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

Kemp SA[1,2,3*], Meng B[2,3*], Ferriera IATM[2,3*], Datir RP[2,3*], Harvey WT[4*], Collier DA[1,2,3], Lytras S[5], Papa G[6], , The COVID-19 Genomics UK (COG-UK) Consortium[7], Carabelli AM[3], Kenyon J[3,8], Lever AML[3,8], James LC[6], Robertson DL[5*], Gupta RK[2,3*]

[1]Division of Infection and Immunity, University College London, London, UK.

[2]Cambridge Institute of Therapeutic Immunology & Infectious Disease (CITIID), Cambridge, UK.

[3]Department of Medicine, University of Cambridge, Cambridge, UK.

[4]Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Glasgow, UK

[5]MRC - University of Glasgow Centre for Virus Research, Glasgow, UK.

[6]MRC – Laboratory of Molecular Biology, Cambridge, UK.

[7]https://www.cogconsortium.uk. Full list of consortium names and affiliations are in Appendix 1.

[8]Department of Microbiology and Immunology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

[9]Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

*Authors contributed equally to this work

Address for correspondence:

Ravindra K. Gupta

Cambridge Institute for Therapeutic Immunology and Infectious Diseases

Jeffrey Cheah Biomedical Centre

Puddicombe Way
Cambridge CB2 0AW, UK
Tel: +44 1223 331491
rkg20@cam.ac.uk

Key words: SARS-CoV-2; COVID-19; antibody escape; neutralising antibodies; mutation; evasion; resistance; fitness; evolution

Abstract

SARS-CoV-2 Spike amino acid replacements in the receptor binding domain (RBD) occur relatively frequently and some have a consequence for immune recognition. Here we report recurrent emergence and significant onward transmission of a six-nucleotide out of frame deletion in the S gene, which results in loss of two amino acids: H69 and V70. We show that ΔH69/V70 and other common NTD deletions occur at loop structures in RNA where polymerase activity is often compromised. In addition, the same two amino acid deletion is also observed in a set of closely related pangolin sarbecoviruses, but not in the bat sarbecovirus RaTG13. We report that in human infections ΔH69/V70 often co-occurs with the receptor binding motif amino acid replacements N501Y, N439K and Y453F, and in the latter two cases has followed the RBD mutation suggesting an epistatic interaction. One of the ΔH69/V70+ N501Y lineages, B.1.1.7, has undergone rapid expansion and includes eight other S gene mutations: RBD (N501Y and A570D), S1 (ΔH69/V70 and Δ144/145) and S2 (P681H, T716I, S982A and D1118H). *In vitro*, we show that ΔH69/V70 does not impact serum neutralisation across multiple convalescent sera. However, the ΔH69/V70 does increase infectivity of Spike and is able to compensate for infectivity defects induced by RBD mutations N501Y, N439K and Y453F. In addition replacement of H69 and V70 in B.1.1.7 Spike substantially reduces its infectivity. Mechanistically, we demonstrate that ΔH69/V70 is associated with two fold increased Spike incorporation into virions, though it does not increase cell-cell fusion. Based on our data ΔH69/V70 increases Spike infectivity, likely acting as a permissive mutation that allows acquisition of otherwise deleterious immune escape mutations. Enhanced surveillance for the ΔH69/V70 deletion with and without RBD mutations should be considered as a global priority not only as a marker for the B.1.1.7 variant, but potentially also for other emerging variants of concern.

Background

SARS-CoV-2's Spike surface glycoprotein engagement of human angiotensin-converting enzyme (hACE2) is essential for virus entry and infection[1], and the receptor is found in respiratory and gastrointestinal tracts[2]. Despite this critical interaction and the constraints it imposes, it appears the RBD, and particularly the receptor binding motif (RBM), is relatively tolerant to mutations[3,4], raising the real possibility of virus escape from vaccine-induced immunity and monoclonal antibody treatments. Spike mutants exhibiting reduced susceptibility to monoclonal antibodies have been identified in *in vitro* screens[3,5,6], and some of these mutations have been found in clinical isolates[7]. Due to the high levels of susceptibility of the human population to this virus, the acute nature of infections and limited use of vaccines to date there has been limited selection pressure placed SARS-CoV-2[8]; as a consequence few mutations that could alter antigenicity have increased significantly in frequency.

The unprecedented scale of whole genome SARS-CoV-2 sequencing has enabled identification and epidemiological analysis of transmission and surveillance, particularly in the UK[9] . As of January 10th, there were 346,447 SARS-CoV-2 sequences available in the GISAID database (https://www.gisaid.org/). However, geographic coverage is very uneven with approximately half of all sequences being provided by the United Kingdom. Indeed, as more countries continue to present data from samples collected up to six months prior, this may result in novel variants with altered biological or antigenic properties evolving and not being detected until they have already become established at high frequency.

Studying SARS-CoV-2 chronic infections can give insight into virus evolution that would require many chains of acute transmission to generate. This is because the majority of infections arise as a result of early transmission during pre or asymptomatic phases prior to peak adaptive responses, and virus adaptation not observed as the virus is usually cleared by the immune response[10,11]. We recently documented *de novo* emergence of antibody evasion mutations mediated by S gene mutations in an individual treated with convalescent plasma (CP)[12]. Dramatic changes in the prevalence of Spike variants **Δ**H69/V70 (an out of frame six-nucleotide deletion) and D796H variant followed repeated use of CP, while *in vitro*

the mutated ΔH69/V70+D796H variant displayed reduced susceptibility to CP using a pseudotyped lentivirus assay [12]. D796H alone resulted in more than five-fold loss of infectivity and the ΔH69/V70 partially rescued this defect. In addition, a chronically infected immune suppressed individual was recently reported in Russia with emergence of Y453F, along with ΔH69/V70[13]. Deletions in other parts of the N-Terminal Domain (NTD) have been reported to arise in chronic infections[7] and to reduce sensitivity to NTD-specific neutralising antibodies[14,15].

Here we analyse the variation in the global SARS-CoV-2 data and find ΔH69/V70 occurs independently often emerging after a significant RBD amino acid replacement such as N501Y, Y453F and N439K, that are known to either facilitate neutralising antibody escape or modulate affinity for the human ACE2 receptor. We show that ΔH69/V70 and other common NTD deletions occur at loop structures in RNA where polymerase activity is often compromised. Functionally, we reveal that ΔH69/V70 does increase Spike infectivity and compensates for infectivity loss resulting from RBD replacements N501Y and Y453F. Importantly, we find that ΔH69/V70 does not confer reduced susceptibility to convalescent sera.

Results

The deletion H69/V70 is present in over 25, 000 sequences worldwide (19,000 from the UK), 2.5% of the available data, and largely in Europe from where most of the sequences in GISAID are derived (Figure 1). ΔH69/V70 occurs in variants observed in different global lineages, representing multiple independent acquisitions of this SARS-CoV-2 deletion (Figure 1). While variants with deletions in this region of Spike are observed in GISAID[13], the earliest unambiguous sequence that includes the ΔH69/V70 was detected in Sweden in April 2020, an independent deletion event relative to other ΔH69/V70 variants. The prevalence of ΔH69/V70 has since increased since August 2020 (Figure 1B). Further analysis of sequences revealed, firstly, that single deletions of either 69 or 70 were uncommon and, secondly, some lineages of ΔH69/V70 alone were present, as well as ΔH69/V70 in the context of other mutations in Spike, specifically those in the RBM (Figure 1).

Polymerase processivity can be affected by physical factors in the template caused by RNA structural and sequence motifs and these can facilitate dissociation events. Stable helix loop motifs are associated with pausing/dissociation events in reverse transcriptase[16] . Since all nucleic acid polymerases have a common ancestor with homologous dNTP binding motifs and similar global structures[17-19] it is probable that all RNA polymerases use similar mechanisms for transcript termination[20]. Analysis of three deletions in the S protein coding sequence observed across new multiply mutated variants demonstrated that each occurred in the terminal loop of a helix loop motif (Figure 2). Although many regions of the structure have a wider structural ensemble of structures that the RNA can adopt, these helix loops were ones in which the ensemble is constrained into a very limited structural range, suggesting that these deletions occur in the more stable stem-loops.

