

1 **Sequence analysis of SARS-CoV-2 in nasopharyngeal samples from patients with COVID-19**  
2 **illustrates population variation and diverse phenotypes, placing the in vitro growth properties**  
3 **of B.1.1.7 and B.1.351 lineage viruses in context.**

4

5 **Running title:** Genetic diversity of SARS-CoV-2 and growth kinetics

6 **Keywords:** SARS-CoV-2, COVID-19, clinical samples, minor variants, growth kinetics

7 Tessa Prince<sup>1,2</sup>, Xiaofeng Dong<sup>1</sup>, Rebekah Penrice-Randal<sup>1</sup>, Nadine Randle<sup>1</sup>, Catherine Hartley<sup>1</sup>,  
8 Hannah Goldswain<sup>1</sup>, Benjamin Jones<sup>1</sup>, Malcolm G. Semple<sup>1,2,3</sup>, J. Kenneth Baillie<sup>4</sup>, Peter J. M.  
9 Openshaw<sup>5</sup>, Lance Turtle<sup>1,2</sup>, ISARIC4C Investigators<sup>1,3,4</sup>, Grant L. Hughes<sup>6</sup>, Enyia R. Anderson<sup>6</sup>,  
10 Edward I. Patterson<sup>6+</sup>, Julian Druce<sup>7</sup>, Gavin Screaton<sup>8</sup>, Miles W. Carroll<sup>2,8,9</sup>, James P. Stewart<sup>1,10</sup>,  
11 and Julian A. Hiscox<sup>1,2,11\*</sup>

12 <sup>1</sup>Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, UK.

13 <sup>2</sup>NIHR Health Protection Research Unit in Emerging and Zoonotic Infections, Liverpool, UK.

14 <sup>3</sup>Department of Respiratory Medicine, Alder Hey Children's Hospital, Liverpool, UK.

15 <sup>4</sup>The Roslin Institute, University of Edinburgh, UK.

16 <sup>5</sup>National Heart and Lung Institute, Imperial College London, UK.

17 <sup>6</sup>Departments of Vector Biology and Tropical Disease Biology, Centre for Neglected Tropical  
18 Diseases, Liverpool School of Tropical Medicine, Liverpool, UK.

19 <sup>7</sup>Virus Identification Laboratory, Doherty Institute, University of Melbourne, Australia.

20 <sup>8</sup>Nuffield Department of Medicine, University of Oxford, UK.

21 <sup>9</sup>Public Health England, Salisbury, UK.

22 <sup>10</sup>Department of Infectious Disease, University of Georgia, Georgia, USA.

23 <sup>11</sup>A\*STAR Infectious Diseases Laboratories (A\*STAR ID Labs), Agency for Science, Technology and

24 Research (A\*STAR), Singapore.

25 <sup>†</sup>Current address: Department of Biological Sciences, Brock University, St. Catharines, Canada.

26

27 <sup>\*</sup>Corresponding author: [julian.hiscox@liverpool.ac.uk](mailto:julian.hiscox@liverpool.ac.uk)

28 **Abstract**

29 New variants of SARS-CoV-2 are continuing to emerge and dominate the regional and global  
30 sequence landscapes. Several variants have been labelled as Variants of Concern (VOCs) because  
31 of perceptions or evidence that these may have a transmission advantage, increased risk of  
32 morbidity and/or mortality or immune evasion in the context of prior infection or vaccination.  
33 Placing the VOCs in context and also the underlying variability of SARS-CoV-2 is essential in  
34 understanding virus evolution and selection pressures. Sequences of SARS-CoV-2 in  
35 nasopharyngeal swabs from hospitalised patients in the UK were determined and virus isolated.  
36 The data indicated the virus existed as a population with a consensus level and non-synonymous  
37 changes at a minor variant. For example, viruses containing the nsp12 P323L variation from the  
38 Wuhan reference sequence, contained minor variants at the position including P and F and other  
39 amino acids. These populations were generally preserved when isolates were amplified in cell  
40 culture. In order to place VOCs B.1.1.7 (the UK 'Kent' variant) and B.1.351 (the 'South African'  
41 variant) in context their growth was compared to a spread of other clinical isolates. The data  
42 indicated that the growth in cell culture of the B.1.1.7 VOC was no different from other variants,  
43 suggesting that its apparent transmission advantage was not down to replicating more quickly.  
44 Growth of B.1.351 was towards the higher end of the variants. Overall, the study suggested that  
45 studying the biology of SARS-CoV-2 is complicated by population dynamics and that these need  
46 to be considered with new variants.

47 **Importance**

48 SARS-CoV-2 is the causative agent of COVID-19. The virus has spread across the planet causing a  
49 global pandemic. In common with other coronaviruses, SARS-CoV-2 genetic material (genomes)  
50 can become quite diverse as a consequence of replicating inside cells. This has given rise to  
51 multiple variants from the original virus that infected humans. These variants may have different  
52 properties and in the context of a widespread vaccination program may render vaccines less  
53 ineffective. Our research confirms the degree of genetic diversity of SARS-CoV-2 in patients. By  
54 isolating viruses from these patients, we show that there is a 100-fold range in growth of even  
55 normal variants. Interestingly, by comparing this to the pattern seen with two Variants of  
56 Concern (UK and South African variants), we show that at least in cells the ability of the B.1.1.7  
57 variant to grow is not substantially different to many of the previous variants.

## 58 Introduction

59 SARS-CoV-2 emerged late 2019 in Wuhan, China and causes COVID-19 (1). This can be a  
60 fatal infection with severe immunopathology in the respiratory system (2). The virus has since  
61 spread worldwide and resulted in more than 2.5 million deaths (3) placing large burdens on  
62 healthcare infrastructures and global economies. Several vaccines have been granted emergency  
63 licensure and these appear to be driving down cases in countries with large scale vaccine roll  
64 outs. However, multiple variants have been identified worldwide and these have the potential  
65 for vaccine evasion and immune escape, leading to the label of Variants of Concern (VOCs).

66 SARS-CoV-2 has a single stranded positive sense RNA genome about 30kb in length. The  
67 first two thirds of the genome is translated to give the viral non-structural proteins (NSP1-16),  
68 which includes the viral RNA dependent RNA polymerase (NSP12). Several viral RNA synthesis  
69 processes occur during infection including replication of the genome and transcription of a  
70 nested set of subgenomic mRNAs (sgmRNAs). This latter process requires discontinuous  
71 transcription during negative strand synthesis (4). As a natural consequence, coronaviruses have  
72 high levels of recombination. This can result in both deletions and insertions and template  
73 switching as well as the formation of defective RNAs. An example of this is the probable insertion  
74 of the furin cleavage site in the spike glycoprotein (5). Although SARS-CoV-2 and other  
75 coronaviruses have some type of proof-reading capability (6), this is generally thought to help  
76 maintain their large genomes, without entering error catastrophe. Otherwise the accumulation  
77 of deleterious mutations would result in a rapid loss of fitness and extinction of a viral population  
78 (7). Additionally, potential genome modifications can result from nucleotide changes through the  
79 action of cellular proteins involved in RNA processing (8). SARS-CoV-2 accumulates mutations at

80 roughly the same frequency as Ebola virus (9). These drivers of genetic diversity and the numbers  
81 of people infected has led to multiple lineages and variants of SARS-CoV-2 being identified  
82 worldwide.

83 The sgRNAs encode for the main structural proteins, including the envelope protein (E)  
84 protein, the membrane (M) protein, the nucleocapsid (N) protein and the spike (S) glycoprotein.  
85 The S protein is a component of the enveloped virion and interacts with the angiotensin  
86 converting enzyme-2 receptor (ACE-2) found on human cells. The S protein is also the major  
87 source of neutralising epitopes and therefore under selection pressure in coronaviruses (and  
88 SARS-CoV-2). Other viral proteins are involved in modulating the innate immune response.

89 Many variations in the coronavirus genome occur in the S gene (10-12) and this also has  
90 been identified for SARS-CoV-2. For example, the D614G substitution in the SARS-CoV-2 S  
91 protein, which emerged by March 2020, demonstrated improved transmissibility compared to  
92 Wuhan variants, and proceeded to dominate worldwide subsequently. This mutation is most  
93 often accompanied with another amino acid substitution in NSP12, P323L (13). In September  
94 2020, a variant of concern, VOC 202012/01 (B.1.1.7 lineage) was detected in Kent in the UK which  
95 possessed 23 mutations distinct from the Wuhan reference sequence, including the N501Y  
96 substitution in the receptor binding domain of the S protein. This may increase the affinity of  
97 spike protein to ACE-2 receptor (14). Initial data suggested this variant could be related to an  
98 increased risk of hospitalisation and death (15). The variant has now spread to several countries  
99 and modelling studies have suggested increased transmissibility (16). Preliminary experiments in  
100 hamsters have identified increased viral shedding compared to the D614G variant (17). However,

101 *in vitro* studies suggest that the B.1.1.7 VOC does not have any replicative advantage in primary  
102 airway epithelial cells (18).

