

 Open access • Posted Content • DOI:10.1101/2020.12.14.422555

Recurrent emergence and transmission of a SARS-CoV-2 Spike deletion Δ H69/V70

— [Source link](#) 

Steven Kemp, Steven Kemp, Bo Meng, Isabella Ferreira ...+14 more authors

Institutions: University College London, University of Cambridge, University of Glasgow, Laboratory of Molecular Biology ...+1 more institutions

Published on: 14 Dec 2020 - bioRxiv (Cold Spring Harbor Laboratory)

Related papers:

- [Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus.](#)
- [Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 \(SARS-CoV-2\) lineage with multiple spike mutations in South Africa](#)
- [Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding.](#)
- [Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants.](#)
- [mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants.](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/recurrent-emergence-and-transmission-of-a-sars-cov-2-spike-4r7mv0ukfv>

1 **Recurrent emergence and transmission of a SARS-CoV-2 spike deletion H69/V70**

2

3 Steven A Kemp^{1,2,3*}, Bo Meng^{1,2*}, Isabella ATM Ferriera^{1,2*}, Rawlings Datir^{1,2*}, William T
4 Harvey^{4*}, Guido Papa⁵, Spyros Lytras⁶, Dami A Collier^{1,2,3}, Ahmed Mohamed⁷, Giulia
5 Gallo⁷, Nazia Thakur⁷, The COVID-19 Genomics UK (COG-UK) Consortium⁸, Alessandro
6 M Carabelli², Julia C Kenyon^{3,9}, Andrew M Lever^{3,10}, Anna De Marco¹¹, Christian Saliba¹¹,
7 Katja Culap¹¹, Elisabetta Cameroni¹¹, Luca Piccoli¹¹, Davide Corti¹¹, Leo C James⁵, Dalan
8 Bailey⁷, David L Robertson^{6*}, Ravindra K. Gupta^{2,3*}

9

10 ¹Cambridge Institute of Therapeutic Immunology & Infectious Disease (CITIID), Cambridge,
11 UK.

12 ²Department of Medicine, University of Cambridge, Cambridge, UK.

13 ³Division of Infection and Immunity, University College London, London, UK.

14 ⁴Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow,
15 Glasgow, UK

16 ⁵MRC – Laboratory of Molecular Biology, Cambridge, UK.

17 ⁶MRC - University of Glasgow Centre for Virus Research, Glasgow, UK.

18 ⁷Pirbright Institute, Woking, Surrey, UK

19 ⁸<https://www.cogconsortium.uk>. Full list of consortium names and affiliations are in
20 Appendix

21 ⁹Department of Microbiology and Immunology, Yong Loo Lin School of Medicine, National
22 University of Singapore, Singapore

23 ¹⁰Department of Medicine, Yong Loo Lin School of Medicine, National University of
24 Singapore, Singapore

25 ¹¹ Humabs Biomed SA, a subsidiary of Vir Biotechnology, 6500 Bellinzona, Switzerland.

26

27

28 *Authors contributed equally to this work

29 Address for correspondence:

30 Ravindra K. Gupta

31 Cambridge Institute for Therapeutic Immunology and Infectious Diseases

32 Jeffrey Cheah Biomedical Centre

33 Puddicombe Way

34 Cambridge CB2 0AW, UK

35 Tel: +44 1223 331491

36 rkg20@cam.ac.uk

37

38 Key words: SARS-CoV-2; COVID-19; antibody escape; neutralising antibodies; infectivity;
39 spike mutation; evasion; resistance; fitness; deletion

40

41 **Highlights**

- 42 • Δ H69/V70 is present in at least 28 SARS-CoV-2 lineages
- 43 • Δ H69/V70 does not confer escape from convalescent sera
- 44 • Δ H69/V70 increases spike infectivity and compensates for RBD mutations
- 45 • Δ H69/V70 is associated with greater spike cleavage
- 46 • B.1.1.7 requires Δ H69/V70 for optimal spike cleavage and infectivity

47

48 **Abstract**

49 SARS-CoV-2 amino acid replacements in the receptor binding domain (RBD) occur
50 relatively frequently and some have a consequence for immune recognition. Here we report
51 recurrent emergence and significant onward transmission of a six-nucleotide out of frame
52 deletion in the S gene, which results in loss of two amino acids: H69 and V70. We report that
53 in human infections Δ H69/V70 often co-occurs with the receptor binding motif amino acid
54 replacements N501Y, N439K and Y453F, and in the latter two cases has followed the RBD
55 mutation. One of the Δ H69/V70+ N501Y lineages, now known as B.1.1.7, has undergone
56 rapid expansion and includes eight S gene mutations: RBD (N501Y and A570D), S1
57 (Δ H69/V70 and Δ Y144) and S2 (P681H, T716I, S982A and D1118H). *In vitro*, we show that
58 Δ H69/V70 does not reduce serum neutralisation across multiple convalescent sera. However,
59 Δ H69/V70 increases infectivity and is associated with increased incorporation of cleaved
60 spike into virions. Δ H69/V70 is able to compensate for small infectivity defects induced by
61 RBD mutations N501Y, N439K and Y453F. In addition, replacement of H69 and V70
62 residues in the B.1.1.7 spike reduces its infectivity and spike mediated cell-cell fusion. Based
63 on our data Δ H69/V70 likely acts as a permissive mutation that allows acquisition of
64 otherwise deleterious immune escape mutations. Enhanced surveillance for the Δ H69/V70
65 deletion with and without RBD mutations should be considered as a global priority not only
66 as a marker for the B.1.1.7 variant, but potentially also for other emerging variants of

67 concern. Vaccines designed to target the deleted spike protein could mitigate against its
68 emergence as increased selective forces from immunity and vaccines increase globally.

69

70

71 **Background**

72 SARS-CoV-2's spike surface glycoprotein engagement of human angiotensin-converting
73 enzyme (hACE2) is essential for virus entry and infection(Zhou et al., 2020), and the receptor
74 is found in respiratory and gastrointestinal tracts(Sungnak et al., 2020). Despite this critical
75 interaction and the constraints it imposes, it appears the RBD, and particularly the receptor
76 binding motif (RBM), is relatively tolerant to mutations(Starr et al., 2020b; Thomson et al.,
77 2020), raising the real possibility of virus escape from vaccine-induced immunity and
78 monoclonal antibody treatments. Spike mutants exhibiting reduced susceptibility to
79 monoclonal antibodies have been identified in *in vitro* screens(Greaney et al., 2020; Starr et
80 al., 2020a), and some of these mutations have been found in clinical isolates(Choi et al.,
81 2020). Due to the high levels of susceptibility of the human population to this virus, the acute
82 nature of infections and limited use of vaccines to date there has been limited selection
83 pressure placed SARS-CoV-2(MacLean et al., 2020); as a consequence few mutations that
84 could alter antigenicity have increased significantly in frequency.

85

86 The unprecedented scale of whole genome SARS-CoV-2 sequencing has enabled
87 identification and epidemiological analysis of transmission and surveillance, particularly in
88 the United Kingdom. As of February 16th there were 544,778 SARS-CoV-2 sequences
89 available in the GISAID database (<https://www.gisaid.org/>). However, geographic coverage
90 is very uneven with two-fifths of all sequences being provided by the United Kingdom.
91 Indeed, as more countries continue to present data from samples collected up to six months
92 prior, this may result in novel variants with altered biological or antigenic properties evolving
93 and not being detected until they have already become established at high frequency.

94

95 Studying SARS-CoV-2 chronic infections can give insight into virus evolution that would
96 require many chains of acute transmission to generate. This is because the majority of
97 infections arise as a result of early transmission during pre or asymptomatic phases prior to
98 peak adaptive responses, and virus adaptation not observed as the virus is usually cleared by
99 the immune response(He et al., 2020; Mlcochova et al., 2020c). We recently documented *de*
100 *novo* emergence of antibody evasion mutations mediated by S gene mutations in an

101 individual treated with convalescent plasma (CP)(Kemp et al., 2021). Dramatic changes in
102 the prevalence of Spike variants Δ H69/V70 (an out of frame six-nucleotide deletion) and
103 D796H variant followed repeated use of CP, while *in vitro* the mutated Δ H69/V70+D796H
104 variant displayed reduced susceptibility to CP using a pseudotyped lentivirus assay. D796H
105 alone resulted in more than five-fold loss of infectivity and the Δ H69/V70 partially rescued
106 this defect. In addition, a chronically infected immune suppressed individual was recently
107 reported in Russia with emergence of Y453F, along with Δ H69/V70(Bazykin et al., 2021).
108 Deletions in other parts of the N-Terminal Domain (NTD) have been reported to arise in
109 chronic infections(Choi et al., 2020) and to reduce sensitivity to NTD-specific neutralising
110 antibodies(McCallum et al., 2021; McCarthy et al., 2021).

111

112 Here we analyse the variation in the global SARS-CoV-2 data and find Δ H69/V70 occurs
113 independently often emerging after a significant RBD amino acid replacement such as
114 N501Y, Y453F and N439K, that are known to either facilitate neutralising antibody escape or
115 modulate affinity for the human ACE2 receptor. We show that Δ H69/V70 and other common
116 NTD deletions occur at loop structures in RNA where polymerase activity is often
117 compromised. Although structural modelling indicates the H69/V70 is in an exposed loop
118 that contracts post deletion, potentially altering an antigenic site, we report that the
119 Δ H69/V70 does not confer reduced susceptibility to convalescent sera. Functionally, we
120 reveal that Δ H69/V70 does increase Spike infectivity and compensates for an infectivity
121 defect resulting from RBD replacements N501Y, N439K and Y453F. The infectivity increase
122 is associated with higher levels of cleaved S in pseudotyped virions. Finally, the deletion is
123 required for optimal infectivity of the 501Y.V1 (B.1.1.7) spike protein and repair of the two
124 amino acids leads to reduced S incorporation into virions. These data support a role for
125 Δ H69/V70 in promoting virus infectivity to balance deleterious escape mutations and should
126 be mitigated against.

127

128 **Results**

129 **Multiple occurrences and transmission of Spike Δ H69/V70 with and without S mutations**

130 The deletion H69/V70 is present in over 87,000 sequences worldwide (71,500 from the UK),
131 and has seen global expansion, particularly across much of Europe, Africa and Asia (Figure
132 1). Δ H69/V70 is observed in at least 28 different global lineages based on PANGO
133 classification, though not all represent multiple independent acquisitions (Figure 1A,

134 Supplementary Table 1). While variants with deletions in this region of Spike are observed in
135 GISAID¹³, the earliest unambiguous sequences that include the Δ H69/V70 were detected on
136 a D614 background in January 2020 (USA and Thailand). The earliest Δ H69/V70 detected on
137 a D614G background was in Sweden in April 2020. The prevalence of Δ H69/V70 has since
138 increased since August 2020 (Figure 1B). Further analysis of sequences revealed, firstly, that
139 single deletions of either 69 or 70 were uncommon and, secondly, some lineages of
140 Δ H69/V70 alone were present, as well as Δ H69/V70 in the context of other mutations in
141 Spike, specifically those in the RBM (Figure 1).

142

143 We next examined in greater detail the SARS-CoV-2 lineages where S gene mutations in the
144 RBD were identified at high frequency and where Δ H69/V70 co-occurs. For example,
145 N439K, an amino acid replacement reported to be defining variants increasing in numbers in
146 Europe and other regions(Thomson et al., 2020) (Figure 1, Supplementary Figure 1) now
147 mostly co-occurs with Δ H69/V70. N439K appears to have reduced susceptibility to some
148 convalescent sera(Hoffmann et al., 2021) as well as monoclonals targeting the RBD, whilst
149 retaining affinity for ACE2 *in vitro*(Thomson et al., 2020). The first lineage possessing
150 N439K (and not Δ H69/V70), B.1.141 is now extinct(Thomson et al., 2020). A second lineage
151 with N439K, B.1.258, later emerged and subsequently acquired Δ H69/V70 leading to the
152 initial rapid increase in the frequency of viruses possessing this deletion, spreading into
153 (Brejová et al., 2021).

154

155 The second significant cluster with Δ H69/V70 and RBD mutants involves Y453F, another
156 Spike RBD mutation that increases binding affinity to ACE2(Starr et al., 2020b) and has been
157 found to be associated with mink-human infections(Munnink et al., 2020). In one SARS-
158 CoV-2 mink-human sub-lineage, termed 'Cluster 5', Y453F and Δ H69/V70 occurred with
159 F486L, N501T and M1229I(Larsen et al., 2021) and was shown to have reduced
160 susceptibility to sera from recovered COVID-19 patients (https://files.ssi.dk/Mink-cluster-5-short-report_AFO2). Y453F has also been described as an escape mutation for mAb
162 REGN10933(Baum et al., 2020; Hoffmann et al., 2021). The Δ H69/V70 was first detected in
163 the Y453F background on August 24th 2020 and thus far appears limited to Danish sequences
164 (Figure 1, Supplementary Figure 2), although an independent acquisition was recently
165 reported within an individual, along with Δ H69/V70 in an immune compromised individual
166 with chronic infection(Bazykin et al., 2021).

167

168 A third lineage containing the same Δ H69/V70 deletion has arisen with another RBD
169 mutation N501Y along with multiple other spike mutations (Figure 1, Supplementary Figure
170 3). These UK sequences were subsequently named as belonging a new lineage (B.1.1.7),
171 termed a variant of concern, VOC 202012/01, by Public Health England as it is associated
172 with faster rates of spread (Volz et al., 2021) (Supplementary Figure 3). In addition to RBD
173 N501Y + NTD Δ H69/V70 this new variant had five further S mutations across the NTD
174 (A570D), S2 (P681H, T716I, S982A and D1118H), and NTD Δ Y144(Rambaut A., 2020).
175 The variant has now been identified in 94 countries. This lineage has a relatively long branch
176 due to 23 mutations (Supplementary figure 3). The available sequences did not enable us to
177 determine whether the B.1.1.7 mutations N501Y + Δ H69/V70 arose as a result of a N501Y
178 virus acquiring Δ H69/V70 or vice versa, though a sequence was identified with N501Y,
179 A570D, Δ H69/V70 and D1118H, indicating that as predicted N501Y + Δ H69/V70 are
180 proximal mutational events potentially in a long term shedding patient(Kemp et al., 2021).
181 N501Y does escape some RBM targeting antibodies as well as binding ACE2 with higher
182 affinity(Collier et al., 2021a). Sequences with N501Y alone were isolated both in the UK,
183 Brazil, and USA in April 2020, as well as in South Africa.

184

185 **Δ H69/V70 and other deletions arise at the terminal end of loop RNA structures**

186 Polymerase processivity can be affected by physical factors in the template caused by RNA
187 structural and sequence motifs and these can facilitate dissociation events. Stable helix loop
188 motifs are associated with pausing/dissociation events in reverse transcriptase(Harrison et al.,
189 1998) . Since all nucleic acid polymerases have a common ancestor with homologous dNTP
190 binding motifs and similar global structures(Delarue et al., 1990; Ollis et al., 1985; Sousa et
191 al., 1993) it is probable that all RNA polymerases use similar mechanisms for transcript
192 termination(Reeder and Lang, 1994). Analysis of three deletions in the S protein coding
193 sequence observed across new multiply mutated variants demonstrated that each occurred in
194 the terminal loop of a helix loop motif (Supplementary Figure 4). Although many regions of
195 the structure have a wider structural ensemble of structures that the RNA can adopt, these
196 helix loops were ones in which the ensemble is constrained into a very limited structural
197 range, suggesting that these deletions occur in the more stable stem-loops. A recent in-cell
198 biochemical analysis of SARS-CoV2 RNA structure showed nucleotide reactivity consistent
199 with our model within these stem-loops (Huston et al., 2021). These analyses provide a

200 rationale for preferential emergence of Δ H69/V70 and other deletions such as the well
201 described NTD-antibody escape deletion Δ Y144 (Chi et al., 2020; McCarthy et al., 2020,
202 2021) (in B.1.1.7 and the recently reported B.1.525) at the terminal loops of helical loop
203 motifs.

204

205

206 **Δ H69/V70 does not confer reduced susceptibility to convalescent sera**

207 We hypothesised that Δ H69/V70 is an antibody escape mechanism. We first examined the
208 protein structural context of Δ H69/V70 for clues regarding alterations in epitopes (Figure 2A,
209 B). In the absence of experimentally derived structural data for Δ H69/V70, the protein
210 structure of the NTD possessing the double deletion was modelled in silico. The Δ H69/V70
211 deletion was predicted to alter the conformation of a protruding loop comprising residues 69
212 to 76, pulling it in towards the NTD (Figure 2B). In the post-deletion structural model, the
213 positions of the alpha carbons of residues either side of the deleted residues, Ile68 and Ser71,
214 were each predicted to occupy positions 2.9Å from the positions of His69 and Val70 in the
215 pre-deletion structure. Concurrently, the positions of Ser71, Gly72, Thr73, Asn74 and Gly75
216 are predicted to have changed by 6.5Å, 6.7Å, 6.0Å, 6.2Å and 8Å, respectively, with the
217 overall effect of these residues moving inwards, resulting in a less dramatically protruding
218 loop.