To further probe the importance of this part of Spike molecule, we examined the 69/70 region of Spike in a set of other known *Sarbecoviruses*, relative to SARS-CoV-2 (Figure 3). We observe substantial variability in the region, resulting in frequent indels, with some viruses including SARS-CoV having 6-7 amino acid deletions (Figure 3B). This is indicative of plasticity in this protein region that could allow the sarbecoviruses to alter their Spike conformation. The second closest relative to SARS-CoV-2 for this region after RaTG13 is the cluster of 5 CoVs sampled in trafficked pangolins in the Guangxi province[21]. Looking at the 69/70 region in these virus sequences raises the interesting observation that one of the five viruses in the cluster, P2V, has amino acids 69H and 70L present, while the other four have a double amino acid deletion (Figure 3B). Given that SARS-CoV-2 and RaTG13 have the homologous HV insertion at these positions, one explanation is that the proximal common ancestor between SARS-CoV-2 and the Guangxi pangolin cluster had the insertion, which was then lost while circulating in the pangolin population, similar to what we now see with SARS-CoV-2 in humans. Yet, the fact that P2V was cultured in Vero E6 cells prior to sequencing (contrary to the other 4, sequenced directly from the pangolin sample) raises the possibility of this being an independent insertion, favoured as a cell line specific adaptation. Interestingly, the double amino acid indel in the pangolin viruses is in-frame in contrast to SARS-CoV-2 (e.g. lineage B.1.1.7, Figure 3C).

157 We next examined in greater detail the SARS-CoV-2 lineages where S gene mutations in the
158 RBD were identified at high frequency and where **Δ**H69/V70 co-occurs. For example, N439K,
159 an amino acid replacement reported to be defining variants increasing in numbers in Europe
160 and other regions[3] (Figure 1, Supplementary figure 1) now mostly co-occurs with **Δ**H69/V70.
161 N439K appears to have reduced susceptibility to some convalescent sera as well as
162 monoclonals targeting the RBD, whilst retaining affinity for ACE2 *in vitro*[3]. The second
163 significant cluster with **Δ**H69/V70 and RBD mutants involves Y453F, another Spike RBD
164 mutation that increases binding affinity to ACE2[4] and has been found to be associated with
165 mink-human infections[22]. In one SARS-CoV-2 mink-human sub-lineage, termed 'Cluster 5',
166 Y453F and **Δ**H69/V70 occurred with F486L, N501T and M1229I and was shown to have
167 reduced susceptibility to sera from recovered COVID-19 patients (https://files.ssi.dk/Mink-
168 cluster-5-short-report_AFO2). Y453F has been described as an escape mutation for mAb
169 REGN10933[23]. The **Δ**H69/V70 was first detected in the Y453F background on August 24th
170 2020 and thus far appears limited to Danish sequences (Figure 1, Supplementary Figure 2),
171 although an independent acquisition was recently reported within an individual, along with
172 **Δ**H69/V70 in an immune compromised individual with chronic infection[13]. A third lineage
173 containing the same **Δ**H69/V70 deletion has arisen with another RBD mutation N501Y
174 (Figure 1, Supplementary Figure 3A).
175
176 On closer inspection these UK sequences were part of a new lineage (B.1.1.7), termed VOC
177 202012/01 by Public Health England as they are associated with relatively high numbers of
178 infections (Supplementary Figure 3A). In addition to RBD N501Y + NTD **Δ**H69/V70 this new
179 variant had five further S mutations across the RBD (N501Y) and S2 (P681H, T716I, S982A
180 and D1118H), as well as NTD **Δ**144[24] (Supplementary Figure 3B). The variant has now been
181 identified in around 80 countries and regions. This lineage has a relatively long branch due
182 to 23 unique mutations (Supplementary figure 3B). The available sequences did not enable
183 us to determine whether the B1.1.7 mutations N501Y + **Δ**H69/V70 arose as a result of a
184 N501Y virus acquiring **Δ**H69/V70 or vice versa, though a sequence was identified with
185 N501Y + **Δ**H69/V70 and D1118H, indicating that as predicted N501Y + **Δ**H69/V70 are
186 proximal mutational events. Based on its location N501Y is expected to escape antibodies
187 similar to COV2-2499[5] (Figure 3C, D). Sequences with N501Y alone were isolated both in the

188 UK, Brazil, and USA in April 2020, and in South Africa (including the recently reported
189 described multi-mutated lineage B.1.351, Figure 1)[25].
190
191 We hypothesised that ΔH69/V70 is an antibody escape mechanism. We first examined the
192 protein structural context of ΔH69/V70 for clues regarding alterations in epitopes (Figure
193 4A, B). In the absence of experimentally derived structural data for ΔH69/V70, the protein
194 structure of the NTD possessing the double deletion was modelled in silico. The ΔH69/V70
195 deletion was predicted to alter the conformation of a protruding loop comprising residues
196 69 to 76, pulling it in towards the NTD (Figure 4A). In the post-deletion structural model, the
197 positions of the alpha carbons of residues either side of the deleted residues, Ile68 and
198 Ser71, were each predicted to occupy positions 2.9Å from the positions of His69 and Val70
199 in the pre-deletion structure. Concurrently, the positions of Ser71, Gly72, Thr73, Asn74 and
200 Gly75 are predicted to have changed by 6.5Å, 6.7Å, 6.0Å, 6.2Å and 8Å, respectively, with the
201 overall effect of these residues moving inwards, resulting in a less dramatically protruding
202 loop. This change in the surface of Spike would be consistent with antibody evasion. To test
203 this we explored whether ΔH69/V70 conferred reduced susceptibility to neutralising
204 antibodies in sera from ten recovered individuals. We performed serial dilutions of sera
205 before mixing with viral particles pseudotyped with Spike proteins with and without
206 ΔH69/V70 (with virus input normalised for infectivity). We plotted infection of target cells
207 as a function of serum dilution across the ten serum samples (Figure 4C). All sera
208 demonstrated clear titratable neutralisation of both wild type and ΔH69/V70 virus. There
209 was no shift in the titration curves and therefore no change in susceptibility to serum
210 neutralisation for ΔH69/V70 relative to wild type. These data show that ΔH69/V70 is
211 unlikely to represent an immune escape mechanism.
212
213 Given the association between ΔH69/V70 and RBD mutations and data from an earlier
214 report in chronic infection[12], we hypothesised that this deletion might alternatively enhance
215 virus infectivity and compensate for potential loss of infectivity due to receptor binding
216 motif mutations Y453F, N510Y and N439K that interact with ACE 2 (Figure 5A-C). In the
217 absence of virus isolates we used a lentiviral pseudotyping approach to test the impact of
218 ΔH69/V70 on virus Spike protein mediated infection. A D614G bearing Spike protein
219 expressing DNA plasmid (wild type, WT) was co-transfected in HEK 293T producer cell lines

220 along with plasmids encoding lentiviral capsid and genome for luciferase. Mutant Spike
221 plasmids bearing RBD mutations N501Y, Y453F, or N439K both with and without **Δ**H69/V70
222 were constructed by site directed mutagenesis. Plasmids were expressed in 293T virus
223 producer cells and infectious titres measured in supernatants from cells by infection of
224 target cells expressing ACE2 receptor. The infectious titres were adjusted for virus input as
225 determined by reverse transcriptase activity in virions (Figure 5D). RBD mutations N501Y,
226 N439K and Y453F all reduced infectivity of Spike relative to WT. **Δ**H69/V70 Spike
227 pseudotyped virus had higher infectivity compared to wild type (Figure 5D). Similarly,
228 addition of **Δ**H69/V70 increased the infectivity of each of the RBD mutations (Figure 5D).
229
230 We used our observations to predict that replacement of H69 and V70 would impair the
231 infectivity of B.1.1.7 (which naturally contains the **Δ**H69/V70 deletion). We compared the
232 infectivity of B.1.1.7 Spike versus B.1.1.7 without the **Δ**H69/V70 deletion in our
233 pseudotyping system. As expected we observed a significant reduction in infectivity for
234 viruses where the H69 and V70 had been re-inserted (Figure 5E). We hypothesised that the
235 mechanism of increased infectivity for **Δ**H69/V70 might involve increased Spike
236 incorporation. We performed western blotting of pseudovirions produced by plasmids
237 expressing lentivirus along with Spike WT with or without **Δ**H69/V70 (Figure 5F). This
238 analysis showed approximately two-fold increased expression of **Δ**H69/V70 Spike versus WT
239 Spike in virus containing cell supernatants where supernatant input was normalised for
240 reverse transcriptase activity (Figure 5F).
241
242 In order to further probe potential mechanisms at play we used a cell fusion assay (Figure
243 6A-D) to monitor kinetics of cell fusion. Previous reports have shown that SARS-CoV-2 Spike
244 protein possesses high fusogenic activity and is able to trigger the formation of large multi-
245 nucleated cells (named syncytia) *in vitro* and *in vivo*[26] (potentially providing an additional
246 and a more rapid route for virus disseminating among neighbour cells). To understand
247 whether the higher infectivity of **Δ**H69/V70 Spike might be related to an increased ability to
248 trigger cell-cell fusion between different cell types, we ectopically overexpressed SARS-CoV-
249 2 Spike variants together with the mCherry fluorescent protein in 293T cells and label Vero
250 cells with a green fluorescent dye. Mixing both cell types and measuring the merged green
251 and red fluorescence allows precise quantification of cell-cell fusion kinetics (Figure 6A, B).

