1 **Convergent evolution of the SARS-CoV-2 Omicron subvariants leading to**
2 **the emergence of BQ.1.1 variant**
3

4 Jumpei Ito[1,#], Rigel Suzuki[2,#], Keiya Uriu[1,3,#], Yukari Itakura[4,#], Jiri Zahradnik[5,6,#],
5 Sayaka Deguchi[7,#], Lei Wang[8,9,#], Spyros Lytras[10,#], Tomokazu Tamura[2], Izumi
6 Kida[11], Hesham Nasser[12,13], Maya Shofa[14,15], MST Monira Begum[12], Masumi
7 Tsuda[8,9], Yoshitaka Oda[8], Shigeru Fujita[1,3], Kumiko Yoshimatsu[16], Hayato Ito[2],
8 Naganori Nao[17,21], Hiroyuki Asakura[18], Mami Nagashima[18], Kenji Sadamasu[18],
9 Kazuhisa Yoshimura[18], Yuki Yamamoto[19], Tetsuharu Nagamoto[19], Gideon
10 Schreiber[5], The Genotype to Phenotype Japan (G2P-Japan) Consortium,
11 Akatsuki Saito[14,15,20], Keita Matsuno[11,21,22], Kazuo Takayama[7,23], Shinya
12 Tanaka[8,9]*, Takasuke Fukuhara[2,23,24]*, Terumasa Ikeda[12]*, Kei
13 Sato[1,3,25,26,27,28,29]*
14

15 [1]Division of Systems Virology, Department of Microbiology and Immunology,
16 The Institute of Medical Science, The University of Tokyo, Tokyo, Japan
17 [2]Department of Microbiology and Immunology, Faculty of Medicine, Hokkaido
18 University, Sapporo, Japan.
19 [3]Graduate School of Medicine, The University of Tokyo, Tokyo, Japan
20 [4]Division of Molecular Pathobiology, International Institute for Zoonosis Control,
21 Hokkaido University, Sapporo, Japan.
22 [5]Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot,
23 Israel
24 [6]First Medical Faculty at Biocev, Charles University, Vestec-Prague, Czechia
25 [7]Center for iPS Cell Research and Application (CiRA), Kyoto University, Kyoto,
26 Japan
27 [8]Department of Cancer Pathology, Faculty of Medicine, Hokkaido University,
28 Sapporo, Japan
29 [9]Institute for Chemical Reaction Design and Discovery (WPI-ICReDD), Hokkaido
30 University, Sapporo, Japan
31 [10]Medical Research Council-University of Glasgow Centre for Virus Research,
32 Glasgow, UK.
33 [11]Division of Risk Analysis and Management, International Institute for Zoonosis
34 Control, Hokkaido University, Sapporo, Japan
35 [12]Division of Molecular Virology and Genetics, Joint Research Center for Human
36 Retrovirus infection, Kumamoto University, Kumamoto, Japan
37 [13]Department of Clinical Pathology, Faculty of Medicine, Suez Canal University,
38 Ismailia, Egypt
39 [14]Department of Veterinary Science, Faculty of Agriculture, University of
40 Miyazaki, Miyazaki, Japan
41 [15]Graduate School of Medicine and Veterinary Medicine, University of Miyazaki,
42 Miyazaki, Japan
43 [16]Institute for Genetic Medicine, Hokkaido University, Sapporo, Japan

44 [17]Division of International Research Promotion, International Institute for
45 Zoonosis Control, Hokkaido University, Sapporo, Japan
46 [18]Tokyo Metropolitan Institute of Public Health, Tokyo, Japan
47 [19]HiLung, Inc., Kyoto, Japan
48 [20]Center for Animal Disease Control, University of Miyazaki, Miyazaki, Japan
49 [21]One Health Research Center, Hokkaido University, Sapporo, Japan
50 [22]International Collaboration Unit, International Institute for Zoonosis Control,
51 Hokkaido University, Sapporo, Japan
52 [23]AMED-CREST, Japan Agency for Medical Research and Development
53 (AMED), Tokyo, Japan
54 [24]Laboratory of Virus Control, Research Institute for Microbial Diseases, Osaka
55 University, Suita, Japan
56 [25]International Research Center for Infectious Diseases, The Institute of Medical
57 Science, The University of Tokyo, Tokyo, Japan
58 [26]International Vaccine Design Center, The Institute of Medical Science, The
59 University of Tokyo, Tokyo, Japan
60 [27]Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa,
61 Japan
62 [28]Collaboration Unit for Infection, Joint Research Center for Human Retrovirus
63 infection, Kumamoto University, Kumamoto, Japan
64 [29]CREST, Japan Science and Technology Agency, Kawaguchi, Japan
65 [#]These authors contributed equally
66
67 *Corresponding authors:
68 tanaka@med.hokudai.ac.jp (Shinya Tanaka),
69 fukut@pop.med.hokudai.ac.jp (Takasuke Fukuhara),
70 ikedat@kumamoto-u.ac.jp (Terumasa Ikeda),
71 KeiSato@g.ecc.u-tokyo.ac.jp (Kei Sato)
72

73 **Conflict of interest**: Yuki Yamamoto and Tetsuharu Nagamoto are founders
74 and shareholders of HiLung, Inc. Yuki Yamamoto is a co-inventor of patents
75 (PCT/JP2016/057254; "Method for inducing differentiation of alveolar epithelial
76 cells", PCT/JP2016/059786, "Method of producing airway epithelial cells"). The
77 other authors declare that no competing interests exist.
78

79 **Short title**: Characteristics of SARS-CoV-2 BQ.1.1 (36/50 characters)
80 **Keywords**: SARS-CoV-2; COVID-19; Omicron; BQ.1.1; convergent evolution;
81 transmissibility; immune resistance; pathogenicity

82   **Abstract** (150 words)

83   In late 2022, although the SARS-CoV-2 Omicron subvariants have highly
84   diversified, some lineages have convergently acquired amino acid substitutions
85   at five critical residues in the spike protein. Here, we illuminated the evolutionary
86   rules underlying the convergent evolution of Omicron subvariants and the
87   properties of one of the latest lineages of concern, BQ.1.1. Our phylogenetic and
88   epidemic dynamics analyses suggest that Omicron subvariants independently
89   increased their viral fitness by acquiring the convergent substitutions.
90   Particularly, BQ.1.1, which harbors all five convergent substitutions, shows the
91   highest fitness among the viruses investigated. Neutralization assays show that
92   BQ.1.1 is more resistant to breakthrough BA.2/5 infection sera than BA.5. The
93   BQ.1.1 spike exhibits enhanced binding affinity to human ACE2 receptor and
94   greater fusogenicity than the BA.5 spike. However, the pathogenicity of BQ.1.1
95   in hamsters is comparable to or even lower than that of BA.5. Our multiscale
96   investigations provide insights into the evolutionary trajectory of Omicron
97   subvariants.

98 **Introduction**

99 As of November 2022, the SARS-CoV-2 Omicron variant (B.1.1.529 and BA

100 lineages) is the only current variant of concern (VOC)[1]. At the end of November

101 2021, Omicron BA.1 rapidly outcompeted Delta, a prior VOC. Soon after the

102 global spread of Omicron BA.1 lineage, Omicron BA.2 became predominant in

103 the world. Thereafter, a variety of BA.2 descendants, such as BA.5 and BA.2.75,

104 emerged and are becoming predominant in certain countries.

105 Omicron BA.2 is highly diversified after emergence and is the origin of

106 other recently emerging Omicron subvariants. Although both BA.5 and BA.2.75

107 diversified from BA.2, these two Omicron subvariants are phylogenetically

108 independent to each other, suggesting that these two variants emerged

109 independently[2]. However, recent studies including ours have demonstrated that

110 the spike (S) proteins of these two variants exhibit similar evolutionary patterns:

111 one is amino acid substitutions to evade antiviral humoral immunity, while the

112 other is the substitution to increase the binding affinity to human angiotensin

113 converting enzyme 2 (ACE2), the receptor for SARS-CoV-2 infection. For BA.5,

114 the F486V substitution contributes to evasion from antiviral humoral immunity[3-6],

115 while the L452R substitution increases ACE2 binding affinity[4,7-9]. For BA.2.75,

116 the G446S substitution is responsible for evasion from antiviral humoral

117 immunity[6,10-14], while the N460K substitution increases ACE2 binding affinity[2,7,11].

118 Because the substitutions in the S protein, that confer resistance to antiviral

119 immunity, such as F486V and G446S, reduce the affinity to human ACE2[2,15], it is

120 conceivable to assume that the additional substitutions, such as L452R and

121 N460K, compensate the decreased ACE2 binding affinity[2,15].

122 Another common feature of BA.5 and BA.2.75 we found is the greater

123 pathogenicity compared to BA.2 in a hamster model[2,15]. Because BA.5 and

124 BA.2.75 are descendants of BA.2, our observations suggest that these two

125 variants evolved to increase intrinsic pathogenicity. Importantly, our previous

126 studies that focused on Delta[16], Omicron BA.1[17], Omicron BA.2[18], Omicron

127 BA.5[15], and Omicron BA.2.75[2] suggested that viral intrinsic pathogenicity in

128 hamsters is closely associated with viral fusogenicity in cell culture system.

129 Because previous studies suggest that higher fusogenicity is partly attributed to

130 the increased affinity to human ACE2[11,15,18,19], the substitutions in the S proteins

131 of BA.5 (L452R) and BA.2.75 (N460K) can result in increasing intrinsic

132 pathogenicity in hamsters.

133 Until the emergence of BA.5, the newly emerging SARS-CoV-2 variant

134 globally outcompeted the previously predominant variant in a few months.

135 However, as of November 2022, although a variety of Omicron BA.2 subvariants,

136 including BA.2.75, have emerged after BA.5, none of them have successfully

137 outcompeted BA.5 yet. Instead of the emergence of an outstanding

138 SARS-CoV-2 variant, recently emerging Omicron subvariants are under

139   convergent evolution: most variants have acquired substitutions at the same site

140   of S, such as R346, K444, L452, N460, or F486. Notably, BQ.1.1 is a

141   descendant of BA.5 and bears all five recent convergent mutations: R346T,

142   K444T, L452R, N460K, or F486V. As of October 12, 2022, the WHO classifies

143   BQ.1.1 as an Omicron subvariant under monitoring[1]. Particularly, the BQ.1.1 S

144   harbors two substitutions, L452R and N460K, that increase ACE2 binding affinity

145   and fusogenicity[2,15]. These observations raise the possibility that BQ.1.1 is more

146   fusogenic and pathogenic than BA.5. In this study, we illuminated the

147   evolutionary principals underlying the current convergent evolution of Omicron

148   lineages and characterized BQ.1.1 in terms of its transmissibility,

149   immunogenicity, fusogenicity and intrinsic pathogenicity. Our results suggest

150   that BQ.1.1 is a newly emerging variant that outcompete BA.5 and will be the

151   globally predominant variant in the near future.

Results

Convergent evolution of Omicron lineages

As of November 2022, various Omicron lineages have continuously emerged, such as Omicron BA.1, BA.2, BA.4, BA.5, and BA.2.75 (**Fig. 1a**). As shown in **Fig. 1a**, BQ.1.1, a latest lineage of concern, emerged from the BA.5 cluster. Notably, the substitutions in S protein, particularly R346X, K444X, L452X, N460X, and F486X, seem to have convergently occurred in a variety of Omicron lineages (hereafter we refer to these five amino acid residues as "convergent sites" and substitutions at the residues as "convergent substitutions")[6]. As described in the Introduction, the BQ.1.1 S harbors these five convergent substitutions, R346T, K444T, L452R, N460K, and F486V (**Fig. 1b, left**). Additionally, BQ.1.1 possesses six substitutions in the non-S region when compared to the parental BA.5 (**Fig. 1b, right**). To investigate the substitutions at these five convergent sites during Omicron evolution in depth, we constructed phylogenetic trees for BA.1, BA.2 (including BA.2.75), BA.4, and BA.5 (including BQ.1.1) and identified the branches on the trees where the convergent substitutions occurred (**Fig. 1c**). The R346 residue showed relatively higher substitution frequency compared to the other residues in all Omicron lineages (**Fig. 1c–e**). Consistent with our previous study[15], the L452 residue in BA.2 showed the highest substitution frequency in that lineage (**Fig. 1c–e**). Importantly, the substitution events were more frequently detected in relatively younger lineages such as BA.4, BA.5, and BA.2.75 when compared to relatively older BA.1 and BA.2 lineages (**Fig. 1c–e, Extended Data Fig. 1a**). For instance, the R346X and K444X substitutions in BA.4 and BA.5 and the R346X and F486X substitutions in BA.2.75 showed substantially higher substitution frequencies compared to those in the other lineages (**Fig. 1c**). The substitution frequencies at R346 and K444 in BA.5 were approximately 10.4- and 9.4-times higher than those in BA.2, respectively (**Fig. 1e and Supplementary Table 1**).

Amino acid substitutions at the convergent sites increase viral fitness

We next hypothesized that the substitutions at these five convergent sites conferred certain selective advantages during the evolution of Omicron. To test this hypothesis, we modeled the relationship between viral epidemic dynamics and the substitutions in S protein and estimated the effects of all S substitutions on viral fitness [i.e., relative effective reproduction number (R_e)]. We classified the viral sequences of Omicron according to the combination of amino acid substitutions (referred to as "S haplotype"). Subsequently, we established a Bayesian hierarchal model, which represents the epidemic dynamics of the S haplotype-based viral groups according to relative R_e represented by a linear combination of the effect of S substitutions (see **Methods**). This model can simultaneously estimate i) the effect of each S substitution on R_e and ii) the

193   relative R_e of a viral group represented by each S haplotype. We analyzed the
194   dataset for 375,121 Omicron sequences collected in the United Kingdom (UK)
195   from March 1st, 2022, to October 15th, 2022, which includes 254 S haplotypes
196   according to the pattern of 107 substitutions in S protein [or substitution clusters
197   (**Extended Data Fig. 1b**)]. We first investigated the effects of S substitution on
198   R_e (**Fig. 1f, Extended Data Fig. 1c and Supplementary Table 2**). As
199   substitutions with negative effects on R_e, the cluster of substitutions specific to
200   BA.1 – the earliest Omicron lineage with the lowest R_e overall – and Q493R –
201   acquired once in the common ancestor of Omicron but subsequently lost in BA.5
202   and BA.2.75 – were identified (**Fig. 1a,f, Extended Data Fig. 1c and**
203   **Supplementary Table 2**). On the other hand, as substitutions with positive
204   effects on R_e, substitutions at the five convergent sites were identified (**Fig. 1f,**
205   **Extended Data Fig. 1c and Supplementary Table 2**). Particularly, i) the L452R
206   and F486V substitutions, acquired by the common ancestor of BA.4 and BA.5, ii)
207   L452Q, acquired by BA.2.12.1[15], and iii) N460K, acquired by BA.2.75 and
208   BQ.1.1 independently, showed higher positive effects (**Fig. 1f, Extended Data**
209   **Fig. 1c and Supplementary Table 2**). Next, we investigated relative R_e values
210   for viral groups represented by respective S haplotypes (**Fig. 1g, Extended**
211   **Data Fig. 1d and Supplementary Table 3**). We found that S haplotypes with
212   substitutions at the convergent sites, particularly with R346T, K444T, L452R,
213   N460K, and F486V, tended to show higher R_e values (**Fig. 1g**). Notably, the S
214   haplotype corresponding to BQ.1.1, harboring all five aforementioned
215   convergent substitutions, showed the highest R_e value, followed by S haplotypes
216   corresponding to BQ.1, harboring all substitutions apart from R346T (**Fig. 1g**
217   **and Supplementary Table 3**).
218   To quantify the impact of substitutions at the convergent sites for viral
219   fitness, we inferred the proportion of the variation of R_e that can be explained by
220   these substitutions in the Omicron lineages. We first calculated the total effect of
221   substitutions at the convergent sites for each S haplotype. Subsequently, we
222   compared this quantity with the relative R_e value for each S haplotype (**Fig. 1h**).
223   These two quantities were strongly correlated (R^2=0.816), suggesting that a
224   larger proportion (81.6%) of the variation of R_e in the Omicron lineages can be
225   explained by substitutions at the convergent sites under our model.
226   Using the Bayesian hierarchal model above, we can predict relative viral
227   fitness for arbitrary S sequences only based on the profile of S substitutions (see
228   **Methods**). Utilizing this property of this model, we next inferred the evolutionary
229   change in viral fitness during BA.5 diversification. We reconstructed the
230   ancestral profile of S substitutions for each node in the BA.5 tree. Subsequently,
231   we predicted relative viral fitness for each node according to the reconstructed S
232   substitution profile using the model above (**Fig. 1i**). This analysis suggested that
233   the viral fitness was independently elevated in multiple lineages during BA.5

