



27 **ABSTRACT**

28 Emergence of SARS-CoV-2 with high transmission and immune evasion potential, the so-called  
29 Variants of Concern (VOC), is a major concern. We describe the early genomic epidemiology of  
30 SARS-CoV-2 recovered from vaccinated healthcare professionals (HCP). Our post-vaccination  
31 COVID-19 symptoms-based surveillance program among HCPs in a 17-hospital network,  
32 identified all vaccinated HCP who tested positive for COVID-19 after routine screening or after  
33 self-reporting. From 01/01/2021 to 04/30/2021, 23,687 HCP received either mRNA-1273 or  
34 BNT162b2 mRNA vaccine. All available post-vaccination SARS-CoV-2 samples and a random  
35 collection from non-vaccinated patients during the similar timeframe were subjected to VOC  
36 screening and whole genome sequencing (WGS). 62% (23,697/37,500) of HCPs received at  
37 least one vaccine dose, with 95% (22,458) fully vaccinated. We detected 138 (0.58%,  
38 138/23,697) COVID-19 cases, 105 among partially vaccinated and 33 (0.15%, 33/22,458)  
39 among fully vaccinated. Five partially vaccinated required hospitalization, four with supplemental  
40 oxygen. VOC screening from 16 fully vaccinated HCPs identified 6 (38%) harboring N501Y and  
41 1 (6%) with E484K polymorphisms; concurrent non-vaccinated samples was 37% (523/1404)  
42 and 20% (284/1394), respectively. There was an upward trend from January to April for  
43 E484K/Q (3% to 26%) and N501Y (1% to 49%). WGS analysis from vaccinated and non-  
44 vaccinated individuals indicated highly congruent phylogenies. We did not detect an increased  
45 frequency of any RBD/NTD polymorphism between groups ( $P>0.05$ ). Our results support robust  
46 protection by vaccination, particularly among recipients of both doses. Despite VOCs  
47 accounting for over 40% of SARS-CoV-2 from fully vaccinated individuals, the genomic diversity  
48 appears to proportionally represent those among non-vaccinated populations.

49

50

51

52 **IMPORTANCE**

53 A number of highly effective vaccines have been developed and deployed to combat the  
54 COVID-19 pandemic. The emergence and epidemiological dominance of SARS-CoV-2 mutants,  
55 with high transmission potential and immune evasion properties, the so-called Variants of  
56 Concern (VOC), continues to be a major concern. Whether these VOCs alter the efficacy of the  
57 administered vaccines is of great concern, and a critical question to study. We describe the  
58 initial genomic epidemiology of SARS-CoV-2 recovered from vaccinated healthcare  
59 professionals and probe specifically for VOC enrichment. Our findings support the high-level of  
60 protection provided by full vaccination despite a steep increase in the prevalence of  
61 polymorphisms associated with increased transmission potential (N501Y) and immune evasion  
62 (E484K) in the non-vaccinated population. Thus, we do not find evidence of VOC enrichment  
63 among vaccinated groups. Overall, the genomic diversity of SARS-CoV-2 recovered post-  
64 vaccination appears to proportionally represent the observed viral diversity within the  
65 community.

66

67 **Key words:** COVID-19, SARS-CoV-2, Variants of Concern, vaccine, breakthrough, spike  
68 protein

69

70

## 71 INTRODUCTION

72 Since the beginning of the pandemic in 2020, SARS-CoV-2 has infected and spread within an  
73 immunologically naïve global population (1, 2). With the application of convalescent plasma and  
74 therapeutic antibody treatment, along with highly effective vaccines, the virus is now starting to  
75 experience the immune pressure that will ultimately shape its evolutionary trajectory (3). Recent  
76 evidence indicates that SARS-CoV-2 can mutate and evade immunity and challenge the  
77 efficacy of emerging vaccines and intrinsic and extrinsic antibodies (i.e., natural by infection and  
78 antibody therapeutics) (3, 4).

79 The reports of mutants, known as Variants of Concerns (VOCs) that exhibit increased  
80 transmission or immune evasion, or both, have been reported in different global regions and  
81 predominant lineages have been genotyped (5). Of note, three VOCs, B.1.1.7 (recently named  
82 Alpha) (6), B.1.351 (Beta) (7), and P.1 (Gamma) (8) containing prominent mutations in the spike  
83 protein, emerged in the UK, South Africa and Brazil, respectively, and more recently, B.1.617.2  
84 (Delta) (9) emerged in India and has spread globally. The mapping of specific mutations in the  
85 spike protein has revealed strong evidence of convergent evolution, and particularly the E484K  
86 polymorphism, which is able to evade certain monoclonal therapy and it is less responsive to  
87 neutralizing antibodies from recovered patients. This mutant shows reduced response to  
88 convalescent plasma and reported reinfections and it has been identified in discrete lineages in  
89 different geographic locations and associated with variant clones with increased incidence (4).  
90 The impact of these antibody escape mutants harboring VOCs on driving community incidence  
91 and whether they alter the efficacy of the administered vaccines is of great concern.

92 In this report, we describe the initial genomic epidemiology of SARS-CoV-2 recovered from  
93 vaccinated healthcare professionals (HCPs) within a large healthcare network in New Jersey.  
94 We specifically probe whether specific SARS-CoV-2 variants skew viral diversity indicative of  
95 vaccine-induced selection; and compare against a random collection of SARS-CoV-2 recovered  
96 from non-vaccinated individuals during the same time period. Although we find steep increases

97 in variants harboring E484K and N501Y among community samples, our early genomic  
98 assessment does not indicate specific variant enrichment among post-vaccinated individuals.  
99 We find that viral genomes among vaccinated individuals largely reflect the viral diversity among  
100 non-vaccinated populations.

101

## 102 **RESULTS**

### 103 ***COVID-19 post-vaccination surveillance program: demographic and clinical*** 104 ***characteristics of COVID-19 cases.***

105 From December 2020 to April 2021, 23,697 of 37,500 HCPs (62%) received at least one dose  
106 of an mRNA vaccine, and 22,458 (59%, 22,458/37,500) received both vaccine doses. Among  
107 vaccinated employees, 12,878 (54%) received mRNA-1273 (Moderna) and 10,819 (46%)  
108 received BNT162b2 (Pfizer/BioNTech) mRNA COVID-19 vaccine. Of the 23,697 vaccinated  
109 (single and both doses) employees, our surveillance program detected 138 (0.58%, 138/23,697)  
110 COVID-19 cases, 105 of these were among employees who received only one dose and 33  
111 (0.15%, 33/22,458) with both doses.

