Check for updates

The furin cleavage site in the SARS-CoV-2 spike protein is required for transmission in ferrets

Thomas P. Peacock[1,5], Daniel H. Goldhill [1,5], Jie Zhou[1,5], Laury Baillon[1,5], Rebecca Frise[1,5], Olivia C. Swann[1], Ruthiran Kugathasan[1], Rebecca Penn[1], Jonathan C. Brown[1], Raul Y. Sanchez-David[1], Luca Braga[2], Maia Kavanagh Williamson[3], Jack A. Hassard[1], Ecco Staller[1], Brian Hanley[4], Michael Osborn[4], Mauro Giacca [2], Andrew D. Davidson[3], David A. Matthews [3] and Wendy S. Barclay [1✉]

SARS-CoV-2 entry requires sequential cleavage of the spike glycoprotein at the S1/S2 and the S2′ cleavage sites to mediate membrane fusion. SARS-CoV-2 has a polybasic insertion (PRRAR) at the S1/S2 cleavage site that can be cleaved by furin. Using lentiviral pseudotypes and a cell-culture-adapted SARS-CoV-2 virus with an S1/S2 deletion, we show that the polybasic insertion endows SARS-CoV-2 with a selective advantage in lung cells and primary human airway epithelial cells, but impairs replication in Vero E6, a cell line used for passaging SARS-CoV-2. Using engineered spike variants and live virus competition assays and by measuring growth kinetics, we find that the selective advantage in lung and primary human airway epithelial cells depends on the expression of the cell surface protease TMPRSS2, which enables endosome-independent virus entry by a route that avoids antiviral IFITM proteins. SARS-CoV-2 virus lacking the S1/S2 furin cleavage site was shed to lower titres from infected ferrets and was not transmitted to cohoused sentinel animals, unlike wild-type virus. Analysis of 100,000 SARS-CoV-2 sequences derived from patients and 24 human postmortem tissues showed low frequencies of naturally occurring mutants that harbour deletions at the polybasic site. Taken together, our findings reveal that the furin cleavage site is an important determinant of SARS-CoV-2 transmission.

I n 2019, SARS-CoV-2 entered the human population and by March 2020 was declared a pandemic by the World Health Organization[1–3]. Coronaviruses enter host cells via their spike glycoprotein which is synthesized as an inactive precursor that must be cleaved to mediate membrane fusion. Depending on the sequence of spike at the S1/S2 junction, the cleavage can occur: (1) during trafficking in the producer cell by host furin-like enzymes; (2) by serine proteases such as the transmembrane protease, serine 2 (TMPRSS2), at the cell surface during attachment; or (3) by cathepsin proteases in the late endosome/endolysosome[4,5]. Upon S1/S2 cleavage and engagement of the host cell receptor with the spike receptor binding domain, a second cleavage site (CS) becomes exposed within the S2 domain, termed the S2′ site[6–8]. Upon S2′ site cleavage by serine proteases or cathepsins, the S2 fusion peptide is liberated and initiates viral–host membrane fusion[7,9].

Similar to the closely related SARS-CoV, the cognate receptor of the SARS-CoV-2 spike is angiotensin-converting enzyme 2 (ACE2)[1,10]. While the SARS-CoV S1/S2 junction is well characterized as being cleaved by serine proteases or cathepsins, the SARS-CoV-2 spike, similarly to the more distantly related Middle Eastern respiratory syndrome-related coronavirus (MERS-CoV), contains a polybasic CS, characterized as being a suboptimal furin CS[6,11–13]. This polybasic CS is absent from the closest relatives of SARS-CoV-2, although similar polybasic CSs are found in more distantly related coronaviruses[14–16]. It has been demonstrated for both MERS-CoV spike and SARS-CoV-2 that the furin CS at the S1/S2 junction promotes entry into lung cells[17–19], and that the furin CS contributes to viral pathogenesis in SARS-CoV-2 animal models[20,21].

SARS-CoV-2 has been repeatedly shown to rapidly lose this polybasic CS upon passage in Vero cells, a popular cell line for isolating and propagating the virus[22–28]. In addition, there are isolated reports of CS mutants sequenced directly from clinical swabs[22,24]. Several different mutants in this region are described, including total deletions of the CS, loss of arginine substitutions within the CS making it less polybasic or deletions of flanking regions leaving the polybasic tract intact but potentially affecting accessibility to protease.

In this study, we use a combination of lentiviral pseudotypes (PVs) with spike CS mutations and Vero-passaged SARS-CoV-2 virus variants to investigate the molecular mechanism by which the polybasic CS of SARS-CoV-2 mediates efficient entry into lung cells. We describe the biological consequences of these mutations and test the effect of these mutations on viral transmission in ferrets.

Results

The polybasic S1/S2 CS of SARS-CoV-2 spike protein allows cleavage during virus packaging. To investigate the importance of the spike polybasic CS of SARS-CoV-2 (PRRAR), a number of spike mutants predicted to modulate the efficiency of furin cleavage were generated (Fig. 1a), including: substituting two upstream arginines to produce a monobasic CS similar to SARS-CoV spike (monoCS), replacing the tribasic CS with the furin CS of a highly pathogenic H5N1 avian influenza haemagglutinin containing seven basic amino acids (H5CS) and two naturally occurring deletions seen following passage in Vero E6 cells and/or in clinical isolates[21,26]. The first deletion removes eight amino acids including all three arginines of the PRRAR site (ΔCS), while the other removes five

[1]Department of Infectious Diseases, Imperial College London, London, UK. [2]British Hearth Foundation Centre of Research Excellence, School of Cardiovascular Medicine and Sciences, King's College London, London, UK. [3]School of Cellular and Molecular Medicine, Faculty of Life Sciences, University of Bristol, Bristol, UK. [4]Department of Cellular Pathology, Northwest London Pathology, Imperial College London NHS Trust, London, UK. [5]These authors contributed equally: Thomas P. Peacock, Daniel H. Goldhill, Jie Zhou, Laury Baillon, Rebecca Frise. ✉e-mail: w.barclay@imperial.ac.uk

899

flanking amino acids but retains the tribasic CS (Δflank). The mutations were engineered into a spike expression plasmid to enable cell surface expression and generation of coronavirus lentiviral PVs. In addition, to study the importance of the PRRAR motif in the context of live virus, we used a naturally occurring Vero-cell-adapted mutant SARS-CoV-2, ΔCS[26]. This variant and the wild-type (WT) virus from which it was derived were cloned by limiting dilution to enable studies using individual genotypes.

Furin cleavage of coronavirus spike proteins has been shown to correlate with syncytia formation when spike is overexpressed at neutral pH (refs. [7,18,29]). Therefore, we transiently expressed the SARS-CoV-2 mutant spike proteins in Vero E6 cells, which do not express TMPRSS2 (refs. [30,31]), and syncytia formation was compared with SARS-CoV and MERS-CoV spikes. As described before[18], SARS-CoV spike expression resulted in poor syncytia formation, while MERS-CoV spike produced appreciable levels of syncytia (Fig. 1b and Extended Data Fig. 1). SARS-CoV-2 WT spike gave an intermediate level of syncytia formation that was ablated in the monoCS or ΔCS/Δflank mutants. The H5CS spike bearing the optimized furin CS produced a higher level of syncytia formation than SARS-CoV-2 WT, similar to MERS-CoV.

To investigate the differences in spike cleavage efficiency in producer cells between the mutants, PVs with each mutant spike protein were concentrated and probed by western blot (Fig. 1c, left panel). Equal amounts of PV particles were loaded as indicated by p24 content. Anti-spike S2 antibody detected two bands in PVs, consistent with cleaved and uncleaved spike. For PVs expressing WT SARS-CoV-2 spike, the stronger band corresponded to the cleaved S2 product. H5CS spike was more efficiently cleaved while SARS-CoV WT spike and SARS-CoV-2 monoCS and deletion mutants were largely uncleaved. Consistent with PV, authentic SARS-CoV-2 virus harboured both uncleaved and cleaved S2, whereas ΔCS mutant virus only contained uncleaved spike (Fig. 1d). Overall, these data are consistent with previous work that has shown the polybasic CS of SARS-CoV-2 is a suboptimal furin CS[11,18,19].

The furin CS of SARS-CoV-2 spike protein promotes entry into epithelial cell lines and cultures but adversely affects entry into Vero and 293T cells. To investigate a role for the S1/S2 furin CS of SARS-CoV-2 in virus replication in different cell types, we performed competition assays, taking a mixed SARS-CoV-2 population containing 70% ΔCS mutant and 30% WT (as determined by

deep sequencing of the S1/S2 CS; Fig. 1e) inoculated onto Vero E6 cells; human intestinal Caco-2 cells; or air–liquid interface, differentiated human airway epithelial cell (HAE) cultures at a low multiplicity of infection (MOI) enabling multicycle replication. By 72 h, the ΔCS mutant outcompeted WT in Vero E6 cells, whereas WT became predominant in the Caco-2 cells. In primary HAE cultures, the WT rapidly outcompeted the ΔCS virus which was almost undetectable after 72 h (Fig. 1f). We also infected Calu-3 (human lung) cells with clonal WT or ΔCS virus at an MOI of 0.1 (Fig. 1g). WT virus replicated robustly and reached peak titres greater than 10^5 plaque-forming units (p.f.u.) after 48 h. Conversely, ΔCS virus was unable to productively infect Calu-3 cells and no infectious titre was detected at any time point.

Next, we probed the ability of PVs with different mutant spike proteins to enter different human cell lines: 293T cells expressing human ACE2, Caco-2 cells or Calu-3 cells (Fig. 2d–f). PVs bearing the envelope of amphotropic murine leukaemia virus (MLV-A) or Indiana vesicular stomatitis virus glycoprotein (VSV-G), or produced without any viral glycoproteins (bald), were used as positive and negative controls throughout. As in the Vero E6 cells (Fig. 1e), a clear negative correlation was seen between efficiency of furin cleavage of the spike and entry in 293T-ACE2 cells (Fig. 1h). PVs with WT SARS-CoV-2 spike entered 293T-ACE2s less efficiently than SARS-CoV, while SARS-CoV-2 spike mutants without furin cleavage (monoCS/ΔCS/Δflank) entered cells significantly more efficiently (>3-fold compared with WT). Introduction of the optimized furin CS (H5CS) dramatically decreased entry (~10-fold lower than WT; $P < 0.001$). In Caco-2 and Calu-3 cells, the opposite trend was observed in accordance with the efficiency of virus replication in Caco-2, Calu-3 and primary HAE cells (Fig. 1i,j). Mutants unable to be cleaved by furin, including ΔCS, entered cells significantly less efficiently than WT and H5CS (>2-fold lower in Caco-2 and ~5-fold lower in Calu-3 cells).

