

1 **Rapid displacement of SARS-CoV-2 variants within Japan correlates with cycle threshold**
2 **values on routine RT-PCR testing**

3 **Authors:**

4 Danelle Wright^{1*}, Carmen Chan¹, Wirawit Chaochaisit¹, Mio Ogawa¹, Junko Tanaka¹, Satoshi
5 Nozaki¹, Shinji Narita¹, Eisuke Shimizu¹, Hideyuki Aoshima¹, Iri Sato Baran¹

6 **Affiliations:**

7 1) Genesis Institute of Genetic Research, Genesis Healthcare Corporation, Yebisu Garden
8 Place Tower 26F 4-20-3 Ebisu, Shibuya-ku, Tokyo, Japan

9

10 **Corresponding Author**

11 Danelle Wright

12 Genesis Institute of Genetic Research, Genesis Healthcare Corporation, Yebisu Garden Place

13 Tower 26F 4-20-3 Ebisu, Shibuya-ku, Tokyo, Japan

14 Email: Danelle_Wright@genesis-healthcare.jp

15

16 **Running head:** Ct values reflect variant prevalence

17 **Word counts:**

18 Abstract: 188

19 Text: 3572

20 **Conflict of Interest**

21 The authors declare that there are no conflicts of interest.

22 **Funding**

23 This work was supported by Genesis Healthcare Corporation. The sponsor had no influence on
24 the study other than providing resources and the facility.

25 **Corresponding Author**

26 Danelle Wright

27 Genesis Institute of Genetic Research, Genesis Healthcare Corporation, Yebisu Garden Place

28 Tower 26F 4-20-3 Ebisu, Shibuya-ku, Tokyo, Japan

29 Email: Danelle_Wright@genesis-healthcare.jp

30

31

32

33

34

35

36

37

38

39

40

41 **Abstract**

42 *Background:* The rapid spread of SARS-CoV-2 worldwide has led to the emergence of new
43 variants due to the presence of mutations that alter viral characteristics, but there have been few
44 studies on trends in viral lineages in Japan, an island country. We hypothesized that changes in
45 cycle threshold (Ct) values on reverse transcription polymerase chain reaction (RT-PCR) reflect
46 the prevalent variants during a given period.

47 *Methods:* We performed next-generation sequencing of positive samples to identify the viral
48 lineages in Japan in 2021 and compared variant prevalence with the average Ct values on routine
49 RT-PCR using 4 primer sets.

50 *Results:* Based on 3 sequencing runs, the highly transmissible Alpha variant, which prevailed
51 over other lineages, such as R.1, from April 2021, was dominated by the even stronger Delta
52 variant between July and August 2021 in Japan. The decrease in our routine RT-PCR Ct values
53 with 4 primer sets correlated with these fluctuations in lineage prevalence over time.

54 *Conclusions:* We confirmed that our RT-PCR protocol reflects the trends in SARS-CoV-2
55 variant prevalence over time regardless of sequence mutation. This may aid in the tracking of
56 new variants in the population.

57

58 **Key words:** SARS-CoV-2, Next-generation sequencing, Cycle threshold values, Variants of
59 concern

60

61

62

63 **Introduction**

64 Following the first report of the novel coronavirus, severe acute respiratory syndrome
65 coronavirus 2 (SARS-CoV-2), in 2019 in Wuhan, China (Zhou et al., 2020; Zhu et al., 2020), its
66 rapid spread worldwide has led to the emergence of new lineages and variants, some of which
67 are labeled variants of interest (VOI) or variants of concern (VOC) due to the presence of
68 mutations that alter viral characteristics such as transmissibility and fitness. In Japan, the first
69 report of a patient with coronavirus disease 2019 (COVID-19) was in January 2020 and multiple
70 clusters were identified in the following months. These clusters found in the first months directly
71 descended from the original Wuhan-Hu-1 strain (Sekizuka et al., 2020) and were considered to
72 have originated from the first wave in China, with PANGO lineage B.1.1.162 being prominent in
73 Japan. From March, the number of cases markedly increased worldwide, especially in the US
74 and Europe, and a similar increase in cases in Japan correlated with the outbreak in Europe
75 (Sekizuka et al., 2020). In September 2020, the B.1.1.7 lineage or Alpha variant became
76 prominent in Europe and was later detected in Japan, followed by the detection of the R.1 lineage
77 carrying the E484K mutation (Hirotsu and Omata, 2021; Sekizuka et al., 2021). Recently, the
78 Delta variant, or B.1.617.2, AY.1, AY.2 and AY.3 lineages, has become the most prevalent
79 among positive cases according to the Japanese Ministry of Health, overtaking Alpha and R.1.
80 The Delta variant is known to have increased transmissibility compared with Alpha and is less
81 sensitive to neutralization by antibodies (Campbell et al., 2021; Lucas et al., 2021; Planas et al.,
82 2021), making it a major VOC attenuating the efficacy of currently available vaccines.

83 This everchanging predominance of variants makes it difficult to screen for SARS-CoV-2 using
84 standard RT-PCR testing because mutations affect primer alignment. We previously verified the

85 use of a set of primers to increase the sensitivity and accuracy of SARS-CoV-2 testing (Tsutae et
86 al., 2020). The JPN-N2 primer set developed by the National Institute for Infectious Diseases,
87 Japan (NIID) (Shirato et al., 2020), and CDC-N1 and CDC-N2 primer sets developed by the US
88 Centers for Disease Control (CDC) (Lu et al., 2020) all detect the nucleocapsid region, but JPN-
89 N2 was reported to have higher sensitivity than CDC-N1 (Tsutae et al., 2020). The fourth primer
90 set included in our screening set, GH-E, detects the envelope region. JPN-N2 is used to surveille
91 positivity in Japan, whereas CDC-N2 is used in the U.S.; however, discrepancy in PCR results
92 has been reported based on which primers are used (Park et al., 2020; Peñarrubia et al., 2020),
93 leading to false-positives or false-negatives (Woloshin et al., 2020). With the continuing increase
94 in the prevalence and emergence of new variants, it is important to assess whether the currently
95 employed testing protocols can accurately detect positive cases. In addition, as the Delta variant
96 is considered to be approximately 60% more infectious than other known variants and has
97 reduced antibody sensitivity (Planas et al., 2021), a testing protocol that reflects the trends in
98 viral load is essential for preventing further infections caused by reduced detection due to
99 mutations affecting RT-PCR.

100 To date, there have been numerous studies assessing the spread of SARS-CoV-2 and the
101 emergence of new variants internationally, such as in the US and the UK (Bolze et al., 2021;
102 Martin et al., 2021; Volz et al., 2021a, 2021b), but few studies from Japan have examined the
103 recent domestic rapid evolution and variant predominance in detail. We performed next-
104 generation sequencing (NGS) of positive cases of SARS-CoV-2 detected by RT-PCR at the
105 Genesis Institute of Genetic Research, and carried out lineage assessment to detect trends in
106 variant predominance and their correlation with RT-PCR results. A previous study reported a
107 correlation between lower cycle threshold (Ct) values and higher viral load (Corman et al., 2020).

