SARS-CoV-2 B.1.617 emergence and sensitivity to vaccine-elicited antibodies

Isabella Ferreira[1,2]*, Rawlings Datir[1,2]*, Steven Kemp[1,2]*, Guido Papa[3]*, Partha Rakshit[4*], Sujeet Singh[4], Bo Meng[1,2], Rajesh Pandey[5], Kalaiarasan Ponnusamy[4], V.S. Radhakrishnan[4], The Indian SARS-CoV-2 Genomics Consortium (INSACOG), The CITIID-NIHR BioResource COVID-19 Collaboration, Kei Sato[7], Leo James[3*], Anurag Agrawal[5*], Ravindra K. Gupta[1,2, 8*]

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

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

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

[4] National Centre for Disease Control, Delhi, India

[5] CSIR Institute of Genomics and Integrative Biology, Delhi, India

[6] Institute of Medical Science, University of Tokyo, Japan

[7] Africa Health Research Institute, Durban, South Africa.

*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; B.1.617; Indian variant; antibody escape; neutralising antibodies; infectivity; spike mutation; evasion; resistance; fitness;

Abstract

The B.1.617 variant emerged in the Indian state of Maharashtra in late 2020 and has spread throughout India and to at least 40 countries. There have been fears that two key mutations seen in the receptor binding domain L452R and E484Q would have additive effects on evasion of neutralising antibodies. Here we delineate the phylogenetics of B.1.617 and spike mutation frequencies, in the context of others bearing L452R. The defining mutations in B.1.617.1 spike are L452R and E484Q in the RBD that interacts with ACE2 and is the target of neutralising antibodies. All B.1.617 viruses have the P681R mutation in the polybasic cleavage site region in spike. We report that B.1.617.1 spike bearing L452R, E484Q and P681R mediates entry into cells with slightly reduced efficiency compared to Wuhan-1. This spike confers modestly reduced sensitivity to BNT162b2 mRNA vaccine-elicited antibodies that is similar in magnitude to the loss of sensitivity conferred by L452R or E484Q alone. Furthermore we show that the P681R mutation significantly augments syncytium formation upon the B.1.617.1 spike protein. These data demonstrate that reduced sensitivity to vaccine elicited neutralising antibodies likely contributes to vaccine breakthrough observed in India, and that polybasic cleavage site mutations potentially contribute to infectivity/transmissibility.

Introduction

Global control of the SARS-CoV-2 pandemic has yet to be realised despite availability of highly effective vaccines[1,2]. Emergence of new variants with multiple mutations is likely the result of chronic infections within individuals who are immune compromised[3]. New variant emergence and transmission has coincided with rollout of vaccines, potentially threatening their success in controlling the pandemic[4,5].

India experienced a wave of infections in mid 2020 and was controlled by a nationwide lockdown. Since easing of restrictions, India has seen expansion in cases of COVID-19 since March 2021. The B.1.1.7 variant has been growing in the north of the country and is known to be more transmissible than previous viruses bearing the D614G spike mutation[6]. The B.1.617 variant emerged in the state of Maharashtra in late 2020/early 2021[7] and has spread throughout India and to at least 40 countries. It was labelled initially as a double mutant since two of the mutations L452R and E484Q were matched to an in-house

screening database for mutations leading to probable evasion of antibodies and/or being linked to increased transmissibility.

The defining mutations in spike are L452R and E484Q in the critical receptor binding domain that interacts with ACE2[8]. L452R was observed in the California variant and is associated with increase in viral load and around 20% increased transmissibility [9]. It was also associated with increased ACE2 binding, increased infectivity[10] and 3-6 fold loss of neutralisation sensitivity to vaccine elicited sera in experiments with pseudotyped virus (PV) particles[10,11]. Little is known about E484Q, though E484K is a defining feature of two VOCs, B.1.351[12] and P.1[13]. E484K is known to confer around 10 fold loss of sensitivity to neutralising antibodies in vaccine and convalescent sera[14,15]. In contrast to L452R, E484Q has not been observed in transmissible variants.

Here we demonstrate three lineages of B.1.617, all bearing the L452R mutation. We report key differences in amino acids between sub-lineages and focus on B.1.617.1 bearing three key mutations: L452R, E484Q and P681R. We report an outbreak of infections dominated by B.1.617 viruses in fully vaccinated HCWs. In vitro, we find modestly reduced sensitivity of the ancestral B.1.617.1 spike protein to BNT162b2 mRNA vaccine-elicited antibodies that is similar in magnitude to the loss of sensitivity conferred by L452R or E484Q alone. Furthermore we show that the P681R mutation significantly augments syncytium formation upon the B.1.617.1 spike protein, potentially contributing to increased pathogenesis observed in hamsters[16] and high growth rates observed in humans.

Results

Three B.1.617 sub-lineages are characterised by L452R and P681R in spike

We downloaded whole genome SARS-CoV-2 sequences (excluding low-quality or >5% N regions) containing L452R from GISAID, subsampled and inferred a maximum likelihood phylogenetic tree (Figure 1A). We annotated the sequences based on the accompanying mutations and observed that B.1.617.1 has three key spike mutations L452R, E484Q and P681R, whereas B.1.617.2 is characterised by L452R, T478K and P681R. This likely signifies loss of E484Q in B.1.617.2 given that B.1.617.3 also bears E484Q in 90% of sequences (Figure 1). The number of sequenced isolates of B.1.617.1 and B.1.617.2 has been steadily

increasing both in India (Figure 1B) and in the UK (Figure 1C), though with the caveat of very low sequencing of prevalent cases. There is already significant diversity in B.1.617.1 as demonstrated in Figure 1D.

B1.617 associated vaccine breakthrough in health care workers

Vaccination of health care workers (HCW) in Delhi was started in early 2021, with the ChdOx-1 vaccine. Surveillance has suggested B.1.1.7 dominance in the Delhi area during early 2021 (Figure 2A), with growth of B.1.617 since late March 2021. During the wave of infections during March and April an outbreak of SARS-CoV-2 was confirmed in 30 vaccinated staff members at a single tertiary centre (age range 27-77 years). Short-read sequencing[17] revealed the majority were B.1.617.2 with a range of other B lineage viruses including B.1.1.7 (Figure 2B). Further analysis of pairwise differences demonstrated a group of highly related, and in some cases, genetically indistinct sequences, differing by only one or two nucleotides (Figure 2C). Maximum likelihood phylogenetic analysis of consensus sequences from symptomatic HCW breakthrough infections revealed that the twelve B.1.617.2 viruses were almost identical and were sampled within one or two days of each other. These data are consistent with a single transmission from an infected individual (Figure 2D). In approximately the same timeframe, numerous other lineages of virus, including the B.1.1.7 variant were detected in the same hospital. Importantly no severe cases were documented in this event. To put the outbreak sequences into context, a further phylogeny was inferred with a random subsample of Indian B.1.617+ sequences downloaded from GISAID (Figure 2E).

B.1.617.1 Spike confers partial evasion of BNT162b2 vaccine elicited antibodies

Spike mutations L452R and E484Q are in the receptor binding domain that not only binds ACE2[8], but is a target for neutralising antibodies[18,19] (Figure 3A-D). We tested the neutralisation sensitivity of B.1.617.1 Spike bearing L452R, E484Q and P681R using a previously reported pseudotyped virus (PV) system with HIV-1 particles bearing SARS-CoV-2 spike (Figure 3E). PV testing against neutralising antibodies has been shown to be highly correlated with live virus systems[20]. We tested nine stored sera from Pfizer BNT162b2 vaccinees against a range of spike mutation bearing PV. As expected E484K conferred a ten-fold reduction in neutralisation by vaccine sera, and E484Q had a slightly milder yet

significant impact. When E484Q and L452R were combined, there was a statistically significant loss of sensitivity as compared to wild type, but the fold change was similar to that observed with each mutation individually with no evidence for an additive effect. We found similar relationships when we tested our mutants convalescent plasma

P681R confers increased syncytium formation capability on B.1.617.1 spike

Spike is known to mediate cell entry via interaction with ACE2 and TMPRSS2[21] and is a major determinant of viral infectivity. Given that that B.1.617.1 does not appear to be highly immune evasive, we hypothesised that it may be have higher infectivity. We tested single round viral entry using the PV system, infecting target 293T cells over-expressing ACE2 and TMPRSS2, as well as Calu-3 lung cells expressing endogenous levels of ACE2 and TMPRSS2 (Supplementary figure 1). We observed similar entry efficiency across mutants, all of which appeared lower than the Wuhan-1 D614G wild type (Figure 4).

SARS-CoV-2 infection in clinically relevant cells is TMPRSS2 dependent and requires fusion at the plasma membrane, potentially to avoid restriction factors in endosomes[22]. The plasma membrane route of entry, and indeed transmissibility in animal models, is critically dependent on the polybasic cleavage site (PBCS) between S1 and S2[22,23]. Mutations at P681 in the PBCS have been observed in multiple SARS-CoV-2 lineages, most notably in the B.1.1.7 variant that likely emerged in the UK.

We previously showed that B.1.1.7 spike, bearing P681H, had significantly higher fusogenic potential than a D614G Wuhan-1 virus[24]. We therefore tested a series of mutations in B.1.617 spike using a split GFP system. We transfected spike bearing plasmids into donor cells and co-cultured them with acceptor cells (Figure 5). We found no significant differences for the single mutants tested L452R, E484Q/K, or indeed the double mutant L452R+E484Q. However, the triple mutant with P681R demonstrated higher fusion activity and syncytium formation, mediated specifically by P681R (Figure 5, supplementary figure 2).

Discussion

Here we have shown that B.1.617 spike bearing L452R, E484Q and P681R has modest ability to avoid neutralising antibodies elicited by BNT162b2 vaccination. The fold reduction for the two RBD mutations L452R and E484Q was no greater than the individual mutations alone, arguing against use of the term 'double mutant'. The loss of neutralisation of B.1.617 has possibly contributed to an epidemic wave in India where background infection to the Wuhan-1 D614G in 2020 was between 20-50%[25]. Despite in vitro data showing only small loss of neutralisation against B.1.617 with the Covaxin vaccine[26], here we also show vaccine breakthrough in health care workers at a single tertiary hospital who were fully vaccinated with ChAdOx-1 vaccine (adenovirus vectored). These infections were predominantly B.1.617, with a mix of other lineages bearing D614G in spike. The dominance of B.1.617 in this outbreak could be explained by prevalence of this lineage in community infection or reflect transmission between HCWs. The data nonetheless raise the possibility of a transmission advantage of B.1.617 in vaccinated individuals.