219

220 This predicted change in the surface of Spike could be consistent with antibody evasion. To
221 test this we explored whether Δ H69/V70 conferred reduced susceptibility to neutralising
222 antibodies in sera from fifteen recovered individuals (Figure 2C, D). We performed serial
223 dilutions of sera before mixing with viral particles pseudotyped with Spike proteins with and
224 without Δ H69/V70 (with virus input normalised for infectivity). We plotted infection of
225 target cells as a function of serum dilution across the fifteen serum samples (Figure 2D,
226 Supplementary Figure 5). All but two sera demonstrated clear titratable neutralisation of both
227 wild type and Δ H69/V70 virus. There was no overall change in susceptibility to serum
228 neutralisation for Δ H69/V70 relative to wild type (Figure 2C), but there were a proportion of
229 individuals with increase in neutralisation to Δ H69/V70 with small shift in the titration
230 curves (Figure 2C, D, Supplementary Figure 5). In these cases the Δ H69/V70 was associated
231 with increased susceptibility to sera. These data suggest that Δ H69/V70 does not represent an
232 important antibody escape mechanism.

233

234

235

236 **Δ H69/V70 spike infectivity correlates with increased cleaved spike in virions**

237 Given the association between Δ H69/V70 and RBD mutations and data from an earlier report
238 in chronic infection¹², we hypothesised that this deletion might alternatively enhance virus
239 infectivity. In the absence of virus isolates we used a lentiviral pseudotyping approach to test
240 the impact of Δ H69/V70 on virus Spike protein mediated infection. A D614G bearing Spike
241 protein expressing DNA plasmid (wild type, WT) was co-transfected in HEK 293T producer
242 cell lines along with plasmids encoding lentiviral capsid and genome for luciferase.
243 Infectivity was adjusted for input reverse transcriptase activity; we observed a two-fold
244 increase in infectivity of Δ H69/V70 as compared to WT (Figure 2E). Western blotting
245 indicated that for Δ H69/V70 there was a higher amount of cleaved spike in the virion
246 containing cell supernatants (unspun) and in the producer cell lysates. It appeared that there
247 was a corresponding reduction in uncleaved spike (Figure 2E). Densitometric analysis of
248 spike and p24 from western blots in multiple experiments showed almost a two-fold increase
249 in spike:p24 ratio for the Δ H69/V70, indicating increased spike cleavage efficiency and
250 incorporation into virions that could explain the increase in infectivity. In order to explore
251 whether D614G was required for this enhanced cleavage efficiency and infectivity, we
252 generated pseudotyped virus bearing D614 spike with and without the Δ H69/V70 and
253 infected target cells. We observed a similar enhancement of infection and increase in cleaved
254 spike in supernatants as we did for D614G spike pseudotyped viruses (Supplementary Figure
255 6).

256

257 **Spike Δ H69/V70 compensates for reduced spike infectivity due to RBD replacements**

258 We hypothesised that Δ H69/V70 might compensate for potential loss of infectivity due to
259 receptor binding motif mutations Y453F, N510Y and N439K that interact with ACE2 (Figure
260 3A-C). We generated mutant spike plasmids bearing RBD mutations N501Y, Y453F, or
261 N439K both with and without Δ H69/V70 (Figure 3A-C) and performed infectivity assays in
262 the lentiviral pseudotyping system. RBD mutations reduced infectivity of Spike relative to
263 WT by 2-3 fold (Figure 3D, E). Based on observations of the impact of Δ H69/V70 on S
264 incorporation, we predicted that the mechanism of increased infectivity for Δ H69/V70 in
265 context of RBD mutations might also involve increased incorporation of cleaved S2 spike as

266 in Figure 2E. The analysis on virions from cell supernatants and cell lysates indeed showed
267 increased cleaved S2 spike when Δ H69/V70 was present with RBD mutants (Figure 3E, F).

268

269 In order to explore the mechanism of increased infectivity, we used a cell fusion assay
270 (Figure 3G, Supplementary Figure 7) to monitor kinetics of cell fusion. Previous reports have
271 shown that SARS-CoV-2 Spike protein possesses high fusogenic activity and is able to
272 trigger the formation of large multi-nucleated cells (named syncytia) *in vitro* and *in vivo*
273 (Papa et al., 2020) (Cattin-Ortola et al. 2020), potentially providing an additional and a more
274 rapid route for virus disseminating among neighbour cells. To understand whether the higher
275 infectivity of Δ H69/V70 spike might be related to an increased ability to trigger cell-cell
276 fusion between different cell types, we ectopically overexpressed SARS-CoV-2 spike
277 variants together with the mCherry fluorescent protein in 293T cells and label Vero cells with
278 a green fluorescent dye. Mixing both cell types and measuring the merged green and red
279 fluorescence allows precise quantification of cell-cell fusion kinetics (Supplementary Figure
280 7). Syncytia formation kinetics appeared slightly lower for N501Y SARS-CoV-2 Spike
281 mutant compared to wild type, Δ H69/V70 and N501Y along with Δ H69/V70 (Figure 3G).

282

283 **B.1.1.7 Spike cleavage efficiency, virion incorporation and infectivity is reduced by re-** 284 **insertion of H69V70**

285 B.1.1.7 naturally contains the Δ H69/V70 deletion (Figure 4A). We predicted that the
286 replacement of H69 and V70 would impair the infectivity of B.1.1.7. To examine this, we
287 compared the infectivity of B.1.1.7 Spike versus B.1.1.7 without the Δ H69/V70 deletion in
288 our pseudotyping system. We observed that infectivity of B.1.1.7 and WT pseudotyped
289 viruses were similar (Figure 4B). As expected, we observed a significant reduction in
290 infectivity for viruses where the H69 and V70 had been re-inserted (Figure 4B, C). When we
291 tested Spike cleavage and incorporation into virions we found that the reduced infectivity of
292 the B.1.1.7 with replaced H69 V70 was associated with reduced cleaved spike S2 protein
293 (Figure 4C, D).

294

295 In order to ascertain whether H69V70 represented a target for neutralising antibodies in the
296 context of B.1.1.7, we tested 13 NTD-specific mAbs isolated from 4 individuals that
297 recovered from WT SARS-CoV-2 infection with an *in-vitro* pseudotyped neutralization assay
298 using VeroE6 target cells expressing Transmembrane protease serine 2 (TMPRSS2,

299 **Supplementary Table 1**). The pseudotyped viruses expressed the WT SARS-CoV-2 S, the
300 B.1.1.7 S or the B.1.1.7 S with reversion of H69/V70 deletions (B.1.1.7 H69/V70). We found
301 that 8 out of 13 (62%) showed a marked decrease or complete loss of neutralizing activity to
302 both B.1.1.7 and B.1.1.7 H69/V70 (>30 fold-change reduction), suggesting that in a sizeable
303 fraction of NTD antibodies the H69/V70 deletion is not responsible for their loss of
304 neutralizing activity (**Supplementary Figure 8**). Only 3 mAbs showed a partial reduction (3-
305 to-8 fold) in B.1.1.7 neutralization that was rescued or improved in one case by reversion of
306 H69/V70 deletions. Neutralization of 2 mAbs (15%) was not affected by both deletion and
307 reversion of H69/V70 in B.1.1.7.

308

309

310 ***ΔH69/V70 in sarbecoviruses closely related to SARS-CoV-2***

311 Finally, to investigate the importance of this part of spike beyond SARS-CoV-2, we
312 examined the 69/70 region of spike in a set of other known *Sarbecoviruses* (Figure 5). We
313 observed substantial variability in the region, resulting in frequent indels, with some viruses
314 including SARS-CoV having 6-7 amino acid deletions (Figure 5B). This is indicative of
315 plasticity in this protein region that could allow the *sarbecoviruses* to alter their Spike
316 conformation. The second closest relative to SARS-CoV-2 for this region after RaTG13 is
317 the cluster of 5 CoVs sampled in trafficked pangolins in the Guangxi province (Lam et al.,
318 2020). Inspection of the 69/70 region in these virus sequences raises the interesting
319 observation that one of the five viruses in the cluster, P2V, has amino acids 69H and 70L
320 present, while the other four have a double amino acid deletion (Figure 5B). Given that
321 SARS-CoV-2 and RaTG13 have the homologous HV insertion at these positions, one
322 explanation is that the proximal common ancestor between SARS-CoV-2 and the Guangxi
323 pangolin cluster had the insertion, which was then lost while circulating in the pangolin
324 population, similar to observations with SARS-CoV-2 in humans. Yet, the fact that P2V was
325 cultured in Vero E6 cells prior to sequencing (contrary to the other 4, sequenced directly
326 from the pangolin sample) raises the possibility of this being an independent insertion,
327 favoured as a cell line specific adaptation. Interestingly, the double amino acid indel in the
328 pangolin viruses is in-frame in contrast to SARS-CoV-2 (e.g. lineage B.1.1.7, Figure 5C).

329

330 Furthermore, the two almost identical bat viruses recently sequenced from Cambodia samples
331 – RShSTT182 and RShSTT200 (Hul et al., 2021) possess an H69V70 insertion despite being
332 more distantly related to SARS-CoV-2 for this region of Spike (Figure 5). This independent

333 occurrence of the insertion is suggestive of conditional selective pressures playing a role in
334 recurring gain and loss of these two residues in the sarbecoviruses. To test whether the
335 beneficial effect of Δ H69/V70 is specific to SARS-CoV-2 and not other *Sarbecovirus* Spike
336 backgrounds, we cloned full length S from RaTG13 with a C terminal FLAG tag and
337 generated pseudotyped lentiviruses expressing RaTG13 Spike protein as well as a RaTG13 S
338 Δ H69V70 counterpart. Interestingly, we observed that cleaved and uncleaved S expression
339 levels in unpelleted supernatants containing pseudotyped virus did not differ between WT
340 and the Δ H69V70 RaTG13 spike, and that there was no difference in infectivity. This
341 indicates that the enhancing effect of Δ H69V70 in SARS-CoV-2 may be virus background
342 specific.

343

344

345 **Discussion.**

346 We have presented data demonstrating multiple, independent, and circulating lineages of
347 SARS-CoV-2 variants bearing spike Δ H69/V70. This recurring deletion spanning six
348 nucleotides is due to an out of frame deletion of six nucleotides, and occurs in the terminal
349 loop of a helix loop motif within the RNA structure, as do other NTD deletions observed in
350 new variants such as the UK B.1.1.7, South African B.1.351, Brazilian P.1. Δ H69/V70 has
351 frequently followed receptor binding amino acid replacements (N501Y, N439K and Y453F
352 that have been shown to increase binding affinity to hACE2 and reduce binding with
353 monoclonal antibodies), and is specifically found in the B.1.1.7 variant known to have higher
354 transmissibility (Volz et al., 2020)

355

356 The Δ H69/V70 deletion was recently shown to emerge following treatment with
357 convalescent plasma¹². In a second case of multiple mutations in context of immune
358 suppression the Δ H69/V70 deletion occurred with Y453F, though without biological
359 characterisation. We show here experimentally that the Δ H69/V70 deletion is indeed able
360 increase infectivity of viruses bearing RBD mutations N501Y, N439K and Y453F,
361 potentially explaining why the deletion is often observed after these RBD mutation in SARS-
362 CoV-2 global phylogenies. We show that the mechanism of enhanced infectivity across the
363 RBD mutations tested is associated with greater spike cleavage and incorporation of cleaved
364 spike into virions where the Δ H69/V70 deletion is present. Importantly, we were able to
365 recapitulate the Δ H69/V70 phenotype in a spike protein that did not have the D614G

366 mutation, indicating that D614G is not involved in the mechanism of S incorporation
367 enhancement. These observations are supported by Δ H69/V70 being observed in D614
368 viruses in Jan 2020 both in the US and Thailand.

369

370 Critically, we show that whilst B.1.1.7 spike has similar infectivity as wild type D614G
371 spike, there is substantial loss of infectivity when the H69/V70 amino acids are replaced, and
372 this is accompanied by reduced S incorporation into virions. These data point to co-evolution
373 of the observed mutations in spike of B.1.1.7, with a balance of mutations that incur fitness
374 cost with those that aid immune evasion. The findings suggest that the Δ H69/V70 may
375 contribute to the higher viral loads and increased transmissibility of B.1.1.7. Taken together
376 these data support a model whereby Δ H69/V70 can act as a ‘permissive’ mutation that
377 enhances infection, with the potential to enhance the ability of SARS-CoV-2 to tolerate
378 immune escape mutations that would have otherwise significantly reduced viral fitness.

379

380 Of note, whilst we showed that RBD mutations reduced infectivity in a pseudotyping system,
381 this reduced infectivity may not translate directly to vivo replication where ACE2 levels on
382 target airway cells are lower. For example the higher affinity of N501Y for ACE2 may be
383 selected for by low ACE2 levels, but incur a fitness defect in higher ACE2 environments.
384 Alternatively, increased ACE2 affinity afforded by 501Y may confer a selective advantage
385 under immune pressure, such as exposure to treatment by convalescent plasma, but require
386 compensatory mutations to transmit optimally among naïve hosts. In the successful lineages
387 B.1.351 and P.1, which also carry 501Y, it is notable that 501Y is accompanied by RBD
388 substitutions K417N and K417T respectively, each of which reduce ACE2 affinity (Starr et
389 al. 2020); potentially acting as alternative compensatory mutations for 501Y.

390

391 However, we do believe that the effect of Δ H69/V70 is robust; firstly, our in vitro work
392 involving removal and insertion of the H69/V70 residues significantly impacts both cleaved
393 spike protein levels in virions and infectivity; secondly, the Δ H69/V70 has emerged and
394 transmitted as a single mutation and as a co-mutation in viruses that have previously acquired
395 an RBD mutation such as Y453F, R439K and N501Y, arguing for a role in transmissibility.

396

397 The potential for SARS-CoV-2 to evolve and fix mutations is exemplified by D614G, an
398 amino acid replacement in S1 that alters linkages between S1 and S2 subunits on adjacent

399 protomers as well as RBD orientation, infectivity, and transmission(Hou et al., 2020; Korber
400 et al., 2020; Yurkovetskiy et al., 2020). While D614G confers a transmissibility advantage,
401 the resulting increased tendency for the open conformation of spike results in increased
402 susceptibility to neutralisation by RBD-binding monoclonal antibodies (Weissmann et al.,
403 2020). The example of D614G also demonstrates that mechanisms directly impacting
404 important biological processes can be indirect. Similarly, a number of possible mechanistic
405 explanations may underlie Δ H69/V70. For example, the fact that it sits on an exposed surface
406 and is estimated to alter the conformation of a particularly exposed loop might be suggestive
407 of immune interactions and escape, though we have presented data to show that Δ H69/V70
408 did not reduce sensitivity of spike to neutralising antibodies in serum from a group of
409 recovered individuals. Indeed in some sera, susceptibility to Δ H69/V70 was increased
410 relative to wild type, raising the hypothesis that, similar to D614G, Δ H69/V70
411 simultaneously increases infectivity and increases susceptibility to neutralising antibodies by
412 a conformational change favouring a more open spike conformation. Consistent with these
413 data, Xie et al recently reported increased susceptibility of N501Y + Δ H69/V70 full length
414 viruses(Xie et al., 2021) in a proportion of post vaccine sera, and our recent work on B.1.1.7
415 spike showed increased susceptibility of a triple mutant bearing N501Y + A570D +
416 Δ H69/V70 to both vaccine and convalescent sera(Collier et al., 2021b).

417

418 Allosteric interactions, as postulated for D614G, could lead to the higher cleavage efficiency
419 and infectivity of spike pseudotyped viruses with Δ H69/V70. It is also possible that the
420 increased infectivity of relates to interactions with receptors other than ACE2 or TMPRSS2,
421 for example L- SIGN/ DC- SIGN. (Soh et al., 2020)Notably, data on increased infectivity
422 conferred by D614G using similar pseudotyped viruses(Yurkovetskiy et al., 2020) was also
423 observed in whole virus and translated to increased viral load and transmission in animal
424 models(Hou et al., 2020). It would be important to investigate whether Δ H69/V70 is
425 associated with higher viral loads or increased transmission, though such epidemiological
426 studies are highly complex and prone to confounding/bias.

427

428 The finding of a lineage (B.1.1.7), termed VOC 202012/01, bearing seven S gene mutations
429 across the RBD (N501Y, A570D), S1 (Δ H69/V70 and Δ I144) and S2 (P681H, T716I, S982A
430 and D1118H) in UK requires further experimental characterisation. The detection of a high
431 number of novel mutations suggests this lineage has either been introduced from a

432 geographic region with very poor sampling or viral evolution may have occurred in a single
433 individual in the context of a chronic infection(Kemp et al., 2020). The emergence of variants
434 with higher numbers of mutations so far in the UK and South Africa may herald an era of re-
435 infection and threaten future vaccine efficacy if left unchecked. Perhaps of greatest concern is
436 the emergence of E484K on the background of B.1.1.7. B.1.1.7 possesses a clear
437 transmissibility advantage(Volz et al., 2021) and considerably changes antigenic profile
438 across spike with Δ Y144 escaping a range of NTD-targeting neutralising antibodies
439 (McCarthy et al. 2021, McCallum et al. 2021) and E484K which escapes a range of RBD-
440 targeting neutralising antibodies (Baum et al. 2020, Liu et al. 2021, Greaney et al. 2021).