234 diversification, coupled with the substitution events at the convergent sites (**Fig.**
235 **1i, left**). Finally, we inferred the evolutionary changes in viral fitness specific to
236 the emergence of BQ.1.1 and revealed that the ancestral lineage of BQ.1.1 has
237 acquired the K444T, N460K, and R346T substitutions in this order (**Fig. 1i,**
238 **right**)[20]. Importantly, our analysis predicted that the ancestral lineage of BQ.1.1
239 has increased its viral fitness step-by-step consistently with the acquisitions of
240 these substitutions (**Fig. 1i, right**). Taken together, our results suggest that the
241 sublineages descending from BA.5, including BQ.1.1, convergently increased
242 viral fitness by consecutively acquiring substitutions at the R346, N460, and
243 K444 residues.
244
245 **Immune resistance of BQ.1.1**
246 It has been recently reported that BQ.1.1 exhibits profound resistance to all
247 therapeutic monoclonal antibodies currently approved by the Food and Drug
248 Administration (FDA) in the United States[6,21]. Also, some substitutions detected
249 in BQ.1.1, such as R346T and K444T, contribute to the resistance to 3-dose
250 treatment of an inactivated vaccine (CoronaVac) and breakthrough infections or
251 prior Omicron subvariants (including BA.1, BA.2 and BA.5) after CoronaVac
252 vaccination[6]. However, the immune resistance of BQ.1.1 to breakthrough
253 infections or prior Omicron subvariants after mRNA vaccine treatment remains
254 unaddressed. To experimentally investigate the virological features of BQ.1.1,
255 we first evaluated the immune resistance of BQ.1.1 using HIV-1-based
256 pseudoviruses. Consistent with our recent study[15], BA.5 (2.5-fold) and BQ.1.1
257 (6.9-fold) were significantly more resistant to breakthrough BA.2 infection sera
258 than BA.2 (**Fig. 2a**). Additionally, BQ.1.1 exhibited more profound resistance to
259 breakthrough BA.2 infection sera than BA.5 (2.7-fold, p=0.0076) (**Fig. 2a**). As
260 shown in **Fig. 1a**, the BQ.1.1 S harbors three additional substitutions in the BA.5
261 S: R346T, K444T and N460K. To determine the responsible substitution(s) for
262 immune resistance of BQ.1.1 to breakthrough BA.2 infection sera, we prepared
263 the BA.5 derivatives bearing either of these three substitutions. However, any
264 BA.5-based derivatives prepared did not exhibit resistance to breakthrough BA.2
265 infection sera compared to BA.5 (**Fig. 2a**), suggesting that multiple substitutions
266 cooperatively contribute to the immune resistance of BQ.1.1 to breakthrough
267 BA.2 infection sera. In the case of breakthrough BA.5 infection sera, BQ.1.1 was
268 significantly (5.6-fold) more resistant to breakthrough BA.5 infection sera than
269 BA.5 (p<0.0001) (**Fig. 2b**). Importantly, the breakthrough BA.5 infection sera
270 obtained from six individuals (five breakthrough infection cases after 3-dose
271 vaccination and a breakthrough infection case after 2-dose vaccination) did not
272 exhibit antiviral effect against BQ.1.1 in this experimental setup. We then
273 assessed the determinant substitutions to be resistant to breakthrough BA.5
274 infection sera. The N460K substitution conferred significant resistance to

8

275　breakthrough BA.5 infection sera (1.6-fold, p=0.016), while the other two
276　substitutions did not affect the immune resistance to breakthrough BA.5 infection
277　sera (**Fig. 2b**). When compared to the immune resistance of BQ.1.1 to
278　breakthrough BA.5 infection sera (5.6-fold), the resistance acquired by the
279　N460K substitution (1.6-fold) is relatively less robust (**Fig. 2b**). Therefore, similar
280　to breakthrough BA.2 infection sera, our results suggest that the immune
281　resistance of BQ.1.1 to breakthrough BA.2 infection sera is attributed by multiple
282　substitutions in the RBD of BQ.1.1 S.

283　　　　To further address the difference in antigenicity among Omicron
284　subvariants including BQ.1.1, we used the sera obtained from infected hamsters
285　at 16 days postinfection (d.p.i.). Consistent with our previous studies[2,15], the
286　hamster sera infected with BA.2, BA.5 or BA.2.75 were most efficiently antiviral
287　against the variant of virus infected, while these antisera were less or no
288　cross-reactive against the other variants (**Fig. 2c**). In the case of BA.5 infection
289　sera, BQ.1.1 was 1.8-fold more resistant to than BA.5 (**Fig. 2c**). To depict the
290　difference of antigenicity among BA.2, BA.5, BA.2.75 and BQ.1.1, we analyzed
291　the neutralization dataset using hamster sera (**Fig. 2c**). As shown in **Fig. 2d**, the
292　cross-reactivity of each Omicron subvariant was well correlated to their
293　phylogenetic relationship (**Fig. 1a**), and the antigenicity of BQ.1.1 is relatively
294　more similar to BA.5 than BA.2 and BA.2.75. Nevertheless, our data show that
295　BQ.1.1 exhibits a profound resistance to the humoral immunity induced by BA.5
296　breakthrough infection (**Fig. 2b**). These observations suggest that the three
297　substitutions in BQ.1.1 S are critical and specific to evade the BA.5
298　infection-induced herd immunity in the human population.

299

300　**ACE2 binding affinity of BQ.1.1 S**
301　We then evaluated the features of BQ.1.1 S that potentially affect viral infection
302　and replication. Yeast surface display assay[2,15,18,19,22-24] showed that the K_D
303　value of BQ.1.1 S RBD (0.66±0.11) to human ACE2 molecule is significantly
304　lower than that of parental BA.5 S RBD (1.08±0.16) (**Fig. 3a**), suggesting that
305　BQ.1.1 increased the binding affinity to human ACE2 during evolution from BA.5.
306　To determine the responsible substitutions in the BQ.1.1 S that enhance ACE2
307　binding affinity, we prepared the RBDs of BA.2 and BA.5 S proteins that possess
308　a BQ.1.1-specific substitution compared to parental BA.5 (i.e., R346T, K444T
309　and N460K). Consistent with our recent study[2], the N460K substitution
310　significantly increased the binding affinity of the S proteins of BA.2 and BA.5 to
311　human ACE2 (**Fig. 3a**). On the other hand, the K444 substitution significantly
312　decreased ACE2 binding affinity regardless of the S backbone (**Fig. 3a**). The
313　R346T substitution increased the ACE2 binding affinity of BA.2 S RBD but not
314　that of BA.5 S RBD (**Fig. 3a**). The *in vitro* observations using yeasts (**Fig. 3a**)
315　were then verified by using an HIV-1-based pseudovirus system. As shown in

9

316 **Fig. 3b**, the infectivity of BQ.1.1 pseudovirus was significantly higher than those
317 of BA.2 (17-fold) and BA.5 (3.2-fold) pseudoviruses. In our recent study[15], at
318 least three mutations detected in the BA.5 S (compared to the BA.2 S),
319 HV69-70del, L452R and F486V contribute to the increase of pseudovirus
320 infectivity. When we particularly focus on the three additional mutations detected
321 in the BQ.1.1 S compared to the BA.5 S, R346T, K444T and N460K, we found
322 that R346T and N460K but not K444T significantly increase pseudovirus
323 infectivity, and it is independent of the S backbone (**Fig. 3b**). To assess the
324 association of TMPRSS2 usage with the increased pseudovirus infectivity of
325 BQ.1.1, we used HEK293-ACE2/TMPRSS2 cells and HEK293-ACE2 cells, on
326 which endogenous surface TMPRSS2 is undetectable[18] as target cells. As
327 shown in **Fig. 3c**, the infectivity of BQ.1.1 pseudovirus was not increased by
328 TMPRSS2 expression, suggesting that TMPRSS2 is not associated with an
329 increase in the infectivity of BQ.1.1 pseudovirus.
330
331 **Fusogenicity of BQ.1.1 S**
332 The fusogenicity of BQ.1.1 S was measured by the SARS-CoV-2 S-based fusion
333 assay[2,15-19,25] using Calu-3 cells. Surface expression level of the BQ.1.1 S was
334 significantly lower than that of BA.2, but the BQ.1.1 and BA.5 expression level
335 were comparable (**Fig. 3d**). In the BA.2 S derivatives, R346T and N460K
336 significantly decreased surface expression (**Fig. 3d**). In the BA.5 S derivatives,
337 N460K significantly decreased surface expression, while K444T increased it
338 (**Fig. 3d**). The fusogenicity of BA.5 S is greater than that of BA.2 S (**Fig. 3e**),
339 which is consistent with our recent studies[2,15]. More importantly, the BQ.1.1 S
340 was significantly more fusogenic than the BA.5 S (**Fig. 3f**). Additional
341 experiments using the S derivatives based on BA.2 and BA.5 showed that the
342 R346T and N460K substitutions significantly increased the S-mediated
343 fusogenicity independently of the S backbone (**Fig. 3e,f**). Together with our
344 recent studies[2], the N460K substitution, which is detected in BA.2.75, increased
345 ACE2 binding affinity (i.e., decrease of the K_D value) (**Fig. 3a**), increased
346 pseudovirus infectivity (**Fig. 3b**) and the S-mediated fusogenicity (**Fig. 3e,f**).
347 Interestingly, the R346T substitution also significantly increased ACE2 binding
348 affinity and the S-based fusogenicity, while the K444T substitution negatively
349 affected these experimental parameters (**Fig. 3b–f**). These results suggest that,
350 compared to BA.5, the virological features of BQ.1.1 S, including increased
351 ACE2 binding affinity, pseudovirus infectivity and fusogenicity, are attributed to
352 the R346T and N460K substitutions.
353
354 **Growth kinetics of BQ.1.1 *in vitro***
355 To investigate the growth kinetics of BQ.1.1 in *in vitro* cell culture systems, we
356 inoculated clinical isolates of BA.2[18], BA.5[2] and BQ.1.1 into multiple cell cultures.

357 The growth of BQ.1.1 in Vero cells (**Fig. 3g**) and VeroE6/TMPRSS2 cells (**Fig. 3h**) was significantly greater than that of BA.5, and the growth of BQ.1.1 and BA.5 was comparable in Calu-3 cells (**Fig. 3i**), human airway organoid-derived air-liquid interface (AO-ALI) system (**Fig. 3j**), and human induced pluripotent stem cell (iPSC)-derived alveolar epithelial cells (**Fig. 3l**). However, the growth of BQ.1.1 in iPSC-derived airway epithelial cells was significantly lower than that of BA.5 (**Fig. 3k**).

364 To evaluate the impact of BQ.1.1 infection on the airway epithelial and endothelial barriers, we used an airway-on-a-chip system[2,26]. By measuring the amount of virus that invades from the top channel (**Fig. 3m, left**) to the bottom channel (**Fig. 3m, middle**), we are able to evaluate the ability of viruses to disrupt the airway epithelial and endothelial barriers. Notably, the percentage of virus that invaded the bottom channel of BQ.1.1-infected airway-on-chip was significantly higher than that of BA.5-infected airway-on-chip (**Fig. 3m, right**). Together with the findings of S-based fusion assay (**Fig. 3f**), these results suggest that BQ.1.1 has higher fusogenic than BA.5.

Virological characteristics of BQ.1.1 *in vivo*

375 To investigate the virological features of BQ.1.1 *in vivo*, we inoculated clinical isolates of Delta[16], BA.5[2], and BQ.1.1. Consistent with our previous studies[2,16], Delta infection resulted in weight loss (**Fig. 4a, left**). On the other hand, the body weights of BA.5- and BQ.1.1-infected hamsters did not increase compared with the negative control and relatively similar (**Fig. 4a, left**). We then analyzed the pulmonary function of infected hamsters as reflected by two parameters, enhanced pause (Penh) and the ratio of time to peak expiratory flow relative to the total expiratory time (Rpef). Among the four groups, Delta infection resulted in significant differences in these two respiratory parameters compared to BA.5 (**Fig. 4a, middle and right**), suggesting that Delta is more pathogenic than BA.5. In contrast, the Penh value of BQ.1.1-infected hamsters was significantly lower than that of BA.5-infected hamsters, and the Rpef value of BQ.1.1-infected hamsters was significantly higher than that of BA.5-infected hamsters (**Fig. 4a, middle and right**). These observations suggest that the pathogenicity of BQ.1.1 is similar to or even less than that of BA.5.

390 To evaluate the viral spread in infected hamsters, we routinely measured the viral RNA load in the oral swab. Although the viral RNA loads of the hamsters infected with Delta were significantly higher than those infected with BA.5, there was no statistical difference between BQ.1.1 and BA.5 (**Fig. 4b, left**). To address the possibility that BQ.1.1 more efficiently spreads in the respiratory tissues, we collected the lungs of infected hamsters at 2 and 5 d.p.i., and the collected tissues were separated into the hilum and periphery regions. However, the viral RNA loads in both lung hilum and periphery of BA.5-infected

11

398 hamsters were comparable to those of Delta- and BQ.1.1-infected hamsters (**Fig.**
399 **4b, middle and right**), suggesting that the spreading efficacy of BQ.1.1 in the
400 lungs is comparable to BA.5. To further investigate the viral spread in the
401 respiratory tissues of infected hamsters, we performed immunohistochemical
402 (IHC) analysis targeting viral nucleocapsid (N) protein. Similar to our previous
403 studies[15,17,18], epithelial cells in the upper tracheae of infected hamsters were
404 sporadically positive for viral N protein at 2 d.p.i., but there were no significant
405 differences among the three viruses, including BQ.1.1 (**Extended Data Fig. 2a**),
406 although tracheal inflammation tended to remain in BQ.1.1-infected hamsters. In
407 the alveolar space around the bronchi/bronchioles at 2 d.p.i., N-positive cells
408 were detected in Delta-infected hamsters (**Fig. 4c, left and Extended Data Fig.**
409 **2b**). In contrast, the percentage of N-positive cells in the lungs of BQ.1.1- and
410 BA.5-infected hamsters were relatively low and comparable (**Fig. 4c, left and**
411 **Extended Data Fig. 2b**). At 5 d.p.i., N-positive cells were detected in the
412 peripheral alveolar space in Delta-infected hamsters, while the N-positive areas
413 of BQ.1.1- and BA.5-infected hamsters were sporadic and faintly detectable (**Fig.**
414 **4c, right and Extended Data Fig. 2b**). These data suggest that the spreading
415 efficiency of BQ.1.1 in the lungs of infected hamsters is comparable to that of
416 BA.5.
417
418 **Intrinsic pathogenicity of BQ.1.1**
419 To investigate the intrinsic pathogenicity of BQ.1.1, we analyzed the
420 formalin-fixed right lungs of infected hamsters at 2 and 5 d.p.i. by carefully
421 identifying the four lobules and main bronchus and lobar bronchi sectioning each
422 lobe along with the bronchial branches (**Fig. 4d**). Histopathological scoring was
423 performed according to the criteria described in our previous studies[27].
424 Consistent with our previous studies[2,16,17], all five histological parameters as well
425 as the total score of the Delta-infected hamsters were significantly greater than
426 those of the BA.5-infected hamsters (**Fig. 4e**). When we compared the
427 histopathological scores of Omicron subvariants, total histopathological scores
428 were comparable between BQ.1.1-infected hamsters and BA.5-infected
429 hamsters with some enhancement of bronchitis/bronchiolitis at 2 d.p.i. and
430 presence of type II pneumocytes at 5 d.p.i. of BQ.1.1 (**Fig. 4e**). Altogether, these
431 histopathological analyses suggest that the intrinsic pathogenicity of BQ.1.1 is
432 lower than that of Delta and comparable to that of BA.5.