108 In this study, we confirmed a significant correlation between variant prevalence and Ct values,
109 with higher viral loads being observed for Delta cases using all primer sets, and the timing of the
110 increases in cases of different variants in Japan correlated with worldwide trends. This study also
111 demonstrated that our 4-primer set RT-PCR protocol is suitable for the accurate assessment of
112 SARS-CoV-2 positivity regardless of lineage. In addition, we report the first detection of
113 geographically specific AY lineages in Japan.

114

115 **Materials and Methods**

116 **Sample collection**

117 Nasopharyngeal swabs and saliva samples for testing were collected at the Genesis Institute of
118 Genetic Research between January and August 2021. Samples were collected from volunteers,
119 testing participants and patients under physician care undergoing clinical testing for SARS-CoV-
120 2 (COVID-19). Samples were anonymized, and laboratory technicians and researchers were
121 blinded to the identity of the patients. Only non-identifying data, such as age, sex and symptoms,
122 were provided. All experimental protocols were approved by the ethics committee of Genesis
123 Healthcare. RNA was extracted using the MGISP-960 system (MGI Tech Co. Ltd.) and RT-PCR
124 was performed using the following primer sets according to the previously reported protocol
125 (Tsutae et al., 2020): JPN-N2, CDC-N1, CDC-N2, GH-E. Primer and probe sequences are listed
126 in Table 1, and binding sites are shown in Fig. 1A. The 730 total samples positive by the JPN-N2
127 primer set in duplicate or all four primer sets were stored at -80°C until sequencing.

128 **SARS-CoV-2 sequencing (COVIDseq)**

129 SARS-CoV-2 RNA from the 730 collected samples was reverse-transcribed into cDNA, and
130 tagged and pooled to create libraries following the illumina COVIDSeq Test Reference
131 Guide. Samples were sequenced using the illumina NovaSeq 6000 system SP flow cell,
132 comprising the NovaSeq 6000 SP reagent kit v.1.5 and NovaSeq Xp 2-lane kit v.1.5, with 100
133 cycles for the first run, 200 cycles for the second run and 200 cycles for the third run. The
134 resultant sequence data in the SampleSheet.csv file were corrected strictly following the
135 SampleSheet COVIDSeq protocol ('illumina-adapter-sequences-1000000002694-15.pdf') and
136 uploaded onto the illumina DRAGEN Bio-IT Platform. The illumina DRAGEN COVIDSeq
137 Test (RUO) App was used for consensus sequence calling and a consensus genome in FASTA
138 format was generated for samples with a result of "SARS-CoV-2 Detected" (for details, see
139 'dragen-covidseq-test-ruo-v130-app-guide-1000000129548-01.pdf'). These FASTA files were
140 next input into the illumina DRAGEN COVID Lineage App. The NextClade tool (Nextstrain,
141 Nextclade: <https://clades.nextstrain.org/>) was used for clade assignment and phylogenetic
142 analyses, and Pangolin (<https://cov-lineages.org/>; corresponding GitHub
143 repository: <https://github.com/cov-lineages>) was used for lineage calling (For more information,
144 see [https://www.illumina.com/products/by-type/informatics-products/basespace-sequence-](https://www.illumina.com/products/by-type/informatics-products/basespace-sequence-hub/apps/dragen-covid-lineage.html)
145 [hub/apps/dragen-covid-lineage.html](https://www.illumina.com/products/by-type/informatics-products/basespace-sequence-hub/apps/dragen-covid-lineage.html)). Sequence quality control (QC) parameters included
146 missing data, mixed sites, private mutations and mutation clusters. Only samples that passed QC
147 for lineage calling and clade assessment were used in subsequent analyses.

148 **Statistical analysis**

149 Ct values for each primer set among variants were compared using the Kruskal-Wallis H test
150 (one-way non-parametric ANOVA) followed by post hoc Conover's test. Values are presented as
151 the mean \pm standard error. Differences between the expected and observed frequencies of

152 symptoms among the Alpha, Delta and ‘others’ variants were assessed using the chi-squared test.
153 A p-value <0.05 was considered significant.

154

155 **Results**

156 **Trends in SARS-CoV-2 infections during 2021 in Japan based on NGS**

157 From January 2021, the R.1 lineage was detected in up to approximately 38% of positive cases in
158 Japan after its first report in November 2020 (Hirotsu and Omata, 2021). As R.1 was first
159 detected in the U.S. a month earlier, this change in R.1 prevalence in Japan and other countries
160 was similar (Hirotsu and Omata, 2021). In addition, an increase in B.1.1.7 or Alpha cases was
161 reported in March 2021 during the fourth wave and B.1.617 or Delta was first detected in April
162 according to the Japanese Ministry of Health. As we perform routine commercial testing of
163 SARS-CoV-2, we hypothesized that our positive cases reflect the prevalent variants during a
164 given period. We therefore performed NGS of positive samples received for SARS-CoV-2
165 screening at the Genesis Institute of Genetic Research to identify the viral lineages causing
166 infection in the Japanese public in 2021. Three separate COVIDseq runs were conducted using
167 samples that were positive for SARS-CoV-2 by all 4 primer sets shown in Fig. 1A collected from
168 January to April in the 1st run, April to June in the 2nd, and July to August in the 3rd. In total, 730
169 samples were sequenced. After QC assessment for lineage calling and clade assignment, 620
170 samples were used for analysis. Of the 620 samples that passed QC, 190 were tested only using
171 the JPN-N2 primer set in duplicate and 430 samples were tested using the four-primer set and
172 were positive with all four sets. Based on phylogenetic clade assignment of known VOC and
173 VOI, 195 samples were determined to be Alpha, 28 were R.1, 255 were Delta and 142 were

174 other lineages (Fig. 1B, Supp. Figs. 1, 2, 3). Alpha accounted for 47% of cases collected between
175 January and June, and 10% of cases during this period were R.1 (Fig. 1C). However, in the 3rd
176 run using samples from July to August, only 17% of cases were Alpha and no cases of R.1 were
177 detected, suggesting that these lineages were overtaken, most likely by Delta. Delta was first
178 detected in our samples in the end of June and accounted for 78% of cases in the 3rd run,
179 demonstrating a rapid increase in prevalence (Fig. 1C). Delta lineages were separated into
180 B.1.617.2 and AY.4-AY.25, with the majority being the former (Fig. 1C). This suggested that the
181 different variants displaced each other within a short period of around two months according to
182 each wave of infection.

183 **SARS-CoV-2 VOC are associated with a higher frequency of symptoms**

184 It is well known that viral mutations can cause more severe cases of COVID-19 (Hahn et al.,
185 2020; Nagy et al., 2021; Voss et al., 2021). We thus analyzed the reported symptoms and
186 compared their frequencies between each VOC and the “other” lineages. At the time of
187 sampling, the presence of the following symptoms was asked: fever, tiredness, chills, headache,
188 muscle pain, dyspnea, cough, sore throat, runny nose, abnormal olfaction/taste, diarrhea, nausea,
189 and rash. The most commonly reported symptom in R.1 and ‘others’ cases was cough in
190 approximately 20% and 15%, respectively, whereas that in Alpha and Delta cases was fever in
191 approximately 24% and 34%, respectively (Fig. 2). Consistent with the reports of more severe
192 cases with VOC infection, Alpha and Delta cases had a significantly higher rate of fever and
193 cough than ‘others’ cases ($p < 0.05$). Furthermore, headache and tiredness were significantly
194 more frequently reported in Alpha and Delta cases ($p < 0.05$), suggesting an increase in severity.
195 Of note, approximately 15% of ‘others’ cases reported abnormal olfaction/taste, which is often
196 one of the first symptoms noted in mild COVID-19 (Spinato et al., 2020). However, it was

197 reported slightly less in Delta and Alpha cases (Fig. 2). Muscle pain and dyspnea were also more
198 common in Delta and Alpha cases, although not significantly.