Entry of SARS-CoV-2 into 293T cells is dependent on cathepsins while entry into Caco-2, Calu-3 and primary HAE cells is dependent on TMPRSS2. As well as at the S1/S2 junction, coronavirus spike proteins require cleavage by host cell proteases at the S2′ site to enable viral–host cell membrane fusion. To investigate whether the different cell entry phenotypes seen in 293T-ACE2/Vero versus Caco-2/Calu-3/HAE cells were due to differences in protease usage, we performed PV entry assays in the presence of protease inhibitors: camostat, which inhibits serine proteases such as TMPRSS2,

Fig. 1 | The suboptimal furin CS of SARS-CoV-2 spike enhances entry into mucosal epithelial and primary human airway cells. a, Amino acid sequence alignment of coronavirus furin CS mutants used in this study. Mutants with potential S1/S2 furin CSs are shown in shades of orange while mutants without furin CSs are shown in shades of blue. **b**, Syncytia formation due to overexpression of different coronavirus spike proteins in Vero E6 cells. Percentage indicates the proportion of nuclei in each field that have formed clear syncytia. Data are plotted as mean + s.d. of three independent repeats. Statistical significance was determined by one-way ANOVA with multiple comparisons against SARS-CoV-2 WT. ****$P < 0.0001$. An extended figure of representative fields shown is in Extended Data Fig. 1. **c**, Western blot analysis of concentrated lentiviral PVs with different coronavirus spike proteins. Levels of lentiviral p24 antigen shown as loading control. Representative blot shown from $n = 3$ independent repeats. **d**, Western blot analysis of concentrated WT and ΔCS SARS-CoV-2 viruses. Levels of nucleocapsid (N) protein shown as loading control. Representative blot shown from $n = 2$ independent repeats. **e**, SARS-CoV-2 competition assay growth curve between WT and ΔCS virus in Vero E6 and Caco-2 cells. Cells were infected at an MOI of 0.1. Starting inoculum ratio is shown on the left-hand bar, while proportions of virus as determined by deep sequencing at 72 h post-inoculation are shown on the right. Virus titres determined by plaque assay at 72 h post-inoculation are shown in superimposed white data points. All results indicate triplicate repeats plotted as mean + s.d. **f**, SARS-CoV-2 competition assay growth curve between WT and ΔCS virus in HAEs. Cells were infected at an MOI of 0.1. Starting inoculum ratio shown at time 0; proportions of virus determined by deep sequencing. All time points were taken from triplicate repeats. Virus replication was determined by plaque assay and is shown as imposed white data points. Data are plotted as mean + s.d. **g**, Head-to-head replication kinetics of clonal WT and ΔCS viruses in Calu-3 human lung cells. Cells were infected at an MOI of 0.1. All time points taken from triplicate repeats are plotted as mean + s.d. Data shown are a representative replicate from (total $n = 2$) repeats. Virus replication was determined by plaque assay. Dotted line indicates limit of detection. **h–j**, Entry of lentiviral PVs containing different viral glycoproteins into 293T-ACE2 (**h**), Caco-2 (**i**) and Calu-3 (**j**) cells. Cells were transduced with different PVs and lysed 48 h later, and analysed by firefly luciferase luminescence. Data are shown as raw luminescence units (RLU). All assays were performed in sextuplicate (coronavirus pseudovirus) or triplicate (non-coronavirus controls) and are plotted as mean + s.d. Data shown are a representative replicate from ($n = 4$) independent repeats. Statistics were determined by one-way ANOVA on log-transformed data (after determining log normality by the Shapiro–Wilk test and QQ plot.) *$0.05 ≥ P > 0.01$; **$0.01 ≥ P > 0.001$; ***$0.001 ≥ P > 0.0001$; ****$P ≤ 0.0001$.

and E-64d, which inhibits cathepsins. Both drugs have been shown to be inhibitory to SARS-CoV and SARS-CoV-2 entry[32,33].

In 293T-ACE2 cells, camostat pretreatment did not inhibit PV entry, whereas E-64d did (Fig. 2a). In Caco-2 cells, a different pattern was seen: camostat had a significant impact on PVs bearing spike proteins with furin CSs, whereas E-64d inhibited only PVs with spikes that were not cleaved by furin (Fig. 2b). In Calu-3 cells, camostat significantly inhibited entry of all coronavirus PVs, while

E-64d also had a modest, but significant ($P < 0.05$), effect on the ΔCS mutant (Fig. 2c). Control PVs expressing MLV-A or VSV-G, which are not reliant on cathepsins or serine proteases for entry[33], were not significantly affected by either drug in any cell line.

To confirm dependence of whole SARS-CoV-2 virus on serine proteases in primary airway cells, we examined multicycle replication of clonal WT and ΔCS viruses in HAE cells in the presence or absence of camostat (Fig. 2d). Consistent with the results of the
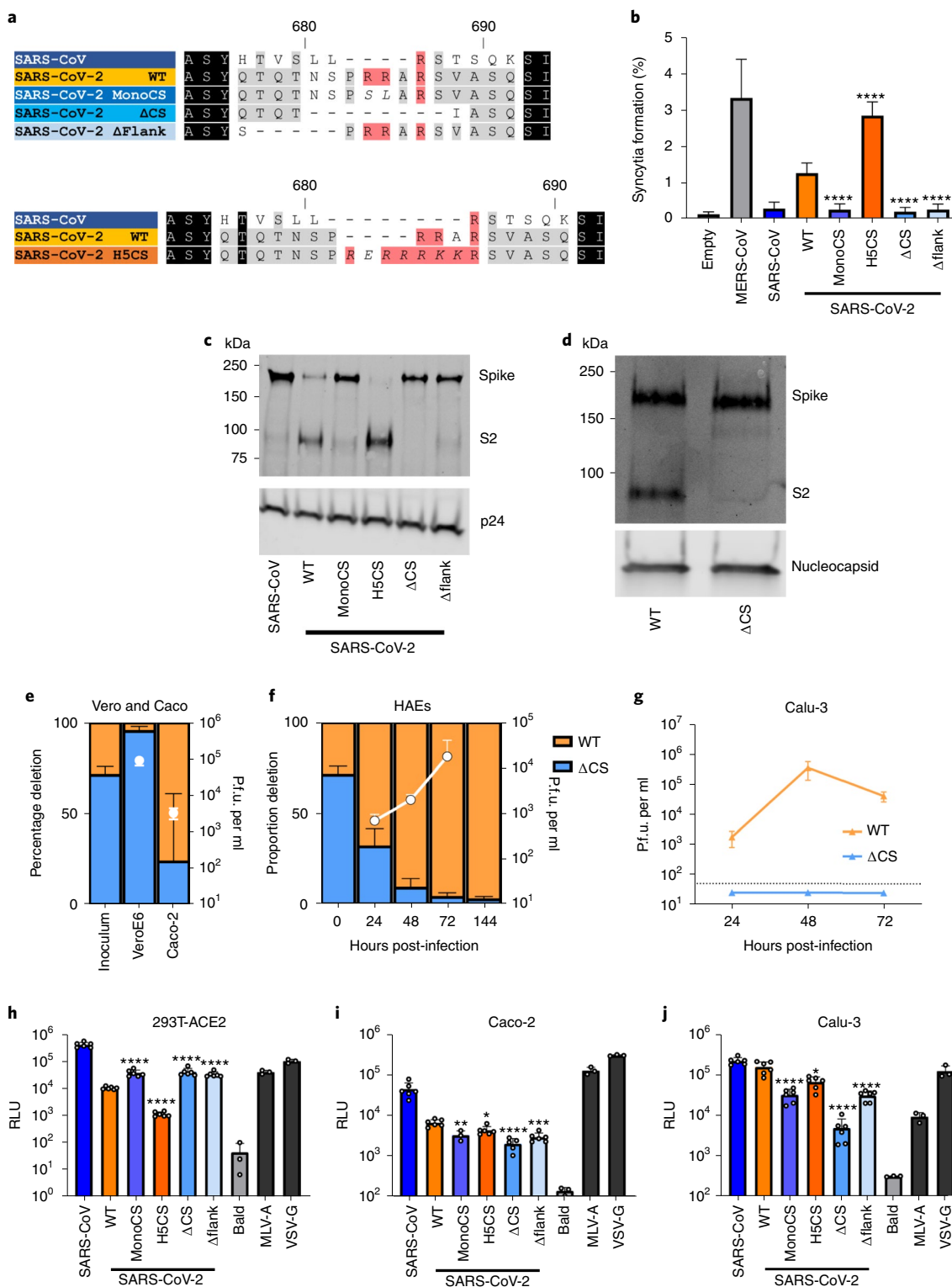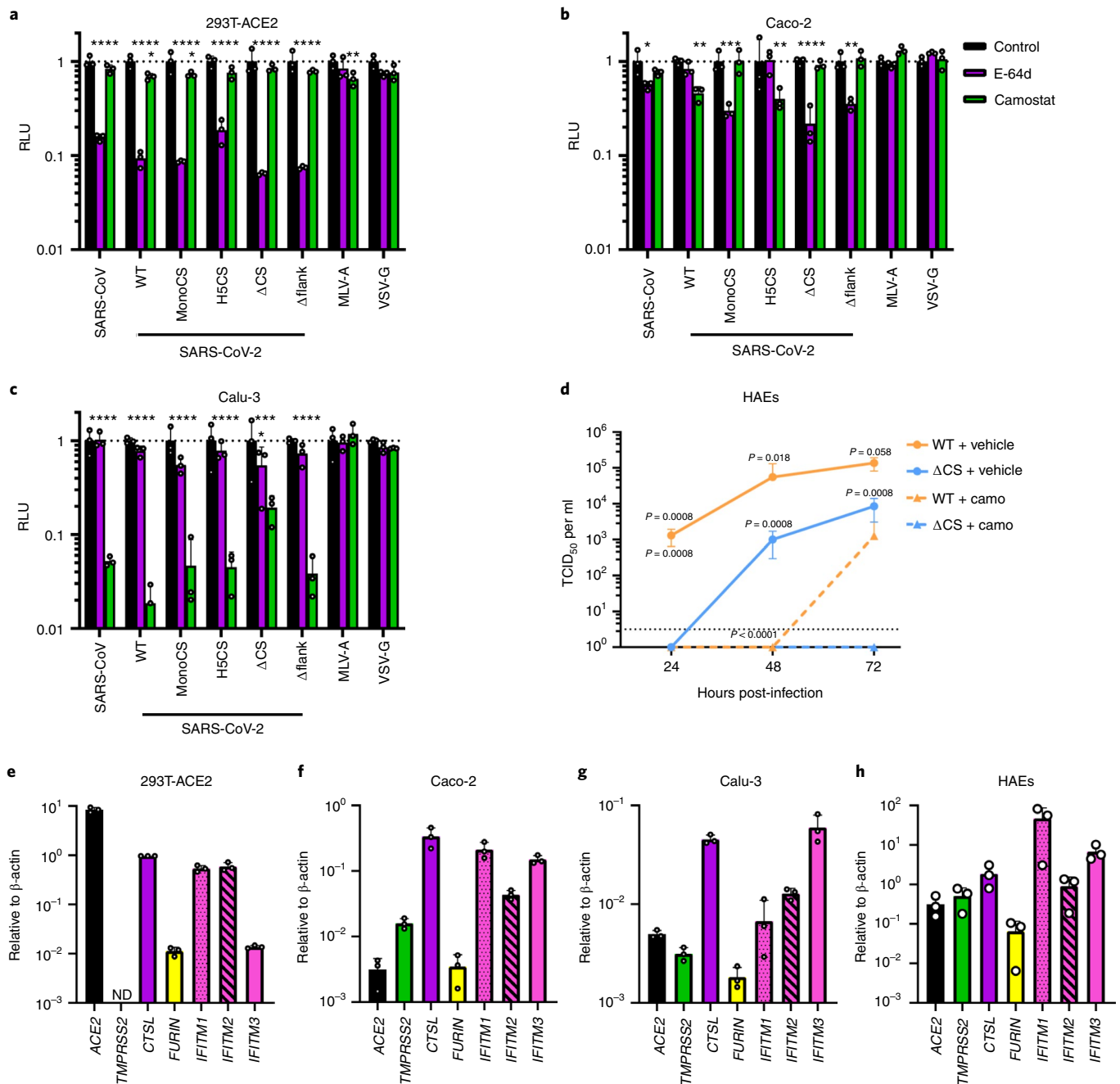
Fig. 2 | The furin CS of SARS-CoV-2 spike allows more efficient serine-protease-dependent entry into airway cells. a–c, Inhibition of entry of lentiviral PVs into 293T-ACE2 (**a**), Caco-2 (**b**) or Calu-3 (**c**) cells by the serine protease inhibitor, camostat (green bars), or the cathepsin inhibitor, E64-d (purple bars). Assays were performed in triplicate and are plotted as mean + s.d. Data shown are a representative replicate ($n = 3$). All data were normalized to no drug control (black bars). Statistics were determined by two-way ANOVA with multiple comparisons against the no drug control. $*0.05 \geq P > 0.01$; $**0.01 \geq P > 0.001$; $***0.001 \geq P > 0.0001$; $****P \leq 0.0001$. **d,** Replication kinetics of SARS-CoV-2 WT and ΔCS viruses in HAE cells. Cells were pretreated with control media or media containing camostat for 1h then infected at an MOI of 0.1. Assays were performed in triplicate and are plotted as mean + s.d. Statistics were determined by one-way ANOVA with multiple comparisons on log-transformed data. Black P values indicate statistical significance between no drug controls of WT and ΔCS, while coloured P values indicate significance between no drug control or camostat. Dotted line indicates limit of detection. **e–h,** Gene expression of selected SARS-CoV-2 entry factors in 293T-ACE2 (**e**), Caco-2 (**f**), Calu-3 (**g**) or HAEs (**h**). Gene expression was determined by RT–qPCR and normalized to β-actin. Assays were performed in triplicate and are plotted as mean + s.d. Data shown are a representative replicate ($n = 2$), except for primary HAEs, where data points represent repeats in ($n = 3$) independent donors. Camo, camostat; ND, not detectable.

competition assay, the ΔCS virus grew to significantly lower titres than WT. Addition of 50 μM camostat severely delayed replication of WT virus and abrogated that of ΔCS, without any loss of HAE integrity, as measured by transepithelial electronic resistance

(Fig. 3d and Extended Data Fig. 2a). Thus, in HAEs, cleavage by serine proteases is required for efficient entry.