112

113 Among the 138 post-vaccinated COVID-19 cases, 74 were vaccinated with the Pfizer/BioNTech  
114 vaccine and 64 were vaccinated with the Moderna preparation. Cases were reported from nine  
115 hospitals. Demographic information was not available for 17 employees. Of the remaining 121  
116 employees, most cases were white (68%) and female (74%) with a median age of 45 (ranging  
117 from 19-77). Twenty-one employees (17%) were documented to have asymptomatic infection.  
118 Of those that were symptomatic, five individuals requiring hospitalization received only one dose  
119 and had recorded BMI of >25 (4 individuals with BMI >30) and four of these patients required  
120 supplemental oxygen. The majority of the vaccine breakthroughs, 76% (105/138), occurred from  
121 1-113 days after the initial dose with a median of 10 days, while 23% (33/138) occurred 4-104  
122 days after the second dose with a median of 22 days.

123

124 ***Screening of post-vaccinated SARS-CoV-2 infections indicate the rise in E484K and***  
125 ***N501Y variants.***

126 To further characterize the SARS-CoV-2 genotypes recovered from the post-vaccinated  
127 individuals, we examined the spread of key mutations underlying VOCs (*i.e.*, B.1.1.7) in New  
128 Jersey using a high-throughput molecular beacon assay designed to screen for polymorphisms  
129 N501Y/T and E484K/Q in the RBD region (10). Among 138 HCPs, 83 swabs were available for  
130 rapid screening providing 76 genotypic results (60 from partial and 16 from fully vaccinated).  
131 There were 3 and 6 viral samples with E484K and N501Y polymorphisms respectively (Figure  
132 1). Among 16 fully vaccinated HCPs, we found 1 (6.3%) and 6 (38%) samples with E484K and  
133 N501Y mutations, respectively.

134 To contextualize the prevalence of E484K and N501Y mutations among the non-vaccinated  
135 population, we randomly sampled 1,404 SARS-CoV-2 positive swabs (from a total of 3,000  
136 positive COVID-19 swabs) from January 2021 to April 2021 within our hospital network. Overall,  
137 we detected 284 E484K/Q (284/1392, 20.4%, 12 samples failed in the E484 detection in  
138 comparison with N501Y) and 523 N501Y (523/1404, 37.3%) mutants (Figure 2). The prevalence  
139 of E484K/Q and N501Y from January to April, was 3.2% to 25.7% and 1.0% to 49.3%,  
140 respectively (Figure 2).

141

142 ***SARS-CoV-2 from vaccinated individuals is diverse and does not indicate predominant***  
143 ***genotypes or genetic characteristics.***

144 To understand the population structure of SARS-CoV-2 post-vaccination, we first analyzed  
145 WGS data for 68 available SARS-CoV-2 isolates from 83 symptomatic COVID-19 cases  
146 included in this study. WGS failed for 15 samples due to high Ct values, suggestive of low viral  
147 burden. We specifically probed whether specific variants or mutations were overrepresented  
148 among SARS-CoV-2 recovered from vaccinated individuals. The 68 genomes were divided into

149 21 different Pangolin lineages with B.1.2 the most common (24/68, 35%) (Figure 3). We noted  
150 distinct amino acid changes or deletions at 20 sites within the N-terminal domain (NTD) or  
151 receptor-binding domain (RBD) regions of the spike protein. Mutations associated with antibody  
152 evasion - L452R, T478K, E484K and S494P - were found in 3, 1, 3 and 2 genomes,  
153 respectively. We did not detect any genomes with more than two of the four polymorphisms.  
154 The N501Y mutation was found in 4% (3/68) of the genomes, and all belonged to the VOC  
155 B.1.1.7 lineage. We recorded 3 viral samples harboring the E484K mutation, two belonged to  
156 the B.1.526 lineage first identified in New York (11, 12) and one from the R.1 lineage.  
157 Importantly, among the 9 individuals fully vaccinated (*i.e.*,  $\geq 14$  days post-second dose) we note  
158 1 and 2 samples harboring E484K and N501Y polymorphisms, respectfully. There were no  
159 samples from fully vaccinated HCPs with both E484K and N501Y.

160 To compare viral diversity from vaccinated and non-vaccinated individuals, we randomly  
161 selected 281 SARS-CoV-2 samples recovered from non-vaccinated patients within the  
162 Hackensack Meridian Health (HMH) network during the same time period for genomic  
163 characterization by WGS (Figure 3). Among them, 38 different Pangolin lineages were  
164 identified, including 18 lineages found in the vaccinated population. B.1.526 (n=33) and its sub-  
165 lineages (B.1.526.1, n=26, and B.1.526.2, n=15) were the most common (74/281, 26.3%),  
166 followed by B.1.2 (35, 12.5%), B.1.1.7 (28, 10%) and B.1.575 (26, 9.3%). In addition, 5 P.1  
167 (20J/501Y.V3) and 1 B.1.617.1 (Kappa) were found among the 281 samples. Mutations or  
168 deletions were detected in 54 sites within the S protein RBD or NTD region, including 18 sites  
169 found in the vaccinated samples. L452R, T478K/I, E484K/Q and S494P mutations were  
170 identified in 32, 8, 40 and 29 genomes, respectively. Compared with the non-vaccinated  
171 samples, we did not detect a significant increase in the frequency of any RBD/NTD mutations or  
172 deletions among the vaccinated group ( $P > 0.05$ ).

173

174 **DISCUSSION**

175 Rapid expansion and contraction of SARS-CoV-2 populations coupled with increasing anti-CoV-  
176 2 immunity dynamics have contributed to pandemic phases accompanied by emergent genomic  
177 signatures (1, 13). Extensive polymorphisms in the spike protein have been documented,  
178 including N501Y and E484K mutations that are particularly concerning (14). In this report, we  
179 present early genomic analysis of SARS-CoV-2 from an ongoing post-vaccination COVID-19  
180 surveillance program among healthcare workers in New Jersey.

181 We detected SARS-CoV-2 in 33 (0.15%) HCPs who were fully vaccinated, none of whom  
182 required hospitalization. All symptomatic cases were noted among partially vaccinated HCPs  
183 and were self-limiting with only 5 requiring hospitalization. Although our passive surveillance  
184 system was intended to capture the HCP vaccinated population who self-reported symptoms or  
185 who were exposed to a COVID-19 positive case, and not able to report true prevalence of  
186 infections (i.e., mild or asymptomatic infections may be missed), our findings are consistent with  
187 emerging reports of incident post-vaccination SARS-CoV-2 infections (15-18). A recent  
188 multicenter prospective cohort study in England found vaccine effectiveness upwards of 85%  
189 (19). Our results support high levels of protection afforded by vaccination, particularly among  
190 recipients of both doses (9, 15, 20-24). The majority of cases in our study were among partially  
191 vaccinated individuals, emphasizing the importance of personal protective measures, and  
192 susceptibility to infection particularly until fully vaccinated (19, 24-27).