252 We did not find a significant difference in Spike protein expression (Figure 6C, D). Syncytia
253 formation kinetics were not significantly different for ΔH69/V70 and N501Y SARS-CoV-2
254 Spike mutants tested (Figure 6B).
255
256
257
258 **Discussion**
259 We have presented data demonstrating multiple, independent, and circulating lineages of
260 SARS-CoV-2 variants bearing a Spike ΔH69/V70. This recurring deletion spanning six
261 nucleotides is due to an out of frame deletion of six nucleotides, and occurs in the terminal
262 loop of a helix loop motif within the RNA structure, as do other NTD deletions observed in
263 new variants such as the UK B.1.1.7, South African B.1.351 and the Brazilian P.1. ΔH69/V70
264 has frequently followed receptor binding amino acid replacements (N501Y, N439K and
265 Y453F that have been shown to increase binding affinity to hACE2 and reduce binding with
266 monoclonal antibodies), and is specifically found in the B.1.1.7 variant known to have higher
267 transmissibility[32].
268
269 The presence of sequence at site 69/70 appears to be unique to SARS-CoV-2, the closest bat
270 sarbecovirus, RaTG13, and one of the Guangxi pangolin sequences. Although we cannot
271 delineate between the indel being lost in the pangolin host, or gained after culturing the
272 virus, both scenarios are suggestive of functional importance at this protein region.
273
274 The ΔH69/V70 deletion was recently shown to emerge following treatment with
275 convalescent plasma[12]. In a second case of multiple mutations in context of immune
276 suppression the ΔH69/V70 deletion occurred with Y453F, though without biological
277 characterisation. We show here experimentally that the ΔH69/V70 deletion is indeed able
278 to rescue an infectivity defect induced by RBD mutations N501Y, N439K and Y453F,
279 potentially explaining why the deletion is often observed after these RBD mutation in SARS-
280 CoV-2 global phylogenies. More importantly perhaps, we demonstrate that the B.1.1.7 Spike
281 loses substantial infectivity when the H69/V70 amino acids are replaced, suggesting that the
282 ΔH69/V70 may contribute to the transmissibility of B.1.1.7. We show that the mechanism of
283 enhanced infectivity is associated with greater Spike incorporation into virions where the

ΔH69/V70 deletion is present. Therefore, ΔH69/V70 can act as a 'permissive' mutation that enhances infection[28], with the potential to enhance the ability of SARS-CoV-2 to tolerate immune escape mutations that would have otherwise significantly reduced viral fitness.

The potential for SARS-CoV-2 to evolve and fix mutations is exemplified by D614G, an amino acid replacement in S1 that alters linkages between S1 and S2 subunits on adjacent protomers as well as RBD orientation, infectivity, and transmission[28-30]. The example of D614G also demonstrates that mechanisms directly impacting important biological processes can be indirect. Similarly, a number of possible mechanistic explanations may underlie ΔH69/V70. For example, the fact that it sits on an exposed surface and is estimated to alter the conformation of a particularly exposed loop might be suggestive of immune interactions and escape, though we have presented data to show that ΔH69/V70 did not impact sensitivity of Spike to neutralising antibodies in serum from a group of recovered individuals. However, allosteric interactions, as postulated for D614G, could lead to the higher infectivity of Spike pseudotyped viruses with ΔH69/V70[31]. It is also possible that the increased infectivity of relates to interactions with receptors other than ACE2 or TMPRSS2, for example L- SIGN/ DC- SIGN. [31]Notably, data on increased infectivity conferred by D614G using similar pseudotyped viruses[29] was also observed in whole virus and translated to increased viral load and transmission in animal models[30]. It would be important to investigate whether ΔH69/V70 is associated with higher viral loads or increased transmission, though such epidemiological studies are highly complex and prone to confounding/bias.

The finding of a lineage (B.1.1.7), termed VOC 202012/01, bearing seven S gene mutations across the RBD (N501Y, A570D), S1 (ΔH69/V70 and Δ144) and S2 (P681H, T716I, S982A and D1118H) in UK requires urgent experimental characterisation. The detection of a high number of novel mutations suggests this lineage has either been introduced from a geographic region with very poor sampling or viral evolution may have occurred in a single individual in the context of a chronic infection[12]. The emergence of variants with higher numbers of mutations so far in the UK and South Africa may herald an era of re-infection and threaten future vaccine efficacy if left unchecked.

The detection and surveillance of B.1.1.7 has been facilitated in the UK by the phenomenon of SGTF (S gene target failure) due to primers in the Thermofisher SARS-CoV-2 diagnostic qPCR assay used by a significant number of testing facilities. The S gene target (binding in the region of H69/V70) is one of three and therefore a marker for the spread of B.1.1.7 has been tracked by the loss of signal in the S gene target[32]. However a recent report from the US cautioned against use of SGTF as a sole marker for B.1.1.7 as there is a significant ΔH69/V70 lineage without other mutations in Spike circulating in the US[33].

Given the emergence of multiple clusters of variants carrying RBD mutations and the ΔH69/V70 deletion, limitation of transmission takes on a renewed urgency. Continued emphasis on testing/tracing, social distancing and mask wearing are essential, with investment in other novel methods to limit transmission[34]. In concert, comprehensive vaccination efforts in the UK and globally should be accelerated in order to further limit transmission and acquisition of further mutations. Finally, detection of the deletion and other key mutations by rapid diagnostics should be a research priority as such tests could be used as a proxy for antibody escape mutations to inform surveillance at global scale.

Limitations

The laboratory virology aspects of this study were conducted with pseudoviruses rather than replication competent viruses.

Acknowledgements

RKG is supported by a Wellcome Trust Senior Fellowship in Clinical Science (WT108082AIA). COG-UK is supported by funding from the Medical Research Council (MRC) part of UK Research & Innovation (UKRI), the National Institute of Health Research (NIHR) and Genome Research Limited, operating as the Wellcome Sanger Institute. SAK is supported by the Bill and Melinda Gates Foundation via PANGEA grant: OPP1175094. DLR is funded by the MRC (MC UU 1201412). WH is funded by the MRC (MR/R024758/1). We thank Dr James Voss for the kind gift of HeLa cells stably expressing ACE2. SL is funded by Medical Research Council MC_UU_12014/12. This study was also partly funded by Rosetrees Trust. AML is

Conflicts of interest

RKG has received consulting fees from UMOVIS lab, Gilead Sciences and ViiV Healthcare, and a research grant from InvisiSmart Technologies.

Methods

Phylogenetic Analysis

All available full-genome SARS-CoV-2 sequences were downloaded from the GISAID database (http://gisaid.org/)[35] on 26th November. Duplicate and low-quality sequences (>5% N regions) were removed, leaving a dataset of 194,265 sequences with a length of >29,000bp. All sequences were realigned to the SARS-CoV-2 reference strain MN908947.3, using MAFFT v7.475 with automatic flavour selection and the --keeplength --addfragments options[36]. Major SARS-CoV-2 clade memberships were assigned to all sequences using the Nextclade server v0.10 (https://clades.nextstrain.org/).

Maximum likelihood phylogenetic trees were produced using the above curated dataset using IQ-TREE v2.1.2[37]. Evolutionary model selection for trees were inferred using ModelFinder[38] and trees were estimated using the GTR+F+I model with 1000 ultrafast bootstrap replicates[39]. All trees were visualised with Figtree v.1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) and ggtree v1.14.6 rooted on the SARS-CoV-2 reference sequence and nodes arranged in descending order. Nodes with bootstraps values of <50 were collapsed using an in-house script.

To reconstruct a phylogeny for the 69/70 Spike region of the 17 *Sarbecoviruses* examined in Fig 2, Rdp5[40] was used on the codon Spike alignment to determine the region between amino acids 1 and 256 as putatively non-recombinant. A tree was reconstructed using the nucleotide alignment of this region under a GTR+Γ substitution model with RAxML-NG[41]. Node support was calculated with 1000 bootstraps. Alignment visualisation was done using BioEdit[42].

RNA secondary structure modelling.