We measured spike mediated entry into target cell lines exogenously or endogenously expressing ACE2 and TMPRSS2 receptors. The E484K, L452R and P681R mutant had reduced entry efficiency relative to wild type. A recent report using a spike B.1.617 with a larger set of mutations found variable entry efficiency relative to wild type across cell types[27].

Virus infectivity and fusogenicity mediated by the PBCS is a key determinant of pathogenicity and transmissibility[22,28] and there are indications that giant cells/syncitia formation are associated with fatal disease[29]. We find that P681R is associated with enhanced capacity to induce cell-cell fusion and syncitia formation, and that P681R alone confers this ability on the B.1.617.1 spike with RBD mutations L452R and E484Q.

It is unclear whether B.1.617 variants will prove more transmissible than B.1.1.7, also circulating in India and now globally dominant. In the absence of published data on transmissibility of B.1.617 we predict that this variant will have a transmission advantage relative to Wuhan-1 with D614G in individuals with pre-existing immunity from vaccines/natural infection as well as in settings where there is low vaccine coverage and low prior exposure. Lower protection against B.1.351, the variant with least sensitivity to neutralising antibodies, has been demonstrated for at least three vaccines[2,30-32]. However,

progression to severe disease and death was low in all studies. Therefore, at population scale, extensive vaccination will likely protect against moderate to severe disease and will reduce transmission of B.1.617 given the *in vitro* neutralisation data we and others have presented.

Methods

Phylogenetic Analysis

All sequences excluding low-quality sequences (>5% N regions) with the L452R mutation were downloaded from https://gisaid.org [33] on the 4th May 2021 and manually aligned to reference strain MN908947.3 with mafft v4.475 [34] using the --keeplength --addfragments option. Sequences were de-duplicated using bbtools dedupe.sh. Due to the high proportion of USA-centric sequences containing L452R, all USA sequences were extracted and saved to a separate fasta file. A random subset of 400 global sequences (excluding USA), and 100 USA sequences were then selected with seqtk and concatenated. Sequence lineages were assigned to all sequences with pangolin v2.4 (https://github.com/cov-lineages/pangolin) and pangolearn (04/05/2021). A secondary set of 30 sequences were donated by collaborators from a hospital outbreak in Delhi. Upon examination, three were discarded due to >50% N regions in the genomes. A further three were excluded from analysis as they could not reliably be assigned a lineage due to numerous gappy regions throughout the genome. In total, 24 sequences remained.

Phylogenies were then inferred using maximum-likelihood in IQTREE v2.1.3 [35] using a GTR+R6 model and the -fast option. Mutations of interest were determined using a local instance of nextclade-cli v0.14.2 (https://github.com/nextstrain/nextclade). The inferred phylogeny was annotated in R v4.04 using ggtree v2.2.4[36] and rooted on the SARS-CoV-2 reference sequence, and nodes arranged in descending order. Major lineages were annotated on the phylogeny, as well as a heatmap indicating which mutations of interest were carried by each viral sequence.

Structural Analyses

The PyMOL Molecular Graphics System v.2.4.0 (https://github.com/schrodinger/pymol-open-source/releases) was used to map the location of the two RDB mutants L452R and

E484Q onto two previously published SARS-CoV-2 spike glycoprotein structures. Th two structures included a closed-conformation spike protein - PDB: 6ZGE [37] and a spike protein in open conformation, bound to nAb H4 [38].

Serum samples and ethical approval

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

Cells

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

Pseudotype virus preparation

Plasmids encoding the spike protein of SARS-CoV-2 D614 with a C terminal 19 amino acid deletion with D614G, were used as a template to produce variants lacking amino acids at position H69 and V70, as well as mutations N439K and Y453F. Mutations were introduced using Quickchange Lightning Site-Directed Mutagenesis kit (Agilent) following the manufacturer's instructions. B.1.1.7 S expressing plasmid preparation was described previously, but in brief was generated by step wise mutagenesis. Viral vectors were prepared by transfection of 293T cells by using Fugene HD transfection reagent (Promega). 293T cells were transfected with a mixture of 11ul of Fugene HD, 1µg of pCDNAΔ19 spike-HA, 1ug of p8.91 HIV-1 gag-pol expression vector and 1.5µg of pCSFLW (expressing the firefly luciferase reporter gene with the HIV-1 packaging signal). Viral supernatant was collected at 48 and 72h after transfection, filtered through 0.45um filter and stored at -80˚C as previously described. Infectivity was measured by luciferase detection in target 293T cells transfected with TMPRSS2 and ACE2.

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

The reverse transcriptase activity of virus preparations was determined by qPCR using a SYBR Green-based product-enhanced PCR assay (SG-PERT) as previously described[39]. Briefly,

10-fold dilutions of virus supernatant were lysed in a 1:1 ratio in a 2x lysis solution (made up of 40% glycerol v/v 0.25% Trition X-100 v/v 100mM KCl, RNase inhibitor 0.8 U/ml, TrisHCL 100mM, buffered to pH7.4) for 10 minutes at room temperature.

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

Plasmids for split GFP system to measure cell-cell fusion

pQCXIP-BSR-GFP11 and pQCXIP-GFP1-10 were from Yutaka Hata [40] Addgene plasmid #68716; http://n2t.net/addgene:68716; RRID:Addgene_68716 and Addgene plasmid #68715; http://n2t.net/addgene:68715; RRID:Addgene_68715)

Generation of GFP1-10 or GFP11 lentiviral particles

Lentiviral particles were generated by co-transfection of 293T cells with pQCXIP-BSR-GFP11 or pQCXIP-GFP1-10 as previously described [41]. Supernatant containing virus particles was harvested after 48 and 72 hours, 0.45 µm filtered, and used to infect 293T or Vero cells to generate stable cell lines. 293T and Vero cells were transduced to stably express GFP1-10 or GFP11 respectively and were selected with 2 µg/ml puromycin.

Cell-cell fusion assay

Cell-cell fusion assay was carried out as previously described [41,42] but using a Split-GFP system. Briefly, 293T-GFP1-10 and Vero-GFP11 cells were seeded at 80% confluence in a 24 multiwell plate the day before. 293T cells were co-transfected with 1.5 µg of spike expression plasmids in pCDNA3 using Fugene 6 and following the manufacturer's instructions (Promega). 293T-GFP1-10 cells were then detached 5 hours post transfection, mixed together with the Vero-GFP11 cells, and plated in a 12 multiwell plate. Cell-cell fusion was measured using an Incucyte and determined as the proportion of green area to total phase area. Data were then analysed using Incucyte software analysis. Graphs were

generated using Prism 8 software. Furin inhibitor CMK (Calbiochem) was added at transfection.

Western blotting

Cells were lysed and supernatants collected 18 hours post transfection. Purified virions were prepared by harvesting supernatants and passing through a 0.45 µm filter. Clarified supernatants were then loaded onto a thin layer of 8.4% optiprep density gradient medium (Sigma-Aldrich) and placed in a TLA55 rotor (Beckman Coulter) for ultracentrifugation for 2 hours at 20,000 rpm. The pellet was then resuspended for western blotting. Cells were lysed with cell lysis buffer (Cell signalling), treated with Benzonase Nuclease (70664 Millipore) and boiled for 5 min. Samples were then run on 4%–12% Bis Tris gels and transferred onto nitrocellulose or PVDF membranes using an iBlot or semidry (Life Technologies and Biorad, respectively).

Membranes were blocked for 1 hour in 5% non-fat milk in PBS + 0.1% Tween-20 (PBST) at room temperature with agitation, incubated in primary antibody (anti-SARS-CoV-2 Spike, which detects the S2 subunit of SARS-CoV-2 S (Invitrogen, PA1-41165), anti-GAPDH (proteintech) or anti-p24 (NIBSC)) diluted in 5% non-fat milk in PBST for 2 hours at 4°C with agitation, washed four times in PBST for 5 minutes at room temperature with agitation and incubated in secondary antibodies anti-rabbit HRP (1:10000, Invitrogen 31462), anti-bactin HRP (1:5000; sc-47778) diluted in 5% non-fat milk in PBST for 1 hour with agitation at room temperature. Membranes were washed four times in PBST for 5 minutes at room temperature and imaged directly using a ChemiDoc MP imaging system (Bio-Rad).

Serum pseudotype neutralisation assay

Spike pseudotype assays have been shown to have similar characteristics as neutralisation testing using fully infectious wild type SARS-CoV-2[20].Virus neutralisation assays were performed on 293T cell transiently transfected with ACE2 and TMPRSS2 using SARS-CoV-2 spike pseudotyped virus expressing luciferase[43]. Pseudotyped virus was incubated with serial dilution of heat inactivated human serum samples or convalescent plasma in duplicate for 1h at 37˚C. Virus and cell only controls were also included. Then, freshly trypsinized 293T

ACE2/TMPRSS2 expressing cells were added to each well. Following 48h incubation in a 5% CO_2 environment at 37°C, the luminescence was measured using Steady-Glo Luciferase assay system (Promega).

Acknowledgments

We would like to thank the Department of Biotechnology, NCDC, 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. This study was supported by the Cambridge NIHRB Biomedical Research Centre. SAK is supported by the Bill and Melinda Gates Foundation via PANGEA grant: OPP1175094. We would like to thank Paul Lehner for Calu-3 cells. We thank Wendy Barclay and Thomas Peacock for helpful discussions and the Geno2pheno UK consortium. This study was also partly funded by Rosetrees Trust.

References

1 Polack, F. P. *et al.* Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New England Journal of Medicine* **383**, 2603-2615, doi:10.1056/NEJMoa2034577 (2020).

2 Abu-Raddad, L. J., Chemaitelly, H. & Butt, A. A. Effectiveness of the BNT162b2 Covid-19 Vaccine against the B.1.1.7 and B.1.351 Variants. *New England Journal of Medicine*, doi:10.1056/NEJMc2104974 (2021).