441

442 The presence of sequence at site 69/70 appears to be unique to SARS-CoV-2, the closest bat
443 sarbecovirus, RaTG13, recently identified bat viruses in Cambodia RShSTT182 and
444 RShSTT200, and one of the Guangxi pangolin sequences. Although we cannot delineate
445 between the indel being lost in the pangolin host, or gained after culturing the virus, both
446 scenarios are suggestive of functional importance at this protein region. In contrast to SARS-
447 CoV-2, deletion of H69/V70 in the bat *sarbecovirus* RaTG13 did not impact S incorporation
448 or infectivity. This could be explained by the lack of a polybasic cleavage site in these
449 *sarbecoviruses*, virus genetic backgrounds, or cellular factors not recapitulated in human cell
450 lines .

451

452 The detection and surveillance of B.1.1.7 has been facilitated in the UK by the phenomenon
453 of SGTF (S gene target failure) due to primers in the Thermofisher SARS-CoV-2 diagnostic
454 qPCR assay used by a significant number of testing facilities. The S gene target (binding in
455 the region of H69/V70) is one of three and therefore a marker for the spread of B.1.1.7 has
456 been tracked by the loss of signal in the S gene target(Volz et al., 2021). However recent
457 reports from the US and central Europe caution against use of SGTF as a sole marker for
458 B.1.1.7 as a significant Δ H69/V70 lineage without other mutations in spike is circulating in
459 the US, and a B.1.258 lineage with N439K with Δ H69/V70, circulating in Slovakia/Czech
460 republic(Brejová et al., 2021; Larsen and Worobey, 2020). Such examples highlight the need
461 for sequencing to accompany novel approaches to diagnostics for variants.

462

463 Given the emergence of multiple clusters of variants carrying RBD mutations and the
464 Δ H69/V70 deletion, limitation of transmission takes on a renewed urgency. Continued

465 emphasis on testing/tracing, social distancing and mask wearing are essential, with
466 investment in other novel methods to limit transmission (Mlcochova et al., 2020a). Detection
467 of the deletion and other key mutations by rapid diagnostics should be a research priority as
468 such tests could be used as a proxy for antibody escape mutations to inform surveillance at
469 global scale. Finally, comprehensive vaccination efforts should be accelerated in order to
470 further limit transmission and acquisition of further mutations, and future vaccines could
471 include Δ H69/V70 in order to close this route for virus evolution, assuming that effective
472 neutralising antibodies to this region are generated.

473

474 **Limitations**

475 The laboratory virology aspects of this study were conducted with pseudoviruses rather than
476 replication competent viruses. We also carried out experiments in cells overexpressing
477 receptors.

478

479 **Acknowledgements**

480 RKG is supported by a Wellcome Trust Senior Fellowship in Clinical Science
481 (WT108082AIA). COG-UK is supported by funding from the Medical Research Council
482 (MRC) part of UK Research & Innovation (UKRI), the National Institute of Health Research
483 (NIHR) and Genome Research Limited, operating as the Wellcome Sanger Institute. This
484 study was supported by the Cambridge NIHRB Biomedical Research Centre. SAK is
485 supported by the Bill and Melinda Gates Foundation via PANGEA grant: OPP1175094. DLR
486 is funded by the MRC (MC_UU_1201412). WH is funded by the MRC (MR/R024758/1). We
487 thank Dr James Voss for the kind gift of HeLa cells stably expressing ACE2. SL is funded by
488 Medical Research Council MC_UU_12014/12. This study was also partly funded by
489 Rosetrees Trust. AML is funded by the Cambridge NIHRB Biomedical Research Centre.

490

491 **Conflicts of interest**

492 A.D.M., C.S., K.C., E.C., L.P. and D.C. are employees of Vir Biotechnology and may hold
493 shares in Vir Biotechnology. RKG has received consulting fees from UMOVIS lab, Gilead
494 Sciences and ViiV Healthcare, and a research grant from InvisiSmart Technologies.

495

496 Author contributions: Designed study and experiments: R.K.G, B.M., R.D., I.A.TM.F,
497 D.L.R, L.C.J., D.B., L.P., A.D.M, D.C. Designed and performed structural analysis: W.T.H,
498 A.M.C. Performed experiments: S.A.K, B.M., A.D.M. Interpreted data: L.P., A.D.M, D.C.,

499 B.M., R.K.G, D.L.R, D.B, R.D., I.A.TM.F, S.A.K. Carried out pseudovirus neutralization
500 assays: A.D.M. Produced pseudoviruses: C.S. Sequencing and expression of antibodies,
501 mutagenesis for mutant expression plasmids: E.C. and K.C. Analysis of the data and
502 manuscript preparation: L.P., D.C., R.D., I.A.TM.F, R.K.G, B.M., S.A.K.

503

504

505 **Methods**

506 *Phylogenetic Analysis*

507 All available full-genome SARS-CoV-2 sequences were downloaded from the GISAID
508 database (<http://gisaid.org/>)(Shu and McCauley, 2017) on 16th February 2021. Low-quality
509 sequences (>5% N regions) were removed, leaving a dataset of 491,395 sequences with a
510 length of >29,000bp. Sequences were deduplicated and then filtered to find the mutations of
511 interest. All sequences were realigned to the SARS-CoV-2 reference strain MN908947.3,
512 using MAFFT v7.475 with automatic strategy selection and the --keeplength --addfragments
513 options(Katoh and Standley, 2013). Major SARS-CoV-2 clade memberships were assigned
514 to all sequences using the Nextclade server v0.13 (<https://clades.nextstrain.org/>), Pangolin
515 v2.2.2(Rambaut et al., 2020) (github.com/cov-lineages/pangolin) and a local instance of the
516 PangoLEARN model, dated 17th Feb 2021 21:49 ([https://github.com/cov-](https://github.com/cov-lineages/pangoLEARN)
517 [lineages/pangoLEARN](https://github.com/cov-lineages/pangoLEARN)).

518

519 Maximum likelihood phylogenetic trees were produced using the above curated dataset using
520 IQ-TREE v2.1.2(Minh et al., 2020). Evolutionary model selection for trees were inferred
521 using ModelFinder(Kalyaanamoorthy et al., 2017) and trees were estimated using the
522 GTR+F+I model with 1000 ultrafast bootstrap replicates(Minh et al., 2013). All trees were
523 visualised with Figtree v.1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>) and ggtree v1.14.6
524 rooted on the SARS-CoV-2 reference sequence and nodes arranged in descending order.
525 Nodes with bootstraps values of <50 were collapsed using an in-house script.

526

527 To reconstruct a phylogeny for the 69/70 spike region of the 20 *Sarbecoviruses* examined in
528 Figure 5, Rdp5(Martin et al., 2015) was used on the codon spike alignment to determine the
529 region between amino acids 1 and 256 as putatively non-recombinant. A tree was
530 reconstructed using the nucleotide alignment of this region under a GTR+ Γ substitution
531 model with RAxML-NG(Kozlov et al., 2019). Node support was calculated with 1000
532 bootstraps. Alignment visualisation was done using BioEdit(Hall et al., 2011).

533

534 *RNA secondary structure modelling.*

535 2990 nucleotides centred around the spike protein amino acids 69-70 from SARS-CoV2
536 sequence from an individual¹² were aligned in CLUSATL-Omega (nucleotides 20277-23265
537 of the Wuhan isolate MN908947.3) and a consensus structure was generated using
538 RNAalifold(Bernhart et al., 2008)).

539

540 *Structural modelling*

541 The structure of the post-deletion NTD (residues 14-306) was modelled using I-
542 TASSER(Roy et al., 2010), a method involving detection of templates from the protein data
543 bank, fragment structure assembly using replica-exchange Monte Carlo simulation and
544 atomic-level refinement of structure using a fragment-guided molecular dynamics simulation.
545 The structural model generated was aligned with the spike structure possessing the pre-
546 deletion conformation of the 69-77 loop (PDB 7C2L(Chi et al., 2020)) using PyMOL
547 (Schrödinger). Figures prepared with PyMOL using PDBs 7C2L, 6M0J(Lan et al., 2020),
548 6ZGE28 and 6ZGG(Wrobel et al., 2020).

549

550 *Cells*

551 HEK 293T CRL-3216, Vero CCL-81 were purchased from ATCC and maintained in
552 Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal calf serum
553 (FCS), 100 U/ml penicillin, and 100mg/ml streptomycin. All cells are regularly tested and are
554 mycoplasma free.

555

556 *Pseudotype virus preparation*

557 Plasmids encoding the spike protein of SARS-CoV-2 D614 with a C terminal 19 amino acid
558 deletion with D614G, were used as a template to produce variants lacking amino acids at
559 position H69 and V70, as well as mutations N439K, Y453F and N501Y. Mutations were
560 introduced using Quickchange Lightning Site-Directed Mutagenesis kit (Agilent) following
561 the manufacturer's instructions. B.1.1.7 S expressing plasmid preparation was described
562 previously, but in brief was generated by step wise mutagenesis. Viral vectors were prepared
563 by transfection of 293T cells by using Fugene HD transfection reagent (Promega). 293T cells
564 were transfected with a mixture of 11ul of Fugene HD, 1µg of pCDNAΔ19 spike-HA, 1ug of
565 p8.91 HIV-1 gag-pol expression vector and 1.5µg of pCSFLW (expressing the firefly
566 luciferase reporter gene with the HIV-1 packaging signal). Viral supernatant was collected at

567 48 and 72h after transfection, filtered through 0.45um filter and stored at -80°C as previously
568 described. Infectivity was measured by luciferase detection in target 293T cells transfected
569 with TMPRSS2 and ACE2.

570

571 *SARS-CoV-2 D614 (Wuhan) and RaTG13 mutant plasmids and infectivity*

572 Plasmids encoding the full-length spike protein of SARS-CoV-2 D614 (Wuhan) and
573 RaTG13, in frame with a C – terminal Flag tag(Conceicao et al., 2020), were used as a
574 template to produce variants lacking amino acids at position H69 and V70. The deletion was
575 introduced using Quickchange Lightning Site-Directed Mutagenesis kit (Agilent) following
576 the manufacturer’s instructions. Viruses were purified by ultracentrifugation; 25mL of crude
577 preparation being purified on a 20% sucrose cushion at 2300rpm for 2 hrs at 4°C. After
578 centrifugation, the supernatant was discarded and the viral pellet resuspended in 600 µL
579 DMEM (10% FBS) and stored at -80°C. Infectivity was examined in HEK293 cells
580 transfected with human ACE2, with RLUs normalised to RT activity present in the
581 pseudotyped virus preparation by PERT assay. Western blots were performed on purified
582 virus with anti-HIV1 p24, 1:1,000 (Abcam) or anti-FLAG, 1:2,000 (Sigma) antibodies used
583 following SDS-PAGE and transfer.

584

585

586 *Standardisation of virus input by SYBR Green-based product-enhanced PCR assay (SG-*
587 *PERT)*

588 The reverse transcriptase activity of virus preparations was determined by qPCR using a
589 SYBR Green-based product-enhanced PCR assay (SG-PERT) as previously
590 described(Vermeire et al., 2012). Briefly, 10-fold dilutions of virus supernatant were lysed in
591 a 1:1 ratio in a 2x lysis solution (made up of 40% glycerol v/v 0.25% Triton X-100 v/v
592 100mM KCl, RNase inhibitor 0.8 U/ml, TrisHCL 100mM, buffered to pH7.4) for 10 minutes
593 at room temperature.

594

595 12µl of each sample lysate was added to thirteen 13µl of a SYBR Green master mix
596 (containing 0.5µM of MS2-RNA Fwd and Rev primers, 3.5pmol/ml of MS2-RNA, and
597 0.125U/µl of Ribolock RNase inhibitor and cycled in a QuantStudio. Relative amounts of
598 reverse transcriptase activity were determined as the rate of transcription of bacteriophage
599 MS2 RNA, with absolute RT activity calculated by comparing the relative amounts of RT to
600 an RT standard of known activity.

601

602 *Cell-cell fusion assay*

603 Cell fusion assay was carried out as previously described. Briefly, Vero cells and 293T cells
604 were seeded at 80% confluency in a 24 multiwell plate. 293T cells were co-transfected with
605 1.5 mg of spike expression plasmids in pCDNA3 and 0.5 mg pmCherry-N1 using Fugene 6
606 and following the manufacturer's instructions (Promega). Vero cells were treated with
607 CellTracker™ Green CMFDA (5-chloromethylfluorescein diacetate) (Thermo Scientific) for
608 20 minutes. 293T cells were then detached 5 hours post transfection, mixed together with the
609 green-labelled Vero cells, and plated in a 12 multiwell plate. Cell-cell fusion was measured
610 using an Incucyte and determined as the proportion of merged area to green area over time.
611 Data were then analysed using Incucyte software analysis. Data were normalised to cells
612 transfected only with mCherry protein and mixed with green labelled Vero cells. Graphs
613 were generated using Prism 8 software.

614

615 *Transfection*

616 HEK 293T cells were transfected with 1.5 mg of spike expression plasmids in pCDNA3 and
617 lysed 18 hours post transfection. Cells were treated with Benzonase Nuclease (70664
618 Millipore) and boiled for 5 min. Samples were then run on 4%–12% Bis Tris gels and
619 transferred onto nitrocellulose membranes using an iBlot (Life Technologies).

620

621 *Western blots*

622 Cells were lysed with cell lysis buffer (Cell signalling) or were treated with Benzonase
623 Nuclease (70664 Millipore) and boiled for 5 min. Samples were then run on 4%–12% Bis
624 Tris gels and transferred onto nitrocellulose or PVDF membranes using an iBlot or semidry
625 (Life Technologies and Biorad, respectively).

626

627 Membranes were blocked for 1 hour in 5% non-fat milk in PBS + 0.1% Tween-20 (PBST) at
628 room temperature with agitation, incubated in primary antibody (anti-SARS-CoV-2 Spike,
629 which detects the S2 subunit of SARS-CoV-2 S (Invitrogen, PA1-41165), anti-GAPDH
630 (proteintech) or anti-p24 (NIBSC)) diluted in 5% non-fat milk in PBST for 2 hours at 4°C
631 with agitation, washed four times in PBST for 5 minutes at room temperature with agitation
632 and incubated in secondary antibody (anti-rabbit or anti-mouse HRP conjugate), anti-bactin
633 HRP (sc-47778) diluted in 5% non-fat milk in PBST for 1 hour with agitation at room

634 temperature. Membranes were washed four times in PBST for 5 minutes at room temperature
635 and imaged directly using a ChemiDoc MP imaging system (Bio-Rad).

636

637 *Serum pseudotype neutralisation assay*

638 Spike pseudotype assays have been shown to have similar characteristics as neutralisation
639 testing using fully infectious wild type SARS-CoV-2 (Schmidt et al., 2020). Virus
640 neutralisation assays were performed on 293T cell transiently transfected with ACE2 and
641 TMPRSS2 using SARS-CoV-2 spike pseudotyped virus expressing luciferase (Mlcochova et
642 al., 2020b). Pseudotyped virus was incubated with serial dilution of heat inactivated human
643 serum samples or convalescent plasma in duplicate for 1h at 37°C. Virus and cell only
644 controls were also included. Then, freshly trypsinized 293T ACE2/TMPRSS2 expressing
645 cells were added to each well. Following 48h incubation in a 5% CO₂ environment at 37°C,
646 the luminescence was measured using Steady-Glo Luciferase assay system (Promega).

647 Ethical approval for use of serum samples. Controls with COVID-19 were enrolled to the
648 NIHR BioResource Centre Cambridge under ethics review board (17/EE/0025).

649

650 *Monoclonal antibody neutralisation of B.1.1.7 or B.1.1.7 H69/V70 pseudotyped viruses*

651 Preparation of B.1.1.7 or B.1.1.7 H69/V70 SARS-CoV-2 S glycoprotein-encoding-plasmid
652 used to produce SARS-CoV-2-MLV based on overlap extension PCR. Briefly, a modification
653 of the overlap extension PCR protocol (Forloni et al., 2018) was used to introduce the 9 or
654 7 mutations of the B.1.1.7 and B.1.1.7 H69/V70 lineages, respectively. In a first step, 9 DNA
655 fragments with overlap sequences were amplified by PCR from a plasmid (phCMV1,
656 Genlantis) encoding the full-length SARS-CoV-2 S gene (BetaCoV/Wuhan-Hu-1/2019,
657 accession number mn908947). The mutations (del-69/70, del-144, N501Y, A570D, D614G,
658 P681H, S982A, T716I and D1118H or K417N, E484K and N501Y) were introduced by
659 amplification with primers with similar T_m. Deletion of the C-terminal 21 amino acids was
660 introduced to increase surface expression of the recombinant S (Case et al., 2020). Next, 3
661 contiguous overlapping fragments were fused by a first overlap PCR (step 2) using the
662 utmost external primers of each set, resulting in 3 larger fragments with overlapping
663 sequences. A final overlap PCR (step 3) was performed on the 3 large fragments using the
664 utmost external primers to amplify the full-length S gene and the flanking sequences
665 including the restriction sites KpnI and NotI. This fragment was digested and cloned into the
666 expression plasmid phCMV1. For all PCR reactions the Q5 Hot Start High fidelity DNA

667 polymerase was used (New England Biolabs Inc.), according to the manufacturer's
668 instructions and adapting the elongation time to the size of the amplicon. After each PCR step
669 the amplified regions were separated on agarose gel and purified using Illustra GFX™ PCR
670 DNA and Gel Band Purification Kit (Merck KGaA).