2990 nucleotides centred around the spike protein amino acids 69-70 from SARS-CoV2 sequence from an individual[12] were aligned in CLUSATL-Omega (nucleotides 20277-23265 of

380 the Wuhan isolate MN908947.3) and a consensus structure was generated using
381 RNAalifold[43]).
382
383 *Structural modelling*
384 The structure of the post-deletion NTD (residues 14-306) was modelled using I-TASSER[44], a
385 method involving detection of templates from the protein data bank, fragment structure
386 assembly using replica-exchange Monte Carlo simulation and atomic-level refinement of
387 structure using a fragment-guided molecular dynamics simulation. The structural model
388 generated was aligned with the spike structure possessing the pre-deletion conformation of
389 the 69-77 loop (PDB 7C2L[45]) using PyMOL (Schrödinger). Figures prepared with PyMOL using
390 PDBs 7C2L, 6M0J[46], 6ZGE28 and 6ZGG[47].
391
392 *Pseudotype virus preparation*
393 Viral vectors were prepared by transfection of 293T cells by using Fugene HD transfection
394 reagent (Promega). 293T cells were transfected with a mixture of 11ul of Fugene HD, 1µg of
395 pCDNAΔ19Spike-HA, 1ug of p8.91 HIV-1 gag-pol expression vector22,23, and 1.5µg of
396 pCSFLW (expressing the firefly luciferase reporter gene with the HIV-1 packaging signal).
397 Viral supernatant was collected at 48 and 72h after transfection, filtered through 0.45um
398 filter and stored at -80˚C as previously described24. Infectivity was measured by luciferase
399 detection in target 293T cells transfected with TMPRSS2 and ACE2.
400
401 *Standardisation of virus input by SYBR Green-based product-enhanced PCR assay (SG-PERT)*
402 The reverse transcriptase activity of virus preparations was determined by qPCR using a
403 SYBR Green-based product-enhanced PCR assay (SG-PERT) as previously described[48]. Briefly,
404 10-fold dilutions of virus supernatant were lysed in a 1:1 ratio in a 2x lysis solution (made up
405 of 40% glycerol v/v 0.25% Trition X-100 v/v 100mM KCl, RNase inhibitor 0.8 U/ml, TrisHCL
406 100mM, buffered to pH7.4) for 10 minutes at room temperature.
407
408 12µl of each sample lysate was added to thirteen 13µl of a SYBR Green master mix
409 (containing 0.5µM of MS2-RNA Fwd and Rev primers, 3.5pmol/ml of MS2-RNA, and
410 0.125U/µl of Ribolock RNAse inhibitor and cycled in a QuantStudio. Relative amounts of
411 reverse transcriptase activity were determined as the rate of transcription of bacteriophage

412 MS2 RNA, with absolute RT activity calculated by comparing the relative amounts of RT to
413 an RT standard of known activity.
414
415 *Cell-cell fusion assay*
416 Cell fusion assay was carried out as previously described (Papa et al. 2021). Briefly, Vero
417 cells and 293T cells were seeded at 80% confluency in a 24 multiwell plate. 293T cells were
418 co-transfected with 1.5 μg of Spike expression plasmids in pCDNA3 and 0.5 μg pmCherry-N1
419 using Fugene 6 and following the manufacturer's instructions (Promega). Vero cells were
420 treated with CellTracker™ Green CMFDA (5-chloromethylfluorescein diacetate) (Thermo
421 Scientific) for 20 minutes. 293T cells were then detached 5 hours post transfection, mixed
422 together with the green-labelled Vero cells, and plated in a 12 multiwell plate. Cell-cell
423 fusion was measured using an Incucyte and determined as the proportion of merged area to
424 green area over time. Data were then analysed using Incucyte software analysis. Data were
425 normalised to cells transfected only with mCherry protein and mixed with green labelled
426 Vero cells. Graphs were generated using Prism 8 software.
427
428 *Cells*
429 HEK 293T CRL-3216, Vero CCL-81 were purchased from ATCC and maintained in Dulbecco's
430 Modified Eagle Medium (DMEM) supplemented with 10% fetal calf serum (FCS), 100 U/ml
431 penicillin, and 100μg/ml streptomycin. All cells are regularly tested and are mycoplasma
432 free.
433
434 *Transfection*
435 HEK 293T cells were transfected with 1.5 μg of Spike expression plasmids in pCDNA3 and
436 lysed 18 hours post transfection. Cells were treated with Benzonase Nuclease (70664
437 Millipore) and boiled for 5 min. Samples were then run on 4%–12% Bis Tris gels and
438 transferred onto nitrocellulose membranes using an iBlot (Life Technologies).
439
440 Membranes were blocked for 1 hour in 5% non-fat milk in PBS + 0.1% Tween-20 (PBST) at
441 room temperature with agitation, incubated in primary antibody (anti-SARS-CoV-2 Spike,
442 which detects the S2 subunit of SARS-CoV-2 S (Invitrogen, PA1-41165)) diluted in 5% non-fat
443 milk in PBST for 2 hours at 4°C with agitation, washed four times in PBST for 5 minutes at

444 room temperature with agitation and incubated in secondary antibody (anti-rabbit HRP
445 conjugate, Invitrogen 31462), anti-bactin HRP (sc-47778) diluted in 5% non-fat milk in PBST
446 for 1 hour with agitation at room temperature. Membranes were washed four times in PBST
447 for 5 minutes at room temperature and imaged directly using a ChemiDoc MP imaging
448 system (Bio-Rad).
449
450 *Serum/plasma pseudotype neutralisation assay*
451 Spike pseudotype assays have been shown to have similar characteristics as neutralisation
452 testing using fully infectious wild type SARS-CoV-2[49].Virus neutralisation assays were
453 performed on 293T cell transiently transfected with ACE2 and TMPRSS2 using SARS-CoV-2
454 Spike pseudotyped virus expressing luciferase[50]. Pseudotyped virus was incubated with
455 serial dilution of heat inactivated human serum samples or convalescent plasma in duplicate
456 for 1h at 37˚C. Virus and cell only controls were also included. Then, freshly trypsinized 293T
457 ACE2/TMPRSS2 expressing cells were added to each well. Following 48h incubation in a 5%
458 CO_2 environment at 37°C, the luminescence was measured using Steady-Glo Luciferase
459 assay system (Promega).

460 Ethical approval for use of serum samples. Controls with COVID-19 were enrolled to the
461 NIHR BioResource Centre Cambridge under ethics review board (17/EE/0025).
462
463 **References**

464 1 Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable
465 bat origin. *nature* **579**, 270-273 (2020).
466 2 Sungnak, W. *et al.* SARS-CoV-2 entry factors are highly expressed in nasal epithelial
467 cells together with innate immune genes. *Nat Med* **26**, 681-687,
468 doi:10.1038/s41591-020-0868-6 (2020).
469 3 Thomson, E. C. *et al.* The circulating SARS-CoV-2 spike variant N439K maintains
470 fitness while evading antibody-mediated immunity. *bioRxiv* (2020).
471 4 Starr, T. N. *et al.* Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain
472 Reveals Constraints on Folding and ACE2 Binding. *Cell* **182**, 1295-+,
473 doi:10.1016/j.cell.2020.08.012 (2020).

474 5 Greaney, A. J. *et al.* Complete mapping of mutations to the SARS-CoV-2 spike
475 receptor-binding domain that escape antibody recognition. *Cell Host & Microbe*
476 (2020).
477 6 Starr, T. N. *et al.* Prospective mapping of viral mutations that escape antibodies used
478 to treat COVID-19. *bioRxiv* (2020).
479 7 Choi, B. *et al.* Persistence and Evolution of SARS-CoV-2 in an Immunocompromised
480 Host. *The New England journal of medicine* **383**, 2291-2293,
481 doi:10.1056/NEJMc2031364 (2020).
482 8 MacLean, O. A. *et al.* Natural selection in the evolution of SARS-CoV-2 in bats, not
483 humans, created a highly capable human pathogen. *BioRxiv* (2020).
484 9 consortiumcontact@cogconsortium.uk, C.-G. U. An integrated national scale SARS-
485 CoV-2 genomic surveillance network. *Lancet Microbe* **1**, e99-e100,
486 doi:10.1016/S2666-5247(20)30054-9 (2020).
487 10 Mlcochova, P. *et al.* Combined Point-of-Care Nucleic Acid and Antibody Testing for
488 SARS-CoV-2 following Emergence of D614G Spike Variant. *Cell Rep Med* **1**, 100099,
489 doi:10.1016/j.xcrm.2020.100099 (2020).
490 11 He, X. *et al.* Temporal dynamics in viral shedding and transmissibility of COVID-19.
491 *Nat Med* **26**, 672-675, doi:10.1038/s41591-020-0869-5 (2020).
492 12 Kemp, S. A. *et al.* Neutralising antibodies in Spike mediated SARS-CoV-2 adaptation.
493 *medRxiv*, 2020.2012.2005.20241927, doi:10.1101/2020.12.05.20241927 (2020).
494 13 Bazykin, G. *et al. Emergence of Y453F and Δ69-70HV mutations in a lymphoma*
495 *patient with long-term COVID-19*, <https://virological.org/t/emergence-of-y453f-
496 and-69-70hv-mutations-in-a-lymphoma-patient-with-long-term-covid-19/580>
497 (2021).
498 14 McCarthy, K. R. *et al.* Natural deletions in the SARS-CoV-2 spike glycoprotein drive
499 antibody escape. *bioRxiv*, 2020.2011.2019.389916, doi:10.1101/2020.11.19.389916
500 (2020).
501 15 McCallum, M. *et al.* N-terminal domain antigenic mapping reveals a site of
502 vulnerability for SARS-CoV-2. *bioRxiv*, 2021.2001.2014.426475,
503 doi:10.1101/2021.01.14.426475 (2021).

