.3 cases with branch lengths represented by divergence of  
334 cases (number of mutations from Wuhan-Hu-1). Each point is a genome colored by country.  
335 Only countries that provide context for Alaska clades and clade origins are included in the  
336 visualization.

337 **Figure 5.** Phylogenetic tree of BA.2.3 cases with branch lengths represented by time. Tree  
338 includes all BA.2.3 cases from Alaska, all global BA.2.3 cases from December 2021, and all  
339 global cases after December 2021 through the first week of February 2022 downsampled for  
340 context. Only cases from Alaska are shown and are colored by economic region. Cases where  
341 the economic region is unknown are colored gray.

342

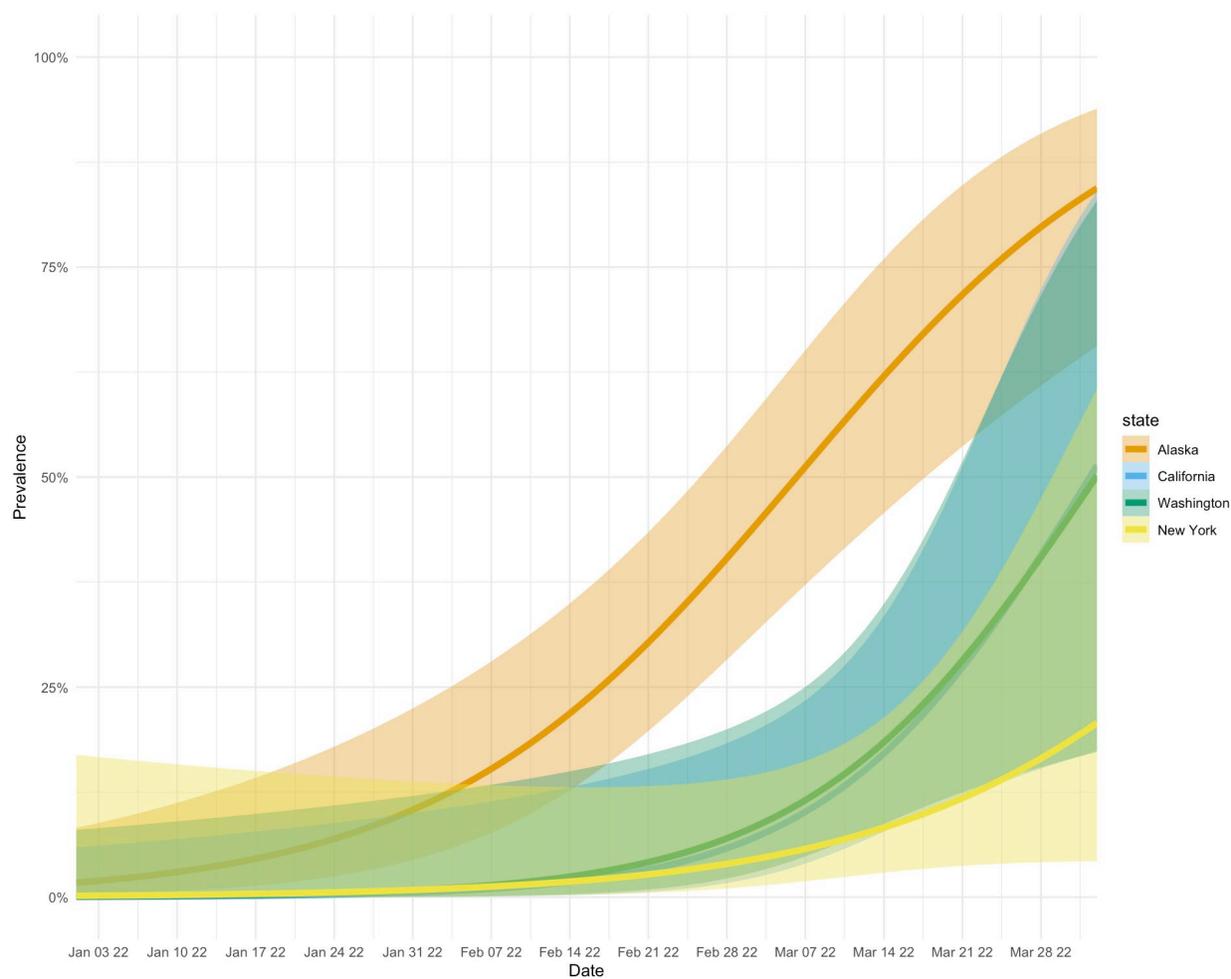


343

344 **Figure 1.** The percent of sequences by week (estimated prevalence) colored by SARS-CoV-2  
345 lineages detected from 2021-11-28 to 2022-04-02 in (A) Alaska and (B) the Lower 48. BA  
346 lineages of Omicron except BA.1.1, BA.2, and BA.2.3 are aggregated into B.1.1.529. BA.2  
347 includes all sublineages of BA.2 detected except BA.2.3.

