Article

Identification of SARS-CoV-2 inhibitors using lung and colonic organoids

https://doi.org/10.1038/s41586-020-2901-9

Received: 5 May 2020

Accepted: 21 October 2020

Published online: 28 October 2020

Check for updates

Yuling Han[1,15], Xiaohua Duan[1,2,15], Liuliu Yang[1,15], Benjamin E. Nilsson-Payant[3,15], Pengfei Wang[4,15], Fuyu Duan[5,15], Xuming Tang[1,15], Tomer M. Yaron[6,7,15], Tuo Zhang[8], Skyler Uhl[3], Yaron Bram[9], Chanel Richardson[10], Jiajun Zhu[1], Zeping Zhao[1], David Redmond[11], Sean Houghton[11], Duc-Huy T. Nguyen[9], Dong Xu[8], Xing Wang[8], Jose Jessurun[12], Alain Borczuk[12], Yaoxing Huang[4], Jared L. Johnson[6], Yuru Liu[13], Jenny Xiang[8], Hui Wang[2]✉, Lewis C. Cantley[6]✉, Benjamin R. tenOever[3]✉, David D. Ho[4]✉, Fong Cheng Pan[1]✉, Todd Evans[1]✉, Huanhuan Joyce Chen[5,6]✉, Robert E. Schwartz[9,14]✉ & Shuibing Chen[1]✉

There is an urgent need to create novel models using human disease-relevant cells to study severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) biology and to facilitate drug screening. Here, as SARS-CoV-2 primarily infects the respiratory tract, we developed a lung organoid model using human pluripotent stem cells (hPSC-LOs). The hPSC-LOs (particularly alveolar type-II-like cells) are permissive to SARS-CoV-2 infection, and showed robust induction of chemokines following SARS-CoV-2 infection, similar to what is seen in patients with COVID-19. Nearly 25% of these patients also have gastrointestinal manifestations, which are associated with worse COVID-19 outcomes[1]. We therefore also generated complementary hPSC-derived colonic organoids (hPSC-COs) to explore the response of colonic cells to SARS-CoV-2 infection. We found that multiple colonic cell types, especially enterocytes, express ACE2 and are permissive to SARS-CoV-2 infection. Using hPSC-LOs, we performed a high-throughput screen of drugs approved by the FDA (US Food and Drug Administration) and identified entry inhibitors of SARS-CoV-2, including imatinib, mycophenolic acid and quinacrine dihydrochloride. Treatment at physiologically relevant levels of these drugs significantly inhibited SARS-CoV-2 infection of both hPSC-LOs and hPSC-COs. Together, these data demonstrate that hPSC-LOs and hPSC-COs infected by SARS-CoV-2 can serve as disease models to study SARS-CoV-2 infection and provide a valuable resource for drug screening to identify candidate COVID-19 therapeutics.

The development of anti-SARS-CoV-2 drugs could change the scope of the ongoing COVID-19 pandemic. In the pursuit of this strategy, high-throughput screens are typically performed in transformed cell lines that fail to capture the physiologically relevant dynamics of human SARS-CoV-2 infection. To overcome the limitations of these cell lines, several adult organoid models have been developed to study SARS-CoV-2 (refs. [2–4]). Here we developed hPSC-LOs and hPSC-COs optimized as in vitro platforms for high-throughput drug screening.

hPSC-LOs are permissive to SARS-CoV-2

We differentiated hPSCs to lung organoids (hPSC-LOs) on the basis of previously reported stepwise strategies[5–13] (Extended Data Fig. 1a–c). Quantitative PCR with reverse transcription (qRT–PCR) and RNA sequencing (RNA-seq) profiling validated the expression of alveolar type II (AT2) cell markers in the hPSC-LOs (Extended Data Fig. 1d, e). Intracellular flow cytometry further confirmed the presence of pro-SP-C$^+$ cells in hPSC-LOs (Extended Data Fig. 1f). Single-cell transcriptomic profiles of hPSC-LOs identified AT2-like cells, which were enriched for adult human lung AT2 cell markers (Fig. 1a–c and Extended Data Fig. 2a–c). Correlation analysis of signature genes further validated the AT2-like cell population in hPSC-LOs showing high similarity to adult human lung AT2 cells (Fig. 1d). The key factors involved in SARS-CoV-2 entry—ACE2 (ref. [14]), the SARS-CoV-2 receptor; TMPRSS2 (ref. [14]), a protease involved in viral entry; and FURIN[15], a pro-protein convertase pre-activating SARS-CoV-2—were enriched

[1]Department of Surgery, Weill Cornell Medicine, New York, NY, USA. [2]State Key Laboratory of Oncogenes and Related Genes, Center for Single-Cell Omics, School of Public Health, Shanghai Jiao Tong University School of Medicine, Shanghai, China. [3]Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [4]Aaron Diamond AIDS Research Center, Columbia University Vagelos College of Physicians and Surgeons, New York, NY, USA. [5]Pritzker School of Molecular Engineering and Ben May Department, University of Chicago, Chicago, IL, USA. [6]Meyer Cancer Center, Weill Cornell Medicine, New York, NY, USA. [7]Englander Institute for Precision Medicine, Institute for Computational Biomedicine, Weill Cornell Medicine, New York, NY, USA. [8]Genomics Resources Core Facility, Weill Cornell Medicine, New York, NY, USA. [9]Division of Gastroenterology and Hepatology, Department of Medicine, Weill Cornell Medicine, New York, NY, USA. [10]Department of Pharmacology, Weill Cornell Medicine, New York, NY, USA. [11]Division of Regenerative Medicine, Ansary Stem Cell Institute, Weill Cornell Medicine, New York, NY, USA. [12]Department of Pathology and Laboratory Medicine, Weill Cornell Medicine, New York, NY, USA. [13]Department of Pharmacology and Regenerative Medicine, University of Illinois College of Medicine, Chicago, IL, USA. [14]Department of Physiology, Biophysics and Systems Biology, Weill Cornell Medicine, New York, NY, USA. [15]These authors contributed equally: Yuling Han, Xiaohua Duan, Liuliu Yang, Benjamin E. Nilsson-Payant, Pengfei Wang, Fuyu Duan, Xuming Tang, Tomer M. Yaron. ✉e-mail: huiwang@shsmu.edu.cn; LCantley@med.cornell.edu; benjamin.tenoever@mssm.edu; dh2994@cumc.columbia.edu; fcp2002@med.cornell.edu; tre2003@med.cornell.edu; joycechen@uchicago.edu; res2025@med.cornell.edu; shc2034@med.cornell.edu

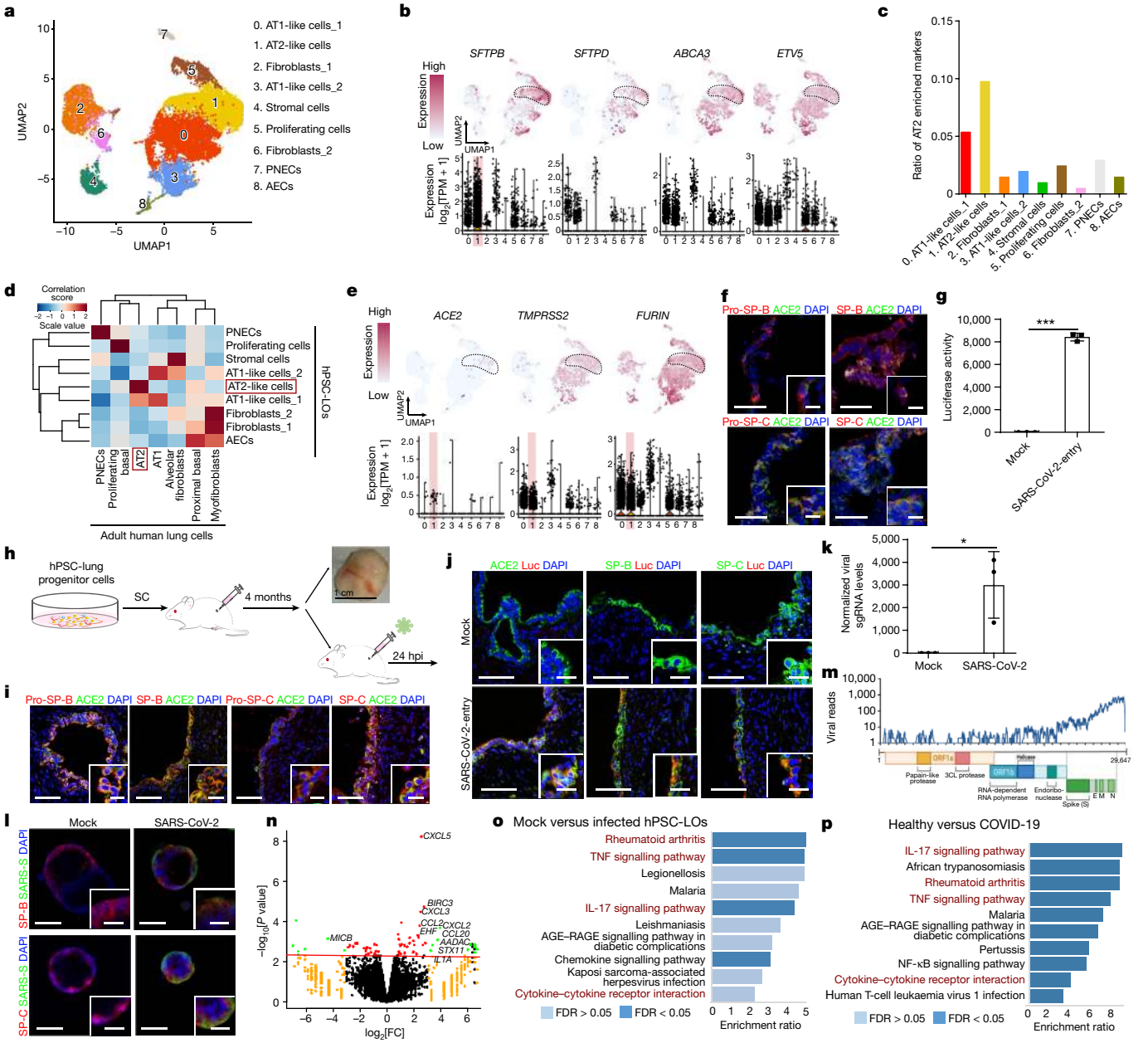


Fig. 1 | hPSC-LOs are permissive to SARS-CoV-2 infection both in vitro and in vivo. a, Uniform manifold approximation and projection (UMAP) of hPSC-LOs. PNECs, pulmonary neuroendocrine cells; AECs, airway epithelial cells. **b**, AT2 cell markers in each cluster in UMAPs. The vertical bars in the bottom panels represent the expression level.The pink bar highlights cluster_1 AT2-like cells. **c**, Enrichment analysis of hPSC-LOs using genes expressed at high levels in adult human AT2 cells. **d**, Correlation analysis of genes with cell fates in hPSC-LOs and adult human lung cells. **e**, UMAPs of *ACE2*, *TMPRSS2* and *FURIN* expression in hPSC-LOs. The vertical bars in the bottom panels represent the expression level. The pink bars highlight cluster_1 AT2-like cells. **f**, Immunostaining of hPSC-LOs. Scale bars, 30 µm (main images) and 10 µm (insets). DAPI, 4′,6-diamidino-2-phenylindole. **g**, Luciferase activity at 24 hpi for hPSC-LOs either mock-infected or infected with SARS-CoV-2-entry virus (multiplicity of infection (MOI) = 0.01). $n = 3$ biologically independent experiments. $^{***}P = 1.62 \times 10^{-6}$. **h**, A schematic of the protocol for in vivo transplantation of hPSC-derived lung xenografts. SC, subcutaneous. **i**, Immunostaining of hPSC-derived lung xenografts. Scale bars, 75 µm (main images) and 10 µm (insets). **j**, Immunostaining of hPSC-derived lung xenografts at 24 hpi (1×10^4 focus-forming units (FFUs)). Scale bars, 75 µm (main images) and 10 µm (insets). Luc, luciferase. **k**, qRT–PCR analysis of total RNA extracted from infected hPSC-LOs (24 hpi, MOI = 0.01) for viral *N* sgRNA. $n = 3$ biologically independent experiments. $^{*}P = 0.0236$. **l**, Immunostaining of hPSC-LOs at 24 hpi (SARS-CoV-2, MOI = 0.01). Scale bars, 50 µm (main images) and 10 µm (insets). **m**, Alignment of the transcriptome with the viral genome in SARS-CoV-2-infected hPSC-LOs. The schematic below shows the SARS-CoV-2 genome. **n**, Volcano plot analysis of differential expression of SARS-CoV-2-infected hPSC-LOs versus mock infection. The red line indicates a *P*-adjusted value < 0.05. **o**, Gene over-representation analysis using the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database of SARS-CoV-2-versus mock-infected hPSC-LOs. $n = 3$ biologically independent experiments. **p**, Gene over-representation analysis using the KEGG pathway database of lung autopsy tissues from patients with COVID-19 versus healthy individuals. $n = 3$ individuals for each group. $^{*}P < 0.05$; $^{***}P < 0.001$. The data were analysed by an unpaired two-tailed Student's *t*-test and are shown as mean ± s.d. The data are representative of at least three independent experiments.

in AT2-like cells (Fig. 1e). Immunostaining analysis further validated that ACE2 is expressed in pro-SP-B$^+$SP-B$^+$pro-SP-C$^+$SP-C$^+$AT2-like cells (Fig. 1f and Extended Data Fig. 2d).

To determine the permissiveness of hPSC-LOs to SARS-CoV-2 entry, we first used a vesicular stomatitis ΔG-luciferase virus pseudotyped with the SARS-CoV-2 spike protein (SARS-CoV-2-entry virus)[16,17]. Robust