3 Kemp, S. A. *et al.* SARS-CoV-2 evolution during treatment of chronic infection. *Nature*, doi:10.1038/s41586-021-03291-y (2021).

4 Gupta, R. K. Will SARS-CoV-2 variants of concern affect the promise of vaccines? *Nature reviews. Immunology*, doi:10.1038/s41577-021-00556-5 (2021).

5 Hacisuleyman, E. *et al.* Vaccine Breakthrough Infections with SARS-CoV-2 Variants. *New England Journal of Medicine*, doi:10.1056/NEJMoa2105000 (2021).

6 Volz, E. *et al.* Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. *Nature*, doi:10.1038/s41586-021-03470-x (2021).

7 Cherian, S. *et al.* Convergent evolution of SARS-CoV-2 spike mutations, L452R, E484Q and P681R, in the second wave of COVID-19 in Maharashtra, India. *bioRxiv*, 2021.2004.2022.440932, doi:10.1101/2021.04.22.440932 (2021).

8 Starr, T. N. *et al.* Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* **182**, 1295-1310 e1220, doi:10.1016/j.cell.2020.08.012 (2020).

9 Deng, X. *et al.* Transmission, infectivity, and antibody neutralization of an emerging SARS-CoV-2 variant in California carrying a L452R spike protein mutation. *medRxiv*, doi:10.1101/2021.03.07.21252647 (2021).

10 Motozono, C. *et al.* An emerging SARS-CoV-2 mutant evading cellular immunity and increasing viral infectivity. *bioRxiv*, 2021.2004.2002.438288, doi:10.1101/2021.04.02.438288 (2021).

11 McCallum, M. *et al.* SARS-CoV-2 immune evasion by variant B.1.427/B.1.429. *bioRxiv*, 2021.2003.2031.437925, doi:10.1101/2021.03.31.437925 (2021).

12 Tegally, H. *et al.* Major new lineages of SARS-CoV-2 emerge and spread in South Africa during lockdown. *medRxiv*, 2020.2010.2028.20221143, doi:10.1101/2020.10.28.20221143 (2020).

13 Faria, N. R. *et al.* Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science*, eabh2644, doi:10.1126/science.abh2644 (2021).

14 Collier, D. A. *et al.* SARS-CoV-2 B.1.1.7 sensitivity to mRNA vaccine-elicited, convalescent and monoclonal antibodies. *Nature*, in press, doi:10.1101/2021.01.19.21249840 (2021).

15 Greaney, A. J. *et al.* Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. *Cell host & microbe* **29**, 463-476 e466, doi:10.1016/j.chom.2021.02.003 (2021).

16 Yadav, P. D. *et al.* SARS CoV-2 variant B.1.617.1 is highly pathogenic in hamsters than B.1 variant. *bioRxiv*, 2021.2005.2005.442760, doi:10.1101/2021.05.05.442760 (2021).

17 Bhoyar, R. C. *et al.* High throughput detection and genetic epidemiology of SARS-CoV-2 using COVIDSeq next-generation sequencing. *PloS one* **16**, e0247115, doi:10.1371/journal.pone.0247115 (2021).

18 Barnes, C. O. *et al.* SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682-687, doi:10.1038/s41586-020-2852-1 (2020).

19 Barnes, C. O. *et al.* Structures of Human Antibodies Bound to SARS-CoV-2 Spike Reveal Common Epitopes and Recurrent Features of Antibodies. *Cell* **182**, 828-842 e816, doi:10.1016/j.cell.2020.06.025 (2020).

20 Schmidt, F. *et al.* Measuring SARS-CoV-2 neutralizing antibody activity using pseudotyped and chimeric viruses. 2020.2006.2008.140871, doi:10.1101/2020.06.08.140871 %J bioRxiv (2020).

21 Hoffmann, M. *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280 e278, doi:10.1016/j.cell.2020.02.052 (2020).

22 Peacock, T. P. *et al.* The furin cleavage site in the SARS-CoV-2 spike protein is required for transmission in ferrets. *Nat Microbiol*, doi:10.1038/s41564-021-00908-w (2021).

23 Papa, G. *et al.* Furin cleavage of SARS-CoV-2 Spike promotes but is not essential for infection and cell-cell fusion. *PLoS pathogens* **17**, e1009246, doi:10.1371/journal.ppat.1009246 (2021).

24 Kemp, S. A. *et al.* Recurrent emergence and transmission of a SARS-CoV-2 spike deletion H69/V70. *bioRxiv*, 2020.2012.2014.422555, doi:10.1101/2020.12.14.422555 (2021).

25 Malani, A. *et al.* Seroprevalence of SARS-CoV-2 in slums versus non-slums in Mumbai, India. *Lancet Glob Health* **9**, e110-e111, doi:10.1016/S2214-109X(20)30467-8 (2021).

26 Yadav, P. D. *et al.* Neutralization of variant under investigation B.1.617 with sera of BBV152 vaccinees. *bioRxiv*, 2021.2004.2023.441101, doi:10.1101/2021.04.23.441101 (2021).

27 Hoffmann, M. *et al.* SARS-CoV-2 variant B.1.617 is resistant to Bamlanivimab and evades antibodies induced by infection and vaccination. *bioRxiv*, 2021.2005.2004.442663, doi:10.1101/2021.05.04.442663 (2021).

28 Johnson, B. A. *et al.* Loss of furin cleavage site attenuates SARS-CoV-2 pathogenesis. *Nature* **591**, 293-299, doi:10.1038/s41586-021-03237-4 (2021).

29 Braga, L. *et al.* Drugs that inhibit TMEM16 proteins block SARS-CoV-2 Spike-induced syncytia. *Nature*, doi:10.1038/s41586-021-03491-6 (2021).

30 Shinde, V. *et al.* Efficacy of NVX-CoV2373 Covid-19 Vaccine against the B.1.351 Variant. *New England Journal of Medicine*, doi:10.1056/NEJMoa2103055 (2021).

31 Madhi, S. A. *et al.* Efficacy of the ChAdOx1 nCoV-19 Covid-19 Vaccine against the B.1.351 Variant. *N Engl J Med*, doi:10.1056/NEJMoa2102214 (2021).

32 Sadoff, J. *et al.* Safety and Efficacy of Single-Dose Ad26.COV2.S Vaccine against Covid-19. *New England Journal of Medicine*, doi:10.1056/NEJMoa2101544 (2021).

33 Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data - from vision to reality. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* **22**, 30494, doi:10.2807/1560-7917.ES.2017.22.13.30494 (2017).

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

35 Minh, B. Q. *et al.* IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *bioRxiv*, 849372, doi:10.1101/849372 (2019).

36 Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T. Y. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* **8**, 28-36 (2017).

37 Wrobel, A. G. *et al.* SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects. *Nat Struct Mol Biol* **27**, 763-767, doi:10.1038/s41594-020-0468-7 (2020).

38 Rapp, M. *et al.* Modular basis for potent SARS-CoV-2 neutralization by a prevalent VH1-2-derived antibody class. *Cell Reports* **35**, 108950, doi:https://doi.org/10.1016/j.celrep.2021.108950 (2021).

39 Vermeire, J. *et al.* Quantification of reverse transcriptase activity by real-time PCR as a fast and accurate method for titration of HIV, lenti- and retroviral vectors. *PloS one* **7**, e50859-e50859, doi:10.1371/journal.pone.0050859 (2012).

40 Kodaka, M. *et al.* A new cell-based assay to evaluate myogenesis in mouse myoblast C2C12 cells. *Experimental cell research* **336**, 171-181 (2015).

41 Papa, G. *et al.* Furin cleavage of SARS-CoV-2 Spike promotes but is not essential for infection and cell-cell fusion. *PLoS Pathogens* **17**, e1009246 (2021).

42 Buchrieser, J. *et al.* Syncytia formation by SARS-CoV-2-infected cells. *The EMBO journal* **39**, e106267 (2020).

43 Mlcochova, P. *et al.* Combined point of care nucleic acid and antibody testing for SARS-CoV-2 following emergence of D614G Spike Variant. *Cell Rep Med*, 100099, doi:10.1016/j.xcrm.2020.100099 (2020).

COG-UK CONSORTIUM

Funding acquisition, Leadership and supervision, Metadata curation, Project administration, Samples and logistics, Sequencing and analysis, Software and analysis tools, and Visualisation:

Samuel C Robson [13].

Funding acquisition, Leadership and supervision, Metadata curation, Project administration, Samples and logistics, Sequencing and analysis, and Software and analysis tools:

Nicholas J Loman [41] and Thomas R Connor [10, 69].

Leadership and supervision, Metadata curation, Project administration, Samples and logistics, Sequencing and analysis, Software and analysis tools, and Visualisation:

Tanya Golubchik [5].

Funding acquisition, Metadata curation, Samples and logistics, Sequencing and analysis, Software and analysis tools, and Visualisation:

Rocio T Martinez Nunez [42].

Funding acquisition, Leadership and supervision, Metadata curation, Project administration, and Samples and logistics:

Catherine Ludden [88].

Funding acquisition, Leadership and supervision, Metadata curation, Samples and logistics, and Sequencing and analysis:

Sally Corden [69].

Funding acquisition, Leadership and supervision, Project administration, Samples and logistics, and Sequencing and analysis:

Ian Johnston [99] and David Bonsall [5].

Funding acquisition, Leadership and supervision, Sequencing and analysis, Software and analysis tools, and Visualisation:

Colin P Smith [87] and Ali R Awan [28].

Funding acquisition, Samples and logistics, Sequencing and analysis, Software and analysis tools, and Visualisation:

Giselda Bucca [87].

Leadership and supervision, Metadata curation, Project administration, Samples and logistics, and Sequencing and analysis:

M. Estee Torok [22, 101].

Leadership and supervision, Metadata curation, Project administration, Samples and logistics, and Visualisation:

Kordo Saeed [81, 110] and Jacqui A Prieto [83, 109].

Leadership and supervision, Metadata curation, Project administration, Sequencing and analysis, and Software and analysis tools:

David K Jackson [99].

Metadata curation, Project administration, Samples and logistics, Sequencing and analysis, and Software and analysis tools:

William L Hamilton [22].

Metadata curation, Project administration, Samples and logistics, Sequencing and analysis, and Visualisation:

Luke B Snell [11].