199 We next examined the relationship between Ct values and the number of reported symptoms at
200 the time of sampling, but no significant trend was observed in our data (Supp. Fig. 4).

201 Based on these results, VOC may cause higher rates of commonly observed symptoms that are
202 not specific to SARS-CoV-2 infection.

203 **Increased viral load is associated with the increased prevalence of Delta cases**

204 The rapid spread of VOC is considered to be due to an increased viral load via faster replication
205 rates (Campbell et al., 2021; Li et al., 2021; Volz et al., 2021a). Alpha was reported to have
206 higher transmissibility than non-VOC and Delta was reported to have even higher
207 transmissibility due to an increased viral load (Li et al., 2021; Wintersdorff et al., 2021).

208 Consistent with this, we noted a significant difference in Ct values with all RT-PCR primer sets
209 among variants. With the JPN-N2 primer set, the mean Ct value for ‘others’ lineages was
210 30.1 ± 0.3 , that for R.1 was 27.5 ± 1.1 , that for Alpha was 25.9 ± 0.3 and that for Delta was 24.4 ± 0.3
211 (Fig. 3A). With the CDC-N1 primer set, the mean Ct values were 29.6 ± 0.5 , 27.2 ± 1.1 , 25.4 ± 0.4
212 and 23.7 ± 0.4 , respectively (Fig. 3B). With the CDC-N2 primer set, the mean Ct values were
213 30.1 ± 0.5 , 27.1 ± 1.1 , 25.8 ± 0.4 and 23.9 ± 0.4 , respectively (Fig. 3C). Similarly, with the GH-E
214 primer set, the mean Ct values were 30.9 ± 0.5 , 26.8 ± 1.2 , 26.9 ± 0.4 and 25.3 ± 0.4 , respectively (Fig.
215 3D). Each variant had significantly lower Ct values than the ‘others’, with Delta having the
216 lowest Ct values among all variants, consistent with its higher viral load and infectivity. As it
217 was previously reported that Ct values >34 are the limit for viral isolation from samples (La
218 Scola et al., 2020), our samples were from contagious individuals.

219 This decrease in Ct correlated with the trends in viral prevalence over time, demonstrating a
220 decline with each successive wave of infection spurred by the introduction of each VOC. As
221 shown in Fig. 4A, the mean Ct using our 4-primer set screening method ranged between
222 approximately 28 and 35 from January to the end of February, during which ‘others’ and R.1
223 lineages comprised the majority of cases (Figs. 1C, 4A). Coinciding with the spread of Alpha in
224 the fourth wave, the mean Ct markedly decreased to around 24-28 during March and April,
225 indicating its higher infectivity. From the last week of May to June, there was a temporary drop
226 in Alpha cases, and the mean Ct increased back to around 30 accordingly. Then, from the first
227 week of June until August, the mean Ct values further dropped to around 24-26 in parallel with
228 the introduction and spread of Delta (Figs. 1C, 4A).

229 To confirm that the changes in Ct values were due to increased viral load rather than
230 mutations affecting primer binding, we examined COVIDseq data for any mutations in the
231 primer binding sites. Among the 4 primer sets, only 3 of the 620 sequenced samples had a
232 mutation at primer binding sites (Table 2). Two of these samples were Delta and one was Alpha,
233 but only one of these mutations, specifically G29179T in the CDC-N2 forward primer binding
234 site, was considered to reduce binding. This confirmed that the lower Ct values in VOC cases
235 were due to increased viral load. As viral loads in Delta cases were reported to be 1000-times
236 higher than those in non-VOC cases (Li et al., 2021), the Ct values observed using our RT-PCR
237 method reflected the real-time trends in the prevalence of VOC.

238 **Delta displaced Alpha as the major variant in Japan within 1 month**

239 Although travel restrictions, quarantine and vaccination measures were implemented worldwide,
240 Delta rapidly established itself as the dominant variant in countries like India, the UK, the
241 Netherlands and Australia, (Campbell et al., 2021; Wintersdorff et al., 2021) and displaced Alpha

242 in the US by July 2021 (Bolze et al., 2021). Alpha was the dominant VOC detected in our
243 samples from March to June (Figs. 1C, 4A). The first reported case of Delta in Japan was in
244 April 2021, whereas we first detected the Delta variant, specifically B.1.617.2, in mid-June in a
245 sample from Chiba prefecture (Figs. 1C, 4A). By mid-July, Delta accounted for 50% of all
246 positive samples sequenced and its prevalence continued to increase to nearly 90% into August,
247 overtaking Alpha (Figs. 1C, 4A). Therefore, although expected based on previous worldwide
248 surveillance and its increased transmission rate, our sequencing results confirmed that the Delta
249 variant rapidly dominated Japan.

250 **Detection of country-specific Delta lineages in Japan**

251 In addition to B.1.617.2, we detected Delta lineages AY.4-AY.25 (Fig. 4B), which are subclades
252 within B.1.617.2 that exhibit significant geographic clustering, although there is no functional
253 biological difference. As of August 21, 2021, AY.1-AY.25 lineages are included in the
254 pangoleARN model, with lineages AY.4-AY.11 being listed as predominant in the UK. During
255 our third sequencing run using samples collected in July and August, we detected 23 cases of the
256 Delta lineage AY.7.1, which is specific to Denmark. There were also 19 cases of AY.4, 4 cases
257 of AY.5, 9 cases of AY.12, 7 cases of AY.25 and 1 case of AY.24 (Fig. 4B). These lineages are
258 most prominent in the UK, Israel, the US and Indonesia, respectively (O'Toole et al., 2021).
259 Considering their geographical diversity, these lineages may have been introduced into Japan
260 through international travel from July 2021, but this point requires further epidemiological study.