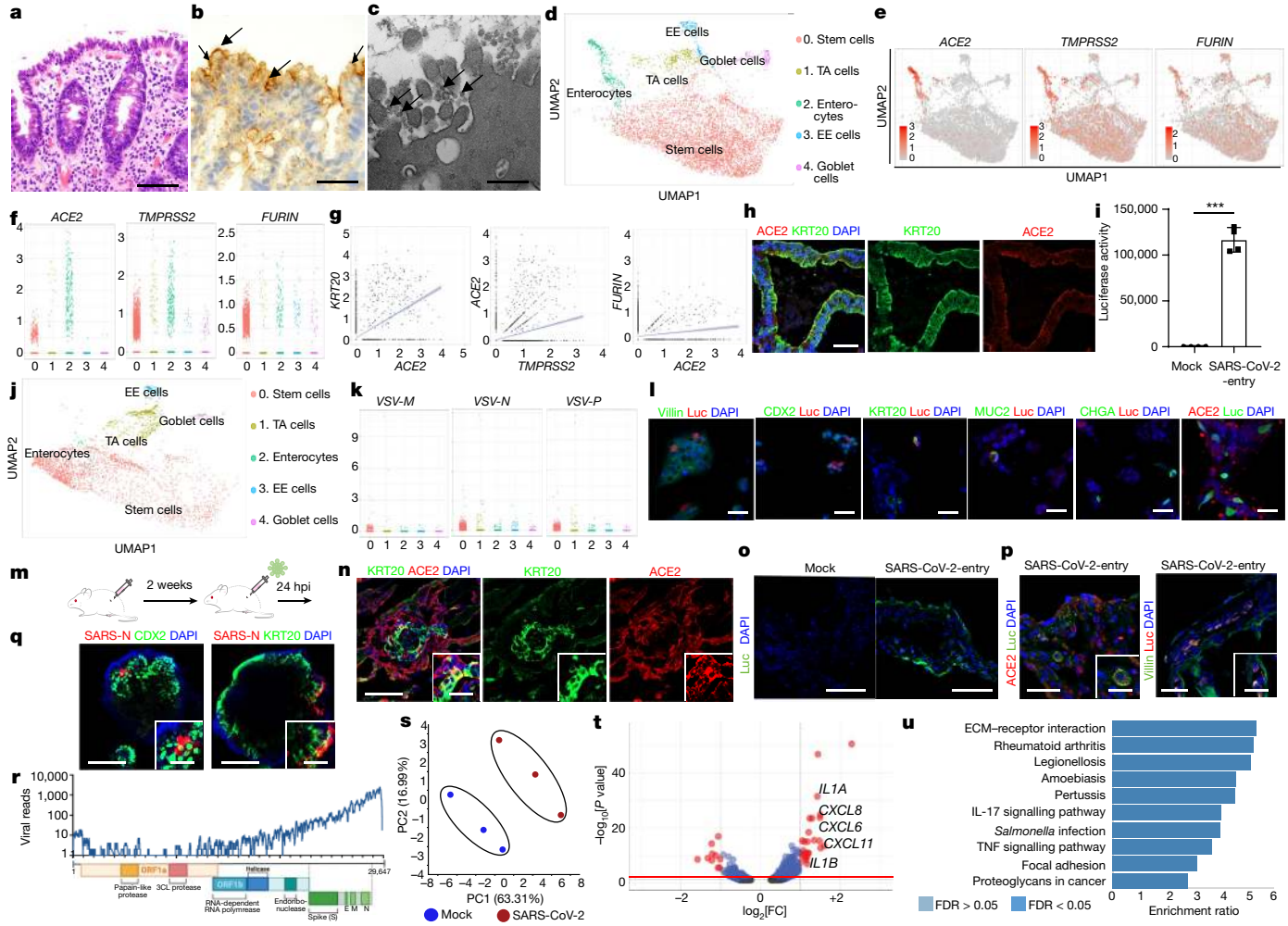


Fig. 2 | hPSC-COs are permissive to SARS-CoV-2 infection. a–**c**, Haematoxylin and eosin staining (**a**), in situ hybridization staining for SARS-CoV-2 RNA (**b**) and electron microscopy (**c**) of colonoscopy biopsy tissue from a patient with COVID-19. Scale bars, 50 µm (**a**), 25 µm (**b**) and 1 µm (**c**). The arrows indicate SARS-CoV-2 RNA (**b**) or SARS-CoV-2 viral particles (**c**). **d**, UMAP of hPSC-CO cell types. EE, enteroendocrine; TA, transit-amplifying. **e**, **f**, UMAPs (**e**) and jitter plots (**f**) of *ACE2*, *TMPRSS2* and *FURIN*. **g**, Correlation of the expression levels for *KRT20* with *ACE2*, *TMPRSS2* and *FURIN*. The blue line with grey shading represents the fitted regression line with 95% confidence interval. **h**, Confocal images of hPSC-COs. Scale bar, 100 µm. **i**, **j**, Luciferase activity (**i**) and UMAP (**j**) of COs infected with SARS-CoV-2-entry virus (24 hpi, MOI = 0.01). $n = 4$ biologically independent experiments (**i**). $^{***}P = 2.25 \times 10^{-6}$. **k**, Jitter plots of *VSV-M*, *VSV-N* and *VSV-P* transcript levels. **l**, Immunostaining of hPSC-COs infected with SARS-CoV-2-entry virus (MOI = 0.01). Scale bars, 50 µm. **m**, A schematic of the protocol for in vivo infection. **n**, Confocal images of a colonic xenograft. Scale bars, 100 µm (main images) and 30 µm (insets). **o**, **p**, Confocal images of colonic xenografts at 24 hpi (1×10^3 FFUs) stained with antibodies against luciferase (**o**) and ACE2 or villin (**p**). Scale bar, 75 µm (main images) and 25 µm (insets). **q**, Immunostaining to detect SARS-CoV-2 nucleocapsid protein (SARS-N) in hPSC-COs. Scale bars, 100 µm (main images) and 40 µm (insets). **r**, RNA-seq read coverage of the viral genome in infected hPSC-COs (24 hpi, MOI = 0.1). **s**–**u**, A PCA plot (**s**), a volcano plot (**t**) and GSEA pathway analysis (**u**) of gene expression profiles from mock-treated and SARS-CoV-2-infected hPSC-COs at 24 hpi (MOI = 0.1). The vertical dotted lines in **t** represent the $\log_2$-transformed fold change (FC) > 1 and $\log_2$-transformed FC < −1, respectively. The horizontal dotted line in **t** represents $P < 0.05$. ECM, extracellular matrix. $n = 3$ biologically independent experiments. The red line indicates $P = 0.05$. $^{***}P < 0.001$. The data were analysed by an unpaired two-tailed Student's *t*-test and are shown as mean ± s.d. The data are representative of at least three independent experiments.

luciferase activity was readily detected in hPSC-LOs infected with SARS-CoV-2-entry virus (Fig. 1g).

To generate an in vivo model using hPSC-LOs, we subcutaneously implanted lung progenitor cells into immunodeficient *NOD-scid IL2Rg*^{null} (NSG) mice (Fig. 1h). The xenografts developed organized distal lung-like structures with AT2-like cells expressing ACE2 (Fig. 1i). At 24 h after intra-xenograft inoculation with SARS-CoV-2-entry virus, luciferase was mainly detected in AT2-like cells (Fig. 1j).

Next, hPSC-LOs were infected with SARS-CoV-2 (isolate USA-WA1/2020). At 24 h post infection (hpi), qRT–PCR confirmed significant amounts of viral replication in infected hPSC-LOs (Fig. 1k). Immunostaining for the spike protein confirmed robust SARS-CoV-2 infection of hPSC-LOs (Fig. 1l and Extended Data Fig. 2e). The viral infection was further confirmed by RNA-seq analysis (Fig. 1m). Moreover, principal component analysis (PCA) demonstrated that the infected hPSC-LOs occupied a distinct transcriptional space compared to mock-infected hPSC-LOs (Extended Data Fig. 2f). Volcano plots of mock- versus SARS-CoV-2-infected hPSC-LOs revealed robust induction of chemokine transcripts (Fig. 1n). Gene set enrichment analysis (GSEA) revealed over-represented pathway networks including rheumatoid arthritis, tumour-necrosis factor signalling, interleukin-17 signalling and cytokine–cytokine receptor interaction (Fig. 1o), which is similar to the pathways enriched in lung autopsy tissues of patients with COVID-19[18] (Fig. 1p and Supplementary Table 1).

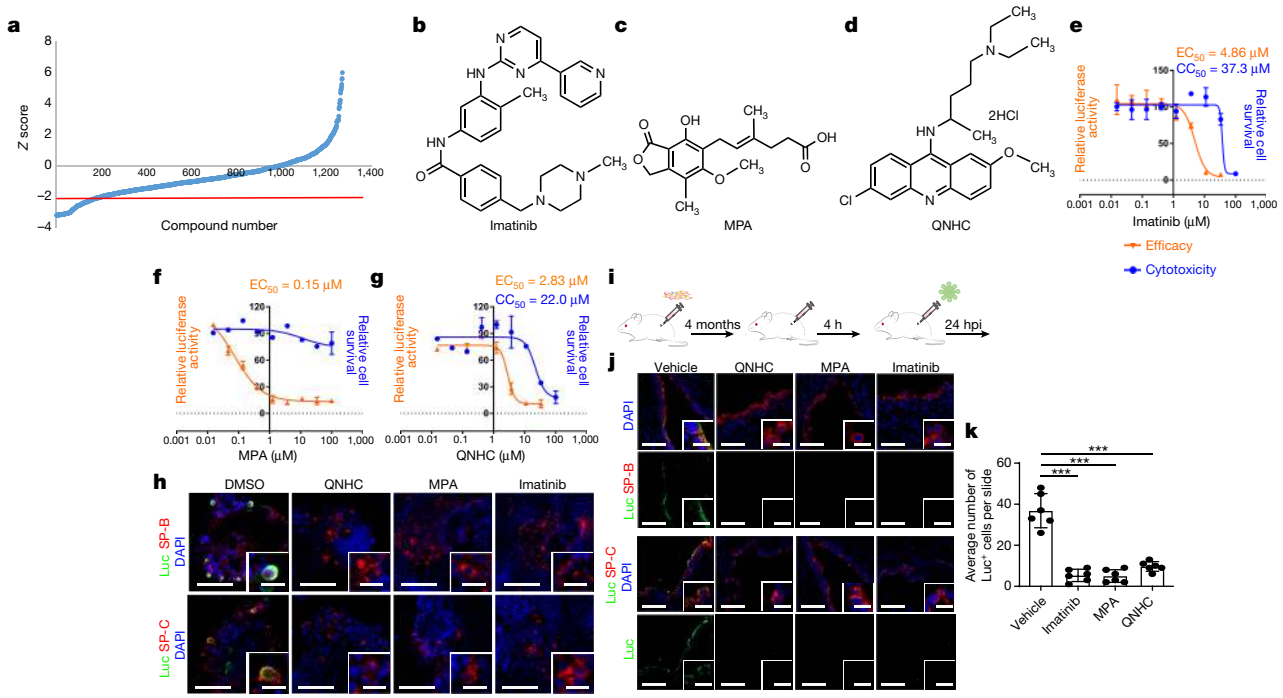


Fig. 3 | An hPSC-LO-based high-throughput chemical screen identifies three FDA-approved drug candidates that block SARS-CoV-2 entry. **a**, Primary screening results. The red line indicates a *Z* score of less than −2, which means the luminescent signal is lower than average − 2 × s.d. **b**–**d**, The chemical structures of imatinib (**b**), MPA (**c**) and QNHC (**d**). **e**–**g**, Efficacy and toxicity curves for imatinib (**e**), MPA (**f**) and QNHC (**g**). $n = 3$ biologically independent experiments. **h**, Immunofluorescent staining of luciferase$^+$ cells in imatinib, MPA and QNHC-treated hPSC-LOs at 24 hpi (MOI = 0.01). Scale bars, 50 μm (main images) and 10 μm (insets). **i**, A scheme of the in vivo drug treatment. **j**, **k**, Immunostaining (**j**) and quantification (**k**) of hPSC-derived lung xenografts of mice treated with 400 mg kg^{-1} imatinib mesylate, 50 mg kg^{-1} MPA and 25 mg kg^{-1} QNHC at 24 hpi (1×10^4 FFUs). Scale bars, 100 μm (main images) and 10 μm (insets). $n = 6$ xenografts for each group. $^{***}P = 5.54 \times 10^{-6}$; $^{***}P = 5.19 \times 10^{-6}$; $^{***}P = 1.66 \times 10^{-5}$ ($^{***}P < 0.001$). The data were analysed by an unpaired two-tailed Student's *t*-test and are shown as mean ± s.d. The data are representative of at least three independent experiments.

hPSC-COs are permissive to SARS-CoV-2

As gastrointestinal complications are associated with worse outcomes of patients with COVID-19 (ref. [1]), we examined whether SARS-CoV-2 can infect colonic cells. First, immunohistochemistry confirmed ACE2 expression in keratin 20 (KRT20)$^+$ enterocytes of human colon tissue (Extended Data Fig. 3a). Pathological analysis of colonoscopy samples from patients with COVID-19 detected significant damage to the colon mucosa marked by injury to the luminal epithelial cells and goblet cell depletion (Fig. 2a), SARS-CoV-2 infection was confirmed by in situ hybridization for viral RNA (Fig. 2b) and by electron microscopy (Fig. 2c).

We next derived hPSC-COs from HUES8 cells using previously established strategies[19,20] (Extended Data Fig. 3b–e). Using single-cell RNA-seq (scRNA-seq), five cell clusters were identified including KRT20$^+$ enterocytes, MUC2$^+$ goblet cells, EPHB2$^+$ transit-amplifying cells, CHGA$^+$ enteroendocrine cells, and LGR5$^+$ or BMI1$^+$ stem cells, and most cells expressed CDX2 and VIL1 (Fig. 2d and Extended Data Fig. 4a–c). ACE2, TMPRSS2 (ref. [14]) and FURIN were also expressed in all five cell clusters, but highly enriched and correlated in KRT20$^+$ enterocytes (Fig. 2e–g). ACE2 expression in KRT20$^+$ enterocytes was further validated by immunostaining (Fig. 2h).

To test the permissiveness of hPSC-COs to SARS-CoV-2, hPSC-COs were infected with SARS-CoV-2-entry virus, and showed strong luciferase activity at 24 hpi (Fig. 2i and Extended Data Fig. 5a). scRNA-seq of the infected hPSC-COs revealed the same five cell populations as in the uninfected condition (Fig. 2j and Extended Data Fig. 5b–d). However, after infection, the KRT20$^+$ enterocyte population decreased significantly (Fig. 2j). This corresponded with increased cellular apoptosis (Extended Data Fig. 5e) and the depletion of the ACE2$^+$ population (Extended Data Fig. 5c, d). Viral RNA from the SARS-CoV-2-entry virus was detected in all five cell populations after SARS-CoV-2-entry virus infection (Fig. 2k), but not in the uninfected hPSC-COs (Extended Data Fig. 5f), suggesting that all cell types within hPSC-COs were permissive to SARS-CoV-2-entry virus. This was also confirmed by detection of luciferase expression in ACE2$^+$, Villin$^+$, CDX2$^+$, KRT20$^+$ and MUC2$^+$ cells (Fig. 2l). ACE2$^+$TMPRSS2$^+$, ACE2$^+$ and TMPRSS2$^+$ cells were all detected in the population of VSV$^+$ cells (Extended Data Fig. 5g). However, the failure to measure *ACE2* or *TMPRSS2* transcripts might reflect limitations in the detection sensitivity using the 10x Genomics scRNA-seq platform. On the basis of immunostaining, all luciferase$^+$ cells were ACE2$^+$ (Fig. 2l).

Humanized mice carrying hPSC-COs in vivo provide a unique platform for modelling COVID-19. In brief, hPSC-COs were transplanted under the kidney capsule of NSG mice (Fig. 2m). ACE2 was detected in hPSC-derived KRT20$^+$ enterocytes (Fig. 2n), consistent with the in vitro culture findings. At 24 h after intra-xenograft inoculation with SARS-CoV-2-entry virus, luciferase was detected in the infected xenografts (Fig. 2o), in particular in ACE2$^+$ and villin$^+$ colonic cells, suggesting that these cells are permissive to SARS-CoV-2-entry virus in vivo (Fig. 2p).

Next, hPSC-COs were infected with SARS-CoV-2 virus and viral nucleocapsid protein was detected in the infected hPSC-COs at 24 hpi, partially colocalizing with CDX2 and KRT20 (Fig. 2q and Extended Data Fig. 5h). Bulk RNA-seq confirmed robust viral infection of hPSC-COs (Fig. 2r), distinct transcriptional profiles of mock-treated and infected hPSC-COs (Fig. 2s) and notable differential gene expression of cytokines and chemokines (Fig. 2t). GSEA revealed over-represented pathway networks in hPSC-COs similar to those of SARS-CoV-2-infected hPSC-LOs and lung autopsy tissues of patients with COVID-19 (Fig. 2u)[18].

An hPSC-LOs-based anti-SARS-CoV-2 screen

To identify drug candidates capable of inhibiting SARS-CoV-2 entry, hPSC-LOs were treated with a library of drugs approved by the FDA,

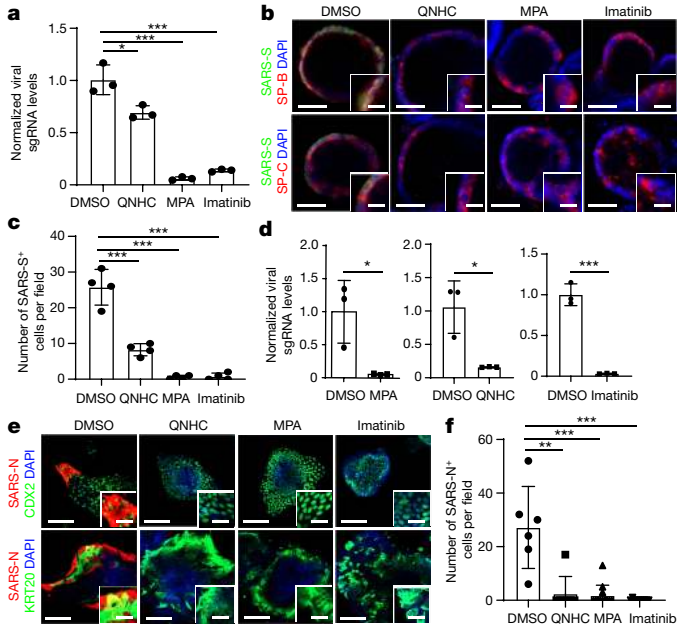


Fig. 4 | Imatinib, MPA and QNHC each block the entry of SARS-CoV-2 in both hPSC-derived LOs and COs. a–**c**, qRT–PCR analysis of total RNA extracted from infected hPSC-LOs for viral *N* sgRNA (**a**; $n = 3$ biologically independent experiments; $^{*}P = 0.0256$; $^{***}P = 0.000333$; $^{***}P = 0.000461$), immunofluorescent staining (**b**) and quantification (**c**; $n = 4$ biological replicates for each group; $^{***}P = 6.36 \times 10^{-5}$; $^{***}P = 5.63 \times 10^{-5}$; $^{***}P = 0.000566$) of SARS-CoV-2 spike protein (SARS-S) and SP-B/SP-C in imatinib-, MPA- or QNHC-treated hPSC-LOs at 24 hpi (MOI = 0.5). Scale bars, 50 μm (main images) and 15 μm (insets). **d**–**f**, qRT–PCR analysis of total RNA extracted from infected hPSC-COs for viral N sgRNA (**d**; $n = 3$ biologically independent experiments; $^{*}P = 0.0260$; $^{*}P = 0.0166$; $^{***}P = 0.000235$), immunofluorescent staining (**e**) and quantification (**f**; $n = 6$ biologically independent experiments; $^{**}P = 0.00242$; $^{***}P = 4.34 \times 10^{-5}$; $^{***}P = 5.26 \times 10^{-5}$) of SARS-CoV-2 nucleocapsid protein (SARS-N) of imatinib-, MPA- or QNHC-treated hPSC-COs at 24 hpi (MOI = 0.5). Scale bars, 50 μm (main images) and 15 μm (insets). $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$. The data were analysed by an unpaired two-tailed Student's *t*-test and are shown as mean ± s.d. The data are representative of at least three independent experiments.

followed by infection with SARS-CoV-2-entry virus. At 24 hpi, the organoids were analysed for luciferase activity. Compounds with a *Z* score of less than −2 were chosen as primary hit drugs (Fig. 3a). Four drugs were confirmed to block luciferase activity in a dose-dependent manner, regardless of cytotoxicity, including imatinib (half-maximum effective concentration (EC_{50}) = 4.86 μM, half-maximum cytotoxic concentration (CC_{50}) = 37.3 μM) (Fig. 3b, e), mycophenolic acid (MPA) (EC_{50} = 0.15 μM) (Fig. 3c, f), quinacrine dihydrochloride (QNHC) (EC_{50} = 2.83 μM, CC_{50} = 22 μM) (Fig. 3d, g) and chloroquine (EC_{50} = 3.85 μM) (Extended Data Fig. 6a). Immunostaining confirmed a significant reduction of luciferase^{+} cells detected among SP-B^{+}SP-C^{+} AT2-like cells in hPSC-LOs treated with 10 μM imatinib, 3 μM MPA or 4.5 μM QNHC (Fig. 3h and Extended Data Fig. 6b). None of the tested drugs showed toxicity-independent inhibition of VSV-G luciferase virus, suggesting that imatinib, MPA and QNHC specifically block SARS-CoV-2-entry virus infection in hPSC-LOs, instead of inhibiting luciferase activity (Extended Data Fig. 6c). Imatinib, MPA and QNHC also inhibited SARS-CoV-1-entry virus infection in a dose-dependent manner (imatinib: IC_{50} = 0.17 μM; MPA: IC_{50} = 0.7 μM; QNHC: IC_{50} = 0.30 μM) (Extended Data Fig. 6d).