193 Steep transmission gradients reported locally and regionally have been partly attributed to  
194 VOCs, including B.1.1.7 in Europe and U.S., P.1 in Brazil, B.1.351 in South Africa, B.1.617.2 in  
195 India and B.1.526 in New York State (6-8, 11, 28). Although the pathogen characteristics  
196 attributable to the spread of these variants at the population-level are difficult to assess,  
197 mathematical modeling, genomic and experimental studies have demonstrated increased  
198 transmissibility and immune evasion properties, including reduced vaccine-induced antibody



199 neutralization *in vitro* (2, 3, 6, 14). Whether the rollout of vaccination would offer selective  
200 advantage to VOCs is a critical question to determine.

201 A recent study from Washington State found that all SARS-CoV-2 vaccine breakthroughs  
202 were VOCs. Of significance, those VOCs harboring immune evasion properties (B.1.351,  
203 B.1.427, B.1.429, and P1) were enriched compared to B.1.1.7 lineages among vaccine  
204 breakthrough cases when compared to cases in the general population (29). Other outbreaks  
205 have been reported among vaccinated individuals implicating SARS-CoV-2 strains harboring  
206 E484K, a noted immune evasion conferring polymorphism (20). A recent study where two VOCs  
207 (B.1.351 and B.1.1.7) dominate the viral population, indicate reduced vaccine effectiveness  
208 against both variants at specific time windows. Reduced effectiveness against B.1.351 at least 7  
209 days after the second BNT162b2 dose, and against B.1.1.7 between 2 weeks after the first dose  
210 and 6 days after the second dose were reported (24). However, among fully vaccinated HCPs  
211 the VOC proportions between cases and controls were comparable, consistent with our  
212 findings. These studies reinforce the need to fully vaccinate to achieve high levels of protection  
213 afforded by vaccination (9).

214 In our study, despite an increasing trend in the proportion of N501Y and E484K variants in  
215 the overall population (Figure 2), we do not find early evidence of genotypic enrichment of  
216 polymorphisms within the NTD or RBD region of the spike protein gene among our fully  
217 vaccinated HCPs. Although 44% of the strains recovered from fully vaccinated HCPs harbored  
218 mutations of concern (6 N501Y and 1 E484K), they were similar in proportion to strains  
219 circulating in the community during the same time period. Our WGS analysis suggests N501Y  
220 harboring strains belong to the B.1.1.7 lineage and the E484K mutants to the B.1.526 lineage,  
221 both highly prevalent lineages in Northeastern US (12, 30). All N501Y mutants were detected  
222 among fully vaccinated HCPs, half of which failed to generate WGS data due to higher Ct  
223 values, a likely consequence of vaccination (31).

224 A recent report showed that B.1.1.7 upregulates key innate immune antagonists likely  
225 increasing the transmission potential of this VOC (32). Interestingly, mRNA vaccines have  
226 shown high levels of efficacy against B.1.1.7 infections (22). Overall, the genomic diversity of  
227 SARS-CoV-2 recovered post-vaccination appears to proportionally represent the observed viral  
228 diversity within the community. We acknowledge that our estimates may be skewed by under-  
229 sampling of asymptomatic infections among fully vaccinated individuals (33), and likely affected  
230 by test seeking behavioral bias (24). However, our results are consistent with those from a large  
231 nationwide surveillance program of COVID-19 vaccine breakthrough infections that found the  
232 proportion of reported vaccine breakthrough infections attributed to VOCs to be similar to the  
233 proportion of these variants circulating throughout the United States (27).

234 There is mounting evidence that currently available vaccines, including those used in our  
235 population, are highly effective against VOCs. The reports of VOC enrichment among post-  
236 vaccinated individuals, despite small sample sizes, warrant surveillance. Our study findings are  
237 preliminary yet highlights the importance of full vaccination against circulating variants, including  
238 VOCs, and the need to continue SARS-CoV-2 genomic surveillance as vaccination coverage  
239 expands, pathogen evolves and immunity wanes.

240

## 241 **METHODS**

### 242 ***Post-vaccination COVID-19 surveillance program.***

243 Since the first mRNA vaccines were administered in December 2020 among the 37,500 HMH  
244 employees, we initiated a symptoms-based surveillance program to identify and report the  
245 status of each vaccinated person in the Electronic Medical Record (EMR). This regularly  
246 updated list was cross-referenced daily against the results of all SARS-Cov-2 nucleic acid  
247 amplification tests (NAAT) in the HMH network. As a result, we were able to identify all  
248 vaccinated employees who tested positive for COVID-19 by NAAT after routine screening or

249 after reporting out sick. Non-team members who were vaccinated at our mass-vaccination site  
250 or by physician report (and were thus in the HMH EMR) and subsequently tested positive for  
251 COVID-19 by NAAT were also similarly identified. The study was approved by Hackensack  
252 Meridian Health Institutional Review Board (IRB).

253

#### 254 ***Whole genome sequencing and phylogenetic analysis***

255 SARS-CoV-2 targeted assay libraries were prepared using the Molecular Loop Viral RNA Target  
256 Capture Kit (Molecular Loop) or QIAseq FX DNA Library UDI Kit (QIAGEN), in accordance with  
257 manufacturer's recommendations. Final libraries were quantified using fluorescent-based  
258 assays including PicoGreen (Life Technologies), Qubit Fluorometer (Invitrogen), and Fragment  
259 Analyzer (Advanced Analytics). Final libraries were sequenced on a NovaSeq 6000 sequencer  
260 (v1 chemistry) with 2x150bp.

261

262 For QIAseq libraries, read pairs that did not contain a single 19bp seed k-mer in common with  
263 the SARS-CoV-2 genome reference (NC\_045512.2) were discarded. Adapter sequences and  
264 low-quality bases ( $Q < 20$ ) were trimmed from the 3' end of the remaining reads, using Cutadapt  
265 v2.10 (34). Sequences corresponding to the amplicon primers were also clipped from the 5' end  
266 of reads. Processed read pairs were then aligned to the SARS-CoV-2 reference genome, using  
267 BWA-MEM v0.7.17 (35) and only read pairs with at least one alignment spanning a minimum of  
268 50 bp in the reference were kept.