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

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

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

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

20 Reeder, R. H. & Lang, W. The mechanism of transcription termination by RNA poiymerase I. *Molecular microbiology* **12**, 11-15 (1994).

21 Lam, T. T. *et al.* Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature* **583**, 282-285, doi:10.1038/s41586-020-2169-0 (2020).

22 Munnink, B. B. O. *et al.* Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. *Science* (2020).

23 Baum, A. *et al.* Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science* **369**, 1014-1018, doi:10.1126/science.abd0831 (2020).

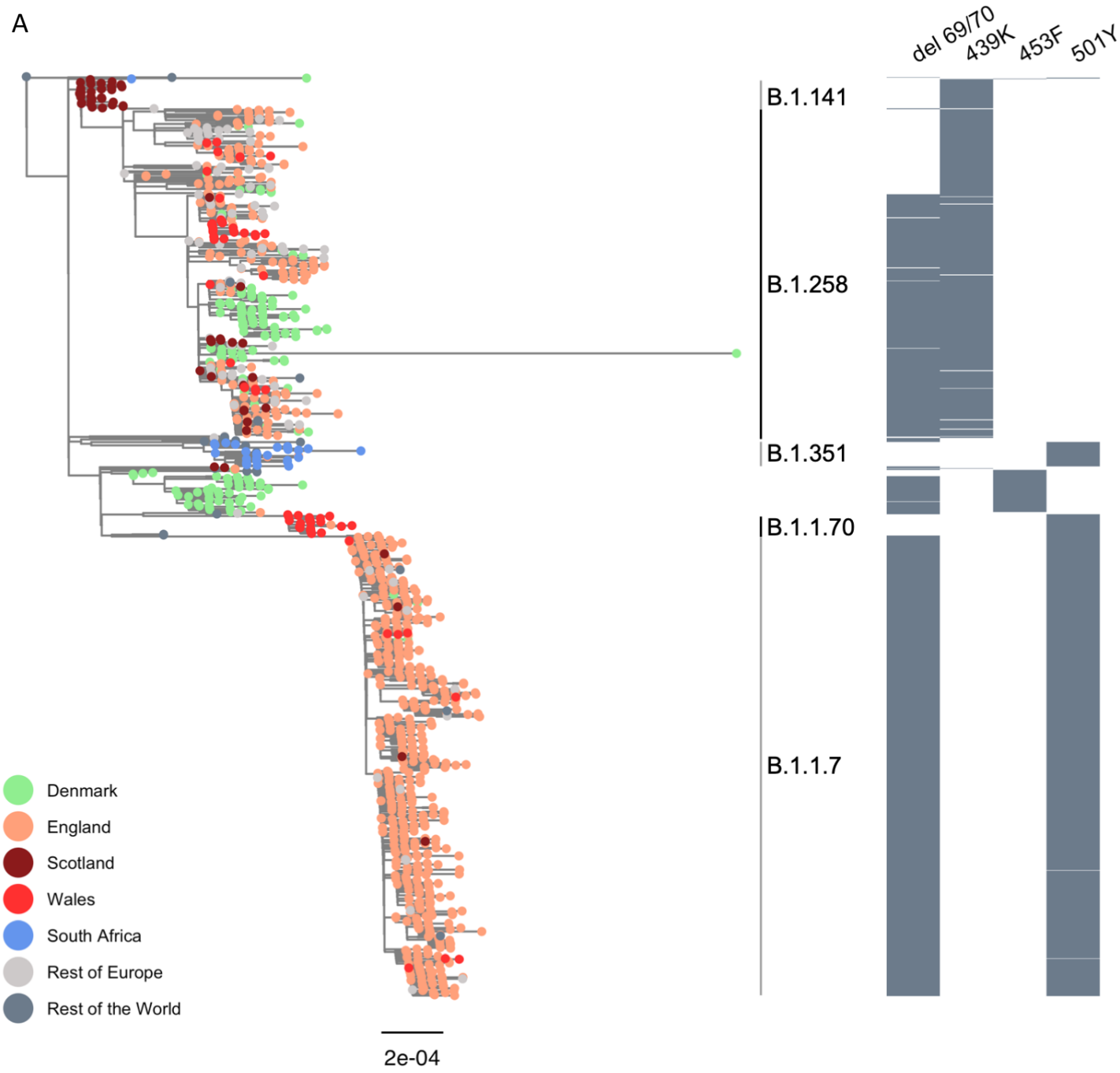
24 Rambaut A., L. N., Pybus O, Barclay W, Carabelli A. C., Connor T., Peacock T., Robertson D. L., Volz E., on behalf of COVID-19 Genomics Consortium UK (CoG-UK). *Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations*, <https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563> (2020).

25 Tegally, H. *et al.* 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. *medRxiv*, 2020.2012.2021.20248640, doi:10.1101/2020.12.21.20248640 (2020).

534 26 Papa, G. *et al.* Furin cleavage of SARS-CoV-2 Spike promotes but is not essential for
535 infection and cell-cell fusion. *bioRxiv*, 2020.2008.2013.243303,
536 doi:10.1101/2020.08.13.243303 (2020).
537 27 Kemp, S. *et al.* Recurrent emergence and transmission of a SARS-CoV-2 Spike
538 deletion ΔH69/V70. *bioRxiv*, 2020.2012.2014.422555,
539 doi:10.1101/2020.12.14.422555 (2021).
540 28 Korber, B. *et al.* Tracking changes in SARS-CoV-2 Spike: evidence that D614G
541 increases infectivity of the COVID-19 virus. *Cell* **182**, 812-827. e819 (2020).
542 29 Yurkovetskiy, L. *et al.* Structural and Functional Analysis of the D614G SARS-CoV-2
543 Spike Protein Variant. *Cell* **183**, 739-751 e738, doi:10.1016/j.cell.2020.09.032 (2020).
544 30 Hou, Y. J. *et al.* SARS-CoV-2 D614G variant exhibits efficient replication ex vivo and
545 transmission in vivo. *Science* **370**, 1464-1468, doi:10.1126/science.abe8499 (2020).
546 31 Soh, W. T. *et al.* The N-terminal domain of spike glycoprotein mediates SARS-CoV-2
547 infection by associating with L-SIGN and DC-SIGN. *bioRxiv*, 2020.2011.2005.369264,
548 doi:10.1101/2020.11.05.369264 (2020).
549 32 Volz, E. *et al.* Transmission of SARS-CoV-2 Lineage B.1.1.7 in England: Insights from
550 linking epidemiological and genetic data. *medRxiv*, 2020.2012.2030.20249034,
551 doi:10.1101/2020.12.30.20249034 (2021).
552 33 Larsen, B. B. & Worobey, M. *Identification of a novel SARS-CoV-2 Spike 69-70*
553 *deletion lineage circulating in the United States*,
554 <https://virological.org/t/identification-of-a-novel-sars-cov-2-spike-69-70-deletion-
555 lineage-circulating-in-the-united-states/577> (2020).
556 34 Mlcochova, P. *et al.* Extended in vitro inactivation of SARS-CoV-2 by titanium dioxide
557 surface coating. *bioRxiv* (2020).
558 35 Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data - from
559 vision to reality. *Euro surveillance : bulletin Europeen sur les maladies transmissibles*
560 *= European communicable disease bulletin* **22**, 30494, doi:10.2807/1560-
561 7917.ES.2017.22.13.30494 (2017).
562 36 Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7:
563 improvements in performance and usability. *Mol Biol Evol* **30**, 772-780,
564 doi:10.1093/molbev/mst010 (2013).