Discussion

In the present study, we highlighted the amino acid substitutions that convergently and frequently occurred in the residues R346, K444, L452, N460, and F486 of the S proteins of relatively younger Omicron lineages, such as BA.4, BA.5, and BA.2.75 (**Fig. 1a,c**). Our modeling analysis suggested that these five frequent substitutions at the convergent sites increase viral fitness, R_e (**Fig. 1f**), and a larger proportion (81.6%) of the R_e variation within Omicron can be explained by substitutions in the convergent sites (**Fig. 1h**). Intriguingly, the viral groups harboring the substitutions in convergent sites showed higher R_e, and BQ.1.1, which harbors all convergent substitutions, showed the highest R_e among viral groups investigated (**Fig. 1g**). Moreover, the reconstruction of ancestral viral fitness suggests that the ancestral lineage of BQ.1.1 has increased viral fitness by acquiring substitutions at the convergent sites in a stepwise manner (**Fig. 1i**). Altogether, our integrated approach of phylogenetic and epidemic dynamics modeling analyses provides insights into the evolutionary rules underlying the outstanding convergent evolution observed in Omicron subvariants.

Our data provide two possibilities that explain the accumulated amino acid substitutions at the convergent sites in relatively younger Omicron lineages, such as BA.4, BA.5 and BA.2.75. One possibility is the epistasis among amino acid substitutions: the fitness of a substitution differs dependent on the presence of the other substitutions and/or the backbone sequence (similar to how the original Omicron genotype likely emerged[28]). In our previous studies, the L452R substitution in the BA.4/5 S[15] and the N460K substitution in the BA.2.75 S[2] increase their binding ability to human ACE2. More importantly, we showed that these substitutions can compensate for the negative effects of the other substitutions that contribute to evasion from antiviral humoral immunity but decrease ACE2 binding ability: for BA.4/5, L452R compensates for the attenuated ACE2 binding affinity by the F486V substitution[4,7-9], while for BA.2.75, N460K compensates for the attenuated ACE2 binding affinity by the G446S substitution[2,7,11]. Acquiring substitutions that potentially increase ACE2 binding ability, such as L452R and N460K, can be one of the factors that increases substitution frequency at the convergent sites in BA.4, BA.5, and BA.2.75. The other possibility is that the effect of substitutions on viral fitness can be changed over time due to changes in immune selective pressures in the human population by vaccinations and/or natural infections with a variety of SARS-CoV-2 variants. The substitutions at R346, K444, and F486 closely associate with the escape from antiviral humoral immunity and monoclonal antibodies[5,6,15]. In fact, here we demonstrated that the antigenicity of BQ.1.1 is different from that of parental BA.5 and BQ.1.1 is more robustly resistant to the antiviral humoral immunity induced by BA.2 and BA.5 breakthrough infections

13

474   than BA.5. Therefore, it is reasonable to assume that the effect of substitutions
475   at these sites on viral fitness likely differs depending on the immune status in the
476   human population. These two possibilities are not mutually exclusive, and these
477   two factors could contribute to the accelerated substitutions at the convergent
478   sites.
479   Based on our experimental results as well as recent reports, it is
480   convincing that BQ.1.1 is one of the variants exhibiting profound resistance to
481   the antiviral humoral immunity induced by breakthrough infections of BA.2 and
482   BA.5 as well as therapeutic monoclonal antibodies[6,21,29]. Additionally, our
483   experiments *in vitro* showed that the BQ.1.1 S exhibits higher affinity to human
484   ACE2 and greater fusogenicity than BA.5, and these abilities are conferred by
485   two substitutions, R346T and N460K. As expected, because BQ.1.1 bears two
486   substitutions, L452R[15] and N460K[2], both of which augment ACE2 binding affinity,
487   BQ.1.1 has evolved to augment its fusogenicity. Moreover, the R346T
488   substitution, while contributing to enhanced ACE2 affinity and augmented
489   fusogenicity, also concomitantly contributes, perhaps unexpectedly, to immune
490   evasion[6,21].
491   Our previous studies focusing on Delta[16], Omicron BA.1[17], BA.5[15], and
492   BA.2.75[2], showed that the intrinsic pathogenicity of SARS-CoV-2 variants is
493   closely related to the fusogenicity of viral S proteins. Therefore, the observations
494   showing the higher fusogenicity of BQ.1.1 S than the BA.5 S based on the
495   S-based fusion assay (**Fig. 3f**) and the experiments using airway-on-a-chip (**Fig.
496   3m**) suggested the increased intrinsic pathogenicity of BQ.1.1 when compared
497   to BA.5. However, it was unexpected that the intrinsic pathogenicity of BQ.1.1 in
498   a hamster model is comparable or even lower than that of BA.5 (**Fig. 4**). This
499   discrepancy between viral fusogenicity and viral intrinsic pathogenicity is
500   reminiscent of the previous two studies on Omicron BA.2 variant. We first
501   showed that the BA.2 S is more fusogenic than the BA.1 S[18]. We then artificially
502   generated a BA.2 S-bearing recombinant SARS-CoV-2, in which the non-S
503   region of viral genome is derived from ancestral SARS-CoV-2 and demonstrated
504   that the BA.2 S-bearing virus is more pathogenic than the BA.1 S-bearing virus
505   in hamsters[18]. On the other hand, Uraki et al. showed that the intrinsic
506   pathogenicity of clinical BA.2 isolates is comparable to clinical BA.1 isolates[30].
507   Because the difference between our study[18] and others[30] is the viral genome
508   sequence in the non-S region, it is suggested that the BA.2 S bears the potential
509   to exhibit augmented pathogenicity when compared to the BA.1 S, whereas the
510   mutations in non-S region of BA.2 genome potentially attenuate viral
511   pathogenicity. In fact, we found there are at least six substitutions in the non-S
512   region of BQ.1.1 when compared to that of BA.5. Therefore, it would be
513   conceivable to assume that there are factors potentially modulate viral intrinsic
514   pathogenicity other than S protein.

14

515   There are two evolutionary scenarios that possibly explain the
516   discrepancy between viral fusogenicity and intrinsic pathogenicity observed in
517   BQ.1.1. One scenario is that BQ.1.1 acquired mutation(s) in the non-S region of
518   viral genome that can attenuate viral pathogenicity and cancel the pathogenicity
519   elevated by the higher fusogenicity compared with the parental BA.5. This is
520   reminiscent to the observations in the two BA.2 studies as mentioned above[18,30].
521   This scenario also provides a possibility that the increased fusogenicity of viral S
522   protein can reduce viral fitness in the human population because greater
523   fusogenicity can result in elevating pathogenicity. Another scenario is brought
524   from a theoretical study by Sasaki, Lion, and Boots[31]. This study provides a
525   possibility that antigenic escape can augment viral pathogenicity[31]. Since we
526   demonstrated that at least two descendants of BA.2, BA.5[15,] and BA.2.75[2],
527   increased their intrinsic pathogenicity, this theory may fit the evolution of
528   Omicron. More importantly, this theory also predicts that there is a limitation to
529   increase viral pathogenicity[31]. Together with our observations, it might be
530   possible to assume that the pathogenicity of Omicron lineage already reaches a
531   plateau.
532   The emergence of SARS-CoV-2 variants with increased intrinsic
533   pathogenicity may not be so critical for the immunized segment of the population.
534   However, the variants bearing greater intrinsic pathogenicity can be a
535   meaningful risk for people who do not have anti-SARS-CoV-2 immunity, most
536   conspicuously the unvaccinated population, including children. Therefore,
537   continued in-depth viral genomic surveillance and real-time evaluation of the risk
538   of newly emerging SARS-CoV-2 variants should be crucial.

Author Contributions

Keiya Uriu, Sayaka Deguchi, Hesham Nasser, Maya Shofa, MST Monira Begum, Shigeru Fujita, Akatsuki Saito, Kazuo Takayama and Terumasa Ikeda performed cell culture experiments.

Rigel Suzuki, Yukari Itakura, Tomokazu Tamura, Izumi Kida, Kumiko Yoshimatsu, Hayato Ito, Naganori Nao, Keita Matsuno and Takasuke Fukuhara performed animal experiments.

Lei Wang, Masumi Tsuda, Yoshitaka Oda and Shinya Tanaka performed histopathological analysis.

Jiri Zahradnik and Gideon Schreiber performed yeast surface display assay.

Sayaka Deguchi and Kazuo Takayama prepared AO-ALI and airway-on-a-chip systems.

Yuki Yamamoto and Tetsuharu Nagamoto performed generation and provision of human iPSC-derived airway and alveolar epithelial cells.

Hiroyuki Asakura, Mami Nagashima, Kenji Sadamasu and Kazuhisa Yoshimura performed viral genome sequencing analysis.

Jumpei Ito and Spyros Lytras performed phylogenetic analyses.

Jumpei Ito performed statistical, modelling, and bioinformatics analyses.

Jumpei Ito, Akatsuki Saito, Keita Matsuno, Kazuo Takayama, Shinya Tanaka, Takasuke Fukuhara, Terumasa Ikeda and Kei Sato designed the experiments and interpreted the results.

Jumpei Ito and Kei Sato wrote the original manuscript.

All authors reviewed and proofread the manuscript.

The Genotype to Phenotype Japan (G2P-Japan) Consortium contributed to the project administration.

Conflict of interest

Yuki Yamamoto and Tetsuharu Nagamoto are founders and shareholders of HiLung, Inc. Yuki Yamamoto is a co-inventor of patents (PCT/JP2016/057254; "Method for inducing differentiation of alveolar epithelial cells", PCT/JP2016/059786, "Method of producing airway epithelial cells"). The other authors declare that no competing interests exist.

Acknowledgments

625

Consortia

627 Saori Suzuki[2], Marie Kato[8], Zannatul Ferdous[8], Hiromi Mouri[8], Kenji Shishido[8],
628 Naoko Misawa[1], Izumi Kimura[1], Yusuke Kosugi[1], Pan Lin[1], Mai Suganami[1], Mika
629 Chiba[1], Ryo Yoshimura[1], Kyoko Yasuda[1], Keiko Iida[1], Naomi Ohsumi[1], Adam P.
630 Strange[1], Daniel Sauter[1,30], So Nakagawa[31] Jiaqi Wu[31], Yukio Watanabe[7],
631 Ayaka Sakamoto[7], Naoko Yasuhara[7], Takao Hashiguchi[32], Tateki Suzuki[32],
632 Kanako Kimura[32], Jiei Sasaki[32], Yukari Nakajima[32], Hisano Yajima[32], Kotaro
633 Shirakawa[32], Akifumi Takaori-Kondo[32], Kayoko Nagata[32], Yasuhiro Kazuma[32],
634 Ryosuke Nomura[32], Yoshihito Horisawa[32], Yusuke Tashiro[32], Yugo Kawa[32],
635 Takashi Irie[33], Ryoko Kawabata[33], Ryo Shimizu[12], Otowa Takahashi[12], Kimiko
636 Ichihara[12], Chihiro Motozono[34], Mako Toyoda[34], Takamasa Ueno[34], Yuki
637 Shibatani[14], Tomoko Nishiuchi[14]

639 [30]University Hospital Tübingen, Tübingen, Germany
640 [31]Tokai University School of Medicine, Isehara, Japan
641 [32]Kyoto University, Kyoto, Japan
642 [33]Hiroshima University, Hiroshima, Japan
643 [34]Kumamoto University, Kumamoto, Japan

644 **Figure legends**

645 **Fig. 1. Convergent evolution of Omicron lineages.**

646 **a,** A maximum likelihood (ML) tree of the Omicron lineages. The tree was rooted
647 using an outgroup sequence (B.1.1). The S substitutions acquired by BA.4/BA.5,
648 BA.2.75, and BQ.1.1 are indicated in the panel, and the five convergent
649 substitutions are indicated in bold. Note that R493Q is a reversion. Bootstrap
650 values, *, ≥ 0.85; **, ≥ 0.9.

651 **b,** Left, amino acid differences in the S proteins of Omicron lineages. The five
652 convergent substitutions are indicated in bold. Right, amino acid differences in
653 the non-S proteins between BA.5 and BQ.1.1.

654 **c,** Left, time-calibrated ML trees for BA.1, BA.2, BA.4, and BA.5. The trees for
655 BA.2 and BA.5 include BA.2.75 and BQ.1.1 lineages, respectively. Dots indicate
656 estimated substitution events at the convergent sites. Branch color indicates the
657 estimated number of additional substitutions at the convergent sites compared to
658 the most recent common ancestor of each lineage. Right, the substitution profile
659 at the convergent sites.

660 **d and e,** The number of substitution events at the convergent sites detected.
661 Raw counts (**d**) and counts per 1 million (1M) analyzed sequences (**d**) are
662 shown. Note that L452 and F486 in BA.4/5 are indicated in gray because BA.4/5
663 harbors the L452R and F486V substitutions.

664 **f,** Effect size of each S substitution on R_e estimated by a hierarchal Bayesian
665 model. Posterior mean value is shown. A group of highly co-occurred
666 substitutions (e.g., L452R and F486V) was treated as substitution clusters. The
667 red and blue dots indicate the substitutions with significant positive and negative
668 effects, respectively. Representative substitutions are annotated.

669 **g,** Relative R_e value for a viral group represented by each S haplotype,
670 assuming a fixed generation time of 2.1-day. Posterior mean value is shown.
671 The R_e of the major S haplotype in BA.2 is set at 1. The substitution profile at the
672 five convergent sites is shown in the left.

673 **h,** Comparison between relative R_e and the total effect of substitutions at the
674 convergent sites on R_e. Dot indicates a viral group represented by a S haplotype.
675 Dots are colored according to the major classification of PANGO lineage.

676 **i,** Change in viral fitness during BA.5 diversification. The lineages indicated with
677 an asterisk, which includes BQ.1.1, are zoomed in the right panel.

678

679 **Fig. 2. Immune resistance of BQ.1.1**

680 Neutralization assays were performed with pseudoviruses harboring the S
681 proteins of BA.2, BA.2.75, BA.5 and BQ.1.1. The BA.5 S-based derivatives are
682 included in **a and b**. The following sera were used.