Funding acquisition, Leadership and supervision, Metadata curation, and Samples and logistics:

Catherine Moore [69].

Funding acquisition, Leadership and supervision, Project administration, and Samples and logistics:

Ewan M Harrison [99, 88].

Leadership and supervision, Metadata curation, Project administration, and Samples and logistics:

Sonia Goncalves [99] and Leigh M Jackson [91].

Leadership and supervision, Metadata curation, Samples and logistics, and Sequencing and analysis:

Ian G Goodfellow [24], Derek J Fairley [3, 72], Matthew W Loose [18] and Joanne Watkins [69].

Leadership and supervision, Metadata curation, Samples and logistics, and Software and analysis tools:

Rich Livett [99].

Leadership and supervision, Metadata curation, Samples and logistics, and Visualisation:

Samuel Moses [25, 106].

Leadership and supervision, Metadata curation, Sequencing and analysis, and Software and analysis tools:

Roberto Amato [99], Sam Nicholls [41] and Matthew Bull [69].

Leadership and supervision, Project administration, Samples and logistics, and Sequencing and analysis:

Darren L Smith [37, 58, 105].

Leadership and supervision, Sequencing and analysis, Software and analysis tools, and Visualisation:

Jeff Barrett [99], David M Aanensen [14, 114].

Metadata curation, Project administration, Samples and logistics, and Sequencing and analysis:

Martin D Curran [65], Surendra Parmar [65], Dinesh Aggarwal [95, 99, 64] and James G Shepherd [48].

Metadata curation, Project administration, Sequencing and analysis, and Software and analysis tools:

Matthew D Parker [93].

Metadata curation, Samples and logistics, Sequencing and analysis, and Visualisation:

Sharon Glaysher [61].

Metadata curation, Sequencing and analysis, Software and analysis tools, and Visualisation:

Matthew Bashton [37, 58], Anthony P Underwood [14, 114], Nicole Pacchiarini [69] and Katie F Loveson [77].

Project administration, Sequencing and analysis, Software and analysis tools, and Visualisation:

Alessandro M Carabelli [88].

Funding acquisition, Leadership and supervision, and Metadata curation:

Kate E Templeton [53, 90].

Funding acquisition, Leadership and supervision, and Project administration:

Cordelia F Langford [99], John Sillitoe [99], Thushan I de Silva [93] and Dennis Wang [93].

Funding acquisition, Leadership and supervision, and Sequencing and analysis:

Dominic Kwiatkowski [99, 107], **And**rew Rambaut [90], Justin O'Grady [70, 89] and Simon Cottrell [69].

Leadership and supervision, Metadata curation, and Sequencing and analysis:

Matthew T.G. Holden [68] and Emma C Thomson [48].

Leadership and supervision, Project administration, and Samples and logistics:

Husam Osman [64, 36], Monique Andersson [59], Anoop J Chauhan [61] and Mohammed O Hassan-Ibrahim [6].

Leadership and supervision, Project administration, and Sequencing and analysis:

Mara Lawniczak [99].

Leadership and supervision, Samples and logistics, and Sequencing and analysis:

Ravi Kumar Gupta [88, 113], Alex Alderton [99], Meera Chand [66], Chrystala Constantinidou [94], Meera Unnikrishnan [94], Alistair C Darby [92], Julian A Hiscox [92] and Steve Paterson [92].

Leadership and supervision, Sequencing and analysis, and Software and analysis tools:

Inigo Martincorena [99], David L Robertson [48], Erik M Volz [39], Andrew J Page [70] and Oliver G Pybus [23].

Leadership and supervision, Sequencing and analysis, and Visualisation:

Andrew R Bassett [99].

Metadata curation, Project administration, and Samples and logistics:

Cristina V Ariani [99], Michael H Spencer Chapman [99, 88], Kathy K Li [48], Rajiv N Shah [48], Natasha G Jesudason [48] and Yusri Taha [50].

Metadata curation, Project administration, and Sequencing and analysis:

Martin P McHugh [53] and Rebecca Dewar [53].

Metadata curation, Samples and logistics, and Sequencing and analysis:

Aminu S Jahun [24], Claire McMurray [41], Sarojini Pandey [84], James P McKenna [3], Andrew Nelson [58, 105], Gregory R Young [37, 58], Clare M McCann [58, 105] and Scott Elliott [61].

Metadata curation, Samples and logistics, and Visualisation:

Hannah Lowe [25].

Metadata curation, Sequencing and analysis, and Software and analysis tools:

Ben Temperton [91], Sunando Roy [82], Anna Price [10], Sara Rey [69] and Matthew Wyles [93].

Metadata curation, Sequencing and analysis, and Visualisation:

Stefan Rooke [90] and Sharif Shaaban [68].

Project administration, Samples and logistics, Sequencing and analysis:

Mariateresa de Cesare [98].

Project administration, Samples and logistics, and Software and analysis tools:

Laura Letchford [99].

Project administration, Samples and logistics, and Visualisation:

Siona Silveira [81], Emanuela Pelosi [81] and Eleri Wilson-Davies [81].

Samples and logistics, Sequencing and analysis, and Software and analysis tools:

Myra Hosmillo [24].

Sequencing and analysis, Software and analysis tools, and Visualisation:

Áine O'Toole [90], Andrew R Hesketh [87], Richard Stark [94], Louis du Plessis [23], Chris Ruis [88], Helen Adams [4] and Yann Bourgeois [76].

Funding acquisition, and Leadership and supervision:

Stephen L Michell [91], Dimitris Grammatopoulos[84, 112], Jonathan Edgeworth [12], Judith Breuer [30, 82], John A Todd [98] and Christophe Fraser [5].

Funding acquisition, and Project administration:

David Buck [98] and Michaela John [9].

Leadership and supervision, and Metadata curation:

Gemma L Kay [70].

Leadership and supervision, and Project administration:

Steve Palmer [99], Sharon J Peacock [88, 64] and David Heyburn [69].

Leadership and supervision, and Samples and logistics:

Danni Weldon [99], Esther Robinson [64, 36], Alan McNally [41, 86], Peter Muir [64], Ian B Vipond [64], John BoYes [29], Venkat Sivaprakasam [46], Tranprit Salluja [75], Samir Dervisevic [54] and Emma J Meader [54].

Leadership and supervision, and Sequencing and analysis:

Naomi R Park [99], Karen Oliver [99], Aaron R Jeffries [91], Sascha Ott [94], Ana da Silva Filipe [48], David A Simpson [72] and Chris Williams [69].

Leadership and supervision, and Visualisation:

Jane AH Masoli [73, 91].

Metadata curation, and Samples and logistics:

Bridget A Knight [73, 91], Christopher R Jones [73, 91], Cherian Koshy [1], Amy Ash [1], Anna Casey [71], Andrew Bosworth [64, 36], Liz Ratcliffe [71], Li Xu-McCrae [36], Hannah M Pymont [64], Stephanie Hutchings [64], Lisa Berry [84], Katie Jones [84], Fenella Halstead [46], Thomas Davis [21], Christopher Holmes [16], Miren Iturriza-Gomara [92], Anita O Lucaci [92], Paul Anthony Randell [38, 104], Alison Cox [38, 104], Pinglawathee Madona [38, 104], Kathryn Ann Harris [30], Julianne Rose Brown [30], Tabitha W Mahungu [74], Dianne Irish-Tavares [74], Tanzina Haque [74], Jennifer Hart [74], Eric Witele [74], Melisa Louise Fenton [75], Steven Liggett [79], Clive Graham [56], Emma Swindells [57], Jennifer Collins [50], Gary Eltringham [50], Sharon Campbell [17], Patrick C McClure [97], Gemma Clark [15], Tim J Sloan [60], Carl Jones [15] and Jessica Lynch [2, 111].

Metadata curation, and Sequencing and analysis:

Ben Warne [8], Steven Leonard [99], Jillian Durham [99], Thomas Williams [90], Sam T Haldenby [92], Nathaniel Storey [30], Nabil-Fareed Alikhan [70], Nadine Holmes [18], Christopher Moore [18], Matthew Carlile [18], Malorie Perry [69], Noel Craine [69], Ronan A Lyons [80], Angela H Beckett [13], Salman Goudarzi [77], Christopher Fearn [77], Kate Cook [77], Hannah Dent [77] and Hannah Paul [77].

Metadata curation, and Software and analysis tools:

Robert Davies [99].

Project administration, and Samples and logistics:

Beth Blane [88], Sophia T Girgis [88], Mathew A Beale [99], Katherine L Bellis [99, 88], Matthew J Dorman [99], Eleanor Drury [99], Leanne Kane [99], Sally Kay [99], Samantha McGuigan [99], Rachel Nelson [99], Liam Prestwood [99], Shavanthi Rajatileka [99], Rahul Batra [12], Rachel J Williams [82], Mark Kristiansen [82], Angie Green [98], Anita Justice [59], Adhyana I.K Mahanama [81, 102] and Buddhini Samaraweera [81, 102].

Project administration, and Sequencing and analysis:

Nazreen F Hadjirin [88] and Joshua Quick [41].

Project administration, and Software and analysis tools:

Radoslaw Poplawski [41].

Samples and logistics, and Sequencing and analysis:

Leanne M Kermack [88], Nicola Reynolds [7], Grant Hall [24], Yasmin Chaudhry [24], Malte L Pinckert [24], Iliana Georgana [24], Robin J Moll [99], Alicia Thornton [66], Richard Myers [66], Joanne Stockton [41], Charlotte A Williams [82], Wen C Yew [58], Alexander J Trotter [70], Amy Trebes [98], George MacIntyre-Cockett [98], Alec Birchley [69], Alexander Adams [69], Amy Plimmer [69], Bree Gatica-Wilcox [69], Caoimhe McKerr [69], Ember Hilvers [69], Hannah Jones [69], Hibo Asad [69], Jason Coombes [69], Johnathan M Evans [69], Laia Fina [69], Lauren Gilbert [69], Lee Graham [69], Michelle Cronin [69], Sara Kumziene-SummerhaYes [69], Sarah Taylor [69], Sophie Jones [69], Danielle C Groves [93], Peijun Zhang [93], Marta Gallis [93] and Stavroula F Louka [93].