To evaluate the drug activities in vivo, we treated humanized mice carrying hPSC-derived lung xenografts with imatinib mesylate, MPA or QNHC before intra-xenograft inoculation with SARS-CoV-2-entry virus (Fig. 3i). As observed for hPSC-LOs, luciferase staining was significantly decreased in SP-B^{+}SP-C^{+} AT2-like cells treated with imatinib mesylate, MPA or QNHC (Fig. 3j, k).

Drugs block SARS-CoV-2 infection

We next treated hPSC-LOs or hPSC-COs with 10 μM imatinib, 3 μM MPA or 4.5 μM QNHC and infected each culture with SARS-CoV-2. At 24 hpi, all three drugs block SARS-CoV-2 infection in a dose-dependent manner (Extended Data Fig. 7a and Supplementary Table 2). In hPSC-LOs, drug treatment before infection resulted in significantly reduced levels of viral subgenomic RNA (sgRNA) (Fig. 4a), as well as spike protein expression (Fig. 4b,c and Extended Data Fig. 7b). Drug treatment before infection of hPSC-COs also led to significantly reduced viral sgRNA levels (Fig. 4d and Extended Data Fig. 7c), as well as nucleocapsid protein expression (Fig. 4e, f and Extended Data Fig. 7d). As most studies on SARS-CoV-2 are performed in the African green monkey Vero E6 cell line, we verified that imatinib, MPA and QNHC block SARS-CoV-2 infection in a toxicity-independent manner in Vero E6 cells as in hPSC-LOs and hPSC-COs (Extended Data Figs. 8 and 9 and Supplementary Table 2).

We briefly explored the effect of imatinib, MPA and QNHC on the key steps of SARS-CoV-2 entry. Surface plasmon resonance binding analysis suggested that both imatinib and QNHC bind with ACE2 (Extended Data Fig. 10a). Treatment with MPA or QNHC decreases the expression levels of FURIN (Extended Data Fig. 10b, c). To explore the inhibitory mechanism of imatinib, RNA-seq analysis of dimethylsulfoxide (DMSO)- and imatinib-treated hPSC-LOs was performed, showing distinct transcriptional profiles (Extended Data Fig. 10d). Volcano plots and GSEA analysis highlighted the change of pathways caused by imatinib, related to fatty acid biosynthesis, steroid biosynthesis and fatty acid metabolism (Extended Data Fig. 10e, f). Viruses have been known to target lipid signalling, synthesis and metabolism to remodel their host cells into an optimal environment for replication.

Discussion

Here, we present an hPSC-LO platform, including AT2-like cells that express ACE2, the receptor for SARS-CoV-2. RNA-seq analysis of infected organoids revealed upregulation of cytokine/chemokine signalling with only a modest interferon signature, which mimics the inflammatory changes observed in primary human COVID-19 pulmonary infections[18]. We also showed that multiple cell types in hPSC-COs can be infected by SARS-CoV-2. Finally, we used the hPSC-LOs in a high-throughput screen of FDA-approved drugs. We identified several drugs that inhibit SARS-CoV-2 entry, including imatinib, MPA and QNHC, both in vitro and in vivo. The anti-viral activities of these drugs were further validated with live SARS-CoV-2. MPA is widely and safely used as an immunosuppressive drug to prevent organ rejection and to treat autoimmune diseases. A recent study predicted that MPA modulates the interaction between inosine-5′-monophosphate dehydrogenase 2 and SARS-CoV-2 non-structural protein 14 (ref. [21]). Imatinib has been shown to be a potent inhibitor of SARS-CoV and Middle East respiratory syndrome coronavirus fusion proteins. Very recently, four clinical trials have been registered to apply imatinib to treat patients with COVID-19. In conclusion, we established hPSC-derived LO and CO models that can be applied to screen for drug candidates for patients with COVID-19.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2901-9.

1. Pan, L. et al. Clinical characteristics of COVID-19 patients with digestive symptoms in Hubei, China: a descriptive, cross-sectional, multicenter study. *Am. J. Gastroenterol*. **115**, 766–773 (2020).
2. Lamers, M. M. et al. SARS-CoV-2 productively infects human gut enterocytes. *Science* **369**, 50–54 (2020).
3. Zhou, J. et al. Infection of bat and human intestinal organoids by SARS-CoV-2. *Nat. Med*. **26**, 1077–1083 (2020).
4. Monteil, V. et al. Inhibition of SARS-CoV-2 infections in engineered human tissues using clinical-grade soluble human ACE2. *Cell* **181**, 905–913 (2020).
5. Chen, Y. W. et al. A three-dimensional model of human lung development and disease from pluripotent stem cells. *Nat. Cell Biol*. **19**, 542–549 (2017).
6. Huang, S. X. et al. Efficient generation of lung and airway epithelial cells from human pluripotent stem cells. *Nat. Biotechnol*. **32**, 84–91 (2014).
7. Jacob, A. et al. Differentiation of human pluripotent stem cells into functional lung alveolar epithelial cells. *Cell Stem Cell* **21**, 472–488 (2017).
8. Mou, H. et al. Generation of multipotent lung and airway progenitors from mouse ESCs and patient-specific cystic fibrosis iPSCs. *Cell Stem Cell* **10**, 385–397 (2012).
9. McCauley, K. B. et al. Efficient derivation of functional human airway epithelium from pluripotent stem cells via temporal regulation of Wnt signaling. *Cell Stem Cell* **20**, 844–857 (2017).
10. Hurley, K. et al. Reconstructed single-cell fate trajectories define lineage plasticity windows during differentiation of human PSC-derived distal lung progenitors. *Cell Stem Cell* **26**, 593–608 (2020).
11. Jacob, A. et al. Derivation of self-renewing lung alveolar epithelial type II cells from human pluripotent stem cells. *Nat. Protoc*. **14**, 3303–3332 (2019).
12. Dye, B. R. et al. In vitro generation of human pluripotent stem cell derived lung organoids. *eLife* **4**, e05098 (2015).
13. Chen, H. J. et al. Generation of pulmonary neuroendocrine cells and SCLC-like tumors from human embryonic stem cells. *J. Exp. Med*. **216**, 674–687 (2019).
14. 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).
15. Shang, J. et al. Cell entry mechanisms of SARS-CoV-2. *Proc. Natl Acad. Sci. USA* **117**, 11727–11734 (2020).
16. Whitt, M. A. Generation of VSV pseudotypes using recombinant ΔG-VSV for studies on virus entry, identification of entry inhibitors, and immune responses to vaccines. *J. Virol. Methods* **169**, 365–374 (2010).
17. Nie, J. et al. Establishment and validation of a pseudovirus neutralization assay for SARS-CoV-2. *Emerg. Microbes Infect*. **9**, 680–686 (2020).
18. Blanco-Melo, D. et al. Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell* **181**, 1036–1045 (2020).
19. Crespo, M. et al. Colonic organoids derived from human induced pluripotent stem cells for modeling colorectal cancer and drug testing. *Nat. Med*. **23**, 878–884 (2017).
20. Munera, J. O. et al. Differentiation of human pluripotent stem cells into colonic organoids via transient activation of BMP signaling. *Cell Stem Cell* **21**, 51–64 (2017).
21. Gordon, D. E. et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020).

Methods

hPSC culture

RUES2 human embryonic stem (ES) cells (provided by WiCell) were cultured on irradiated mouse embryonic fibroblasts (Global Stem, catalogue no. GSC-6001G) at a density of 20,000–25,000 cells cm^{-2} in Dulbecco's modified Eagle medium (DMEM)/F12, 20% knockout serum replacement (Life Technologies), 0.1 mM β-mercaptoethanol (Sigma Aldrich) and 20 ng ml^{-1} bFGF (R&D Systems), and the medium was changed daily. Human ES cell cultures were maintained in an undifferentiated state at 37 °C in a 5% CO_2/air environment until stem cells reached about 90% confluence. H1 human ES cells (provided by WiCell) and HUES8 human ES cells (provided by Harvard University) were grown and maintained on 1% Matrigel (Corning)-coated six-well plates in StemFlex medium (Gibco) at 37 °C with 5% CO_2. All human ES cell lines were authenticated by short tandem repeat profiling and tested for mycoplasma contamination every six months.

hPSC lung differentiation

Protocols for the maintenance of hPSCs and the generation of lung organoids were slightly modified from previous studies[6,13]. hPSC differentiation into endoderm was performed in serum-free differentiation (SFD) medium of DMEM/F12 (3:1) (Life Technologies) supplemented with 1× N2 (Life Technologies), 1× B27 (Life Technologies), 50 μg ml^{-1} ascorbic acid, 2 mM Glutamax (Gibco), 0.4 μM monothioglycerol and 0.05% BSA at 37 °C in a 5% CO_2/5% O_2/95% N_2 environment. hPSCs were treated with Accutase and plated onto low-attachment 6-well plates (Corning), and then resuspended in endoderm induction medium containing 10 μM Y-27632, 0.5 ng ml^{-1} human BMP4 (R&D Systems), 2.5 ng ml^{-1} human bFGF and 100 ng ml^{-1} human activin A (R&D Systems), for 72–76 h dependent on the formation rates of endoderm cells. On day 3 or 3.5, the endoderm bodies were dissociated into single cells using 0.05% trypsin/0.02% EDTA and plated onto fibronectin-coated, 24-well tissue culture plates (about 100,000–150,000 cells per well). For induction of anterior foregut endoderm, the endoderm cells were cultured in SFD medium supplemented with 1.5 μM dorsomorphin dihydrochloride (R&D Systems) and 10 μM SB431542 (R&D Systems) for 36 h, and then switched for 36 h to 10 μM SB431542 and 1 μM IWP2 (R&D Systems) treatment. For induction of early stage lung progenitor cells (day 6–15), the resulting anterior foregut endoderm was treated with 3 μM CHIR99021 (CHIR, Stem-RD), 10 ng ml^{-1} human FGF10, 10 ng ml^{-1} human KGF, 10 ng ml^{-1} human BMP4 and 50–60 nM all-*trans* retinoic acid (ATRA), in SFD medium on day 8–10. The day 10–15 culture was maintained in a 5% CO_2/air environment. On days 15 and 16, the lung field progenitor cells were replated after 1-min trypsinization onto fibronectin-coated plates, in the presence of SFD containing 3 μM CHIR99021, 10 ng ml^{-1} human FGF10, 10 ng ml^{-1} human FGF7, 10 ng ml^{-1} human BMP4 and 50 nM all-*trans* retinoic acid. Day 16–25 cultures of late-stage lung progenitor cells were maintained in SFD medium containing 3 μM CHIR99021, 10 ng ml^{-1} human FGF10 and 10 ng ml^{-1} human KGF, in a 5% CO_2/air environment. For differentiation of mature lung cells (day 25 to 55) in three-dimensional culture, cells were replated and embedded in 90% Matrigel after brief trypsinization in SFD medium containing 3 μM CHIR99021, 10 ng ml^{-1} human FGF10, 10 ng ml^{-1} human KGF, 50 nM dexamethasone, 0.1 mM 8-bromo-cAMP (Sigma Aldrich) and 0.1 mM IBMX (3,7-dihydro-1-methyl-3-(2-methylpropyl)-1H-purine-2,6-dione; Sigma Aldrich).

hPSC colonic lineage differentiation

For definitive endoderm differentiation, hPSCs were cultured to achieve 80–90% confluency, and treated with 3 μM CHIR99021 and 100 ng ml^{-1} activin A in basal medium RPMI1640 (Cellgro) supplemented with 1× penicillin–streptomycin (Gibco) for 1 day, and changed to the basal medium containing 100 ng ml^{-1} activin A the next day. To induce $CDX2^+$ hindgut endoderm, definitive endoderm cells were treated with 3 μM CHIR99021 and 500 ng ml^{-1} FGF4 (Peprotech) in RPMI1640 supplemented with 1× B27 supplement and 1× Pen-Strep for 4 days with daily changes to fresh medium. The hindgut endoderm was then subjected to colonic lineage induction by treatment with 100 ng ml^{-1} BMP2 (Peprotech), 3 μM CHIR99021 and 100 ng ml^{-1} hEGF (Peprotech) in Advance DMEM F12 medium supplemented with 1× B27 supplement, 1× GlutaMax, 10 mM HEPES (Gibco) and 1× Pen-Strep for 3 days with daily changes to fresh medium. The colon spheroids were collected from the initial two-dimensional cultures and embedded in a 100% Matrigel dome in a 24-well plate. Differentiation to mature colonic cell types was achieved by culturing these colon spheroids in differentiation medium containing 600 nM LDN193189 (Axon), 3 μM CHIR99021 and 100 ng ml^{-1} hEGF in Advance DMEM F12 medium supplemented with 1× B27 supplement, 1× GlutaMax, 10 mM HEPES and 1× Pen-Strep. The differentiation medium was refreshed every 3 days for at least 40 days to achieve full colonic differentiation. The hPSC-COs were passaged and expanded every 10–14 days at 1:6 density. To passage the organoids, the Matrigel domes containing the organoids were scraped off the plate and resuspended in cold splitting medium (Advance DMEM F12 medium supplemented with 1× GlutaMax, 10 mM HEPES and 1× Pen-Strep). The organoids were mechanically dislodged from the Matrigel dome and fragmented by pipetting in cold splitting medium. The old Matrigel and splitting medium were removed after pelleting cells and the organoids were resuspended in 100% Matrigel. A 50-μl volume of Matrigel containing fragmentized colon organoids was plated in 1 well of a pre-warmed 24-well plate.

Cell lines

HEK293T (human (*Homo sapiens*) fetal kidney) and Vero E6 (African green monkey (*Chlorocebus aethiops*) kidney) cells were obtained from ATCC. Cells were cultured in DMEM supplemented with 10% FBS, 100 IU ml^{-1} penicillin and 100 μg ml^{-1} streptomycin. All cell lines were incubated at 37 °C with 5% CO_2.

SARS-CoV-2-entry viruses

Recombinant Indiana VSV for the expression of SARS-CoV-1 or SARS-CoV-2 spike proteins was generated as previously described[22]. HEK293T cells were grown to 80% confluency before transfection with pCMV3-SARS-CoV-1-spike and pCMV3-SARS-CoV-2-spike (provided by P. Wang) using FuGENE 6 (Promega). Cells were cultured overnight at 37 °C with 5% CO_2. The next day, medium was removed and VSV-G pseudotyped ΔG-luciferase (G*ΔG-luciferase, Kerafast) was used to infect the cells in DMEM at an MOI of 3 for 1 h before washing the cells with 1× Dulbecco's phosphate-buffered saline (PBS) three times. DMEM supplemented with anti-VSV-G (I1, mouse hybridoma supernatant from CRL-2700; ATCC) was added to the infected cells and they were cultured overnight as previously described[23]. The next day, the supernatant was collected and clarified by centrifugation at 300*g* for 10 min and aliquots were stored at −80 °C.

hPSC-LOs or hPSC-COs were seeded in 24-well plates, SARS-CoV-2-entry virus was added at the indicated MOIs and the plates were centrifuged at 1,200*g* for 1 h. Then, the organoids were cultured at 37 °C with 5% CO_2. At 24 hpi, organoids were fixed for immunohistochemistry or collected for luciferase assay following the Luciferase Assay System protocol (E1501, Promega)

SARS-CoV-2 virus infections

SARS-CoV-2, isolate USA-WA1/2020 (NR-52281), was deposited by the Center for Disease Control and Prevention and obtained through BEI Resources, NIAID, NIH. SARS-CoV-2 was propagated in Vero E6 cells in DMEM supplemented with 2% FBS, 4.5 g l^{-1} D-glucose, 4 mM L-glutamine, 10 mM non-essential amino acids, 1 mM sodium pyruvate and 10 mM HEPES as previously described[18].

All work involving live SARS-CoV-2 was performed in the CDC/USDA-approved BSL-3 facility of the Global Health and Emerging

Pathogens Institute at the Icahn School of Medicine at Mount Sinai in accordance with institutional biosafety requirements.

hPSC-LOs or hPSC-COs were infected with SARS-CoV-2 at the indicated MOI and incubated for 24 h at 37 °C. hPSC-LOs or hPSC-COs were pretreated with DMSO, imatinib, MPA or QNHC at the indicated concentration for 4 h before infection. At 24 hpi, hPSC-LOs or hPSC-COs were washed three times in PBS and lysed in TRIzol for RNA analysis or fixed for 24 h in 5% formaldehyde for immunofluorescence staining.