103         As variants are likely to continue to emerge on a background of incomplete vaccination  
104 globally, understanding the significance of such variants both *in vitro* and *in vivo* is important to  
105 provide biological mechanistic data rather than rely on *in silico* modelling to determine their  
106 potential threat to vaccines or transmission advantage. To investigate the genetic and phenotypic  
107 diversity of SARS-CoV-2 in patients and in the context of the emergence of the B.1.1.7 and B.1.351  
108 lineage viruses and concerns around potential higher viral loads, the growth of these viruses was  
109 bench marked against the Victoria isolate and clinical isolates from other samples taken during  
110 the outbreak.

## 111 **Results**

112           Although consensus genomes for SARS-CoV-2 are reported on global databases from the  
113 sequencing of clinical specimens, in reality the virus will exist as a population within an individual  
114 and may also include defective RNAs. Likewise, in some pipelines, viral genomes or variants  
115 containing out of place stop codons within ORFs will not be returned as consensus even though  
116 they may be dominant. In this case, at a minor variant level, which could represent 49% of other  
117 genomes within the same individual, the wildtype protein may be expressed, and counterbalance  
118 any aberrantly functioning proteins. To investigate the sequence diversity of SARS-CoV-2 within  
119 a patient and to compare the growth of these viral populations to recent VOCs, nasopharyngeal  
120 swabs were taken from patients with COVID-19, sequenced and the genotypes and variants of  
121 isolated viruses and their growth properties compared in cell culture (Figure 1).

122

### 123 **Sequence variation of SARS-CoV-2 in clinical swabs compared to Wuhan reference strain**

124           Nine swabs representing different time points in the outbreak in the UK contained  
125 recoverable virus that could be isolated and grown. The virus population in these swabs was  
126 sequenced and both consensus genomes and minor variants determined. Consensus sequence  
127 variation was compared to the reference genome (NC\_045512; Wuhan-Hu-1) to see how far the  
128 isolates had diverged and with minor variants listed for the secondary and tertiary positions  
129 (Supplementary Table 1). Most viruses demonstrated a few amino acid variations compared to  
130 the reference sequence. For example, SCV2-006, a lineage B virus sequenced from the swab of  
131 a patient from the Diamond Princess cruise ship (February 2020) had only one substitution  
132 present, R203K in the N protein (Figure 2). In comparison, sequence analysis of SCV2-009, a virus



133 isolated from a swab sampled from a patient in the UK in March 2020 (Figure 2) now possessed  
134 the D614G and P323L substitutions in the spike glycoprotein and NSP12, respectively. These are  
135 in contrast to the B.1.1.7 variant, which emerged later in 2020 and is characterised by the  
136 presence of 23 amino acid differences from the reference genome. Analysis of the virus  
137 population present in the nasopharyngeal isolate of SCV2-009 illustrated the diversity associated  
138 with the virus. For example, taking the P323L substitution in NSP12, out of an amino acid  
139 coverage of 202, 170 amino acids mapped to L, 12 to P and 9 to F. For the D614G substitution in  
140 the spike glycoprotein, out of an amino acid coverage of 3452, 3360 amino acids mapped to G,  
141 24 to S and 21 to V. This general pattern is reflected in other clinical isolates. For example, in  
142 isolate SCV2-010, in NSP12, out of an amino acid coverage of 285, 273 mapped to L, 50 to I and  
143 3 to P. In isolate SCV2-008, in NSP12, out of an amino acid coverage of 153, 130 mapped to L, 9  
144 to P and 7 to F. This suggests, for NSP12, that at the minor variant level the reference sequence  
145 amino acid is still present, but other amino acids such as F may be common (Supplementary Table  
146 1), and subject to selection pressure. In some clinical swabs, for example in N at position 204, the  
147 second most common feature is a stop codon. SCV2-011 and SCV2-018 were variants isolated  
148 from clinical swabs taken from the same patient but three days apart, these did not vary at the  
149 consensus between each other, but did at the minor variant level. SCV2-007 and SCV2-017 were  
150 also variants isolated from clinical swabs taken from the same patient but three days apart and  
151 did not vary at the consensus between each other in swabs, but did at the minor variant level.

152

153 **Comparison of sequence variation in stocks and after 72 hours in hACE2-A549 cells**

154           In order to assess the biology of the viruses isolated from the clinical swabs and compare  
155 their growth to B.1.1.7 and B.1.351, sufficient stocks had to be grown. To isolate SARS-CoV-2  
156 from the clinical swabs, the nasopharyngeal sample was filtered and placed on VeroE6 cells with  
157 antibiotics and antifungals until CPE was observed. The supernatant was collected from these  
158 cells to generate sufficient stocks for infectivity assays and comparisons.

159           Growing virus for stocks may have introduced or selected for specific variants. One of  
160 these, that has been characterised for SARS-CoV-2, is a deletion of the furin cleavage site in the  
161 spike glycoprotein when grown in Vero E6 cell (19). Therefore, viral stocks were sequenced to  
162 ensure they did not possess the deletion and to determine if variation occurred compared to  
163 when the virus was sequenced directly from clinical swabs. Comparator viruses of known  
164 provenance were obtained from collaborators. The comparator viruses were the B.1.1.7 ('Kent'  
165 UK VOC) virus (termed SCV2-019 in this study) obtained at P4, the SARS-CoV-2/Victoria/01/2020  
166 (an isolate from Australia) obtained at P3 (termed SCV2-021), and the B.1.351 virus ('South  
167 African' VOC) (termed SCV2-022 in this study). These were grown in Vero/hSLAMS as a precaution  
168 to prevent selection for the furin deletion. These were also sequenced to ensure they had the  
169 variant defining mutations present. For these three comparator viruses, the sequencing showed  
170 at the consensus level the furin cleavage site was intact and the other defining variations  
171 separating these variants from the Wuhan reference sequence were present (Figure 3).

172           Analysis of the genome diversity between viruses sequenced in swabs from patients and  
173 the virus stock used to infect cells indicated that most consensus variations from the Wuhan  
174 reference sequence were still present (Supplementary Table 1). The minor variants at selected  
175 positions were also still present. For example, in the stock preparation for SCV2-009, at position

176 323 in NSP12, this was read with an amino acid depth of 517. The L was present at a depth of  
177 499, P with a depth of 6 and F with a depth of 5, indicating that the consensus level amino acid  
178 was still present with P and F at a minor level. For some stock viruses, variation from the  
179 reference sequence was lost during preparation of the stock virus. The growth of these viruses  
180 from the stocks was compared to a B.1.1.7. and a B.1.351 lineage virus and SARS-CoV-  
181 2/Victoria/01/2020, obtained from near the start of the COVID-19 pandemic.

182

### 183 **Growth comparison of different SARS-CoV-2 variants to Variants of Concern (VOCs).**

184 In order to identify whether the B.1.1.7 (SCV2-019) and B.1.351 (SCV2-022) displayed a  
185 growth advantage over less recent strains of the virus, three different cell lines were infected  
186 with the viruses at an MOI of 0.01 over the course of 72 hours and the resultant supernatants, at  
187 24, 48 and 72 hours titrated by plaque assay on Vero E6 cells. The three different cell lines were  
188 Vero E6 (commonly used to grow viral stocks and initial isolates from clinical samples),  
189 Vero/hSLAM (reported to prevent deletion of the furin cleavage site in the spike glycoprotein)  
190 and hACE2-A549 cells. This latter cell line is based on A549 cells, which are respiratory epithelium  
191 in origin, commonly used to study respiratory viruses in cell culture but overexpress the ACE2  
192 protein. A549 cells mount an interferon response to virus infection.

193 In Vero E6 cells, eleven SARS-CoV-2 variants followed a similar pattern of growth, with  
194 the exception of SCV2-021 (SARS-CoV-2/Victoria/01/2020) which grew at significantly reduced  
195 levels compared to other variants by 72 hours post infection (mean  $5.7 \times 10^4$  PFU/ml,  $p=0.006$ )  
196 (Figure 4A). A similar pattern of growth was observed in all twelve variants in Vero/hSLAM cells,  
197 but there was no significant difference ( $p>0.05$ ) in the titres of any of the viruses produced by 72

198 hours post infection (Figure 4B). In contrast, in hACE2-A549 cells, there was more heterogeneity  
199 observed between variants, with the range of viral titres being much lower ( $7.9 \times 10^1$ - $3.01 \times 10^4$   
200 PFU/ml) than that observed in Vero cells at 72 hours post infection. The B.1.1.7 variant (SCV2-  
201 019) had the lowest titre at 24 hours post-infection in growth assays in hACE2-A549 cells, before  
202 growing to reach a final titre  $2.71 \times 10^3$  PFU/ml at 72 hours post-infection (Figure 4C). We  
203 observed that between variants at the growth extremes at 72 hrs post-infection in hACE2-A549  
204 cells there was an approximately <2 log difference between titres of SCV2-007 and SCV2-018,  
205 despite the same amount of virus being used as the inoculum (MOI=0.01). For the VOC B.1.1.7  
206 the growth at 72 hrs post-infection in hACE2-A549 cells was in the middle of the other variants  
207 tested ( $2.71 \times 10^3$  PFU/ml) while the VOC B.1.351 had the second highest final titre of  $2.62 \times 10^4$   
208 PFU/ml. (Figure 4C). In addition, we note that the B.1.351 lineage variant had the highest viral  
209 titres in both VeroE6 and Vero/hSLAMs at 24 and 48 hrs post-infection.