261 **Discussion**

262 During 2020, the first (March to April 2020), second (July to September 2020), and third
263 (October 2020 to January 2021) waves of COVID-19 in Japan were characterized by B.1.1.162,

264 B.1.1.284, and B.1.1.214 lineages, respectively (Sekizuka et al., 2020). According to a previous
265 study, variants of SARS-CoV-2 evolve through natural selection and dominate the population
266 (Hou et al., 2020). NGS analysis has made it possible to trace the spread of SARS-CoV-2 and
267 identify specific lineages infecting a population. This is particularly important as the emergence
268 of mutations has led to the rapid spread of VOC and VOI that may be less sensitive to available
269 vaccines. We used Ct values from RT-PCR and NGS data to investigate the trends in variant
270 prevalence in Japan based on samples collected from January to August 2021. The most
271 prevalent strain changed within an approximately two-month period between sequencing runs.
272 From the period of January to March, ‘others’ and R.1 lineages were most prevalent among our
273 samples. R.1 was first reported in the US in October 2020 and in Japan in November, but the
274 reported isolates were phylogenetically similar; therefore, it is possible that the Japanese isolate
275 originated from the US isolate (Sekizuka et al., 2020). By mid-April 2021, nearly 40% of the R.1
276 lineage entries in the Global Initiative on Sharing All Influenza Data (GISAID) database were
277 from Japan (Hirotsu and Omata, 2021). We observed the highest number of cases at the end of
278 March and into April, which was consistent with the peak incidence in Japan and the US
279 according to international lineage tracking at cov-lineages.org (O’Toole et al., 2021).

280 However, despite the rapid spread of R.1, it was quickly dominated by the more
281 transmissible Alpha in Japan and overseas (Washington et al., 2021). We first noted a significant
282 drop in Ct values mid-March, which coincided not only with an increase in Alpha cases detected
283 by our sequencing, but also with the worldwide increase based on lineage tracking. The use of Ct
284 values to assess mortality, prognosis, disease severity and infectivity has been reviewed in many
285 previous studies (Rabaan et al., 2021; Rao et al., 2020), but it is well known that they are
286 affected by many factors, including primer design. Our RT-PCR assay uses 4 primer sets, 3 of

287 which detect the nucleocapsid region, in order to increase the detection sensitivity and specificity
288 while using the same PCR program. We previously reported that the CDC-N1 primer set is prone
289 to false-positives, and should not be used alone without CDC-N2 or JPN-N2, which have a
290 higher accuracy in the Japanese population (Tsutae et al., 2020). In addition, among all 620
291 sequenced samples, only one Alpha case carried a mutation that may have reduced CDC-N2
292 primer binding, leading to a single higher Ct value (Table 2). The overall observed changes in Ct
293 values with all primer sets were congruent with changes in the lineages isolated over time, with
294 lower Ct values correlating with the increase in cases of more infective Alpha and Delta variants,
295 indicating that the varying mutations in these lineages did not significantly affect our RT-PCR.
296 Of note, we only observed a significant correlation between the number of symptoms and
297 presence of VOC, but not between Ct values and symptoms. This is consistent with a study that
298 reported similar Ct values between asymptomatic and preymptomatic cases (Singanayagam et al.,
299 2020), as well as a study in the UK stating that Ct values ultimately depend on the virus strain in
300 the population (Walker et al., 2021). Therefore, our study revealed that Ct values of multiple
301 primer sets can be a good reflection of the transmissibility of the most prominent SARS-CoV-2
302 lineage in a population at a given time. Although our study has the limitations of symptom onset
303 data being unavailable and the timing of sample collection varying throughout the course of
304 infection, it suggested that a rapid decrease in Ct values in routine testing indicates the
305 introduction of a new, more infective variant, whereas an increase may indicate its fading out
306 from a population.

307 The R.1 variant, which was reported to have potential escape mutations in the N-terminal
308 domain (W152L) and spike protein (E484K) (Hirotsu and Omata, 2021), was recently re-
309 detected in the US amidst the rampant spread of Delta (Cavanaugh et al., 2021). This variant was

310 not detected in our samples after June, suggesting that variants can reemerge in different
311 populations worldwide. Variants harboring E484K, L452R, or N501Y are of particular concern
312 because they are related to escape from antibody neutralization (Cherian et al., 2021; Lucas et al.,
313 2021). As such, even after vaccination or previous infection, if these variants are reintroduced to
314 a population, they may become dominant again. As nearly 80% of the Japanese population has
315 received two doses of the SARS-CoV-2 vaccine as of February 2022, it will be of interest to
316 investigate how the presence of antibodies affects the average Ct values and transmissibility of
317 known and new variants such as Omicron. Furthermore, as the third vaccine is just now being
318 administered in Japan at present, it is important to identify mutations in highly transmissible
319 strains in order to prevent the spread of even more infectious variants that escape antibody
320 neutralization. Our NGS data will be highly useful for investigating which mutations were
321 conserved in the SARS-CoV-2 sequence throughout the rapid fluctuation of lineages present in
322 Japan, aiding in the estimation of future VOC through machine learning.

323

324 **Competing Interests**

325 The authors declare that there are no competing interests.

326

327 **Ethics Statement**

328 Samples were collected from volunteers, testing participants and patients under physician care
329 undergoing clinical testing for SARS-CoV-2 (COVID-19). Samples were anonymized, and
330 laboratory technicians and researchers were blinded to the identity of the patients. Only non-
331 identifying data, such as age, sex and symptoms, were provided. All experimental protocols were
332 approved by the ethics committee of Genesis Healthcare.

333

334 **Data availability**

335 NGS data from the first and second COVIDseq runs were uploaded to GISAID.

336

337 **Funding**

338 This study was supported by Genesis Healthcare Corporation. The sponsor had no influence on
339 the study other than providing resources and the facility.

340 **Acknowledgements**

341 None.

342

343 **Author Contributions**

344 I.S.B. conceived the study. D.W., J.T., and H.A. processed samples and performed RT-PCR.
345 D.W, M.O., S.N., E.S., and J.T. performed next-generation sequencing. C.C. and W.O. carried
346 out data analysis and lineage assignment. D.W. wrote the manuscript. All authors reviewed the
347 data and approved the manuscript.

348

349 **References**

350 Bolze A, Cirulli ET, Luo S, White S, Wyman D, Dei Rossi A, et al. SARS-CoV-2 variant Delta
351 rapidly displaced variant Alpha in the United States and led to higher viral loads. *Infectious*
352 *Diseases (except HIV/AIDS)*; 2021. <https://doi.org/10.1101/2021.06.20.21259195>.