269

270 For Molecular Loop libraries, the two 5bp UMIs located at the 5' end of each mate were first  
271 clipped and combined into a single UMI tag. Additional 25 bp, corresponding to the molecular  
272 inversion probes, were also clipped from the 5' end of each read. Next, read pairs that did not  
273 contain a single 19bp seed k-mer in common with the SARS-CoV-2 genome reference  
274 (NC\_045512.2) were discarded and adapter sequences and low-quality bases ( $Q < 20$ ) were

275 trimmed from the 3' end of the remaining reads, using Cutadapt v2.10 (34). Processed reads  
276 pairs were then merged using NGmerge v0.2 (36), allowing for dovetailed alignments. The  
277 resulting single end reads were mapped against the SARS-CoV-2 genome reference using  
278 BWA-MEM v0.7.17 (35) and the resulting alignments filtered using the following criteria: 1)  
279 reference span  $\geq$  50bp, 2) quality  $\geq$  60, and 3) maximum soft-clip length on either end  $\leq$  30bp.  
280 Next, reads representing the same original molecule were identified based on their shared UMI  
281 and alignment position and used to draw the molecule consensus sequence, taking into account  
282 base quality scores. Molecule sequences were then realigned to the SARS-CoV-2 genome  
283 reference using BWA-MEM v0.7.17 (35). Finally, genome sequences were determined either  
284 by read (QIAseq) or molecule (Molecular Loop) alignment pileup consensus calling with a  
285 minimum support of 5 reads/molecules.

286  
287 The resulting SARS-CoV-2 viral genome sequences were characterized by Nextclade CLI (v1)  
288 (<https://clades.nextstrain.org/>) to assign Nextstrain clades (37). SARS-CoV-2 lineages were  
289 determined using Pangolin v3.1.3 (<https://github.com/cov-lineages/pangolin>) and GISAID clades  
290 were determined based upon the clade-specific marker variants from <https://www.gisaid.org>  
291 (38). The genomes were aligned using nextalign v1.0.0 (<https://github.com/neherlab/nextalign>),  
292 using default setting. A maximum likelihood phylogenetic tree was constructed using IQ-TREE  
293 v2.1.2 (39) with automatic model selection and 1000-bootstrap replicates. The resulting tree  
294 was annotated using ITOL v6 (40).

295

### 296 ***Population sampling and screening of SARS-CoV-2***

297 SARS-CoV-2 positive swabs collected from seven HMH network hospitals were shipped to CDI  
298 on a weekly basis and stored at  $-80^{\circ}\text{C}$ . 1,404 swabs collected during the same time frame as  
299 the vaccinated samples were randomly selected and subjected to a molecular beacon-based  
300 real-time asymmetric PCR and melting curve analysis to identify the SARS-CoV-2 E484K/Q and

301 N501Y mutations, as previously described (10). In brief, 50  $\mu$ l aliquots of SARS-CoV-2 swab  
302 specimens were treated by proteinase K and heated at 95°C for 5 min. 5ul of heat inactivated  
303 sample was used as template for the asymmetric RT-PCR testing on a Mic Real Time PCR  
304 Cyclor (Bio Molecular Systems, software micPCRv2.8.13). Based on the different melting  
305 profiles ( $T_m$ ) of E484 and N501 molecular beacons in each sample, the E484K and N501Y  
306 mutations were determined.

307

### 308 **Statistical analysis**

309 Fisher's exact or chi-square tests, as appropriate, were used to examine whether the spike  
310 protein NTD/RBD mutations were enriched in the vaccinated HCPs (n=68) in comparison with  
311 the non-vaccinated patients (n=281). SPSS, version 17.0 (IBM Corp., Armonk, NY) was used  
312 for statistical analyses and two-tailed P values  $\leq 0.05$  were considered statistically significant.

313

### 314 **Data availability**

315 The SARS-CoV-2 genomes sequenced in this study were deposited in GISAID  
316 ([www.gisaid.org](http://www.gisaid.org)). Sequences can be accessed by searching records from both the originating  
317 lab at Hackensack Medical Center and the submitting lab at the New York Genome Center.

318

319 **FIGURE 1.** Clinical characteristics of 83 COVID-19 vaccinated HCPs and their SARS-CoV-2 S  
320 protein E484 and N501 genotypes. The samples were ordered by the time of diagnosis, with the  
321 the first case detection week denoted as “0”.

322

323

324

325

326

327 **FIGURE 2.** Prevalence of E484K/Q and N501Y mutants among SARS-Cov-2 samples from  
328 non-vaccinated individuals from January to April 2021.

329

330

331

332

333 **FIGURE 3.** Maximum Likelihood Phylogenetic tree of 349 SARS-CoV-2 genomes isolated from  
334 vaccinated and non-vaccinated populations. The tree is rooted to the SARS-CoV-2 Wuhan-Hu-1  
335 reference genome (NC\_045512.2). The scale represents 0.0001 nucleotide substitutions per  
336 site. The vaccination conditions are color coded at the tree tips: purple, black and white tips  
337 denote single-dose, double-dose vaccine and non-vaccination, respectively. The SARS-CoV-2  
338 Pangolin lineage is illustrated as a color bar and the S protein NTD/RBD mutations are shown  
339 as a heatmap on the right panel.

340

341 **References:**

342

- 343 1. Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M. 2020. Projecting the  
344 transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science*  
345 368:860-868.
- 346 2. Hodcroft EB, Zuber M, Nadeau S, Vaughan TG, Crawford KHD, Althaus CL, Reichmuth  
347 ML, Bowen JE, Walls AC, Corti D, Bloom JD, Veessler D, Mateo D, Hernando A, Comas  
348 I, Candelas FG, Seq C-Sc, Stadler T, Neher RA. 2021. Spread of a SARS-CoV-2 variant  
349 through Europe in the summer of 2020. *Nature* doi:10.1038/s41586-021-03677-y.
- 350 3. Starr TN, Greaney AJ, Addetia A, Hannon WW, Choudhary MC, Dingens AS, Li JZ,  
351 Bloom JD. 2021. Prospective mapping of viral mutations that escape antibodies used to  
352 treat COVID-19. *Science* 371:850-854.
- 353 4. Greaney AJ, Loes AN, Crawford KHD, Starr TN, Malone KD, Chu HY, Bloom JD.  
354 2021. Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding  
355 domain that affect recognition by polyclonal human plasma antibodies. *Cell Host*  
356 *Microbe* 29:463-476 e6.
- 357 5. Boehm E, Kronig I, Neher RA, Eckerle I, Vetter P, Kaiser L, Geneva Centre for  
358 Emerging Viral D. 2021. Novel SARS-CoV-2 variants: the pandemics within the  
359 pandemic. *Clin Microbiol Infect* doi:10.1016/j.cmi.2021.05.022.
- 360 6. Volz E, Mishra S, Chand M, Barrett JC, Johnson R, Geidelberg L, Hinsley WR, Laydon  
361 DJ, Dabrera G, O'Toole A, Amato R, Ragonnet-Cronin M, Harrison I, Jackson B, Ariani  
362 CV, Boyd O, Loman NJ, McCrone JT, Goncalves S, Jorgensen D, Myers R, Hill V,  
363 Jackson DK, Gaythorpe K, Groves N, Sillitoe J, Kwiatkowski DP, consortium C-GU,