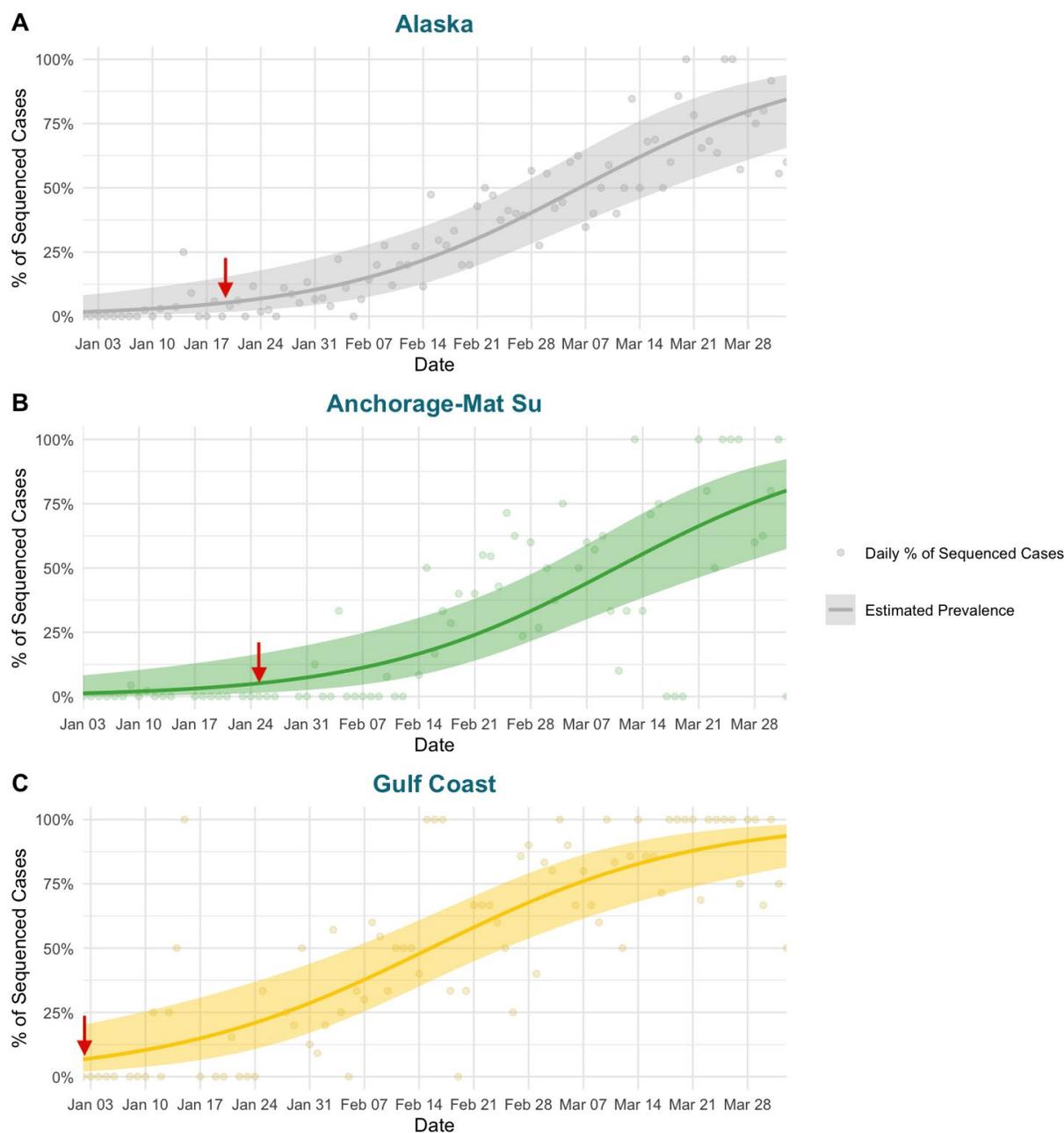
348

349

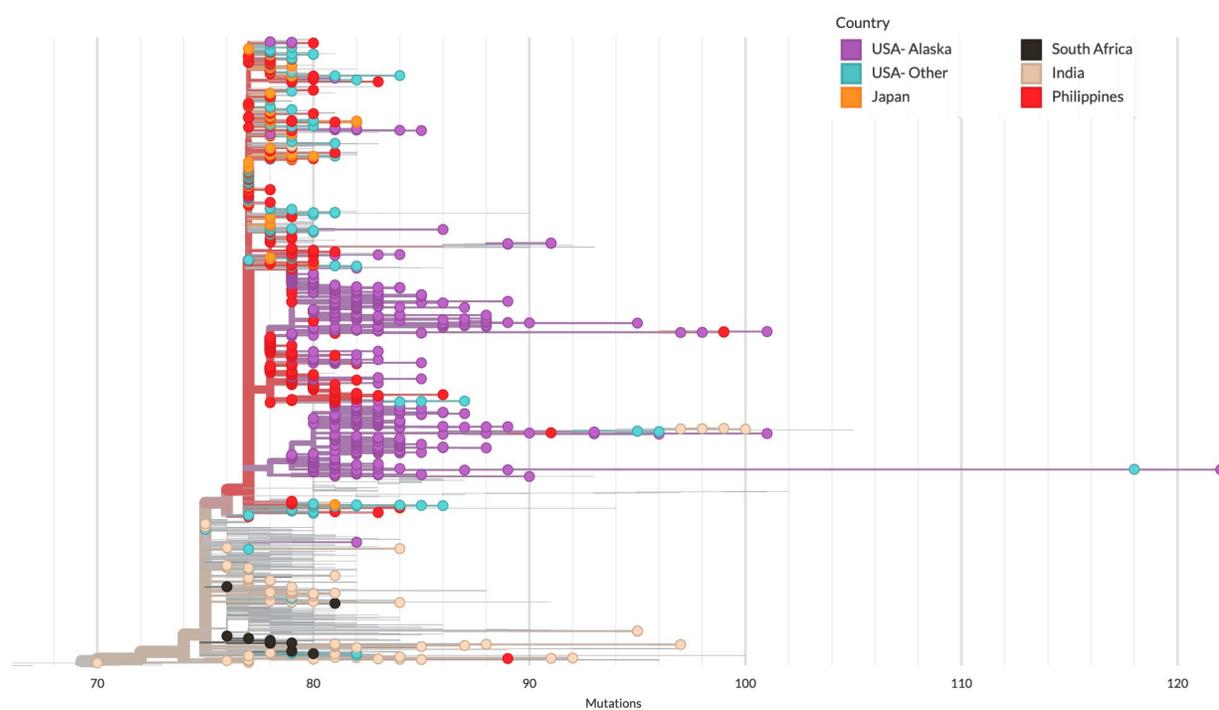


350

351 **Figure 2.** The percent of sequences by week (estimated prevalence) belonging to BA.2.3  
352 colored by states including Alaska, California, New York, and Washington detected from 2021-  
353 11-28 to 2022-04-03.