353 Campbell F, Archer B, Laurenson-Schafer H, Jinnai Y, Konings F, Batra N, et al. Increased
354 transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021.
355 *Eurosurveillance* 2021;26:2100509. <https://doi.org/10.2807/1560-7917.ES.2021.26.24.2100509>.
356 Cavanaugh AM, Fortier S, Lewis P, Arora V, Johnson M, George K, et al. COVID-19 Outbreak
357 Associated with a SARS-CoV-2 R.1 Lineage Variant in a Skilled Nursing Facility After
358 Vaccination Program — Kentucky, March 2021. *Morb Mortal Wkly Rep* 2021;70:639–43.
359 <https://doi.org/10.15585/mmwr.mm7017e2>.
360 Cherian S, Potdar V, Jadhav S, Yadav P, Gupta N, Das M, et al. Convergent evolution of SARS-
361 CoV-2 spike mutations, L452R, E484Q and P681R, in the second wave of COVID-19 in
362 Maharashtra, India. 2021. <https://doi.org/10.1101/2021.04.22.440932>.
363 Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, et al. Detection of 2019
364 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance* 2020;25.
365 <https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045>.
366 Hahn G, Wu CM, Lee S, Hecker J, Lutz SM, Haneuse S, et al. Two mutations in the SARS-
367 CoV-2 spike protein and RNA polymerase complex are associated with COVID-19 mortality
368 risk. 2020. <https://doi.org/10.1101/2020.11.17.386714>.
369 Hirotsu Y, Omata M. Detection of R.1 lineage severe acute respiratory syndrome coronavirus 2
370 (SARS-CoV-2) with spike protein W152L/E484K/G769V mutations in Japan. *PLoS Pathog*
371 2021;17:e1009619. <https://doi.org/10.1371/journal.ppat.1009619>.
372 Hou YJ, Chiba S, Halfmann P, Ehre C, Kuroda M, Dinno KH, et al. SARS-CoV-2 D614G
373 variant exhibits efficient replication ex vivo and transmission in vivo. *Science* 2020;370:1464–8.
374 <https://doi.org/10.1126/science.abe8499>.
375 La Scola B, Le Bideau M, Andreani J, Hoang VT, Grimaldier C, Colson P, et al. Viral RNA load
376 as determined by cell culture as a management tool for discharge of SARS-CoV-2 patients from
377 infectious disease wards. *Eur J Clin Microbiol Infect Dis* 2020;39:1059–61.
378 <https://doi.org/10.1007/s10096-020-03913-9>.
379 Li B, Deng A, Li K, Hu Y, Li Z, Xiong Q, et al. Viral infection and transmission in a large, well-
380 traced outbreak caused by the SARS-CoV-2 Delta variant. 2021.
381 <https://doi.org/10.1101/2021.07.07.21260122>.
382 Lu X, Wang L, Sakthivel SK, Whitaker B, Murray J, Kamili S, et al. US CDC Real-Time
383 Reverse Transcription PCR Panel for Detection of Severe Acute Respiratory Syndrome
384 Coronavirus 2 - Volume 26, Number 8—August 2020 - *Emerging Infectious Diseases journal* -
385 CDC. *Emerg Infect Dis* 2020;26. <https://doi.org/10.3201/eid2608.201246>.
386 Lucas C, Vogels CBF, Yildirim I, Rothman JE, Lu P, Monteiro V, et al. Impact of circulating
387 SARS-CoV-2 variants on mRNA vaccine-induced immunity in uninfected and previously
388 infected individuals. 2021. <https://doi.org/10.1101/2021.07.14.21260307>.
389 Martin DP, Weaver S, Tegally H, San JE, Shank SD, Wilkinson E, et al. The emergence and
390 ongoing convergent evolution of the SARS-CoV-2 N501Y lineages. *Cell* 2021;184:5189-
391 5200.e7. <https://doi.org/10.1016/j.cell.2021.09.003>.
392 Nagy Á, Pongor S, Györfy B. Different mutations in SARS-CoV-2 associate with severe and
393 mild outcome. *Int J Antimicrob Agents* 2021;57:106272.
394 <https://doi.org/10.1016/j.ijantimicag.2020.106272>.
395 O'Toole Á, Hill V, Pybus OG, Watts A, Bogoch II, Khan K, et al. Tracking the international
396 spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2 with grinch 2021.
397 <https://doi.org/10.12688/wellcomeopenres.16661.2>.

398 Park M, Won J, Choi BY, Lee CJ. Optimization of primer sets and detection protocols for
399 SARS-CoV-2 of coronavirus disease 2019 (COVID-19) using PCR and real-time PCR. *Exp Mol*
400 *Med* 2020;52:963–77. <https://doi.org/10.1038/s12276-020-0452-7>.
401 Peñarrubia L, Ruiz M, Porco R, Rao SN, Juanola-Falgarona M, Manissero D, et al. Multiple
402 assays in a real-time RT-PCR SARS-CoV-2 panel can mitigate the risk of loss of sensitivity by
403 new genomic variants during the COVID-19 outbreak. *Int J Infect Dis* 2020;97:225–9.
404 <https://doi.org/10.1016/j.ijid.2020.06.027>.
405 Planas D, Veyer D, Baidaliuk A, Staropoli I, Guivel-Benhassine F, Rajah MM, et al. Reduced
406 sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. *Nature* 2021;596:276–80.
407 <https://doi.org/10.1038/s41586-021-03777-9>.
408 Rabaan AA, Tirupathi R, Sule AA, Aldali J, Mutair AA, Alhumaid S, et al. Viral Dynamics and
409 Real-Time RT-PCR Ct Values Correlation with Disease Severity in COVID-19. *Diagnostics*
410 2021;11:1091. <https://doi.org/10.3390/diagnostics11061091>.
411 Rao SN, Manissero D, Steele VR, Pareja J. A Narrative Systematic Review of the Clinical
412 Utility of Cycle Threshold Values in the Context of COVID-19. *Infect Dis Ther* 2020;9:573–86.
413 <https://doi.org/10.1007/s40121-020-00324-3>.
414 Sekizuka T, Itokawa K, Hashino M, Kawano-Sugaya T, Tanaka R, Yatsu K, et al. A genome
415 epidemiological study of SARS-CoV-2 introduction into Japan. 2020.
416 <https://doi.org/10.1101/2020.07.01.20143958>.
417 Sekizuka T, Itokawa K, Hashino M, Okubo K, Ohnishi A, Goto K, et al. A Discernable Increase
418 in the Severe Acute Respiratory Syndrome Coronavirus 2 R.1 Lineage Carrying an E484K Spike
419 Protein Mutation in Japan. 2021. <https://doi.org/10.1101/2021.04.04.21254749>.
420 Shirato K, Nao N, Katano H, Takayama I, Saito S, Kato F, et al. Development of Genetic
421 Diagnostic Methods for Detection for Novel Coronavirus 2019(nCoV-2019) in Japan. *Jpn J*
422 *Infect Dis* 2020;73:304–7. <https://doi.org/10.7883/yoken.JJID.2020.061>.
423 Singanayagam A, Patel M, Charlett A, Lopez Bernal J, Saliba V, Ellis J, et al. Duration of
424 infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19,
425 England, January to May 2020. *Eurosurveillance* 2020;25:2001483.
426 <https://doi.org/10.2807/1560-7917.ES.2020.25.32.2001483>.
427 Spinato G, Fabbris C, Polesel J, Cazzador D, Borsetto D, Hopkins C, et al. Alterations in Smell
428 or Taste in Mildly Symptomatic Outpatients With SARS-CoV-2 Infection. *JAMA*
429 2020;323:2089–90. <https://doi.org/10.1001/jama.2020.6771>.
430 Tsutae W, Chaochaisit W, Aoshima H, Ida C, Miyakawa S, Sekine H, et al. Detecting and
431 isolating false negatives of SARS-CoV-2 primers and probe sets among the Japanese Population:
432 A laboratory testing methodology and study. *Infectious Diseases (except HIV/AIDS)*; 2020.
433 <https://doi.org/10.1101/2020.10.07.20208264>.
434 Volz E, Mishra S, Chand M, Barrett JC, Johnson R, Geidelberg L, et al. Transmission of SARS-
435 CoV-2 Lineage B.1.1.7 in England: Insights from linking epidemiological and genetic data.
436 *Infectious Diseases (except HIV/AIDS)*; 2021a. <https://doi.org/10.1101/2020.12.30.20249034>.
437 Volz E, Mishra S, Chand M, Barrett JC, Johnson R, Geidelberg L, et al. Assessing
438 transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature* 2021b;593:266–9.
439 <https://doi.org/10.1038/s41586-021-03470-x>.
440 Voss JD, Skarzynski M, McAuley EM, Maier EJ, Gibbons T, Fries AC, et al. Variants in SARS-
441 CoV-2 associated with mild or severe outcome. *Evol Med Public Health* 2021;9:267–75.
442 <https://doi.org/10.1093/emph/eoab019>.