683 **a,b,** Convalescent sera from fully vaccinated individuals who had been infected
684 with BA.2 after full vaccination (9 2-dose vaccinated and 5 3-dose vaccinated. 14

19

685   donors in total) (**a**), and BA.5 after full vaccination (2 2-dose vaccinated donors,
686   17 3-dose vaccinated donors and 1 4-dose vaccinated donors. 20 donors in
687   total) (**b**).
688   **c,** Sera from hamsters infected with BA.2 (12 hamsters; left), BA.2.75 (12
689   hamsters; middle), and BA.5 (12 hamsters; right).
690   **d,** Principal component analysis to representing the antigenicity of the S proteins.
691   The analysis is based on the results of neutralization assays using hamster sera
692   (**Fig. 3c**).
693   Assays for each serum sample were performed in triplicate to determine the
694   50% neutralization titer (NT_{50}). Each dot represents one NT_{50} value, and the
695   geometric mean and 95% CI are shown. Statistically significant differences were
696   determined by two-sided Wilcoxon signed-rank tests. The *P* values versus BA.2
697   (**a and c, left**) BA.2.75 (**c, middle**) or BA.5 (**b and c, right**) are indicated in the
698   panels.
699   For the BA.5 derivatives (**a and b**), statistically significant differences ($P < 0.05$)
700   versus BA.5 are indicated with asterisks.
701   The horizontal dashed line indicates the detection limit (120-fold). Information on
702   the convalescent donors is summarized in **Supplementary Table 4**.
703
704   **Fig. 3. Virological characteristics of BQ.1.1 *in vitro***
705   **a,** Binding affinity of the RBD of SARS-CoV-2 S protein to ACE2 by yeast
706   surface display. The K_D value indicating the binding affinity of the RBD of the
707   SARS-CoV-2 S protein to soluble ACE2 when expressed on yeast is shown.
708   **b,** Pseudovirus assay. HOS-ACE2-TMPRSS2 cells were infected with
709   pseudoviruses bearing each S protein. The amount of input virus was
710   normalized based on the amount of HIV-1 p24 capsid protein. The percent
711   infectivity compared to that of the virus pseudotyped with the BA.2 S protein are
712   shown.
713   **c,** Fold increase in pseudovirus infectivity based on TMPRSS2 expression.
714   **d–f,** S-based fusion assay. **d,** S protein expression on the cell surface. The
715   summarized data are shown. **e,f,** S-based fusion assay in Calu-3 cells. The
716   recorded fusion activity (arbitrary units) is shown. The dashed green line (**e**) and
717   the dashed brown line (**f**) indicate the results of BA.2 and BA.5, respectively. The
718   red number in each panel indicates the fold difference between BA.2 (**e**) or BA.5
719   (**f**) and the derivative tested at 24 h post coculture.
720   **g–l,** Growth kinetics of BQ.1.1. Clinical isolates of BA.2, BA.5, BQ.1.1 and Delta
721   (only in **m**) were inoculated into Vero cells (**g**), VeroE6/TMPRSS2 cells (**h**),
722   Calu-3 cells (**i**), AO-ALI (**j**), iPSC-derived airway epithelial cells (**k**), iPSC-derived
723   alveolar epithelial cells (**l**) and an airway-on-a-chip system (**m**). The copy
724   numbers of viral RNA in the culture supernatant (**g–i**), the apical sides of
725   cultures (**j–l**), and the top (**m, left**) and bottom (**m, middle**) channels of an

726   airway-on-a-chip were routinely quantified by RT–qPCR. In **m, right**, the
727   percentage of viral RNA load in the bottom channel per top channel during 3–6
728   d.p.i. (i.e., % invaded virus from the top channel to the bottom channel) is shown.
729   Assays were performed in triplicate (**a,d,m**) or quadruplicate (**b,c,e–l**). The
730   presented data are expressed as the average ± SD (**a–f**) or SEM (**g–m**). In **a–d**,
731   each dot indicates the result of an individual replicate. In **a–c**, the dashed
732   horizontal lines indicate the value of BA.5. In **a–d**, statistically significant
733   differences (*, $P < 0.05$) versus each parental S and those between BA.5 and
734   BQ.1.1 were determined by two-sided Student's t tests. In **e–m**, statistically
735   significant differences versus BA.2 (**e**) and BA.5 (**f–m**) across timepoints were
736   determined by multiple regression. The FWERs calculated using the Holm
737   method are indicated in the figures. NA, not applicable.
738
739   **Fig. 4. Virological characteristics of BQ.1.1 *in vivo***
740   Syrian hamsters were intranasally inoculated with BA.5, BQ.1.1 and Delta. Six
741   hamsters of the same age were intranasally inoculated with saline (uninfected).
742   Six hamsters per group were used to routinely measure the respective
743   parameters (**a**). Four hamsters per group were euthanized at 2 and 5 d.p.i. and
744   used for virological and pathological analysis (**b–e**).
745   **a,** Body weight, Penh, and Rpef values of infected hamsters (n = 6 per infection
746   group).
747   **b,** (Left) Viral RNA loads in the oral swab (n=6 per infection group). (Middle and
748   right) Viral RNA loads in the lung hilum (middle) and lung periphery (right) of
749   infected hamsters (n=4 per infection group).
750   **c,** IHC of the viral N protein in the lungs at 2 d.p.i. (left) and 5 d.p.i. (right) of
751   infected hamsters. Representative figures (N-positive cells are shown in brown)
752   and the percentage of N-positive cells in whole lung lobes (n=4 per infection
753   group) are respectively shown. The raw data are shown in **Extended Data Fig.
754   2b**.
755   **d,e,** H&E staining of the lungs of infected hamsters. Representative figures are
756   shown in **d**. Uninfected lung alveolar space and bronchioles are also shown. **e,**
757   Histopathological scoring of lung lesions (n=4 per infection group).
758   Representative pathological features are reported in our previous studies[27]. In
759   **a–c,e**, data are presented as the average ± SEM. In **c**, each dot indicates the
760   result of an individual hamster.
761   In **a,b,e**, statistically significant differences between BA.5 and other variants
762   across timepoints were determined by multiple regression. In **a**, the 0 d.p.i. data
763   were excluded from the analyses. The FWERs calculated using the Holm
764   method are indicated in the figures.
765   In **c**, the statistically significant differences between BA.5 and other variants
766   were determined by a two-sided Mann–Whitney U test.

21

767 In **c and d**, each panel shows a representative result from an individual infected
768 hamster. Scale bars, 500 μm (**c**); 200 μm (**d**).
769
770 **Supplementary Table 1.** Information on the number of estimated substitution
771 events in each Omicron lineage
772
773 **Supplementary Table 2.** Effect size of each S substitution on R_e estimated by a
774 hierarchal Bayesian model
775
776 **Supplementary Table 3.** Relative R_e value for a viral group represented by
777 each S haplotype
778
779 **Supplementary Table 4.** Human sera used in this study
780
781 **Supplementary Table 5.** Primers used in this study
782
783 **Supplementary Table 6.** Summary of unexpected amino acid mutations
784 detected in the working virus stocks
785
786 **Extended Data Fig. 1. Convergent evolution of the Omicron lineages.**
787 **a,** Detected substitution events at the convergent sites (related to **Fig. 1d,e**).
788 Raw counts (left) and counts per 1 million (M) analyzed sequences (right) are
789 shown. Unlike **Fig. 1d,e**, results for BA.2.75 are shown in addition to BA.1, BA.2,
790 BA.4, and BA.5.
791 **b,** The co-occurrence network of S substitutions in the Omicron lineages. In the
792 S haplotype dataset, a pair of substitutions with Pearson's correlation > 0.9 is
793 considered co-occurring substitutions and indicated as a link in the network. In
794 the modeling analysis, a group of co-occurring substitutions was clustered, and
795 one effect value was estimated for each substitution cluster.
796 **c,** Effect size of each S substitution on R_e (related to **Fig. 1f**). A dot and line
797 indicate the posterior mean and the 95% Bayesian confidential interval (CI),
798 respectively.
799 **d,** Relative R_e value for a viral group represented by each S haplotype,
800 assuming a fixed generation time of 2.1 days (related to **Fig. 1g**). A dot and line
801 indicate the posterior mean and the 95% Bayesian CI, respectively. Unlike **Fig.**
802 **1g**, the profile of all S substitutions analyzed is shown on the left side.
803
804 **Extended Data Fig. 2. Histological observations in infected hamsters**
805 **a,** IHC of the viral N protein in the middle portion of the tracheas of all infected
806 hamsters at 2 d.p.i (4 hamsters per infection group). Each panel shows a
807 representative result from an individual infected hamster.

22

808   **b,** IHC of the SARS-CoV-2 N protein in the lungs of infected hamsters at 2 d.p.i.
809   (left) and 5 d.p.i (right) (4 hamsters per infection group). In each panel, IHC
810   staining (top) and the digitalized N-positive area (bottom, indicated in red) are
811   shown. The red numbers in the bottom panels indicate the percentage of the
812   N-positive area. Summarized data are shown in **Fig.4c**. In **a and b**, N-positive
813   cells are shown in brown. Scale bars, 1 mm (**a**); 5 mm (**b**).

814 **Methods**

815

816 **Ethics statement**

817 All experiments with hamsters were performed in accordance with the Science
818 Council of Japan's Guidelines for the Proper Conduct of Animal Experiments.
819 The protocols were approved by the Institutional Animal Care and Use
820 Committee of National University Corporation Hokkaido University (approval ID:
821 20-0123 and 20-0060). All protocols involving specimens from human subjects
822 recruited at Interpark Kuramochi Clinic was reviewed and approved by the
823 Institutional Review Board of Interpark Kuramochi Clinic (approval ID:
824 G2021-004). All human subjects provided written informed consent. All protocols
825 for the use of human specimens were reviewed and approved by the Institutional
826 Review Boards of The Institute of Medical Science, The University of Tokyo
827 (approval IDs: 2021-1-0416 and 2021-18-0617) and University of Miyazaki
828 (approval ID: O-1021).

829

830 **Human serum collection**

831 Convalescent sera were collected from fully vaccinated individuals who had
832 been infected with BA.2 (9 2-dose vaccinated and 5 3-dose vaccinated; 11–61
833 days after testing. n=14 in total; average age: 47 years, range: 24–84 years,
834 64% male) (**Fig. 2a**), and fully vaccinated individuals who had been infected with
835 BA.5 (2 2-dose vaccinated, 17 3-dose vaccinated and 1 4-dose vaccinated;
836 10–23 days after testing. n=20 in total; average age: 51 years, range: 25–73
837 years, 45% male) (**Fig. 2b**). The SARS-CoV-2 variants were identified as
838 previously described[2,15,18]. Sera were inactivated at 56°C for 30 minutes and
839 stored at –80°C until use. The details of the convalescent sera are summarized
840 in **Supplementary Table 4**.

841

842 **Cell culture**

843 HEK293T cells (a human embryonic kidney cell line; ATCC, CRL-3216),
844 HEK293 cells (a human embryonic kidney cell line; ATCC, CRL-1573) and
845 HOS-ACE2/TMPRSS2 cells (HOS cells stably expressing human ACE2 and
846 TMPRSS2)[32,33] were maintained in DMEM (high glucose) (Sigma-Aldrich, Cat#
847 6429-500ML) containing 10% fetal bovine serum (FBS, Sigma-Aldrich Cat#
848 172012-500ML) and 1% penicillin–streptomycin (PS) (Sigma-Aldrich, Cat#
849 P4333-100ML). HEK293-ACE2 cells (HEK293 cells stably expressing human
850 ACE2)[19] were maintained in DMEM (high glucose) containing 10% FBS, 1 µg/ml
851 puromycin (InvivoGen, Cat# ant-pr-1) and 1% PS. HEK293-ACE2/TMPRSS2
852 cells (HEK293 cells stably expressing human ACE2 and TMPRSS2)[19] were
853 maintained in DMEM (high glucose) containing 10% FBS, 1 µg/ml puromycin,
854 200 µg/ml hygromycin (Nacalai Tesque, Cat# 09287-84) and 1% PS. Vero cells

24

855   [an African green monkey (*Chlorocebus sabaeus*) kidney cell line; JCRB Cell
856   Bank, JCRB0111] were maintained in Eagle's minimum essential medium
857   (EMEM) (Sigma-Aldrich, Cat# M4655-500ML) containing 10% FBS and 1% PS.
858   VeroE6/TMPRSS2 cells (VeroE6 cells stably expressing human TMPRSS2;
859   JCRB Cell Bank, JCRB1819)[34] were maintained in DMEM (low glucose) (Wako,
860   Cat# 041-29775) containing 10% FBS, G418 (1 mg/ml; Nacalai Tesque, Cat#
861   G8168-10ML) and 1% PS. Calu-3 cells (ATCC, HTB-55) were maintained in
862   Eagle's minimum essential medium (EMEM) (Sigma-Aldrich, Cat#
863   M4655-500ML) containing 10% FBS and 1% PS. Calu-3/DSP$_{1-7}$ cells (Calu-3
864   cells stably expressing DSP$_{1-7}$)[35] were maintained in EMEM (Wako, Cat#
865   056-08385) containing 20% FBS and 1% PS. Human airway and lung epithelial
866   cells derived from human induced pluripotent stem cells (iPSCs) were
867   manufactured according to established protocols as described below (see
868   "Preparation of human airway and lung epithelial cells from human iPSCs"
869   section) and provided by HiLung Inc. AO-ALI model was generated according to
870   established protocols as described below (see "AO-ALI model" section).
871

Viral genome sequencing

873   Viral genome sequencing was performed as previously described[15]. Briefly, the
874   virus sequences were verified by viral RNA-sequencing analysis. Viral RNA was
875   extracted using a QIAamp viral RNA mini kit (Qiagen, Cat# 52906). The
876   sequencing library employed for total RNA sequencing was prepared using the
877   NEBNext Ultra RNA Library Prep Kit for Illumina (New England Biolabs, Cat#
878   E7530). Paired-end 76-bp sequencing was performed using a MiSeq system
879   (Illumina) with MiSeq reagent kit v3 (Illumina, Cat# MS-102-3001). Sequencing
880   reads were trimmed using fastp v0.21.0[36] and subsequently mapped to the viral
881   genome sequences of a lineage B isolate (strain Wuhan-Hu-1; GenBank
882   accession number: NC_045512.2)[34] using BWA-MEM v0.7.17[37]. Variant calling,
883   filtering, and annotation were performed using SAMtools v1.9[38] and snpEff
884   v5.0e[39].
885

Phylogenetic reconstruction

887   A total of 5,345,749 SARS-CoV-2 genome sequences labelled as 'Omicron' and
888   their corresponding metadata were retrieved from the GISAID database on
889   October 3, 2022 (https://www.gisaid.org/)[40]. The dataset was then filtered based
890   on the following criteria: i) only 'original passage' sequences, ii) collection date in
891   2022, iii) host labelled as 'Human', iv) sequence length above 28,000 base pairs
892   and v) proportion of ambiguous bases below 2%. This filtering reduced the
893   dataset to a total of 3,840,308 sequences. To ensure that PANGO lineage
894   definitions in our dataset's metadata included the latest circulating lineages, the

25

895   GISAID metadata were downloaded again on October 15th, 2022, and PANGO
896   lineages of our sequences were updated accordingly.

897   To construct an ML tree of Omicron lineages (**Fig. 1a**), we randomly
898   sampled 100 sequences from BA.1, BA.2, BA.4, and BA.5 and 20 sequences
899   from BA.2.75 and BQ.1.1. In addition, an outgroup sequence, EPI_ISL_466615,
900   representing the oldest isolate of B.1.1 obtained in the UK was added to the
901   dataset. The viral genome sequences were mapped to the reference sequence
902   of Wuhan-Hu-1 (GenBank accession number: NC_045512.2) using Minimap2
903   v2.17[41] and subsequently converted to a multiple sequence alignment according
904   to the GISAID phylogenetic analysis pipeline
905   (https://github.com/roblanf/sarscov2phylo). The alignment sites corresponding to
906   the 1–265 and 29674–29903 positions in the reference genome were masked
907   (i.e., converted to NNN). Alignment sites at which >50% of sequences contained
908   a gap or undetermined/ambiguous nucleotide were trimmed using trimAl v1.2[42].
909   Phylogenetic tree construction was performed via a three-step protocol: i) the
910   first tree was constructed; ii) tips with longer external branches (Z score > 4)
911   were removed from the dataset; iii) and the final tree was constructed. Tree
912   reconstruction was performed by RAxML v8.2.12[43] under the GTRCAT
913   substitution model. The node support value was calculated by 100 times
914   bootstrap analysis.