210 Comparing viruses grown from the same patient but sampled three days apart (SCV2-007  
211 and SCV2-017, and SCV2-011 and SCV2-018 at day one and day three respectively, showed  
212 differences in their growth in hACE2-A549 cells. (Supplementary Table 2). There was a 2-log  
213 difference in the growth of SCV2-007 and SCV2-017, while there was little difference in the  
214 growth of SCV2-011 and SCV2-018.

215

## 216 **The phenotype of the variants differed widely between cell lines, displaying mixed plaque** 217 **morphology and growth characteristics**

218 The phenotype of the plaques formed by each virus stock was observed in the three  
219 different cell lines used at 72 hours post-infection. The appearance of the plaques from the

220 variants differed (Figure 5). SCV2-006, SCV2-011, SCV2-016, SCV2-018 and SCV2-022 had a larger  
221 plaque phenotype after growth in Vero E6 cells, compared with SCV2-019 (B.1.1.7) and SCV2-021  
222 (SARS-CoV-2/Victoria/01/2020). Equally, some variants displayed a mixed phenotype of both  
223 large and small plaques in Vero E6 cells, as seen for SCV2-011, suggesting mixed viral species  
224 were present (Figure 5). After growth in Vero/hSLAMS, SCV2-011 and SCV2-018 showed a mixed  
225 phenotype after plaque assay. SCV2-019 (B.1.1.7) and SCV2-021 had the smallest plaque  
226 phenotypes. After growth in hACE2-A549 cells, SCV2-006 and SCV2-022 (B.1.351) had the largest  
227 plaque phenotypes, while SCV2-021 had the smallest plaque phenotype. SCV2-006 and SCV2-016  
228 had mixed morphology of both large and small plaques. This illustrates the potential diversity  
229 with a viral population.

230

### 231 **Genetic diversity of variants after passage in the three different cell types**

232 We hypothesised that differences in the phenotypic appearance of viruses and their  
233 reproduction might reflect the presence of minor variants and stop codons in their underlying  
234 sequences and this was investigated at 72 hrs post-infection in hACE2-A549 cells. All variants  
235 have a consensus level genome but also minor variants. In SCV2-016, at 72 hrs post-infection  
236 there was a stop codon at consensus level in ORF3A that was also present in the viral stock  
237 (Supplementary Table 1). In SCV2-019 (the B.1.1.7. lineage virus) there was a stop codon in ORF8  
238 at the consensus level at 72 hr post-infection, which was also present in the stock (Supplementary  
239 Table 1). We note that both of these were low read depth, and other amino acids were present  
240 at the minor variant level. Stop codons were also present in the variants at a minor variant level  
241 (Supplementary Table 1).

## 242 Discussion

243 Sequence analysis of SARS-CoV-2 in clinical swabs from patients revealed a heterogenous  
244 and diverse population from the Wuhan reference sequence. At the minor variant level, a  
245 number of variants had genomes which contained premature stop codons. Examples of SARS-  
246 CoV-2 genomes encoding non-functioning proteins have been previously identified in the human  
247 population. For example, a cohort of patients in Singapore were identified with a deletion in ORF8  
248 that was associated with a milder infection (20), although the variant disappeared either through  
249 control measures or lack of fitness. This potential disconnect is not restricted to SARS-CoV-2. The  
250 balance between consensus and minor variants and the presence of stop codons in virus  
251 populations within individual patients has been shown to influence the activity of the Ebola virus  
252 RNA dependent RNA polymerase and correlate with outcome in patients with Ebola virus disease  
253 (21). Within an individual person with SARS-CoV-2, these mixtures of functioning and presumably  
254 non-functioning viral proteins will potentially influence viral load.

255 Recent VOCs include SARS-CoV-2 variants from Nigeria (B.1.525). One of the differences  
256 in this variant from the Wuhan consensus sequence is P323F in NSP12. This variation was also  
257 identified in a cluster of patients in Northern Nevada in the USA (22). The analysis of variants in  
258 this study, isolated earlier in 2020, indicated that an F at position 323 existed at a minor variant  
259 level. Therefore, we hypothesize that if an F at this position was advantageous (e.g. altered RdRp  
260 activity) then the variant would have been selected during passage. However, this was not the  
261 case, and therefore we speculate that the emergence of an F at position 323 in NSP12 may be  
262 through founder effect.

263           The growth of different variants, and a variant from near the start of the COVID-19  
264 pandemic, were compared in three different cell lines to the growth of two VOCs and the  
265 Australia Victoria variant. These VOCs were from the B.1.1.7 and B.1.351 lineages and represent  
266 viruses that have an apparent transmission advantage in the general population and/or may be  
267 less refractive to currently approved vaccines. There was an approximately 2 log difference in  
268 growth at 72 hrs in the hACE2-A549 cells by the different variants. The growth of VOCs in this cell  
269 line were within these limits. Extrapolating this observation to the perceived transmission  
270 advantage of B.1.1.7 in the human population, would suggest this is not down to the VOC growing  
271 to higher titres in cells *in vivo* compared with other variants, though we acknowledge that out *in*  
272 *vitro* experiments may not correlate exactly with growth rates *in vivo*. With B.1.1.7, the same  
273 heterogenous patterns of disease in humans as other variants, although we note that B.1.1.7 has  
274 been associated with a decrease in Ct values from nasopharyngeal swabs and also an increase in  
275 mortality in some populations. In contrast, different variants of the coronavirus infectious  
276 bronchitis (IBV) can individually cause different spectrums of organ specific disease, and  
277 therefore current variations in the genome of SARS-CoV-2 may not automatically equate to  
278 radically different disease as observed with IBV. Due to the promiscuous nature of coronavirus  
279 RNA synthesis, variants have and will occur all of the time. This emphasises the need for genotype  
280 to phenotype studies to place newly emerged variants that have perceived differences in context.

281           Comparison of viruses isolated from the same patients at different time points revealed  
282 intriguing differences. While the viruses SCV2-007 and SCV2-017 differed by 2-logs in hACE2-  
283 A549 cells at 72 hours, there was little difference observed between the twin viruses SCV2-011  
284 and SCV2-018. Notably, the SCV2-017 virus had picked up an additional mutation at 72 hours

285 post-infection – a change from A in the reference genome at position 1120 in NSP3 to a V. It is  
286 possible this may be responsible for the difference in growth of this virus to its founder, SCV2-  
287 007, and could reflect viral adaptation to the immune response in this individual over the course  
288 of infection (Supplementary table 2).

289         The analysis of virus in the nasopharyngeal swabs clearly paints a picture of a diverse  
290 population of SARS-CoV-2. When studying isolates, even when grown in cell culture, that  
291 population still continues. Thus, whilst lineage defining variations are present at a consensus  
292 level, minor variants are present underneath that may have an impact on biology. This would  
293 suggest that the study of specific genotypes requires either plaque purification or reverse  
294 genetics. However, the study suggests that the viral population (consensus and minor variants)  
295 should be taken into account when studying the transmission of SARS-CoV-2.



296 **Methods**

297

298 **Cells.** African green monkey kidney C1008 (Vero E6) cells (Public Health England, PHE) were  
299 cultured in Dulbecco's minimal essential medium (DMEM) (Sigma) with 10% foetal bovine serum  
300 (FBS) (Sigma) and 0.05mg/ml gentamicin at 37°C/5% CO<sub>2</sub>. Vero/hSLAM cells (PHE) were grown in  
301 DMEM with 10% FBS and 0.05mg/ml gentamicin (Merck) with the addition of 0.4mg/ml Geneticin  
302 (G418; Thermofisher) at 37°C/5% CO<sub>2</sub>. Human ACE2-A549 (hACE2-A549), a lung epithelial cell line  
303 which overexpresses the ACE-2 receptor, were the kind gift of Oliver Schwartz (23) and were  
304 cultured in DMEM with 10% FBS and 0.05mg/ml gentamicin with the addition of 10µg/ml  
305 Blastidin (Invitrogen). Only passage 3-10 cultures were used for experiments.

306

307 **Virus isolation.** The SARS-CoV-2/human/Liverpool/REMRQ0001/2020 isolate (Genbank ID  
308 MW041156.1), was used at passage 3. The fourth passage of virus (here named SCV2-  
309 006\_stockP4) was cultured in Vero E6 cells with DMEM containing 4% FBS and 0.05mg/ml  
310 gentamicin at 37°C/5% CO<sub>2</sub> and harvested 48 hours post inoculation. Virus stocks were aliquoted  
311 and stored at -80°C.

312 Viruses named SCV2-007 to SCV2-018 were grown from nasopharyngeal swabs of  
313 patients using the following method. One hundred microlitres of viral transport media from the  
314 swab was mixed with 100µl DMEM with 4% FBS, 0.05mg/ml gentamicin, 25µg/ml plasmocin  
315 (Invivogen) and 2.5µg/ml amphotericin B (Merck). These were then filtered using ultrapure MC  
316 0.22µm filters (Merck) and the filtrate placed onto cells in a 24 well plate of Vero E6 cells for 1  
317 hour. After one hour, the media was topped up with DMEM (2% FBS, 0.05 mg/ml gentamicin,

318 25µg/ml plasmocin, 2.5 µg/ml amphotericin B). Cells were observed daily for cytopathic effect  
319 (CPE) and the cell supernatant harvested once CPE was evident. This provided the first passage  
320 virus. Stocks of these were then grown in Vero E6 as described above and frozen down in aliquots  
321 at -80°C and named SCV2-007 to SCV2-018\_stockP2.