Approximately 2.5×10^5 Vero E6 cells were treated with DMSO, imatinib, MPA or QNHC at the indicated concentration, followed by infection with SARS-CoV-2 at an MOI of 0.01 in DMEM supplemented with 2% FBS, 4.5 g l^{-1} D-glucose, 4 mM L-glutamine, 10 mM non-essential amino acids, 1 mM sodium pyruvate and 10 mM HEPES. At 24 hpi, cells were washed three times in PBS and lysed in TRIzol for RNA analysis, lysed in radioimmunoprecipitation assay (RIPA) buffer for protein analysis or fixed for 24 h in 5% formaldehyde for immunofluorescence staining.

Xenograft formation

One million hPSC-derived cells at the lung progenitor stage (at day 25) were subcutaneously injected into male and female 6- to 8-week-old *NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ* (NSG) mice (Jackson Laboratory). The mice were maintained at 22 °C with 40–60% humidity and a 12 light/12 dark cycle. Four months later, the mice were used for drug evaluation. The mice were randomly separated into different groups for different treatments. The researchers analysing the data were blind to treatment.

To determine drug activity in vivo, the mice were retreated with 400 mg kg^{-1} imatinib mesylate, 50 mg kg^{-1} MPA and 25 mg kg^{-1} QNHC in 10% DMSO/90% corn oil by intraperitoneal injection. Four hours after drug administration, SARS-CoV-2-entry virus was inoculated directly to the xenograft at 1×10^3 FFUs. At 24 hpi, the mice were euthanized and used for immunohistochemistry analysis.

All animal work was conducted in agreement with NIH guidelines and approved by the Weill Cornell Medicine Institutional Animal Care and Use Committee and the Institutional Biosafety Committee.

Immunohistochemistry

Histology on tissues from mice was performed on paraffin-embedded or frozen sections from xenografts. Tissues were fixed overnight in 10% buffered formalin and transferred to 70% ethanol, followed by paraffin embedding, or tissues were fixed in 10% buffered formalin and transferred to 30% sucrose, followed by snap-freezing in O.C.T (Fisher Scientific). Adjacent sections stained with haematoxylin and eosin were used for comparison. Living cells in culture were directly fixed in 4% paraformaldehyde for 25 min, followed by 15-min permeabilization in 0.1% Triton X-100. Formalin-fixed paraffin-embedded human colon tissue for ACE2 staining was purchased from VitroVivo Biotech (SKU no. HuPS-07005A). For immunofluorescence, cells or tissue sections were immunostained with primary antibodies at 4 °C overnight and secondary antibodies at room temperature for 1 h. The information for the primary and secondary antibodies is provided in Supplementary Table 3. Nuclei were counterstained by DAPI. The figures were processed using Adobe Illustrator CC2017.

Intracellular flow cytometry analysis

Flow cytometry intracellular staining was performed following the instruction manual for the Fixation/Permeabilization Solution Kit (BD Biosciences). In brief, cells were resuspended in Fixation/Permeabilization solution at 4 °C for 20 min, washed twice in 1× Perm/Wash buffer, incubated with primary antibody at 4 °C for 30 min in the dark, washed twice and then incubated with secondary antibody at 4 °C for 30 min in the dark. The cells were then washed twice before flow cytometry analysis. The information for the primary antibodies and antibodies is provided in Supplementary Table 3.

Western blot

Protein was extracted from cells in RIPA lysis buffer containing 1× Complete Protease Inhibitor Cocktail (Roche) and 1× phenylmethylsulfonyl fluoride (Sigma Aldrich) before safe removal from the BSL-3 facility. Samples were analysed by SDS–polyacrylamide gel electrophoresis and transferred onto nitrocellulose membranes. Proteins were detected using rabbit polyclonal anti-GAPDH (Sigma Aldrich, G9545, 1:1,000), mouse monoclonal anti-SARS-CoV-2 nucleocapsid (1C7; 1:1,000) and mouse monoclonal anti-SARS-CoV-2 spike (2B3E5) protein (1:1,000) (a gift from T. Moran). Endogenous TMPRSS2 and FURIN were detected using TMPRSS2 antibody (H-4; Santa Cruz, sc-515727, 1:500) and anti-FURIN (Abcam, ab183495, 1:1,000). Primary antibodies were detected using fluorophore-conjugated secondary goat anti-mouse (IRDye 680RD, 926-68070, 1:25,000) and goat anti-rabbit (IRDye 800CW, 926-32211, 1:25,000) antibodies. Antibody-mediated fluorescence was detected on a LI-COR Odyssey CLx imaging system and analysed using Image Studio software (LI-COR).

Surface plasmon resonance

The binding of imatinib, MPA and QHC to human ACE2 protein was detected by a Biacore T200 surface plasmon resonance system (Cytiva). All experiments were performed at 25 °C in HBS-EP+ buffer (10 mM HEPES, pH 7.4; 150 mM NaCl; 3.4 mM EDTA; and 0.005% (v/v) surfactant P20). The human ACE2 protein, diluted at 50 μg ml^{-1} in 10 mM sodium acetate, pH 4.5, was immobilized on the surface of a CM5 sensor chip using the amine coupling method and served as the active surface. A blank immobilized flow cell was used as the reference surface. Then 1.5-fold dilutions of compound concentrations from 30 to 1.17 μM were injected over the reference and active surfaces, and the surface was regenerated after each concentration using 10 mM glycine, pH 1.5. Background binding to reference flow cells was subtracted and compound binding levels to ACE2 were calculated using Biacore T200 evaluation software.

qRT–PCR

Total RNA samples were prepared from cells/organoids using TRIzol and the Direct-zol RNA Miniprep Plus kit (Zymo Research) according to the manufacturer's instructions. To quantify viral replication, measured by the accumulation of subgenomic N transcripts, one-step quantitative real-time PCR was performed using the SuperScript III Platinum SYBR Green One-Step qRT–PCR Kit (Invitrogen) with primers specific for the TRS-L and TRS-B sites for the *N* gene as well as *ACTB* as an internal reference as previously described[24]. Quantitative real-time PCR reactions were performed on a LightCycler 480 Instrument II (Roche). The delta-delta-cycle threshold ($\Delta\Delta C_T$) was determined relative to *ACTB* and mock-infected/treated samples. Error bars indicate the standard deviation of the mean from three biological replicates. The sequences of primers/probes are provided in Supplementary Table 4.

Single-cell colonic organoid preparation for scRNA-seq

hPSC-COs cultured in Matrigel domes were dissociated into single cells using 0.25% trypsin (Gibco) at 37 °C for 10 min, and the trypsin was then neutralized with DMEM F12 supplemented with 10% FBS. The dissociated organoids were pelleted and resuspended with L15 medium (Gibco) supplemented with 10 mM HEPES, and 10 ng ml^{-1} DNaseI (Sigma). The resuspended organoids were then placed through a 40-μm filter to obtain a single-cell suspension, and stained with DAPI followed by sorting of live cells using an ARIA II flow cytometer (BD Biosciences). The live colonic single-cell suspension was proceeded with the Chromium Single Cell 3′ Reagent Kit v3 (10x Genomics, product code 1000075) using a 10x Genomics Chromium Controller. A total of 10,000 cells were loaded into each channel of the Single-Cell A Chip to target 8,000 cells. In brief, according to the manufacturer's instructions, the sorted cells were washed with 1× PBS + 0.04% BSA, counted

by a Bio-Rad TC20 Cell Counter, and the cell viability was assessed and visualized. A total of 10,000 cells and Master Mixes were loaded into each channel of the cartridge to generate the droplets on the Chromium Controller. Beads-in-emulsion (GEMs) were transferred and GEMs-RT was undertaken in droplets by PCR incubation. GEMs were then broken and pooled fractions were recovered. After purification of the first-strand complementary DNA from the post GEM-RT reaction mixture, barcoded, full-length cDNA was amplified via PCR to generate sufficient mass for library construction. Enzymatic fragmentation and size selection were used to optimize the cDNA amplicon size. TruSeq Read 1 (read 1 primer sequence) was added to the molecules during GEM incubation. P5, P7, a sample index and TruSeq Read 2 (read 2 primer sequence) were added via end repair, A-tailing, adaptor ligation and PCR. The final libraries were assessed by an Agilent Technology 2100 Bioanalyzer and sequenced on an Illumina NovaSeq sequencer with a pair-end 100 cycle kit (28+8+91).

Sequencing and gene expression unique molecular identifier count matrix generation

The FASTQ files were imported to a 10x Cell Ranger data analysis pipeline (v.3.0.2) to align reads, generate feature–barcode matrices and perform clustering and gene expression analysis. In a first step, cellranger mkfastq demultiplexed samples and generated .fastq files; and in the second step, cellranger count aligned fastq files to the reference genome and extracted gene expression unique molecular identifier (UMI) count matrix. To measure viral gene expression in colonic organoids, we built a custom reference genome by integrating the four virus genes and luciferase into the 10x pre-built human reference (GRCh38 v.3.0.0) using cellranger mkref. The sequences of four viral genes (*VSV-N*, *VSV-P*, *VSV-M* and *VSV-L*) were retrieved from NCBI (https://www.ncbi.nlm.nih.gov/nuccore/335873), and the sequence of the luciferase was retrieved from HIV-luciferase. For lung organoids, we aligned reads to the 10x pre-built human reference (GRCh38 v3.0.0).

scRNA-seq data analysis for hPSC-COs

We filtered cells with fewer than 300 or more than 8,000 genes detected as well as cells with mitochondrial gene content greater than 30%, and used the remaining cells (6,175 cells for the uninfected sample and 2,962 cells for the infected sample) for downstream analysis. We normalized the gene expression UMI counts for each sample separately using a deconvolution strategy[25] implemented by the R scran package (v.1.14.1). In particular, we pre-clustered cells in each sample using the quickCluster function; we computed the size factor per cell within each cluster and rescaled the size factors by normalization between clusters using the computeSumFactors function; and we normalized the UMI counts per cell by the size factors and took a logarithm transform using the normalize function. We further normalized the UMI counts across samples using the multiBatchNorm function in the R batchelor package (v.1.2.1). We identified highly variable genes using the FindVariableFeatures function in the R Seurat (v.3.1.0)[26], and selected the top 3,000 variable genes after excluding mitochondria genes, ribosomal genes and dissociation-related genes. The list of dissociation-related genes was originally built on mouse data[27]; we converted them to human orthologue genes using Ensembl BioMart. We aligned the two samples on the basis on their mutual nearest neighbours (MNNs) using the fastMNN function in the R batchelor package; this was achieved by performing PCA on the highly variable genes and then correcting the principal components (PCs) according to their MNNs. We selected the corrected top 50 PCs for downstream visualization and clustering analysis. We ran the UMAP dimensional reduction using the RunUMAP function in the R Seurat[26] package with training epochs setting to 2,000. We clustered cells into eight clusters by constructing a shared-nearest-neighbour graph and then grouping cells of similar transcriptome profiles using the FindNeighbours function and the FindClusters function (resolution set to 0.2) in the R Seurat package. We identified marker genes for each cluster by performing differential expression analysis between cells inside and outside that cluster using the FindMarkers function in the R Seurat package. After reviewing the clusters, we merged four clusters that were probably from the stem cell population into a single cluster ($LGR5^+$ or $BMI1^+$ stem cells) and kept the other four clusters ($KRT20^+$ epithelial cells, $MUC2^+$ goblet cells, $EPHB2^+$ transit-amplifying cells and $CHGA^+$ neuroendocrine cells) for further analysis. We reidentified marker genes for the merged five clusters and selected the top ten positive marker genes per cluster for the heatmap plot using the DoHeatmap function in the R Seurat package[26].

scRNA-seq data analysis for lung organoids

For downstream analysis, we filtered cells with fewer than 200 or more than 6,000 unique molecular identifiers (UMIs), and genes detected (UMI count > 0) in fewer than two cells were removed. In addition, cells were excluded if more than 30% of sequences mapped to mitochondrial genes. In total, 14,263 cells passed these filters for quality. Following the R Seurat package suggestions, we normalized the gene expression using the NormalizeData function, and the differentially expressed genes were found by the vst method using the FindVariableFeatures function. The top 3,000 variable genes were selected for PCA analysis. We used an elbow plot to determine the number of PCs, and the top 20 PCs were used for each group of cells. Next, we scaled the data and performed linear dimensional reduction using the ScaleData and RunPCA functions, respectively. The clustering resolution was set to 0.2 using the FindClusters function. We determined the biomarkers in each cluster using the FindAllMarkers function and renamed the nine clusters as the following cell types: AT1-like cells_1, AT2-like cells, fibroblast cells_1, AT1-like cells_2, stromal cells, proliferating cells, fibroblast cells_2, PNECs and AECs. We reidentified marker genes and selected top differentially expressed marker genes per cluster for the heatmap plot using the DoHeatmap function in the R Seurat package. UMAP plots and violin plots were generated by the Seurat toolkit FeaturePlot and VlnPlot functions.

Gene enrichment and correlation

For gene enrichment analysis, we compared the enriched genes in each cluster in hPSC-LOs to 205 marker genes reported to be expressed at high levels in AT2 cells in the human lung cell dataset[28]. A bar graph was generated using the ratio of overlapping genes in each cluster. For correlation analysis, we compared the marker genes in hPSC-LOs in each cluster (cluster 0. AT1-like cells_1 (171 genes), cluster 1. AT2-like cells (144 genes), cluster 2. fibroblast cells_1 (323 genes), cluster 3. AT1-like cells_2 (288 genes), cluster 4. stromal cells (261 genes), cluster 5. proliferating cells (332 genes), cluster 6. fibroblast cells_2 (157 genes), cluster 7. PNECs (313 genes) and cluster 8. AECs (206 genes)) with the human lung cells (AT1 cells (1,087 genes), AT2 cells (205 genes), proliferating basal cells (984 genes), proximal basal cells (630 genes), alveolar fibroblasts (423 genes), myofibroblasts (290 genes) and PNECs (1,585 genes)). The heatmap plot on correlation of genes with cell fate was generated using the R heatmap.2 package.

RNA-seq before and following viral infections

Organoid infections were performed at the indicated MOI and cells were collected at 24 hpi in DMEM supplemented with 0.3% BSA, 4.5 g l^{-1} D-glucose, 4 mM L-glutamine and 1 μg ml^{-1} TPCK trypsin. Total RNA was extracted and DNase-treated using TRIzol (Invitrogen) and the Directzol RNA Miniprep Plus kit (Zymo Research) according to the manufacturer's instructions. RNA-seq libraries of polyadenylated RNA were prepared using TruSeq RNA Library Prep Kit v2 (Illumina) or the TruSeq Stranded mRNA Library Prep Kit (Illumina) according to the manufacturer's instructions. cDNA libraries were sequenced using an Illumina NextSeq 500 platform. For viral RNA analysis, sequencing

reads were aligned to the SARS-CoV-2/human/USA/WA-CDC-WA1/2020 genome (GenBank: MN985325.1) using Bowtie2 and were visualized using IGV software.

After further filtering and quality control, the R package edgeR[29] was used to calculate reads per kilobase of transcript per million mapped reads (RPKM) and $\log_2$[counts per million] matrices as well as to perform differential expression analysis. PCA was performed using $\log_2$[counts per million] values and gene set analysis was run with WebGestalt[30]. Heatmaps and bar plots were generated using Graphpad Prism software, v.7.0d. In the volcano plots, differentially expressed genes (P-adjusted value < 0.05) with a $\log_2$[fold change] > 1 are indicated in red.

High-throughput chemical screening

hPSC-LOs were dissociated using TrypLE for 10 min in a 37 °C water bath and replated into 10% Matrigel-coated 384-well plates at 10,000 cells per 40 µl medium per well. Six hours after plating, compounds from an in-house FDA-approved drug library (Prestwick) were added at 10 µM. DMSO treatment was used as a negative control. hPSC-LOs were further infected with SARS-CoV-2-entry virus (MOI = 0.01). After 24 hpi, hPSC-LOs were collected for luciferase assay following the Luciferase Assay System protocol (Promega).

To calculate the EC_{50}, the luciferase activity was normalized to the DMSO-treated condition. To calculate the CC_{50}, the cell survival was monitored by Cell-Titer Glo (Promega) and normalized to the DMSO-treated condition. The efficacy and cytotoxicity curves were calculated using Prism GrapdPad 7.0.

Human studies

Tissue samples were provided by the Weill Cornell Medicine Department of Pathology. The Tissue Procurement Facility operates under the Institutional Review Board (IRB)-approved protocol and follows the guidelines set by Health Insurance Portability and Accountability Act. Experiments using samples from human subjects were conducted in accordance with local regulations and with the approval of the IRB at Weill Cornell Medicine. The autopsy samples were collected under IRB protocol 20-04021814. Consent is not required for obtaining autopsy samples as per the IRB protocol. The colonoscopy sample was collected as part of a study on inflammatory bowel disease. Colonoscopy tissue samples were collected, with informed consent, for the diagnosis of inflammatory bowel disease under the IRB protocol 1409015468. As standard clinical practice during the pandemic, patient stool samples were tested for SARS-CoV-2. Residual colonoscopy tissue was analysed for SARS-CoV-2 staining in patients whose stool sample was positive for the virus. No extra tissue was collected for this study.

Tissue specimens were fixed in 10% formalin or 2.5% glutaraldehyde for 48–72 h before the subsequent procedures. Haematoxylin-and-eosin staining was performed on sections from paraffin blocks.

RNA in situ hybridization

Formalin-fixed paraffin-embedded colon sections of 7-µm thickness were prepared according to the manufacturer's instructions (Advanced Cell Diagnostics). The 2.5 HD Reagent Brown Kit was used for detection with probes directed against SARS-CoV-2 (MT020880.1, region 28,274-28,882) using 3,3'-diaminobenzidine (DAB) for chromogenic readout.