- 364 Flaxman S, Ratmann O, Bhatt S, Hopkins S, Gandy A, Rambaut A, Ferguson NM. 2021.  
365 Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature* 593:266-  
366 269.
- 367 7. Tegally H, Wilkinson E, Giovanetti M, Iranzadeh A, Fonseca V, Giandhari J, Doolabh D,  
368 Pillay S, San EJ, Msomi N, Mlisana K, von Gottberg A, Walaza S, Allam M, Ismail A,  
369 Mohale T, Glass AJ, Engelbrecht S, Van Zyl G, Preiser W, Petruccione F, Sigal A,  
370 Hardie D, Marais G, Hsiao NY, Korsman S, Davies MA, Tyers L, Mudau I, York D,  
371 Maslo C, Goedhals D, Abrahams S, Laguda-Akingba O, Alisoltani-Dehkordi A, Godzik  
372 A, Wibmer CK, Sewell BT, Lourenco J, Alcantara LCJ, Kosakovsky Pond SL, Weaver S,  
373 Martin D, Lessells RJ, Bhiman JN, Williamson C, de Oliveira T. 2021. Detection of a  
374 SARS-CoV-2 variant of concern in South Africa. *Nature* 592:438-443.
- 375 8. Faria NR, Mellan TA, Whittaker C, Claro IM, Candido DDS, Mishra S, Crispim MAE,  
376 Sales FCS, Hawryluk I, McCrone JT, Hulswit RJG, Franco LAM, Ramundo MS, de  
377 Jesus JG, Andrade PS, Coletti TM, Ferreira GM, Silva CAM, Manuli ER, Pereira RHM,  
378 Peixoto PS, Kraemer MUG, Gaburo N, Jr., Camilo CDC, Hoeltgebaum H, Souza WM,  
379 Rocha EC, de Souza LM, de Pinho MC, Araujo LJT, Malta FSV, de Lima AB, Silva  
380 JDP, Zauli DAG, Ferreira ACS, Schnekenberg RP, Laydon DJ, Walker PGT, Schluter  
381 HM, Dos Santos ALP, Vidal MS, Del Caro VS, Filho RMF, Dos Santos HM, Aguiar RS,  
382 Proenca-Modena JL, Nelson B, Hay JA, Monod M, Miscouridou X, et al. 2021.  
383 Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science*  
384 372:815-821.
- 385 9. Liu J, Liu Y, Xia H, Zou J, Weaver SC, Swanson KA, Cai H, Cutler M, Cooper D, Muik  
386 A, Jansen KU, Sahin U, Xie X, Dormitzer PR, Shi PY. 2021. BNT162b2-elicited



- 387 neutralization of B.1.617 and other SARS-CoV-2 variants. *Nature* doi:10.1038/s41586-  
388 021-03693-y.
- 389 10. Zhao Y, Lee A, Composto K, Cunningham MH, Mediavilla JR, Fennessey S, Corvelo A,  
390 Chow KF, Zody M, Chen L, Kreiswirth BN, Perlin DS. 2021. A novel diagnostic test to  
391 screen SARS-CoV-2 variants containing E484K and N501Y mutations. *Emerg Microbes*  
392 *Infect* 10:994-997.
- 393 11. Annavajhala MK, Mohri H, Zucker JE, Sheng Z, Wang P, Gomez-Simmonds A, Ho DD,  
394 Uhlemann AC. 2021. A Novel SARS-CoV-2 Variant of Concern, B.1.526, Identified in  
395 New York. medRxiv doi:10.1101/2021.02.23.21252259.
- 396 12. Thompson CN, Hughes S, Ngai S, Baumgartner J, Wang JC, McGibbon E, Devinney K,  
397 Luoma E, Bertolino D, Hwang C, Kepler K, Del Castillo C, Hopkins M, Lee H, DeVito  
398 AK, Rakeman JL, PhD, Fine AD. 2021. Rapid Emergence and Epidemiologic  
399 Characteristics of the SARS-CoV-2 B.1.526 Variant - New York City, New York,  
400 January 1-April 5, 2021. *MMWR Morb Mortal Wkly Rep* 70:712-716.
- 401 13. Cobey S, Larremore DB, Grad YH, Lipsitch M. 2021. Concerns about SARS-CoV-2  
402 evolution should not hold back efforts to expand vaccination. *Nat Rev Immunol* 21:330-  
403 335.
- 404 14. Wang P, Nair MS, Liu L, Iketani S, Luo Y, Guo Y, Wang M, Yu J, Zhang B, Kwong PD,  
405 Graham BS, Mascola JR, Chang JY, Yin MT, Sobieszczyk M, Kyratsous CA, Shapiro L,  
406 Sheng Z, Huang Y, Ho DD. 2021. Antibody resistance of SARS-CoV-2 variants B.1.351  
407 and B.1.1.7. *Nature* 593:130-135.
- 408 15. Bouton TC, Lodi S, Turcinovic J, Weber SE, Quinn E, Korn C, Steiner J, Schechter-  
409 Perkins EM, Duffy E, Ragan EJ, Taylor BP, Schaeffer B, Miller N, Davidoff R, Hanage