Samples and logistics, and Software and analysis tools:

Igor Starinskij [48].

Sequencing and analysis, and Software and analysis tools:

Chris J Illingworth [47], Chris Jackson [47], Marina Gourtovaia [99], Gerry Tonkin-Hill [99], Kevin Lewis [99], Jaime M Tovar-Corona [99], Keith James [99], Laura Baxter [94], Mohammad T. Alam [94], Richard J Orton [48], Joseph Hughes [48], Sreenu Vattipally [48], Manon Ragonnet-Cronin [39], Fabricia F. Nascimento [39], David Jorgensen [39], Olivia Boyd [39], Lily Geidelberg [39], Alex E Zarebski [23], Jayna Raghwani [23], Moritz UG Kraemer [23], Joel Southgate [10, 69], Benjamin B Lindsey [93] and Timothy M Freeman [93].

Software and analysis tools, and Visualisation:

Jon-Paul Keatley [99], Joshua B Singer [48], Leonardo de Oliveira Martins [70], Corin A Yeats [14], Khalil Abudahab [14, 114], Ben EW Taylor [14, 114] and Mirko Menegazzo [14].

Leadership and supervision:

John Danesh [99], Wendy Hogsden [46], Sahar Eldirdiri [21], Anita Kenyon [21], Jenifer Mason [43], Trevor I Robinson [43], Alison Holmes [38, 103], James Price [38, 103], John A Hartley [82], Tanya Curran [3], Alison E Mather [70], Giri Shankar [69], Rachel Jones [69], Robin Howe [69] and Sian Morgan [9].

Metadata curation:

Elizabeth Wastenge [53], Michael R Chapman [34, 88, 99], Siddharth Mookerjee [38, 103], Rachael Stanley [54], Wendy Smith [15], Timothy Peto [59], David Eyre [59], Derrick Crook [59], Gabrielle Vernet [33], Christine Kitchen [10], Huw Gulliver [10], Ian Merrick [10], Martyn Guest [10], Robert Munn [10], Declan T Bradley [63, 72] and Tim Wyatt [63].

Project administration:

Charlotte Beaver [99], Luke Foulser [99], Sophie Palmer [88], Carol M Churcher [88], Ellena Brooks [88], Kim S Smith [88], Katerina Galai [88], Georgina M McManus [88], Frances Bolt [38, 103], Francesc Coll [19], Lizzie Meadows [70], Stephen W Attwood [23], Alisha Davies [69], Elen De Lacy [69], Fatima Downing [69], Sue Edwards [69], Garry P Scarlett [76], Sarah Jeremiah [83] and Nikki Smith [93].

Samples and logistics:

Danielle Leek [88], Sushmita Sridhar [88, 99], Sally Forrest [88], Claire Cormie [88], Harmeet K Gill [88], Joana Dias [88], Ellen E Higginson [88], Mailis Maes [88], Jamie Young [88], Michelle Wantoch [7], Sanger Covid Team (www.sanger.ac.uk/covid-team) [99], Dorota Jamrozy [99], Stephanie Lo [99], Minal Patel [99], Verity Hill [90], Claire M Bewshea [91], Sian Ellard [73, 91], Cressida Auckland [73], Ian Harrison [66], Chloe Bishop [66], Vicki Chalker [66], Alex Richter [85], Andrew Beggs [85], Angus Best [86], Benita Percival [86], Jeremy Mirza [86], Oliver Megram [86], Megan Mayhew [86], Liam Crawford [86], Fiona Ashcroft [86], Emma Moles-Garcia [86], Nicola Cumley [86], Richard Hopes [64], Patawee

Asamaphan [48], Marc O Niebel [48], Rory N Gunson [100], Amanda Bradley [52], Alasdair Maclean [52], Guy Mollett [52], Rachel Blacow [52], Paul Bird [16], Thomas Helmer [16], Karlie Fallon [16], Julian Tang [16], Antony D Hale [49], Louissa R Macfarlane-Smith [49], Katherine L Harper [49], Holli Carden [49], Nicholas W Machin [45, 64], Kathryn A Jackson [92], Shazaad S Y Ahmad [45, 64], Ryan P George [45], Lance Turtle [92], Elaine O'Toole [43], Joanne Watts [43], Cassie Breen [43], Angela Cowell [43], Adela Alcolea-Medina [32, 96], Themoula Charalampous [12, 42], Amita Patel [11], Lisa J Levett [35], Judith Heaney [35], Aileen Rowan [39], Graham P Taylor [39], Divya Shah [30], Laura Atkinson [30], Jack CD Lee [30], Adam P Westhorpe [82], Riaz Jannoo [82], Helen L Lowe [82], Angeliki Karamani [82], Leah Ensell [82], Wendy Chatterton [35], Monika Pusok [35], Ashok Dadrah [75], Amanda Symmonds [75], Graciela Sluga [44], Zoltan Molnar [72], Paul Baker [79], Stephen Bonner [79], Sarah Essex [79], Edward Barton [56], Debra Padgett [56], Garren Scott [56], Jane Greenaway [57], Brendan Al Payne [50], Shirelle Burton-Fanning [50], Sheila Waugh [50], Veena Raviprakash [17], Nicola Sheriff [17], Victoria Blakey [17], Lesley-Anne Williams [17], Jonathan Moore [27], Susanne Stonehouse [27], Louise Smith [55], Rose K Davidson [89], Luke Bedford [26], Lindsay Coupland [54], Victoria Wright [18], Joseph G Chappell [97], Theocharis Tsoleridis [97], Jonathan Ball [97], Manjinder Khakh [15], Vicki M Fleming [15], Michelle M Lister [15], Hannah C Howson-Wells [15], Louise Berry [15], Tim Boswell [15], Amelia Joseph [15], Iona Willingham [15], Nichola Duckworth [60], Sarah Walsh [60], Emma Wise [2, 111], Nathan Moore [2, 111], Matilde Mori [2, 108, 111], Nick Cortes [2, 111], Stephen Kidd [2, 111], Rebecca Williams [33], Laura Gifford [69], Kelly Bicknell [61], Sarah Wyllie [61], Allyson Lloyd [61], Robert Impey [61], Cassandra S Malone [6], Benjamin J Cogger [6], Nick Levene [62], Lynn Monaghan [62], Alexander J Keeley [93], David G Partridge [78, 93], Mohammad Raza [78, 93], Cariad Evans [78, 93] and Kate Johnson [78, 93].

Sequencing and analysis:

Emma Betteridge [99], Ben W Farr [99], Scott Goodwin [99], Michael A Quail [99], Carol Scott [99], Lesley Shirley [99], Scott AJ Thurston [99], Diana Rajan [99], Iraad F Bronner [99], Louise Aigrain [99], Nicholas M Redshaw [99], Stefanie V Lensing [99], Shane McCarthy [99], Alex Makunin [99], Carlos E Balcazar [90], Michael D Gallagher [90], Kathleen A Williamson [90], Thomas D Stanton [90], Michelle L Michelsen [91], Joanna Warwick-Dugdale [91], Robin Manley [91], Audrey Farbos [91], James W Harrison [91], Christine M Sambles [91], David J Studholme [91], Angie Lackenby [66], Tamyo Mbisa [66], Steven Platt [66], Shahjahan Miah [66], David Bibby [66], Carmen Manso [66], Jonathan Hubb [66], Gavin Dabrera [66], Mary Ramsay [66], Daniel Bradshaw [66], Ulf Schaefer [66], Natalie Groves [66], Eileen Gallagher [66], David Lee [66], David Williams [66], Nicholas Ellaby [66], Hassan Hartman [66],

Nikos Manesis [66], Vineet Patel [66], Juan Ledesma [67], Katherine A Twohig [67], Elias Allara [64, 88], Clare Pearson [64, 88], Jeffrey K. J. Cheng [94], Hannah E. Bridgewater [94], Lucy R. Frost [94], Grace Taylor-Joyce [94], Paul E Brown [94], Lily Tong [48], Alice Broos [48], Daniel Mair [48], Jenna Nichols [48], Stephen N Carmichael [48], Katherine L Smollett [40], Kyriaki Nomikou [48], Elihu Aranday-Cortes [48], Natasha Johnson [48], Seema Nickbakhsh [48, 68], Edith E Vamos [92], Margaret Hughes [92], Lucille Rainbow [92], Richard Eccles [92], Charlotte Nelson [92], Mark Whitehead [92], Richard Gregory [92], Matthew Gemmell [92], Claudia Wierzbicki [92], Hermione J Webster [92], Chloe L Fisher [28], Adrian W Signell [20], Gilberto Betancor [20], Harry D Wilson [20], Gaia Nebbia [12], Flavia Flaviani [31], Alberto C Cerda [96], Tammy V Merrill [96], Rebekah E Wilson [96], Marius Cotic [82], Nadua Bayzid [82], Thomas Thompson [72], Erwan Acheson [72], Steven Rushton [51], Sarah O'Brien [51], David J Baker [70], Steven Rudder [70], Alp Aydin [70], Fei Sang [18], Johnny Debebe [18], Sarah Francois [23], Tetyana I Vasylyeva [23], Marina Escalera Zamudio [23], Bernardo Gutierrez [23], Angela Marchbank [10], Joshua Maksimovic [9], Karla Spellman [9], Kathryn McCluggage [9], Mari Morgan [69], Robert Beer [9], Safiah Afifi [9], Trudy Workman [10], William Fuller [10], Catherine Bresner [10], Adrienn Angyal [93], Luke R Green [93], Paul J Parsons [93], Rachel M Tucker [93], Rebecca Brown [93] and Max Whiteley [93].

Software and analysis tools:

James Bonfield [99], Christoph Puethe [99], Andrew Whitwham [99], Jennifier Liddle [99], Will Rowe [41], Igor Siveroni [39], Thanh Le-Viet [70] and Amy Gaskin [69].

