

TITLE PAGE

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

Full-length title:

A 21L/BA.2-21K/BA.1 “MixOmicron” SARS-CoV-2 hybrid undetected by qPCR that screen for variant in routine diagnosis

Short title (for the running head):

Omicron BA.2-BA.1 SARS-CoV-2 recombinant

Author list: Philippe COLSON^{1,2,3}*, Jeremy DELERCE¹, Elise MARION-PARIS⁴, Jean-Christophe LAGIER^{1,2,3}, Anthony LEVASSEUR^{1,2}, Pierre-Edouard FOURNIER^{1,2,5}, Bernard LA SCOLA^{1,2,3}, Didier RAOULT^{1,2}*

Affiliations: ¹ IHU Méditerranée Infection, 19-21 boulevard Jean Moulin, 13005 Marseille, France; ² Aix-Marseille Univ., Institut de Recherche pour le Développement (IRD), Microbes Evolution Phylogeny and Infections (MEPHI), 27 boulevard Jean Moulin, 13005 Marseille, France; ³ Assistance Publique-Hôpitaux de Marseille (AP-HM), 264 rue Saint-Pierre, 13005 Marseille, France; ⁴ Assistance Publique-Hôpitaux de Marseille (AP-HM), Hôpital Timone, Service de médecine du travail, 264 rue Saint-Pierre, 13005 Marseille, France; ⁵ Aix-Marseille Univ., Institut de Recherche pour le Développement (IRD), Vecteurs – Infections Tropicales et Méditerranéennes (VITROME), 27 boulevard Jean Moulin, 13005 Marseille, France.

*** Corresponding authors:** Didier Raoult, IHU Méditerranée Infection, 19-21 boulevard Jean Moulin, 13005 Marseille, France. Tel.: +33 413 732 401, Fax: +33 413 732 402; email: didier.raoult@gmail.com; Philippe Colson, IHU Méditerranée Infection, 19-21 boulevard Jean Moulin, 13005 Marseille, France. Tel.: +33 413 732 401, Fax: +33 413 732 402; email: philippe.colson@univ-amu.fr

Key words: SARS-CoV-2; recombinant; variant; lineage; Omicron; 21L/BA.2; 21K/BA.1; ; epidemic

Word counts: abstract, 250; text, 2,039; **Figures:** 2; **Table:** 1; **References:** 51

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

ABSTRACT

26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48

Among the multiple SARS-CoV-2 variants identified since summer 2020, several have co-circulated, creating opportunities for coinfections and potentially genetic recombinations that are common in coronaviruses. Viral recombinants are indeed beginning to be reported more frequently. Here, we describe a new SARS-CoV-2 recombinant genome that is mostly that of a Omicron 21L/BA.2 variant but with a 3' tip originating from a Omicron 21K/BA.1 variant. Two such genomes were obtained in our institute from adults sampled in February 2022 in university hospitals of Marseille, southern France, by next-generation sequencing carried out with the Illumina or Nanopore technologies. The recombination site was located between nucleotides 26,858-27,382. In the two genomic assemblies, mean sequencing depth at mutation-harboring positions was 271 and 1,362 reads and mean prevalence of the majoritary nucleotide was $99.3\pm 2.2\%$ and $98.8\pm 1.6\%$, respectively. Phylogeny generated trees with slightly different topologies according to whether genomes were depleted or not of the 3' tip. This 3' terminal end brought in the Omicron 21L/BA.2 genome a short transposable element of 41 nucleotides named S2m that is present in most SARS-CoV-2 except a few variants among which the Omicron 21L/BA.2 variant and may be involved in virulence. Importantly, this recombinant is not detected by currently used qPCR that screen for variants in routine diagnosis. The present observation emphasizes the need to survey closely the genetic pathways of SARS-CoV-2 variability by whole genome sequencing, and it could contribute to gain a better understanding of factors that lead to observed differences between epidemic potentials of the different variants.

TEXT

49

50

51 INTRODUCTION

52 Multiple SARS-CoV-2 variants have been identified since summer 2020 (Lemey et al.,
53 2021; Colson et al., 2022a; Hodcroft et al., 2021; Harvey et al., 2021). There were periods during
54 which two distinct variants co-circulated with a crossing of their incidence when an old variant
55 was vanishing and a new one was rising. Substantial rates of co-incidence over periods of several
56 weeks have thus been reported recently worldwide for the Delta [WHO denomination
57 (<https://www.who.int/fr/activities/tracking-SARS-CoV-2-variants>)]/21J [Nextclade classification
58 (Aksamentov et al., 2021) (<https://nextstrain.org/>)] and Omicron 21K/BA.1 [Pangolin
59 classification (Rambaut et al., 2020) (<https://cov-lineages.org/resources/pangolin.html>)] variants,
60 and thereafter, to a lesser extent, for the Omicron BA.1 and BA.2 variants
61 (<https://covariants.org/per-country>) (Hodcroft, 2021; Hadfield et al., 2018), including in our
62 geographical area (Figure 1). This created the opportunity for coinfections (Rockett et al., 2022;
63 Hosch et al., 2022; Bolze et al., 2022; Belen Pisano et al., 2022), and consequently for
64 homologous genetic recombinations, which constitute a major and very common mechanism of
65 evolution in viruses (Bentley and Evans, 2018). Such genetic recombinations are extremely
66 frequent for viruses of family *Coronaviridae*, and they have already been identified for endemic
67 human coronaviruses (Lai, 1996; Zhang et al., 2015; So et al., 2019; Gribble et al., 2020).
68 Regarding SARS-CoV-2, the occurrence of recombinations has been reported or suspected (Yi,
69 2019; Yeh and Contreras, 2020; Haddad et al., 2021; Ignatieva et al., 2021; Jackson et al., 2021;
70 Taghizadeh et al., 2021; Varabyou et al., 2021; Kreier, 2022; Wertheim et al., 2022; He et al.,
71 2022; Sekizuka et al., 2022; Colson et al., 2022b; Lacek et al., 2022; Lohrasbi-Nejad, 2022;
72 Bolze et al., 2022; Ou et al., 2022; Belen Pisano et al., 2022; Burel et al., 2022). Very recently,
73 we described the identification and culture of two SARS-CoV-2 recombinants, one between the

74 B.1.160 and Alpha/20I variants in a patient chronically-infected with SARS-CoV-2 (Burel et al.,
75 2022), and another between the Delta/21J AY.4 and Omicron 21K/BA.1 variants in patients
76 infected approximately 10 weeks after the start of the period of co-detection of these two variants
77 in our geographical area (Colson et al., 2022b). Here, we describe a new hybrid genome, which
78 consists of a Omicron 21L/BA.2 genome whose 3' tip originates from a Omicron 21K/BA.1.