915   A separate phylogenetic tree was reconstructed for each Omicron
916   lineage (BA.1, BA.2, BA.4 and BA.5) including all their descendant sublineages
917   (**Fig. 1c**). To remove redundant sequences and reduce the volume of data for
918   each reconstruction, a representative subsampling approach was used. 3,000
919   sequences from each Omicron lineage that had no substitutions at the
920   convergent sites in S: 346, 444, 452, 460 and 486 for BA.1 and BA.2 or no
921   substitutions in sites 346, 444 and 460 for BA.4 and BA.5 were randomly
922   sampled from each dataset, weighting the sampling by the frequency of each
923   PANGO lineage in the dataset. In this way we included a sample of background
924   sequences with no 'additional' substitutions in the sites of interest with PANGO
925   lineage frequencies representative of the full dataset. It was also ensured that
926   the selected sequences had no ambiguous bases in the S gene (checked
927   between sequence positions 21,000 to 26,000) to avoid ambiguous residues in
928   the sites of interest. Recombinant PANGO lineages were excluded from the
929   analysis.

930   After collecting the subsampled set of background sequences for each
931   lineage, a maximum of 30 randomly selected sequences of each PANGO
932   sublineage with at least one additional substitution at the convergent sites were
933   added to the dataset. This subsampling approach aimed to capture sequences
934   of all sublineages that have acquired additional mutations at the convergent
935   sites, while maintaining a large set of background lineages that reflects

26

936 circulating lineage distribution. One SARS-CoV-2 sequence from the sister
937 lineage of each set with a recent collection date was also added to each dataset
938 to be used as an outgroup of the phylogeny [for the BA.1 tree,
939 EPI_ISL_15170885 (BA.2); for the BA.2 tree, EPI_ISL_15148193 (BA.1); for the
940 BA.4 tree, EPI_ISL_15192101 (BA.5); and for BA.5 the tree, EPI_ISL_15174939
941 (BA.4)].
942   Each lineage sequence dataset was aligned using the
943 'global_profile_alignment.sh' from the SARS-CoV-2 global phylogeny pipeline
944 [www.doi.org/10.5281/zenodo.3958883], utilising MAFFT[44]. Phylogenies were
945 reconstructed using iqtree2 (v2.1.3)[45] under a GTR+I+F+G4 model with 1000
946 ultrafast bootstrap replicates to determine node support. Trees were manually
947 rerooted on the branch leading to the outgroup sequence and time-calibrated
948 with TreeTime[46] (with the '--keep-root' option to preserve the outgroup rooting).
949 Branches leading to tips with dates not matching the root-to-tip regression model
950 were removed from the phylogeny using the ete3 python package[47]. The final
951 trees for BA.1, BA.2, BA.4 and BA.5 contain 3,901, 5,343, 3,328, and 5,197
952 sequences, respectively.
953
954 **Ancestral node reconstruction of site substitutions**
955 To infer the branches where substitution events occurred at the five convergent
956 sites (positioned at 346, 444, 452, 460, 486) in the trees of Omicron lineages, we
957 reconstructed the ancestral state of the substitution profile at the convergent
958 sites in each node using a parsimony method, implemented by the phangorn
959 package (https://github.com/KlausVigo/phangorn). Internal nodes with
960 substitution probabilities above or equal to 0.5 were annotated as having the
961 substitution. Branches where substitutions took place for each site were denoted
962 as branches connecting an ancestral internal node with no substitution to an
963 internal node that has a substitution. Additionally, 70% of tips descending from
964 that internal node were also required to have the substitution and at least 3 tips
965 needed to be descended from the node, to avoid picking up branches with low
966 support or clades that reverted back to the original residue. The analysis was
967 performed on R v4.1.2 (https://www.r-project.org/).
968
969 **Modeling the relationship between viral epidemic dynamics and S**
970 **substitutions**
971 Motivated by the model established by Obermeyer et al[48], we developed a
972 method to model the relationship between viral epidemic dynamics and S
973 substitutions. This model can simultaneously estimate i) the effect of each S
974 substitution on R_e and ii) the relative R_e of a viral group represented by each S
975 haplotype. The key concept of the model used in this study is the same as the
976 one in Obermeyer et al[48]. However, our method is independent of the predefined

27

977 viral classification such as PANGO lineage but based on the viral classification
978 according to the profile of S substitutions. Therefore, our method can link the
979 effect of S substitutions to viral epidemic dynamics in a more direct manner. Also,
980 in our method, a Markov Chain Monte Carlo (MCMC) method is used for
981 parameter estimation instead of variational inference, an approximation method.
982　　　　The data used in this analysis were downloaded from the GISAID
983 database (https://www.gisaid.org/) on November 7, 2022. For quality control, we
984 excluded the data of viral sequences with the following features from the
985 analysis: i) a lack of collection date information; ii) sampling in animals other
986 than humans; iii) >1% undetermined nucleotide characters; or iv) sampling by
987 quarantine. Furthermore, in this analysis, we analyzed viral sequences of the
988 Omicron lineages collected in the UK from March 1, 2022, to October 15th, 2022.
989　　　　We selected S substitutions (including insertions and deletions) to be
990 analyzed and classified Omicron sequences into S haplotypes according to the
991 profile of the selected S substitutions: We analyzed S substitutions observed in
992 ≥200 sequences in the dataset we used. We excluded S substitutions commonly
993 (≥90%) detected in sequences analyzed. According to the criteria above, 123 S
994 substitutions were retrieved. Subsequently, we classified the sequences
995 according to the profile of S substitutions above (referred to as S haplotype). We
996 excluded S haplotypes with ≤30 sequences from the downstream analyses.
997 According to the criterion above, 254 S haplotypes, composed of 375,121
998 sequences, were retrieved. The substitution profile was represented as a matrix,
999 where the rows and columns depict S haplotypes and S substitutions,
1000 respectively. An element in the matrix represents the status [presence (1) or
1001 absence (0)] of one S substitution in one S haplotype. Next, we identified a
1002 group of highly co-occurring substitutions (i.e., a pair of substitutions with >0.9
1003 Pearson's correlation in the substitution profile matrix) and clustered these
1004 substitutions as a substitution cluster (**Extended Data Fig. 1b**). For example,
1005 the L452R:F486V cluster represents the L452R and F486V substitutions. For
1006 one substitution cluster, the mean value of the substitution statuses (0 or 1) of
1007 the members of substitutions was calculated for each S haplotype, and the mean
1008 value was used as the substitution status of the substitution cluster. For example,
1009 if one S haplotype has L452R but not F486V, the substitution status of the
1010 L452R:F486V cluster of the haplotype was set at 0.5. Consequently, our dataset
1011 included the profile of 107 S substitutions/substitution clusters for 254 S
1012 haplotypes. Next, to set the major S haplotype of BA.2 as the reference S
1013 haplotype (or lineage) in the statistical model described below, we transformed
1014 the S substitution profile matrix by subtracting the substitution profile of the major
1015 S haplotype of BA.2 from those for all S haplotypes. Consequently, elements in
1016 the transformed S substitution profile matrix were converted to -1, 0, or 1: The
1017 zero value means that the status of a substitution in one haplotype is the same

1018   as that in the reference haplotype. The one value means that a substitution is
1019   present in one haplotype but not in the reference haplotype. The minus one
1020   value means that a substitution is absent in one haplotype but present in the
1021   reference haplotype. As a consequence of the transformation, the relative R_e
1022   value for the reference haplotype was set at 1 in the parameter estimation in the
1023   statistical model described below. Finally, the number of viral sequences
1024   belonging to each S haplotype collected on each day was counted, and the
1025   count matrix was constructed as an input for the statistical model described
1026   below.

1027   We assigned one major lineage classification (i.e., BA.1 BA.2, BA.4,
1028   BA.5, and BA.2.75) to each S haplotype: We examined the major lineage
1029   classification of respective viral sequences belonging to one S haplotype, and
1030   the classification of the S haplotype was determined according to the majority
1031   vote system.

1032   We constructed a Bayesian hierarchal model, which represents the
1033   epidemic dynamics of each S haplotype according to growth rate parameters for
1034   each S haplotype, which is represented by a linear combination of the effect of S
1035   substitutions. Arrays in the model index over one or more indices: L = 254 viral
1036   lineages (i.e., S haplotypes) l; S = 107 substitutions/substitution clusters s; and
1037   T = 229 days t. The model is:

$$\sigma_1 \sim Student\_t^+(5,0,10)$$

$$f_m \sim Laplace(0,10)$$

$$\beta_l \sim Student\_t\left(5, \sum_m f_m X_{lm}, \sigma_1\right)$$

$$y_{.t} \sim Multinomial\left(\sum_l y_{lt}, softmax(\alpha_. + \beta_. t)\right)$$

1038   The count of viral lineage l at time t, y_{lt}, is modeled as a hierarchal
1039   Multinomial logistic regression with intercept α_l and slope β_l parameters for
1040   lineage l. The slope (or viral lineage growth) parameter β_l is generated from
1041   Student's t distribution with five degrees of freedom, the mean value represented
1042   by $f_m X_{lm}$, and standard deviation, σ_1. $f_m X_{lm}$ denotes the linear combination of
1043   the effect of each substitution, where f_m and X_{lm} are the effect of substitution
1044 m and the profile of substitution m in lineage l (i.e., the substitution profile
1045   matrix constructed in the above paragraph), respectively. As a prior of f_m, the
1046   Laplace distribution with the mean 0 and the standard deviation 10 was set. In
1047   other words, we estimated the parameter f_m in the framework of Bayesian least
1048   absolute shrinkage and selection operator (LASSO). As a prior of σ_1, a half
1049   Student's t distribution with the mean 0 and the standard deviation 10 was set.
1050   For the other parameters, non-informative priors were set.

1051 The relative R_e of each viral lineage, r_l, was calculated according to the
1052 slope parameter β_l as:

$$r_l = exp(\gamma\beta_l)$$

1053 where γ is the average viral generation time (2.1 days)
1054 (http://sonorouschocolate.com/covid19/index.php?title=Estimating_Generation_
1055 Time_Of_Omicron). Similarly, the effect size of substitution m on the relative R_e,
1056 F_l, was calculated according to the coefficient f_l as:

$$F_l = exp(\gamma f_l)$$

1057 Parameter estimation was performed via the MCMC approach
1058 implemented in CmdStan v2.30.1 (https://mc-stan.org) with CmdStanr v0.5.3
1059 (https://mc-stan.org/cmdstanr/). Four independent MCMC chains were run with
1060 500 and 2,000 steps in the warmup and sampling iterations, respectively. We
1061 confirmed that all estimated parameters showed <1.01 R-hat convergence
1062 diagnostic values and >200 effective sampling size values, indicating that the
1063 MCMC runs were successfully convergent. The above analyses were performed
1064 in R v4.2.1 (https://www.r-project.org/). Information on the estimated effect size
1065 of each substitution or substitution cluster on relative R_e and relative R_e for each
1066 S haplotype are summarized in **Supplementary Table 2,3**.
1067 Since our model simply represents the viral lineage growth parameter
1068 (β_l) as the linear combination of the effects of S substitutions, the model can
1069 predict the total effect of a set of substations on relative R_e. Using this property
1070 of the model, we predicted the total effect of substitutions at the convergent sites
1071 (**Fig. 1h**) and the ancestral relative viral fitness for each node in the BA.5 tree
1072 (**Fig. 1i**). We calculated these values as the sum of the posterior means of the
1073 effects of substitutions of interest. To reconstruct the ancestral relative viral
1074 fitness of each node of the BA.5 tree, we first reconstructed the ancestral state of
1075 the S substitution profile in each node of the tree using a parsimony method,
1076 implemented by the phangorn package. Subsequently, we predicted the relative
1077 viral fitness for each node according to the reconstructed ancestral mutation
1078 profile for the node. The above analyses were performed in R v4.2.1
1079 (https://www.r-project.org/).
1080
1081 **Plasmid construction**
1082 Plasmids expressing the codon-optimized SARS-CoV-2 S proteins of B.1.1 (the
1083 parental D614G-bearing variant), BA.2 and BA.5, and BA.2.75 were prepared in
1084 our previous studies[27]. Plasmids expressing the codon-optimized S proteins of
1085 BQ.1.1, BA.5 S-based derivatives and BA.2 S-based derivatives were generated
1086 by site-directed overlap extension PCR using the primers listed in
1087 **Supplementary Table 5**. The resulting PCR fragment was digested with KpnI
1088 (New England Biolabs, Cat# R0142S) and NotI (New England Biolabs, Cat#
1089 R1089S) and inserted into the corresponding site of the pCAGGS vector[49].

1090 Nucleotide sequences were determined by DNA sequencing services (Eurofins),
1091 and the sequence data were analyzed by Sequencer v5.1 software (Gene
1092 Codes Corporation). Plasmids for yeast surface display were constructed by
1093 restriction enzyme-free cloning by incorporation of RBD genes ["construct 3" in
1094 ref.[23], covering residues 330–528] into the pJYDC1 plasmid (Addgene, Cat#
1095 162458). The primers are listed in **Supplementary Table 6**. The non-mutated
1096 RBD genes (BA.2, BA.5, and BQ.1) were purchased from Twist Biosciences.
1097

Neutralization assay

1099 Pseudoviruses were prepared as previously described[2,15,18,19,24,33,50-52]. Briefly,
1100 lentivirus (HIV-1)-based, luciferase-expressing reporter viruses were
1101 pseudotyped with SARS-CoV-2 S proteins. HEK293T cells (1,000,000 cells)
1102 were cotransfected with 1 μg psPAX2-IN/HiBiT[32], 1 μg pWPI-Luc2[32], and 500 ng
1103 plasmids expressing parental S or its derivatives using PEI Max (Polysciences,
1104 Cat# 24765-1) according to the manufacturer's protocol. Two days
1105 posttransfection, the culture supernatants were harvested and centrifuged. The
1106 pseudoviruses were stored at –80°C until use.
1107   The neutralization assay (**Fig. 2**) was prepared as previously
1108 described[2,15,18,24,33,50-52]. Briefly, the SARS-CoV-2 S pseudoviruses (counting
1109 ~20,000 relative light units) were incubated with serially diluted (120-fold to
1110 87,480-fold dilution at the final concentration) heat-inactivated sera at 37°C for 1
1111 hour. Pseudoviruses without sera were included as controls. Then, a 40 μl
1112 mixture of pseudovirus and serum/antibody was added to
1113 HOS-ACE2/TMPRSS2 cells (10,000 cells/50 μl) in a 96-well white plate. At 2
1114 d.p.i., the infected cells were lysed with a One-Glo luciferase assay system
1115 (Promega, Cat# E6130), a Bright-Glo luciferase assay system (Promega, Cat#
1116 E2650), or a britelite plus Reporter Gene Assay System (PerkinElmer, Cat#
1117 6066769), and the luminescent signal was measured using a GloMax explorer
1118 multimode microplate reader 3500 (Promega) or CentroXS3 (Berthhold
1119 Technologies). The assay of each serum sample was performed in triplicate,
1120 and the 50% neutralization titer (NT_{50}) was calculated using Prism 9 software
1121 v9.1.1 (GraphPad Software).
1122

SARS-CoV-2 preparation and titration

1124 The working virus stocks of SARS-CoV-2 were prepared and titrated as
1125 previously described[15,18,19,53]. In this study, clinical isolates of B.1.1 (strain
1126 TKYE610670; GISAID ID: EPI_ISL_479681)[17], Delta (B.1.617.2, strain
1127 TKYTK1734; GISAID ID: EPI_ISL_2378732)[16], BA.2 (strain TY40-385; GISAID
1128 ID: EPI_ISL_9595859)[15] and BA.5 (strain TKYS14631; GISAID ID:
1129 EPI_ISL_12812500)[2,54], and BQ.1.1 (strain TY41-796-P1; GISAID ID:
1130 EPI_ISL_15579783) were used. In brief, 20 μl of the seed virus was inoculated

31

1131   into VeroE6/TMPRSS2 cells (5,000,000 cells in a T-75 flask). One h.p.i., the
1132   culture medium was replaced with DMEM (low glucose) (Wako, Cat#
1133   041-29775) containing 2% FBS and 1% PS. At 3 d.p.i., the culture medium was
1134   harvested and centrifuged, and the supernatants were collected as the working
1135   virus stock.