Electron microscopy

For electron microscopy examination, after osmium tetroxide post-fixation and gradient dehydration, 'semi-thin' sections were examined, and selected areas were chosen for thin sections. Thin sections were then cut and stained with uranyl acetate and lead citrate. Electron microscopy grids were then viewed with a transmission electron microscope.

Statistics and reproducibility

No statistical methods were used to predetermine sample size. $n = 3$ independent biological replicates were used for all experiments unless otherwise indicated. NS indicates a non-significant difference. The P values were calculated by unpaired two-tailed Student's t-tests unless otherwise indicated. $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$. Each experiment was repeated independently three times with similar results. For Figs. 1f, i, j, l, 3h, j and 4b, e, h, k and Extended Data Figs. 1b, c, 2d, e, 3a, c–e, 5e, h, 6b, 7a–d, 8c, e and 10b, each experiment was repeated independently three times with similar results.

Reporting summary

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

Data availability

RNA-seq data for hPSC-LOs and lung autopsy samples are publicly available in the GEO repository database under the accession number GSE155241. scRNA-seq data for hPSC-LOs are publicly available in the GEO repository database under the accession number GSE148113. scRNA-seq data for hPSC-COs are publicly available on the GEO repository database under the accession number GSE147975. RNA-seq data for hPSC-derived endocrine cells and liver organoids are available on the GEO repository database under the accession number GSE151803. Source data are provided with this paper.

Code availability

Code for processing scRNA-seq data can be directly downloaded from GitHub at https://github.com/shuibingchen/SARS_COV2.

22. Zhao, X. et al. Immunization-elicited broadly protective antibody reveals ebolavirus fusion loop as a site of vulnerability. *Cell* **169**, 891–904 (2017).
23. Liu, L. et al. Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature* **584**, 450–456 (2020).
24. Yang, L. et al. A human pluripotent stem cell-based platform to study SARS-CoV-2 tropism and model virus infection in human cells and organoids. *Cell Stem Cell* **27**, 125–136 (2020).
25. Lun, A. T., Bach, K. & Marioni, J. C. Pooling across cells to normalize single-cell RNA sequencing data with many zero counts. *Genome Biol.* **17**, 75 (2016).
26. Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902 (2019).
27. van den Brink, S. C. et al. Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat. Methods* **14**, 935–936 (2017).
28. Travaglini, K. J. et al. A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* **587**, 619–625 (2020).
29. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
30. Liao, Y., Wang, J., Jaehnig, E. J., Shi, Z. & Zhang, B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.* **47**, W199–W205 (2019).

Acknowledgements This work was supported by the Department of Surgery, Weill Cornell Medicine (T.E., F.C.P. and S.C.), the American Diabetes Association (7-20-COVID-211 to S.C.), NIDDK (R01 DK124463, DP3 DK111907-01, R01 DK116075-01A1 and R01 DK119667-01A1 to S.C.), NCI (R01CA234614), NIAID (2R01AI107301), NIDDK (R01DK121072 and 1RO3DK117252), the Department of Medicine, Weill Cornell Medicine (R.E.S.), NIH (4R00CA226353-02 to H.J.C), the Defense Advanced Research Projects Agency under Cooperative Agreement W911NF-16-C-0050, the Marc Haas Foundation (B.R.tO) and the Jack Ma Foundation (D.D.H.). S.C and R.E.S. are supported as Irma Hirschl Trust Research Award Scholars. We thank H. Varmus at Weill Cornell Medicine for his support and T. Moran, Center for Therapeutic Antibody Discovery at the Icahn School of Medicine at Mount Sinai, for providing anti-SARS-CoV-spike. We thank S. V. Shehan for completing the electron microscopy analysis and B. He for performing the in situ hybridization. We are also grateful for technical support and advice provided by L. Cohen-Gould and R. L. Furler at the Cell Screening Core Facility of Weill Cornell Medicine.

Author contributions S.C., R.E.S., T.E., F.C.P., H.J.C., B.R.tO., D.D.H., L.C.C. and H.W. conceived and designed the experiments. Y. Han, X.D., L.Y., X.T., Y.B., C.R., J.Z., Z.Z., D.-H.T.N., J.L.J., Y.L. and T.M.Y. performed organoid differentiation, in vivo transplantation, SARS-CoV-2 virus infection and drug screening. J.J. and A.B. collected human tissues. P.W. and Y. Huang performed SARS-CoV-2 pseudo-entry virus-related experiments. B.E.N.-P., S.U. and B.R.tO. performed SARS-CoV-2-related experiments. F.D., T.Z., J.X., D.X., X.W., D.R. and S.H. performed the scRNA-seq and bioinformatics analyses.

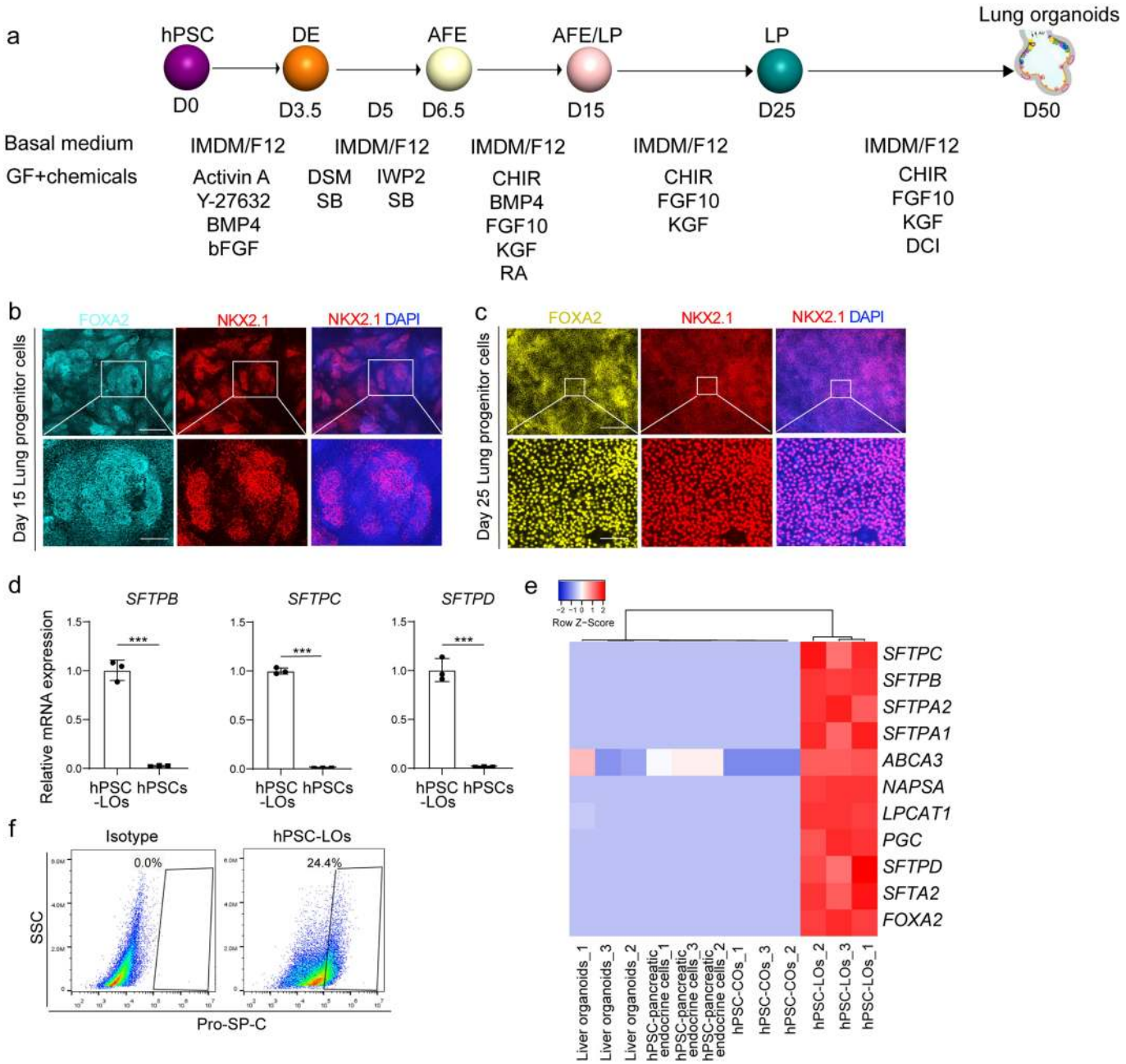
Competing interests R.E.S. is on the scientific advisory board of Miromatrix Inc. and is a consultant and speaker for Alnylam Inc. L.C.C. is a founder and member of the board of directors of Agios Pharmaceuticals and is a founder and receives research support from Petra Pharmaceuticals. L.C.C. is an inventor on patents (pending) for Combination Therapy for PI3K-associated Disease or Disorder, and The Identification of Therapeutic Interventions to Improve Response to PI3K Inhibitors for Cancer Treatment. L.C.C. is a co-founder and shareholder in Faeth Therapeutics. T.M.Y. is a stockholder and on the board of directors of DESTROKE, Inc., an early stage start-up developing mobile technology for automated clinical stroke detection. The other authors declare no competing interests.

Additional information

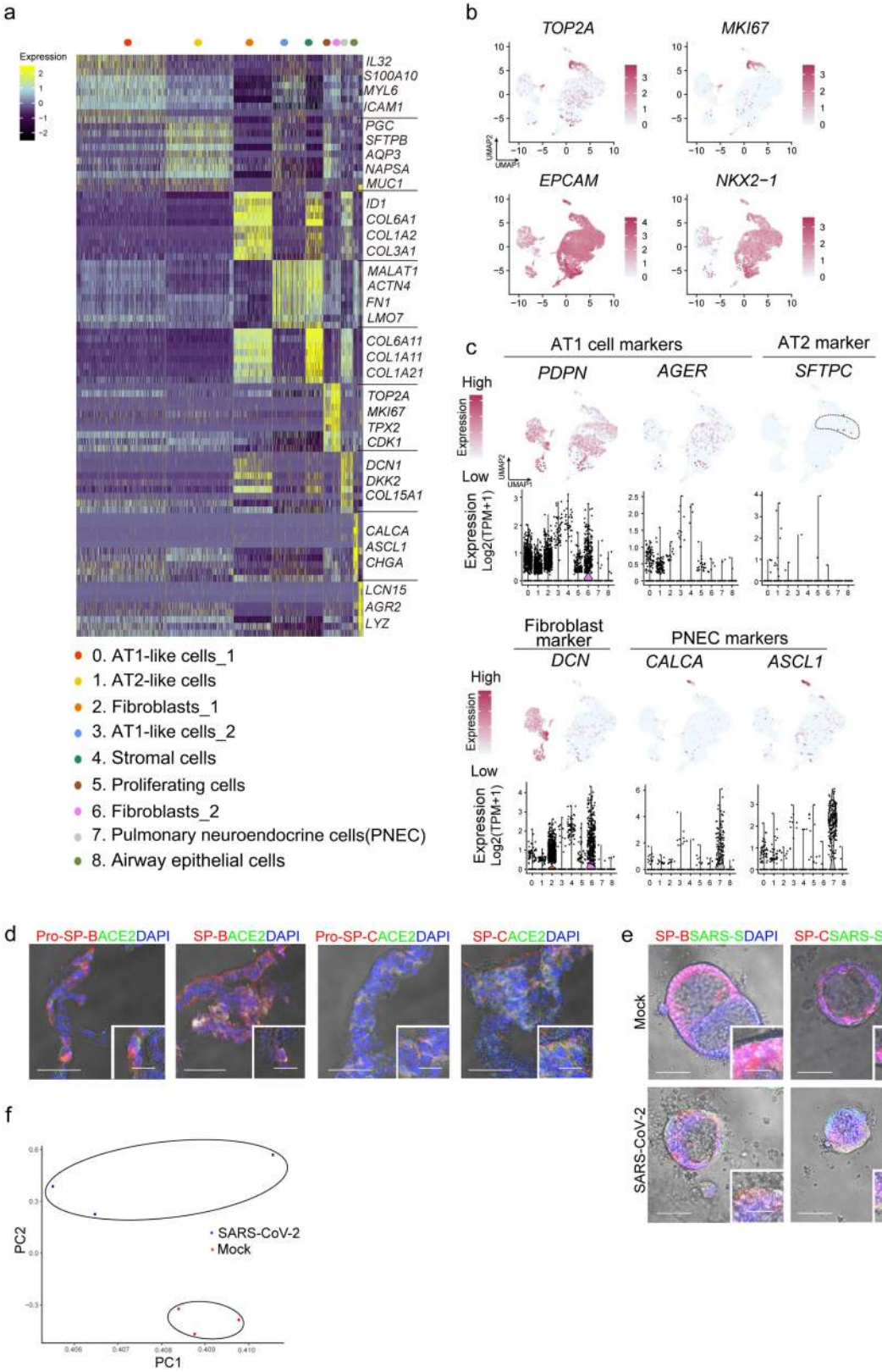
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2901-9.

Correspondence and requests for materials should be addressed to H.W., L.C.C., B.R.t., D.D.H., F.C.P., T.E., H.J.C., R.E.S. or S.C.

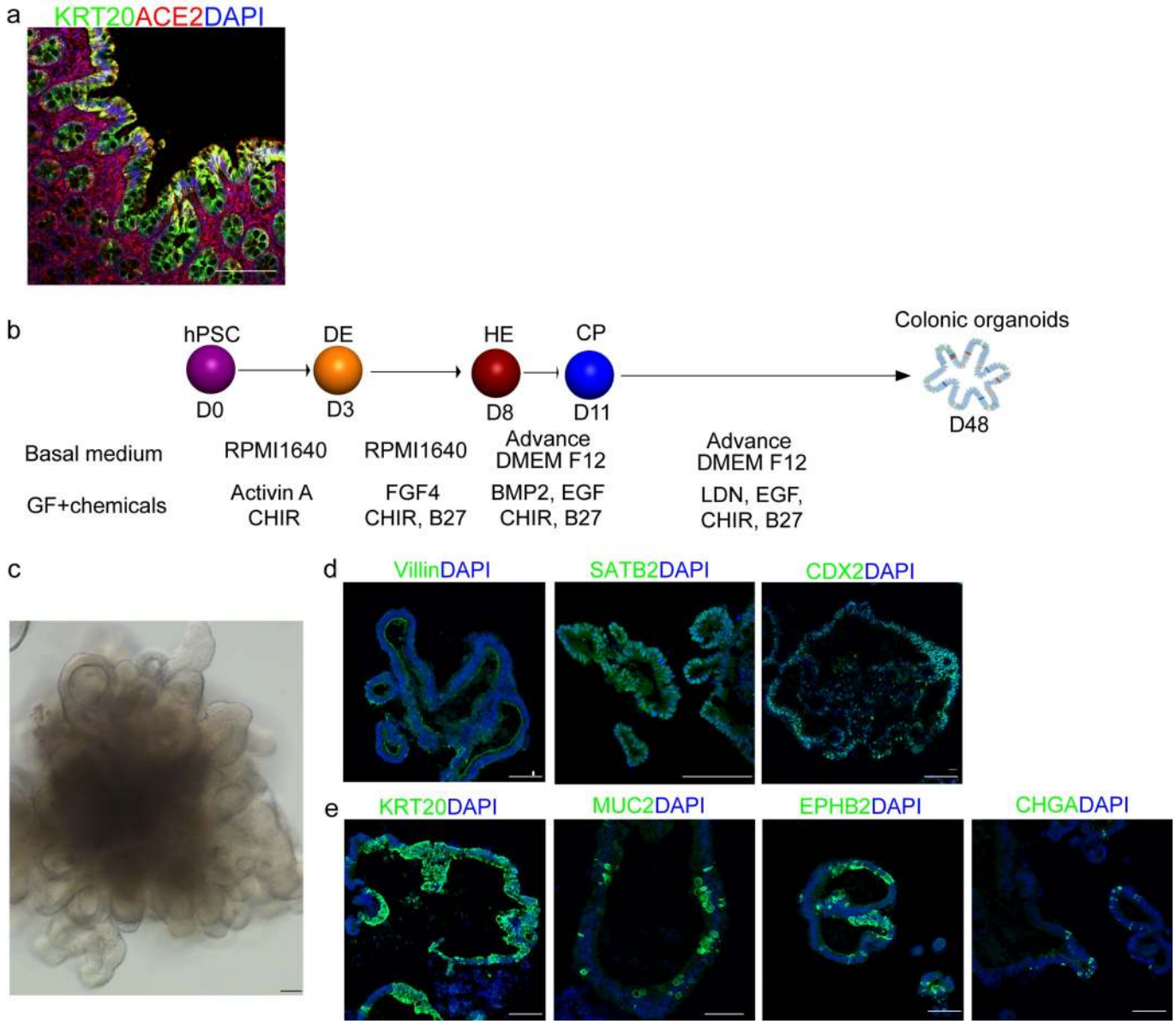
Peer review information *Nature* thanks Yunjeong Kim, Toshiro Sato and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.