To investigate whether differences in endogenous levels of receptor or proteases accounted for different entry pathways
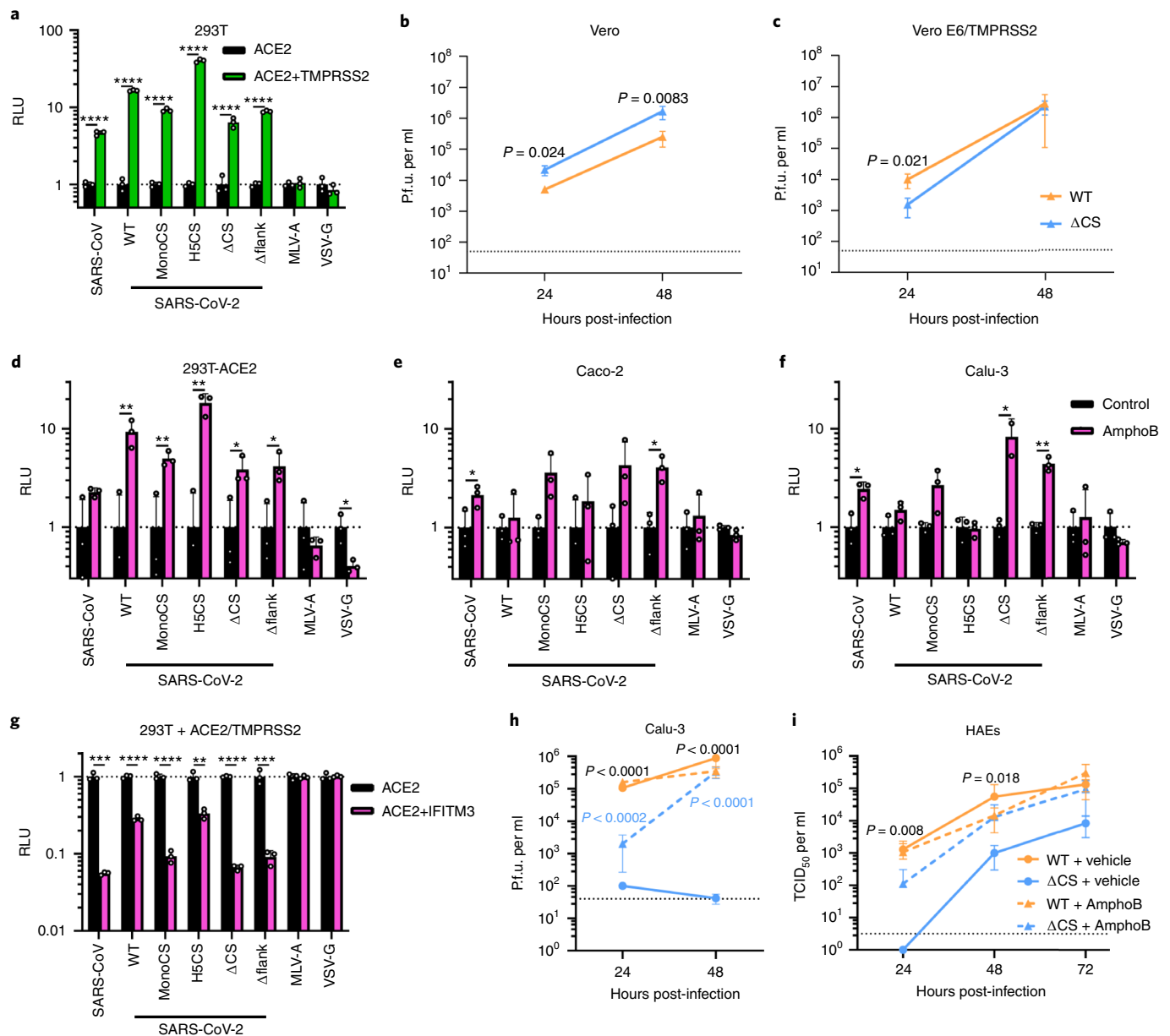
Fig. 3 | The efficient furin CS-dependent entry of SARS-CoV-2 is due to TMPRSS2 and allows for subsequent escape from IFITM3. a, Relative lentiviral PV entry into 293T cells expressing ACE2-FLAG with or without co-expression of TMPRSS2. Entry into cells not transfected with TMPRSS2 was normalized to 1. Assays were performed in triplicate and are plotted as mean + s.d. Data shown are a representative replicate ($n = 4$). Statistics were determined by multiple t-tests. ****$P \leq 0.0001$. **b,c**, Replication kinetics of SARS-CoV-2 WT and ΔCS viruses in Vero E6 (**b**) and Vero E6/TMPRSS2 (**c**) cells at an MOI of 0.1. Assays were performed in triplicate and are plotted as mean + s.d. Statistics were determined by one-way ANOVA with multiple comparisons on log-transformed data. **d–f**, Relative PV entry into 293T-ACE2 (**d**), Caco-2 (**e**) or Calu-3 (**f**) cells pretreated with amphoB (pink bars). Entry into untreated cells was normalized to 1 (black bars). Assays were performed in triplicate and are plotted as mean + s.d. Data shown are a representative replicate ($n = 4$). Statistics were determined by multiple two-tailed t-tests. *$0.05 \geq P > 0.01$; **$0.01 \geq P > 0.001$. **g**, Relative PV entry into 293T cells overexpressing ACE2-FLAG and TMPRSS2, with or without IFITM3. Entry into cells not transfected with IFITM3 was normalized to 1 (black bars). Assays were performed in triplicate and are plotted as mean + s.d. Data shown are a representative replicate ($n = 4$). Statistics were determined by multiple two-tailed t-tests. **$0.01 \geq P > 0.001$; ***$0.001 \geq P > 0.0001$; ****$P \leq 0.0001$. **h,i**, Replication kinetics of SARS-CoV-2 WT and ΔCS viruses in Calu-3 (**h**) or HAE (**i**) cells. Cells were pretreated with control media or media containing amphoB for 1 h then infected at an MOI of 0.05 (Calu-3) or 0.1 (HAE). Assays were performed in triplicate and are plotted as mean + s.d. The HAE assay was performed with three separate donors, with data from a representative donor shown. Statistics were determined by one-way ANOVA with multiple comparisons on log-transformed data. Black P values indicate statistical significance between no drug controls of WT and ΔCS, while coloured P values indicate significance between no drug control or amphoB. Vehicle controls were the same as from **d**. Dotted lines in **b**, **c**, **h** and **i** indicate limits of detection.

for SARS-CoV-2 in different human cell types, we quantified expression of ACE2, TMPRSS2 and cathepsin messenger RNAs (Fig. 2e–h). All three human cell lines and the primary HAE

cultures expressed ACE2 and cathepsin L to varying degrees. However, 293T-ACE2 cells lacked any detectable TMPRSS2 expression, explaining why camostat had little effect in these cells.

Previous studies have shown Vero E6 cells express no endogenous TMPRSS2 (refs. [30,31]).

Expression of TMPRSS2 promotes entry of SARS-CoV-2 with a polybasic CS. To investigate whether TMPRSS2 expression enhanced entry of viruses with furin-cleavable spike proteins, we compared the entry of PVs transiently in 293T cells co-expressing ACE2 with or without TMPRSS2 (Fig. 3a). TMPRSS2 promoted entry of all coronavirus PVs, even though TMPRSS2 expression led to lower levels of cell-associated ACE2 due to ACE2 being a substrate of TMPRSS2 (Extended Data Fig. 2b)[34]. The TMPRSS2-mediated enhancement of PV entry was particularly potent for the PVs harbouring furin CS-containing spike (>15-fold), compared with the non-furin-cleaved mutants (<10-fold), indicating that expression of TMPRSS2 favours the entry into cells of PVs with furin CSs.

We further tested the ability of authentic WT and ΔCS SARS-CoV-2 to replicate in Vero E6 cells compared with Vero E6 cells constitutively expressing TMPRSS2 (refs. [35,36]). In line with the previous competition assay (Fig. 1e), clonal ΔCS replicated significantly more rapidly than WT in Vero E6 cells (Fig. 3b). Conversely, in Vero E6/TMPRSS2 cells, WT virus grew to significantly higher titres by 24 h post-infection and both viruses reached similar peak titres at 48 h post-infection (Fig. 3c). Overall, these data confirm that TMPRSS2 expression gives WT SARS-CoV-2 a replication advantage over viruses lacking the furin CS.

The furin CS of SARS-CoV-2 allows escape from endosomal IFITM proteins in TMPRSS2-expressing cells. TMPRSS2 cleavage has been proposed to allow other coronaviruses to avoid restriction by endosomal IFITM proteins, such as IFITM2/3 (refs. [37,38]). Therefore, we hypothesized that the furin CS may enable SARS-CoV-2 also to avoid these IFITM proteins, known to restrict SARS-CoV and SARS-CoV-2 entry in the absence of TMPRSS2 (refs. [38–40]). The antifungal agent amphotericin B (amphoB) is well described as inhibiting the restriction imposed by endosomal/endolysosomal IFITM proteins, potentially through modulating the host membrane fluidity[38,39,41]. All the human cell lines used herein, 293T-ACE2, Caco-2, Calu-3 and HAEs, constitutively expressed all three antiviral IFITM paralogues, even in the absence of exogenous interferon (Fig. 2e–h). We pretreated cells with amphoB and investigated the effect on PV entry. In 293T-ACE2 cells, entry of all coronavirus PVs was improved by amphoB pretreatment, showing that all PVs entered these cells through endosomes (Fig. 3d). Conversely, in Caco-2 and Calu-3 cells, entry of PVs with uncleaved spikes was boosted by amphoB treatment, whereas there was little or no effect on the entry of PVs with furin CS-containing spikes (Fig. 3e,f).

Next, we co-expressed ACE2 and TMPRSS2 with or without IFITM3 in 293T cells. Entry of PVs with furin CS-containing spikes was less inhibited by IFITM3 than for those with spikes that could not be furin cleaved (Fig. 3e and Extended Data Fig. 2c).

Finally, we investigated the effect of amphoB treatment on SARS-CoV-2 replication in Calu-3 and HAEs. In both cell types amphoB had no effect on WT virus replication, but greatly increased the replication of the ΔCS mutant (Fig. 3h,i). This implies that

endosomal IFITM proteins are a major block for entry of viruses without furin CSs in TMPRSS2-expressing cells.

The SARS-CoV-2 polybasic CS promotes replication in the respiratory tract and transmission in a ferret model. To investigate whether the furin CS plays a role in the transmission of SARS-CoV-2, we used ferrets as an in vivo model. Ferrets are commonly used in transmission studies of respiratory pathogens such as influenza and, more recently, SARS-CoV-2 (refs. [42–44]). Furthermore, mink, which are closely related to ferrets, are highly susceptible to reverse zoonotic infection and outbreaks in mink farms show that SARS-CoV-2 is efficiently transmitted between these animals[45,46]. Four ferrets per group were each infected intranasally with 10^5 p.f.u. of clonal WT or ΔCS mutant SARS-CoV-2. After 24 h, naïve contact ferrets were cohoused with each donor. Ferrets were nasal-washed daily for the following 2 weeks and virus shedding was titrated by quantitative PCR with reverse transcription (RT–qPCR) and by median tissue culture infectious dose (TCID$_{50}$) (Fig. 4a,b and Extended Data Fig. 3a,b). All eight directly inoculated ferrets shed virus robustly for 9–12 d (Fig. 4a). The WT-infected group shed more virus than ferrets infected with ΔCS virus, indicated by higher infectious virus and E gene copy numbers, the latter significant at days 2–4. In the WT group, two of four contact ferrets became productively infected, indicated by infectious virus, E gene loads and seroconversion, whereas no transmission from donor ferret infected with ΔCS mutant virus was detected (Fig. 4b and Extended Data Fig. 3a–c). In nasal washes of the two remaining ferrets exposed to donors infected with WT virus, low E gene copy numbers were detected but no infectious virus was measured, and these animals remained seronegative at 14 d post exposure, implying these ferrets were not infected (Extended Data Fig. 3c).