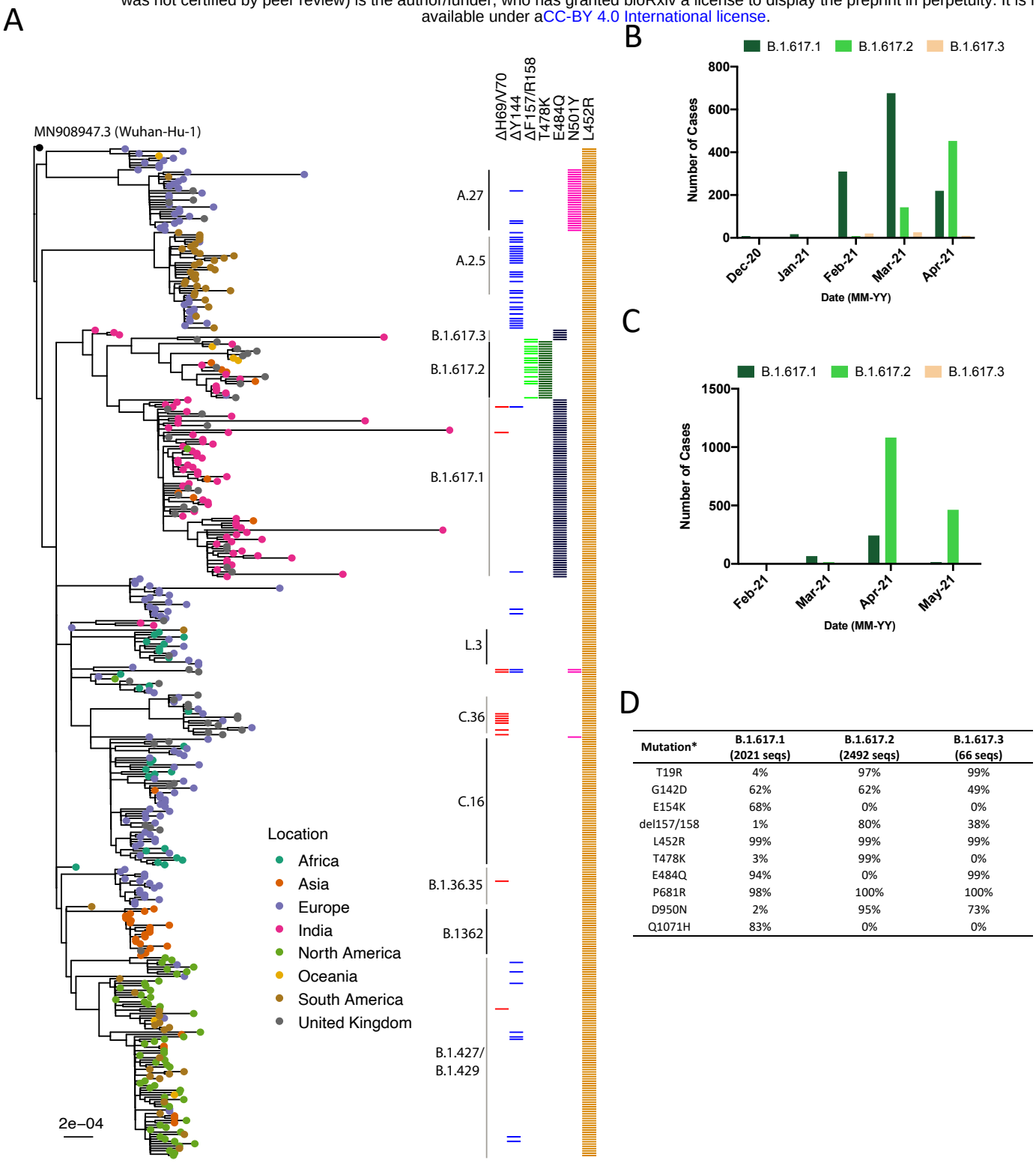
Visualisation:

Rob Johnson [39].

1 Barking, Havering and Redbridge University Hospitals NHS Trust, **2** Basingstoke Hospital, **3** Belfast Health & Social Care Trust, **4** Betsi Cadwaladr University Health Board, **5** Big Data Institute, Nuffield Department of Medicine, University of Oxford, **6** Brighton and Sussex University Hospitals NHS Trust, **7** Cambridge Stem Cell Institute, University of Cambridge, **8** Cambridge University Hospitals NHS Foundation Trust, **9** Cardiff and Vale University Health Board, **10** Cardiff University, **11** Centre for Clinical Infection & Diagnostics Research, St. Thomas' Hospital and Kings College London, **12** Centre for Clinical Infection and Diagnostics

Research, Department of Infectious Diseases, Guy's and St Thomas' NHS Foundation Trust, **13** Centre for Enzyme Innovation, University of Portsmouth (PORT), **14** Centre for Genomic Pathogen Surveillance, University of Oxford, **15** Clinical Microbiology Department, Queens Medical Centre, **16** Clinical Microbiology, University Hospitals of Leicester NHS Trust, **17** County Durham and Darlington NHS Foundation Trust, **18** Deep Seq, School of Life Sciences, Queens Medical Centre, University of Nottingham, **19** Department of Infection Biology, Faculty of Infectious & Tropical Diseases, London School of Hygiene & Tropical Medicine, **20** Department of Infectious Diseases, King's College London, **21** Department of Microbiology, Kettering General Hospital, **22** Departments of Infectious Diseases and Microbiology, Cambridge University Hospitals NHS Foundation Trust; Cambridge, UK, **23** Department of Zoology, University of Oxford, **24** Division of Virology, Department of Pathology, University of Cambridge, **25** East Kent Hospitals University NHS Foundation Trust, **26** East Suffolk and North Essex NHS Foundation Trust, **27** Gateshead Health NHS Foundation Trust, **28** Genomics Innovation Unit, Guy's and St. Thomas' NHS Foundation Trust, **29** Gloucestershire Hospitals NHS Foundation Trust, **30** Great Ormond Street Hospital for Children NHS Foundation Trust, **31** Guy's and St. Thomas' BRC, **32** Guy's and St. Thomas' Hospitals, **33** Hampshire Hospitals NHS Foundation Trust, **34** Health Data Research UK Cambridge, **35** Health Services Laboratories, **36** Heartlands Hospital, Birmingham, **37** Hub for Biotechnology in the Built Environment, Northumbria University, **38** Imperial College Hospitals NHS Trust, **39** Imperial College London, **40** Institute of Biodiversity, Animal Health & Comparative Medicine, **41** Institute of Microbiology and Infection, University of Birmingham, **42** King's College London, **43** Liverpool Clinical Laboratories, **44** Maidstone and Tunbridge Wells NHS Trust, **45** Manchester University NHS Foundation Trust, **46** Microbiology Department, Wye Valley NHS Trust, Hereford, **47** MRC Biostatistics Unit, University of Cambridge, **48** MRC-University of Glasgow Centre for Virus Research, **49** National Infection Service, PHE and Leeds Teaching Hospitals Trust, **50** Newcastle Hospitals NHS Foundation Trust, **51** Newcastle University, **52** NHS Greater Glasgow and Clyde, **53** NHS Lothian, **54** Norfolk and Norwich University Hospital, **55** Norfolk County Council, **56** North Cumbria Integrated Care NHS Foundation Trust, **57** North Tees and Hartlepool NHS Foundation Trust, **58** Northumbria University, **59** Oxford University Hospitals NHS Foundation Trust, **60** PathLinks, Northern Lincolnshire & Goole NHS Foundation Trust, **61** Portsmouth Hospitals University NHS Trust, **62** Princess Alexandra Hospital Microbiology Dept., **63** Public Health Agency, **64** Public

Health England, **65** Public Health England, Clinical Microbiology and Public Health Laboratory, Cambridge, UK, **66** Public Health England, Colindale, **67** Public Health England, Colindale, **68** Public Health Scotland, **69** Public Health Wales NHS Trust, **70** Quadram Institute Bioscience, **71** Queen Elizabeth Hospital, **72** Queen's University Belfast, **73** Royal Devon and Exeter NHS Foundation Trust, **74** Royal Free NHS Trust, **75** Sandwell and West Birmingham NHS Trust, **76** School of Biological Sciences, University of Portsmouth (PORT), **77** School of Pharmacy and Biomedical Sciences, University of Portsmouth (PORT), **78** Sheffield Teaching Hospitals, **79** South Tees Hospitals NHS Foundation Trust, **80** Swansea University, **81** University Hospitals Southampton NHS Foundation Trust, **82** University College London, **83** University Hospital Southampton NHS Foundation Trust, **84** University Hospitals Coventry and Warwickshire, **85** University of Birmingham, **86** University of Birmingham Turnkey Laboratory, **87** University of Brighton, **88** University of Cambridge, **89** University of East Anglia, **90** University of Edinburgh, **91** University of Exeter, **92** University of Liverpool, **93** University of Sheffield, **94** University of Warwick, **95** University of Cambridge, **96** Viapath, Guy's and St Thomas' NHS Foundation Trust, and King's College Hospital NHS Foundation Trust, **97** Virology, School of Life Sciences, Queens Medical Centre, University of Nottingham, **98** Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, **99** Wellcome Sanger Institute, **100** West of Scotland Specialist Virology Centre, NHS Greater Glasgow and Clyde, **101** Department of Medicine, University of Cambridge, **102** Ministry of Health, Sri Lanka, **103** NIHR Health Protection Research Unit in HCAI and AMR, Imperial College London, **104** North West London Pathology, **105** NU-OMICS, Northumbria University, **106** University of Kent, **107** University of Oxford, **108** University of Southampton, **109** University of Southampton School of Health Sciences, **110** University of Southampton School of Medicine, **111** University of Surrey, **112** Warwick Medical School and Institute of Precision Diagnostics, Pathology, UHCW NHS Trust, **113** Wellcome Africa Health Research Institute Durban and **114** Wellcome Genome Campus.



D

Mutation*	B.1.617.1 (2021 seqs)	B.1.617.2 (2492 seqs)	B.1.617.3 (66 seqs)
T19R	4%	97%	99%
G142D	62%	62%	49%
E154K	68%	0%	0%
del157/158	1%	80%	38%
L452R	99%	99%	99%
T478K	3%	99%	0%
E484Q	94%	0%	99%
P681R	98%	100%	100%
D950N	2%	95%	73%
Q1071H	83%	0%	0%

Figure 1. Context of SARS-CoV-2 B.1.617 variant emerging in India A. Maximum-likelihood phylogeny of lineages bearing L452R in spike. All sequences with the L452R mutation were downloaded from https://gisaid.org and manually aligned to reference strain MN908947.3 with mafft. Sequences were de-duplicated and a random subset of 400 global sequences, and 100 USA sequences were then selected with seqtk. All sequence lineages were assigned using pangolin v2.4. Major lineages are indicated as straight lines adjacent to the heatmap, alongside mutations of current interest. The phylogeny was inferred with IQTREE2 v2.1.3. **B, C, D.** The number of B.1617+ cases per month **B.** in India and **C.** the UK. **D.** Table of spike mutations in B.1.617+ lineages. All data collected on 13th May 2021.