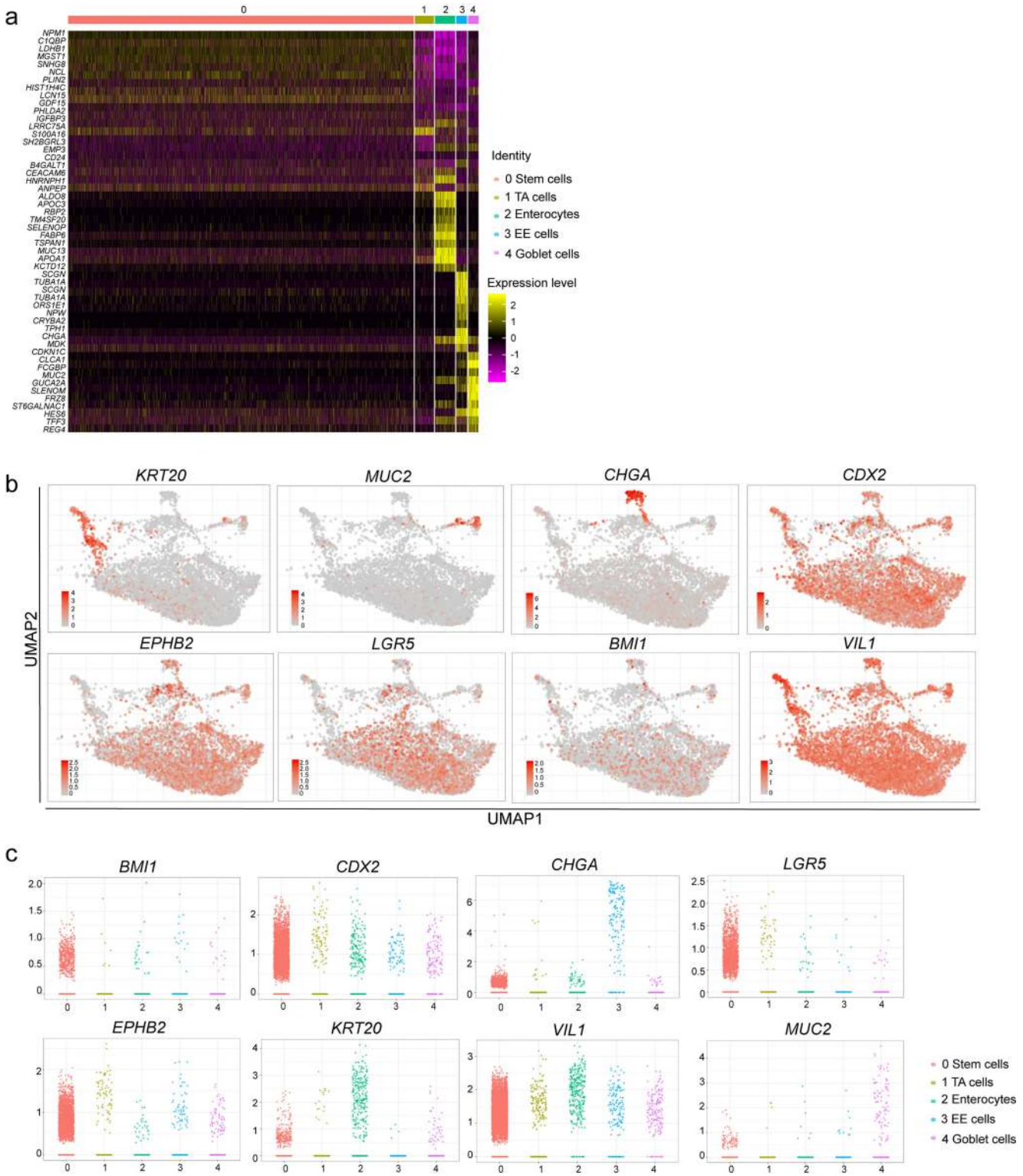
Extended Data Fig. 1 | Characterization of hPSC-LOs. a, Scheme of protocol for differentiation of hPSCs to lung organoids. **b**, **c**, Immunostaining was performed in the hPSC-derived cell cultures at day 15 (b) and day 25 (c). Scale bars = 100 µm. Microscale bars = 20 µm. **d**, qRT–PCR of hPSCs and hPSC-LOs. $n = 3$ biological independent experiments. ***$P = 8.44E-05$, ***$P = 7.05E-07$, ***$P = 0.000130$. **e**, Heatmap from RNA-seq data of AT2 cell markers in hPSC-derived LOs, COs, pancreatic endocrine cells, and liver organoids. **f**, Intra-cellular flow cytometry analysis of pro-SP-C expression in hPSC-LOs. ***$P < 0.001$. Data were analysed by an unpaired two-tailed Student's *t*-test and shown as mean ± STDEV. Data are representative of at least three independent experiments.



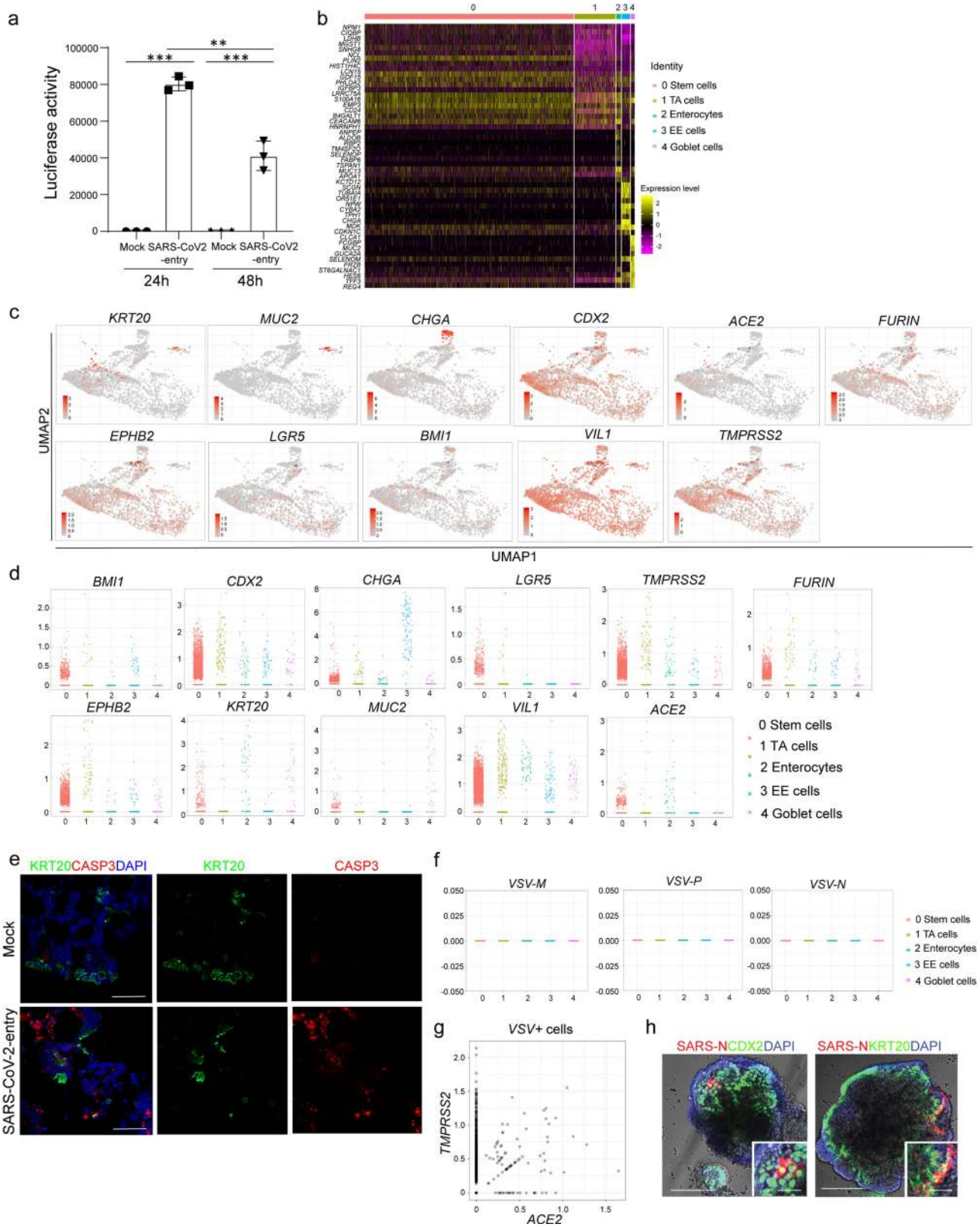
Extended Data Fig. 2 | Single cell RNA-seq of hPSC-LOs. a, Heatmap of enriched genes in each cluster of scRNA profiles in hPSC-LOs. Each row represents one top differentially expressed gene and each column represents a single cell. **b**, UMAP of genes highly expressed in proliferating cells. **c**, Putative AT2, fibroblast and PNEC markers in each cluster in UMAPs. Relative expression level of each marker gene ranges from low (light blue) to high (pink) as indicated. Individual cells positive for lung cell markers are donated by red dots. The violin plot shows the expression level (log2(TPM+1)) of indicated gene in each cluster. **d**, Bright field+immunostaining images of cryo-section of hPSC-LOs. Scale bars = 30 μm. Microscale bars = 10 μm. **e**, Bright field+immunostaining images of SARS-CoV-2 infected hPSC-LOs. Scale bars = 75 μm. Microscale bars = 25 μm. **f**, PCA plot of RNA-seq data from mock-infected or SARS-CoV-2 infected hPSC-LOs.



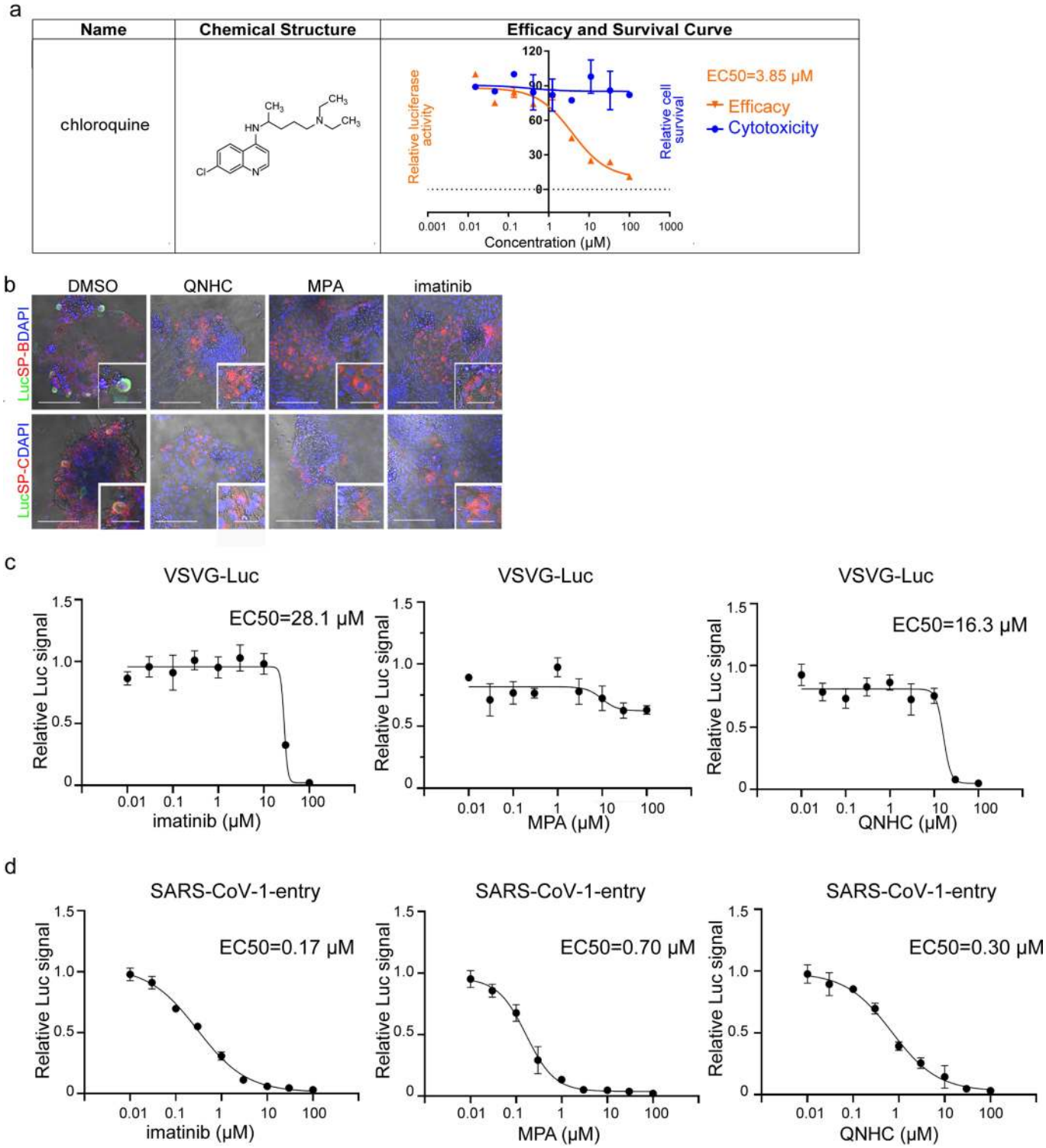
Extended Data Fig. 3 | Directed differentiation of hPSCs to COs. **a**, Immunohistochemistry staining of human colon tissue. Scale bar = 30 μm. **b**, Schematic of protocol and conditions for hPSC differentiation to generate colonic organoids. **c**, Phase contrast image of a representative hPSC-COs. Scale bar = 100 μm. **d**, **e**, Confocal imaging of hPSC-COs stained with antibodies against (**d**) markers for colon cell fate, including Villin, SATB2, CDX2, or (**e**) KRT20, MUC2, EPHB2, and CHGA; Scale bar = 100 μm.



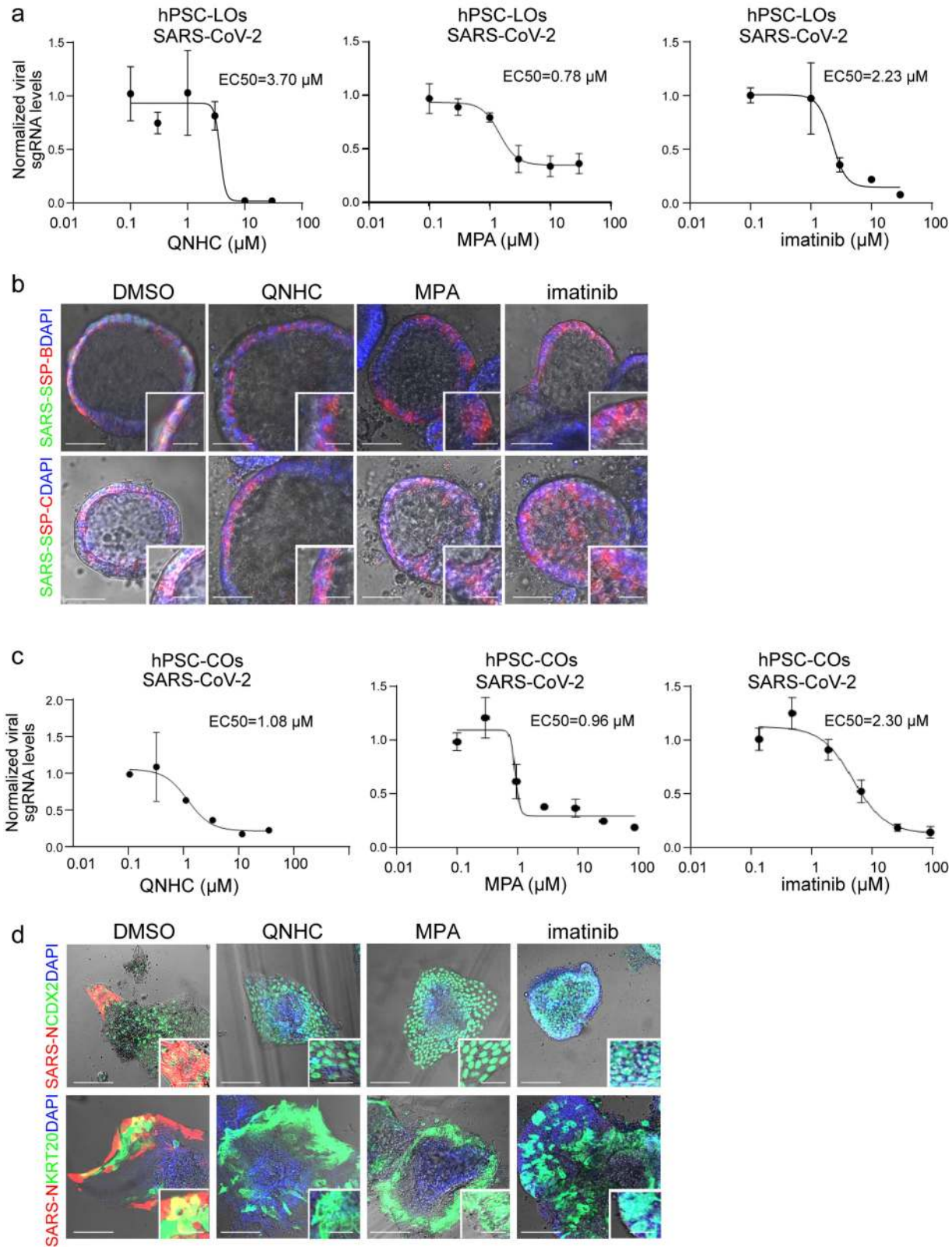
Extended Data Fig. 4 | Single cell RNA-seq analysis of hPSC-COs. a, Heatmap of top 10 differentially expressed genes in each cluster of single cell RNA-seq data. **b**, UMAP of the expression levels of colonic cell markers. **c**, Jitter plots for expression levels of colonic cell markers.