79

80 **MATERIALS AND METHODS**

81 Nasopharyngeal samples were tested for the presence of SARS-CoV-2 RNA by real-time
82 reverse transcription-PCR (qPCR) using the BGI real-time fluorescent RT-PCR assay (BGI
83 Genomics, Shanghai Fosun Long March Medical Science Co., Ltd., Shenzhen, China), as
84 previously described (Burel et al., 2022). Subsequently, qPCR assays that screen for SARS-
85 CoV-2 variants were carried out with the detection of mutations among which spike substitution
86 K417N (Thermo Fisher Scientific, Waltham, USA) and the targeting of viral genes S (spike), N
87 (nucleocapsid) and ORF1 with the TaqPath COVID-19 kit (Thermo Fisher Scientific), as
88 previously reported (Colson et al., 2022a; Colson et al., 2022b).

89 SARS-CoV-2 genomes were obtained and analyzed as described previously (Colson et
90 al., 2022a; Colson et al., 2022b). Briefly, next-generation sequencing was carried out with the
91 Illumina COVID-seq protocol on the NovaSeq 6000 instrument (Illumina Inc., San Diego, CA,
92 USA), or with the Oxford Nanopore technology (ONT) on a GridION instrument (Oxford
93 Nanopore Technologies Ltd., Oxford, UK) following multiplex PCR amplification with the
94 ARTIC nCoV-2019 Amplicon Panel v4.1 of primers (IDT, Coralville, IA, USA) and the ARTIC
95 procedure (<https://artic.network/>).

96 Sequence read processing and genome analysis were performed as described previously
97 (Colson et al., 2022a; Colson et al., 2022b). Briefly, for Illumina NovaSeq reads, base calling
98 was carried out using the Dragen Bcl Convert pipeline [v3.9.3;

99 https://emea.support.illumina.com/sequencing/sequencing_software/bcl-convert.html (Illumina
100 Inc.]), mapping was carried out using the bwa-mem2 tool (v2.2.1; [https://github.com/bwa-
101 mem2/bwa-mem2](https://github.com/bwa-mem2/bwa-mem2)) on the Wuhan-Hu-1 isolate genome (GenBank accession no. NC_045512.2)
102 before cleaning with the SAMtools program (v. 1.13; <https://www.htslib.org/>) (Danecek et al.,
103 2021). Variant calling was performed using FreeBayes (v1.3.5;
104 <https://github.com/freebayes/freebayes>) (Garrison et al., 2012), and consensus genomes were
105 built with the Bcftools program (v1.13; <https://samtools.github.io/bcftools/bcftools.html>). ONT
106 reads were processed with the ARTIC-nCoV-bioinformaticsSOP pipeline (v1.1.0;
107 <https://github.com/artic-network/fieldbioinformatics>). Nucleotide and amino acid changes
108 compared to the Wuhan-Hu-1 isolate genome were determined using the Nextclade tool
109 (<https://clades.nextstrain.org/>) (Hadfield et al., 2018; Aksamentov et al., 2021). Nextstrain clades
110 and Pangolin lineages were identified with the Nextclade web application
111 (<https://clades.nextstrain.org/>) (Hadfield et al., 2018; Aksamentov et al., 2021) and the Pangolin
112 tool (<https://cov-lineages.org/pangolin.html>) (Rambaut et al., 2020), respectively.

113 Phylogenetic analyses were performed the MEGA X software (v10.2.5;
114 <https://www.megasoftware.net/>) (Kumar et al., 2018) following sequence alignment with
115 MAFFT (<https://mafft.cbrc.jp/alignment/server/>) (Kato et al., 2002), and trees were visualized
116 with MEGA X. We built two separate trees, a first one for the whole recombinant genome and a
117 second one for its part originating from an Omicron 21L/BA.2 variant. The 10 genomes the most
118 similar to these sequences among genomes of the Omicron 21L/BA.2 and 21K/BA.1 variants of
119 the sequence database of our institute were selected by a BLAST search (Altschul et al., 1990)
120 then incorporated in the phylogenies together with the sequence of the Wuhan-Hu-1 isolate.

121 Genome sequences obtained and analyzed here were deposited in the NCBI GenBank
122 nucleotide sequence database (<https://www.ncbi.nlm.nih.gov/genbank/>) (Sayers et al., 2022)
123 (Accession no. OM993515 and OM993473), on the IHU Méditerranée Infection website

124 (<https://www.mediterranee-infection.com/sars-cov-2-recombinant/>) (IHUCOVID-063942 and
125 IHUCOVID-068136), and in the GISAID database (<https://www.gisaid.org/>) (Elbe et al., 2017;
126 Alm et al., 2020) (EPI_ISL_10843457, EPI_ISL_10047082).

127

128 **RESULTS**

129 The recombinant genomes were obtained from two adult patients sampled in February
130 2022 in the university hospitals of Marseille, southern France. Cycle threshold values (Ct) of the
131 diagnosis qPCR performed in our institute were 13 and 19. The TaqPath COVID-19 assay
132 showed positivity for all targeted genes including the S gene, and the spike K417N mutation was
133 detected, which led to suspect an infection with the Omicron 21L/BA.2 variant. However, next-
134 generation sequencing performed in our laboratory allowed obtaining SARS-CoV-2 Omicron
135 21L/BA.2-21K/BA.1 genomic hybrid forms. These hybrid genomes have a Omicron 21L/BA.2
136 backbone but their approximately 2,500-3,000 nucleotide-long 3' terminal region is that of a
137 Omicron 21K/BA.1 (Figure 2). The recombination site is located between nucleotides 26,858
138 and 27,382 (in reference to the genome of the Wuhan-Hu-1 isolate). This 3' terminal end in the
139 genomes of the Omicron BA.1 variant does not harbor any signature mutations, and therefore a
140 "gain" of mutation could not be observed here. However, we noted the presence and integrity of
141 a sequence corresponding to a short transposable element of 41 nucleotides named S2m (Tengs
142 et al., 2021), which is present in the Omicron 21K/BA.1 variant but is truncated of 26
143 nucleotides in its central part in the Omicron 21L/BA.2 variant. Current strategies of variant
144 screening by qPCR therefore fail to detect this recombinant as they target mostly mutations in
145 the spike gene, or other genomic regions that do not allow identifying it either. In the two
146 genomic assemblies, the mean (\pm standard deviation) sequencing depth at positions harboring
147 mutations relatively to the genome of the Wuhan-Hu-1 isolate was 271 ± 164 and $1,362\pm 1,146$
148 reads and the mean prevalence of the majoritary nucleotide was $99.3\pm 2.2\%$ and $98.8\pm 1.6\%$,

149 respectively, which rules out the concomitant presence of the sequences of two variants in the
150 samples, either due to co-infection or to contamination.