671

672

673 *Ab discovery and recombinant expression*

674 Human mAbs were isolated from plasma cells or memory B cells of SARS-CoV or SARS-
675 CoV-2 immune donors, as previously reported. Recombinant antibodies were expressed in
676 ExpiCHO cells at 37°C and 8% CO₂. Cells were transfected using ExpiFectamine.
677 Transfected cells were supplemented 1 day after transfection with ExpiCHO Feed and
678 ExpiFectamine CHO Enhancer. Cell culture supernatant was collected eight days after
679 transfection and filtered through a 0.2 µm filter. Recombinant antibodies were affinity
680 purified on an ÄKTA xpress FPLC device using 5 mL HiTrap™ MabSelect™ Prisma
681 columns followed by buffer exchange to Histidine buffer (20 mM Histidine, 8% sucrose, pH
682 6) using HiPrep 26/10 desalting columns.

683

684 *MAbs pseudovirus neutralization assay*

685 MLV-based SARS-CoV-2 S-glycoprotein-pseudotyped viruses were prepared as previously
686 described (Pinto et al., 2020). HEK293T/17cells were cotransfected with a WT, B.1.1.7 or
687 B.1.1.7 H69/V70 SARS-CoV-2 spike glycoprotein-encoding-plasmid, an MLV Gag-Pol
688 packaging construct and the MLV transfer vector encoding a luciferase reporter using X-
689 tremeGENE HP transfection reagent (Roche) according to the manufacturer's instructions.
690 Cells were cultured for 72 h at 37°C with 5% CO₂ before harvesting the supernatant. VeroE6
691 stably expressing human TMPRSS2 were cultured in Dulbecco's Modified Eagle's Medium
692 (DMEM) containing 10% fetal bovine serum (FBS), 1% penicillin–streptomycin (100 I.U.
693 penicillin/mL, 100 µg/mL), 8 µg/mL puromycin and plated into 96-well plates for 16–24 h.
694 Pseudovirus with serial dilution of mAbs was incubated for 1 h at 37°C and then added to the
695 wells after washing 2 times with DMEM. After 2–3 h DMEM containing 20% FBS and 2%
696 penicillin–streptomycin was added to the cells. Following 48-72 h of infection, Bio-Glo
697 (Promega) was added to the cells and incubated in the dark for 15 min before reading
698 luminescence with Synergy H1 microplate reader (BioTek). Measurements were done in
699 duplicate and relative luciferase units were converted to percent neutralization and plotted
700 with a non-linear regression model to determine IC₅₀ values using GraphPad PRISM
701 software (version 9.0.0).

702

703

704

705 **References**

706 Baum, A., Fulton, B.O., Wloga, E., Copin, R., Pascal, K.E., Russo, V., Giordano, S., Lanza, K.,
707 Negron, N., Ni, M., *et al.* (2020). Antibody cocktail to SARS-CoV-2 spike protein prevents
708 rapid mutational escape seen with individual antibodies. *Science* 369, 1014-1018.

709 Bazykin, G., Stanevich, O., Danilenko, D., Fadeev, A., Komissarova, K., Ivanova, A., Sergeeva,
710 M., Safina, K., Nabieva, E., Klink, G., *et al.* (2021). Emergence of Y453F and Δ 69-70HV
711 mutations in a lymphoma patient with long-term COVID-19.

712 Bernhart, S.H., Hofacker, I.L., Will, S., Gruber, A.R., and Stadler, P.F. (2008). RNAalifold:
713 improved consensus structure prediction for RNA alignments. *BMC bioinformatics* 9, 1-13.

714 Brejová, B., Hodorová, V., Boršová, K., Čabanová, V., Reizigová, I., Paul, E.D., Čekan, P.,
715 Klempa, B., Nosek, J., and Vinař, T. (2021). B.1.258 Δ , a SARS-CoV-2 variant with Δ H69/ Δ V70
716 in the Spike protein circulating in the Czech Republic and Slovakia. In *Virological*

717 Case, J.B., Rothlauf, P.W., Chen, R.E., Liu, Z., Zhao, H., Kim, A.S., Bloyet, L.M., Zeng, Q.,
718 Tahan, S., Droit, L., *et al.* (2020). Neutralizing Antibody and Soluble ACE2 Inhibition of a
719 Replication-Competent VSV-SARS-CoV-2 and a Clinical Isolate of SARS-CoV-2. *Cell Host*
720 *Microbe* 28, 475-485 e475.

721 Chi, X., Yan, R., Zhang, J., Zhang, G., Zhang, Y., Hao, M., Zhang, Z., Fan, P., Dong, Y., Yang, Y.,
722 *et al.* (2020). A neutralizing human antibody binds to the N-terminal domain of the Spike
723 protein of SARS-CoV-2. *Science* 369, 650-655.

724 Choi, B., Choudhary, M.C., Regan, J., Sparks, J.A., Padera, R.F., Qiu, X., Solomon, I.H., Kuo,
725 H.H., Boucau, J., Bowman, K., *et al.* (2020). Persistence and Evolution of SARS-CoV-2 in an
726 Immunocompromised Host. *The New England journal of medicine* 383, 2291-2293.

727 Collier, D., De Marco, A., Ferreira, I., Meng, B., Datir, R., Walls, A.C., Kemp S, S.A., Bassi, J.,
728 Pinto, D., Silacci Fregni, C., *et al.* (2021a). SARS-CoV-2 B.1.1.7 sensitivity to mRNA vaccine-
729 elicited, convalescent and monoclonal antibodies. medRxiv, 2021.2001.2019.21249840.

730 Collier, D.A., Marco, A.D., Ferreira, I.A.T.M., Meng, B., Datir, R., Walls, A.C., Kemp S, S.A.,
731 Bassi, J., Pinto, D., Fregni, C.S., *et al.* (2021b). SARS-CoV-2 B.1.1.7 escape from mRNA
732 vaccine-elicited neutralizing antibodies. medRxiv, 2021.2001.2019.21249840.

733 Conceicao, C., Thakur, N., Human, S., Kelly, J.T., Logan, L., Bialy, D., Bhat, S., Stevenson-
734 Leggett, P., Zagrajek, A.K., Hollinghurst, P., *et al.* (2020). The SARS-CoV-2 Spike protein has a
735 broad tropism for mammalian ACE2 proteins. *PLoS Biol* 18, e3001016.

736 Delarue, M., Poch, O., Tordo, N., Moras, D., and Argos, P. (1990). An attempt to unify the
737 structure of polymerases. *Protein Engineering, Design and Selection* 3, 461-467.

738 Forloni, M., Liu, A.Y., and Wajapeyee, N. (2018). Creating Insertions or Deletions Using
739 Overlap Extension Polymerase Chain Reaction (PCR) Mutagenesis. *Cold Spring Harb Protoc*
740 2018.

741 Greaney, A.J., Starr, T.N., Gilchuk, P., Zost, S.J., Binshtein, E., Loes, A.N., Hilton, S.K.,
742 Huddleston, J., Eguia, R., and Crawford, K.H. (2020). Complete mapping of mutations to the
743 SARS-CoV-2 spike receptor-binding domain that escape antibody recognition. *Cell Host &*
744 *Microbe*.

745 Hall, T., Biosciences, I., and Carlsbad, C. (2011). BioEdit: an important software for molecular
746 biology. *GERF Bull Biosci* 2, 60-61.

747 Harrison, G.P., Mayo, M.S., Hunter, E., and Lever, A.M. (1998). Pausing of reverse
748 transcriptase on retroviral RNA templates is influenced by secondary structures both 5' and
749 3' of the catalytic site. *Nucleic acids research* 26, 3433-3442.

750 He, X., Lau, E.H.Y., Wu, P., Deng, X., Wang, J., Hao, X., Lau, Y.C., Wong, J.Y., Guan, Y., Tan, X.,
751 *et al.* (2020). Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med*
752 26, 672-675.

753 Hoffmann, M., Zhang, L., Krüger, N., Graichen, L., Kleine-Weber, H., Hofmann-Winkler, H.,
754 Kempf, A., Nessler, S., Riggert, J., Winkler, M.S., *et al.* (2021). SARS-CoV-2 mutations
755 acquired in mink reduce antibody-mediated neutralization. *bioRxiv*,
756 2021.2002.2012.430998.

757 Hou, Y.J., Chiba, S., Halfmann, P., Ehre, C., Kuroda, M., Dinno, K.H., 3rd, Leist, S.R., Schafer,
758 A., Nakajima, N., Takahashi, K., *et al.* (2020). SARS-CoV-2 D614G variant exhibits efficient
759 replication ex vivo and transmission in vivo. *Science* 370, 1464-1468.

760 Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., and Jermini, L.S. (2017).
761 ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 14,
762 587-589.

763 Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version
764 7: improvements in performance and usability. *Mol Biol Evol* 30, 772-780.

765 Kemp, S.A., Collier, D.A., Datir, R., Ferreira, I., Gayed, S., Jahun, A., Hosmillo, M., Rees-Spear,
766 C., Mlcochova, P., Lumb, I.U., *et al.* (2020). Neutralising antibodies in Spike mediated SARS-
767 CoV-2 adaptation. medRxiv, 2020.2012.2005.20241927.

768 Kemp, S.A., Collier, D.A., Datir, R.P., Ferreira, I.A.T.M., Gayed, S., Jahun, A., Hosmillo, M.,
769 Rees-Spear, C., Mlcochova, P., Lumb, I.U., *et al.* (2021). SARS-CoV-2 evolution during
770 treatment of chronic infection. Nature.

771 Korber, B., Fischer, W.M., Gnanakaran, S., Yoon, H., Theiler, J., Abfalterer, W., Hengartner,
772 N., Giorgi, E.E., Bhattacharya, T., and Foley, B. (2020). Tracking changes in SARS-CoV-2 Spike:
773 evidence that D614G increases infectivity of the COVID-19 virus. Cell 182, 812-827. e819.

774 Kozlov, A.M., Darriba, D., Flouri, T., Morel, B., and Stamatakis, A. (2019). RAxML-NG: a fast,
775 scalable and user-friendly tool for maximum likelihood phylogenetic inference.
776 Bioinformatics (Oxford, England) 35, 4453-4455.

777 Lam, T.T., Jia, N., Zhang, Y.W., Shum, M.H., Jiang, J.F., Zhu, H.C., Tong, Y.G., Shi, Y.X., Ni, X.B.,
778 Liao, Y.S., *et al.* (2020). Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins.
779 Nature 583, 282-285.

780 Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., Zhang, Q., Shi, X., Wang, Q., Zhang, L., *et al.*
781 (2020). Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2
782 receptor. Nature 581, 215-220.

783 Larsen, B.B., and Worobey, M. (2020). Identification of a novel SARS-CoV-2 Spike 69-70
784 deletion lineage circulating in the United States ([https://virological.org/t/identification-of-a-
785 novel-sars-cov-2-spike-69-70-deletion-lineage-circulating-in-the-united-states/577](https://virological.org/t/identification-of-a-novel-sars-cov-2-spike-69-70-deletion-lineage-circulating-in-the-united-states/577)).

786 Larsen, H.D., Fonager, J., Lomholt, F.K., Dalby, T., Benedetti, G., Kristensen, B., Urth, T.R.,
787 Rasmussen, M., Lassauniere, R., Rasmussen, T.B., *et al.* (2021). Preliminary report of an
788 outbreak of SARS-CoV-2 in mink and mink farmers associated with community spread,
789 Denmark, June to November 2020. Euro Surveill 26.

790 MacLean, O.A., Lytras, S., Weaver, S., Singer, J.B., Boni, M.F., Lemey, P., Pond, S.L.K., and
791 Robertson, D.L. (2020). Natural selection in the evolution of SARS-CoV-2 in bats, not
792 humans, created a highly capable human pathogen. BioRxiv.

793 Martin, D.P., Murrell, B., Golden, M., Khoosal, A., and Muhire, B. (2015). RDP4: Detection
794 and analysis of recombination patterns in virus genomes. Virus evolution 1.

795 McCallum, M., Marco, A.D., Lempp, F., Tortorici, M.A., Pinto, D., Walls, A.C., Beltramello, M.,
796 Chen, A., Liu, Z., Zatta, F., *et al.* (2021). N-terminal domain antigenic mapping reveals a site
797 of vulnerability for SARS-CoV-2. *bioRxiv*, 2021.2001.2014.426475.

798 McCarthy, K.R., Rennick, L.J., Nambulli, S., Robinson-McCarthy, L.R., Bain, W.G., Haidar, G.,
799 and Duprex, W.P. (2020). Natural deletions in the SARS-CoV-2 spike glycoprotein drive
800 antibody escape. *bioRxiv*, 2020.2011.2019.389916.

801 McCarthy, K.R., Rennick, L.J., Nambulli, S., Robinson-McCarthy, L.R., Bain, W.G., Haidar, G.,
802 and Duprex, W.P. (2021). Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive
803 antibody escape. *Science*, eabf6950.

804 Minh, B.Q., Nguyen, M.A., and von Haeseler, A. (2013). Ultrafast approximation for
805 phylogenetic bootstrap. *Mol Biol Evol* 30, 1188-1195.

806 Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A.,
807 and Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic
808 Inference in the Genomic Era. *Mol Biol Evol* 37, 1530-1534.

809 Mlcochova, P., Chadha, A., Hesselhoj, T., Fraternali, F., Ramsden, J., and Gupta, R. (2020a).
810 Extended in vitro inactivation of SARS-CoV-2 by titanium dioxide surface coating. *bioRxiv*.

811 Mlcochova, P., Collier, D., Ritchie, A., Assennato, S.M., Hosmillo, M., Goel, N., Meng, B.,
812 Chatterjee, K., Mendoza, V., Temperton, N., *et al.* (2020b). Combined point of care nucleic
813 acid and antibody testing for SARS-CoV-2 following emergence of D614G Spike Variant. *Cell*
814 *Rep Med*, 100099.

815 Mlcochova, P., Collier, D., Ritchie, A., Assennato, S.M., Hosmillo, M., Goel, N., Meng, B.,
816 Chatterjee, K., Mendoza, V., Temperton, N., *et al.* (2020c). Combined Point-of-Care Nucleic
817 Acid and Antibody Testing for SARS-CoV-2 following Emergence of D614G Spike Variant. *Cell*
818 *Rep Med* 1, 100099.

819 Munnink, B.B.O., Sikkema, R.S., Nieuwenhuijse, D.F., Molenaar, R.J., Munger, E.,
820 Molenkamp, R., Van Der Spek, A., Tolsma, P., Rietveld, A., and Brouwer, M. (2020).
821 Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans.
822 *Science*.

823 Ollis, D., Brick, P., Hamlin, R., Xuong, N., and Steitz, T. (1985). Structure of large fragment of
824 *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature* 313, 762-766.

825 Papa, G., Mallery, D.L., Albecka, A., Welch, L., Cattin-Ortolá, J., Luptak, J., Paul, D.,
826 McMahon, H.T., Goodfellow, I.G., Carter, A., *et al.* (2020). Furin cleavage of SARS-CoV-2

827 Spike promotes but is not essential for infection and cell-cell fusion. bioRxiv,
828 2020.2008.2013.243303.

829 Rambaut, A., Holmes, E.C., O'Toole, A., Hill, V., McCrone, J.T., Ruis, C., du Plessis, L., and
830 Pybus, O.G. (2020). A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist
831 genomic epidemiology. *Nat Microbiol* 5, 1403-1407.

832 Rambaut A., L.N., Pybus O, Barclay W, Carabelli A. C., Connor T., Peacock T., Robertson D. L.,
833 Volz E., on behalf of COVID-19 Genomics Consortium UK (CoG-UK). (2020). Preliminary
834 genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel
835 set of spike mutations.

836 Reeder, R.H., and Lang, W. (1994). The mechanism of transcription termination by RNA
837 polymerase I. *Molecular microbiology* 12, 11-15.

838 Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated
839 protein structure and function prediction. *Nat Protoc* 5, 725-738.

840 Schmidt, F., Weisblum, Y., Muecksch, F., Hoffmann, H.-H., Michailidis, E., Lorenzi, J.C.C.,
841 Mendoza, P., Rutkowska, M., Bednarski, E., Gaebler, C., *et al.* (2020). Measuring SARS-CoV-2
842 neutralizing antibody activity using pseudotyped and chimeric viruses.
843 2020.2006.2008.140871.