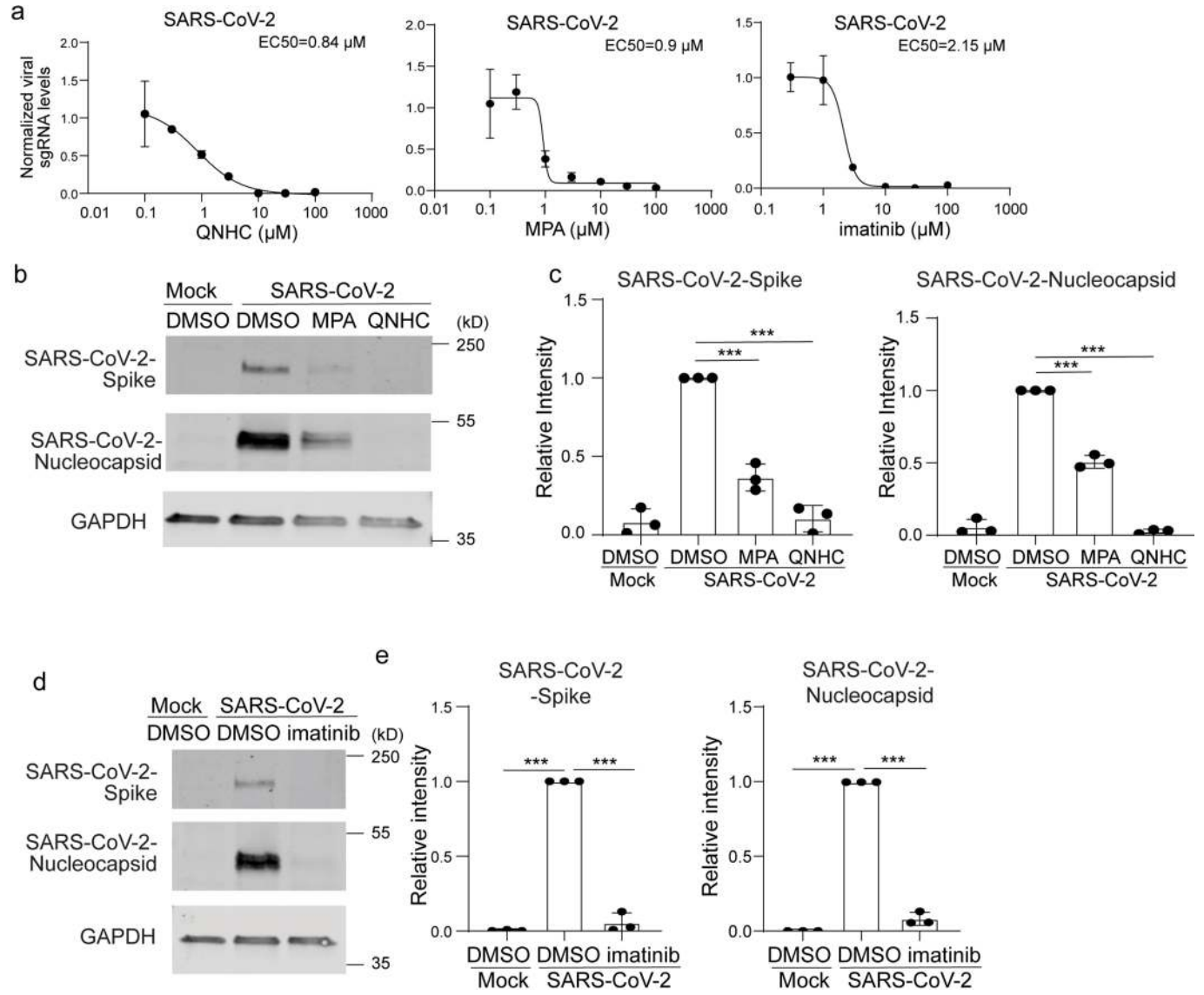
Extended Data Fig. 5 | Single cell RNA-seq analysis of hPSC-COs at 24 hpi with SARS-CoV-2-entry virus. a, Relative luciferase levels in lysates derived from hPSC-COs inoculated with SARS-CoV-2-entry virus at 24 or 48 hpi (MOI = 0.01). $n = 3$ biological independent experiments. ***$P = 4.52E-08$. Data were analysed by ordinary one-way ANOVA and shown as Sidak's multiple comparisons. **b**, Heatmap of top 10 differentially expressed genes in each cluster of single cell RNA-seq data. **c**, UMAP of *ACE2, TMPRSS2, FURIN* and colonic markers. **d**, Jitter plots for transcript levels of *ACE2, TMPRSS2, FURIN* and colonic markers. **e**, Representative immunostaining of infected hPSC-COs co-stained for KRT20 and CASP3. Scale bar = 50 µm. **f**, Jitter plots of transcript levels for *VSV-M, VSV-N* and *VSV-P* from hPSC-COs without SARS-COV-2 infection (mock). **g**, 2D correlation of expression levels for *ACE2* and *TMPRSS2* in VSV$^+$ cells. **h**, Bright field+immunostaining images of SARS-CoV-2 infected hPSC-COs. Scale bars = 100 µm. Microcale bars = 40 µm. **$P < 0.01$, ***$P < 0.001$. Data are representative of at least three independent experiments.



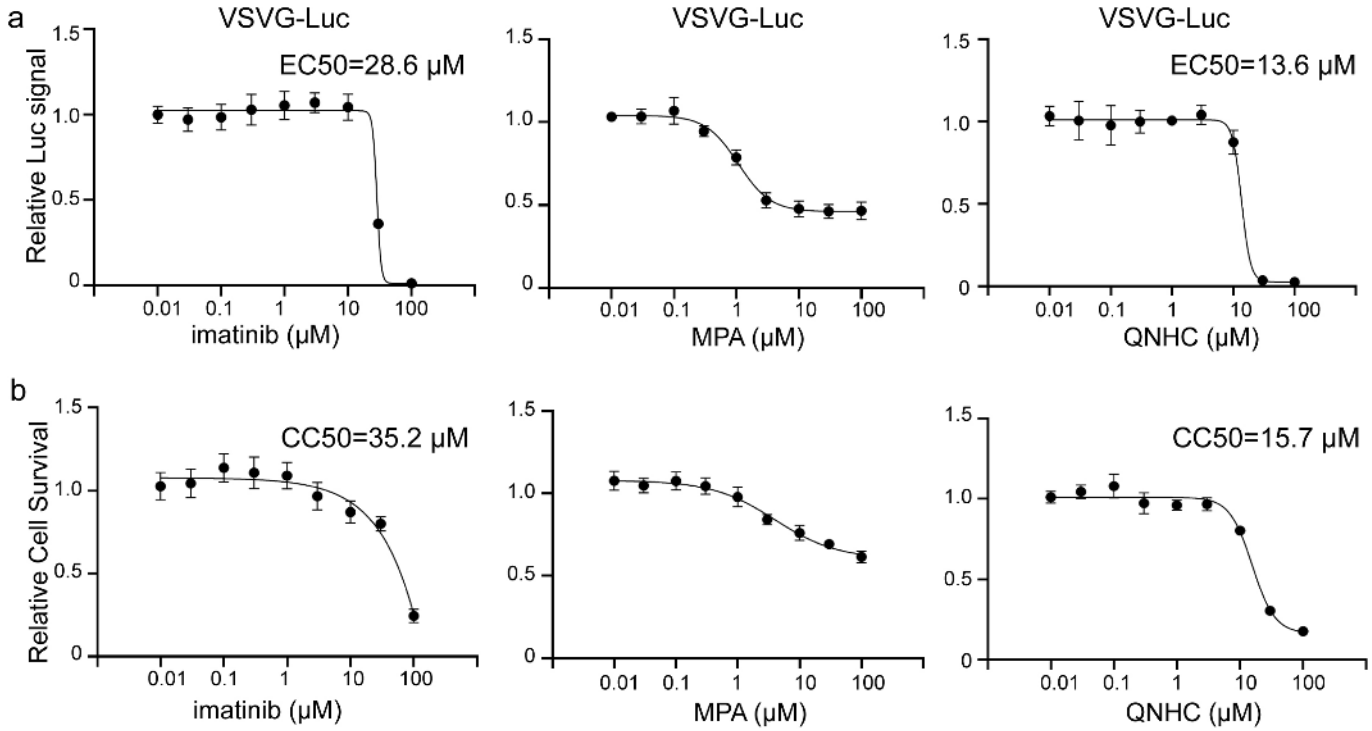
Extended Data Fig. 6 | Efficacy curve of imatinib, MPA and QNHC comparing VSVG and SARS-CoV-1-entry virus on hPSC-LOs. a, Chemical structure, efficacy curve and toxicity curve of two primary hit drug candidates, choloroquine and prochlorperazine. *n* = 3 biological independent experiments. **b**, Bright field+immunostaining images of SARS-CoV-2-entry virus infected hPSC-LOs. Scale bars = 50 μm. Microscale bars = 10 μm. **c**, Efficacy curve of imatinib, MPA and QNHC on VSVG virus. *n* = 3 biological independent experiments. **d**, Efficacy curve of imatinib, MPA and QNHC on SARS-CoV-1-entry virus. *n* = 3 biological independent experiments. Data are representative of at least three independent experiments.



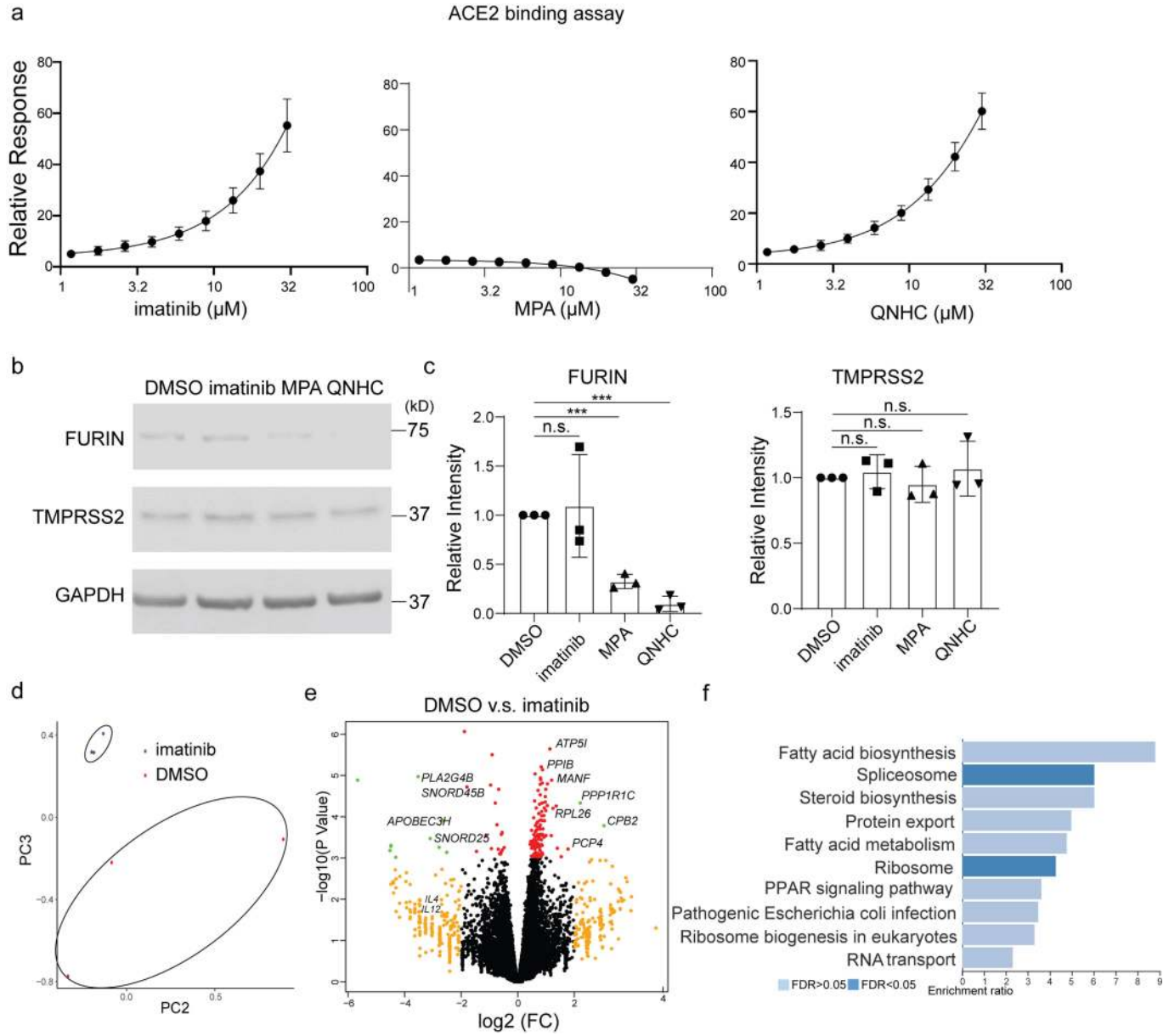
Extended Data Fig. 7 | Efficacy curve of imatinib, MPA and QNHC on hPSC-LOs and hPSC-COs. a, qRT–PCR based dose curve of imatinib, MPA, and QNHC on hPSC-LOs at 24 h post-SARS-CoV-2 infection (SARS-CoV-2, MOI = 0.1). $n = 3$ biological independent experiments. **b**, Bright field+immunostaining images of SARS-CoV-2 Spike protein (SARS-S) and SP-B/SP-C in imatinib, MPA, or QNHC treated hPSC-LOs at 24 hpi (MOI = 0.5). Scale bar = 50 μm. Microscale bars = 15 μm. **c**, qRT–PCR based dose curve of imatinib, MPA, and QNHC on hPSC-COs at 24 hpi of SARS-CoV-2 (SARS-CoV-2, MOI = 0.1). $n = 3$ biological independent experiments. **d**, Bright field+immunostaining images of SARS-S and SP-B/SP-C at 24 hpi of hPSC-COs infected with SARS-CoV-2 virus (MOI = 0.5) and three hours later followed by 10 μM imatinib, 3 μM MPA or 4.5 μM QNHC treatment. Scale bar = 50 μm. Microscale bars = 15 μm. Data are representative of at least three independent experiments.



Extended Data Fig. 8 | Imatinib, MPA, and QNHC inhibit SARS-CoV-2 on Vero cells. a. qRT–PCR based dose curve of imatinib, MPA and QNHC on Vero cells at 24 h post-SARS-CoV-2 infection (SARS-CoV-2, MOI = 0.01). $n = 3$ biological independent experiments. **b**, **c**, Western blotting (**b**) and quantification (**c**) of 3 μM MPA, 4.5 μM QNHC or DMSO-treated Vero cells at 24 h post-SARS-CoV-2 infection (SARS-CoV-2, MOI = 0.01). $n = 3$ biological independent experiments, ***$P = 0.000223$, ***$P = 5.09E\text{-}05$, ***$P = 4.32E\text{-}05$. ***$P = 3.72E\text{-}08$. **d**, **e**, Western blotting (**d**) and quantification (**e**) of DMSO or 10 μM imatinib treated Vero cells at 24 hpi (SARS-CoV-2, MOI = 0.01). $n = 3$ biological independent experiments, ***$P = 7.41E\text{-}18$, ***$P = 8.06E\text{-}07$, ***$P = 7.41E\text{-}18$, ***$P = 3.39E\text{-}06$. ***$P < 0.001$. Data were analysed by an unpaired two-tailed Student's t-test and shown as mean ± STDEV. Data are representative of at least three independent experiments.



Extended Data Fig. 9 | Efficacy and survival curve of imatinib, MPA and QNHC on VSVG virus on Vero cells. a, Inhibition curve of imatinib, MPA and QNHC on VSVG virus. $n = 3$ biological independent experiments. **b**, Cell survival curve of imatinib, MPA and QNHC. $n = 3$ biological independent experiments. Data are representative of at least three independent experiments.