Next, a competition assay was performed whereby four ferrets were inoculated intranasally with 10^5 p.f.u. of the previously described mixture of WT and ΔCS virus at a 30:70 ratio (Fig. 4c,d). At 1 d post-inoculation, naïve contact ferrets were cohoused with each donor animal and all animals were nasal-washed daily. All directly inoculated ferrets became productively infected, shedding infectious virus and detectable E gene between days 1 and 12 (Fig. 4c). Interestingly, which virus genotype became dominant in the nasal washes of the directly infected ferrets appeared to vary stochastically; in two animals the WT virus became predominant by day 2; these animals shed the highest levels of infectious virus and E gene RNA. In the remaining two directly inoculated animals, the ΔCS virus remained the majority species or outcompeted the WT over the course of the experiment. Productive transmission was only recorded in a single contact which was co-caged with one of the animals shedding predominantly WT virus (Fig. 4d). The ΔCS genotype was detectable in this single contact at low levels on days 3, 8 and 9, but at all times the WT virus clearly predominated. This animal was the only contact ferret to seroconvert, confirming the other three contact ferrets were not productively infected (Extended Data Fig. 4d). No ferrets from either experiment showed appreciable fever or weight loss (Extended Data Figs. 3d,e and 4a–c). Together, these results strongly suggest that the furin CS of SARS-CoV-2 spike is a determinant of transmission in the ferret model.

Fig. 4 | The furin CS of SARS-CoV-2 allows for efficient replication and transmission in a ferret model. a,b, Head-to-head transmission experiment of WT and ΔCS SARS-CoV-2 in ferrets. In each group, four individually housed donor ferrets were infected with 10^5 p.f.u. of either WT or ΔCS SARS-CoV-2. At 1 d post-inoculation, naïve contact ferrets were added to each donor ferret. Ferrets were sampled by nasal wash daily and donor (**a**) and direct contact (**b**) ferret virus titres were determined by E gene quantitative PCR. Statistics were determined by multiple two-tailed *t*-tests of the log-transformed E gene copy numbers between each group. *$P \leq 0.05$. **c,d,** Competition transmission experiment of SARS-CoV-2 mix of WT and ΔCS in ferrets. Four individually housed donor ferrets (**c**) were infected with 10^5 p.f.u. of virus mix containing ~70% ΔCS and ~30% WT. At 1 d post-inoculation, naïve contact ferrets (**d**) were added to each donor ferret. Ferrets were sampled by nasal wash daily and virus titres were determined by E gene quantitative PCR. For donors on days 2, 3, 5, 6 and 8, and for contacts on days 3, 6, 8 and 9, viral RNA across the S1/S2 CS was deep sequenced. Where sequencing data were obtained, bars showing the ratios of WT and deletion are shown. Inoc., inoculum. Dotted lines indicate limits of detection.

SARS-CoV-2 spike variants with deletions or mutations in the polybasic CS are detectable in human tissues. Finally, we investigated whether spike deletion mutants were present in human clinical samples and, if so, whether they were more likely to be found in a particular organ. Initially we downloaded 100,000 genome sequences from GISAID and found only two sequences from clinical swabs with CS deletions (Supplementary Table 1). Next, we deep sequenced the S1/S2 CS from 24 previously described
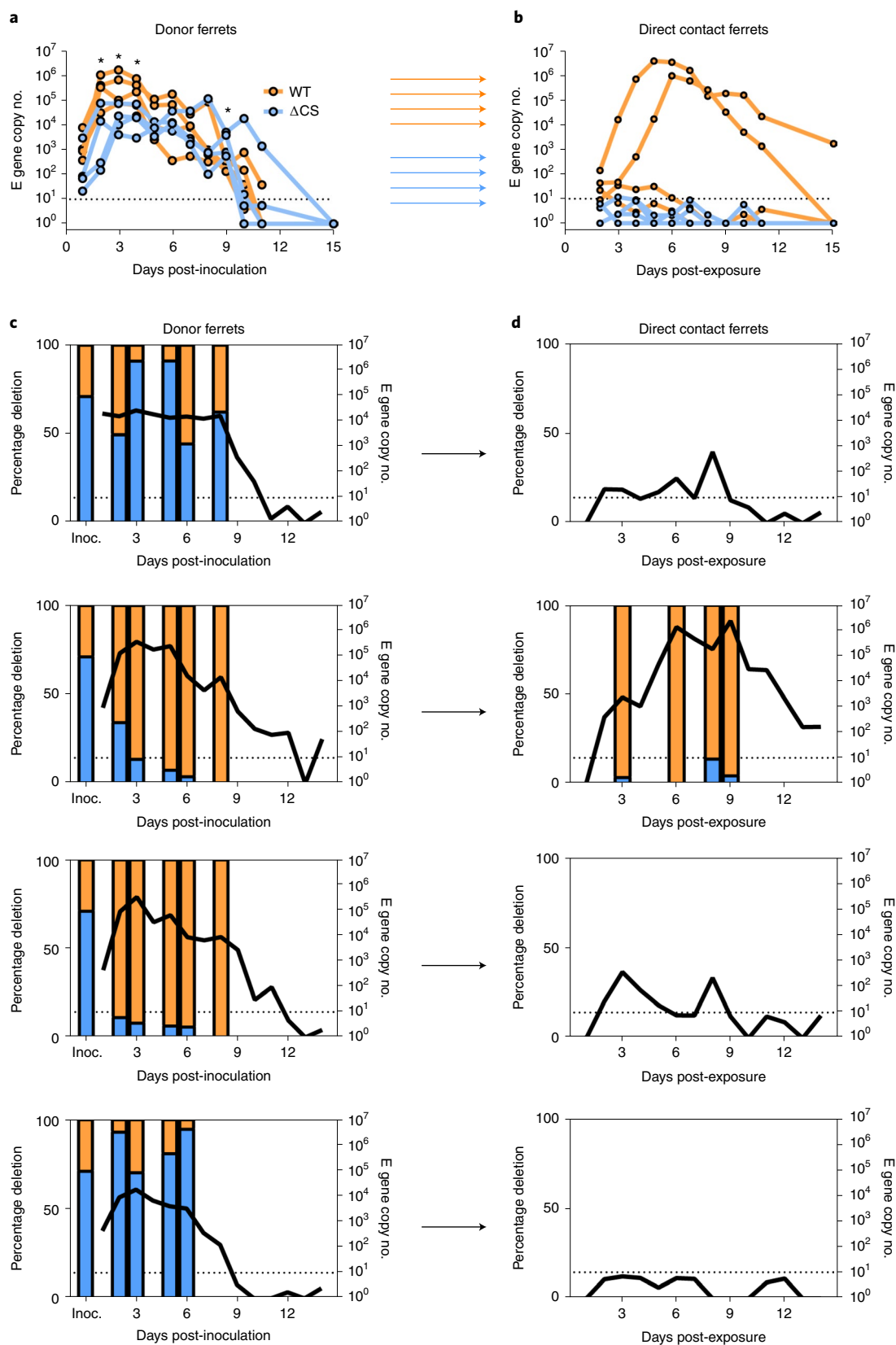
samples taken from five different postmortem cases, including tissues from the respiratory and gastrointestinal tract, the brain, heart, bone marrow, kidney, tongue and spleen[47]. Sequencing revealed very low levels of viral RNA bearing different S1/S2 CS deletions (<1%) from heart and spleen tissue from two separate patients (Supplementary Table 2). The three deletions reported in Supplementary Table 1 are similar to those seen upon passage in Vero E6 cells. OS5 deletes four amino acids after the CS, similar to a deletion reported in a recent study[27]; OS19-1 overlaps with most of Δflank and OS19-2 completely removes the S1/S2 site, similar to ΔCS. We have also observed identical deletions to OS19-2 upon passaging the clonal WT virus in Vero E6 cells. The S1/S2 CS of SARS-CoV-2 lies on an exposed, flexible loop, therefore any deletion, whether directly of the polybasic site or the flanking region, probably results in a reduced ability for furin to cleave this region[13,48], as demonstrated throughout this study by the Δflank and ΔCS mutants showing identical phenotypes. These results are consistent with the conclusion that S1/S2 CS deletions can arise in vivo, albeit at a very low rate.

Discussion

An insertion of four amino acids in the SARS-CoV-2 spike protein occurred during its emergence from an animal reservoir and created a suboptimal furin CS[11]. Here, we propose a mechanism by which this conferred an advantage to the virus in the human airway, enabling efficient human-to-human transmission. We confirm that precleavage of the spike during viral egress enhances entry of progeny virions into TMPRSS2-expressing cells such as those abundant in respiratory tissue[18,19]. TMPRSS2 cleaves spike at S2′ and facilitates early entry at or near the cell surface, as opposed to late entry through the endosome. This allows virus to avoid the potent endosomal/endolysosomal restriction factors, the IFITM proteins, which inhibit viral membrane fusion. Indeed, Winstone et al. recently showed that IFITM2 can potently restrict entry of SARS-CoV-2 variants that lack the spike polybasic site, and that this restriction accounts for the majority of type I interferon-induced inhibition of SARS-CoV-2 (ref. [49]). Viruses that lack a furin CS are forced to enter cells through the IFITM-containing endosome where the spike can be cleaved at S1/S2 and S2′ sites by cathepsins. However, furin-cleaved spike is not always advantageous: in cell types such as Vero, lacking TMPRSS2 expression, viruses without the furin CS gain an advantage, potentially because they are more stable, since spike cleavage may result in premature shedding of the S1 subunit altogether and abrogate receptor binding[50]. We show that, in contrast with WT SARS-CoV-2, a virus with a deleted furin CS did not replicate to high titres in the upper respiratory tract of ferrets and did not transmit to cohoused sentinel animals, in agreement with similar experiments using hamsters[28]. We have also found that furin CS deletions arise naturally at very low levels across different human organs during severe infection. Indeed, we note only two recorded genomes on GISAID out of 100,690 (as of 16 September 2020) with furin CS deletions (Supplementary Table 1). Given the ease of loss of the furin CS in cell culture, the lack of these mutants in sequenced isolates is further evidence that the furin CS is essential for sustained transmission of SARS-CoV-2 in humans.

Our study confirms TMPRSS2 as a potential drug target. Whilst inhibition of TMPRSS2 protease activity would not prevent infection via the endosome, using this pathway is detrimental to virus replication in airway cells. We have shown in this study that the protease inhibitor camostat is highly efficient at blocking SARS-CoV-2 replication in human airway cells and we note that clinical trials are ongoing (ClinicalTrials.gov Identifier: NCT04455815). Our study also confirms the limitations of relying on Vero E6 cells as a system for developing classes of drugs such as entry inhibitors as they do not accurately reflect the preferred entry mechanism of SARS-CoV-2 into human airway cells[51,52]. Indeed, the data here explain why

chloroquine is ineffective in clinic against SARS-CoV-2 (ref. [51]), since during replication in the human airway WT SARS-CoV-2 has evolved to enter cells without the need for endosomal acidification.

Presence of a furin CS at the S1/S2 junction is not uncommon in human coronaviruses; while half of human seasonal coronaviruses as well as MERS-CoV contain furin CSs, the remaining strains and SARS-CoV do not[6,16]. Thus, furin-mediated cleavage of spike is not an absolute requirement for efficient human respiratory transmission. Monitoring animal coronaviruses will probably be important in predicting and preventing future pandemics. We suggest that gain of a furin CS in the wider SARS-related coronaviruses is a cause for concern. The polybasic insertion to the S1/S2 CS provides a significant fitness advantage in TMPRSS2-expressing cells and is probably essential for efficient human transmission. We also note that the SARS-CoV-2 CS remains suboptimal for furin cleavage. It is unclear if this is a trade-off (that is, with stability of spike) or whether further optimization of this site could result in higher transmissibility. In this regard, multiple SARS-CoV-2 variants have recently emerged and spread rapidly, including some, such as the B.1.1.7 'UK' variant, that have mutations proximal to the S1/S2 CS predicted to enhance furin cleavage. This further emphasizes the role of this site for virus transmission and the importance of continued monitoring as SARS-CoV-2 circulates in the human population[53].