322 The B.1.1.7 and B.1.351 isolates were used at passage 4. The fifth passage (here named  
323 SCV2-019\_stockP5 and SCV2-022\_stockP5) were cultured in Vero/hSLAM cells with DMEM  
324 containing 4% FBS, 0.05mg/ml gentamicin and 0.4mg/ml geneticin and harvested 72 hours post  
325 inoculation. Virus stocks were aliquoted and stored at -80°C. SARS-CoV-2 Victoria/01/2020 was  
326 passaged three times in Vero/hSLAM cells. The fourth passage stock (here named SCV2-  
327 021\_stockP4) was cultured in Vero/hSLAM cells DMEM containing 4% FBS, 0.05mg/ml  
328 gentamicin and 0.4mg/ml geneticin and harvested 72 hours post inoculation. Virus stocks were  
329 aliquoted and stored at -80°C (Supplementary table 2).

330

331 **Virus titration.** Viral titres of stocks were calculated using plaque assays. Briefly, confluent 24-  
332 well plates of Vero E6 cells were inoculated with serial ten-fold dilutions of the stocks in duplicate  
333 for one hour at 37°C/5% CO<sub>2</sub>. Plates were overlaid with DMEM containing 2% FBS, 0.05mg/ml  
334 gentamicin and 2% low melting point agarose (Lonza) and incubated at 37°C/5% CO<sub>2</sub> for 72 hours.  
335 Plates were fixed using 10% formalin, the overlay removed, and plates stained using crystal violet  
336 solution (Sigma). Virus titre was measured in plaque forming units per ml (PFU/ml).

337

338 **Virus growth kinetics.** Vero E6, Vero/hSLAM and hACE2-A549 cells were grown in 96 well plates  
339 for viral growth kinetic experiments. For infection, media was removed from plates and virus

340 inoculum added at an MOI of 0.01 in DMEM containing 2% FBS, 0.05mg/ml gentamicin and the  
341 respective selective antibiotics for each cell line (6 wells per timepoint). Plates were incubated at  
342 37°C/5% CO<sub>2</sub> for one hour. The inoculum was removed, and cells were washed once with PBS  
343 (Sigma). The respective media with 2% FBS (100µl) was added to each well. The cell supernatant  
344 was removed from wells and combined (0hrs post infection) and plates incubated further.  
345 Supernatants were likewise removed at 24, 48 and 72 hours post infection. Approximately 250µl  
346 of the supernatants were aliquoted directly into tubes containing 750µl Trizol LS (Fisher) to  
347 inactivate the virus. All supernatants and inactivated supernatants were stored at -80°C until viral  
348 titration and RNA extraction could be performed. All infections were performed at least three  
349 times in independent experiments.

350

351 **RNA extraction and amplification of viral nucleic acids.** RNA from clinical samples was extracted  
352 and DNase treated as described previously. Samples from patients were sequenced using the  
353 RLSA approach (24). RNA from viral stocks and from 72-hour post infection cultures were  
354 sequenced by Oxford Nanopore long read length sequencing on flow cells run on MinION or  
355 GridION.

356

357 **Nanopore sequencing.** Sequencing libraries for amplicons generated by RSLA (24) or ARTIC were  
358 prepared following the 'PCR tiling of SARS-CoV-2 virus with Native Barcoding' protocol provided  
359 by Oxford Nanopore Technologies using LSK109 and EXP-NBD104/114.

360

361 **Variant calling.** The artic-ncov2019 pipeline v1.2.1 (<https://artic.network/ncov-2019/ncov2019->  
362 [bioinformatics-sop.html](https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html)) was used to filter the passed Fastq files produced by Nanopore  
363 sequencing with lengths between 800 and 1600 for RSLA, and 400 and 700 for ARTIC. This  
364 pipeline was then used to map the filtered reads on the reference SARS-CoV-2 genome  
365 (NC\_045512.2) by minimap2 and assigned each read alignment to a derived amplicon and  
366 excluded primer sequences based on the RSLA and ARTIC V3 primer schemes in the bam files.  
367 These bam files were further analysed using DiversiTools  
368 (<http://josephhughes.github.io/btctools/>) with the “-orfs” function to generate the ratio of  
369 amino acid change in the reads and coverage at each site of protein in comparison to the  
370 reference SARS-CoV-2 genome (NC\_045512.2). The amino acids with highest ratio and coverage  
371 > 10 were used to assemble the consensus protein sequences.

372

373 **Statistics.** Viral titre data was log transformed and one-way ANOVAs performed with post-hoc  
374 Bonferroni tests performed to determine if any significant difference at T=72 hours post infection  
375 occurred between the SCV2-019 (B.1.1.7) and other viruses in different cell lines.

376

377 **Ethics and clinical information.** The patients from which the virus samples were obtained gave  
378 informed consent and were recruited under the International Severe Acute Respiratory and  
379 emerging Infection Consortium (ISARIC) WHO Clinical Characterisation Protocol CCP. Ethical  
380 approval for data collection and analysis by ISARIC4C was given by the South Central-Oxford C  
381 Research Ethics Committee in England (reference 13/SC/0149), and by the Scotland A Research  
382 Ethics Committee (reference 20/SS/0028). Samples were use with consent from patients or

383 consultees. The ISARIC WHO CCP-UK study was registered at  
384 <https://www.isrctn.com/ISRCTN66726260> and designated an Urgent Public Health Research  
385 Study by NIHR. Protocol, patient information sheets, consents, case report forms and process of  
386 data and sample access request are available at <https://ISARIC4C.net>.

387

388 **Biosafety.** All work was performed in accordance with risk assessments and standard operating  
389 procedures approved by the University of Liverpool Biohazards Sub- Committee and by the UK  
390 Health and Safety Executive. Work with SARS-CoV-2 was performed at containment level 3 by  
391 personnel equipped with respirator airstream units with filtered air supply.

392

393 **Funding.** This work was funded by U.S. Food and Drug Administration Medical Countermeasures  
394 Initiative contract (75F40120C00085) awarded to JAH. The article reflects the views of the  
395 authors and does not represent the views or policies of the FDA. This work was also supported  
396 by the MRC (MR/W005611/1) G2P-UK: A national virology consortium to address phenotypic  
397 consequences of SARS-CoV-2 genomic variation (co-I JAH). JAH is also funded by the Centre of  
398 Excellence in Infectious Diseases Research (CEIDR) and the Alder Hey Charity. The ISARIC4C  
399 sample collection and sequencing in this study was supported by a grants from the Medical  
400 Research Council (grant MC\_PC\_19059), the National Institute for Health Research (NIHR; award  
401 CO-CIN-01) and the Medical Research Council (MRC; grant MC\_PC\_19059). JAH, GLH, MWC and  
402 LT are supported by the NIHR Health Protection Research Unit (HPRU) in Emerging and Zoonotic  
403 Infections at University of Liverpool in partnership with Public Health England (PHE), in  
404 collaboration with Liverpool School of Tropical Medicine and the University of Oxford (award

405 200907). LT is supported by a Wellcome Trust fellowship [205228/Z/16/Z]. For the purpose of  
406 Open Access, the authors have applied a CC BY public copyright licence to any Author Accepted  
407 Manuscript version arising from this submission. The views expressed are those of the authors  
408 and not necessarily those of the funders.

409 **Author Contributions**

410 TP, XD, RP-R, NR, CH, HG, BJ, JD, GLH, GS, ERA and EIP performed the experiments,  
411 sequencing, bioinformatics and isolated virus. Data was analysed by TP, XD, RP-R and  
412 JAH. MWC, LT, JPS and JAH supervised the project. MGS, JKB and PJMO established  
413 the ISARIC4C consortium that was used to obtain some of the UK clinical isolates used  
414 in the study. TP, XD, MWC and JAH wrote the manuscript, all authors provided editing  
415 and final approval.