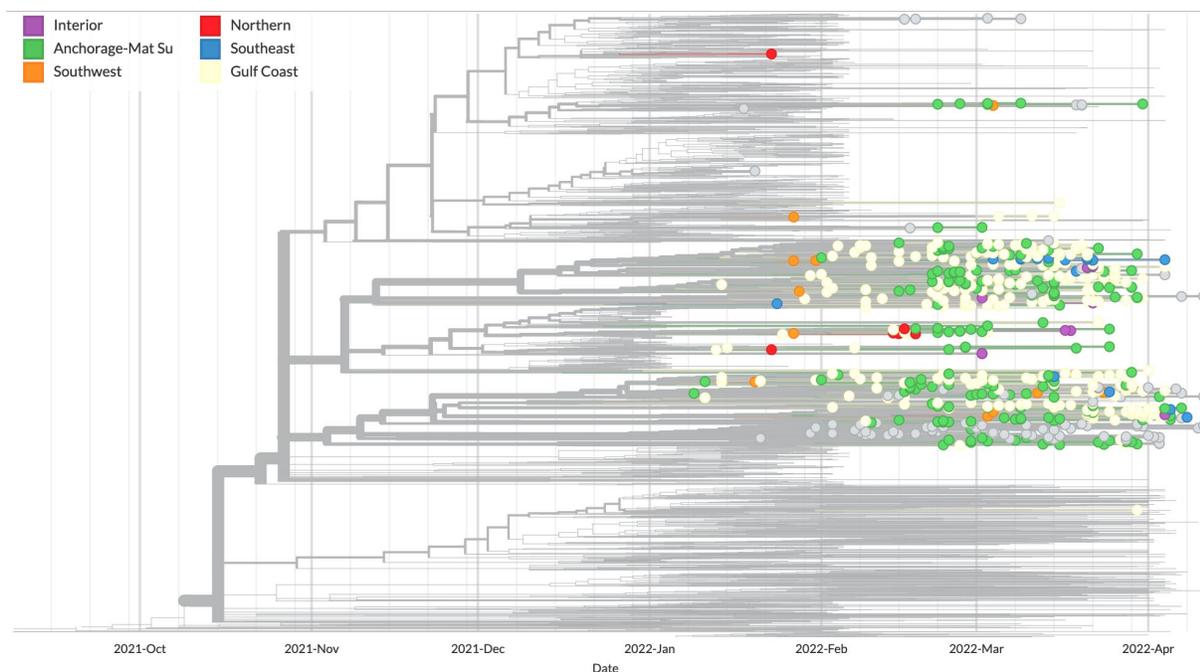
354  
355 **Figure 3.** Logistic regression (line = regression; shaded region = standard error) estimating the  
356 prevalence of BA.2.3 over time in A) Alaska and the two economic regions of Alaska with a  
357 deep enough coverage of cases including (B) the Anchorage-Mat Su and (C) Gulf Coast. Points  
358 represent the daily percent of cases assigned to BA.2.3 used to calculate the regression. The  
359 red arrow highlights when the regression estimated BA.2.3 was at greater than 5% prevalence.  
360



361

362 **Figure 4.** Phylogenetic tree of BA.2.3 cases with branch lengths represented by divergence of  
363 cases (number of mutations from Wuhan-Hu-1). Each point is a genome colored by country.  
364 Only countries that provide context for Alaska clades and clade origins are included in the  
365 visualization.

366



367

368 **Figure 5.** Phylogenetic tree of BA.2.3 cases with branch lengths represented by time. Tree  
369 includes all BA.2.3 cases from Alaska, all global BA.2.3 cases from December 2021, and  
370 downsampled global cases after December 2021 through the first week of February 2022 for  
371 context. Only cases from Alaska are shown and are colored by economic region. Cases where  
372 the economic region is unknown are colored gray.

373