Methods

Biosafety and ethics statement. All work performed was approved by the local genetic manipulation safety committee of Imperial College London, St. Mary's Campus (centre number GM77), and the Health and Safety Executive of the United Kingdom, under reference CBA1.77.20.1. Animal research was carried out under a United Kingdom Home Office License, P48DAD9B4.

Human samples used in this research project were obtained from the Imperial College Healthcare Tissue Bank (ICHTB). ICHTB is supported by the National Institute for Health Research Biomedical Research Centre based at Imperial College Healthcare NHS Trust and Imperial College London. ICHTB is approved by Wales REC3 to release human material for research (17/WA/0161), and the samples for this project (R20012) were issued from subcollection reference number MED_MO_20_011.

Cells and viruses. African green monkey kidney cells (Vero E6; ATCC CRL-1586) and human embryonic kidney cells (293T; ATCC CRL-11268) were maintained in DMEM, 10% FCS, 1% non-essential amino acids (NEAA) and 1% penicillin-streptomycin (P/S). Human epithelial colorectal adenocarcinoma cells (Caco-2; ATCC HTB-37) and human lung cancer cells (Calu-3; ATCC HTB-55) were maintained in DMEM, 20% FCS, 1% NEAA and 1% P/S. Vero E6/TMPRSS2 cells were obtained from the Centre for AIDS Reagents (National Institute for Biological Standards and Control; 100978)[35,36], and maintained in DMEM, 10% FCS, 1% NEAA, 1% P/S and 1 mg ml⁻¹ Geneticin (G418). Air–liquid interface HAEs were purchased from Epithelix and maintained in Mucilair cell culture medium (Epithelix). All cell lines were maintained at 37 °C, 5% CO_2. HAE integrity in the presence of drugs was measured by transepithelial electrical resistance. Cell lines were not tested for mycoplasma contamination.

293T-hACE2 cells were generated by transducing 293T cells with an ACE2-expressing lentiviral vector, MT126, and selecting with 2 µg ml⁻¹ puromycin; after selection, cells were subsequently maintained with 1 µg ml⁻¹ puromycin.

The mixed SARS-CoV-2 WT/deletion virus mix was produced as previously described[26]. Briefly, the mix was generated by passaging the strain England/2/2020 (VE6-T), isolated by Public Health England, in Vero E6 cells whereby the deletion spontaneously arose[26]. The WT SARS-CoV-2 strain, SARS-CoV-2 strain England/2/2020 (VE6-T) and the ΔCS mutant present in the original mixed stock were purified by serially diluting the stock (tenfold dilutions) in MEM supplemented with 2% FCS and adding the dilutions to either Vero E6 or Caco-2 cells in a 96-well plate. After 5 d of incubation at 37 °C in 5% CO_2, the culture supernatants in wells showing cytopathic effect at the highest dilution were again diluted and passaged on the same cells. After a further 5 d of incubation, a 20-µl aliquot of culture supernatant from wells showing cytopathic effect at the highest dilution was used for RNA extraction and RT–PCR using a primer set designed to discriminate the WT and ΔCS mutant viruses. Culture supernatants containing either the WT or the ΔCS mutant virus, with no sign of a mixed virus population, were used to produce large-scale stocks in Vero E6 cells. The presence of the expected virus in the stocks was verified by direct RNA sequencing using an Oxford Nanopore flow cell as previously described[26]. Clonally pure viruses were then further amplified by one additional passage in Vero E6/TMPRSS2 cells to make the working stocks of the viruses used throughout this study.

For plaque assays, Vero E6 cells were used at 70–80% confluence. Cells were washed with PBS then serial dilutions of inoculum, diluted in serum-free DMEM, 1% NEAA and 1% P/S, were overlayed onto cells for 1 h at 37 °C. Inoculum was then removed and replaced with SARS-CoV-2 overlay medium (1× MEM, 0.2% w/v BSA, 0.16% w/v NaHCO$_3$, 10 mM HEPES, 2 mM L-Glutamine, 1× P/S, 0.6% w/v agarose). Plates were incubated for 3 d at 37 °C before overlay was removed and cells were stained for 1 h at room temperature in crystal violet solution.

To titrate virus by TCID$_{50}$, Vero E6 cells were used at 70–80% confluence. Serial dilutions of virus, diluted in serum-free DMEM, 1% NEAA and 1% P/S, were added to each well and cells were left for 5 d before they were fixed with 2× crystal violet solution and analysed. Four replicates of each sample were performed in tandem. TCID$_{50}$ titres were determined by the Spearman–Kärbar method[54].

Plasmids and cloning. Lentiviral packaging constructs pCSLW and pCAGGs-GAGPOL were made as previously described[55]. The codon-optimized spike proteins of SARS-CoV-2, SARS-CoV and MERS-CoV were a gift from P. McKay, Imperial College London[56]. Mutant SARS-CoV-2 expression plasmids were generated by site-directed mutagenesis. The lentiviral expression vector for human ACE2, MT126, was a gift from C. Goujon, University of Montpellier[57]. ACE was further cloned into pCAGGs with the addition of a C-terminal FLAG-tag. TMPRSS2 expression plasmid was a gift from R. Reeves (Addgene plasmid no. 53887; http://n2t.net/addgene:53887; RRID: Addgene_53887 (ref. [58])).

Syncytia formation assay. Vero E6 cells were seeded in 96-well plates (6.5×10^3 cells per well) to reach 70–80% confluency on the subsequent day. Transfection was performed using 100 ng of expression plasmid and 0.3 µl of FuGENE HD Transfection Reagent (Promega E2311) in 20 µl of Opti-MEM medium (Life Technologies). At 48 h after drug treatment, plates were washed in 100 µl per well of 1× PBS and fixed in 40 µl of 4% PFA for 10 min at room temperature. After fixation, cells were permeabilized in 0.1% Triton X-100 for 10 min at room temperature. Nuclei were stained using Hoechst 33342 (H3570 Thermo Fisher), according to the manufacturer's instructions.

Image acquisition was performed using the Operetta CLS high-content screening microscope (PerkinElmer) with a Zeiss ×20 (numerical aperture = 0.80) objective. A total of 25 fields per well were imaged for the Hoechst 33342 channel (excitation 365–385 nm, emission 430–500 nm). Images were subsequently analysed using Harmony software (PerkinElmer). Images were first flatfield-corrected and nuclei were segmented using the 'Find Nuclei' analysis module (Harmony). The thresholds for image segmentation were adjusted according to the signal-to-background ratio. The splitting coefficient was set to avoid splitting of overlapping nuclei (fused cells). All the cells that had a nuclear area greater than three times the average area of a single nucleus were considered as fused. Data were expressed as a percentage of fused cells by calculating the average number of fused cells normalized on the total number of cells per well.

Lentiviral PV assays. Lentiviral PVs were generated as previously described[55,59]. Briefly, 100 mm dishes of 293T cells were cotransfected with a mixture of 1 µg of the HIV packaging plasmid pCAGGs-GAGPOL, 1.5 µg of the luciferase reporter construct(pCSLW), and 1 µg of each envelope protein in pcDNA3.1. PV-containing supernatants were collected at 48 and 72 h post transfection, passed through a 0.45-µm filter, aliquoted and frozen at −80 °C. Concentrated PVs were produced by ultracentrifugation at 100,000g for 2 h over a 20% sucrose cushion.

Cells were transduced by PV for 48 h before lysis with cell culture lysis buffer (Promega). Luciferase luminescence was read on a FLUOstar Omega plate reader (BMF Labtech) using the Luciferase Assay System (Promega). The cathepsin inhibitor E-64d (Sigma-Aldrich), the serine protease inhibitor camostat mesylate (Abcam), or the antifungal and IFITM3 inhibitor amphoB (Sigma-Aldrich) was pre-applied to cells for 1 h at a concentration of 50 µM before addition of PV.

Overexpression experiments in 293T cells were performed by cotransfecting pCAGGS-ACE2-FLAG (1 µg) with TMPRSS2 (4 µg) or pCAGGs-IFITM3 (2.5 µg) into 100 mm dishes of 293T cells. Controls were transfected with equal amounts of empty vector instead of the named plasmid. After 24 h, cells were washed, resuspended into fresh media and added to PV, or spun down for analysis by western blot.

Deep sequencing using Primer ID. RNA was extracted from ferret nasal washes or cell supernatants using the QIAamp Viral RNA Mini Kit (Qiagen) with carrier RNA. RNA was reverse transcribed using Superscript IV (Invitrogen) and a barcoded primer for Primer ID (TGCGTTGATACCACTGCTTTNNNNANNNN ANNNNAACTGAATTTTCTGCACCAAG). Primer ID attaches a unique barcode to each complementary DNA molecule during reverse transcription and allows for PCR and sequencing error correction[60–62]. PCR was performed using KOD polymerase (Merck) and the following primers (CAACTTACTCCTACTTGGCGT and XXXXTGCGTTGATACCACTGCTTT), giving a 272-base pair (bp) amplicon. XXXX was a 4-base barcode (CACA, GTTG, AGGA or TCTC) to allow for additional multiplexing. Samples were pooled and prepared for sequencing using NebNext Ultra II (NEB), then sequenced on an Illumina MiSeq with 300-bp paired-end reads. Sequences were analysed in Geneious (v.11) and a pipeline in R. Forward and reverse reads were paired using FLASh (https://ccb.jhu.edu/software/ FLASH) before being mapped to a reference sequence and consensus sequences made for each barcode. A minimum cut-off of 3 reads per barcode was chosen. Raw sequences were deposited at www.ebi.ac.uk/ena, project number PRJEB40394. The analysis pipeline can be found at www.github.com/Flu1/Corona.

Deep sequencing from postmortem samples. RNA samples from human postmortem tissues from patients with SARS-CoV-2 where COVID-19 was listed clinically as the cause of death were sourced and processed as previously described[47]. Briefly, fresh tissue was processed within biosafety level 3 facilities and total RNA was extracted using TRIzol (Invitrogen)–chloroform extraction followed by precipitation and purification using an RNeasy Mini Kit (Qiagen). RNA was reverse transcribed using Superscript IV (Invitrogen) and the following primer (GTCTTGGTCATAGACACTGGTAG). PCR was performed using KOD polymerase (Merck) and the following primers (GTCTTGGTCATAGACACTGGTAG and GGCTGTTTAATAGGGGCTGAAC), giving a 260-bp amplicon. Samples were prepared for sequencing using NebNext Ultra II (NEB), then sequenced on an Illumina MiSeq with 300-bp paired-end reads. Sequences were analysed in Geneious (v.11) and a pipeline in R. Forward and reverse reads were paired using FLASh (https://ccb.jhu.edu/software/FLASH) before being mapped to a reference sequence. Raw sequences were deposited at www.ebi.ac.uk/ena, project number PRJEB40394. The analysis pipeline can be found at www.github.com/Flu1/Corona.

Human clinical samples. A total of 100,690 SARS-CoV-2 genomes were downloaded from GISAID on 16 September 2020 and aligned to 234 bp from the spike protein using Geneious. In-frame deletions were identified in R and analysed to ensure that samples had not been passaged before sequencing. Code for this analysis can be found at www.github.com/Flu1/Corona.

Ferret transmission studies. Ferret (*Mustela putorius furo*) transmission studies were performed in a containment level 3 laboratory, using a bespoke isolator system (Bell Isolation Systems). Outbred female ferrets (16–20 weeks old) weighing 750–1,000 g were used.

Before the study, ferrets were confirmed to be seronegative against SARS-CoV-2. Four donor ferrets were inoculated intranasally with 200 µl of 10^5 p.f.u. of virus mix while lightly anaesthetized with ketamine (22 mg kg^{-1}) and xylazine (0.9 mg kg^{-1}). To assess direct contact transmission, one naïve direct contact ferret was introduced into each cage at 1 d post initial inoculation.

All animals were nasal-washed daily, while conscious, by instilling 2 ml of PBS into the nostrils; the expectorate was collected into disposable 250-ml sample pots. Ferrets were weighed daily post infection, and body temperature was measured daily via subcutaneous IPTT-300 transponder (Plexx B.V).