416

417 **Acknowledgements**

418 Clinical isolates used in this study were gathered under the auspices of the ISARIC Coronavirus  
419 Clinical Characterisation Consortium (ISARIC4C) and processed at the University of Liverpool. We  
420 would like to acknowledge all members of the consortia. Consortium Lead Investigator: J.  
421 Kenneth Baillie; Chief Investigator: Malcolm G. Semple; Co-Lead Investigator: Peter J.M.  
422 Openshaw; ISARIC Clinical Coordinator: Gail Carson; Co-Investigators: Beatrice Alex, Benjamin  
423 Bach, Wendy S. Barclay, Debby Bogaert, Meera Chand, Graham S. Cooke, Annemarie B. Docherty,  
424 Jake Dunning, Ana da Silva Filipe, Tom Fletcher, Christopher A. Green, Ewen M. Harrison, Julian  
425 A. Hiscox, Antonia Ying Wai Ho, Peter W. Horby, Samreen Ijaz, Saye Khoo, Paul Klenerman,  
426 Andrew Law, Wei Shen Lim, Alexander J. Mentzer, Laura Merson, Alison M. Meynert, Mahdad  
427 Noursadeghi, Shona C. Moore, Massimo Palmarini, William A. Paxton, Georgios Pollakis, Nicholas  
428 Price, Andrew Rambaut, David L. Robertson, Clark D. Russell, Vanessa Sancho-Shimizu, Janet T.  
429 Scott, Thushan de Silva, Louise Sigfrid, Tom Solomon, Shiranee Srisakandan, David Stuart,

430 Charlotte Summers, Richard S. Tedder, Emma C. Thomson, A.A. Roger Thompson, Ryan S.  
431 Thwaites, Lance C.W. Turtle, and Maria Zambon; Project Managers: Hayley Hardwick, Chloe  
432 Donohue, Ruth Lyons, Fiona Griffiths, and Wilna Oosthuyzen; Data Analysts: Lisa Norman, Riinu  
433 Pius, Tom M. Drake, Cameron J. Fairfield, Stephen Knight, Kenneth A. Mclean, Derek Murphy,  
434 and Catherine A. Shaw; Data and Information System Managers: Jo Dalton, James Lee, Daniel  
435 Plotkin, Michelle Girvan, Egle Saviciute, Stephanie Roberts, Janet Harrison, Laura Marsh, Marie  
436 Connor, Sophie Halpin, Clare Jackson, and Carrol Gamble; Data Integration and Presentation:  
437 Gary Leeming, Andrew Law, Murray Wham, Sara Clohisey, Ross Hendry, and James Scott-Brown;  
438 Material Management: William Greenhalf, Victoria Shaw, and Sarah McDonald; Patient  
439 Engagement: Seán Keating; Outbreak Laboratory Staff and Volunteers: Katie A. Ahmed, Jane A.  
440 Armstrong, Milton Ashworth, Innocent G. Asimwe, Siddharth Bakshi, Samantha L. Barlow, Laura  
441 Booth, Benjamin Brennan, Katie Bullock, Benjamin W.A. Catterall, Jordan J. Clark, Emily A. Clarke,  
442 Sarah Cole, Louise Cooper, Helen Cox, Christopher Davis, Oslem Dincarslan, Chris Dunn, Philip  
443 Dyer, Angela Elliott, Anthony Evans, Lorna Finch, Lewis W.S. Fisher, Terry Foster, Isabel Garcia-  
444 Dorival, William Greenhalf, Philip Gunning, Catherine Hartley, Antonia Ho, Rebecca L. Jensen,  
445 Christopher B. Jones, Trevor R. Jones, Shadia Khandaker, Katharine King, Robyn T. Kiy, Chrysa  
446 Koukorava, Annette Lake, Suzannah Lant, Diane Latawiec, L. Lavelle-Langham, Daniella Lefteri,  
447 Lauren Lett, Lucia A. Livoti, Maria Mancini, Sarah McDonald, Laurence McEvoy, John McLauchlan,  
448 Soeren Metelmann, Nahida S. Miah, Joanna Middleton, Joyce Mitchell, Shona C. Moore, Ellen G.  
449 Murphy, Rebekah Penrice-Randal, Jack Pilgrim, Tessa Prince, Will Reynolds, P. Matthew Ridley,  
450 Debby Sales, Victoria E. Shaw, Rebecca K. Shears, Benjamin Small, Krishanthi S. Subramaniam,  
451 Agnieska Szemiel, Aislynn Taggart, Jolanta Tanianis-Hughes, Jordan Thomas, Erwan Trochu, Libby



452 van Tonder, Eve Wilcock, and J. Eunice Zhang; Local Principal Investigators: Kayode Adeniji, Daniel  
453 Agranoff, Ken Agwuh, Dhiraj Ail, Ana Alegria, Brian Angus, Abdul Ashish, Dougal Atkinson,  
454 Shahedal Bari, Gavin Barlow, Stella Barnass, Nicholas Barrett, Christopher Bassford, David Baxter,  
455 Michael Beadsworth, Jolanta Bernatoniene, John Berridge, Nicola Best, Pieter Bothma, David  
456 Brealey, Robin Brittain-Long, Naomi Bulteel, Tom Burden, Andrew Burtenshaw, Vikki Caruth,  
457 David Chadwick, Duncan Chambler, Nigel Chee, Jenny Child, Srikanth Chukkambotla, Tom Clark,  
458 Paul Collini, Catherine Cosgrove, Jason Cupitt, Maria-Teresa Cutino-Moguel, Paul Dark, Chris  
459 Dawson, Samir Dervisevic, Phil Donnison, Sam Douthwaite, Ingrid DuRand, Ahilanadan  
460 Dushianthan, Tristan Dyer, Cariad Evans, Chi Eziefula, Christopher Fegan, Adam Finn, Duncan  
461 Fullerton, Sanjeev Garg, Sanjeev Garg, Atul Garg, Effrossyni Gkrania-Klotsas, Jo Godden, Arthur  
462 Goldsmith, Clive Graham, Elaine Hardy, Stuart Hartshorn, Daniel Harvey, Peter Havalda, Daniel B.  
463 Hawcutt, Maria Hobrok, Luke Hodgson, Anil Hormis, Michael Jacobs, Susan Jain, Paul Jennings,  
464 Agilan Kaliappan, Vidya Kasipandian, Stephen Kegg, Michael Kelsey, Jason Kendall, Caroline  
465 Kerrison, Ian Kerslake, Oliver Koch, Gouri Koduri, George Koshy, Shondipon Laha, Steven Laird,  
466 Susan Larkin, Tamas Leiner, Patrick Lillie, James Limb, Vanessa Linnett, Jeff Little, Michael  
467 MacMahon, Emily MacNaughton, Ravish Mankregod, Huw Masson, Elijah Matovu, Katherine  
468 McCullough, Ruth McEwen, Manjula Meda, Gary Mills, Jane Minton, Mariyam Mirfenderesky,  
469 Kavya Mohandas, Quen Mok, James Moon, Elinoor Moore, Patrick Morgan, Craig Morris,  
470 Katherine Mortimore, Samuel Moses, Mbiye Mpenge, Rohinton Mulla, Michael Murphy, Megan  
471 Nagel, Thapas Nagarajan, Mark Nelson, Igor Otahal, Mark Pais, Selva Panchatsharam, Hassan  
472 Paraiso, Brij Patel, Natalie Pattison, Justin Pepperell, Mark Peters, Mandeep Phull, Stefania  
473 Pintus, Jagtur Singh Pooni, Frank Post, David Price, Rachel Prout, Nikolas Rae, Henrik Reschreiter,

474 Tim Reynolds, Neil Richardson, Mark Roberts, Devender Roberts, Alistair Rose, Guy Rousseau,  
475 Brendan Ryan, Taranprit Saluja, Aarti Shah, Prad Shanmuga, Anil Sharma, Anna Shawcross,  
476 Jeremy Sizer, Manu Shankar-Hari, Richard Smith, Catherine Snelson, Nick Spittle, Nikki Staines,  
477 Tom Stambach, Richard Stewart, Pradeep Subudhi, Tamas Szakmany, Kate Tatham, Jo Thomas,  
478 Chris Thompson, Robert Thompson, Ascanio Tridente, Darell Tupper-Carey, Mary Twagira,  
479 Andrew Ustianowski, Nick Vallotton, Lisa Vincent-Smith, Shico Visuvanathan, Alan Vuylsteke,  
480 Sam Waddy, Rachel Wake, Andrew Walden, Ingeborg Welters, Tony Whitehouse, Paul Whittaker,  
481 Ashley Whittington, Meme Wijesinghe, Martin Williams, Lawrence Wilson, Sarah Wilson,  
482 Stephen Winchester, Martin Wiselka, Adam Wolverson, Daniel G. Wooton, Andrew Workman,  
483 Bryan Yates, and Peter Young.