443 Walker AS, Pritchard E, House T, Robotham JV, Birrell PJ, Bell I, et al. Ct threshold values, a
444 proxy for viral load in community SARS-CoV-2 cases, demonstrate wide variation across
445 populations and over time. *ELife* 2021;10:e64683. <https://doi.org/10.7554/eLife.64683>.
446 Washington NL, Gangavarapu K, Zeller M, Bolze A, Cirulli ET, Barrett KMS, et al. Genomic
447 epidemiology identifies emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United
448 States. 2021. <https://doi.org/10.1101/2021.02.06.21251159>.
449 Wintersdorff C von, Dingemans J, Alphen L van, Wolffs P, Veer B van der, Hoebe C, et al.
450 Infections caused by the Delta variant (B.1.617.2) of SARS-CoV-2 are associated with increased
451 viral loads compared to infections with the Alpha variant (B.1.1.7) or non-Variants of Concern.
452 2021. <https://doi.org/10.21203/rs.3.rs-777577/v1>.
453 Woloshin S, Patel N, Kesselheim AS. False Negative Tests for SARS-CoV-2 Infection —
454 Challenges and Implications. *N Engl J Med* 2020. <https://doi.org/10.1056/NEJMp2015897>.
455 Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak
456 associated with a new coronavirus of probable bat origin. *Nature* 2020;579:270–3.
457 <https://doi.org/10.1038/s41586-020-2012-7>.
458 Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with
459 Pneumonia in China, 2019. *N Engl J Med* 2020;382:727–33.
460 <https://doi.org/10.1056/NEJMoa2001017>.

461

462

463

464

465 **Figure Legends**

466 **Figure 1. Trends in SARS-CoV-2 lineage prevalence over time**

467 A) Binding sites of each primer set in the SARS-CoV-2 sequence. JPN-N2, CDC-N1 and CDC-
468 N2 sets bind in the nucleocapsid region. The GH-E set binds in the envelope region. B) Detailed
469 lineage breakdown of all collected samples. C) Percentage of detected variants among sequenced
470 samples per week of 2021. Brackets indicate the period during which samples were collected for
471 each COVIDseq run.

472 **Figure 2. Reported symptoms associated with each variant**

473 Frequencies of reported symptoms in cases of each variant. Differences were assessed by the chi-
474 squared test. * indicates significantly more often compared with the ‘others’ group.

475 **Figure 3. Differences in Ct values among variants by RT-PCR primer set**

476 A) Ct values with the JPN-N2 primer set for each variant. B) The mean \pm standard error Ct value
477 for each lineage with the JPN-N2 primer set. C) Ct values with the CDC-N1 primer set for each
478 variant. D) The mean \pm standard error Ct value for each lineage with the CDC-N1 primer set. E)
479 Ct values with the CDC-N2 primer set for each variant. F) The mean \pm standard error Ct value
480 for each lineage with the CDC-N2 primer set. G) Ct values with the GH-E primer set for each
481 variant. H) The mean \pm standard error Ct value for each lineage with the GH-E primer set. Ct
482 values were compared among variants using the Kruskal-Wallis H test (one-way non-parametric
483 ANOVA) followed by post hoc Conover’s test. P-values with an * indicate significantly different
484 from the ‘others’ group.

485 **Figure 4. Correlation between SARS-CoV-2 variants and RT-PCR Ct values**

486 A) Graph of the mean Ct value with each primer set for each variant over time. B) Detailed
487 breakdown of lineages detected during COVIDseq run 3.

488 **Supplementary Figure 1. Phylogenetic tree for COVIDseq run 1**

489 Phylogenetic tree for samples collected between January and April 2021 generated by
490 <https://clades.nextstrain.org/tree>.

491 **Supplementary Figure 2. Phylogenetic tree for COVIDseq run 2**

492 Phylogenetic tree for samples collected between April and June 2021 generated by
493 <https://clades.nextstrain.org/tree>.

494 **Supplementary Figure 3. Phylogenetic tree for COVIDseq run 3**

495 Phylogenetic tree for samples collected between June and August 2021 generated by

496 <https://clades.nextstrain.org/tree>.

497 **Supplementary Figure 4. Plot of the number of symptoms at sampling versus Ct values**

498 Graph of the average number of symptoms reported at the time of sampling versus rounded Ct

499 value. No significant trend was observed.