- 410 WP, Connor J, Pierre C, Jacobson KR. 2021. COVID-19 vaccine impact on rates of  
411 SARS-CoV-2 cases and post vaccination strain sequences among healthcare workers at  
412 an urban academic medical center: a prospective cohort study. medRxiv  
413 doi:10.1101/2021.03.30.21254655.
- 414 16. Daniel W, Nivet M, Warner J, Podolsky DK. 2021. Early Evidence of the Effect of  
415 SARS-CoV-2 Vaccine at One Medical Center. *N Engl J Med* 384:1962-1963.
- 416 17. Keehner J, Horton LE, Pfeffer MA, Longhurst CA, Schooley RT, Currier JS, Abeles SR,  
417 Torriani FJ. 2021. SARS-CoV-2 Infection after Vaccination in Health Care Workers in  
418 California. *N Engl J Med* 384:1774-1775.
- 419 18. Teran RA, Walblay KA, Shane EL, Xydis S, Gretsch S, Gagner A, Samala U, Choi H,  
420 Zelinski C, Black SR. 2021. Postvaccination SARS-CoV-2 Infections Among Skilled  
421 Nursing Facility Residents and Staff Members - Chicago, Illinois, December 2020-March  
422 2021. *MMWR Morb Mortal Wkly Rep* 70:632-638.
- 423 19. Hall VJ, Foulkes S, Saei A, Andrews N, Oguti B, Charlett A, Wellington E, Stowe J,  
424 Gillson N, Atti A, Islam J, Karagiannis I, Munro K, Khawam J, Chand MA, Brown CS,  
425 Ramsay M, Lopez-Bernal J, Hopkins S, Group SS. 2021. COVID-19 vaccine coverage in  
426 health-care workers in England and effectiveness of BNT162b2 mRNA vaccine against  
427 infection (SIREN): a prospective, multicentre, cohort study. *Lancet* 397:1725-1735.
- 428 20. Cavanaugh AM, Fortier S, Lewis P, Arora V, Johnson M, George K, Tobias J, Lunn S,  
429 Miller T, Thoroughman D, Spicer KB. 2021. COVID-19 Outbreak Associated with a  
430 SARS-CoV-2 R.1 Lineage Variant in a Skilled Nursing Facility After Vaccination  
431 Program - Kentucky, March 2021. *MMWR Morb Mortal Wkly Rep* 70:639-643.

- 432 21. Dagan N, Barda N, Kepten E, Miron O, Perchik S, Katz MA, Hernan MA, Lipsitch M,  
433 Reis B, Balicer RD. 2021. BNT162b2 mRNA Covid-19 Vaccine in a Nationwide Mass  
434 Vaccination Setting. *N Engl J Med* 384:1412-1423.
- 435 22. Haas EJ, Angulo FJ, McLaughlin JM, Anis E, Singer SR, Khan F, Brooks N, Smaja M,  
436 Mircus G, Pan K, Southern J, Swerdlow DL, Jodar L, Levy Y, Alroy-Preis S. 2021.  
437 Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections  
438 and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination  
439 campaign in Israel: an observational study using national surveillance data. *Lancet*  
440 397:1819-1829.
- 441 23. Angel Y, Spitzer A, Henig O, Saiag E, Sprecher E, Padova H, Ben-Ami R. 2021.  
442 Association Between Vaccination With BNT162b2 and Incidence of Symptomatic and  
443 Asymptomatic SARS-CoV-2 Infections Among Health Care Workers. *JAMA*  
444 doi:10.1001/jama.2021.7152.
- 445 24. Kustin T, Harel N, Finkel U, Perchik S, Harari S, Tahor M, Caspi I, Levy R, Leshchinsky  
446 M, Ken Dror S, Bergerzon G, Gadban H, Gadban F, Eliassian E, Shimron O, Saleh L,  
447 Ben-Zvi H, Keren Taraday E, Amichay D, Ben-Dor A, Sagas D, Strauss M, Shemer Avni  
448 Y, Huppert A, Kepten E, Balicer RD, Netzer D, Ben-Shachar S, Stern A. 2021. Evidence  
449 for increased breakthrough rates of SARS-CoV-2 variants of concern in BNT162b2-  
450 mRNA-vaccinated individuals. *Nat Med* doi:10.1038/s41591-021-01413-7.
- 451 25. Abu-Raddad LJ, Chemaitelly H, Butt AA, National Study Group for C-V. 2021.  
452 Effectiveness of the BNT162b2 Covid-19 Vaccine against the B.1.1.7 and B.1.351  
453 Variants. *N Engl J Med* doi:10.1056/NEJMc2104974.

- 454 26. Chodick G, Tene L, Rotem RS, Patalon T, Gazit S, Ben-Tov A, Weil C, Goldshtein I,  
455 Twig G, Cohen D, Muhsen K. 2021. The effectiveness of the TWO-DOSE BNT162b2  
456 vaccine: analysis of real-world data. *Clin Infect Dis* doi:10.1093/cid/ciab438.
- 457 27. Team CC-VBCI. 2021. COVID-19 Vaccine Breakthrough Infections Reported to CDC -  
458 United States, January 1-April 30, 2021. *MMWR Morb Mortal Wkly Rep* 70:792-793.
- 459 28. Candido DS, Claro IM, de Jesus JG, Souza WM, Moreira FRR, Dellicour S, Mellan TA,  
460 du Plessis L, Pereira RHM, Sales FCS, Manuli ER, Theze J, Almeida L, Menezes MT,  
461 Voloch CM, Fumagalli MJ, Coletti TM, da Silva CAM, Ramundo MS, Amorim MR,  
462 Hoeltgebaum HH, Mishra S, Gill MS, Carvalho LM, Buss LF, Prete CA, Jr., Ashworth J,  
463 Nakaya HI, Peixoto PS, Brady OJ, Nicholls SM, Tanuri A, Rossi AD, Braga CKV,  
464 Gerber AL, de CGAP, Gaburo N, Jr., Alencar CS, Ferreira ACS, Lima CX, Levi JE,  
465 Granato C, Ferreira GM, Francisco RS, Jr., Granja F, Garcia MT, Moretti ML, Perroud  
466 MW, Jr., Castineiras T, Lazari CS, et al. 2020. Evolution and epidemic spread of SARS-  
467 CoV-2 in Brazil. *Science* doi:10.1126/science.abd2161.
- 468 29. McEwen AE, Cohen S, Bryson-Cahn C, Liu C, Pergam SA, Lynch J, Schippers A, Strand  
469 K, Whimbey E, Mani NS, Zelikoff AJ, Makarewicz VA, Brown ER, Mohamed Bakhsh  
470 SA, Baker NR, Castor J, Livingston RJ, Huang M-L, Jerome KR, Greninger AL,  
471 Roychoudhury P. 2021. Variants of concern are overrepresented among post-vaccination  
472 breakthrough infections of SARS-CoV-2 in Washington State. *medRxiv*  
473 doi:10.1101/2021.05.23.21257679:2021.05.23.21257679.
- 474 30. West AP, Wertheim JO, Wang JC, Vasylyeva TI, Havens JL, Chowdhury MA, Gonzalez  
475 E, Fang CE, Di Lonardo SS, Hughes S, Rakeman JL, Lee HH, Barnes CO,  
476 Gnanapragasam PNP, Yang Z, Gaebler C, Caskey M, Nussenzweig MC, Keeffe JR,