484 **References**

485

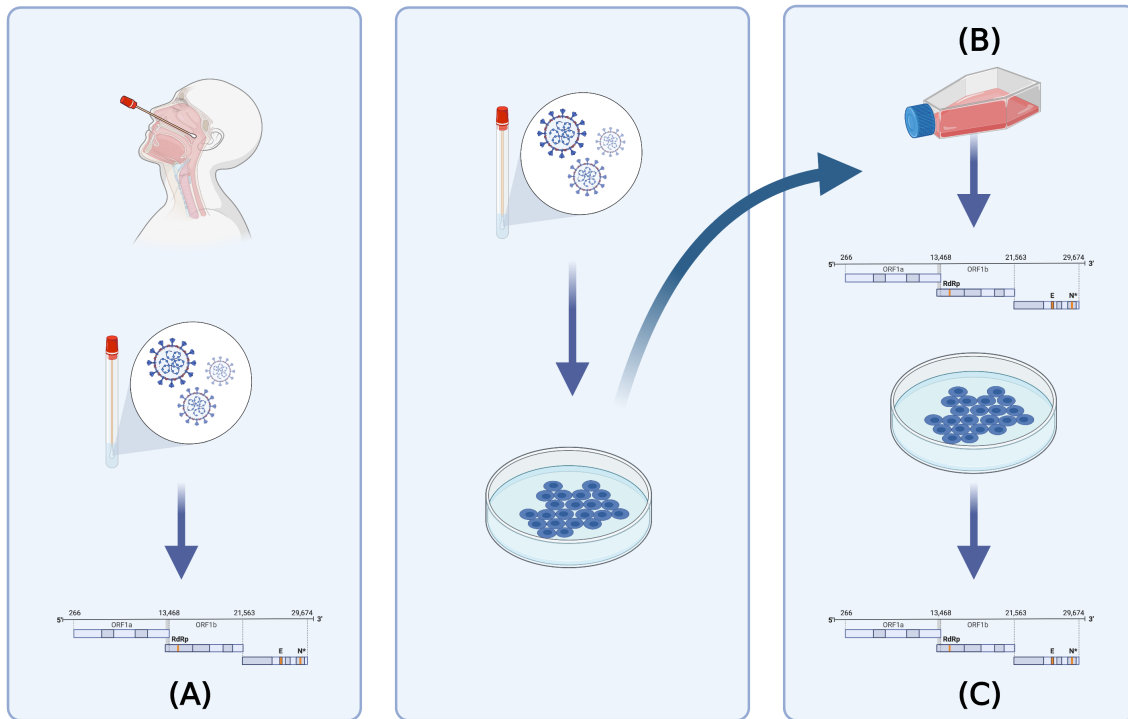
- 486 1. Gorbalenya AE, Baker SC, Baric RS, de Groot RJ, Drosten C, Gulyaeva AA, Haagmans BL,  
487 Lauber C, Leontovich AM, Neuman BW, Penzar D, Perlman S, Poon LLM, Samborskiy DV,  
488 Sidorov IA, Sola I, Ziebuhr J, Coronaviridae Study Group of the International Committee  
489 on Taxonomy of V. 2020. The species Severe acute respiratory syndrome-related  
490 coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiology*  
491 5:536-544.
- 492 2. Dorward DA, Russell CD, Um IH, Elshani M, Armstrong SD, Penrice-Randal R, Millar T,  
493 Lerpiniere CEB, Tagliavini G, Hartley CS, Randle NP, Gachanja NN, Potey PMD, Dong X,  
494 Anderson AM, Campbell VL, Duguid AJ, Al Qsous W, BouHaidar R, Baillie JK, Dhaliwal K,  
495 Wallace WA, Bellamy COC, Prost S, Smith C, Hiscox JA, Harrison DJ, Lucas CD. 2021. Tissue-  
496 Specific Immunopathology in Fatal COVID-19. *American journal of respiratory and critical*  
497 *care medicine* 203:192-201.
- 498 3. (WHO) WHO. 2021. WHO Coronavirus Disease (COVID-19) Dashboard, *on* WHO.  
499 <https://covid19.who.int/>. Accessed 03.02.21.
- 500 4. V'kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. 2021. Coronavirus biology and  
501 replication: implications for SARS-CoV-2. *Nature Reviews Microbiology* 19:155-170.
- 502 5. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. 2020. The proximal origin of  
503 SARS-CoV-2. *Nat Med* 26:450-452.

- 504 6. Gribble J, Stevens LJ, Agostini ML, Anderson-Daniels J, Chappell JD, Lu X, Pruijssers AJ,  
505 Routh AL, Denison MR. 2021. The coronavirus proofreading exoribonuclease mediates  
506 extensive viral recombination. *PLoS Pathog* 17:e1009226.
- 507 7. Robson F, Khan KS, Le TK, Paris C, Demirbag S, Barfuss P, Rocchi P, Ng W-L. 2020.  
508 Coronavirus RNA Proofreading: Molecular Basis and Therapeutic Targeting. *Molecular*  
509 *Cell* 79:710-727.
- 510 8. Di Giorgio S, Martignano F, Torcia MG, Mattiuz G, Conticello SG. 2020. Evidence for host-  
511 dependent RNA editing in the transcriptome of SARS-CoV-2. *Science Advances*  
512 6:eabb5813.
- 513 9. Carroll MW, Matthews DA, Hiscox JA, Elmore MJ, Pollakis G, Rambaut A, Hewson R,  
514 García-Dorival I, Bore JA, Koundouno R, Abdellati S, Afrough B, Aiyepada J, Akhilomen P,  
515 Asogun D, Atkinson B, Badusche M, Bah A, Bate S, Baumann J, Becker D, Becker-Ziaja B,  
516 Bocquin A, Borremans B, Bosworth A, Boettcher JP, Cannas A, Carletti F, Castilletti C, Clark  
517 S, Colavita F, Diederich S, Donatus A, Duraffour S, Ehichioya D, Ellerbrok H, Fernandez-  
518 Garcia MD, Fizet A, Fleischmann E, Gryseels S, Hermelink A, Hinzmann J, Hopf-Guevara U,  
519 Ighodalo Y, Jameson L, Kelterbaum A, Kis Z, Kloth S, Kohl C, Korva M, et al. 2015. Temporal  
520 and spatial analysis of the 2014–2015 Ebola virus outbreak in West Africa. *Nature* 524:97-  
521 101.
- 522 10. Cavanagh D, Picault JP, Gough R, Hess M, Mawditt K, Britton P. 2005. Variation in the spike  
523 protein of the 793/B type of infectious bronchitis virus, in the field and during alternate  
524 passage in chickens and embryonated eggs. *Avian Pathol* 34:20-5.

- 525 11. Jaimes JA, Whittaker GR. 2018. Feline coronavirus: Insights into viral pathogenesis based  
526 on the spike protein structure and function. *Virology* 517:108-121.
- 527 12. Millet JK, Goldstein ME, Labitt RN, Hsu HL, Daniel S, Whittaker GR. 2016. A camel-derived  
528 MERS-CoV with a variant spike protein cleavage site and distinct fusion activation  
529 properties. *Emerg Microbes Infect* 5:e126.
- 530 13. Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, Hengartner N,  
531 Giorgi EE, Bhattacharya T, Foley B, Hastie KM, Parker MD, Partridge DG, Evans CM,  
532 Freeman TM, de Silva TI, Angyal A, Brown RL, Carrilero L, Green LR, Groves DC, Johnson  
533 KJ, Keeley AJ, Lindsey BB, Parsons PJ, Raza M, Rowland-Jones S, Smith N, Tucker RM, Wang  
534 D, Wyles MD, McDanal C, Perez LG, Tang H, Moon-Walker A, Whelan SP, LaBranche CC,  
535 Sapphire EO, Montefiori DC. 2020. Tracking Changes in SARS-CoV-2 Spike: Evidence that  
536 D614G Increases Infectivity of the COVID-19 Virus. *Cell* 182:812-827.e19.
- 537 14. Liu H, Zhang Q, Wei P, Chen Z, Aviszus K, Yang J, Downing W, Peterson S, Jiang C, Liang B,  
538 Reynoso L, Downey GP, Frankel SK, Kappler J, Marrack P, Zhang G. 2021. The basis of a  
539 more contagious 501Y.V1 variant of SARS-COV-2. *bioRxiv : the preprint server for biology*  
540 [doi:10.1101/2021.02.02.428884](https://doi.org/10.1101/2021.02.02.428884):2021.02.02.428884.
- 541 15. NERVTAG. 2021. NERVTAG paper on COVID-19 variant of concern B.1.1.7, *on* NERVTAG  
542 update. [https://www.gov.uk/government/publications/nervtag-paper-on-covid-19-](https://www.gov.uk/government/publications/nervtag-paper-on-covid-19-variant-of-concern-b117)  
543 [variant-of-concern-b117](https://www.gov.uk/government/publications/nervtag-paper-on-covid-19-variant-of-concern-b117). Accessed 03.02.21.
- 544 16. Volz E, Mishra S, Chand M, Barrett JC, Johnson R, Hopkins S, Gandy A, Rambaut A,  
545 Ferguson NM. 2021. Transmission of SARS-CoV-2 Lineage B.1.1.7 in England: Insights  
546 from linking epidemiological and genetic data. <https://virological.org/t/transmission-of->

- 547 [sars-cov-2-lineage-b-1-1-7-in-england-insights-from-linking-epidemiological-and-](#)  
548 [genetic-data/576](#). Accessed 03/02/21.
- 549 17. Mohandas S, Yadav PD, Nyayanit D, Deshpande G, Shete-Aich A, Sapkal G, Kumar S, Jain  
550 R, Kadam M, Kumar A, Patil DY, Sarkale P, Gawande P, Abraham P. 2021. Comparison of  
551 the pathogenicity and virus shedding of SARS CoV-2 VOC 202012/01 and D614G variant  
552 in hamster model. bioRxiv doi:10.1101/2021.02.25.432136:2021.02.25.432136.
- 553 18. Brown JC, Goldhill DH, Zhou J, Peacock TP, Frise R, Goonawardane N, Baillon L, Kugathasan  
554 R, Pinto A, McKay PF, Hassard J, Moshe M, Singanayagam A, Burgoyne T, Barclay WS.  
555 2021. Increased transmission of SARS-CoV-2 lineage B.1.1.7 (VOC 202012/01) is not  
556 accounted for by a replicative advantage in primary airway cells or antibody escape.  
557 bioRxiv doi:10.1101/2021.02.24.432576:2021.02.24.432576.
- 558 19. Davidson AD, Williamson MK, Lewis S, Shoemark D, Carroll MW, Heesom KJ, Zambon M,  
559 Ellis J, Lewis PA, Hiscox JA, Matthews DA. 2020. Characterisation of the transcriptome and  
560 proteome of SARS-CoV-2 reveals a cell passage induced in-frame deletion of the furin-like  
561 cleavage site from the spike glycoprotein. *Genome Med* 12:68.
- 562 20. Young BE, Fong SW, Chan YH, Mak TM, Ang LW, Anderson DE, Lee CY, Amrun SN, Lee B,  
563 Goh YS, Su YCF, Wei WE, Kalimuddin S, Chai LYA, Pada S, Tan SY, Sun L, Parthasarathy P,  
564 Chen YYC, Barkham T, Lin RTP, Maurer-Stroh S, Leo YS, Wang LF, Renia L, Lee VJ, Smith  
565 GJD, Lye DC, Ng LFP. 2020. Effects of a major deletion in the SARS-CoV-2 genome on the  
566 severity of infection and the inflammatory response: an observational cohort study.  
567 *Lancet* 396:603-611.