844 Shu, Y., and McCauley, J. (2017). GISAID: Global initiative on sharing all influenza data - from
845 vision to reality. *Euro surveillance : bulletin Europeen sur les maladies transmissibles =*
846 *European communicable disease bulletin* 22, 30494.

847 Soh, W.T., Liu, Y., Nakayama, E.E., Ono, C., Torii, S., Nakagami, H., Matsuura, Y., Shioda, T.,
848 and Arase, H. (2020). The N-terminal domain of spike glycoprotein mediates SARS-CoV-2
849 infection by associating with L-SIGN and DC-SIGN. bioRxiv, 2020.2011.2005.369264.

850 Sousa, R., Chung, Y.J., Rose, J.P., and Wang, B.-C. (1993). Crystal structure of bacteriophage
851 T7 RNA polymerase at 3.3 Å resolution. *Nature* 364, 593-599.

852 Starr, T.N., Greaney, A.J., Addetia, A., Hannon, W.H., Choudhary, M.C., Dingens, A.S., Li, J.Z.,
853 and Bloom, J.D. (2020a). Prospective mapping of viral mutations that escape antibodies
854 used to treat COVID-19. bioRxiv.

855 Starr, T.N., Greaney, A.J., Hilton, S.K., Ellis, D., Crawford, K.H.D., Dingens, A.S., Navarro, M.J.,
856 Bowen, J.E., Tortorici, M.A., Walls, A.C., *et al.* (2020b). Deep Mutational Scanning of SARS-
857 CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* 182,
858 1295-+.

859 Sungnak, W., Huang, N., Becavin, C., Berg, M., Queen, R., Litvinukova, M., Talavera-Lopez,
860 C., Maatz, H., Reichart, D., Sampaziotis, F., *et al.* (2020). SARS-CoV-2 entry factors are highly
861 expressed in nasal epithelial cells together with innate immune genes. *Nat Med* 26, 681-
862 687.

863 Thomson, E.C., Rosen, L.E., Shepherd, J.G., Spreafico, R., da Silva Filipe, A., Wojcechowskyj,
864 J.A., Davis, C., Piccoli, L., Pascall, D.J., and Dillen, J. (2020). The circulating SARS-CoV-2 spike
865 variant N439K maintains fitness while evading antibody-mediated immunity. *bioRxiv*.

866 Vermeire, J., Naessens, E., Vanderstraeten, H., Landi, A., Iannucci, V., Van Nuffel, A., Taghon,
867 T., Pizzato, M., and Verhasselt, B. (2012). Quantification of reverse transcriptase activity by
868 real-time PCR as a fast and accurate method for titration of HIV, lenti- and retroviral vectors.
869 *PloS one* 7, e50859-e50859.

870 Volz, E., Mishra, S., Chand, M., Barrett, J.C., Johnson, R., Geidelberg, L., Hinsley, W.R.,
871 Laydon, D.J., Dabrera, G., O'Toole, Á., *et al.* (2021). Transmission of SARS-CoV-2 Lineage
872 B.1.1.7 in England: Insights from linking epidemiological and genetic data. *medRxiv*,
873 2020.2012.2030.20249034.

874 Wrobel, A.G., Benton, D.J., Xu, P., Roustan, C., Martin, S.R., Rosenthal, P.B., Skehel, J.J., and
875 Gamblin, S.J. (2020). SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on
876 virus evolution and furin-cleavage effects. *Nat Struct Mol Biol* 27, 763-767.

877 Xie, X., Liu, Y., Liu, J., Zhang, X., Zou, J., Fontes-Garfias, C.R., Xia, H., Swanson, K.A., Cutler,
878 M., Cooper, D., *et al.* (2021). Neutralization of SARS-CoV-2 spike 69/70 deletion, E484K and
879 N501Y variants by BNT162b2 vaccine-elicited sera. *Nature medicine*.

880 Yurkovetskiy, L., Wang, X., Pascal, K.E., Tomkins-Tinch, C., Nyalile, T.P., Wang, Y., Baum, A.,
881 Diehl, W.E., Dauphin, A., Carbone, C., *et al.* (2020). Structural and Functional Analysis of the
882 D614G SARS-CoV-2 Spike Protein Variant. *Cell* 183, 739-751 e738.

883 Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., Si, H.-R., Zhu, Y., Li, B., and
884 Huang, C.-L. (2020). A pneumonia outbreak associated with a new coronavirus of probable
885 bat origin. *nature* 579, 270-273.

886

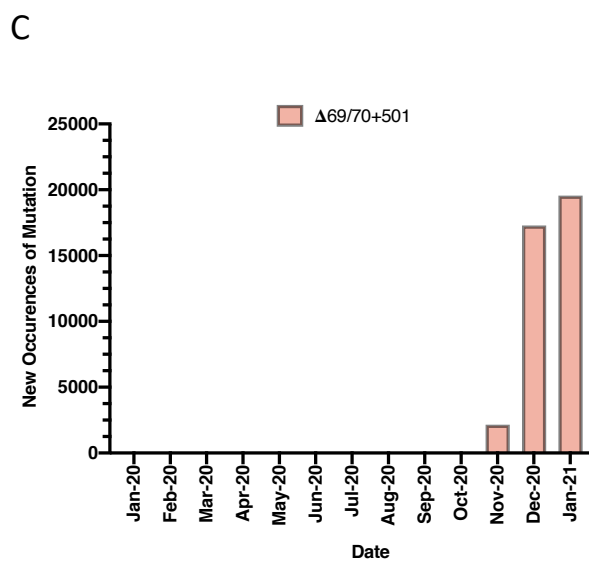
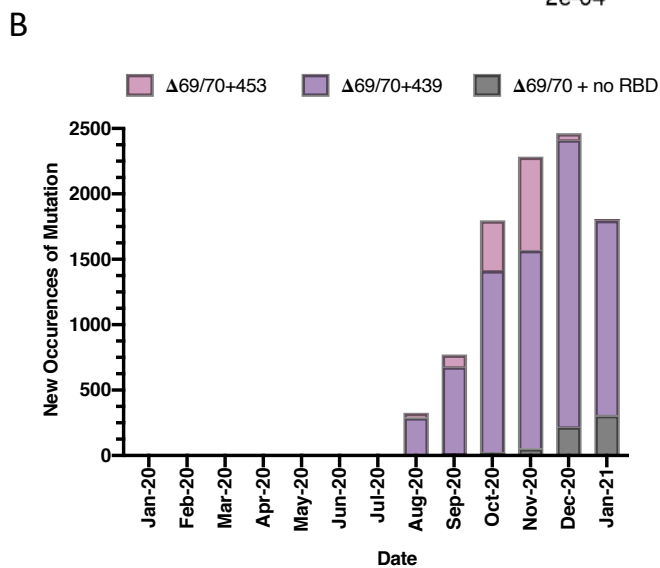
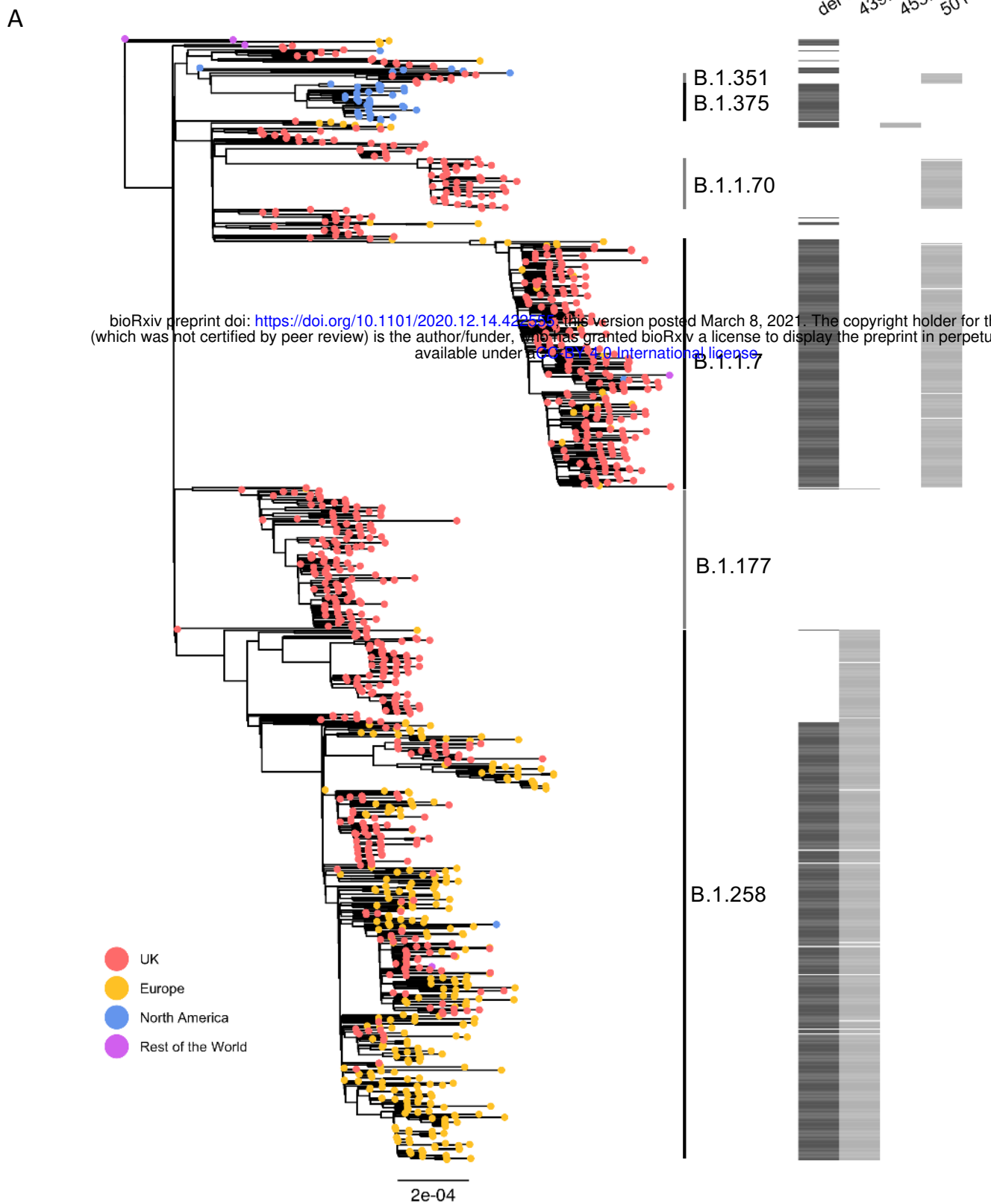
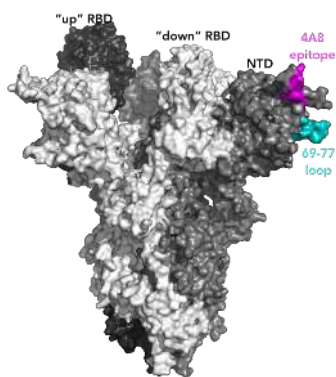
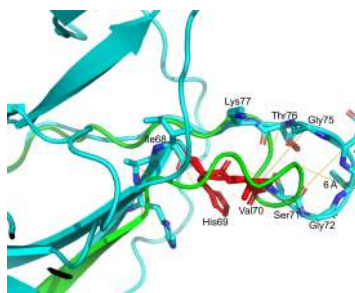


Figure 1. A. Global phylogeny of SARS-CoV-2 whole genome sequences highlighting those with specific mutations in Spike: Δ H69/V70, N439K, Y453F and N501Y. Tree is subsampled and tips are coloured by geographic region (see key). Grey bars on the right show the presence or absence of the deletion Δ H69/V70 and amino acid variants N439K, Y453F, and N501Y. Lineages from Rambaut et al. 2020 are shown. **New occurrences of SARS-CoV-2 sequences with the Δ H69/V70 deletion by month for B. Δ H69/V70 with or without N439K/ Y453F and C. Δ H69/V70 with or without N501Y. Indicated frequencies by month of the Δ H69/V70 deletion are from the GISAID database (accessed 18th Feb 2021) by reporting country and sampling date.**

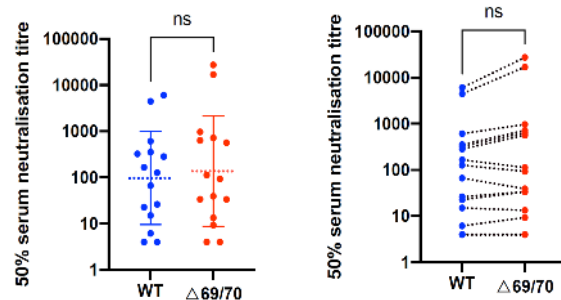
A



B

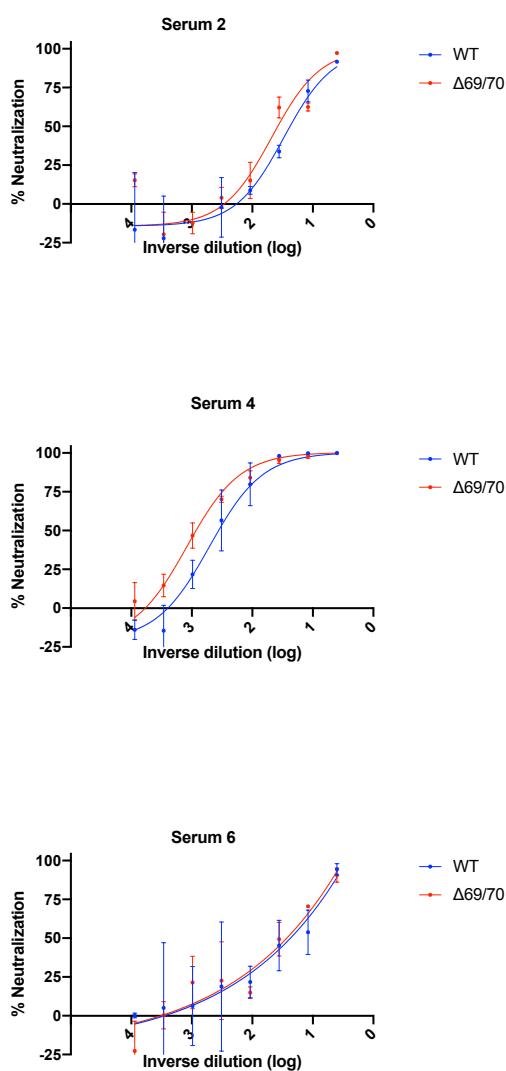


C

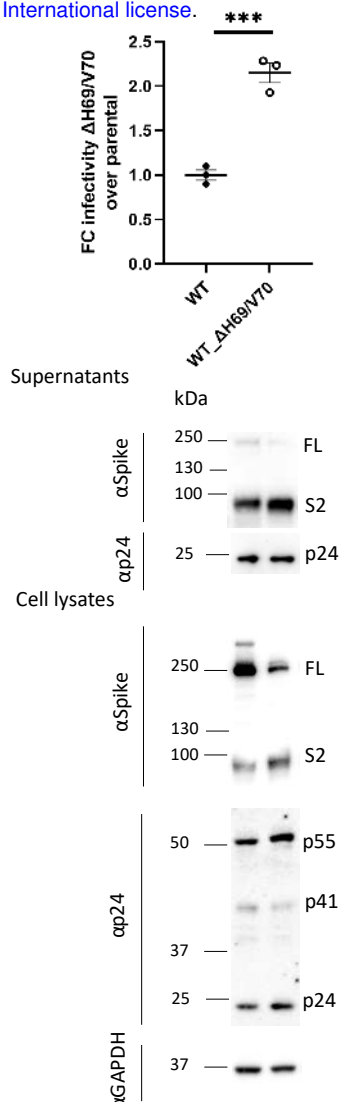


bioRxiv preprint doi: <https://doi.org/10.1101/2020.12.14.422555>; this version posted March 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

D



E



F

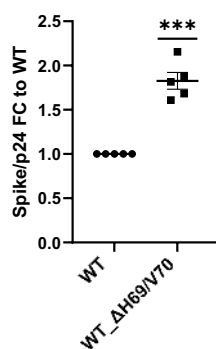


Figure 2. Spike Δ H69/V70 does not reduce sensitivity to neutralising serum antibodies but does increase infectivity

A. Surface representation of spike homotrimer in open conformation (PDB: 7C2L) with each monomer shown in different shades of grey. On the monomer shown positioned to the right, the exposed loop consisting of residues 69-77 is shown in cyan and the neutralising antibody (4A8) binding NTD epitope in magenta. **B.** Prediction of conformational change in the spike N-terminal domain due to deletion of residues His69 and Val70. The pre-deletion structure is shown in cyan, except for residues 69 and 70, which are shown in red. The predicted post-deletion structure is shown in green. Residues 66-77 of the pre-deletion structure are shown in stick representation and coloured by atom (nitrogen in blue, oxygen in coral). Yellow lines connect aligned residues 66-77 of the pre- and post-deletion structures and the distance of 6 Å between aligned alpha carbons of Thr73 in the pre- and post-deletion conformation is labelled. **C.** Neutralisation of spike Δ H69/V70 pseudotyped virus and wild type (D614G background) by convalescent sera from 15 donors. GMT (geometric mean titre) with s.d presented of two independent experiments each with two technical repeats. Wilcoxon matched-pairs signed rank test, ns not significant. **D.** Example neutralisation curves. Indicated is serum \log_{10} inverse dilution against % neutralisation. Data points represent means of technical replicates and error bars represent standard deviation. **E.** Single round infection by spike Δ H69/V70 pseudotyped virus. Data from three experiments shown with mean and standard error of mean (SEM). Representative western blot of unspun supernatant and cell lysates probed with antibodies for HIV-1 p24 and SARS-Cov-2 S2. **F.** Quantification of cleaved S2 spike:p24 ratio for wild type virus with Δ H69/V70 deletion versus WT alone across multiple replicate experiments. Mean and SEM are shown Students t-test *** $p < 0.001$.