Virus neutralization assay. The ability of ferret sera to neutralize WT SARS-CoV-2 virus was assessed by neutralization assay on Vero E6 cells. Heat-inactivated sera were serially diluted in assay diluent consisting of DMEM (Gibco, Thermo Fisher Scientific) with 1% P/S (Thermo Fisher Scientific) and 0.3% BSA fraction V (Thermo Fisher Scientific). Serum dilutions were incubated with 100 TCID$_{50}$ per well of virus in assay diluent for 1 h at room temperature and transferred to 96-well plates preseeded with Vero E6 cells. Serum dilutions were performed in duplicate. Plates were incubated at 37 °C, 5% CO$_2$ for 5 d before adding an equal volume of 2× crystal violet stain to wells for 1 h. Plates were washed, wells were scored for cytopathic effect and a neutralization titre was calculated as the reciprocal of the highest serum dilution at which full virus neutralization occurred.

Quantitative PCR. Viral RNA was extracted from ferret nasal washes using the Qiagen Viral RNA Mini Kit, according to the manufacturer's instructions.

RT–qPCR was performed using the 7500 Real Time PCR system (ABI) in 20-µl reactions using AgPath-ID One-Step RT–PCR Reagents: 10 µl of RT–PCR buffer (2×) (Thermo Fisher), 5 µl of RNA, 1 µl of forward (5′ ACAGGTACGTTAATAGTTAATAGCGT 3′) and reverse primers (5′ ATATTGCAGCAGTACGCACACA 3′), and 0.5 µl of probe (5′ FAM-ACACTAGCCATCCTTACTGCGCTTCG -BBQ 3′). The following conditions were used: 45 °C for 10 min, 1 cycle; 95 °C for 15 min, 1 cycle; 95 °C for 15 s then 58 °C for 30 s, 45 cycles. For each sample, the threshold cycle (C_t) value for the target E gene was determined. Based on the standard curves, absolute E gene copy numbers were calculated.

RNA was extracted from cells using RNA extraction kits (Qiagen, RNeasy Mini Kit, catalogue 74106) following the manufacturer's instructions. cDNA was synthesized in a reverse transcription step using Oligo-dT (RevertAid First Strand cDNA Synthesis, ThermoScientific, catalogue K1621). To quantify mRNA levels, quantitative PCR analysis with a gene-specific primer pair using SYBR green PCR mix (Applied Biosystems, catalogue 4385612) was performed and data were analysed on the Applied Biosystems ViiATM 7 Real-Time PCR System. Primers were used as described elsewhere[63–67], and can be found in Supplementary Table 3.

Western blotting. To investigate transfected protein expression, after 24 h transfected cells were lysed in RIPA buffer (150 mM NaCl, 1% NP-40, 0.5% sodium

deoxycholate, 0.1% SDS, 50 mM Tris, pH 7.4) supplemented with an EDTA-free protease inhibitor cocktail tablet (Roche). Cell lysates were then mixed with 4× Laemmli sample buffer (Bio-Rad) with 10% β-mercaptoethanol. Concentrated PVs as described above were also diluted in Laemmli buffer.

Membranes were probed with mouse anti-FLAG (diluted 1/2,000; F1804; Sigma), mouse anti-tubulin (diluted 1/5,000; abcam; ab7291), mouse anti-p24 (diluted 1/2,000; abcam; ab9071), rabbit anti-TMPRSS2 (diluted 1/2,000; abcam; ab92323), rabbit anti-Fragilis/IFITM3 (diluted 1/2,000; abcam; ab109429), rabbit anti-SARS spike protein (diluted 1/2,000; NOVUS; NB100-56578) or rabbit anti-SARS-CoV-2 nucleocapsid (diluted 1/3,000; SinoBiological; 40143-R019). Near infra-red secondary antibodies, IRDye 680RD Goat anti-mouse (diluted 1/10,000; abcam; ab216776), IRDye 680RD Goat anti-rabbit (diluted 1/10,000; abcam; ab216777), IRDye 800CW Goat anti-mouse (diluted 1/10,000; abcam; ab216772) or IRDye 800CW Goat anti-rabbit (diluted 1/10,000; abcam; ab216773)), were subsequently used. Western blots were visualized using an Odyssey Imaging System (LI-COR Biosciences).

Statistics and reproducibility. Statistics throughout this study were performed using one-way analysis of variance (ANOVA) or Student's *t*-test and are described in the figure legends. No statistical method was used to predetermine sample size. Two postmortem samples were excluded from the analysis as they had low number of reads similar to the negative control sample, and therefore were probably contamination. The experiments were not randomized, and the investigators were not blinded to allocation during experiments and outcome assessment.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
The data that support the findings of this study are available from the corresponding author upon reasonable request. Raw sequences were deposited at www.ebi.ac.uk/ena, project number PRJEB40394. The analysis pipeline can be found at www.github.com/Flu1/Corona. Source data are provided with this paper.

Code availability
All code is available at www.github.com/Flu1/corona.

References

1. Zhou, P. et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
2. Zhu, N. et al. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).
3. World Health Organisation. *WHO Coronavirus Disease (COVID-19) Dashboard* https://covid19.who.int/ (2021).
4. Matsuyama, S. et al. Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2. *J. Virol.* **84**, 12658–12664 (2010).
5. Simmons, G. et al. Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. *Proc. Natl Acad. Sci. USA* **102**, 11876–11881 (2005).
6. Millet, J. K. & Whittaker, G. R. Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. *Proc. Natl Acad. Sci. USA* **111**, 15214–15219 (2014).
7. Belouzard, S., Chu, V. C. & Whittaker, G. R. Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. *Proc. Natl Acad. Sci. USA* **106**, 5871–5876 (2009).
8. Benton, D. J. et al. Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion. *Nature* https://doi.org/10.1038/s41586-020-2772-0 (2020).
9. Madu, I. G., Roth, S. L., Belouzard, S. & Whittaker, G. R. Characterization of a highly conserved domain within the severe acute respiratory syndrome coronavirus spike protein S2 domain with characteristics of a viral fusion peptide. *J. Virol.* **83**, 7411–7421 (2009).
10. Li, W. et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450–454 (2003).
11. Shang, J. et al. Cell entry mechanisms of SARS-CoV-2. *Proc. Natl Acad. Sci. USA* https://doi.org/10.1073/pnas.2003138117 (2020).
12. Coutard, B. et al. The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. *Antiviral Res.* **176**, 104742 (2020).
13. Jaimes, J. A., Millet, J. K. & Whittaker, G. R. Proteolytic cleavage of the SARS-CoV-2 spike protein and the role of the novel S1/S2 site. *iScience* **23**, 101212 (2020).
14. Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C. & Garry, R. F. The proximal origin of SARS-CoV-2. *Nat. Med.* **26**, 450–452 (2020).
15. Boni, M. F. et al. Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. *Nat. Microbiol.* https://doi.org/10.1038/s41564-020-0771-4 (2020).
16. Le Coupanec, A. et al. Cleavage of a neuroinvasive human respiratory virus spike glycoprotein by proprotein convertases modulates neurovirulence and virus spread within the central nervous system. *PLoS Pathog.* **11**, e1005261 (2015).
17. Park, J. E. et al. Proteolytic processing of Middle East respiratory syndrome coronavirus spikes expands virus tropism. *Proc. Natl Acad. Sci. USA* **113**, 12262–12267 (2016).
18. Hoffmann, M., Kleine-Weber, H. & Pohlmann, S. A multibasic cleavage site in the spike protein of SARS-CoV-2 is essential for infection of human lung cells. *Mol. Cell* https://doi.org/10.1016/j.molcel.2020.04.022 (2020).
19. Bestle, D. et al. TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. *Life Sci. Alliance* https://doi.org/10.26508/lsa.202000786 (2020).
20. Johnson, B. A. et al. Loss of furin cleavage site attenuates SARS-CoV-2 pathogenesis. *Nature* https://doi.org/10.1038/s41586-021-03237-4 (2021).
21. Lau, S. Y. et al. Attenuated SARS-CoV-2 variants with deletions at the S1/S2 junction. *Emerg. Microbes Infect.* **9**, 837–842 (2020).
22. Liu, Z. et al. Identification of common deletions in the spike protein of SARS-CoV-2. *J. Virol.* https://doi.org/10.1128/JVI.00790-20 (2020).
23. Ogando, N. S. et al. SARS-coronavirus-2 replication in Vero E6 cells: replication kinetics, rapid adaptation and cytopathology. *J. Gen. Virol.* https://doi.org/10.1099/jgv.0.001453 (2020).
24. Wong, Y. C. et al. Natural transmission of bat-like SARS-CoV-2PRRA variants in COVID-19 patients. *Clin. Infect Dis.* https://doi.org/10.1093/cid/ciaa953 (2020).
25. Klimstra, W. B. et al. SARS-CoV-2 growth, furin-cleavage-site adaptation and neutralization using serum from acutely infected hospitalized COVID-19 patients. *J. Gen. Virol.* https://doi.org/10.1099/jgv.0.001481 (2020).
26. Davidson, A. D. et al. Characterisation of the transcriptome and proteome of SARS-CoV-2 reveals a cell passage induced in-frame deletion of the furin-like cleavage site from the spike glycoprotein. *Genome Med.* **12**, 68 (2020).
27. Sasaki, M. et al. SARS-CoV-2 variants with mutations at the S1/S2 cleavage site are generated in vitro during propagation in TMPRSS2-deficient cells. *PLoS Pathog.* **17**, e1009233 (2021).
28. Zhu, Y. et al. A genome-wide CRISPR screen identifies host factors that regulate SARS-CoV-2 entry. *Nat. Commun.* **12**, 961 (2021).
29. Xia, S. et al. The role of furin cleavage site in SARS-CoV-2 spike protein-mediated membrane fusion in the presence or absence of trypsin. *Signal Transduct. Target. Ther.* **5**, 92 (2020).
30. Shirato, K., Kawase, M. & Matsuyama, S. Middle East respiratory syndrome coronavirus infection mediated by the transmembrane serine protease TMPRSS2. *J. Virol.* **87**, 12552–12561 (2013).
31. Bertram, S. et al. TMPRSS2 and TMPRSS4 facilitate trypsin-independent spread of influenza virus in Caco-2 cells. *J. Virol.* **84**, 10016–10025 (2010).
32. Ou, X. et al. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat. Commun.* **11**, 1620 (2020).
33. 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 (2020).
34. Shulla, A. et al. A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry. *J. Virol.* **85**, 873–882 (2011).
35. Matsuyama, S. et al. Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. *Proc. Natl Acad. Sci. USA* **117**, 7001–7003 (2020).
36. Nao, N. et al. Consensus and variations in cell line specificity among human metapneumovirus strains. *PLoS ONE* **14**, e0215822 (2019).
37. Bertram, S. et al. TMPRSS2 activates the human coronavirus 229E for cathepsin-independent host cell entry and is expressed in viral target cells in the respiratory epithelium. *J. Virol.* **87**, 6150–6160 (2013).
38. Zheng, M. et al. Bat SARS-lke WIV1 coronavirus uses the ACE2 of multiple animal species as receptor and evades IFITM3 restriction via TMPRSS2 activation of membrane fusion. *Emerg. Microbes Infect.* **9**, 1567–1579 (2020).
39. Zhao, X. et al. LY6E restricts the entry of human coronaviruses, including the currently pandemic SARS-CoV-2. *J. Virol.* https://doi.org/10.1128/JVI.00562-20 (2020).
40. Huang, I. C. et al. Distinct patterns of IFITM-mediated restriction of filoviruses, SARS coronavirus, and influenza A virus. *PLoS Pathog.* **7**, e1001258 (2011).
41. Lin, T. Y. et al. Amphotericin B increases influenza A virus infection by preventing IFITM3-mediated restriction. *Cell Rep.* **5**, 895–908 (2013).
42. Kim, Y. I. et al. Infection and rapid transmission of SARS-CoV-2 in ferrets. *Cell Host Microbe* **27**, 704–709 (2020).
43. Richard, M. et al. SARS-CoV-2 is transmitted via contact and via the air between ferrets. *Nat. Commun.* **11**, 3496 (2020).

44. Belser, J. A. et al. Ferrets as models for influenza virus transmission studies and pandemic risk assessments. *Emerg. Infect. Dis.* **24**, 965–971 (2018).