565 37 Minh, B. Q. *et al.* IQ-TREE 2: New Models and Efficient Methods for Phylogenetic
566 Inference in the Genomic Era. *Mol Biol Evol* **37**, 1530-1534,
567 doi:10.1093/molbev/msaa015 (2020).
568 38 Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S.
569 ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*
570 **14**, 587-589, doi:10.1038/nmeth.4285 (2017).
571 39 Minh, B. Q., Nguyen, M. A. & von Haeseler, A. Ultrafast approximation for
572 phylogenetic bootstrap. *Mol Biol Evol* **30**, 1188-1195, doi:10.1093/molbev/mst024
573 (2013).
574 40 Martin, D. P., Murrell, B., Golden, M., Khoosal, A. & Muhire, B. RDP4: Detection and
575 analysis of recombination patterns in virus genomes. *Virus evolution* **1** (2015).
576 41 Kozlov, A. M., Darriba, D., Flouri, T., Morel, B. & Stamatakis, A. RAxML-NG: a fast,
577 scalable and user-friendly tool for maximum likelihood phylogenetic inference.
578 *Bioinformatics (Oxford, England)* **35**, 4453-4455 (2019).
579 42 Hall, T., Biosciences, I. & Carlsbad, C. BioEdit: an important software for molecular
580 biology. *GERF Bull Biosci* **2**, 60-61 (2011).
581 43 Bernhart, S. H., Hofacker, I. L., Will, S., Gruber, A. R. & Stadler, P. F. RNAalifold:
582 improved consensus structure prediction for RNA alignments. *BMC bioinformatics* **9**,
583 1-13 (2008).
584 44 Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein
585 structure and function prediction. *Nat Protoc* **5**, 725-738, doi:10.1038/nprot.2010.5
586 (2010).
587 45 Chi, X. *et al.* A neutralizing human antibody binds to the N-terminal domain of the
588 Spike protein of SARS-CoV-2. *Science* **369**, 650-655, doi:10.1126/science.abc6952
589 (2020).
590 46 Lan, J. *et al.* Structure of the SARS-CoV-2 spike receptor-binding domain bound to
591 the ACE2 receptor. *Nature* **581**, 215-220, doi:10.1038/s41586-020-2180-5 (2020).
592 47 Wrobel, A. G. *et al.* SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform
593 on virus evolution and furin-cleavage effects. *Nat Struct Mol Biol* **27**, 763-767,
594 doi:10.1038/s41594-020-0468-7 (2020).

595 48 Vermeire, J. *et al.* Quantification of reverse transcriptase activity by real-time PCR as
596 a fast and accurate method for titration of HIV, lenti- and retroviral vectors. *PloS one*
597 **7**, e50859-e50859, doi:10.1371/journal.pone.0050859 (2012).
598 49 Schmidt, F. *et al.* Measuring SARS-CoV-2 neutralizing antibody activity using
599 pseudotyped and chimeric viruses. 2020.2006.2008.140871,
600 doi:10.1101/2020.06.08.140871 %J bioRxiv (2020).
601 50 Mlcochova, P. *et al.* Combined point of care nucleic acid and antibody testing for
602 SARS-CoV-2 following emergence of D614G Spike Variant. *Cell Rep Med*, 100099,
603 doi:10.1016/j.xcrm.2020.100099 (2020).
604



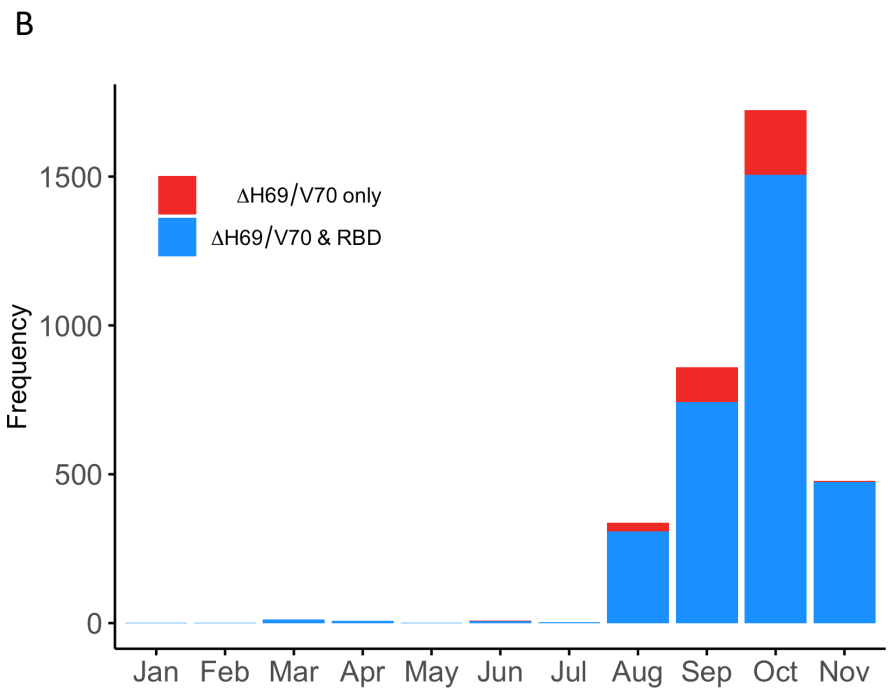


Figure 1. A. Sub-sampled global phylogeny of SARS-CoV-2 whole genome sequences highlighting those with specific mutations in Spike: ΔH69/V70, N439K, Y453F and N501Y. Tree tips are coloured by geographic region (see key). Grey bars on the right show the presence or absence of the deletion H69/V70 and amino acid variants N439K, Y453F, and N501Y. Lineages from Rambaut et al. 2020 are shown. **Number of new occurrences of SARS-CoV-2 sequences with the ΔH69/V70 deletion.** Frequency of the ΔH69/V70 deletion from the GISAID database (accessed 16th Dec 2020) by reporting country and sampling date: **B.** Worldwide carriage of ΔH69/V70; most sequences preferentially carry RBD mutations alongside the deletion.

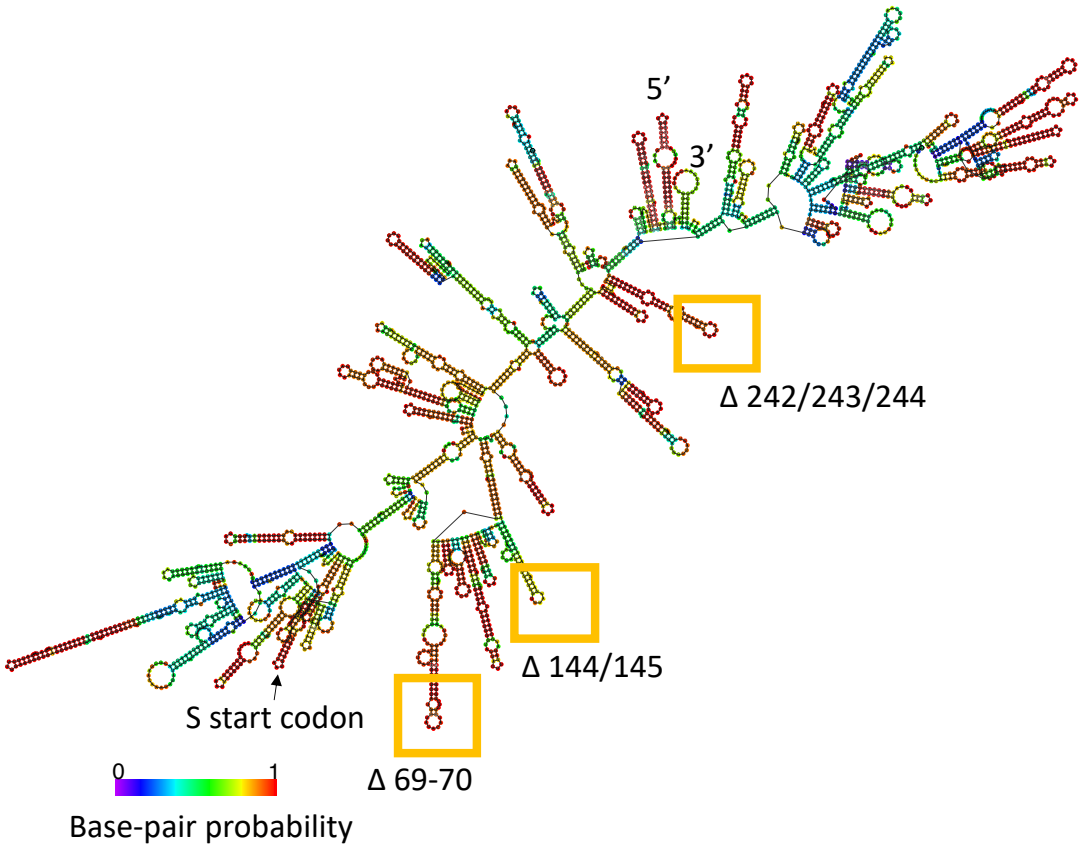


Figure 2: The positions of common deletion mutations on the RNA structure of the Spike Δ69-70 region of the gRNA. The optimal secondary structure was generated from a consensus alignment of human SARS-CoV2 RNAs using RNAalifold. Figure shows nucleotides 20277-23265. Base-pair probability, representative of the breadth of the structural ensemble that could be adopted by the RNA, is shown in colour according to the key.