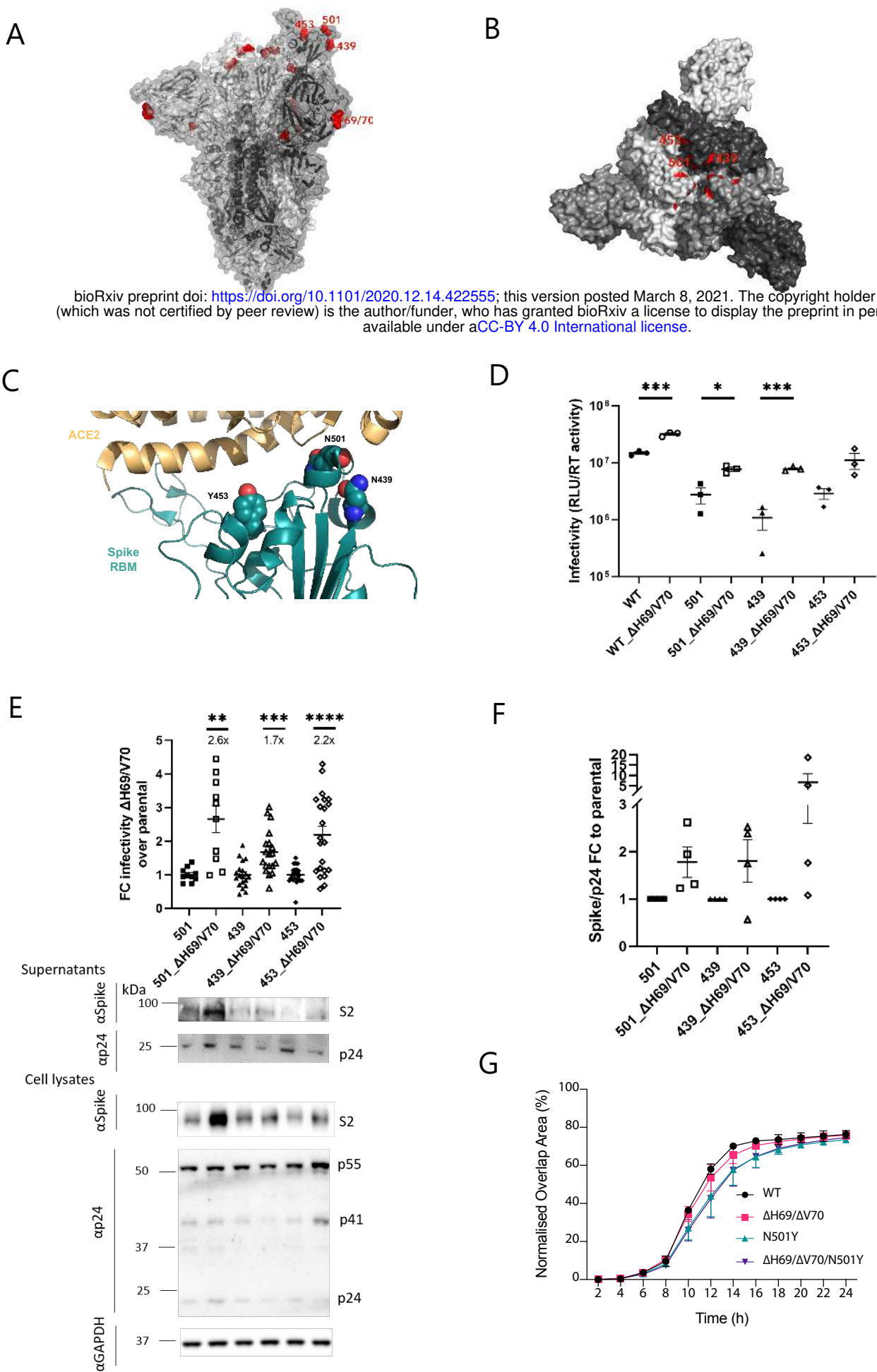


Figure 3: Δ H69/V70 compensates for reduced infectivity of RBD mutations N439K, Y453F and N501Y.

A. Spike in open conformation with a single erect RBD (PDB: 6ZGG) in trimer axis vertical view with the location of Δ H69/V70 in the N-terminal domain and RBD mutations highlighted as red spheres and labelled on the monomer with erect RBD. **B.** Surface representation of spike in closed conformation (PDB: 6ZGE) viewed in a ‘top-down’ view along the trimer axis. The residues associated with RBD substitutions N439K, Y453F and N501Y are highlighted in red and labelled on a single monomer. **C.** Representation of Spike RBM:ACE2 interface (PDB: 6MOJ) with residues N439, Y453 and N501 highlighted as spheres coloured by element. **D-F.** Spike mutant Δ H69/V70 compensates for infectivity defect of Spike RBD mutations and is associated with increased Spike incorporation into virions. **D.** Infectivity of Spike (D614G) Δ H69/V70 deletion in absence and presence of Spike RBD mutations. Single round infection by luciferase expressing lentivirus pseudotyped with SARS-CoV-2 spike protein on HeLa cells transduced with ACE2. **E.** Fold change infectivity over multiple experiments comparing RBD mutants with and without Δ H69/V70, with mean and SEM shown. Representative western blot of unspun supernatant and cell lysates probed with antibodies against HIV-1 p24 and SARS-CoV-2 spike S2. **F.** Quantification of cleaved S2 spike:p24 ratio for wild type virus with Δ H69/V70 deletion versus WT alone across multiple replicate experiments. Mean and SEM are shown. **G.** Cell-cell fusion kinetics for mutant spike proteins. Student's t-test *** $p < 0.001$. Data for D and E show infectivity normalized for virus input using reverse transcriptase activity in virus supernatants. RLU – relative light units; U – unit of reverse transcriptase activity (RT). Data are technical replicates and are representative of 2 independent experiments. Student t test * $p < 0.05$, *** $p < 0.001$

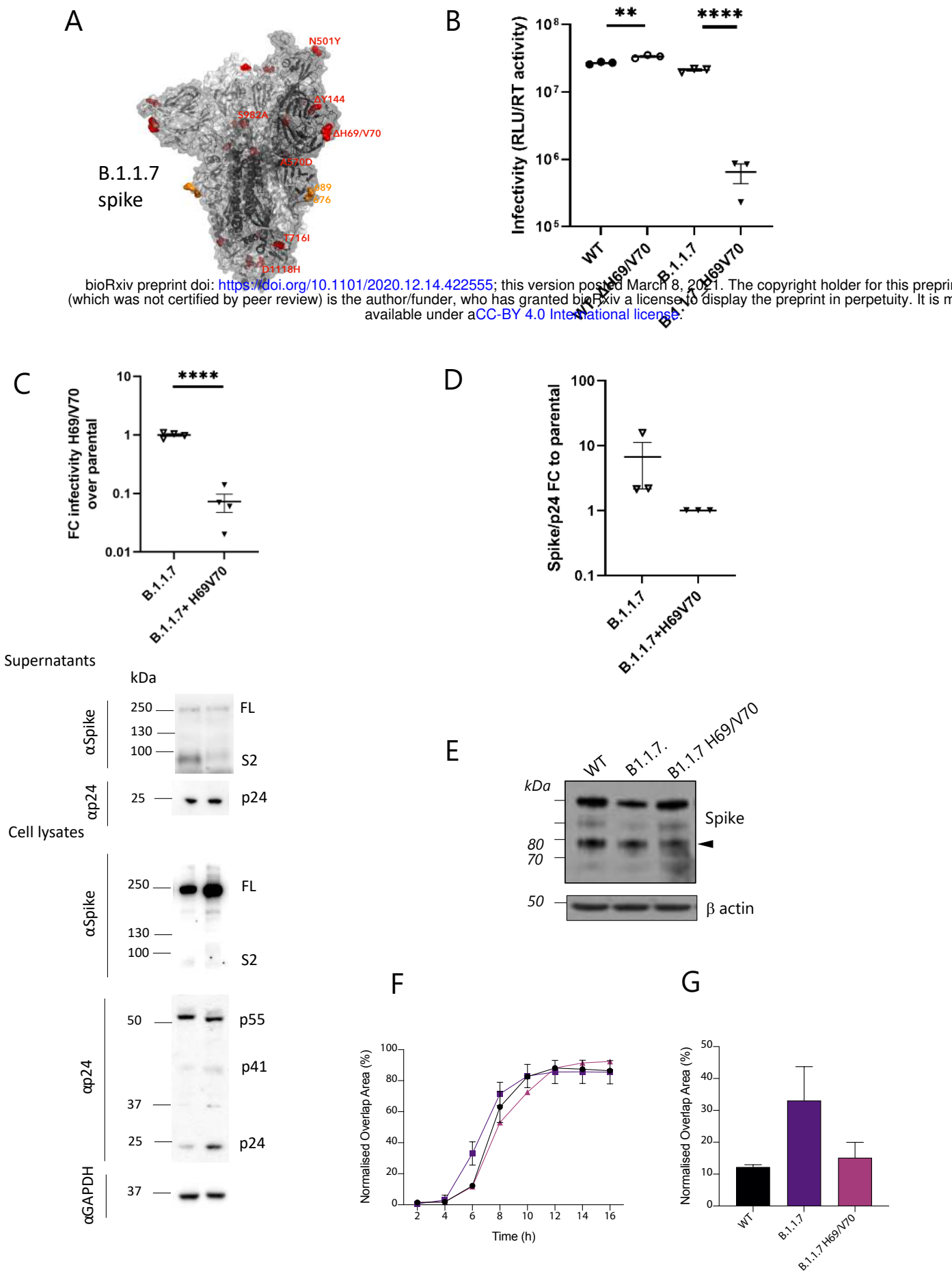
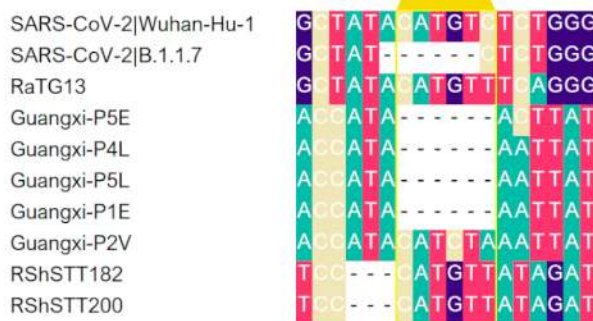
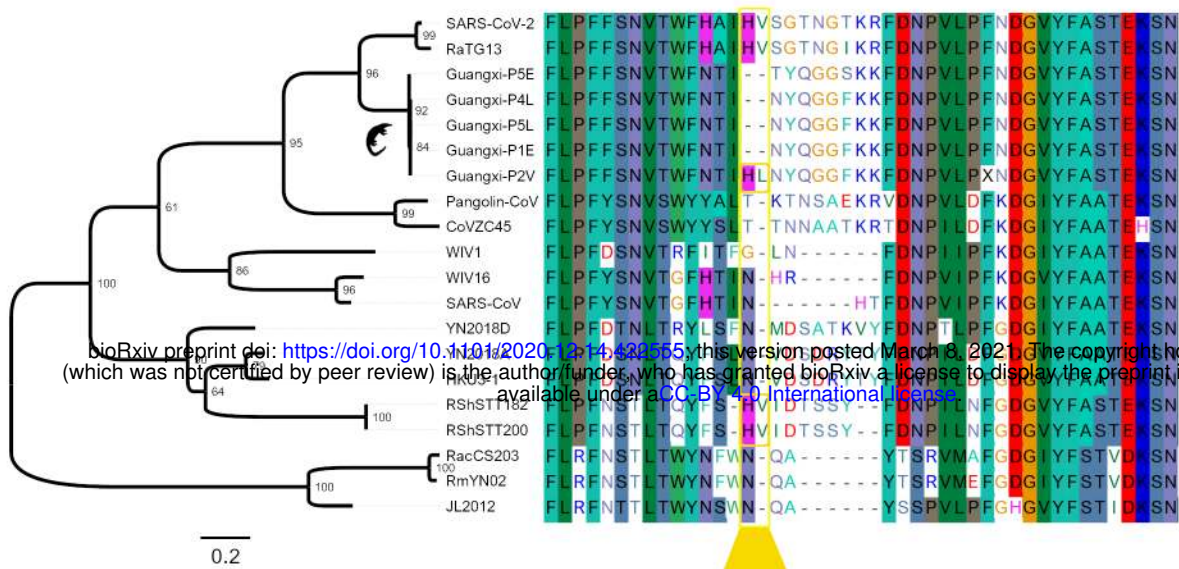
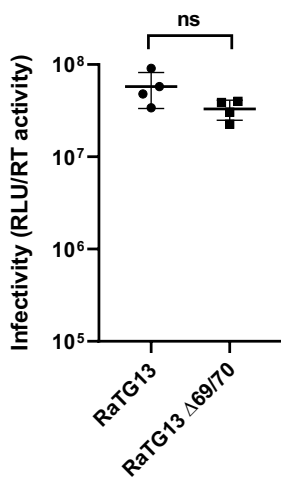


Figure 4. Spike Δ H69/V70 in B.1.1.7 enhances spike infectivity and cell fusogenicity. **A.** Surface representation of spike homotrimer in open conformation with one upright RBD overlaid with ribbon representation (PDB: 6ZGG, Wrobel et al., 2020), with different monomers shown in shades of grey. The deleted residues H69 and V70 and the residues involved in amino acid substitutions (501, 570, 716, 982 and 1118) and the deletion at position 144 are coloured red on each monomer and labelled on the monomer with an upright RBD. The location of an exposed loop containing the furin cleavage site and including residue 681 is absent from the structure, though modelled residues either side of this loop, 676 and 689, are coloured orange. **B.** Infectivity of B.1.1.7 with replacement of H69 and V70 versus B.1.1.7 containing Spike Δ H69/V70 and wild type (D614G) spike. Single round infection by luciferase expressing lentivirus pseudotyped with SARS-CoV-2 Spike protein on HeLa cells transduced with ACE2. **C.** Infectivity of B.1.1.7 without the Δ H69/V70 expressed as fold change compared to B.1.1.7, and representative western blot analysis following transfection of cells with spike and lentiviral plasmids. Supernatant (unspun) loading was normalized for input virus input using reverse transcriptase activity in virus supernatants. **D.** Quantification of Spike:p24 ratio for wild type virus with Δ H69/V70 deletion versus WT alone across replicate experiments. Antibodies against HIV-1 p24 and Spike S2 were used. Data are representative of at least two independent experiments. **E-F. Cell-cell fusion.** Representative western blot of cells transfected with the indicated Spike mutants. The cleaved Spike identifies the S2 subunit and is indicated with the arrowhead. Quantification of cell-cell fusion kinetics showing percentage of green and red overlap area over time, indicative of successful cell-cell fusion. **G.** Data for cell-cell fusion at 6hrs post transfection. Data are representative of two independent experiments