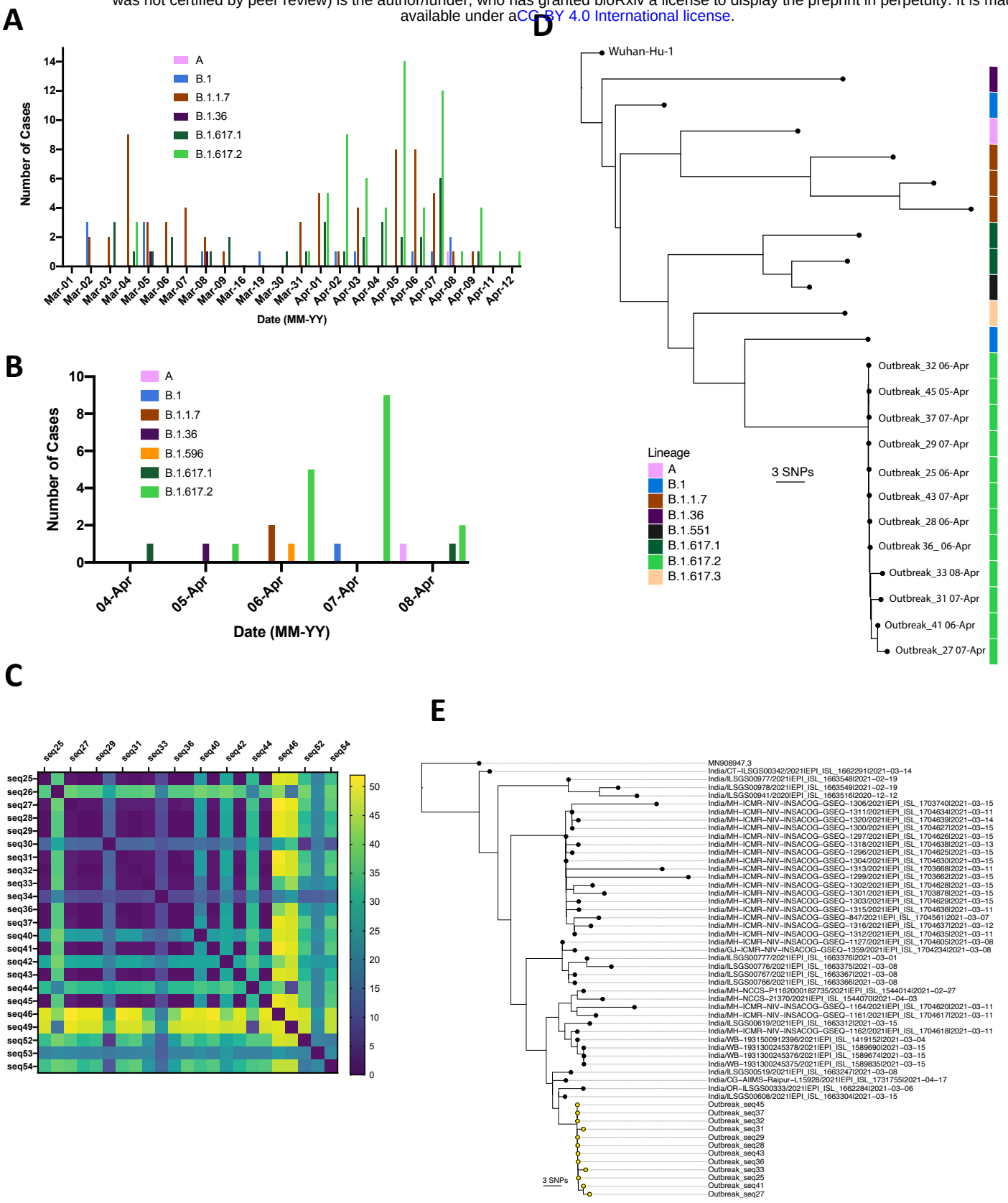


Figure 2. SARS-CoV-2 B.1.617.2 infection and transmission to fully vaccinated HCW in a health care centre located in Delhi, India. A. Case frequencies of SARS CoV-2 lineages over time for A. Delhi and **B**. fully vaccinated HCW at a single centre **C.** A heatmap of pairwise SARS-CoV-2 SNP differences of vaccinated HCW **samples.** The B.1.617.2 lineage is in the upper-left quarter, with fewer than 2 SNP difference between them. **D.** Maximum likelihood phylogeny of vaccine breakthrough SARS-CoV-2 sequences. Phylogeny was inferred with IQTREE2 with 1000 bootstrap replicates. Rooted on Wuhan-Hu-1 and annotated with the lineage designated by pangolin v.2.4.2. **E.** Maximum likelihood phylogeny of vaccine breakthrough SARS-CoV-2 B.1.617.2 sequences in context of closest Indian B.1.617.2 sequences.

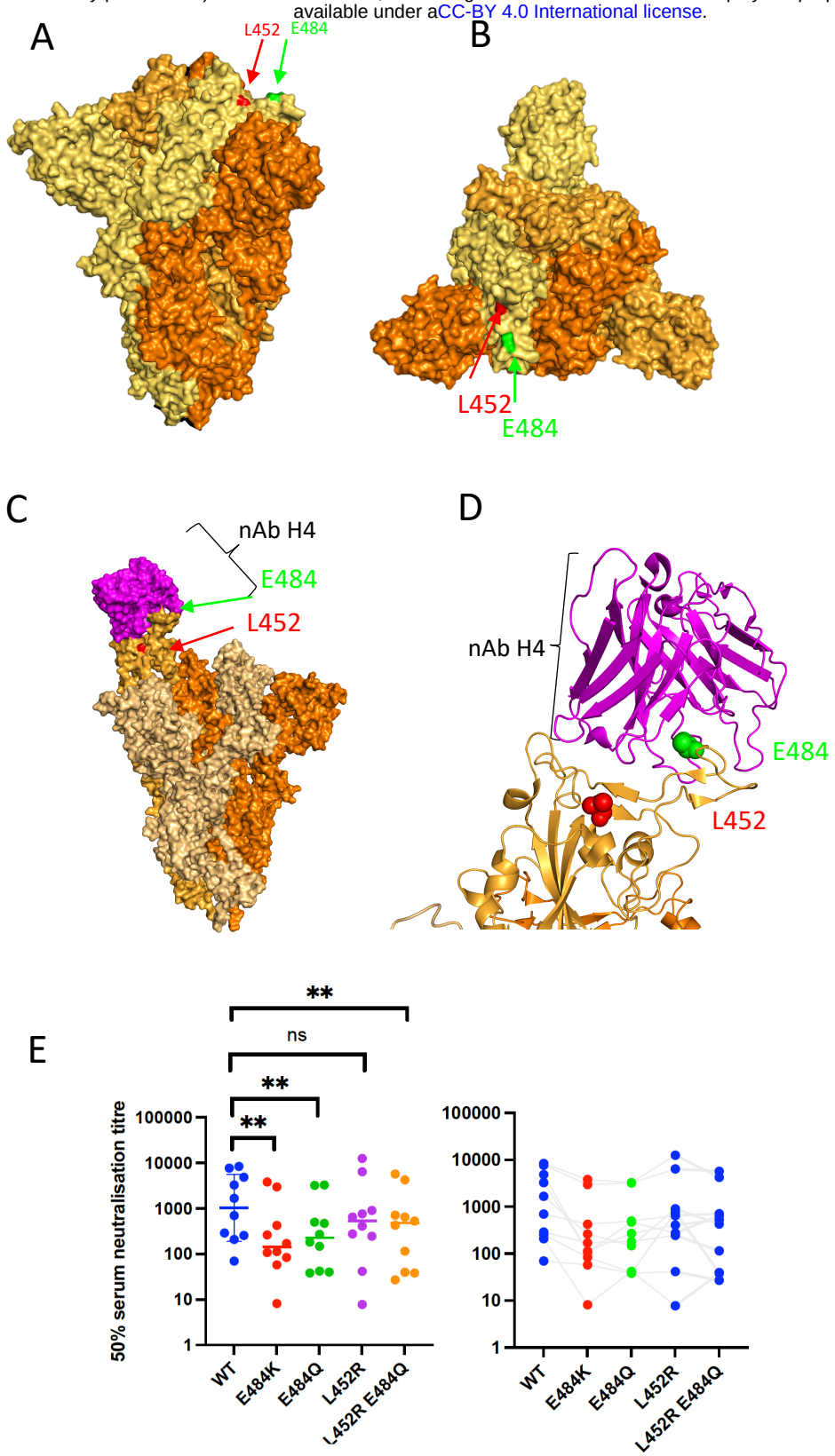


Figure 3. Presence and impact of mutations L452R and E484Q in the RDB of SARS-CoV-2 spike protein. A. Surface representation of the spike protein in closed formation (PDB: 6ZGE) in a vertical view with the location of L452 and E484 mutations highlighted as red and green sphere, respectively. Each monomer in the homotrimer is coloured accordingly **B.** Surface representation of the same spike protein in closed confirmation in a 'top-down' view along the trimer axis. The residues associated with RBD substitutions L452R, and E484Q are highlighted in red and green spheres respectively, on a single monomer. **C.** Ribbon representation of a single monomer of the same Spike, with residues L452 and E484 highlighted as spheres coloured by element. **D.** Surface representation of the spike protein in open formation with neutralising antibody H4 (pink spheres, PDB: 7L58, Rapp et al, 2021) bound to one monomer of the spike protein. Residues L452 and E484 are indicated with red and green sphere, respectively. Note that E484 is partially occluded by the bound monoclonal antibody **E.** Ribbon representation of the interaction between the neutralising antibody H4 and the RBD of a spike monomer. Residue E484 has direct interaction with the antibody, suggesting that mutations at this site may be involved in immune escape. **Neutralisation by mRNA vaccine-elicited sera E against wild type SARS-CoV-2 spike pseudotyped viruses bearing RBD mutations observed in B.1.617.1.** Serial dilutions of sera used and approx. 250,000 RLU of virus used for each mutant at each dilution. Geometric Mean Titre (GMT) shown with 95% CI. ** p<0.01.