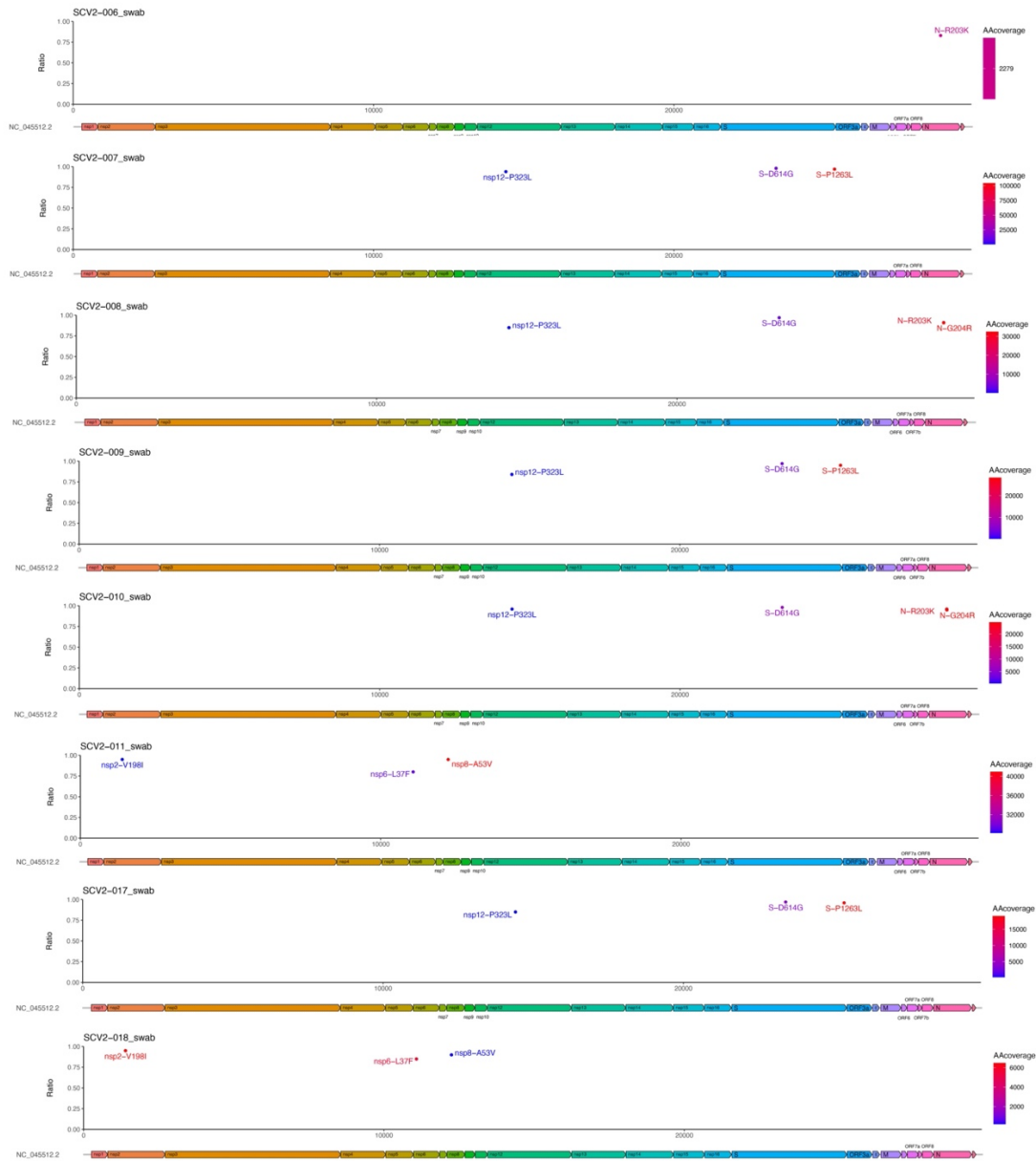
- 568 21. Dong X, Munoz-Basagoiti J, Rickett NY, Pollakis G, Paxton WA, Gunther S, Kerber R, Ng  
569 LFP, Elmore MJ, Magassouba N, Carroll MW, Matthews DA, Hiscox JA. 2020. Variation  
570 around the dominant viral genome sequence contributes to viral load and outcome in  
571 patients with Ebola virus disease. *Genome Biol* 21:238.
- 572 22. Hartley PD, Tillett RL, AuCoin DP, Sevinsky JR, Xu Y, Gorzalski A, Pandori M, Buttery E,  
573 Hansen H, Picker MA, Rossetto CC, Verma SC. 2021. Genomic surveillance of Nevada  
574 patients revealed prevalence of unique SARS-CoV-2 variants bearing mutations in the  
575 RdRp gene. *Journal of Genetics and Genomics*  
576 doi:<https://doi.org/10.1016/j.jgg.2021.01.004>.
- 577 23. Buchrieser J, Dufloo J, Hubert M, Monel B, Planas D, Rajah MM, Planchais C, Porrot F,  
578 Guivel-Benhassine F, Van der Werf S, Casartelli N, Mouquet H, Bruel T, Schwartz O. 2020.  
579 Syncytia formation by SARS-CoV-2-infected cells. *The EMBO Journal* 39:e106267.
- 580 24. Moore SC, Penrice-Randal R, Alruwaili M, Randle N, Armstrong S, Hartley C, Haldenby S,  
581 Dong X, Alrezaihi A, Almsaud M, Bentley E, Clark J, Garcia-Dorival I, Gilmore P, Han X,  
582 Jones B, Luu L, Sharma P, Shawli G, Sun Y, Zhao Q, Pullan ST, Carter DP, Bewley K, Dunning  
583 J, Zhou EM, Solomon T, Beadsworth M, Cruise J, Crook DW, Matthews DA, Davidson AD,  
584 Mahmood Z, Aljabr W, Druce J, Vipond R, Ng L, Renia L, Openshaw PJM, Baillie JK, Carroll  
585 MW, Stewart J, Darby A, Semple M, Turtle L, Hiscox JA. 2020. Amplicon-Based Detection  
586 and Sequencing of SARS-CoV-2 in Nasopharyngeal Swabs from Patients With COVID-19  
587 and Identification of Deletions in the Viral Genome That Encode Proteins Involved in  
588 Interferon Antagonism. *Viruses* 12.
- 589



590

591 **Figure 1.** Testing strategy. (A) Nasopharyngeal swabs from patients with COVID-19 recruited to  
592 the ISARIC-4C study were sequenced using an amplicon based approach on the Oxford Nanopore  
593 MinION (P0). Virus was isolated from the same nasopharyngeal swabs(P1). (B) Viral isolates from  
594 the ISARIC-4C study, B.1.1.7, B.1.351 and Victoria isolates were grown up into stocks which were  
595 then sequenced. (C) Viral stocks were titrated and used to infected hACE2-A549 cells, and 72-  
596 hour post infection supernatants were sequenced.