151 The two recombinant genomes harbor 65 mutations relatively to the genome of the
152 Wuhan-Hu-1 isolate, including 3 or 4 that are not signature mutations of the parental Omicron
153 21L/BA.2 or Omicron 21K/BA.1 genomes (Table 1). The spike genes of the two genomes only
154 harbor signature mutations of the Omicron 21L/BA.2 variant, apart from a synonymous mutation
155 in one genome. Besides, these two genomes differ between each other by a single mutation in the
156 ORF1 gene (C2790U; T841I).

157 The two phylogenetic trees that were built, based either on the whole genomes or only on
158 the region of these genomes that originate from the Omicron 21L/BA.2 variant in the
159 recombinants, did not show exactly the same topology (Figure 2c, d). The trees incorporated the
160 10 genomes of Omicron 21L/BA.2 and 21K/BA.1 variants from the sequence database of our
161 institute that were the most similar to the sequences of the recombinants. In both trees, sequences
162 from the recombinants were clustered with sequences of the Omicron 21L/BA.2 variant.
163 Nonetheless, whole recombinant genomes were clustered separately, apart from their best hits, in
164 the Omicron 21L/BA.2 clade that encompassed two sister groups. In contrast, the two partial
165 recombinant genomes were clustered together but nested inside the Omicron 21L/BA.2 clade.

166

167 **DISCUSSION**

168 In the present work we show a recombination between Omicron BA.2 and BA.1 variant
169 viruses. In two previous studies, 16 (0.006%) of 279,000 genomes and 1,175 (0.2%) of 537,360
170 genomes were identified as being recombinants (Jackson et al., 2021; VanInsberghe et al., 2021).
171 In addition, it was deemed that up to 5% of SARS-CoV-2 that circulated in the USA and UK
172 might have been recombinants (VanInsberghe et al., 2021), and estimated that $\approx 2.7\%$ of
173 sequenced SARS-CoV-2 genomes might have recombinant ancestry (Turkahia et al., 2021). As a

174 matter of fact, cases of detection of recombinant genomes are increasingly reported in 2022
175 (Wertheim et al., 2022; Sekizuka et al., 2022; Colson et al., 2022b; Lacey et al., 2022; Bolze et
176 al., 2022; Ou et al., 2022; Belen Pisano et al., 2022; Burel et al., 2022). Such recombinants are
177 all the more likely to be generated when different variants co-circulate with high incidence
178 levels. In our geographical area, >15,000 infections with the Omicron 21K/BA.1 variant and
179 >1,000 infections with the Omicron 21L/BA.2 variant were diagnosed over the same period of
180 11 weeks between late December and mid-March (Figure 1). Moreover, recombinants becomes
181 more easily identifiable during bioinformatic analyses due to the accumulation over time of
182 mutations along SARS-CoV-2 genomes, at intervals of increasingly reduced size. The two
183 phylogeny reconstructions based on whole genomes or on their region that originate from the
184 Omicron 21L/BA.2 variant in the recombinants exhibited slightly different topologies. This
185 emphasizes that phylogenetic analyses do not accurately handle genomes that are hybrids of
186 sequences with different evolutionary histories, which prompts building separate trees for
187 sequences of different origins in the case of recombinants.

188 The spike gene of the recombinant virus identified here is typical of that of the Omicron
189 21L/BA.2 variant, which suggests similar phenotypic features regarding immune escape (Yu et
190 al., 2022; Iketani et al., 2022). However, the acquisition by a Omicron 21L/BA.2 genome of the
191 3' terminal part of a Omicron 21K/BA.1 genome is of very particular interest as this Omicron
192 21K/BA.1 fragment contains a short transposable element named S2m. S2m is a 41-nucleotide
193 long stem loop motif that is present in four different families of positive-sense single-stranded
194 RNA viruses (Robertson et al., 1998; Imperatore et al., 2022) and also shows high levels of
195 similarity with sequences of insects (Tengs et al., 2021). This element is present in
196 sarbecoviruses and in most of the SARS-CoV-2 genomes. It was proposed to be involved in
197 RNA interference pathways, in hijacking of host protein synthesis, and in RNA recombination
198 events (Imperatore et al., 2022; Gallaher et al., 2022), and it could have initiated viral infection

199 and pathogenicity in various animal hosts. For instance, the s2m element was reported to interact
200 with cellular miRNA-1307-3p in humans, being putatively able to manipulate the host immune
201 response. Moreover, in SARS-CoV-2, S2m is absent or truncated in a few variants including the
202 Eta (B.1.525), Iota (B.1.526) and B.1.640.1 lineages, which all had a low epidemic spread; and it
203 is also truncated in the Omicron 21L/BA.2 variant. Taken together these data suggest that S2m
204 could be considered as the equivalent of a virulence factor. The consequence of the S2m
205 acquisition by an Omicron 21L/BA.2 genome as reported here is unknown. A possibility would
206 be a gain in transmissibility leading to a larger epidemic, which should be investigated by
207 genotypic surveillance among SARS-CoV-2 diagnoses and phenotypic *in vitro* experiments.
208 Thus, the present observation may contribute to gain a better understanding of factors that
209 enhance SARS-CoV-2 spread and lead to the observed differences between the epidemic
210 potential of the variants (Tao et al., 2021; Campbell et al., 2021).