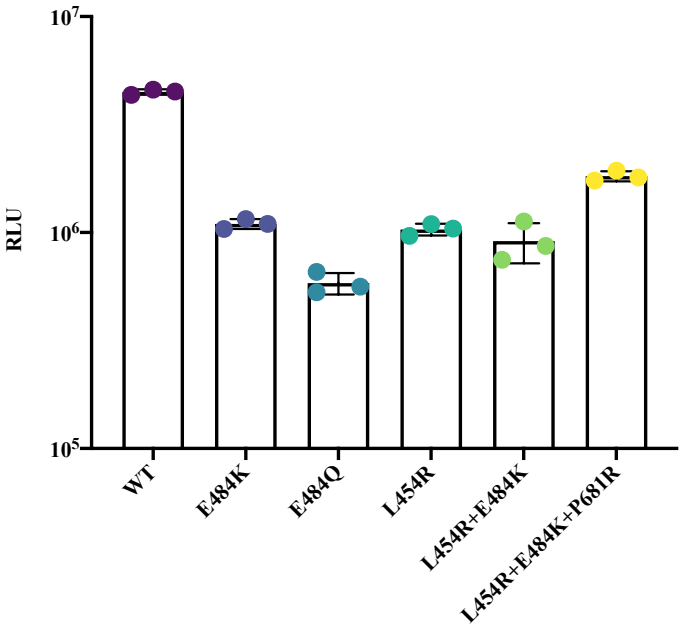


Figure 4: Entry efficiency of B.1.617 spike PV in 293T cells over expressing ACE2 TMPRSS2. PV were generated in 293T cells filtered and then used to infect target cells. Luciferase was measured 48 hours after infection. Input virus inoculum was corrected for genome copy input using SG-PERT. Mean is plotted with error bars representing SEM. Data are representative of two experiments.

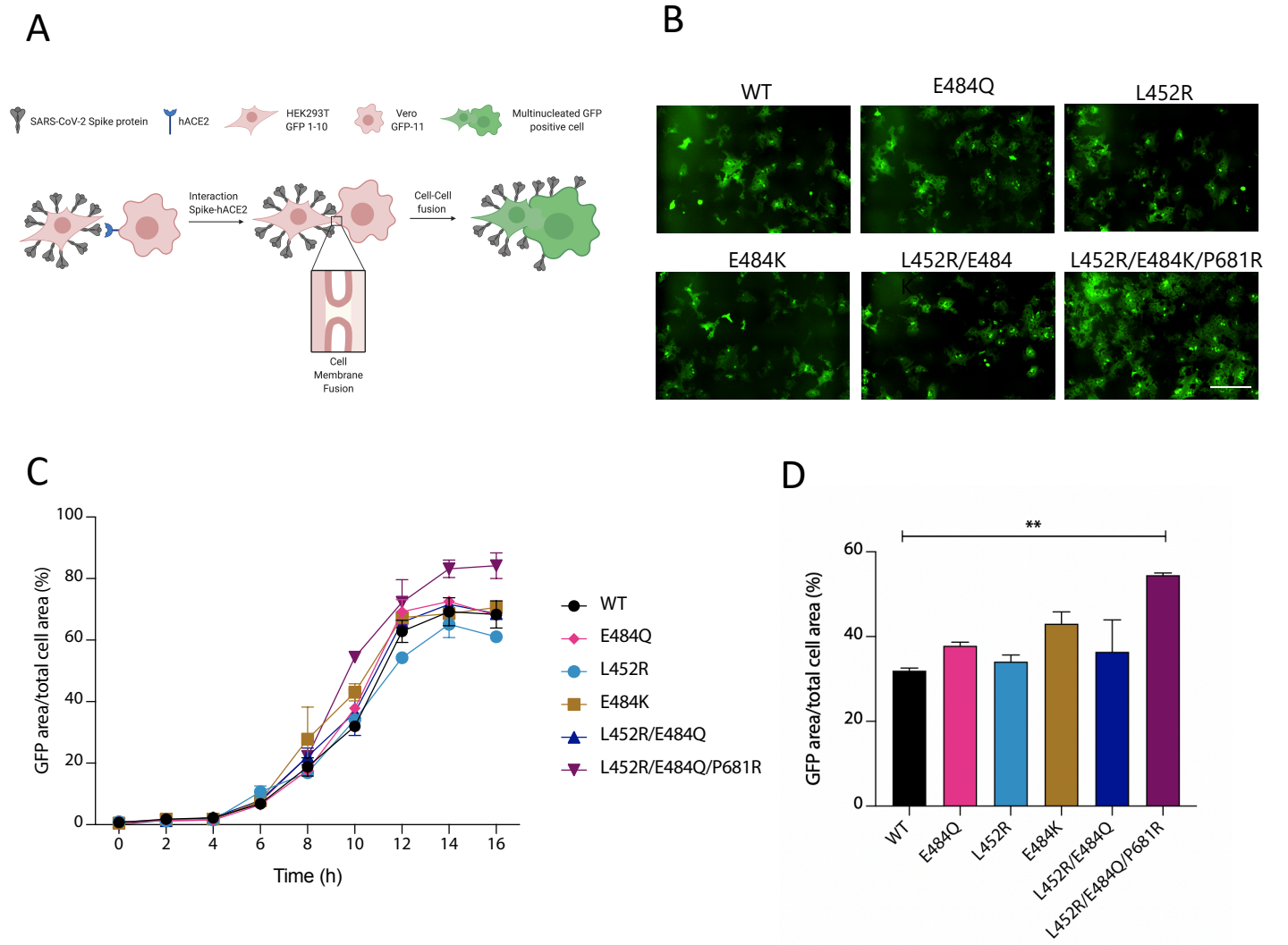
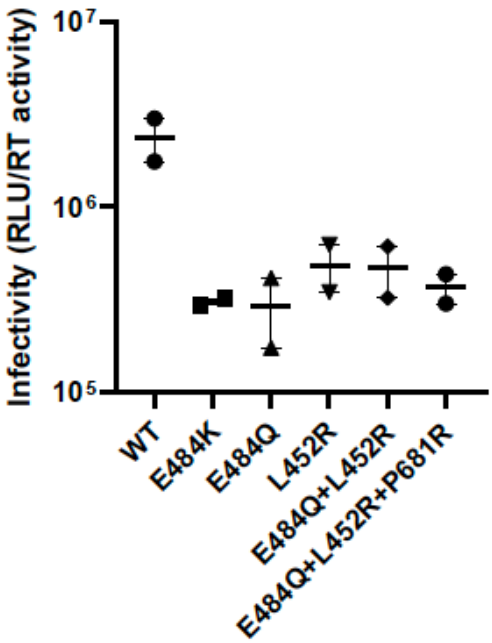
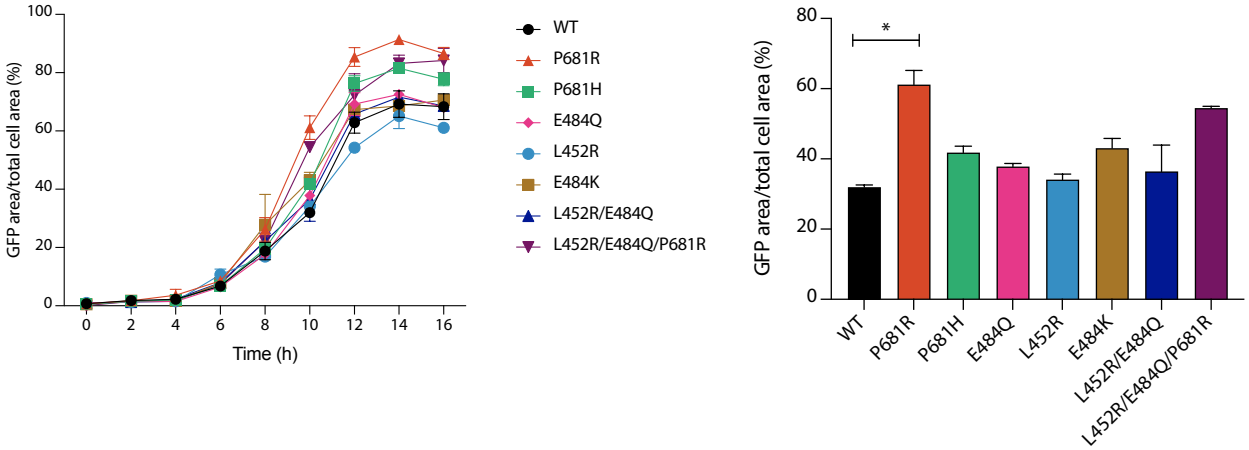


Figure 5: B.1.617 spike has accelerated cell-cell fusion activity dependent on furin cleavage at the polybasic cleavage site. A. Schematic of cell-cell fusion assay. B Reconstructed images at 10 hours of GFP positive syncytia formation. Scale bars represent 400 mm. C. Quantification of cell-cell fusion kinetics showing percentage of green area to total cell area over time. Mean is plotted with error bars representing SEM D. Quantification of cell-cell fusion of the indicated Spike mutants at 10 hours post transfection. D. Mean is plotted with error bars representing SEM. **p<0.005 Unpaired Student t test. E. Western blotting of donor cell lysates using an antibody against S2.



Supplementary figure 1: Entry efficiency of B.1.617 spike PV in CaLu cells. PV were generated in 293T cells filtered and then used to infect target cells. Luciferase was measured 48 hours after infection. Infectivity data were corrected for genome copy input using SG-PERT. Mean is plotted with error bars representing SEM.



Supplementary Figure 2: P681R accelerates cell-cell fusion activity. A. Quantification of cell-cell fusion kinetics showing percentage of green area to total cell area over time. Mean is plotted with error bars representing SEM **B.** Quantification of cell-cell fusion of the indicated Spike mutants at 10 hours post transfection. Mean is plotted with error bars representing SEM. **p<0.005 Unpaired Student t test