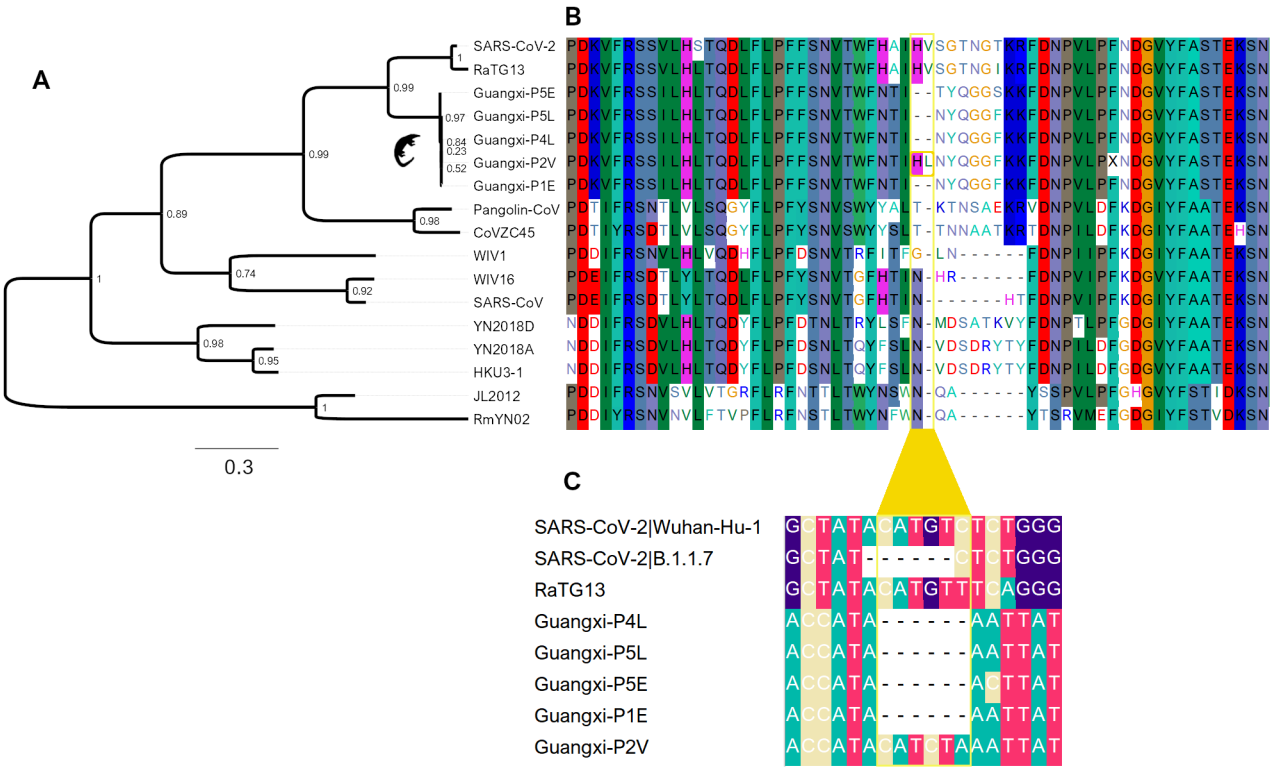


Figure 3. Comparison of the H69/V70 deletion site to other *Sarbecoviruses*. A. phylogeny for the Spike peptide region 1-256 **B**. protein sequences from 17 Sarbecoviruses, including SARS-CoV-2 (Wuhan-Hu-1) and SARS-CoV (HSZ-Cc), with distinct genotypes at the Spike region around amino acid positions 69 and 70 (highlighted in yellow box). The 69/70 HL insertion in the P2V sequence from the Guangxi pangolin virus cluster is highlighted. **C**. The nucleotide alignment between SARS-CoV-2 Wuhan-Hu-1, B.1.1.7, the bat sarbecovirus RaTG13 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.

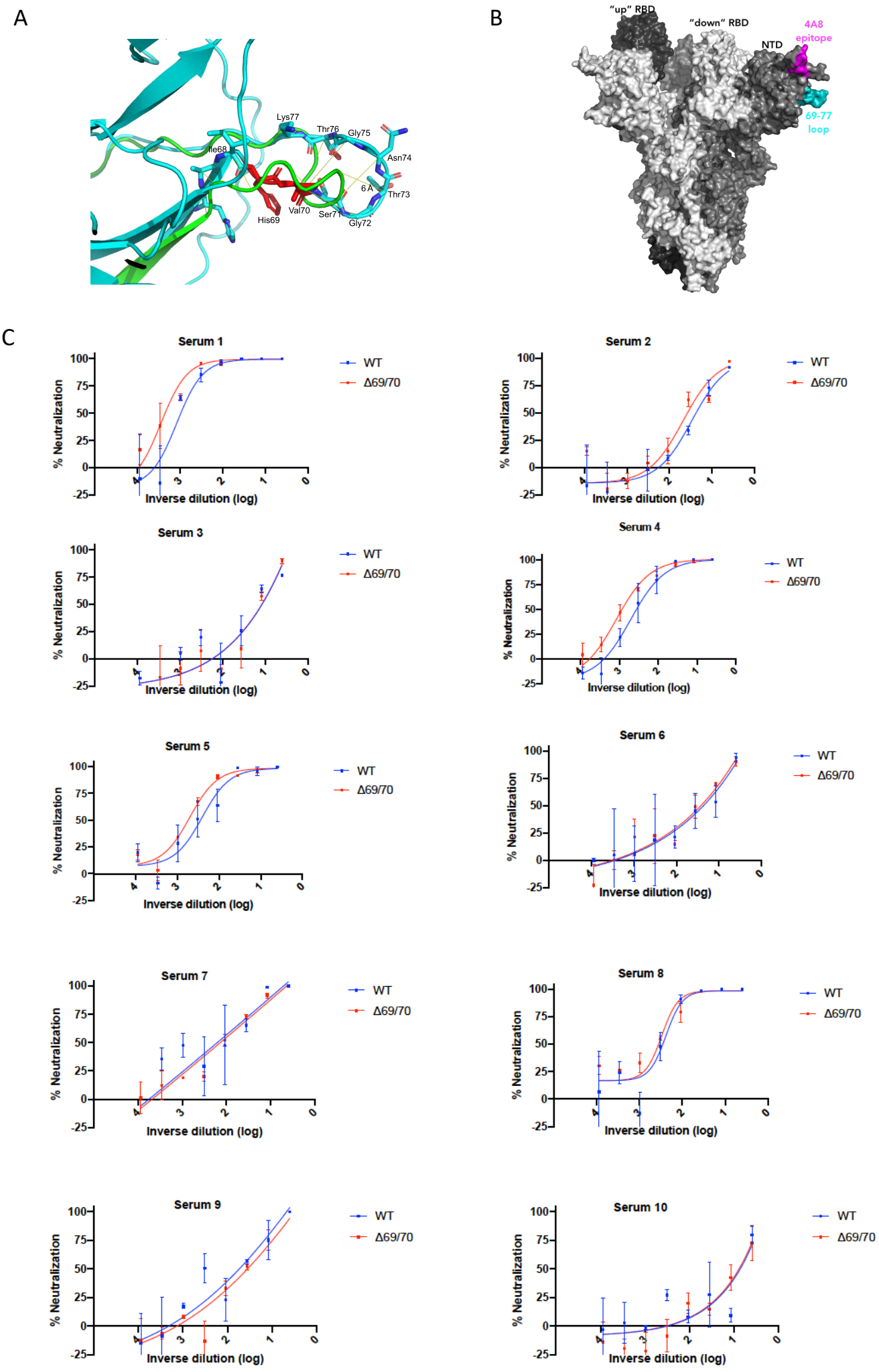


Figure 4. Structural Characteristics of viruses carrying deletion at H69/V70 and sensitivity to neutralising antibodies. **A.** Prediction of conformational change in the spike N-terminal domain due to deletion of residues His69 and Val70. The pre-deletion structure is shown in cyan, except for residues 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. **B.** 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 described by Chi et al. is shown in magenta. B. **Neutralization of pseudovirus virus bearing Spike ΔH69/V70 and wild type (all In Spike D614G background) by convalescent sera from 10 donors.** 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. Data points represent means of technical replicates and error bars represent standard deviation.