211 Finally, the increasing identification of recombinant SARS-CoV-2 genomes worldwide
212 highlights the unpredictable nature of the genetic variability of this virus. The recombinant
213 described here is not detected by current strategies that screen for variants in routine diagnosis by
214 qPCR. This emphasizes the interest of the most exhaustive whole-genome based surveillance
215 possible to allow deciphering the genetic pathways of the variability and investigating their
216 phenotypic consequences regarding transmissibility, clinical severity, and escape from
217 neutralizing antibodies.

218

219

220 **Acknowledgments**

221 We are very grateful to Raphael Tola, Ludivine Bréchar, and Claudia Andrieu for their
222 technical help.

223 **Author contributions**

224 Study conception and design: Philippe Colson, Pierre-Edouard Fournier, Bernard La
225 Scola, Didier Raoult. Materials, data and analysis tools: Philippe Colson, Jeremy Delerce, Elise
226 Marion-Paris, Jean-Christophe Lagier, Anthony Levasseur. Data analyses: Philippe Colson,
227 Jeremy Delerce, Elise Marion-Paris, Jean-Christophe Lagier, Anthony Levasseur, Pierre-
228 Edouard Fournier, Bernard La Scola, Didier Raoult. Writing of the first draft of the manuscript:
229 Philippe Colson, Pierre-Edouard Fournier, Bernard La Scola, Didier Raoult. All authors read,
230 commented on, and approved the final manuscript.

231 **Funding**

232 This work was supported by the French Government under the “Investments for the
233 Future” program managed by the National Agency for Research (ANR) (Méditerranée-Infection
234 10-IAHU-03), by the Région Provence Alpes Côte d’Azur and European funding FEDER
235 PRIMMI (Fonds Européen de Développement Régional-Plateformes de Recherche et
236 d’Innovation Mutualisées Méditerranée Infection) (FEDER PA 0000320 PRIMMI), and by the
237 French Ministry of Higher Education, Research and Innovation (Ministère de l’Enseignement
238 supérieur, de la Recherche et de l’Innovation) and the French Ministry of Solidarity and Health
239 (Ministère des Solidarités et de la Santé).

240 **Data availability**

241 Genome sequences generated and analyzed in the present study are available from the
242 NCBI GenBank nucleotide sequence database (<https://www.ncbi.nlm.nih.gov/genbank/>) (Sayers
243 et al., 2022) (Accession no. OM993515 and OM993473), from the IHU Méditerranée Infection
244 website (<https://www.mediterranee-infection.com/sars-cov-2-recombinant/>) (IHUCOVID-
245 063942 and IHUCOVID-068136), and from the GISAID database (<https://www.gisaid.org/>)
246 (Elbe et al., 2017; Alm et al., 2020) (EPI_ISL_10843457, EPI_ISL_10047082).

247 **Conflicts of interest**

248 The authors have no conflicts of interest to declare relative to the present study. Didier

249 Raoult was a consultant for the Hitachi High-Technologies Corporation, Tokyo, Japan from
250 2018 to 2020. He is a scientific board member of the Eurofins company and a founder of a
251 microbial culture company (Culture Top). Funding sources had no role in the design and con-
252 duct of the study, the collection, management, analysis, and interpretation of the data, and the
253 preparation, review, or approval of the manuscript.

254 **Ethics**

255 This study has been approved by the ethics committee of the University Hospital Institute
256 Méditerranée Infection (No. 2022-008). Access to the patients' biological and registry data
257 issued from the hospital information system was approved by the data protection committee of
258 Assistance Publique-Hôpitaux de Marseille (APHM) and was recorded in the European General
259 Data Protection Regulation registry under number RGPD/APHM 2019-73.

260

261

262

263

REFERENCES

- 264 Aksamentov, I.; Roemer, C.; Hodcroft, E.B.; Neher, R.A. Nextclade: clade assignment, mutation
265 calling and quality control for viral genomes. *Zenodo* **2021**,
266 <https://doi.org/10.5281/zenodo.5607694>.
- 267 Alm, E.; Broberg, EK.; Connor, T.; et al. Geographical and temporal distribution of SARS-CoV-
268 2 clades in the WHO European Region, January to June 2020. *Euro Surveill.* **2020**, *25*,
269 2001410.
- 270 Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search
271 tool. *J Mol Biol.* **1990**, *215*, 403-10.
- 272 Belén Pisano, M.; Sicilia, P.; Zeballos, M.; et al. SARS-CoV-2 genomic surveillance enables the
273 identification of Delta/Omicron co-infections in Argentina. *medRxiv* **2022**, doi:
274 <https://doi.org/10.1101/2022.03.08.22270920>
- 275 Bentley, K.; Evans, D.J. Mechanisms and consequences of positive-strand RNA virus
276 recombination. *J Gen Virol.* **2018**, *99*, 1345-1356. doi: 10.1099/jgv.0.001142.
- 277 Bolze, A.; White, S.; Basler, T.; et al. Evidence for SARS-CoV-2 Delta and Omicron co-
278 infections and recombination. *medRxiv* **2022**, doi:
279 <https://doi.org/10.1101/2022.03.09.22272113>
- 280 Burel, E.; Colson, P.; Lagier, J.C.; et al. Sequential appearance and isolation of a SARS-CoV-2
281 recombinant between two major SARS-CoV-2 variants in a chronically infected
282 immunocompromised patient. *medRxiv* **2022**, doi:
283 <https://doi.org/10.1101/2022.03.21.22272673>.
- 284 Campbell, F.; Archer, B.; Laurenson-Schafer, H.; et al. Increased transmissibility and global
285 spread of SARS-CoV-2 variants of concern as at June 2021. *Euro Surveill.* **2021**, *26*,
286 2100509.
- 287 Colson, P.; Fournier, PE.; Chaudet, H.; et al. Analysis of SARS-CoV-2 Variants From 24,181
288 Patients Exemplifies the Role of Globalization and Zoonosis in Pandemics. *Front*
289 *Microbiol.* **2022a**, *12*, 786233.
- 290 Colson, P.; Fournier, P.E.; Delerce, J.; et al. Culture and identification of a “Deltamicon”
291 SARS-CoV-2 in a three cases cluster in southern France *medRxiv* **2022b**, doi:
292 <https://doi.org/10.1101/2022.03.03.22271812>
- 293 Danecek, P.; Bonfield, J.K.; Liddle, J.; et al. Twelve years of SAMtools and BCFtools,
294 *GigaScience* **2021**, *10*, giab008.
- 295 Elbe, S.; Buckland-Merrett, G. Data, disease and diplomacy: GISAID’s innovative contribution
296 to global health. *Global Challenges* **2017**, *1*, 33-46.
- 297 Gallaher, W.R. A palindromic RNA sequence as a common breakpoint contributor to copy-
298 choice recombination in SARS-CoV-2. *Arch Virol.* **2020**, *165*, 2341-2348.
- 299 Garrison, E.; Marth, G. Haplotype-based variant detection from short-read sequencing. *arXiv.org*
300 **2012**, <https://arxiv.org/abs/1207.3907>.