Figure 1

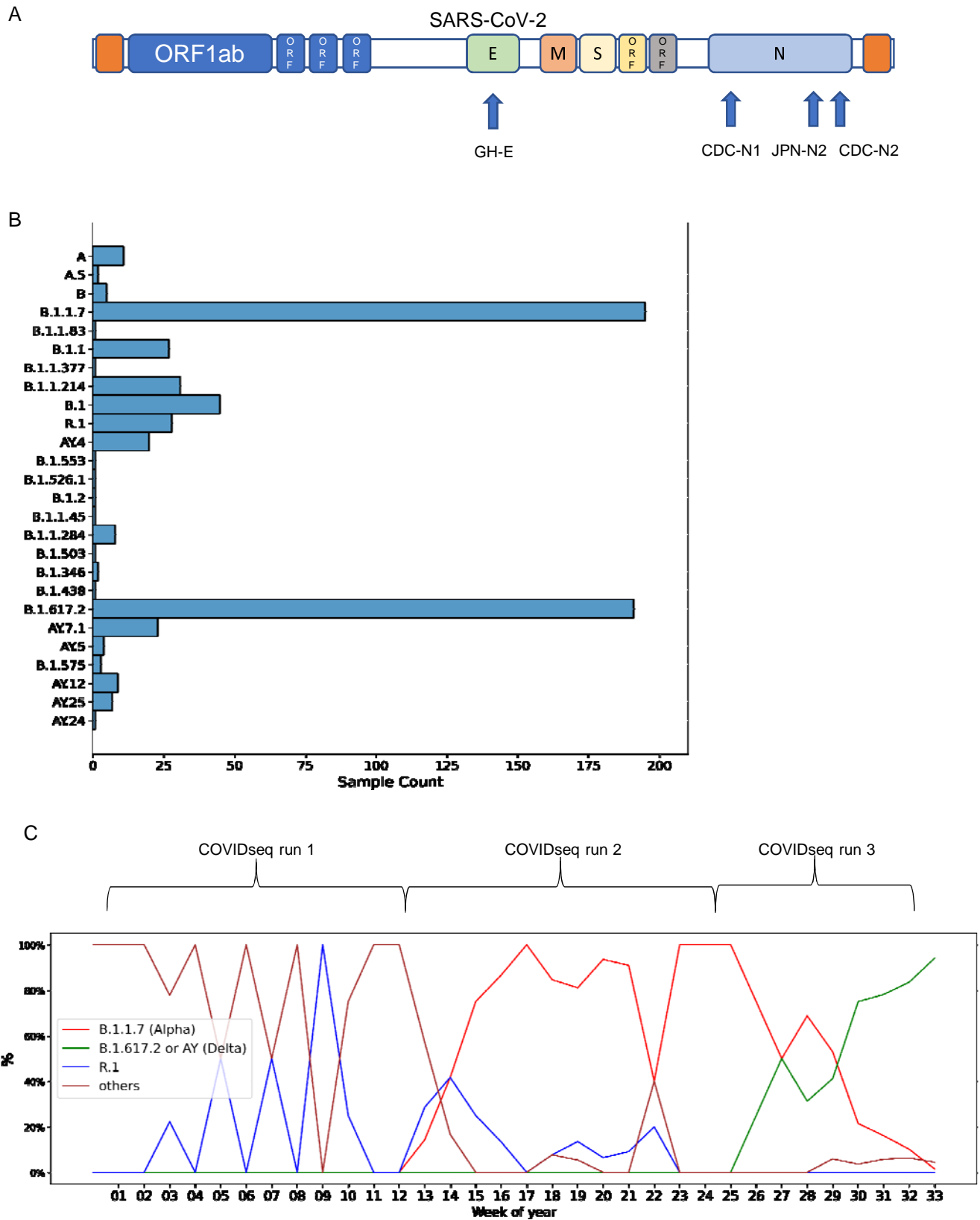


Figure 2

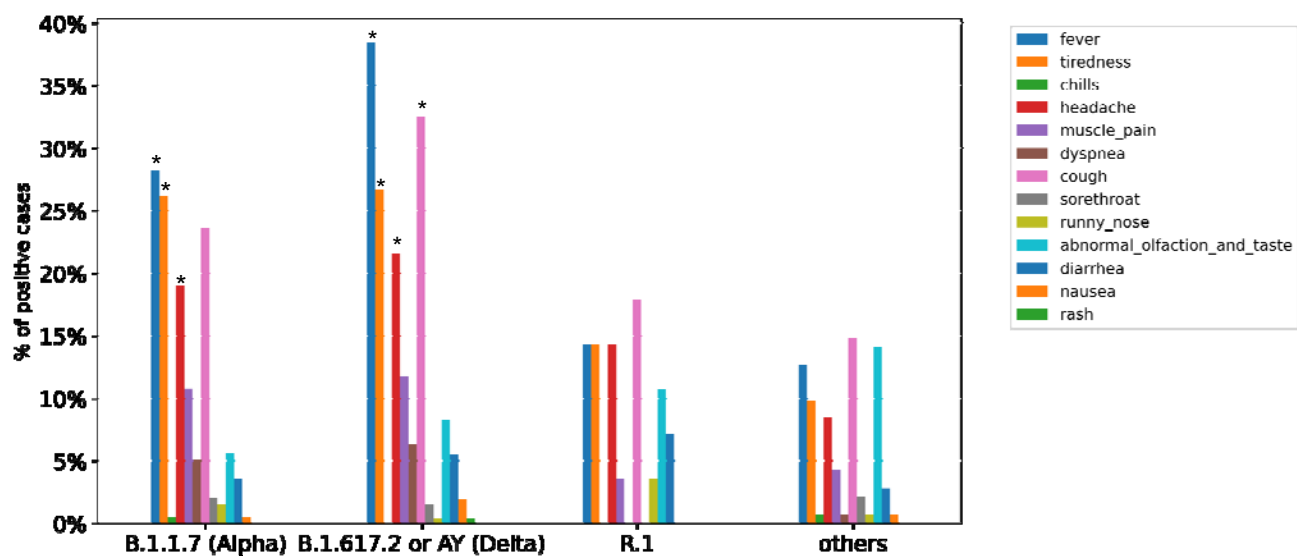


Figure 3

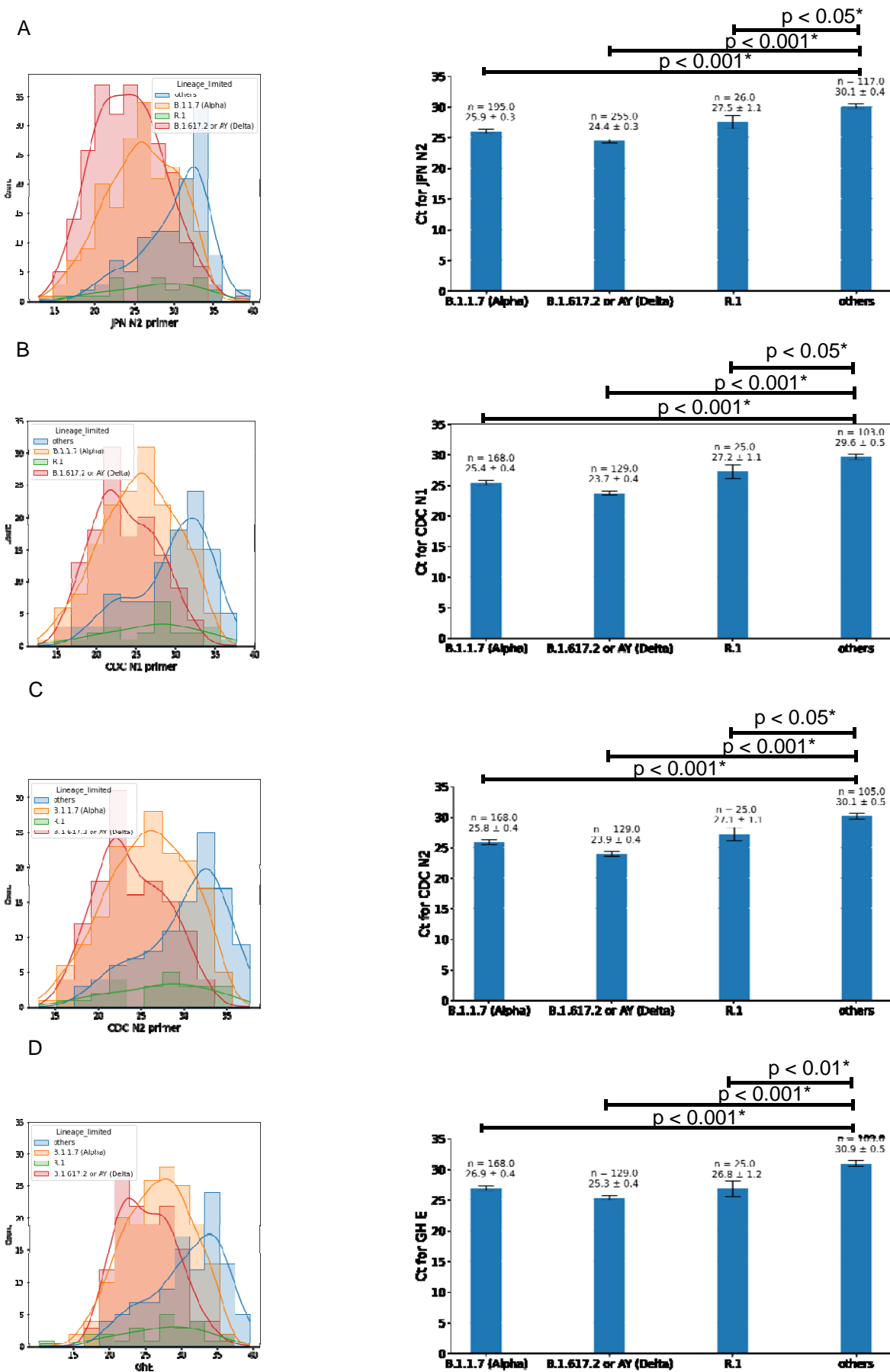
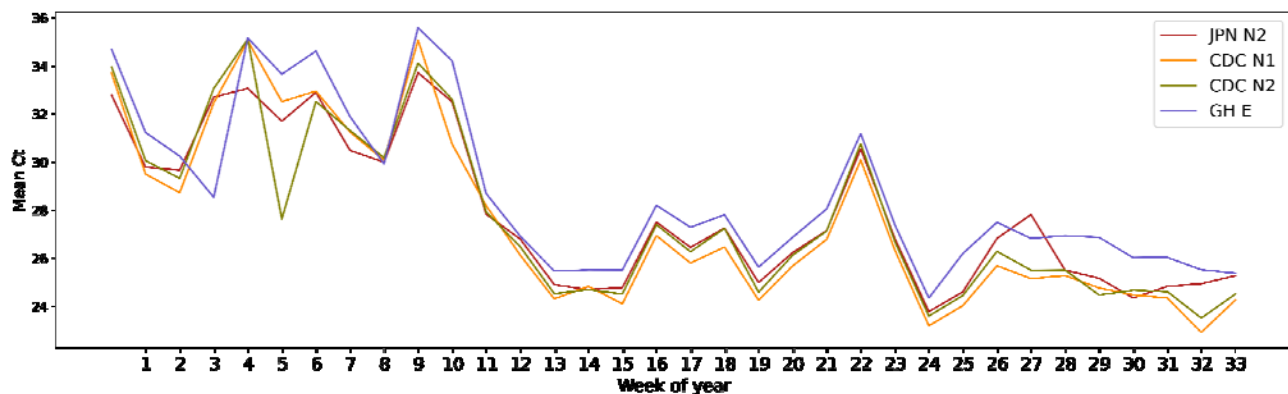