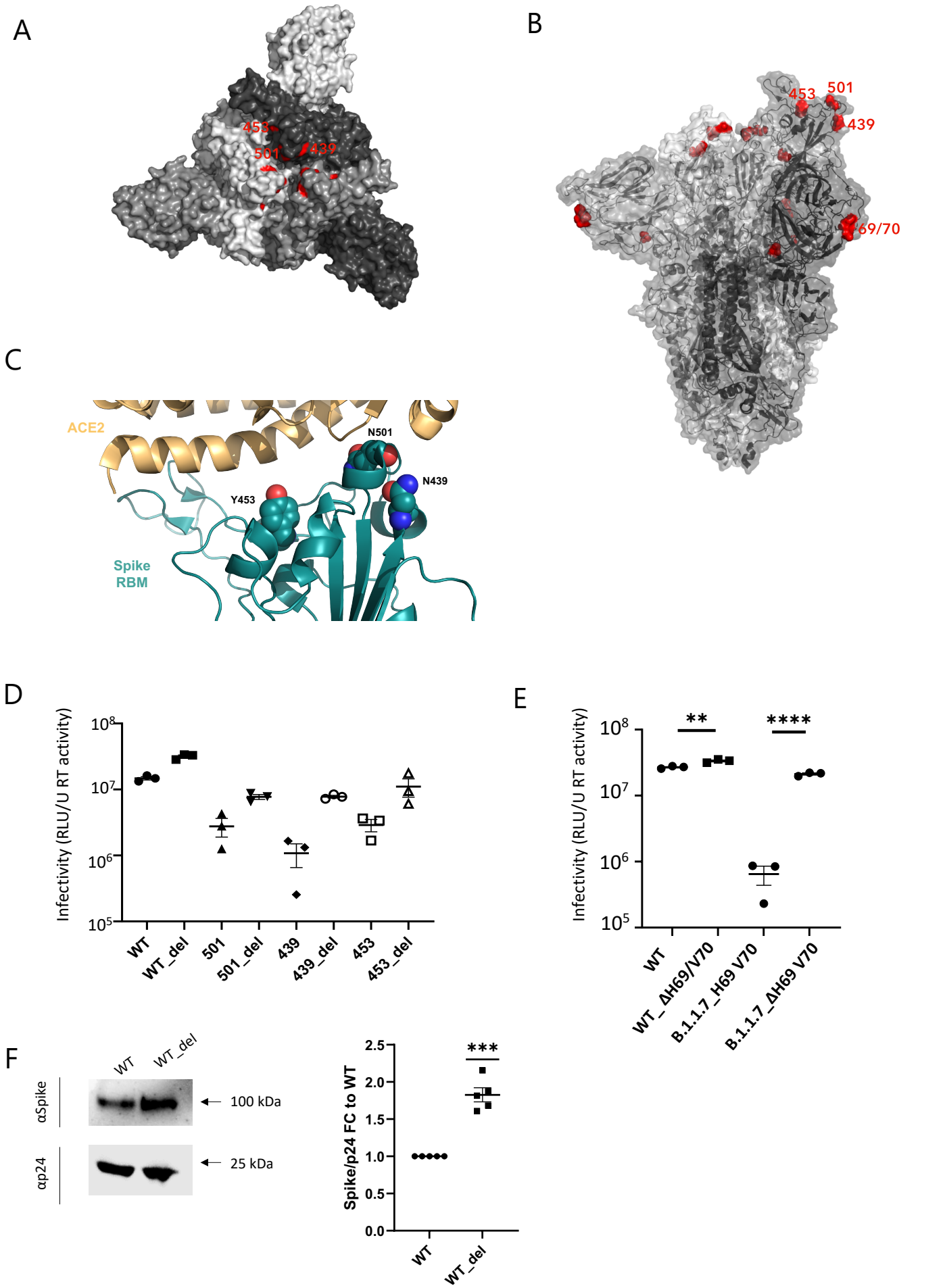


Figure 5: Structure and infectivity of Spike proteins carrying deletion at H69/V70 with RBD mutations A. Surface representation of spike in closed conformation (PDB: 6ZGE) viewed in a 'top-down' view along the trimer axis. The residues associated with RBD substitutions N439K, Y453F and N501Y are highlighted in red and labelled on a single monomer. **B.** 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. Residues 71-75, which form the exposed loop undergoing conformational change in **A**, are omitted from this structure. **C.** Representation of Spike RBM:ACE2 interface (PDB: 6M0J) with residues N439, Y453 and N501 highlighted as spheres coloured by element **D-G**. **Spike mutant ΔH69/V70 compensates for infectivity defect of Spike RBD mutations and is associated with increased Spike incorporation into virions**. **D. Infectivity of Spike ΔH69/V70 deletion (del) 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. Data showing 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. **E. Infectivity of B.1.1.7 with replacement of H69 and V70 versus B.1.1.7 containing Spike ΔH69/V70 deletion.** Single round infection by luciferase expressing lentivirus pseudotyped with SARS-CoV-2 Spike protein on HeLa cells transduced with ACE2. **F. Western blot analysis of supernatants following transfection with Spike and HIV-1 gag-pol expressing plasmids.** Supernatant loading was normalized for input virus input using reverse transcriptase activity in virus supernatants. **Right panel: Q**uantification of Spike:p24 ratio for wild type virus with ΔH69/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.

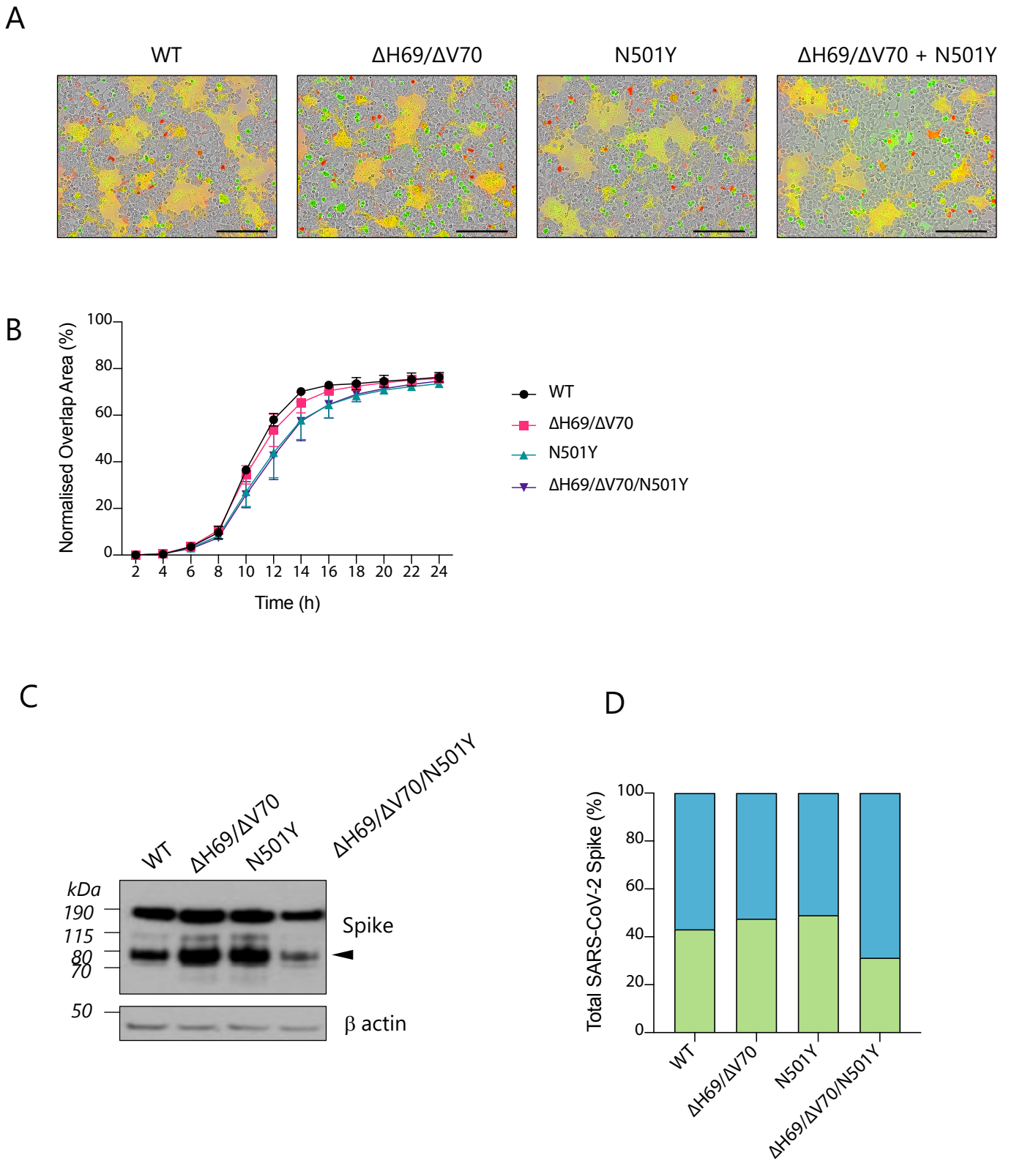


Figure 6*:* Cell Fusogenicity in the presence of Spike mutants ΔH69/V70 and N501Y. A. Reconstructed images at 12 hours of 293T cells co-transfected with the indicated Spike mutants and mCherry expressing plasmid mixed with green dye-labelled Vero cells. Scale bars represent 200 mm. Green colour identifies the acceptor cells while red colour marks donor cells. Merged green-red colours indicate the syncytia. **B.** Quantification of cell-cell fusion kinetics in (A) showing percentage of green and red overlap area over time. **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. b actin is shown as loading control. **D.** Quantification analysis of cleaved (green bars) and uncleaved (blue bars) Spike shown in (C) normalised by b actin.