1136   The titer of the prepared working virus was measured as the 50%
1137   tissue culture infectious dose (TCID$_{50}$). Briefly, one day before infection,
1138   VeroE6/TMPRSS2 cells (10,000 cells) were seeded into a 96-well plate. Serially
1139   diluted virus stocks were inoculated into the cells and incubated at 37°C for 4
1140   days. The cells were observed under a microscope to judge the CPE
1141   appearance. The value of TCID$_{50}$/ml was calculated with the Reed–Muench
1142   method[55].

1143   For verification of the sequences of SARS-CoV-2 working viruses, viral
1144   RNA was extracted from the working viruses using a QIAamp viral RNA mini kit
1145   (Qiagen, Cat# 52906) and viral genome sequences were analyzed as described
1146   above (see "Viral genome sequencing" section). Information on the unexpected
1147   substitutions detected is summarized in **Supplementary Table 6**, and the raw
1148   data   are   deposited   in   the   GitHub   repository
1149   (https://github.com/TheSatoLab/BQ.1).

1150

1151   **Yeast surface display**
1152   Yeast surface display (**Fig. 3a**) was performed as previously described[2,22,23].
1153   Briefly, the *S. cerevisiae* EBY100 yeasts were transformed with RBD expression
1154   plasmid and grown (220 rpm, 30°C, SD-CAA media). The expression media 1/9
1155   (ref.[56]) was inoculated to starting OD$_{600}$ 0.7–1 by overnight grown culture and
1156   cultivated for 24 hours at 20°C. The expression media was supplemented with
1157   10 nM DMSO solubilized bilirubin (Sigma-Aldrich, Cat# 14370-1G) for activation
1158   of eUnaG2 fluorescence (excitation at 498 nm, emission at 527 nm).

1159   Yeast cells were washed in ice-cold PBSB buffer (PBS with 1 mg/ml
1160   BSA), liquated (100 µl), transferred in an analysis solution and incubated for 8
1161   hours. The analysis solutions consisted of a series of CF®640R succinimidyl
1162   ester labeled (Biotium, Cat# 92108) ACE2 peptidase domain (residues 18–740)
1163   concentrations, PBSB buffer and 1 nM bilirubin. Incubated samples were
1164   washed twice with PBSB buffer and transferred into a 96-well plate (Thermo
1165   Fisher Scientific, Cat# 268200) for automated data acquisition by a CytoFLEX S
1166   Flow Cytometer (Beckman Coulter, USA, Cat#. N0-V4-B2-Y4). The gating and
1167   analysis strategies were described previously[23]. The titration curves were fitted
1168   with nonlinear least-squares regression using Python v3.7 and two additional
1169   parameters to describe the titration curve[23].

1170

1171   **Pseudovirus infection**

32

1172 Pseudovirus infection (**Fig. 3b**) was performed as previously
1173 described[2,15,18,19,24,33,50-52]. Briefly, the amount of pseudoviruses prepared was
1174 quantified by the HiBiT assay using a Nano Glo HiBiT lytic detection system
1175 (Promega, Cat# N3040) as previously described[32,57]. For measurement of
1176 pseudovirus infectivity, the same amount of pseudoviruses (normalized to the
1177 HiBiT value, which indicates the amount of HIV-1 p24 antigen) was inoculated
1178 into HOS-ACE2/TMPRSS2 cells, HEK293-ACE2 cells or
1179 HEK293-ACE2/TMPRSS2 cells and viral infectivity was measured as described
1180 above (see "Neutralization assay" section). For analysis of the effect of
1181 TMPRSS2 on pseudovirus infectivity (**Fig. 3c**), the fold change of the values of
1182 HEK293-ACE2/TMPRSS2 to HEK293-ACE2 was calculated.
1183

1184 **SARS-CoV-2 S-based fusion assay**
1185 A SARS-CoV-2 S-based fusion assay (**Fig. 3d–f**) was performed as previously
1186 described[2,15-19,25,53]. Briefly, on day 1, effector cells (i.e., S-expressing cells) and
1187 target cells (Calu-3/DSP$_{1-7}$ cells) were prepared at a density of 0.6–0.8 $\times$ 10^6
1188 cells in a 6-well plate. On day 2, for the preparation of effector cells, HEK293
1189 cells were cotransfected with the S expression plasmids (400 ng) and pDSP$_{8-11}$
1190 (ref.[58]) (400 ng) using TransIT-LT1 (Takara, Cat# MIR2300). On day 3 (24 hours
1191 posttransfection), 16,000 effector cells were detached and reseeded into a
1192 96-well black plate (PerkinElmer, Cat# 6005225), and target cells were reseeded
1193 at a density of 1,000,000 cells/2 ml/well in 6-well plates. On day 4 (48 hours
1194 posttransfection), target cells were incubated with EnduRen live cell substrate
1195 (Promega, Cat# E6481) for 3 hours and then detached, and 32,000 target cells
1196 were added to a 96-well plate with effector cells. *Renilla* luciferase activity was
1197 measured at the indicated time points using Centro XS3 LB960 (Berthhold
1198 Technologies). For measurement of the surface expression level of the S protein,
1199 effector cells were stained with rabbit anti-SARS-CoV-2 S S1/S2 polyclonal
1200 antibody (Thermo Fisher Scientific, Cat# PA5-112048, 1:100). Normal rabbit IgG
1201 (Southern Biotech, Cat# 0111-01, 1:100) was used as a negative control, and
1202 APC-conjugated goat anti-rabbit IgG polyclonal antibody (Jackson
1203 ImmunoResearch, Cat# 111-136-144, 1:50) was used as a secondary antibody.
1204 The surface expression level of S proteins (**Fig. 3d**) was measured using a
1205 FACS Canto II (BD Biosciences) and the data were analyzed using FlowJo
1206 software v10.7.1 (BD Biosciences). For calculation of fusion activity, *Renilla*
1207 luciferase activity was normalized to the MFI of surface S proteins. The
1208 normalized value (i.e., *Renilla* luciferase activity per the surface S MFI) is shown
1209 as fusion activity.
1210

1211 **AO-ALI model**

1212 An airway organoid (AO) model was generated according to our previous
1213 report[2,59]. Briefly, normal human bronchial epithelial cells (NHBEs, Cat#
1214 CC-2540, Lonza) were used to generate AOs. NHBEs were suspended in 10
1215 mg/ml cold Matrigel growth factor reduced basement membrane matrix (Corning,
1216 Cat# 354230). Fifty microliters of cell suspension were solidified on prewarmed
1217 cell culture-treated multiple dishes (24-well plates; Thermo Fisher Scientific,
1218 Cat# 142475) at 37°C for 10 min, and then, 500 µl of expansion medium was
1219 added to each well. AOs were cultured with AO expansion medium for 10 days.
1220 For maturation of the AOs, expanded AOs were cultured with AO differentiation
1221 medium for 5 days.
1222   The AO-ALI model (**Fig. 3j**) was generated according to our previous
1223 report[2,59]. For generation of AO-ALI, expanding AOs were dissociated into single
1224 cells, and then were seeded into Transwell inserts (Corning, Cat# 3413) in a
1225 24-well plate. AO-ALI was cultured with AO differentiation medium for 5 days to
1226 promote their maturation. AO-ALI was infected with SARS-CoV-2 from the apical
1227 side.
1228
1229 **Preparation of human airway and alveolar epithelial cells from human**
1230 **iPSCs**
1231 The air-liquid interface culture of airway and alveolar epithelial cells (**Fig. 3k,l**)
1232 was differentiated from human iPSC-derived lung progenitor cells as previously
1233 described[2,15,54,60-62]. Briefly, alveolar progenitor cells were induced stepwise from
1234 human iPSCs according to a 21-day and 4-step protocol[60]. At day 21, alveolar
1235 progenitor cells were isolated with the specific surface antigen carboxypeptidase
1236 M and seeded onto the upper chamber of a 24-well Cell Culture Insert (Falcon,
1237 #353104), followed by 28-day and 7-day differentiation of airway and alveolar
1238 epithelial cells, respectively. Alveolar differentiation medium with
1239 dexamethasone (Sigma-Aldrich, Cat# D4902), KGF (PeproTech, Cat# 100-19),
1240 8-Br-cAMP (Biolog, Cat# B007), 3-isobutyl 1-methylxanthine (IBMX) (Fujifilm
1241 Wako, Cat# 095-03413), CHIR99021 (Axon Medchem, Cat# 1386), and
1242 SB431542 (Fujifilm Wako, Cat# 198-16543) was used for the induction of
1243 alveolar epithelial cells. PneumaCult ALI (STEMCELL Technologies, Cat#
1244 ST-05001) with heparin (Nacalai Tesque, Cat# 17513-96) and Y-27632 (LC
1245 Laboratories, Cat# Y-5301) hydrocortisone (Sigma-Aldrich, Cat# H0135) was
1246 used for induction of airway epithelial cells.
1247
1248 **Airway-on-a-chips**
1249 Airway-on-a-chips (**Fig. 3m**) were prepared as previously described[2,26,54].
1250 Human lung microvascular endothelial cells (HMVEC-L) were obtained from
1251 Lonza (Cat# CC-2527) and cultured with EGM-2-MV medium (Lonza, Cat#
1252 CC-3202). For preparation of the airway-on-a-chip, first, the bottom channel of a

1253   polydimethylsiloxane (PDMS) device was precoated with fibronectin (3 µg/ml,
1254   Sigma-Aldrich, Cat# F1141). The microfluidic device was generated according to
1255   our previous report[63]. HMVEC-L cells were suspended at 5,000,000 cells/ml in
1256   EGM2-MV medium. Then, 10 µl of suspension medium was injected into the
1257   fibronectin-coated bottom channel of the PDMS device. Then, the PDMS device
1258   was turned upside down and incubated. After 1 hour, the device was turned over,
1259   and the EGM2-MV medium was added into the bottom channel. After 4 days,
1260   AOs were dissociated and seeded into the top channel. AOs were generated
1261   according to our previous report[59]. AOs were dissociated into single cells and
1262   then suspended at 5,000,000 cells/ml in the AO differentiation medium. Ten
1263   microliter suspension medium was injected into the top channel. After 1 hour, the
1264   AO differentiation medium was added to the top channel. In the infection
1265   experiments (**Fig. 3m**), the AO differentiation medium containing either BA.2,
1266   BA.5, BQ.1.1 or Delta isolate (500 $TCID_{50}$) was inoculated into the top channel.
1267   At 2 h.p.i., the top and bottom channels were washed and cultured with AO
1268   differentiation and EGM2-MV medium, respectively. The culture supernatants
1269   were collected, and viral RNA was quantified using RT–qPCR (see "RT–qPCR"
1270   section above).
1271
1272   **Microfluidic device**
1273   A microfluidic device was generated according to our previous report[2,63]. Briefly,
1274   the microfluidic device consisted of two layers of microchannels separated by a
1275   semipermeable membrane. The microchannel layers were fabricated from
1276   PDMS using a soft lithographic method. PDMS prepolymer (Dow Corning, Cat#
1277   SYLGARD 184) at a base to curing agent ratio of 10:1 was cast against a mold
1278   composed of SU-8 2150 (MicroChem, Cat# SU-8 2150) patterns formed on a
1279   silicon wafer. The cross-sectional size of the microchannels was 1 mm in width
1280   and 330 µm in height. Access holes were punched through the PDMS using a
1281   6-mm biopsy punch (Kai Corporation, Cat# BP-L60K) to introduce solutions into
1282   the microchannels. Two PDMS layers were bonded to a PET membrane
1283   containing 3.0-µm pores (Falcon, Cat# 353091) using a thin layer of liquid PDMS
1284   prepolymer as the mortar. PDMS prepolymer was spin-coated (4000 rpm for 60
1285   sec) onto a glass slide. Subsequently, both the top and bottom channel layers
1286   were placed on the glass slide to transfer the thin layer of PDMS prepolymer
1287   onto the embossed PDMS surfaces. The membrane was then placed onto the
1288   bottom layer and sandwiched with the top layer. The combined layers were left
1289   at room temperature for 1 day to remove air bubbles and then placed in an oven
1290   at 60°C overnight to cure the PDMS glue. The PDMS devices were sterilized by
1291   placing them under UV light for 1 hour before the cell culture.
1292
1293   **SARS-CoV-2 infection**

35

1294 One day before infection, Vero cells (10,000 cells), VeroE6/TMPRSS2 cells
1295 (10,000 cells) and Calu-3 cells (10,000 cells) were seeded into a 96-well plate.
1296 SARS-CoV-2 [1,000 $TCID_{50}$ for Vero cells (**Fig. 3g**); 100 $TCID_{50}$ for
1297 VeroE6/TMPRSS2 cells (**Fig. 3h**) and Calu-3 cells (**Fig. 3i**)] was inoculated and
1298 incubated at 37°C for 1 hour. The infected cells were washed, and 180 µl of
1299 culture medium was added. The culture supernatant (10 µl) was harvested at the
1300 indicated timepoints and used for RT–qPCR to quantify the viral RNA copy
1301 number (see "RT–qPCR" section below). In the infection experiments using
1302 human iPSC-derived airway and lung epithelial cells (**Fig. 3k,l**), working viruses
1303 were diluted with Opti-MEM (Thermo Fisher Scientific, Cat# 11058021). The
1304 diluted viruses (1,000 $TCID_{50}$ in 100 µl) were inoculated onto the apical side of
1305 the culture and incubated at 37 °C for 1 hour. The inoculated viruses were
1306 removed and washed twice with Opti-MEM. For collection of the viruses, 100 µl
1307 Opti-MEM was applied onto the apical side of the culture and incubated at
1308 37 °C for 10 minutes. The Opti-MEM was collected and used for RT–qPCR to
1309 quantify the viral RNA copy number (see "RT–qPCR" section below). The
1310 infection experiments using an airway-on-a-chip system (**Fig. 3m**) were
1311 performed as described above (see "Airway-on-a-chips" section).
1312
1313 **RT–qPCR**
1314 RT–qPCR was performed as previously described[2,15-19,53,54,64]. Briefly, 5 µl
1315 culture supernatant was mixed with 5 µl of 2 × RNA lysis buffer [2% Triton X-100
1316 (Nacalai Tesque, Cat# 35501-15), 50 mM KCl, 100 mM Tris-HCl (pH 7.4), 40%
1317 glycerol, 0.8 U/µl recombinant RNase inhibitor (Takara, Cat# 2313B)] and
1318 incubated at room temperature for 10 min. RNase-free water (90 µl) was added,
1319 and the diluted sample (2.5 µl) was used as the template for real-time RT-PCR
1320 performed according to the manufacturer's protocol using One Step TB Green
1321 PrimeScript PLUS RT-PCR kit (Takara, Cat# RR096A) and the following
1322 primers: Forward *N*, 5'-AGC CTC TTC TCG TTC CTC ATC AC-3'; and Reverse
1323 *N*, 5'-CCG CCA TTG CCA GCC ATT C-3'. The viral RNA copy number was
1324 standardized with a SARS-CoV-2 direct detection RT-qPCR kit (Takara, Cat#
1325 RC300A). Fluorescent signals were acquired using a QuantStudio 1 Real-Time
1326 PCR system (Thermo Fisher Scientific), QuantStudio 3 Real-Time PCR system
1327 (Thermo Fisher Scientific), QuantStudio 5 Real-Time PCR system (Thermo
1328 Fisher Scientific), StepOne Plus Real-Time PCR system (Thermo Fisher
1329 Scientific), CFX Connect Real-Time PCR Detection system (Bio-Rad), Eco
1330 Real-Time PCR System (Illumina), qTOWER3 G Real-Time System (Analytik
1331 Jena) Thermal Cycler Dice Real Time System III (Takara) or 7500 Real-Time
1332 PCR System (Thermo Fisher Scientific).
1333
1334 **Animal experiments**