Figure 4

A



B

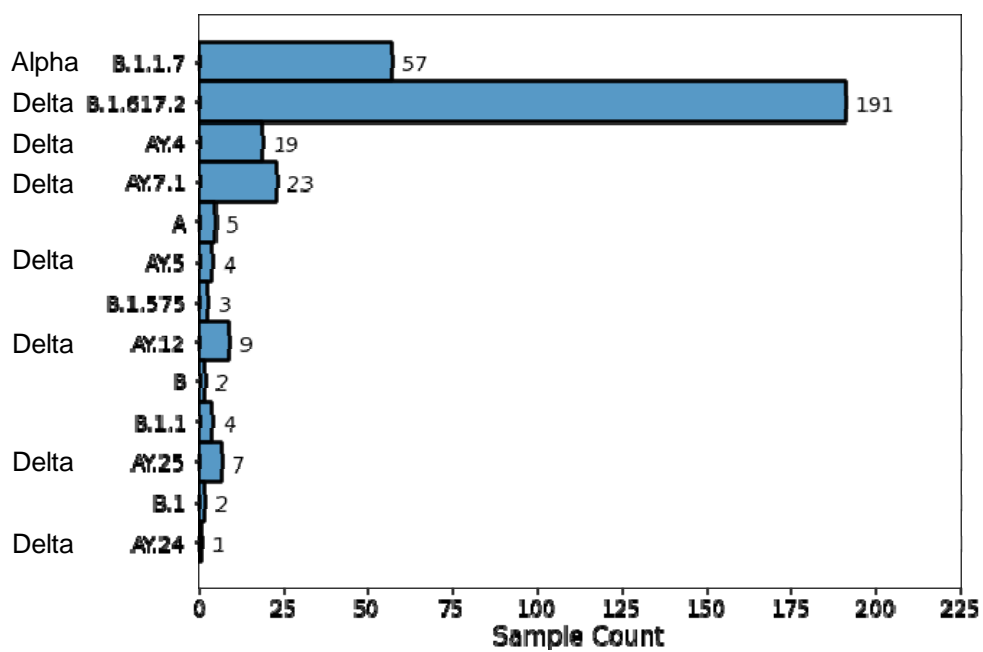


Table 1. Primer/Probe sets used to detect SARS-CoV-2

Primer/Probe	Name	Sequence (5'→3')	Label
NIID_2019-nCoV_N_F2	JPN-F2	AAATTTTGGGGACCAGGAAC	-
NIID_2019-nCoV_N_R2	JPN-R2	TGGCAGCTGTGTAGGTCAAC	-
NIID_2019-nCoV_N_P2	JPN-P2	ATGTCGCGCATTGGCATGGA	5' FAM / 3' BHQ1
nCoV_N1 Forward Primer	CDC-N1-F	GACCCCAAAATCAGCGAAAT	-
nCoV_N1 Reverse Primer	CDC-N1-R	TCTGGTTACTGCCAGTTGAATCTG	-
nCoV_N1 Probe	CDC-N1-P	ACCCCGCATTACGTTTGGTGGACC	5' FAM / 3' BHQ1
nCoV_N2 Forward Primer	CDC-N2-F	TTACAAACATTGGCCGCAA	-
nCoV_N2 Reverse Primer	CDC-N2-R	GCGCGACATTCCGAAGAA	-
nCoV_N2 Probe	CDC-N2-P	ACAATTTGCCCCAGCGCTTCAG	5' FAM / 3' BHQ1
GH_E-F2	GH-E-F2	GTACTCATTCGTTTCGGAAGAGA	-
GH_E-R2	GH-E-R2	AAGGATGGCTAGTGTAAGTAGCAA	-
GH-E-P	GH-E-P	ACCACGAAAGCAAGAAAAAGAAGTACGC	5' FAM / 3' BHQ1

Primer/probe sequences used to screen for SARS-CoV-2 by RT-PCR. Probes were labeled with FAM at the 5' end and BHQ1 at the 3' end. RT-PCR: reverse transcription polymerase chain reaction, FAM: fluorescein amidite, BHQ1: Black Hole Quencher-1

Table 2. Mutations in primer binding sites

Sample	Clade	Changes in primer site	JPN-N2	CDC-N1	CDC-N2	GH-E
20210827_A113	21A (Delta)	USCDC_N1_F:C28298A	29.49	28.61	28.84	29.58
20210827_A222	21A (Delta)	GHE_F:G26272T	20.51			
20210712_A135	20I (Alpha, V1)	USCDC_N2_F:G29179T	27.63	26.76	31.39	22.65

Mutations identified in primer binding sites by COVIDseq. Listed Ct values are those from the initial SARS-CoV-2 RT-PCR screening.