- 477 Bjorkman PJ. 2021. Detection and characterization of the SARS-CoV-2 lineage B.1.526  
478 in New York. bioRxiv doi:10.1101/2021.02.14.431043.
- 479 31. Levine-Tiefenbrun M, Yelin I, Katz R, Herzel E, Golan Z, Schreiber L, Wolf T, Nadler  
480 V, Ben-Tov A, Kuint J, Gazit S, Patalon T, Chodick G, Kishony R. 2021. Initial report of  
481 decreased SARS-CoV-2 viral load after inoculation with the BNT162b2 vaccine. *Nat*  
482 *Med* 27:790-792.
- 483 32. Thorne LG, Bouhaddou M, Reuschl AK, Zuliani-Alvarez L, Polacco B, Pelin A, Batra J,  
484 Whelan MVX, Ummadi M, Rojc A, Turner J, Obernier K, Braberg H, Soucheray M,  
485 Richards A, Chen KH, Harjai B, Memon D, Hosmillo M, Hiatt J, Jahun A, Goodfellow  
486 IG, Fabius JM, Shokat K, Jura N, Verba K, Noursadeghi M, Beltrao P, Swaney DL,  
487 Garcia-Sastre A, Jolly C, Towers GJ, Krogan NJ. 2021. Evolution of enhanced innate  
488 immune evasion by the SARS-CoV-2 B.1.1.7 UK variant. bioRxiv  
489 doi:10.1101/2021.06.06.446826.
- 490 33. Hay JA, Kennedy-Shaffer L, Kanjilal S, Lennon NJ, Gabriel SB, Lipsitch M, Mina MJ.  
491 2021. Estimating epidemiologic dynamics from cross-sectional viral load distributions.  
492 *Science* doi:10.1126/science.abh0635.
- 493 34. Kechin A, Boyarskikh U, Kel A, Filipenko M. 2017. cutPrimers: A New Tool for  
494 Accurate Cutting of Primers from Reads of Targeted Next Generation Sequencing. *J*  
495 *Comput Biol* 24:1138-1143.
- 496 35. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler  
497 transform. *Bioinformatics* 25:1754-60.
- 498 36. Gaspar JM. 2018. NGmerge: merging paired-end reads via novel empirically-derived  
499 models of sequencing errors. *BMC Bioinformatics* 19:536.

- 500 37. Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, Sagulenko P,  
501 Bedford T, Neher RA. 2018. Nextstrain: real-time tracking of pathogen evolution.  
502 *Bioinformatics* 34:4121-4123.
- 503 38. Elbe S, Buckland-Merrett G. 2017. Data, disease and diplomacy: GISAID's innovative  
504 contribution to global health. *Glob Chall* 1:33-46.
- 505 39. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and  
506 effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol*  
507 *Evol* 32:268-74.
- 508 40. Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new  
509 developments. *Nucleic Acids Res* 47:W256-W259.
- 510