45. Enserink, M. Coronavirus rips through Dutch mink farms, triggering culls. *Science* **368**, 1169 (2020).

46. Oude Munnink, B. B. et al. Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. *Science* https://doi.org/10.1126/science.abe5901 (2020).

47. Hanley, B. et al. Histopathological findings and viral tropism in UK patients with severe fatal COVID-19: a post-mortem study. *Lancet Microbe* https://doi.org/10.1016/S2666-5247(20)30115-4 (2020).

48. Lemmin, T., Kalbermatter, D., Harder, D., Plattet, P. & Fotiadis, D. Structures and dynamics of the novel S1/S2 protease cleavage site loop of the SARS-CoV-2 spike glycoprotein. *J. Struct. Biol. X* **4**, 100038 (2020).

49. Winstone, H. et al. The polybasic cleavage site in the SARS-CoV-2 spike modulates viral sensitivity to type I interferon and IFITM2. *J. Virol.* https://doi.org/10.1128/JVI.02422-20 (2021).

50. Zhang, L. et al. SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. *Nat. Commun.* **11**, 6013 (2020).

51. Hoffmann, M. et al. Chloroquine does not inhibit infection of human lung cells with SARS-CoV-2. *Nature* https://doi.org/10.1038/s41586-020-2575-3 (2020).

52. Ou, T. et al. Hydroxychloroquine-mediated inhibition of SARS-CoV-2 entry is attenuated by TMPRSS2. *PLoS Pathog.* **17**, e1009212 (2021).

53. Rambaut, A. et al. Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. virological.org https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563 (2020).

54. Kärber, G. Beitrag zur kollektiven Behandlung pharmakologischer Reihenversuche. *Naunyn Schmiedebergs Arch. Exp. Pathol. Pharmakol.* **162**, 480–483 (1931).

55. Long, J., Wright, E., Molesti, E., Temperton, N. & Barclay, W. Antiviral therapies against Ebola and other emerging viral diseases using existing medicines that block virus entry. *F1000Res* **4**, 30 (2015).

56. McKay, P. F. et al. Self-amplifying RNA SARS-CoV-2 lipid nanoparticle vaccine candidate induces high neutralizing antibody titers in mice. *Nat. Commun.* **11**, 3523 (2020).

57. Rebendenne, A. et al. SARS-CoV-2 triggers an MDA-5-dependent interferon response which is unable to control replication in lung epithelial cells. *J. Virol.* https://doi.org/10.1128/jvi.02415-20 (2021).

58. Edie, S. et al. Survey of human chromosome 21 gene expression effects on early development in *Danio rerio*. *G3 (Bethesda)* **8**, 2215–2223 (2018).

59. Sumner, R. P. et al. Disrupting HIV-1 capsid formation causes cGAS sensing of viral DNA. *EMBO J.* https://doi.org/10.15252/embj.2019103958 (2020).

60. Goldhill, D. H. et al. Determining the mutation bias of favipiravir in influenza virus using next-generation sequencing. *J. Virol.* https://doi.org/10.1128/JVI.01217-18 (2019).

61. Goldhill, D. H. et al. The mechanism of resistance to favipiravir in influenza. *Proc. Natl Acad. Sci. USA* **115**, 11613–11618 (2018).

62. Jabara, C. B., Jones, C. D., Roach, J., Anderson, J. A. & Swanstrom, R. Accurate sampling and deep sequencing of the HIV-1 protease gene using a Primer ID. *Proc. Natl Acad. Sci. USA* **108**, 20166–20171 (2011).

63. Ma, D. et al. Expression of SARS-CoV-2 receptor ACE2 and TMPRSS2 in human primary conjunctival and pterygium cell lines and in mouse cornea. *Eye (Lond.)* **34**, 1212–1219 (2020).

64. Xu, Q. F. et al. Ultraviolet a enhances cathepsin L expression and activity via JNK pathway in human dermal fibroblasts. *Chin. Med. J. (Engl.)* **129**, 2853–2860 (2016).

65. Zhou, Z. et al. The cAMP-responsive element binding protein (CREB) transcription factor regulates furin expression during human trophoblast syncytialization. *Placenta* **35**, 907–918 (2014).

66. Xu-yang, Z. et al. Interferon-induced transmembrane protein 3 inhibits Hantaan virus infection, and its single nucleotide polymorphism rs12252 influences the severity of hemorrhagic fever with renal syndrome. *Front. Immunol.* https://doi.org/10.3389/fimmu.2016.00535 (2017).

67. Li, H. et al. Internal genes of a highly pathogenic H5N1 influenza virus determine high viral replication in myeloid cells and severe outcome of infection in mice. *PLoS Pathog.* **14**, e1006821 (2018).

Author contributions

T.P.P., D.H.G., A.D.D., D.A.M. and W.S.B. conceived and planned experiments. T.P.P., D.H.G., J.Z., L.B., R.F., O.C.S., R.K., R.P., J.C.B., R.Y.S.-D., L.B., M.K.W., J.A.H. and E.S. performed the experiments. L.B., O.C.S., R.P., B.H. and M.O. processed the autopsy samples. M.O., M.G., A.D.D., D.A.M. and W.S.B. provided supervision. T.P.P., D.H.G. and W.S.B. wrote the manuscript with input from all other authors.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41564-021-00908-w.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41564-021-00908-w.

Correspondence and requests for materials should be addressed to W.S.B.

Peer review information *Nature Microbiology* thanks Lars Dolken and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