36

1335 Animal experiments (**Fig. 4 and Extended Data Fig. 2**) were performed as
1336 previously described[27]. Syrian hamsters (male, 4 weeks old) were purchased
1337 from Japan SLC Inc. (Shizuoka, Japan). For the virus infection experiments,
1338 hamsters were anesthetized by intramuscular injection of a mixture of 0.15
1339 mg/kg medetomidine hydrochloride (Domitor®, Nippon Zenyaku Kogyo), 2.0
1340 mg/kg midazolam (Dormicum®, Fujifilm Wako, Cat# 135-13791) and 2.5 mg/kg
1341 butorphanol (Vetorphale®, Meiji Seika Pharma) or 0.15 mg/kg medetomidine
1342 hydrochloride, 4.0 mg/kg alphaxaone (Alfaxan®, Jurox) and 2.5 mg/kg
1343 butorphanol. BA.5, BQ1.1 and Delta (10,000 $TCID_{50}$ in 100 µl) or saline (100 µl)
1344 was intranasally inoculated under anesthesia. Oral swabs were collected at the
1345 indicated timepoints. Body weight was recorded daily by 7 d.p.i. Enhanced
1346 pause (Penh), the ratio of time to peak expiratory follow relative to the total
1347 expiratory time (Rpef) were measured every day until 7 d.p.i. (see below). Lung
1348 tissues were anatomically collected at 2 and 5 d.p.i. The viral RNA load in the
1349 oral swabs and respiratory tissues was determined by RT–qPCR. These tissues
1350 were also used for IHC and histopathological analyses (see below).
1351

1352 **Lung function test**
1353 Lung function tests (**Fig. 4a**) were routinely performed as previously
1354 described[2,15,17,18,54]. The two respiratory parameters (Penh and Rpef) were
1355 measured by using a Buxco Small Animal Whole Body Plethysmography system
1356 (DSI) according to the manufacturer's instructions. In brief, a hamster was
1357 placed in an unrestrained plethysmography chamber and allowed to acclimatize
1358 for 30 seconds. Then, data were acquired over a 2.5-minute period by using
1359 FinePointe Station and Review software v2.9.2.12849 (DSI).
1360

1361 **Immunohistochemistry**
1362 Immunohistochemistry (IHC) (**Fig. 4c and Extended Data Fig. 2**) was
1363 performed as previously described[2,15,17,18,54] using an Autostainer Link 48 (Dako).
1364 The deparaffinized sections were exposed to EnVision FLEX target retrieval
1365 solution high pH (Agilent, Cat# K8004) for 20 minutes at 97°C for activation, and
1366 a mouse anti-SARS-CoV-2 N monoclonal antibody (clone 1035111, R&D
1367 Systems, Cat# MAB10474-SP, 1:400) was used as a primary antibody. The
1368 sections were sensitized using EnVision FLEX for 15 minutes and visualized by
1369 peroxidase-based enzymatic reaction with 3,3'-diaminobenzidine
1370 tetrahydrochloride (Dako, Cat# DM827) as substrate for 5 minutes. The N
1371 protein positivity was evaluated by certificated pathologists as previously
1372 described[2,15,17,18,54]. Images were incorporated as virtual slides by NDP.scan
1373 software v3.2.4 (Hamamatsu Photonics). The N-protein positivity was measured
1374 as the area using Fiji software v2.2.0 (ImageJ).
1375

H&E staining

H&E staining (**Fig. 4d**) was performed as previously described[2,15,17,18,54]. Briefly, excised animal tissues were fixed with 10% formalin neutral buffer solution and processed for paraffin embedding. The paraffin blocks were sectioned at a thickness of 3 μm and then mounted on MAS-GP-coated glass slides (Matsunami Glass, Cat# S9901). H&E staining was performed according to a standard protocol.

Histopathological scoring

Histopathological scoring (**Fig. 4e**) was performed as previously described[2,15,17,18,54]. Pathological features, including (i) bronchitis or bronchiolitis, (ii) hemorrhage with congestive edema, (iii) alveolar damage with epithelial apoptosis and macrophage infiltration, (iv) hyperplasia of type II pneumocytes, and (v) the area of hyperplasia of large type II pneumocytes, were evaluated by certified pathologists, and the degree of these pathological findings was arbitrarily scored using a four-tiered system as 0 (negative), 1 (weak), 2 (moderate), and 3 (severe). The "large type II pneumocytes" are type II pneumocytes with hyperplasia exhibiting more than 10-μm-diameter nuclei. We described "large type II pneumocytes" as one of the notable histopathological features of SARS-CoV-2 infection in our previous studies[2,15,17,18,54]. The total histological score is the sum of these five indices.

Statistics and reproducibility

Statistical significance was tested using a two-sided Mann–Whitney U test, a two-sided Student's t test, a two-sided Welch's t test, or a two-sided paired t-test unless otherwise noted. The tests above were performed using Prism 9 software v9.1.1 (GraphPad Software).

In the time-course experiments (**Fig. 3e–m, 4a–b,e**), a multiple regression analysis including experimental conditions (i.e., the types of infected viruses) as explanatory variables and timepoints as qualitative control variables was performed to evaluate the difference between experimental conditions thorough all timepoints. The initial time point was removed from the analysis. The P value was calculated by a two-sided Wald test. Subsequently, familywise error rates (FWERs) were calculated by the Holm method. These analyses were performed on R v4.1.2 (https://www.r-project.org/).

Principal component analysis to representing the antigenicity of the S proteins was performed (**Fig. 2d**). The NT50 values for biological replicates were scaled, and subsequently, principal component analysis was performed using the prcomp function on R v4.1.2 (https://www.r-project.org/).

In **Fig. 4c,d, and Extended Data Fig. 2**, photographs shown are the representative areas of at least two independent experiments by using four

1417   hamsters at each timepoint.

1418

1419   **Data availability**

1420   All databases/datasets used in this study are available from the GISAID

1421   database (https://www.gisaid.org; EPI_SET_221203cz, EPI_SET_221203ep,

1422   EPI_SET_221203qr, EPI_SET_221203se, and EPI_SET_221203vk) and GenBank

1423   database (https://www.ncbi.nlm.nih.gov/genbank/). Viral genome sequencing

1424   data for working viral stocks are available in the GitHub repository

1425   (https://github.com/TheSatoLab/BQ.1).

1426

1427   **Code availability**

1428   The computational codes used in the present study and the GISAID

1429   supplemental tables for EPI_SET_221203cz, EPI_SET_221203ep,

1430   EPI_SET_221203qr, EPI_SET_221203se, and EPI_SET_221203vk are available
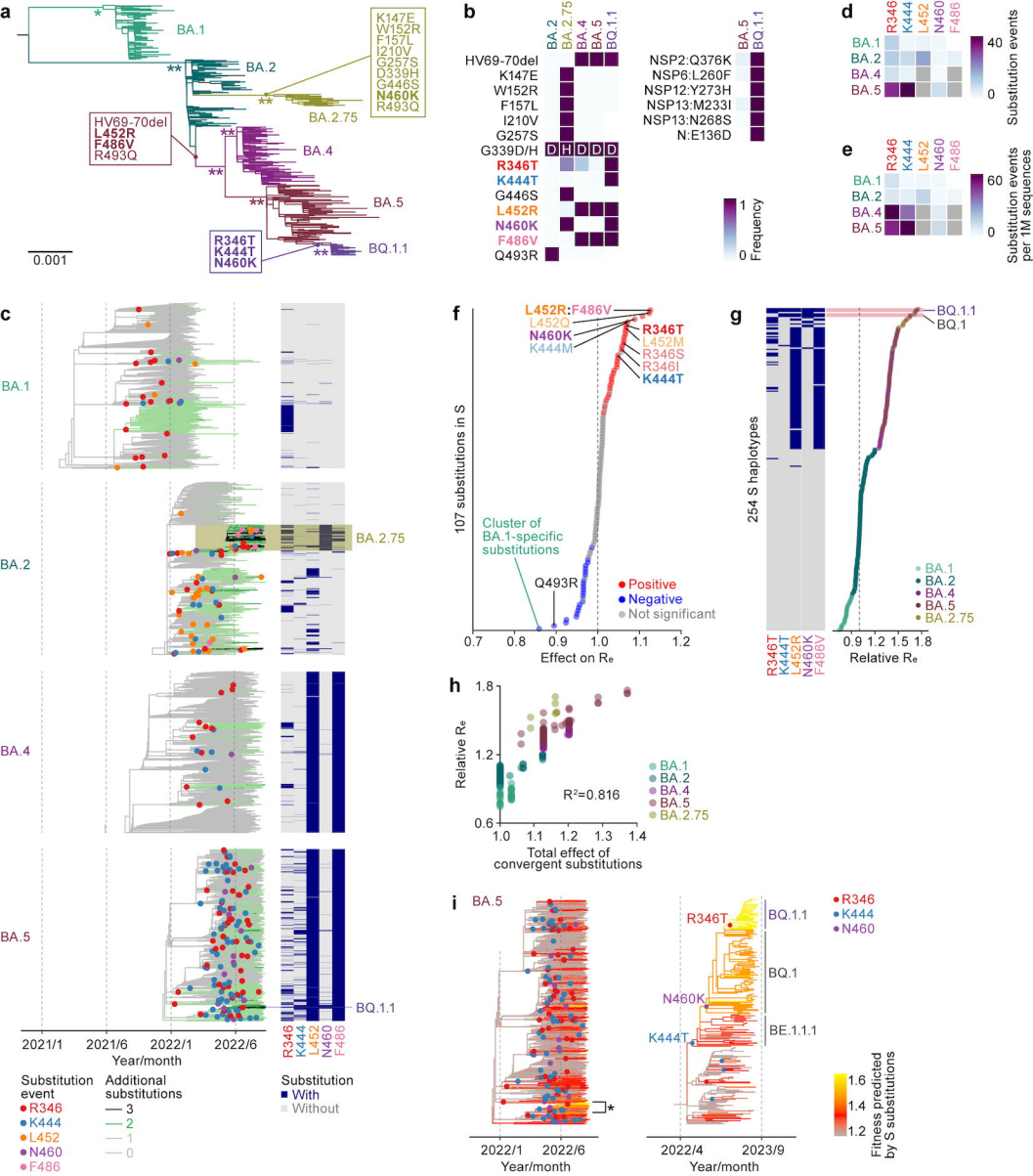
1431   in the GitHub repository (https://github.com/TheSatoLab/BQ.1).

References

1432

1433   1   WHO. "Tracking SARS-CoV-2 variants (October 28, 2022)"
1434   https://www.who.int/en/activities/tracking-SARS-CoV-2-variants. (2022).

1435   2   Saito, A. *et al.* Virological characteristics of the SARS-CoV-2 Omicron
1436   BA.2.75 variant. *Cell Host Microbe*, doi:10.1016/j.chom.2022.10.003
1437   (2022).

1438   3   Wang, Q. *et al.* Antibody evasion by SARS-CoV-2 Omicron subvariants
1439   BA.2.12.1, BA.4, & BA.5. *Nature*, doi:10.1038/s41586-022-05053-w
1440   (2022).

1441   4   Tuekprakhon, A. *et al.* Antibody escape of SARS-CoV-2 Omicron BA.4
1442   and BA.5 from vaccine and BA.1 serum. *Cell* **185**, 2422-2433 e2413,
1443   doi:10.1016/j.cell.2022.06.005 (2022).

1444   5   Cao, Y. *et al.* BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by
1445   Omicron infection. *Nature*, doi:10.1038/s41586-022-04980-y (2022).

1446   6   Cao, Y. *et al.* Imprinted SARS-CoV-2 humoral immunity induces
1447   convergent Omicron RBD evolution. *BioRxiv*, doi:
1448   https://doi.org/10.1101/2022.1109.1115.507787 (2022).

1449   7   Makowski, E. K., Schardt, J. S., Smith, M. D. & Tessier, P. M. Mutational
1450   analysis of SARS-CoV-2 variants of concern reveals key tradeoffs
1451   between receptor affinity and antibody escape. *PLoS Comput Biol* **18**,
1452   e1010160, doi:10.1371/journal.pcbi.1010160 (2022).

1453   8   Aggarwal, A. *et al.* Mechanistic insights into the effects of key mutations
1454   on SARS-CoV-2 RBD-ACE2 binding. *Phys Chem Chem Phys* **23**,
1455   26451-26458, doi:10.1039/d1cp04005g (2021).

1456   9   Deshpande, A., Harris, B. D., Martinez-Sobrido, L., Kobie, J. J. & Walter,
1457   M. R. Epitope Classification and RBD Binding Properties of Neutralizing
1458   Antibodies Against SARS-CoV-2 Variants of Concern. *Front Immunol* **12**,
1459   691715, doi:10.3389/fimmu.2021.691715 (2021).

1460   10   Cao, Y. *et al.* Characterization of the enhanced infectivity and antibody
1461   evasion of Omicron BA.2.75. *Cell Host Microbe* **30**, 1527-1539 e1525,
1462   doi:10.1016/j.chom.2022.09.018 (2022).

1463   11   Qu, P. *et al.* Evasion of neutralizing antibody responses by the
1464   SARS-CoV-2 BA.2.75 variant. *Cell Host Microbe* **30**, 1518-1526 e1514,
1465   doi:10.1016/j.chom.2022.09.015 (2022).

1466   12   Wang, Q. *et al.* Antigenic characterization of the SARS-CoV-2 Omicron
1467   subvariant BA.2.75. *Cell Host Microbe* **30**, 1512-1517 e1514,
1468   doi:10.1016/j.chom.2022.09.002 (2022).

1469   13   Zhou, T. *et al.* Structural basis for potent antibody neutralization of
1470   SARS-CoV-2 variants including B.1.1.529. *Science* **376**, eabn8897,
1471   doi:10.1126/science.abn8897 (2022).