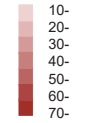
### Isolation time



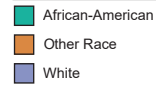
### Sex



### Age



### Race



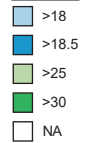
### Hospital



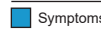
### Vaccine



### BMI



### Symptoms



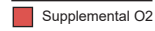
### Fever



### Hospitalized



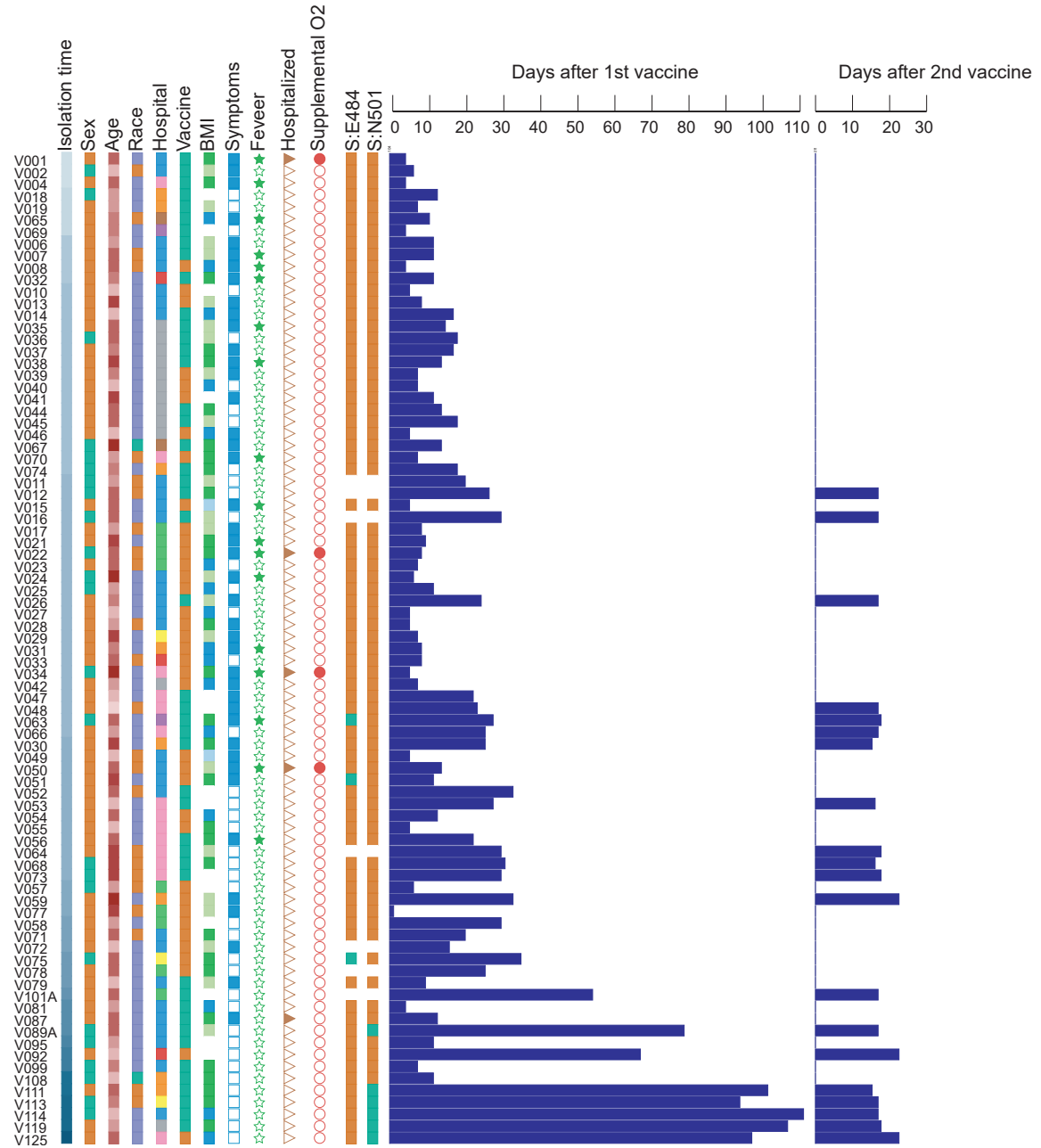
### Supplemental O2

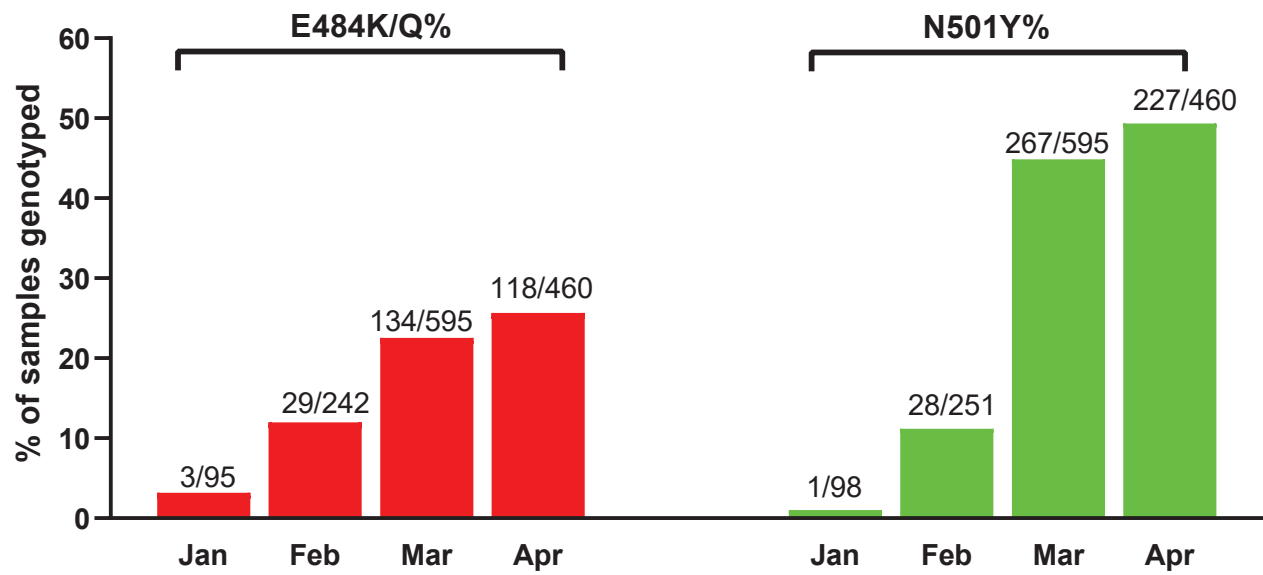


### S:484



### S:501





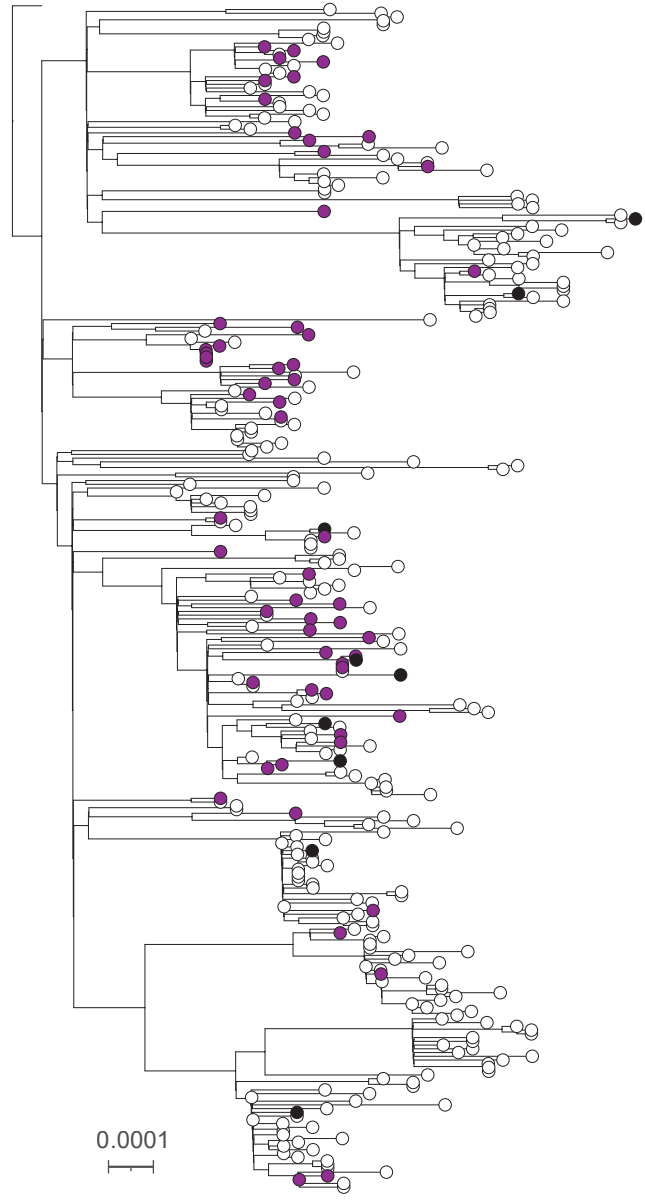


### Lineage

- B.1.1.434
- B.1.1.7
- B.1.2
- B.1.234
- B.1.243
- B.1.429
- B.1.526
- B.1.526.1
- B.1.526.2
- B.1.575
- B.1.596
- B.1.617.1
- P.1
- R.1
- Other

### NTD/RBD mutation

- Wild-type
- Mutation allele 1
- Mutation allele 2
- Mutation allele 3



### Lineage