- 301 Gribble, J.; Stevens, L.J.; Agostini, M.L.; et al. The coronavirus proofreading exoribonuclease
302 mediates extensive viral recombination. *PLoS Pathog.* **2021**, *17*, e1009226.
- 303 Haddad, D.; John, S.E.; Mohammad, A.; et al. SARS-CoV-2: Possible recombination and
304 emergence of potentially more virulent strains. *PLoS One.* **2021**, *16*, e0251368.
- 305 Hosch, S.; Mpina, M.; Nyakurungu, E.; et al. Genomic Surveillance Enables the Identification of
306 Co-infections With Multiple SARS-CoV-2 Lineages in Equatorial Guinea. *Front Public*
307 *Health.* **2022**, *9*, 818401.
- 308 Hadfield, J.; Megill, C.; Bell, S.M.; et al. Nextstrain: real-time tracking of pathogen evolution.
309 *Bioinformatics.* **2018**, *34*, 4121-4123.
- 310 Harvey, W.T.; Carabelli, A.M.; Jackson, B.; et al. SARS-CoV-2 variants, spike mutations and
311 immune escape. *Nat Rev Microbiol.* **2021**, *19*, 409–424.
- 312 He, Y.; Ma, W.; Dang, S.; et al. Possible recombination between two variants of concern in a
313 COVID-19 patient. *Emerg Microbes Infect.* **2022**, *11*, 552-555. doi:
314 10.1080/22221751.2022.2032375.
- 315 Hodcroft E. 2021. CoVariants: SARS-CoV-2 mutations and variants of interest. Available from:
316 <https://covariants.org/https://covariants.org/>.
- 317 Hodcroft, E.B.; Zuber, M.; Nadeau, S.; et al. Spread of a SARS-CoV-2 variant through Europe in
318 the summer of 2020. *Nature.* **2021**, *595*, 707-712.
- 319 Ignatieva, A.; Hein, J.; Jenkins, P.A. Ongoing Recombination in SARS-CoV-2 Revealed through
320 Genealogical Reconstruction. *Mol Biol Evol.* **2022**, *39*, msac028.
- 321 Iketani, S.; Liu, L.; Guo, Y.; et al. Antibody evasion properties of SARS-CoV-2 Omicron
322 sublineages. *Nature.* **2022** Mar 3. doi: 10.1038/s41586-022-04594-4. Epub ahead of print.
323 PMID: 35240676.
- 324 Imperatore, J.A.; Cunningham, C.L.; Pellegrine, K.A.; et al. Highly conserved s2m element of
325 SARS-CoV-2 dimerizes via a kissing complex and interacts with host miRNA-1307-3p.
326 *Nucleic Acids Res.* **2022**, *50*, 1017-1032.
- 327 Jackson, B.; Boni, M.F.; Bull, M.J.; et al. Generation and transmission of interlineage
328 recombinants in the SARS-CoV-2 pandemic. *Cell.* **2021**, *184*, 5179-5188.e8.
- 329 Ou, J.; Lan, W.; Wu, X.; et al. Tracking SARS-CoV-2 Omicron diverse spike gene mutations
330 identifies multiple inter-variant recombination events. *bioRxiv* **2022**, doi:
331 <https://doi.org/10.1101/2022.03.13.484129>.
- 332 Katoh, K.; Misawa, K.; Kuma, K.; et al. MAFFT: a novel method for rapid multiple sequence
333 alignment based on fast Fourier transform. *Nucleic Acids Res.* **2002**, *30*, 3059-66.
- 334 Kreier, F. Deltacron: the story of the variant that wasn't. *Nature.* **2022**, *602*, 19.
- 335 Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary
336 Genetics Analysis across Computing Platforms. *Mol Biol Evol.* **2018**, *35*, 1547-1549.

- 337 Lacek, K.A.; Rambo-Martin, B.L.; Batra, D.; et al. Identification of a Novel SARS-CoV-2 Delta-
338 Omicron Recombinant Virus in the United States. *bioRxiv* **2022**, doi:
339 <https://doi.org/10.1101/2022.03.19.484981>.
- 340 Lai, M.M.C. Recombination in large RNA viruses: Coronaviruses. *Semin Virol.* **1996**, 7, 381-
341 388.
- 342 Lemey, P.; Ruktanonchai, N.; Hong, S.L.; et al.. Untangling introductions and persistence in
343 COVID-19 resurgence in Europe. *Nature.* **2021**, 595, 713-717.
- 344 Lohrasbi-Nejad, A. Detection of homologous recombination events in SARS-CoV-2. *Biotechnol*
345 *Lett.* **2022**, 17, 1–16.
- 346 Rambaut, A.; Holmes, E.C.; O'Toole, A.; et al. A dynamic nomenclature proposal for SARS-
347 CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol.* **2020**, 5, 1403-1407.
- 348 Robertson, M.P.; Igel, H.; Baertsch, R.; et al. The structure of a rigorously conserved RNA
349 element within the SARS virus genome. *PLoS Biol.* **2005**, 3, 86–94.
- 350 Rockett, J.D.; Gall, M.; Sim, E.M.; et al. Co-infection with SARS-CoV-2 Omicron and Delta
351 variants revealed by genomic surveillance. *medRxiv* **2022**, doi:
352 <https://doi.org/10.1101/2022.02.13.22270755>.
- 353 Sayers, E.W.; Cavanaugh, M.; Clark, K.; et al. GenBank. *Nucleic Acids Res.* **2022**, 50, D161-
354 D164.
- 355 Sekizuka, T.; Itokawa, K.; Saito, M.; et al. Genome Recombination between Delta and Alpha
356 Variants of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). *Jpn J*
357 *Infect Dis.* **2022** Feb 28. doi: 10.7883/yoken.JJID.2021.844. Online ahead of print.
- 358 So, RTY.; Chu, DKW.; Miguel, E.; et al. Diversity of Dromedary Camel Coronavirus HKU23 in
359 African Camels Revealed Multiple Recombination Events among Closely Related
360 Betacoronaviruses of the Subgenus Embecovirus. *J Virol.* **2019**, 93, e01236-19.
- 361 Taghizadeh, P.; Salehi, S.; Heshmati, A.; et al. Study on SARS-CoV-2 strains in Iran reveals
362 potential contribution of co-infection with and recombination between different strains to
363 the emergence of new strains. *Virology.* **2021**, 562, 63-73.
- 364 Tao, K.; Tzou, P.L.; Nouhin, J.; et al. The biological and clinical significance of emerging
365 SARS-CoV-2 variants. *Nat Rev Genet.* **2021**, 22, 757-773.
- 366 Tengs, T.; Delwiche, C.F.; Monceyron, J.C. A genetic element in the SARS-CoV-2 genome is
367 shared with multiple insect species. *J Gen Virol.* **2021**, 102, 001551.
- 368 Turkahia, Y.; Thornlow, B.; Hinrichs, A.; et al. Pandemic-Scale Phylogenomics Reveals
369 Elevated Recombination Rates in the SARS-CoV-2 Spike Region. *bioRxiv* **2021**, doi:
370 <https://doi.org/10.1101/2021.08.04.455157>.
- 371 VanInsberghe, D.; Neish, A.S.; Lowen, A.C.; Koelle, K. Recombinant SARS-CoV-2 genomes
372 are currently circulating at low levels. *bioRxiv.* **2021**, doi: 10.1101/2020.08.05.238386.