D



E

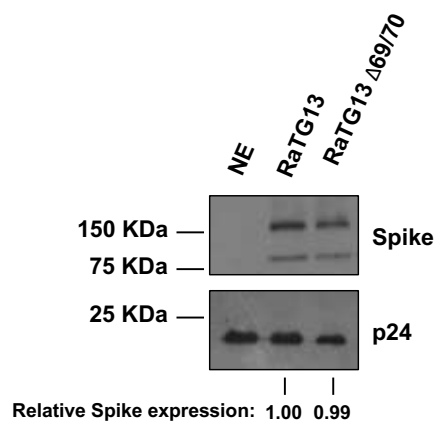
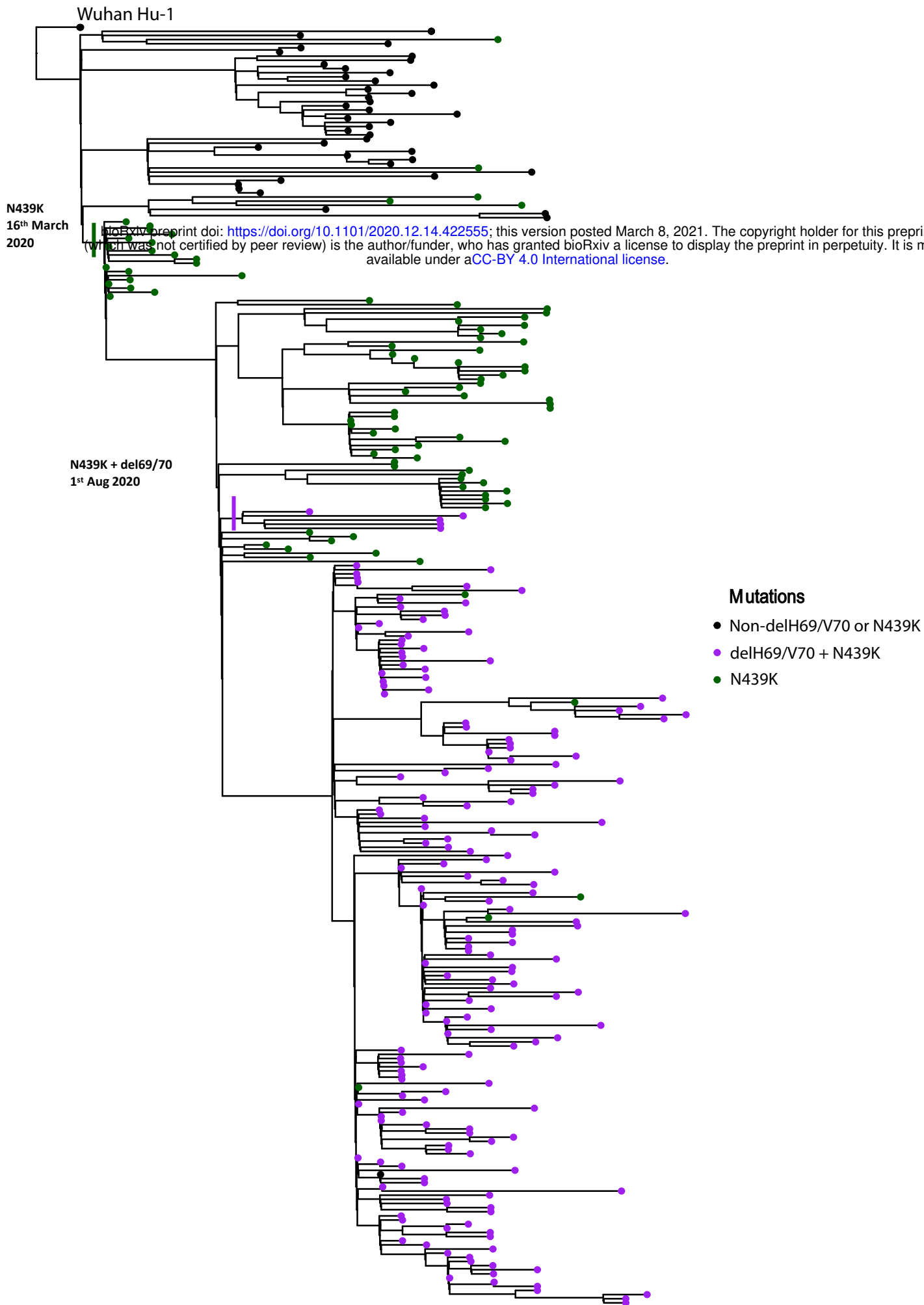
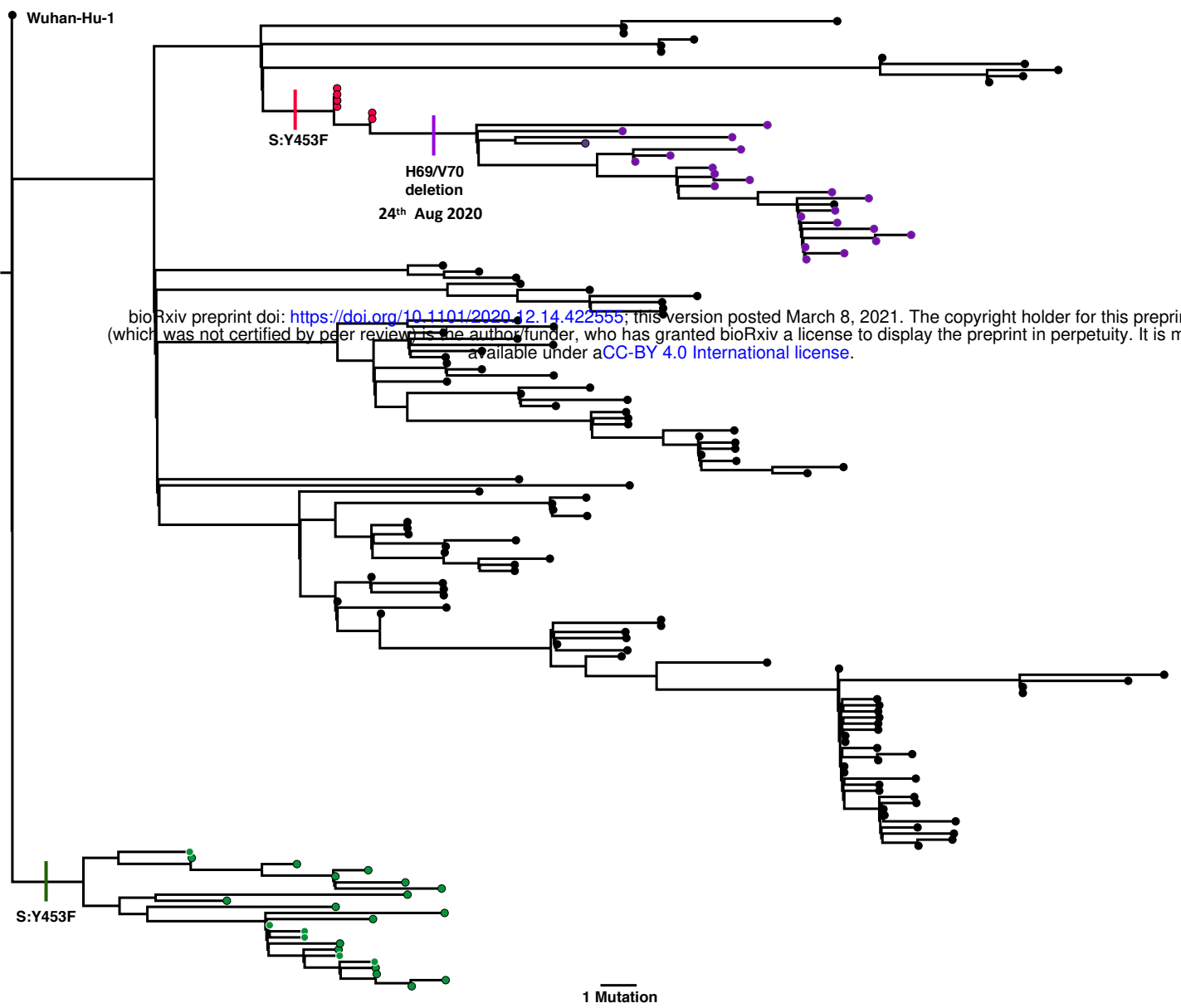


Figure 5. Comparison of the H69/V70 deletion site to other *Sarbecoviruses*. **A.** phylogeny for the Spike peptide region 1-256 **B.** protein sequences from 20 Sarbecoviruses, including SARS-CoV-2 (Wuhan-Hu-1) and SARS-CoV (HSZ-Cc), with distinct genotypes at the Spike region around amino acid positions 69 and 70 (highlighted in yellow box). The 69/70 HL insertion in the P2V sequence from the Guangxi pangolin virus cluster and the HV convergent insertion in the RShSTT182/200 bat virus sequences are highlighted. **C.** The nucleotide alignment between SARS-CoV-2 Wuhan-Hu-1, B.1.1.7, the bat sarbecovirus RaTG13 RShSTT182/200 and the Guangxi pangolin viruses shows the difference between the out-of-frame deletion observed in the former and the in-frame deletion in the latter. **D.** Single round infection by luciferase expressing lentivirus pseudotyped with RaTG13 Spike protein on 293T cells transduced with ACE2. Experiments were performed in biological quadruplicate with the mean and standard deviation plotted. Results are representative of experiments performed two times. Statistical significance was assessed using an unpaired t-test (ns; non-significant, ***; <0.005). **E.** Representative western blot of supernatant from virus producer cells. Spike and HIV pseudotype abundances were assessed using Flag and p24 antibodies, respectively. Relative spike expression was calculated by densitometry using Image J. NE: no envelope.



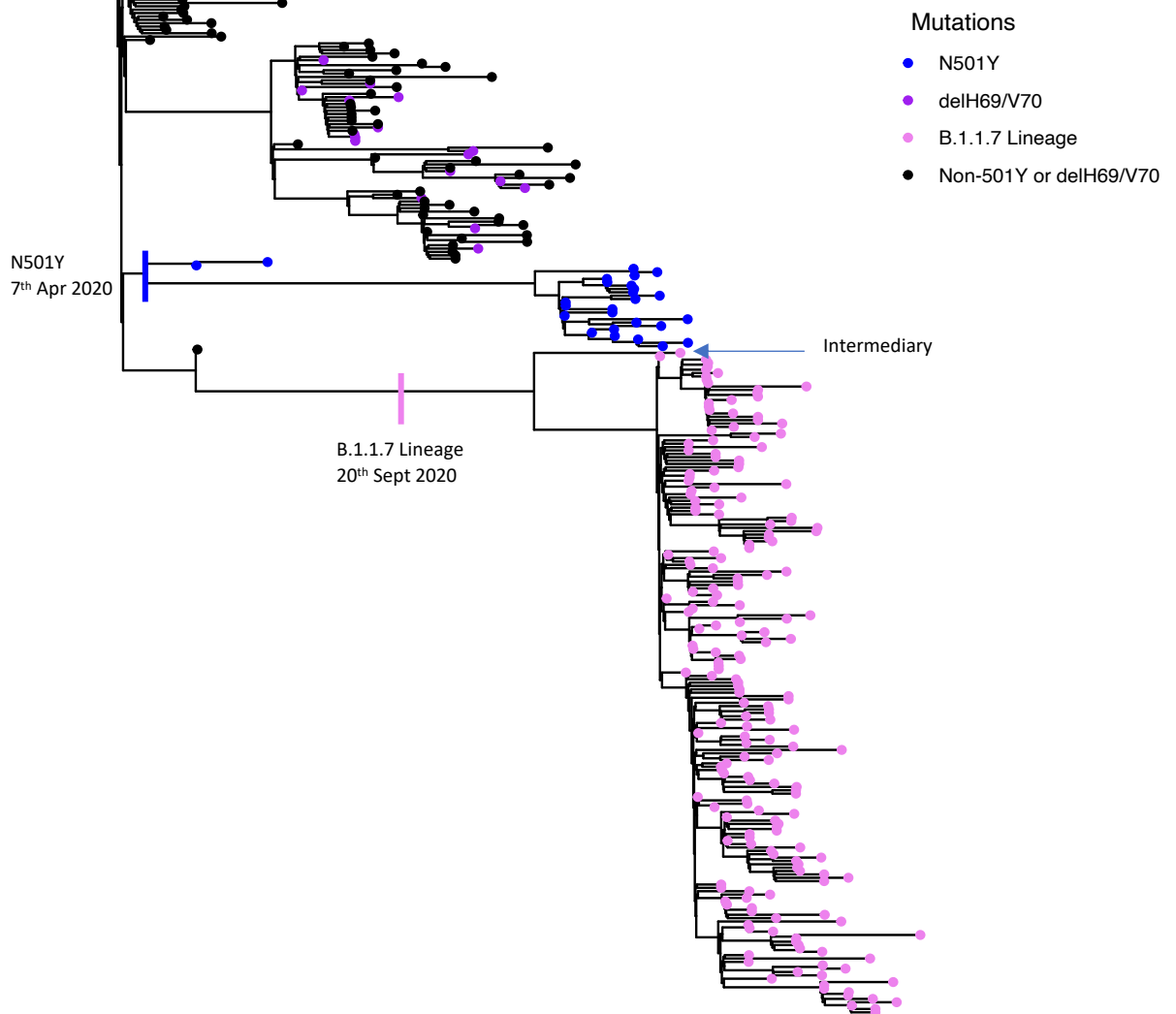
Supplementary Figure 1. Maximum likelihood phylogeny of global sequences carrying Spike mutant N439K. All sequences in the GISAID database containing S:439K (12094 sequences, 18th February 2021) were downloaded, realigned to Wuhan-Hu-1 using MAFFT and deduplicated. Viruses carrying the Spike deletion Δ H69/V70 emerged and expanded from viruses with S:439K, predominantly across the United Kingdom and Europe.



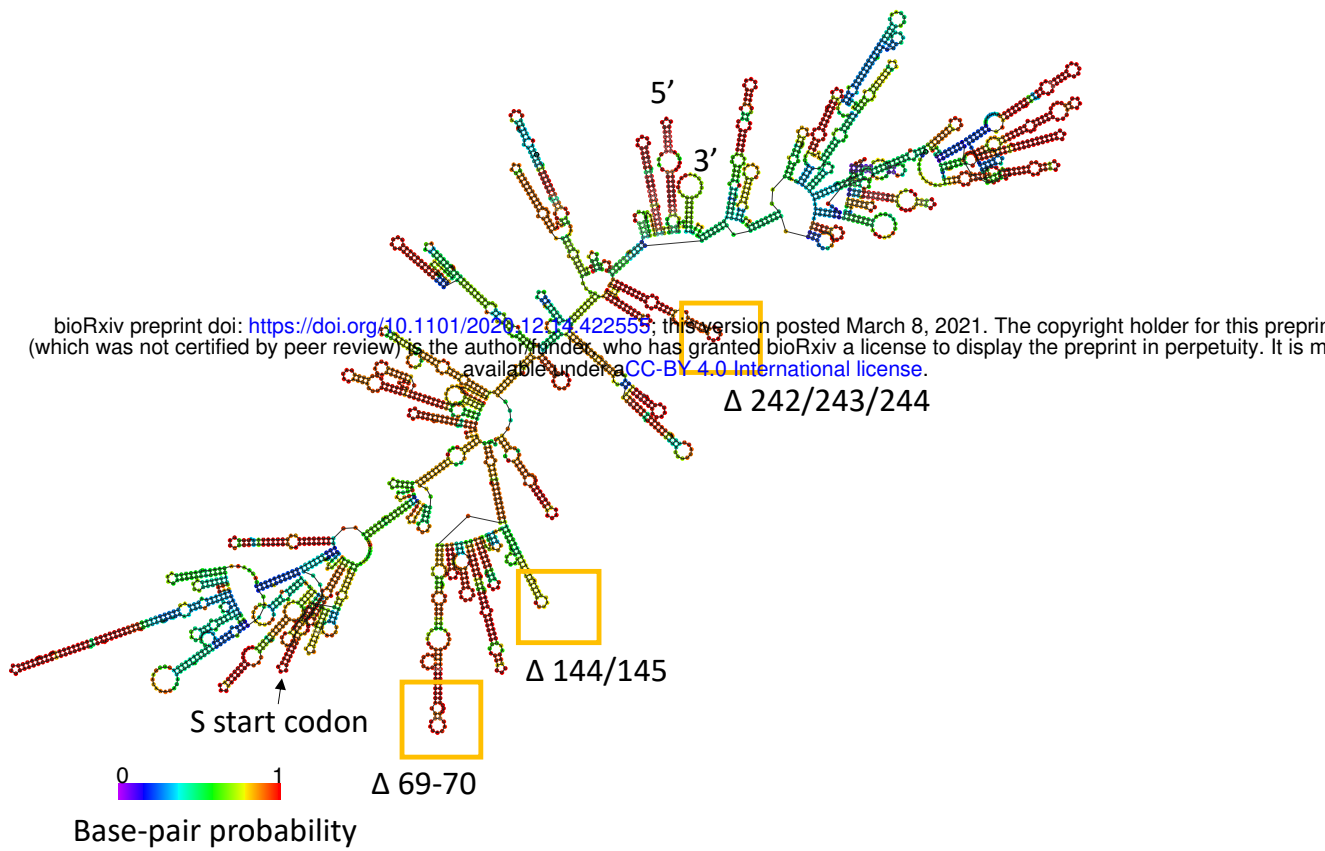
Supplementary Figure 2: Maximum likelihood phylogenetic tree of SARS-CoV-2 sequences carrying Spike mutant Y453F. 753 sequences in the GISAID database (accessed 14th February 2021) were downloaded and realigned to Wuhan-Hu-1 using MAFFT. Two distinct lineages carrying the mink-associated Spike Y453F mutations can be seen in Danish (red) sequences, with a separate lineage isolated only in Netherlands (green). After acquiring the Y453F mutation, Danish mink also appeared to acquire the Spike deletion Δ H69/V70 (purple).