Empty Vector

MERS-CoV

SARS-CoV

SARS-CoV-2 WT

MonoCS

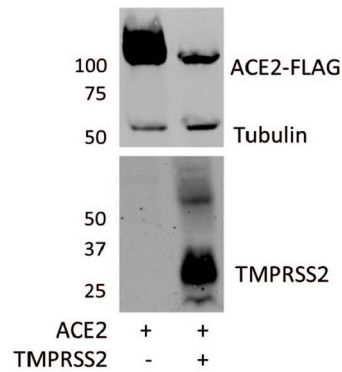**H5CS**

ΔCS

Δflank
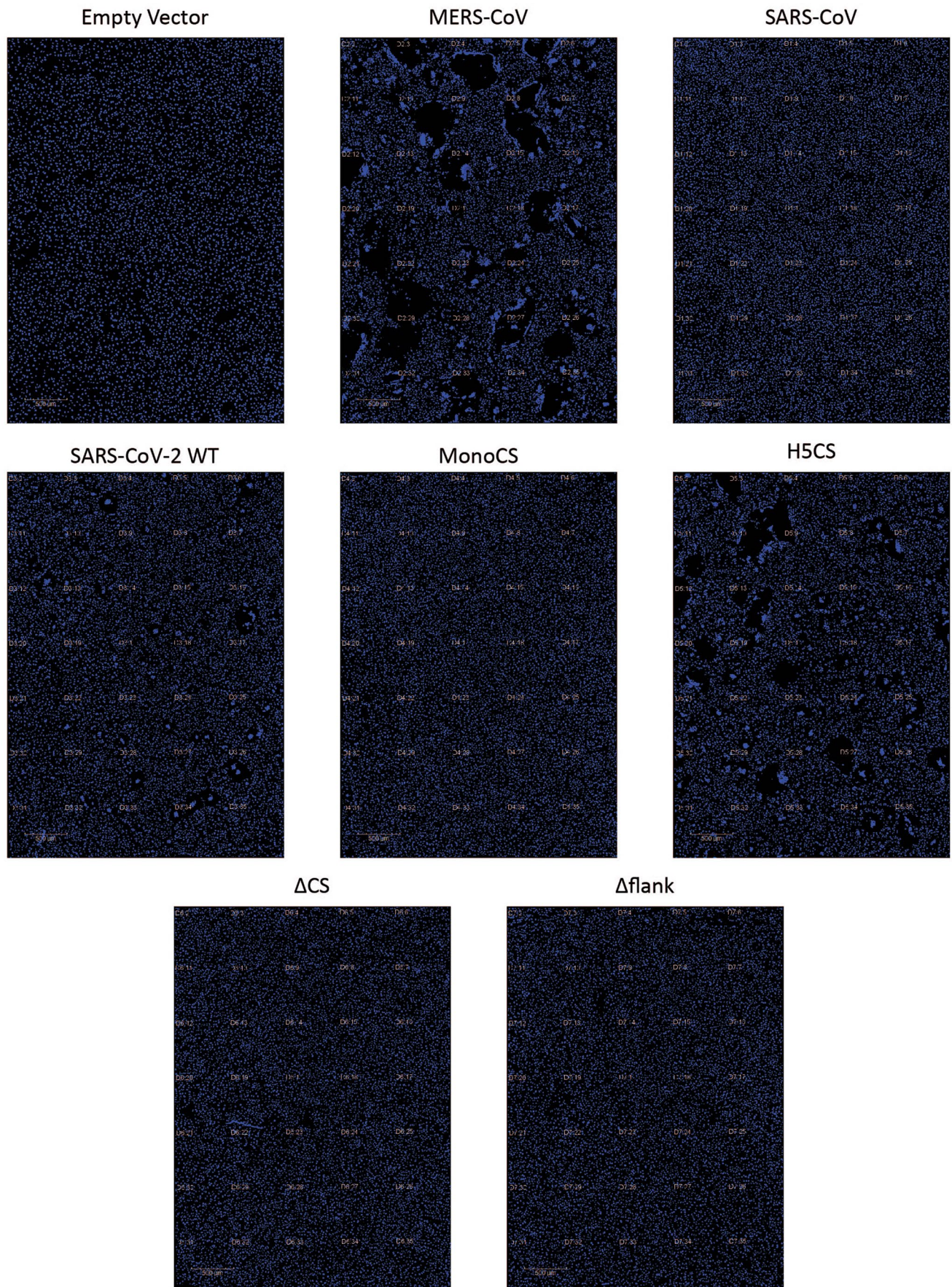
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Extended data for Fig. 1b – Spike-mediated syncytia formation in Vero cells. Representative images of Vero cells transfected with different spike proteins used for the analysis in main Fig. 1b.
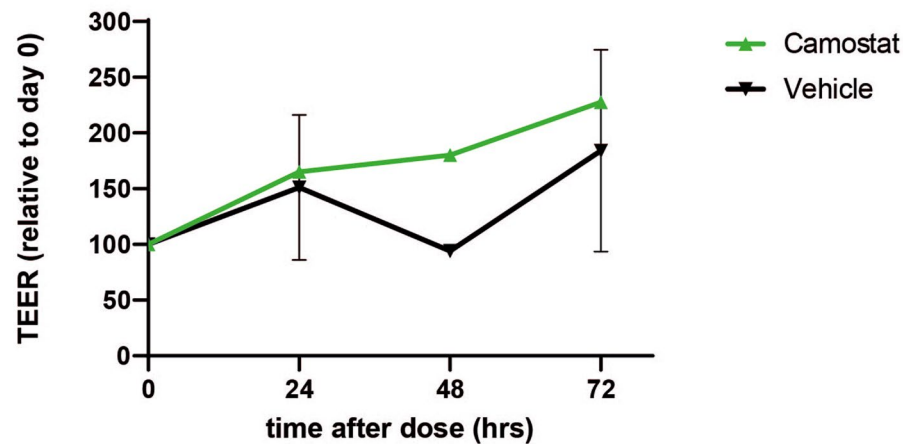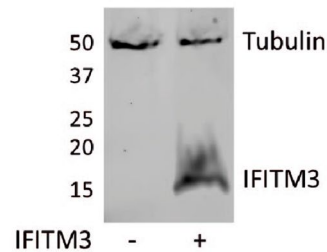
Extended Data Fig. 2 | Expression of transfected ACE2, TMPRSS2 and IFITM3. (**a**) Relative transepithelial electrical resistance (TEER) measurements of HAEs in the presence or absence of camostat. Assays performed in duplicate and plotted as mean + s.d. (**b**) Expression of TMPRSS2 and ACE2-FLAG in 293 T cells shown by western blot (from Fig. 3a). Representative western blot shown from N = 2 independent repeats. (**c**) Expression of IFITM3 in 293 T cells shown by western blot (from Fig. 3g). Representative western blot shown from N = 2 independent repeats.
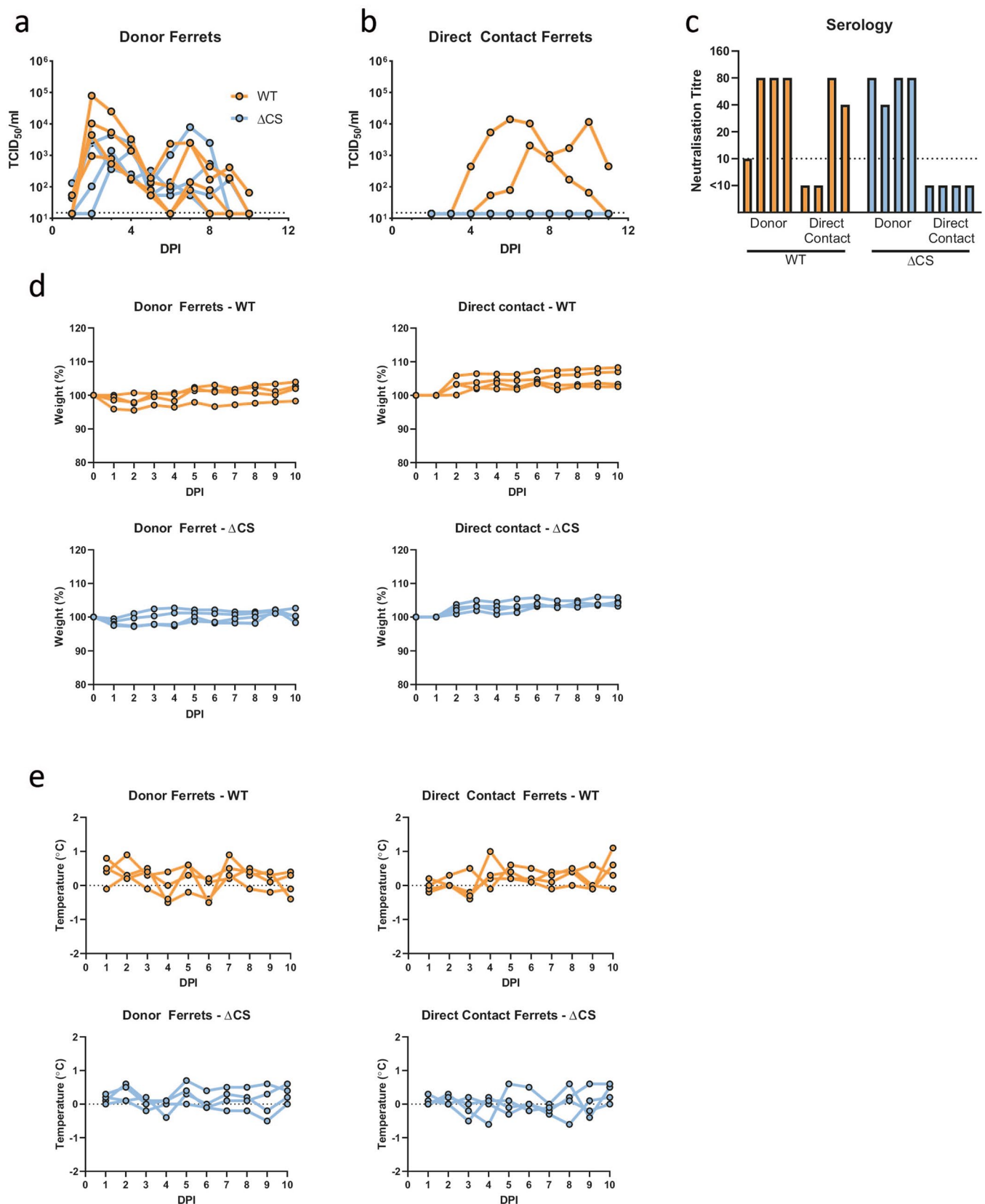
Extended Data Fig. 3 | Ferret head-to-head transmission experiment addition shedding data and clinical data. (a, b) Infectious virus titres of WT (orange) or ΔCS (blue) taken from donor ferret (a) and direct contact ferret (b) nasal washes as determined by $TCID_{50}$. Dotted line indicates limit of detection of infectious virus. **(c)** Microneutralisation assay showing ferret post-infection serology. Threshold of detection was a neutralisation titre of 10 (dotted line). Serum taken 14 days post-infection. **(d)** Changes in donor and direct contact ferret body weights during the duration of the infection. **(e)** Changes in donor and direct contact ferret body temperatures during the duration of the infection.
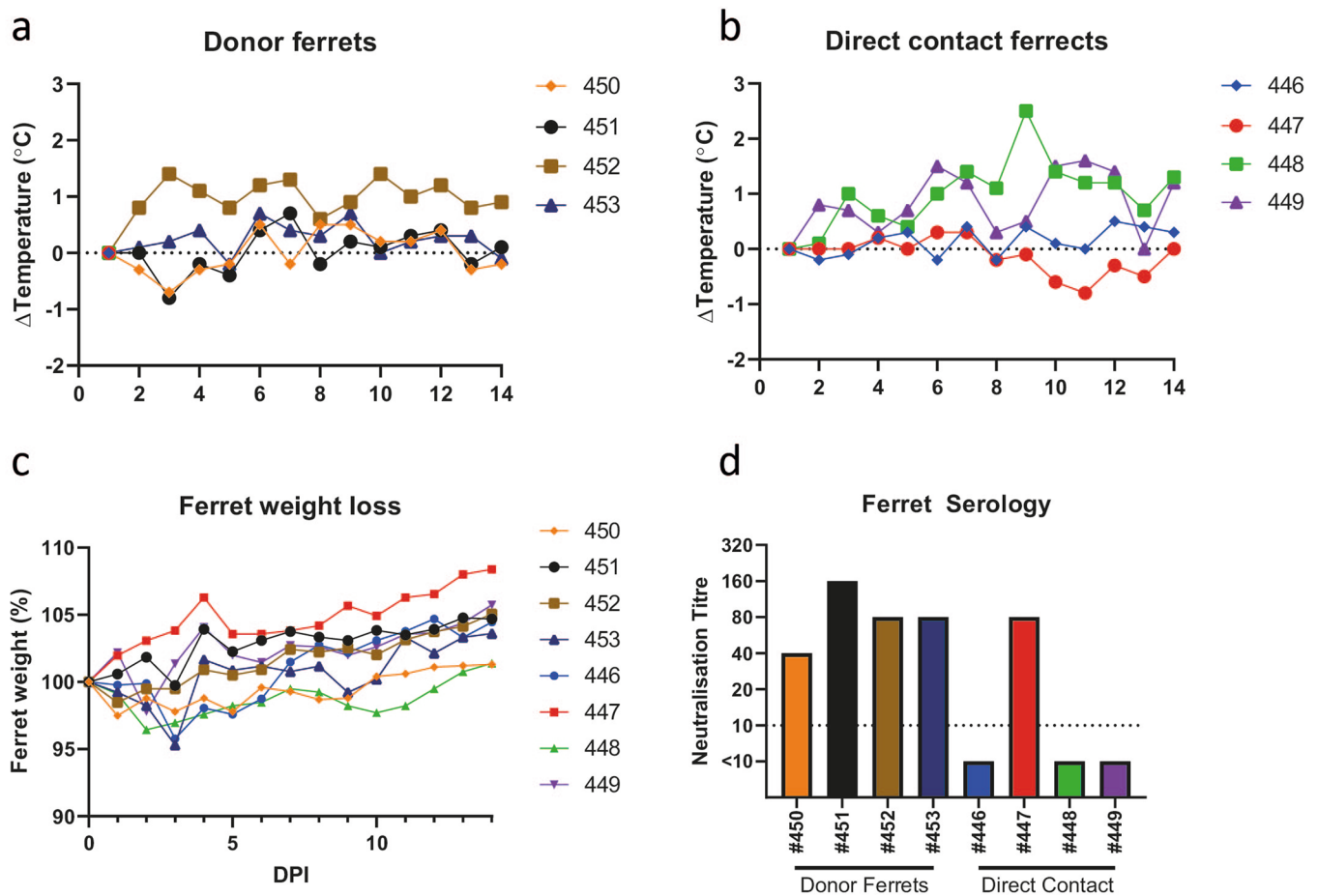
Extended Data Fig. 4 | Ferret competition transmission experiment clinical data. (**a, b**) Changes in donor (a) and direct contact (b) ferret body temperatures during the duration of the infection. (**c**) Changes in ferret body weight (in percentage) during the duration of the infection. (**d**) Microneutralisation assay showing ferret post-infection serology. Threshold of detection was a neutralisation titre of 10 (dotted line). Serum taken 14 days post-infection. Representative repeat shown from (N = 2) independent repeats.

Corresponding author(s):   Wendy Barclay

Last updated by author(s):   Mar 17, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	All analysis was performed in Geneious v11.1.5, R v3.5.0 or Harmony software (PerkinElmer). All code can be found at github.com/Flu1/Corona.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Raw sequences were deposited at www.ebi.ac.uk/ena, project number PRJEB40394.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was chosen for ferret experiments at the maximum available given the financial and logistical constraints involved in the experiment. All assays throughout were repeated on the day with at least 3 replicates, which has proven in the past to be sufficient to see a significant effect. Assays were further performed on at least 2 separate occasions (details provided in figure legends)
Data exclusions	2 post-mortem samples were excluded from the analysis as they had low number of reads similar to the negative control sample and therefore were likely contamination.
Replication	Ferret experiments were not replicated due to cost limitations. Other experimental data shown is representative of multiple experiments which replicated the findings. All assays throughout were repeated on the day with at least 3 replicates. Assays were further performed on at least 2 seperate occasions (details provided in figure legends)
Randomization	Randomization was not relevant for this study as assays were read in an unbiased manner - ie there was little or no interpretation involved that could be influenced by unconscious bias.
Blinding	Blinding was not possible/relevant for this study as assays were read in an unbiased manner - ie there was little or no interpretation involved that could be influenced by unconscious bias.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	mouse anti-FLAG (diluted 1/2000; F1804, Sigma), mouse anti-tubulin (diluted 1/5000; abcam; ab7291), mouse anti-p24 (diluted 1/2000; abcam; ab9071), rabbit anti-TMPRSS2 (diluted 1/2000; abcam; ab92323), rabbit anti-Fragilis/IFITM3 (diluted 1/2000; abcam; ab109429), rabbit anti-SARS spike protein (diluted 1/2000; NOVUS; NB100-56578) or rabbit anti-SARS-CoV-2 nucleocapsid (diluted 1/3000; SinoBiological; 40143-R019), IRDye® 680RD Goat anti-mouse (diluted 1/10,000; abcam; ab216776), IRDye® 680RD Goat anti-rabbit (diluted 1/10,000; abcam; ab216777), IRDye® 800CW Goat anti-mouse (diluted 1/10,000; abcam; ab216772), IRDye® 800CW Goat anti-rabbit (diluted 1/10,000; abcam; ab216773))
Validation	These are all commercially available antibodies with validation data available on the manufacturer's website. All antibodies used were specifically against the human version of antigens and used in this manuscript against the human version - therefore species validation is irrelevent. All antibodies were specifically chosen as they had been validated by Western Blot by the manufacturers.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Vero E6 - ATCC® CRL-1586; HEK 293T - ATCC® CRL-11268; Caco-2 - ATCC® HTB-37; Calu-3 - ATCC® HTB-55; VeroE6/TMPRSS2 - Centre for AIDS Reagents, NIBSC (100978).

Authentication	We did not authenticate the cell lines.
Mycoplasma contamination	We did not test for mycoplasma.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in the study.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Outbred female ferrets (16-20 weeks old), weighing 750-1000 g were used.
Wild animals	Study did not use wild animals.
Field-collected samples	Study did not use field-collected samples.
Ethics oversight	Animal research was carried out under a United Kingdom Home Office License, P48DAD9B4.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Information is irrelevent to the study which used a small set of human post-mortem samples from the Imperial College Healthcare Tissue Bank (ICHTB).
Recruitment	Human samples used in this research project were obtained from the Imperial College Healthcare Tissue Bank (ICHTB).
Ethics oversight	ICHTB is approved by Wales REC3 to release human material for research (17/WA/0161), and the samples for this project (R20012) were issued from subcollection reference number MED_MO_20_011.

Note that full information on the approval of the study protocol must also be provided in the manuscript.