- 373 Varabyou, A.; Pockrandt, C.; Salzberg, S.L.; Pertea, M. Rapid detection of inter-clade
374 recombination in SARS-CoV-2 with Bolotie. *Genetics*. **2021**, *218*, iyab074.
- 375 Wertheim, J.O.; Wang, J.C.; Leelawong, M.; et al. Capturing intrahost recombination of SARS-
376 CoV-2 during superinfection with Alpha and Epsilon variants in New York City. *medRxiv*
377 **2022**, doi: <https://doi.org/10.1101/2022.01.18.22269300>.
- 378 Yeh, T.Y.; Contreras, G.P. Emerging viral mutants in Australia suggest RNA recombination
379 event in the SARS-CoV-2 genome. *Med J Aust*. **2020**, *213*, 44-44.e1.
- 380 Yi, H. 2019 Novel Coronavirus Is Undergoing Active Recombination. *Clin Infect Dis*. **2020**, *71*,
381 884-887.
- 382 Yu, J.; Collier, A.Y.; Rowe, M.; et al. Neutralization of the SARS-CoV-2 Omicron BA.1 and
383 BA.2 Variants. *N Engl J Med*. **2022**, Mar 16. doi: 10.1056/NEJMc2201849. Epub ahead of
384 print.
- 385 Zhang, Y.; Li, J.; Xiao, Y.; et al. Genotype shift in human coronavirus OC43 and emergence of a
386 novel genotype by natural recombination. *J Infect*. **2015**, *70*, 641-50.
- 387
- 388
- 389

FIGURE LEGEND

390

391

392 **Figure 1. Incidence and co-incidence of Omicron 21K/BA.1 and Omicron 21L/BA.2**

393 **variants among cases diagnosed in our institute.**

394

395 **Figure 2. Schematic of the SARS-CoV-2 Omicron 21L/BA.2-21K/BA.1 recombinant**

396 **genomes (a, b) and phylogeny reconstruction based on SARS-CoV-2 Omicron 21L/BA.2**

397 **and Omicron 21K/BA.1 genomes (c, d).**

398 a. Map of the SARS-CoV-2 genome.

399 b. Schematic representation of parental and recombinant genomes and mutations in the
400 recombinant genomes. Adapted from screenshots of the nextclade web application output
401 (<https://clades.nextstrain.org>) (Aksamentov et al., 2021; Hadfield et al., 2018). Color codes for
402 nucleotide mutations are as follows: Green: U; yellow: G; blue: C; red: A; light grey: deletions;
403 dark grey: regions uncovered by sequencing reads.

404 c, d. Phylogeny reconstructions based on the whole genomes (c) or only on the region of these
405 genomes that corresponds to the region of the recombinant genome that originates from the
406 Omicron 21L/BA.2 variant.

407

408

409

TABLE

410 **Table 1. Nucleotide and amino acid changes in the MixOmicron recombinant according to**
411 **their presence/absence in the Omicron 21L/BA.2 and Omicron 21K/BA.1 variants**

Gene or genome region	Nucleotide changes	Amino acid changes	Omicron BA.2 genomes	IHUCOVID-068136 (or OM993515, or EPI_ISL_10843457)	IHUCOVID-063942 (or OM993473, or EPI_ISL_10047082)	Omicron BA.1 genomes
5UTR	C241T	/	X	X	X	
ORF1a	T670G	S135R	X	X	X	
ORF1a	C2790T	T842I	X		X	
ORF1a	A2832G	K856R				X
ORF1a	C3037T	/	X	X	X	X
ORF1a	G4184A	G1307S	X	X	X	
ORF1a	C4321T	/	X	X	X	
ORF1a	T5386G	/				X
ORF1a	AGTT6512A	SL2083I				X
ORF1a	G8393A	A2710T				X
ORF1a	C9344T	L3027F	X	X	X	
ORF1a	A9424G	/	X	X	X	
ORF1a	C9534T	T3090I	X	X	X	
ORF1a	C9866T	L3201F	X	X	X	
ORF1a	C10029T	T3255I	X	X	X	X
ORF1a	C10198T	/				
ORF1a	G10447A	/				X
ORF1a	C10449A	P3395H	X	X	X	X
ORF1a	AGTTTGCTCG11282A	LSG3674-				X
ORF1a	GTCTGGTTTT11287G	SGF3675-	X	X	X	
ORF1a	A11537G	I3758V				X
ORF1a	C11650T	/		X	X	
ORF1a	C12880T	/	X	X	X	
ORF1a	T13195C	/				X
ORF1b	C14408T	P314L	X	X	X	X
ORF1b	C15240T	/				X
ORF1b	C15714T	/	X	X	X	
ORF1b	C17410T	R1315C	X	X	X	
ORF1b	A18163G	I1566V	X	X	X	X
ORF1b	C19955T	T2163I	X	X	X	
ORF1b	A20055G	/	X	X	X	
S	C21618T	T19I	X	X	X	
S	TTACCCCTG21632T	LPPA24S	X	X	X	
S	C21762T	A67V				X
S	ATACATG21764A	HV69-				X
S	C21846T	T95I				X
S	GGTGTTTATT21986G	GVYY142D				X
S	G21987A	G142D	X	X	X	
S	AATT22193A	NL211I				X
S	T22200G	V213G	X	X	X	
S	T22204TGAGCCAGAA	INS214EPE				X
S	G22578A	G339D	X	X	X	X
S	T22673C	/				X
S	C22674T	S371F	X	X	X	X
S	T22679C	S373P				X
S	C22686T	S375F	X	X	X	X