Wuhan-Hu-1

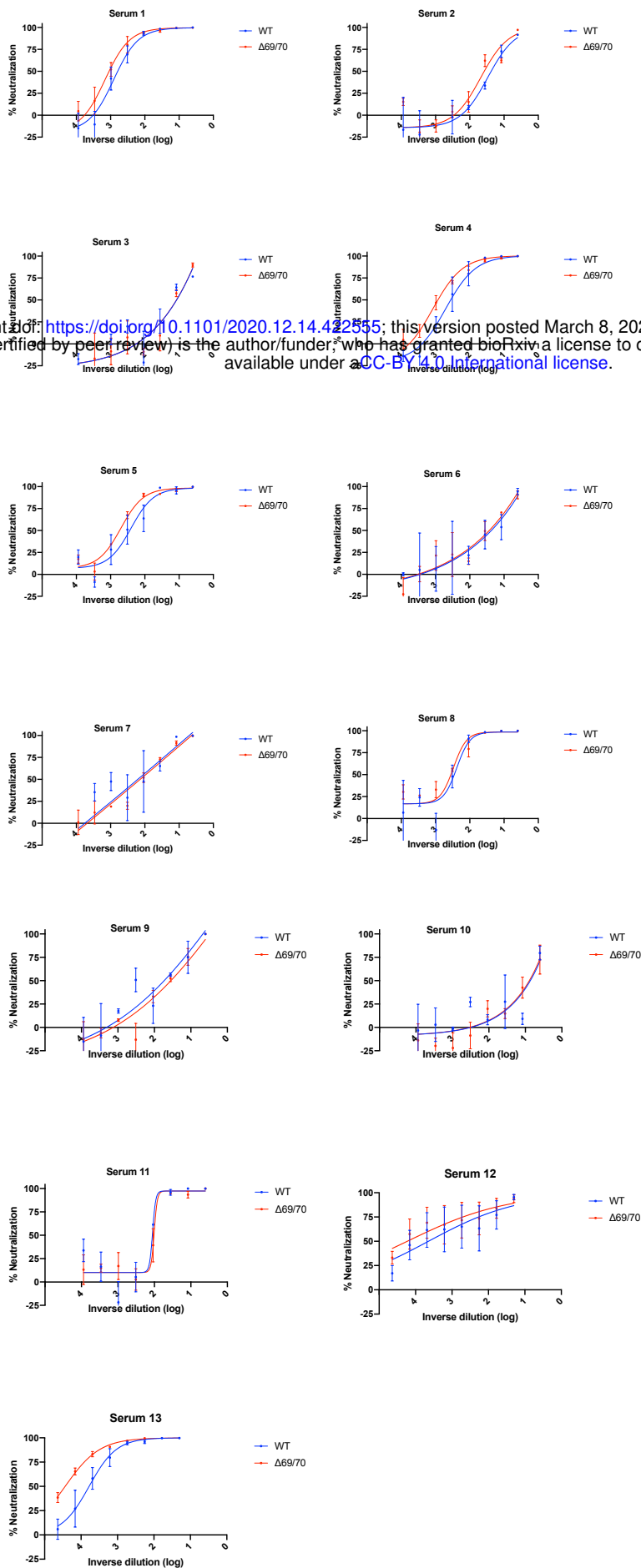
bioRxiv preprint doi: <https://doi.org/10.1101/2020.12.14.492555>; this version posted March 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



Supplementary Figure 3: Maximum likelihood phylogeny of UK viruses bearing Δ 69/70 and N501Y mutations. Two distinct lineages of the Δ H69/V70 were observed to expand in the UK, separately from the 501Y lineage. Prior to expansion of the B.1.1.7 lineage, clusters of infections bearing either N501Y or Δ H69/V70 were observed. Alongside expansion of the B.1.1.7 lineage, is a population in Wales that carries 501Y, but no Δ H69/V70. An intermediary was detected alongside the B.1.1.7 lineage (indicated on phylogeny) which had only a subset of the mutations that make up B.1.1.7 (Δ H69/V70, N501Y, A570D and D1118H).

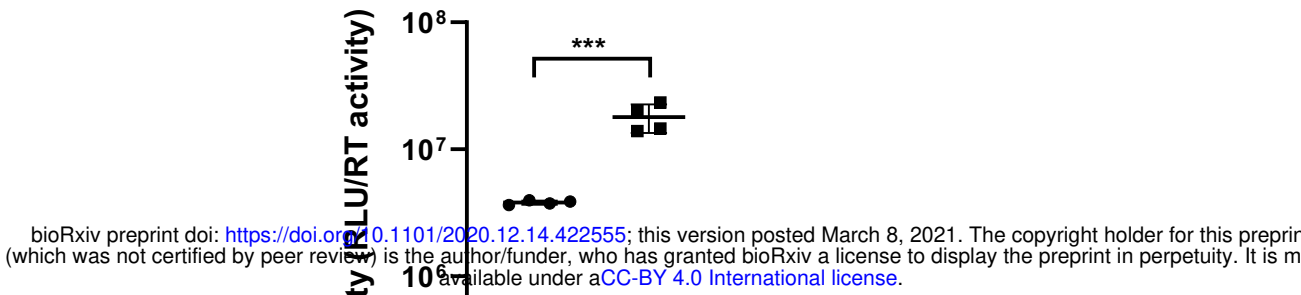


Supplementary Figure 4: The positions of common deletion mutations on the RNA structure of the Spike Δ 69-70 region of the gRNA. The optimal secondary structure was generated from a consensus alignment of human SARS-CoV2 RNAs using RNAalifold. Figure shows nucleotides 20277-23265. Base-pair probability, representative of the breadth of the structural ensemble that could be adopted by the RNA, is shown in colour according to the key.

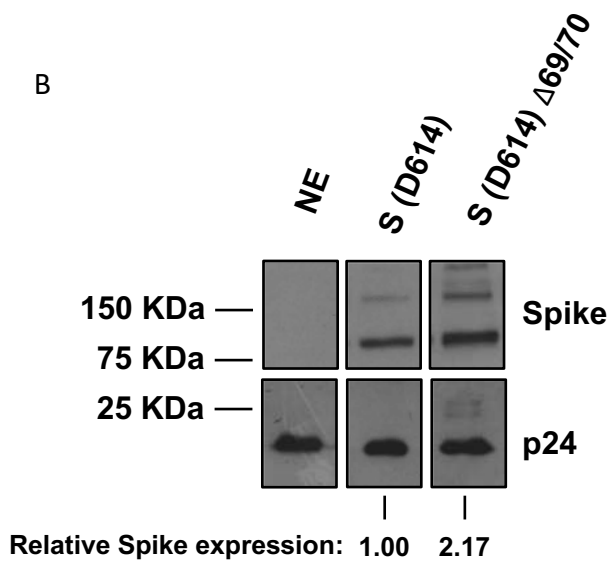


Supplementary Figure 5: Neutralization of pseudovirus virus bearing Spike Δ H69/V70 and wild type (all In Spike D614G background) by convalescent sera from 13 donors who showed neutralisation. Indicated is serum \log_{10} inverse dilution against % neutralisation. Where a curve is shifted to the right this indicates the virus is less sensitive to the neutralising antibodies in the serum. Two of fifteen original sera were non neutralising. Data points represent means of technical replicates and error bars represent standard deviation.

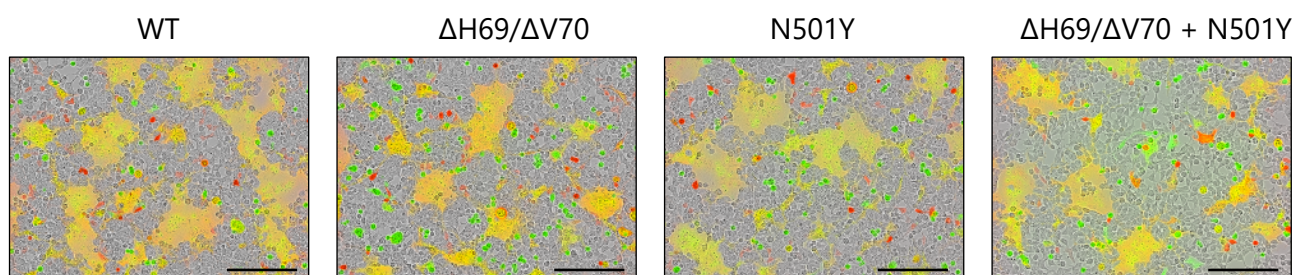
A



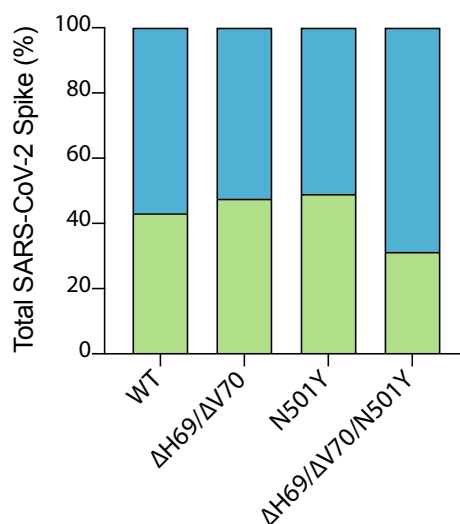
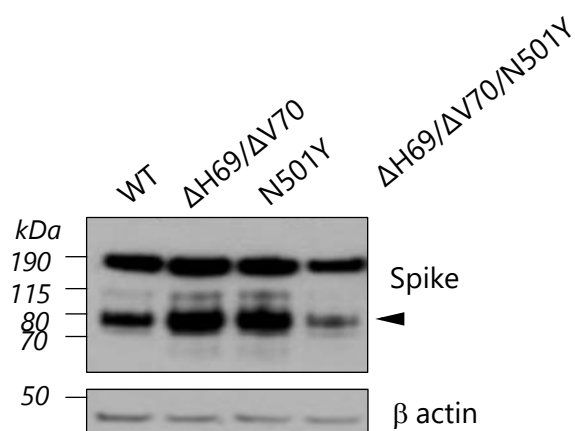
B



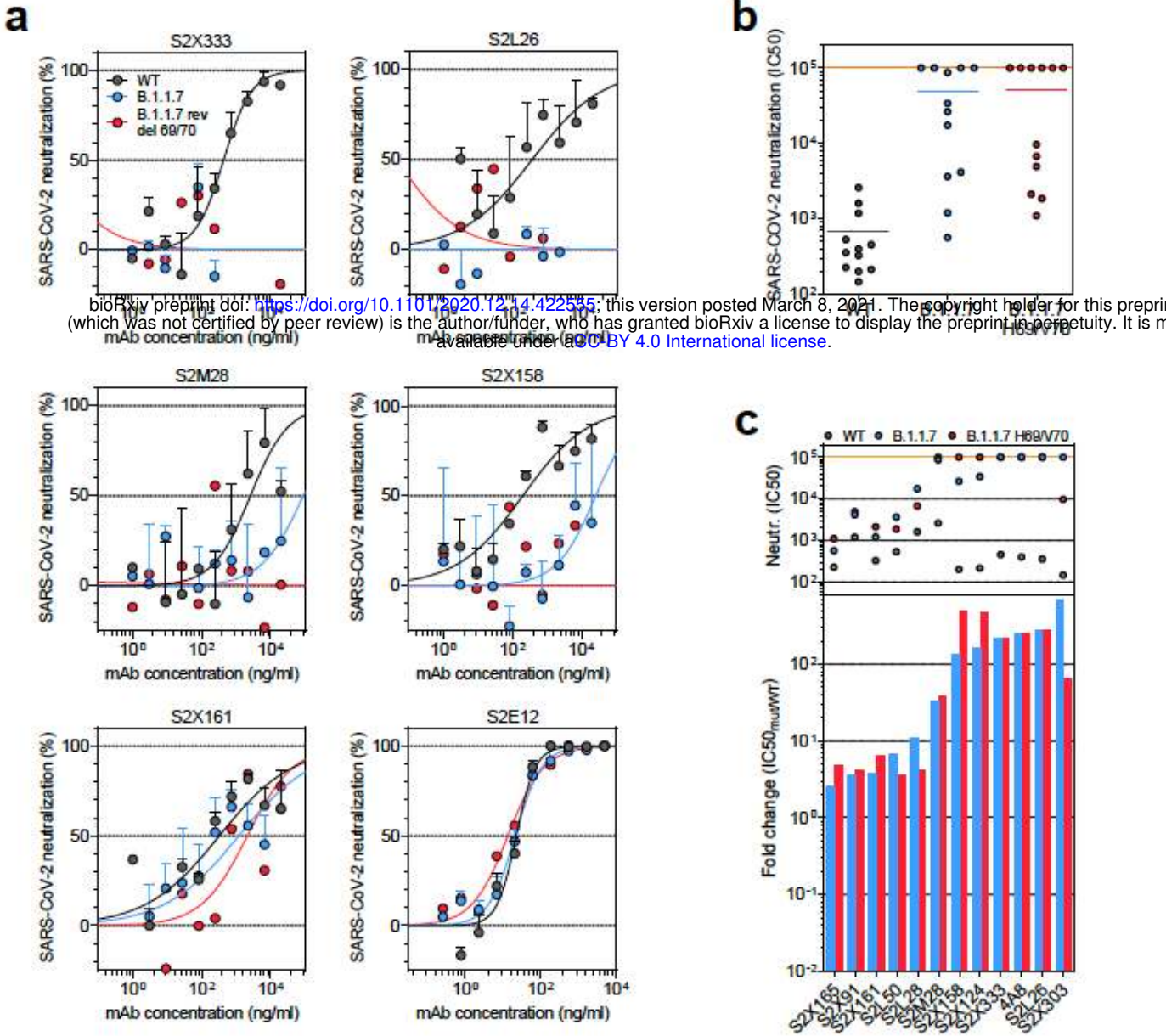
Supplementary Figure 6: Infectivity of spike Δ H69/V70 in a background of D614 (Wuhan). Purified pseudotypes, as indicated, were used to infect human ACE2 expressing HEK293 cells, with luciferase readings read at 72 hours post infection. Experiments were performed in biological quadruplicate with the mean and standard deviation plotted. Results are representative of experiments performed two times. Statistical significance was assessed using an unpaired t-test (ns; non-significant, ***; <0.005). Western blot of purified pseudotype virus. Spike and HIV pseudotype abundances were assessed using Flag and p24 antibodies, respectively. Relative spike expression was calculated by densitometry using Image J. Briefly, inverted pixel intensities for spike and p24 bands were first normalised to a background region of the gel. Spike protein intensities were then normalised to p24 intensity before mutant protein expression was calculated as a factor of wild-type protein. NE: no envelope/spike

A**B**

bioRxiv preprint doi: <https://doi.org/10.1101/2020.12.14.422555>; this version posted March 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

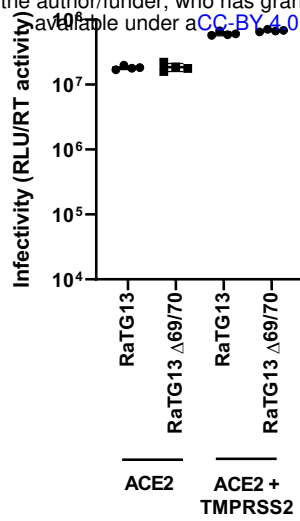


Supplementary Figure 7: Cell Fusogenicity in the presence of Spike mutants ΔH69/V70 and N501Y. **A.** Reconstructed images at 12 hours of 293T cells co-transfected with the indicated Spike mutants and mCherry expressing plasmid mixed with green dye-labelled Vero cells. Scale bars represent 200 μm. Green colour identifies the acceptor cells while red colour marks donor cells. Merged green-red colours indicate the syncytia. **C.** Representative western blot of cells transfected with the indicated Spike mutants (detected with anti-Spike antibody). The cleaved Spike identifies the S2 subunit and is indicated with the arrowhead. β actin is shown as loading control. **D.** Quantification analysis of cleaved (green bars) and uncleaved (blue bars) Spike shown in (C) normalised by β actin.



Supplementary Figure 8: Neutralisation and binding by a panel of NTD-specific mAbs against WT, B.1.1.7 and B.1.1.7 H69/V70 mutant SARS-CoV-2 viruses. A. Neutralisation of WT (black), B.1.1.7 (blue) and B.1.1.7 H69/V70 mutant (red) pseudotyped SARS-CoV-2-MLVs by 6 selected mAbs from one experiment. **B.** Neutralisation of WT, B.1.1.7 and B.1.1.7 H69/V70 SARS-CoV-2-MLVs by 13 mAbs targeting NTD. Shown are the mean IC₅₀ values (ng/ml) from one experiment. The higher the IC₅₀ the less sensitive the virus to antibodies. **C.** Neutralisation shown as mean IC₅₀ values (upper panel) and mean fold change of B.1.1.7 (blue) or B.1.1.7 H69/V70 (red) relative to WT (lower panel) of the 13 NTD mAbs tested. Lower panel shows IC₅₀ values from one experiment.

bioRxiv preprint doi: <https://doi.org/10.1101/2020.12.14.422555>; this version posted March 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



Supplementary Figure 9: Deletion of H69V70 in the bat sarbecovirus RaTG13 spike protein does not increase infectivity in cells overexpressing ACE2 or TMPRSS2. Single round infection by luciferase expressing lentivirus pseudotyped with RaTG13 Spike protein on 293T cells transduced with ACE2 or ACE2 and TMPRSS2. Experiments were performed in biological quadruplicate with the mean and standard deviation plotted. Results are a single experiment. tNE: no envelope.