597

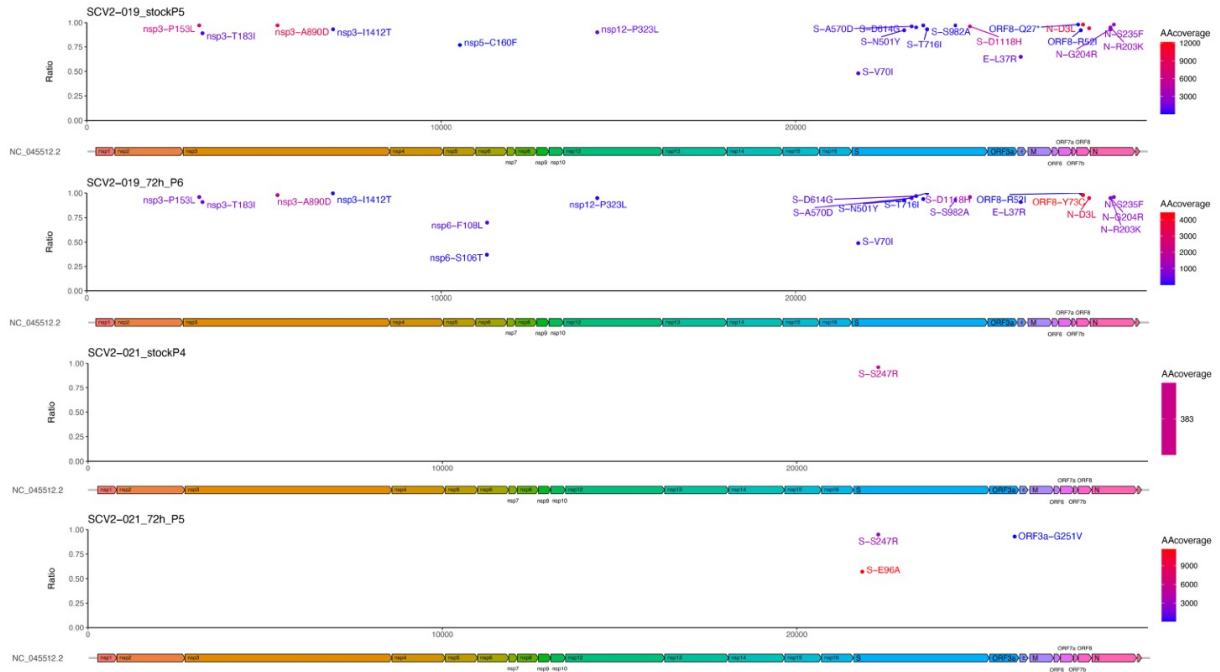
598 **Figure 2.** (A) Comparison of the consensus sequence SARS-CoV-2 in ISARIC4C swabs collected

599 from patients in hospital with COVID-19. Variations from the Wuhan reference sequence are

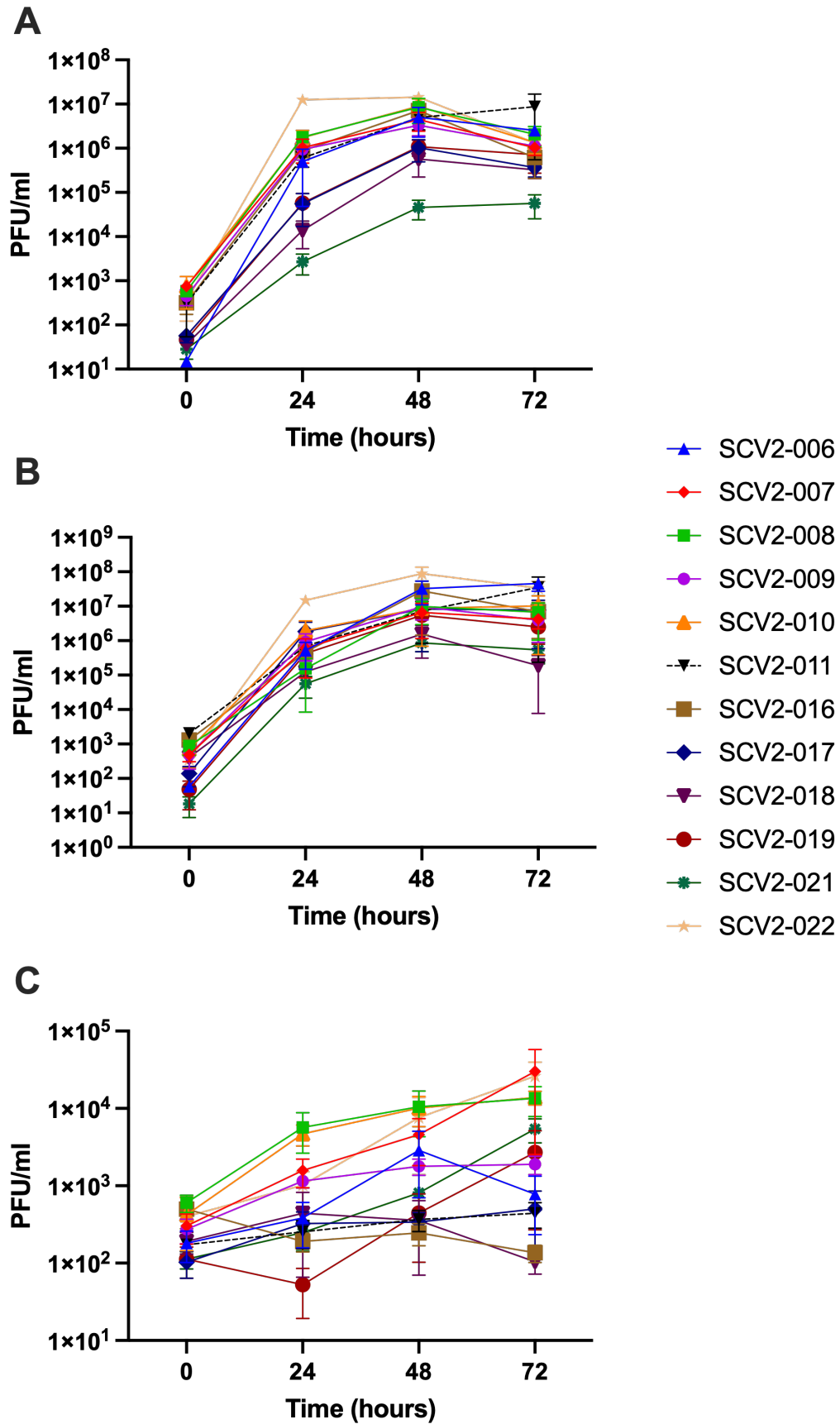
600 plotted with the location on the viral genome against the percentage (ratio) of reads where the

601 variation is observed. Variants are coloured to demonstrate the number of reads achieved (>10).

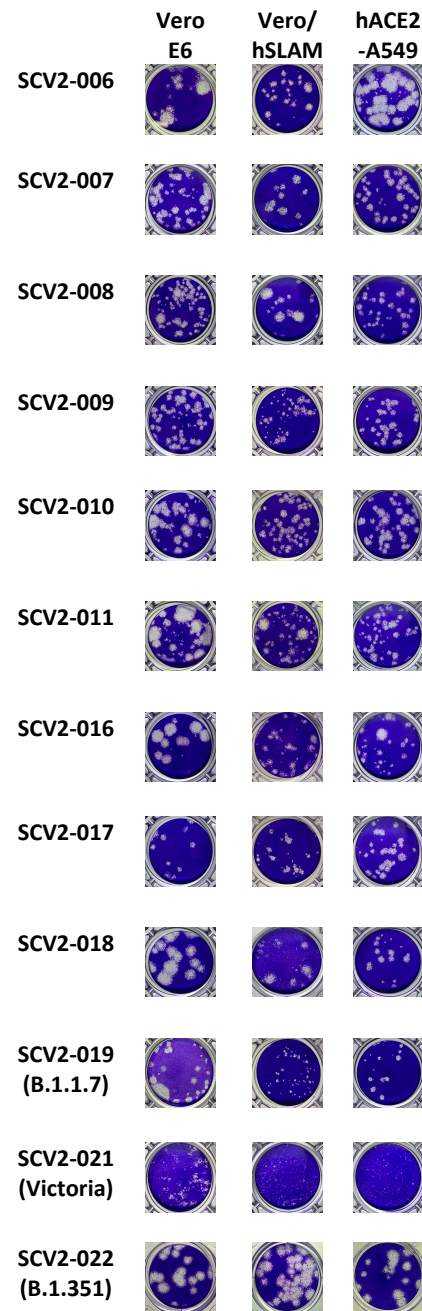
602



603  
604 **Figure 3.** Comparison of the UK 'Kent' VOC (SCV2-019) and the Australian Victoria isolate (SCV2-  
605 021) with the Wuhan reference sequence. Variations from the Wuhan reference sequence are  
606 plotted with the location on the viral genome against the percentage (ratio) of reads where the  
607 variation is observed. Variants are coloured to demonstrate the number of reads achieved (> 10).



609 **Figure 4.** Growth over time of 11 different viral isolates in three different cell lines compared to  
610 the Variants of Concern SCV2-019 (UK 'Kent' VOC) and SCV2-022 ('South African' VOC).  
611 Comparison viruses included the Australian Victoria variant (SCV2-021). (A) Growth of viruses in  
612 plaque-forming units (PFU) per ml over times in Vero E6 African green monkey kidney cells. (B)  
613 Growth of viruses in Vero cells expressing the human signalling lymphocytic activation module  
614 (SLAM) gene (Vero/hSLAM). (C) Growth of viruses in human ACE-2 expressing A549 cells (hACE2-  
615 A549). All experiments were repeated in triplicate using supernatant from 6 wells (n=3).



617 **Figure 5.** Phenotypic appearance of plaque assays from variants grown in three different cell  
618 lines; (i) Vero E6, (ii) Vero/hSLAM and (iii) hACE2-A549 cells. Plaque assays were performed on  
619 VeroE6 cells. Variants of Concern are SCV2-019 (UK 'Kent' VOC) and SCV2-022 ('South African'  
620 VOC). Comparison viruses include the Australian Victoria variant (SCV2-021).