412

413
414

Table 1 – continued

Gene or genome region	Nucleotide changes	Amino acid changes	Omicron BA.2 genomes	IHUCOVID-068136 (or OM993515, or EPI_ISL_10843457)	IHUCOVID-063942 (or OM993473, or EPI_ISL_10047082)	Omicron BA.1 genomes
S	A22688G	T376A	X	X	X	
S	G22775A	D405N	X	X	X	
S	A22786C	R408S	X	X	X	
S	C22792T	/		X	X	
S	G22813T	K417N	X	X	X	X
S	T22882G	N440K	X	X	X	X
S	G22898A	G446S				X
S	G22992A	S477N	X	X	X	X
S	C22995A	T478K	X	X	X	X
S	A23013C	E484A	X	X	X	X
S	A23040G	Q493R	X	X	X	X
S	G23048A	G496S				X
S	A23055G	Q498R	X	X	X	X
S	A23063T	N501Y	X	X	X	X
S	T23075C	Y505H	X	X	X	X
S	C23202A	T547K				X
S	A23403G	D614G	X	X	X	X
S	C23525T	H655Y	X	X	X	X
S	T23599G	N679K	X	X	X	X
S	C23604A	P681H	X	X	X	X
S	C23854A	N764K	X	X	X	X
S	G23948T	D796Y	X	X	X	X
S	C24130A	N856K				X
S	A24424T	Q954H	X	X	X	X
S	T24469A	N969K	X	X	X	X
S	C24503T	L981F				X
S	C25000T	/	X	X	X	X
ORF3a	C25584T	/	X	X	X	X
ORF3a	C25624T	H78Y		X	X	
ORF3a	C26060T	T223I	X	X	X	
E	C26270T	T9I	X	X	X	X
M	A26530G	D3G				X
M	C26577G	Q19E	X	X	X	X
M	G26709A	A63T	X	X	X	X
M	C26858T	/	X	X	X	
ORF6	A27259C	/	X	X	X	X
ORF6	GAT27382CTC	D61L	X			
ORF7b	C27807T	/	X	X	X	X
/	A28271T	/	X	X	X	X
N	C28311T	P13L	X	X	X	X
N	GGAGAACGCA28361G	ERS31-	X	X	X	X
N	GGG28881AAC	RG203KR	X	X	X	X
N	A29510C	S413R	X			
3UTR	CGAGGCCACGCGGAGT ACGATCGAGTG29733C	/	X			

415
416
417
418
419
420

/, no change; UTR, untranslated region; S gene region is indicated by a grey background; X, present. Mutations with a bold font are those that are not signature mutations of the Omicron BA.1 or Omicron BA.2 variants.

Fig. 1

medRxiv preprint doi: <https://doi.org/10.1101/2022.03.28.22273010>; this version posted March 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