1472   14   Cao, Y. *et al.* Omicron escapes the majority of existing SARS-CoV-2

40

1473   neutralizing antibodies. *Nature* **602**, 657-663,
1474   doi:10.1038/s41586-021-04385-3 (2022).
1475 15 Kimura, I. *et al.* Virological characteristics of the novel SARS-CoV-2
1476   Omicron variants including BA.4 and BA.5. *Cell* **185**, 3992-4007 e3916
1477   (2022).
1478 16 Saito, A. *et al.* Enhanced fusogenicity and pathogenicity of SARS-CoV-2
1479   Delta P681R mutation. *Nature* **602**, 300-306,
1480   doi:10.1038/s41586-021-04266-9 (2022).
1481 17 Suzuki, R. *et al.* Attenuated fusogenicity and pathogenicity of
1482   SARS-CoV-2 Omicron variant. *Nature*, doi:10.1038/s41586-022-04462-1
1483   (2022).
1484 18 Yamasoba, D. *et al.* Virological characteristics of the SARS-CoV-2
1485   Omicron BA.2 spike. *Cell*, doi:10.1016/j.cell.2022.04.035 (2022).
1486 19 Motozono, C. *et al.* SARS-CoV-2 spike L452R variant evades cellular
1487   immunity and increases infectivity. *Cell Host Microbe* **29**, 1124-1136,
1488   doi:10.1016/j.chom.2021.06.006 (2021).
1489 20 GitHub. "BE.1.1.1 sublineage with Orf1b:Y264H and S:N460K (69
1490   sequences) emerged in Nigeria (14 seqs) (August 26, 2022)".
1491   https://github.com/cov-lineages/pango-designation/issues/993. (2022).
1492 21 Arora, P. *et al.* Omicron sublineage BQ.1.1 resistance to monoclonal
1493   antibodies. *Lancet Infect Dis*, doi:10.1016/S1473-3099(22)00733-2
1494   (2022).
1495 22 Dejnirattisai, W. *et al.* SARS-CoV-2 Omicron-B.1.1.529 leads to
1496   widespread escape from neutralizing antibody responses. *Cell* **185**,
1497   467-484 e415, doi:10.1016/j.cell.2021.12.046 (2022).
1498 23 Zahradnik, J. *et al.* SARS-CoV-2 variant prediction and antiviral drug
1499   design are enabled by RBD in vitro evolution. *Nat Microbiol* **6**, 1188-1198,
1500   doi:10.1038/s41564-021-00954-4 (2021).
1501 24 Kimura, I. *et al.* The SARS-CoV-2 Lambda variant exhibits enhanced
1502   infectivity and immune resistance. *Cell Rep* **38**, 110218,
1503   doi:10.1016/j.celrep.2021.110218 (2022).
1504 25 Nasser, H. *et al.* Monitoring fusion kinetics of viral and target cell
1505   membranes in living cells using a SARS-CoV-2 spike-protein-mediated
1506   membrane fusion assay. *STAR Protoc* **3**, 101773,
1507   doi:10.1016/j.xpro.2022.101773 (2022).
1508 26 Hashimoto, R. *et al.* SARS-CoV-2 disrupts the respiratory vascular barrier
1509   by suppressing Claudin-5 expression. *Sci Adv* **8**, eabo6783, doi:doi:
1510   10.1126/sciadv.abo6783 (2022).
1511 27 . (!!! INVALID CITATION !!! 2,15-18).
1512 28 Martin, D. P. *et al.* Selection Analysis Identifies Clusters of Unusual
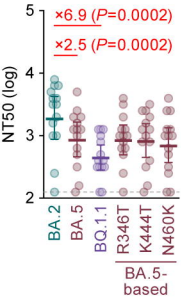1513   Mutational Changes in Omicron Lineage BA.1 That Likely Impact Spike

41

1514   Function. *Mol Biol Evol* **39**, doi:10.1093/molbev/msac061 (2022).

1515   29   Wang, Q. *et al.* Alarming antibody evasion properties of rising
1516   SARS-CoV-2 BQ and XBB subvariants. *BioRxiv*, doi:
1517   https://doi.org/10.1101/2022.1111.1123.517532 (2022).

1518   30   Uraki, R. *et al.* Characterization and antiviral susceptibility of
1519   SARS-CoV-2 Omicron/BA.2. *Nature*, doi:10.1038/s41586-022-04856-1
1520   (2022).

1521   31   Sasaki, A., Lion, S. & Boots, M. Antigenic escape selects for the evolution
1522   of higher pathogen transmission and virulence. *Nat Ecol Evol* **6**, 51-62,
1523   doi:10.1038/s41559-021-01603-z (2022).

1524   32   Ozono, S. *et al.* SARS-CoV-2 D614G spike mutation increases entry
1525   efficiency with enhanced ACE2-binding affinity. *Nat Commun* **12**, 848,
1526   doi:10.1038/s41467-021-21118-2 (2021).

1527   33   Ferreira, I. *et al.* SARS-CoV-2 B.1.617 mutations L452R and E484Q are
1528   not synergistic for antibody evasion. *J Infect Dis* **224**, 989-994,
1529   doi:10.1093/infdis/jiab368 (2021).

1530   34   Matsuyama, S. *et al.* Enhanced isolation of SARS-CoV-2 by
1531   TMPRSS2-expressing cells. *Proc Natl Acad Sci U S A* **117**, 7001-7003,
1532   doi:10.1073/pnas.2002589117 (2020).

1533   35   Yamamoto, M. *et al.* The anticoagulant nafamostat potently inhibits
1534   SARS-CoV-2 S protein-mediated fusion in a cell fusion assay system and
1535   viral infection *in vitro* in a cell-yype-dependent manner. *Viruses* **12**,
1536   doi:10.3390/v12060629 (2020).

1537   36   Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one
1538   FASTQ preprocessor. *Bioinformatics* **34**, i884-i890,
1539   doi:10.1093/bioinformatics/bty560 (2018).

1540   37   Li, H. & Durbin, R. Fast and accurate short read alignment with
1541   Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760,
1542   doi:10.1093/bioinformatics/btp324 (2009).

1543   38   Li, H. *et al.* The sequence alignment/map format and SAMtools.
1544   *Bioinformatics* **25**, 2078-2079, doi:10.1093/bioinformatics/btp352 (2009).

1545   39   Cingolani, P. *et al.* A program for annotating and predicting the effects of
1546   single nucleotide polymorphisms, SnpEff: SNPs in the genome of
1547   Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly (Austin)* **6**, 80-92,
1548   doi:10.4161/fly.19695 (2012).

1549   40   Khare, S. *et al.* GISAID's role in pandemic response. *China CDC Wkly* **3**,
1550   1049-1051, doi:10.46234/ccdcw2021.255 (2021).

1551   41   Li, H. Minimap2: pairwise alignment for nucleotide sequences.
1552   *Bioinformatics* **34**, 3094-3100, doi:10.1093/bioinformatics/bty191 (2018).

1553   42   Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: a tool
1554   for automated alignment trimming in large-scale phylogenetic analyses.

1555   *Bioinformatics* **25**, 1972-1973, doi:10.1093/bioinformatics/btp348 (2009).

1556   43   Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313, doi:10.1093/bioinformatics/btu033 (2014).

1559   44   Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**, 772-780, doi:10.1093/molbev/mst010 (2013).

1562   45   Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**, 268-274, doi:10.1093/molbev/msu300 (2015).

1566   46   Sagulenko, P., Puller, V. & Neher, R. A. TreeTime: Maximum-likelihood phylodynamic analysis. *Virus Evol* **4**, vex042, doi:10.1093/ve/vex042 (2018).

1569   47   Huerta-Cepas, J., Serra, F. & Bork, P. ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Mol Biol Evol* **33**, 1635-1638, doi:10.1093/molbev/msw046 (2016).

1572   48   Obermeyer, F. *et al.* Analysis of 6.4 million SARS-CoV-2 genomes identifies mutations associated with fitness. *Science* **376**, 1327-1332, doi:10.1126/science.abm1208 (2022).

1575   49   Niwa, H., Yamamura, K. & Miyazaki, J. Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene* **108**, 193-199, doi:10.1016/0378-1119(91)90434-d (1991).

1578   50   Uriu, K. *et al.* Neutralization of the SARS-CoV-2 Mu variant by convalescent and vaccine serum. *N Engl J Med* **385**, 2397-2399, doi:10.1056/NEJMc2114706 (2021).

1581   51   Uriu, K. *et al.* Characterization of the immune resistance of SARS-CoV-2 Mu variant and the robust immunity induced by Mu infection. *J Infect Dis*, doi:10.1093/infdis/jiac053 (2022).

1584   52   Yamasoba, D. *et al.* Neutralisation sensitivity of SARS-CoV-2 omicron subvariants to therapeutic monoclonal antibodies. *Lancet Infect Dis* **22**, 942-943, doi:10.1016/S1473-3099(22)00365-6 (2022).

1587   53   Kimura, I. *et al.* SARS-CoV-2 spike S375F mutation characterizes the Omicron BA.1 variant. *BioRxiv*, doi: https://doi.org/10.1101/2022.1104.1103.486864 (2022).

1590   54   Tamura, T. *et al.* Comparative pathogenicity of SARS-CoV-2 Omicron subvariants including BA.1, BA.2, and BA.5. *BioRxiv*, doi: https://doi.org/10.1101/2022.1108.1105.502758 (2022).

1593   55   Reed, L. J. & Muench, H. A simple method of estimating fifty percent endpoints. *Am J Hygiene* **27**, 493-497 (1938).

1595   56   Zahradnik, J. *et al.* A protein-engineered, enhanced yeast display

1596   platform for rapid evolution of challenging targets. *ACS Synth Biol* **10**,
1597   3445-3460, doi:10.1021/acssynbio.1c00395 (2021).

1598 57 Ozono, S., Zhang, Y., Tobiume, M., Kishigami, S. & Tokunaga, K.
1599   Super-rapid quantitation of the production of HIV-1 harboring a
1600   luminescent peptide tag. *J Biol Chem* **295**, 13023-13030,
1601   doi:10.1074/jbc.RA120.013887 (2020).

1602 58 Kondo, N., Miyauchi, K. & Matsuda, Z. Monitoring viral-mediated
1603   membrane fusion using fluorescent reporter methods. *Curr Protoc Cell*
1604   *Biol* **Chapter 26**, Unit 26 29, doi:10.1002/0471143030.cb2609s50 (2011).

1605 59 Sano, E. *et al.* Cell response analysis in SARS-CoV-2 infected bronchial
1606   organoids. *Commun Biol* **5**, 516, doi:10.1038/s42003-022-03499-2
1607   (2022).

1608 60 Yamamoto, Y. *et al.* Long-term expansion of alveolar stem cells derived
1609   from human iPS cells in organoids. *Nat Methods* **14**, 1097-1106,
1610   doi:10.1038/nmeth.4448 (2017).

1611 61 Konishi, S. *et al.* Directed induction of functional multi-ciliated cells in
1612   proximal airway epithelial spheroids from human pluripotent stem cells.
1613   *Stem Cell Reports* **6**, 18-25, doi:10.1016/j.stemcr.2015.11.010 (2016).

1614 62 Gotoh, S. *et al.* Generation of alveolar epithelial spheroids via isolated
1615   progenitor cells from human pluripotent stem cells. *Stem Cell Reports* **3**,
1616   394-403, doi:10.1016/j.stemcr.2014.07.005 (2014).

1617 63 Deguchi, S. *et al.* Usability of polydimethylsiloxane-based microfluidic
1618   devices in pharmaceutical research using human hepatocytes. *ACS*
1619   *Biomater Sci Eng* **7**, 3648-3657, doi:10.1021/acsbiomaterials.1c00642
1620   (2021).

1621 64 Meng, B. *et al.* Altered TMPRSS2 usage by SARS-CoV-2 Omicron
1622   impacts tropism and fusogenicity. *Nature*,
1623   doi:10.1038/s41586-022-04474-x (2022).

1624

a Phylogenetic tree showing BA.1, BA.2, BA.2.75, BA.4, BA.5, and BQ.1.1 lineages with key substitutions:
- BA.2.75: K147E, W152R, F157L, I210V, G257S, D339H, G446S, N460K, R493Q
- BA.4: HV69-70del, L452R, F486V, R493Q
- BQ.1.1: R346T, K444T, N460K
- Scale: 0.001

b Frequency heatmaps for substitutions across BA.2, BA.2.75, BA.4, BA.5, BQ.1.1 and BA.5, BQ.1.1:
- HV69-70del, K147E, W152R, F157L, I210V, G257S, G339D/H, R346T, K444T, G446S, L452R, N460K, F486V, Q493R
- NSP2:Q376K, NSP6:L260F, NSP12:Y273H, NSP13:M233I, NSP13:N268S, N:E136D
- Frequency scale: 0 to 1

c Phylogenetic trees for BA.1, BA.2, BA.4, BA.5 showing substitution events (R346, K444, L452, N460, F486) over time (2021/1–2022/6). BA.2.75 and BQ.1.1 highlighted.
- Substitution event legend: R346 (red), K444 (blue), L452 (orange), N460 (purple), F486 (pink)
- Additional substitutions: 3, 2, 1, 0 (line shades)
- Substitution: With (dark blue), Without (gray)

d Substitution events heatmap for BA.1, BA.2, BA.4, BA.5 (R346, K444, L452, N460, F486). Scale 0–40.

e Substitution events per 1M sequences heatmap for BA.1, BA.2, BA.4, BA.5. Scale 0–60.

f Plot of 107 substitutions in S versus effect on R_e (0.7–1.2):
- L452R:F486V, L452Q, N460K, K444M
- R346T, L452M, R346S, R346I, K444T
- Cluster of BA.1-specific substitutions, Q493R
- Positive (red), Negative (blue), Not significant (gray)

g 254 S haplotypes versus relative R_e (colored by BA.1, BA.2, BA.4, BA.5, BA.2.75). BQ.1.1 and BQ.1 highlighted. Columns: R346T, K444T, L452R, N460K, F486V.

h Relative R_e versus total effect of convergent substitutions. $R^2=0.816$. Colored by BA.1, BA.2, BA.4, BA.5, BA.2.75.

i BA.5 phylogenetic trees with substitutions R346T, K444T, N460K over time (2022/1–2023/9). Lineages: BQ.1.1, BQ.1, BE.1.1.1. Fitness predicted by S substitutions scale 0.6–1.6.
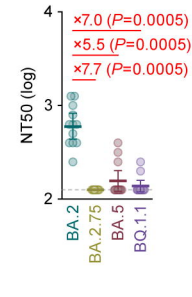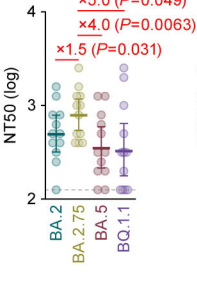
a BA.2 breakthrough infection sera (n=14)
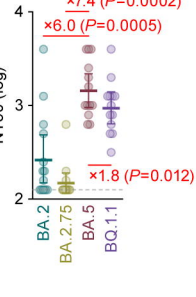**b** BA.5 breakthrough infection sera (n=20)
c BA.2 infection sera (hamster) (n=12)　BA.2.75 infection sera (hamster) (n=12)　BA.5 infection sera (hamster) (n=12)
d Antigenic cartography
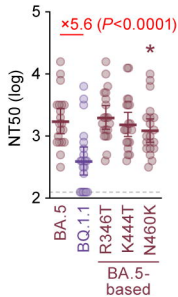
a

% Body Weight — Uninfected, BA.5, BQ.1.1, Delta
- : 2.2E−20
- : 0.040
- : 2.3E−11

Penh
- : 8.1E−6
- : 0.0029
- : 4.2E−8

Rpef
- : 1.3E−8
- : 2.6E−7
- : 1.3E−8

b

Oral swab
- : 0.21
- : 0.020

Lung hilum
- : 0.16
- : 0.16

Lung periphery
- : 0.56
- : 0.56

BA.5, BQ.1.1, Delta

c

BA.5 | BQ.1.1 | Delta

% N-positive cells in the lung at 2 d.p.i. — NS *

2 d.p.i.

BA.5 | BQ.1.1 | Delta

% N-positive cells in the lung at 5 d.p.i. — NS *

5 d.p.i.

d

BA.5 | BQ.1.1 | Delta | BA.5 | BQ.1.1 | Delta | Uninfected

2 d.p.i. — 5 d.p.i.

e

Bronchitis/bronchiolitis
- : 1.0
- : 0.097

Hemorrhage/congestion
- : 0.63
- : 0.037

Alveolar damage
- : 0.45
- : 0.020

Type II pneumocytes
- : 0.20
- : 0.035

Area of large type II pneumocytes
- : 0.51
- : 0.080

Total histology score
- : 1.0
- : 0.0017

BA.5, BQ.1.1, Delta