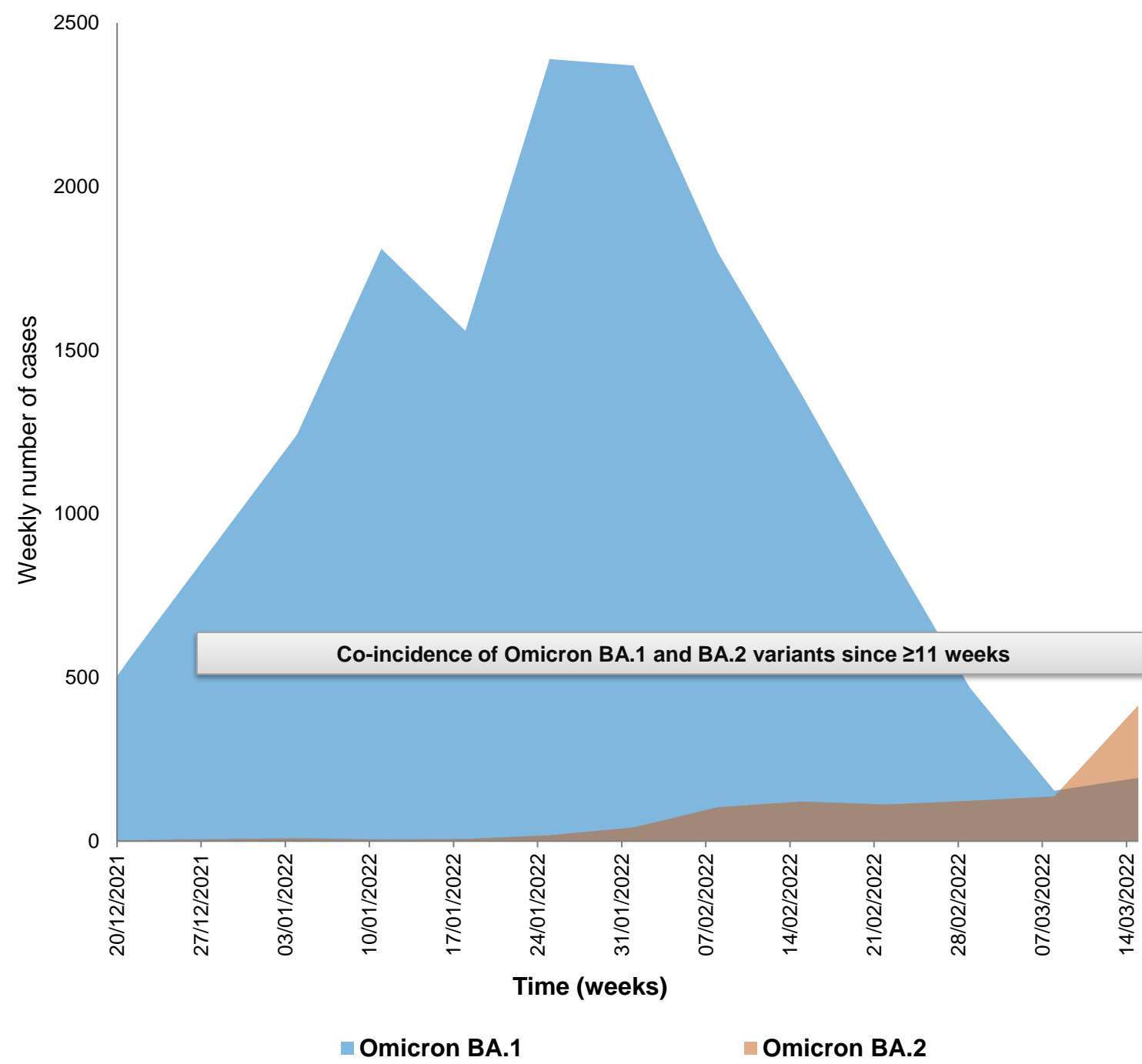
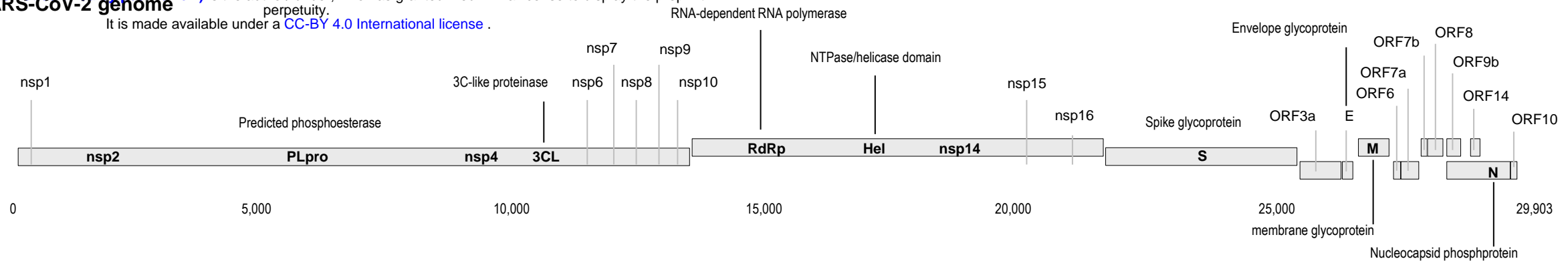
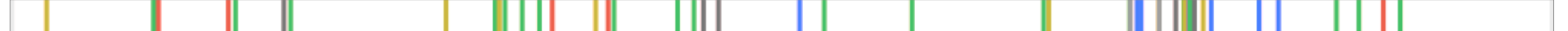
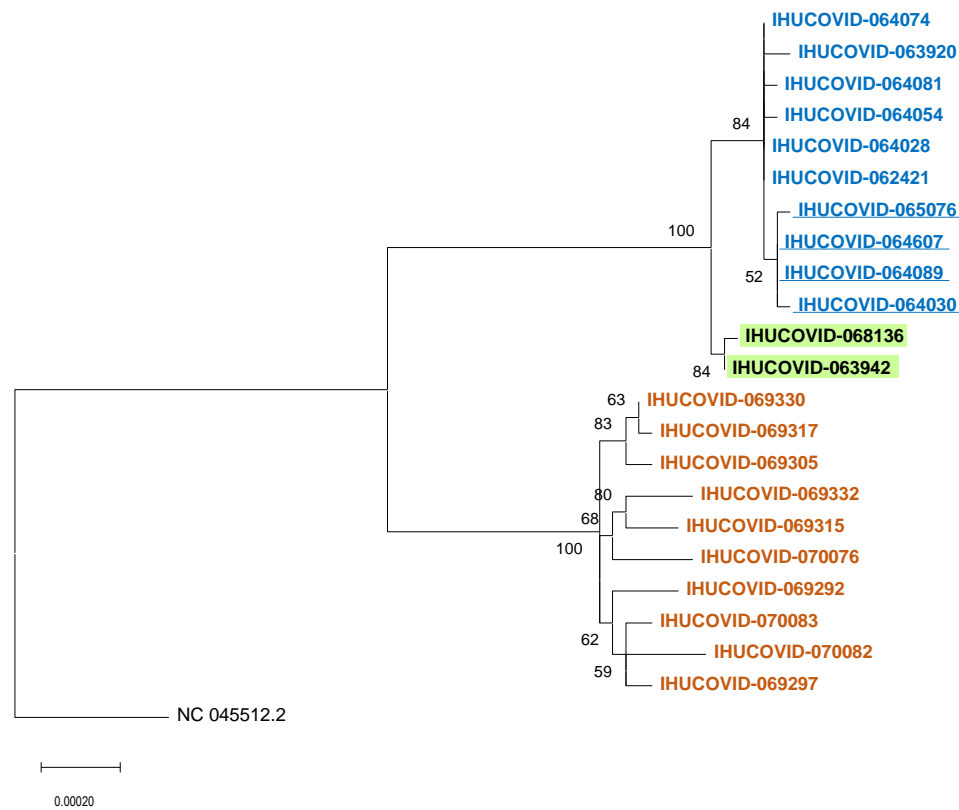


Fig. 2

medRxiv preprint doi: <https://doi.org/10.1101/2022.03.28.22273010>; this version posted March 31, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a [CC-BY 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

a. Map of the SARS-CoV-2 genome**b. Schematic of parental and recombinant genomes, and nucleotide mutations in the Omicron 21L/BA2 - Omicron 21K/BA1 recombinant genome****Parental Omicron 21L/BA.2 genome****Omicron 21L/BA.2 - Omicron 21K/BA1 recombinant genomes****Nucleotide mutations relatively to Wuhan-Hu-1 isolate genome****Nucleotide mutations relatively to Omicron 21L/BA.2 variant genome****Nucleotide mutations relatively to Omicron 21K/BA.1 variant genome****Parental Omicron 21K/BA.1 genome****c. Phylogeny based on the recombinant whole genomes****d. Phylogeny based on the recombinant genomes excluding their Omicron 21K/BA.1 region**