Extended Data Fig. 10 | The impact of imatinib, MPA and QNHC on different steps of viral entry. a. ACE2 binding assay. **b**, **c**, Western blotting (**b**) and quantification (**c**) of TMPRSS2 and FURIN of DMSO, imatinib, MPA and QNHC treated hPSC-LOs. $n = 3$ biological independent experiments, $P = 0.771$, $^{***}P = 8.86E\text{-}05$, $^{***}P = 3.86E\text{-}05$. **d**, PCA plot of RNA-seq data from hPSC-LOs treated with DMSO or 10 μM imatinib at 24 hpi of SARS-CoV-2 virus. **e**, Volcano plot analysis of differential gene expression of hPSC-LOs treated with DMSO or 10 μM imatinib at 24 hpi of SARS-CoV-2 virus. Individual genes are denoted by gene name. **f**, Gene over-representation analysis on KEGG pathway database of differential expression of hPSC-LOs pretreated with DMSO or 10 μM imatinib at 24 hpi of SARS-CoV-2 virus. $n = 3$ biological independent experiments. $^{***}P < 0.001$. Data were analysed by an unpaired two-tailed Student's *t*-test and shown as mean ± STDEV. Data are representative of at least three independent experiments.

nature research

Corresponding author(s): Shuibing Chen, Robert E. Schwartz, Huanhuan Joyce Chen, Todd Evans, Fong Cheng Pan, David D. Ho, Benjamin tenOever, Lewis C. Cantley, Hui Wang

Last updated by author(s): Aug 12, 2020

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	Cell Ranger 10X Genomics https://support.10xgenomics.com/single-cell-gene-expression/software/overview/welcome
Data analysis	Scran Lun ATL, McCarthy DJ, Marioni JC (2016). "A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor." F1000Res., 5, 2122. doi: 10.12688/f1000research.9501.2. https://bioconductor.org/packages/release/bioc/html/scran.html Adobe illustrator CC2017 Adobe https://www.adobe.com/product/photoshop.html Graphpad Prism 7 Graphpad software https://www.graphpad.com FlowJo v10.3 https://www.flowjo.com/solutions/flowjo/downloads Image Studio software 5.2.5 https://www.licor.com/bio/image-studio-lite/download Biacore T200 evaluation software 1.0 https://www.biacore.com/lifesciences/service/downloads/software_and_associated_documents/index.html 10x Cell Ranger - data analysis pipeline (v3.0.2) https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/what-is-cell-ranger R scran package v1.14.1 https://bioconductor.org/packages/scran/ R batchelor package v1.2.1 https://bioconductor.org/packages/batchelor/ R seurat package v3.1.0 https://satijalab.org/seurat/ Bowtie2 v2.4.1 http://bowtie-bio.sourceforge.net/bowtie2/index.shtml IGV software v2.8.9 http://software.broadinstitute.org/software/igv/download

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

The authors declare that the data supporting the findings of this study are available within the paper and it's supplementary information files.
RNA-seq data of hPSC-LOs and lung autopsy samples is publicly available on the GEO repository database under the accession number GSE155241.
https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE155241
scRNA-seq data of hPSC-LOs is publicly available on the GEO repository database under the accession number GSE148113.
https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE148113
scRNA-seq data of hPSC-COs are publicly available on the GEO repository database, accession number GSE147975.
https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE147975
RNA-seq of hPSC-derived endocrine cells and liver organoids are available on the GEO repository database, accession number GSE151803.
https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE151803
DAVID6.8 LHRI https://david.ncifcrf.gov/home.jsp
ToppCell Atlas Toppgene https://toppgene.cchmc.org/
Viral genes (VSV-N VSV-P, VSV-M and VSV-L) are retrieved from NCBI (https://www.ncbi.nlm.nih.gov/nuccore/335873)
SARS-CoV-2/human/USA/WA-CDC-WA1/2020 genome (GenBank: MN985325.1) are retrieved from NCBI (https://www.ncbi.nlm.nih.gov/nuccore/MN985325.1/)

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 for animal experiments was determined based on criteria set by institutional ACUC.
Data exclusions	No data were excluded.
Replication	All attempts at replication are successful and have been included in the studies.
Randomization	The cells or mice were randomly separated into different groups for different treatment.
Blinding	The researchers analyzing the data are blind to sample name and treatment.

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	Immunocytochemistry Human ACE-2 Antibody Polyclonal Goat AF933 R&D Systems 1:200 Immunocytochemistry ACE2 Polyclonal Rabbit ab15348 Abcam 1:500

Immunocytochemistry Firefly luciferase Monoclonal Antibody (CS 17) CS 17 Mouse 35-6700 Thermo Fisher Scientific 1:200
Immunocytochemistry Recombinant Anti-Firefly Luciferase antibody EPR17790 Rabbit ab185924 Abcam 1:100
Immunocytochemistry Anti-NKX2.1 Antibody Polyclonal Rabbit WRAB-1231 Seven Hills Bioreagents 1:500
Immunocytochemistry Anti-SOX2 Antibody Y-17 Goat sc-17320 Santa Cruz 1:150
Immunocytochemistry Anti-FOXA2 Antibody M-20 Goat sc-6554 Santa Cruz 1:150
Immunocytochemistry Anti-SP-C Antibody Polyclonal Rabbit WRAB-76694 Seven Hills Bioreagents 1:500
Immunocytochemistry Anti-SP-B Antibody Polyclonal Rabbit WRAB-48604 Seven Hills Bioreagents 1:500
Intracellular flow cytometry Anti-Pro-SP-C Polyclonal Rabbit WRAB-9337 Seven Hills Bioreagents 1:200
Immunocytochemistry Anti-Pro-SP-B - Rabbit Polyclonal Rabbit WRAB-55522 Seven Hills Bioreagents 1:500
Immunocytochemistry CDX2 CDX2-88 Mouse MU392A-UC Biogenex 1:500
Immunocytochemistry Villin C-19 Goat sc-7672 Santa Cruz 1:200
Immunocytochemistry Villin 1D2C3 Mouse sc-58897 Santa Cruz 1:200
Immunocytochemistry Chromogranin A SP-1 Rabbit 20086 Immunostar 1:300
Immunocytochemistry SATB2 Polyclonal Rabbit HPA001042 Sigma 1:50
Immunocytochemistry Cytokeratin-20 SPM140 Mouse sc-56522 Santa Cruz 1:100
Immunocytochemistry Mucin2 Polyclonal Rabbit ab76774 Abcam 1:100
Immunocytochemistry EPHB2 Polyclonal Goat AF467 R&D 1:100
Immunocytochemistry Cleaved Caspase-3 Asp175 Rabbit 9661S Cell Signaling 1:200
Immunocytochemistry SARS-CoV/SARS-CoV-2 Nucleocapsid Antibody R001 Rabbit 40143-R001 Sino Biological 1:200
Immunocytochemistry Donkey anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 Polyclonal Donkey #A-21202 Thermo Fisher Scientific 1:500
Immunocytochemistry Donkey anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor 594 Polyclonal Donkey #A-21207 Thermo Fisher Scientific 1:500
Immunocytochemistry Donkey anti-Goat IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 647 Polyclonal Donkey #A-21447 Thermo Fisher Scientific 1:500
Immunocytochemistry Donkey anti-Goat IgG Secondary Antibody, Alexa Fluor 594 Polyclonal Donkey A32816 Thermo Fisher 1:500
Immunocytochemistry Donkey anti-Rabbit IgG Secondary Antibody, Alexa Fluor 647 Polyclonal Donkey A32795 Thermo Fisher 1:500
Immunocytochemistry Cy3-conjugated anti rabbit Polyclonal Donkey 711-165-152 Jackson Immunoresearch 1:500
Immunocytochemistry Cy3-conjugated anti goat Polyclonal Donkey 705-165-003 Jackson Immunoresearch 1:500
Immunocytochemistry Cy3-conjugated anti mouse Polyclonal Donkey 715-165-150 Jackson Immunoresearch 1:500
Immunocytochemistry Cy2-conjugated anti rabbit Polyclonal Donkey 711-225-152 Jackson Immunoresearch 1:500
Immunocytochemistry Cy2-conjugated anti goat Polyclonal Donkey 705-225-147 Jackson Immunoresearch 1:500
Immunocytochemistry Cy2-conjugated anti mouse Polyclonal Donkey 715-225-150 Jackson Immunoresearch 1:500
Western Blot Recombinant Anti-Furin antibody EPR14674 Rabbit ab183495 Abcam 1:1000
Western Blot TMPRSS2 H-4 Mouse sc-515727 Santa Cruz 1:500
Western Blot rabbit polyclonal anti-GAPDH polyclonal rabbit G9545 Sigma Aldrich 1:1000
Western Blot mouse monoclonal anti-SARS-CoV-2 Nucleocapsid 1C7 mouse N/A N/A 1:1000
Western Blot mouse monoclonal anti-SARS-CoV-2 Spike 2B3E5 mouse N/A N/A 1:1000
Western Blot Fluorophore-conjugated secondary goat anti-mouse IRDye 680RD Polyclonal Goat 926-68070 IRDye 1:25000
Western Blot Fluorophore-conjugated secondary goat anti-rabbit IRDye 800RD Polyclonal Goat 926-32211 IRDye 1:25000
Western Blot GAPDH 14C10 Rabbit 2118 Cell Signaling 1:1000
Western Blot 680RD Donkey anti-mouse Polyclonal Donkey 926-68072 IRDye 1:25000
Western Blot 680RD Donkey anti-Rabbit Polyclonal Donkey 926-68073 IRDye 1:25000
mouse monoclonal anti-SARS-CoV-2 Nucleocapsid [1C7], and mouse monoclonal anti-SARS-CoV-2 Spike [2B3E5] was a kind gift by Dr. T. Moran, Center for Therapeutic Antibody Discovery at the Icahn School of Medicine at Mount Sinai

Validation

The antibodies were validated by either western blotting or immunostaining by vendors.
Immunocytochemistry Human ACE-2 Antibody Polyclonal Goat AF933 R&D Systems 1:200
Western Blot 1 µg/mL Simple Western 10 µg/mL Immunohistochemistry 3-15 µg/mL
https://www.rndsystems.com/products/human-hamster-ace-2-antibody_af933
Immunocytochemistry ACE2 Polyclonal Rabbit ab15348 Abcam 1:500
Western Blot 1 - 2 µg/ml IHC-P 2 µg/ml
https://www.abcam.com/ace2-antibody-ab15348.html
Immunocytochemistry Firefly luciferase Monoclonal Antibody (CS 17) CS 17 Mouse 35-6700 Thermo Fisher Scientific 1:200
Immunofluorescence (IF) 10-30 µg/mL Western Blot (WB) 1-3 µg/mL Immunohistochemistry (Frozen) (IHC (F)) 15-50 µg/mL https://www.thermofisher.com/antibody/product/Firefly-luciferase-Antibody-clone-CS-17-Monoclonal/35-6700
Immunocytochemistry Recombinant Anti-Firefly Luciferase antibody EPR17790 Rabbit ab185924 Abcam 1:100
ICC/IF 1/500 WB 1/20000
https://www.abcam.com/firefly-luciferase-antibody-epr17790-ab185924.html
Immunocytochemistry Anti-NKX2.1 Antibody Polyclonal Rabbit WRAB-1231 Seven Hills Bioreagents 1:500
Immunoblots - 1:5000, Immunohistochemistry - 1:1000 to 1:2000
https://www.sevenhillsbioreagents.com/products/anti-ttf1-rabbit
Immunocytochemistry Anti-SOX2 Antibody Y-17 Goat sc-17320 Santa Cruz 1:150
IP WB IHC https://www.scbt.com/p/sox-2-antibody-y-17
Immunocytochemistry Anti-FOXA2 Antibody M-20 Goat sc-6554 Santa Cruz 1:150
IP WB IHC https://www.scbt.com/p/hnf-3beta-antibody-m-20?requestFrom=search
Immunocytochemistry Anti-SP-C Antibody Polyclonal Rabbit WRAB-76694 Seven Hills Bioreagents 1:500
Immunoblots - 1:5000 https://www.sevenhillsbioreagents.com/products/anti-mature-sp-c-rabbit
Immunocytochemistry Anti-SP-B Antibody Polyclonal Rabbit WRAB-48604 Seven Hills Bioreagents 1:500
Immunoblots - 1:5000 ELISA - 1:500-1:1000 Immunohistochemistry - 1:1000
https://www.sevenhillsbioreagents.com/products/anti-mature-sp-b-rabbit?_pos=1&_sid=9152d85fc&_ss=r
Intracellular flow cytometry Anti-Pro-SP-C Polyclonal Rabbit WRAB-9337 Seven Hills Bioreagents 1:200
Immunoblots - 1:5000 Immunohistochemistry - 1:3000
https://www.sevenhillsbioreagents.com/products/anti-pro-sp-c-rabbit-n-terminal?_pos=1&_sid=cf74b5282&_ss=r
Immunocytochemistry Anti-Pro-SP-B - Rabbit Polyclonal Rabbit WRAB-55522 Seven Hills Bioreagents 1:500
Immunoblots - 1:5000 ELISA - 1:500-1:1000 Immunohistochemistry - 1:1000

https://www.sevenhillsbioreagents.com/products/anti-pro-sp-b-rabbit?_pos=1&_sid=c905fd6c2&_ss=r
Immunocytochemistry CDX2 CDX2-88 Mouse MU392A-UC Biogenex 1:500
IHC 1:100
http://store.biogenex.com/us/applications/ihc/controls/controls/anti-cdx-2-clone-cdx2-88.html
Immunocytochemistry Villin C-19 Goat sc-7672 Santa Cruz 1:200
IP WB IHC https://www.scbt.com/p/villin-antibody-c-19
Immunocytochemistry Villin 1D2C3 Mouse sc-58897 Santa Cruz 1:200
IP WB IHC https://www.scbt.com/p/villin-antibody-1d2c3?requestFrom=search
Immunocytochemistry Chromogranin A SP-1 Rabbit 20086 Immunostar 1:300
IHC 1/500–1/1,000 https://www.immunostar.com/shop/content/pdf/20086-1317001.pdf
Immunocytochemistry SATB2 Polyclonal Rabbit HPA001042 Sigma 1:50
immunoblotting: 0.04-0.4 µg/mL immunohistochemistry: 1:1000-1:2500
https://www.sigmaaldrich.com/catalog/product/sigma/hpa001042?
lang=en&region=US&gclid=CjwKCAjw4MP5BRBtEiwASfwAL3TV7vo1JGAVcC3Po9cnqh-
eeTUrHAstZtzlfA8hOisZ8dDbCQfg9xoCQ4EQAvD_BwE
Immunocytochemistry Cytokeratin-20 SPM140 Mouse sc-56522 Santa Cruz 1:100
WB, IP, IF and IHC(P) https://www.scbt.com/p/cytokeratin-20-antibody-spm140
Immunocytochemistry Mucin2 Polyclonal Rabbit ab76774 Abcam 1:100
IHC-P 1/50 - 1/100 IHC-Fr 1/50 - 1/100 https://www.abcam.com/muc2-antibody-ab76774.html
Immunocytochemistry EPHB2 Polyclonal Goat AF467 R&D 1:100
Western Blot 0.1 µg/mL Immunohistochemistry 1-15 µg/mL
https://www.rndsystems.com/products/human-mouse-ephb2-antibody_af467
Immunocytochemistry Cleaved Caspase-3 Asp175 Rabbit 9661S Cell Signaling 1:200
Immunohistochemistry (Paraffin) 1:200 Immunofluorescence (Immunocytochemistry) 1:400
https://www.cellsignal.com/products/primary-antibodies/cleaved-caspase-3-asp175-antibody/9661
Immunocytochemistry SARS-CoV/SARS-CoV-2 Nucleocapsid Antibody R001 Rabbit 40143-R001 Sino Biological 1:200
WB,ELISA https://www.sinobiological.com/antibodies/cov-nucleocapsid-40143-r001
Western Blot Recombinant Anti-Furin antibody EPR14674 Rabbit ab183495 Abcam 1:1000
ICC/IF 1/250 IHC-P 1/50 WB 1/1000 - 1/10000
https://www.abcam.com/furin-antibody-epr14674-ab183495.html
Western Blot TMPRSS2 H-4 Mouse sc-515727 Santa Cruz 1:500
WB, IP, IF,ELISA and IHC(P) https://www.scbt.com/p/tmprss2-antibody-h-4
Western Blot rabbit polyclonal anti-GAPDH polyclonal rabbit G9545 Sigma Aldrich 1:1000
WB 0.1-0.2µg/mL
https://www.sigmaaldrich.com/catalog/product/sigma/g9545?
lang=en&region=US&gclid=CjwKCAjw4MP5BRBtEiwASfwAL63LJDcVebbDpSEH4Br5m5x3AsUutPZwkYqHte0_cjsJegbpHEsJFBoCQo4QA
vD_BwE
Western Blot mouse monoclonal anti-SARS-CoV-2 Nucleocapsid 1C7 mouse N/A N/A 1:1000
Ref PMID: 32416070
Western Blot mouse monoclonal anti-SARS-CoV-2 Spike 2B3E5 mouse N/A N/A 1:1000
Ref PMID: 32416070
Western Blot GAPDH 14C10 Rabbit 2118 Cell Signaling 1:1000
Western Blotting 1:1000 Immunohistochemistry (Paraffin) 1:800 Immunofluorescence 1:100 Flow Cytometry 1:200
https://www.cellsignal.com/products/primary-antibodies/gapdh-14c10-rabbit-mab/2118?
fromPage=plp&productId=9101&bvstate=pg:2/ct:r

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK293T (human [Homo sapiens] fetal kidney) and Vero E6 (African green monkey [Chlorocebus aethiops] kidney) were obtained from ATCC (https://www.atcc.org/). RUES2 hESCs, 0013, provided by WiCell, H1 hESCs, 0043, provided WiCell; HUES8 hESCs, 0021, provided by Harvard University.
Authentication	The hESCs are authenticated by Genewiz Company every six months utilizing Short Tandem Repeat (STR) profiling according to the standard ANSI/ATCC ASN-0002-2011. HEK293T and Vero E6 cells were not authenticated.
Mycoplasma contamination	Mycoplasma was checked every three months using Mycoplasma PCR ELISA (Sigma). All cell lines tested negative for mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	No

Animals and other organisms

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

Laboratory animals	Species: mouse Strain: NODscid IL2Rgnull mouse Sex: male Age: 7-9 weeks old

Wild animals	No wild animals were used in the study.
Field-collected samples	No field collected samples were used in the study.
Ethics oversight	Animal protocol was conducted in agreement with NIH guidelines and approved by the WCM Institutional Animal Care and Use Committee (IACUC) and the Institutional Biosafety Committee.

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	The autopsies were collected from COVID-19 patients and non-COVID-19 patients. COVID 64 Male White COVID 47 Male White COVID 87 Female Hispanic Healthy 56 Female White Healthy 66 Male Black Healthy 38 Female Hispanic
Recruitment	There is no patient recruitment. All dead COVID-19 patients that were collected for autopsy examination can enroll in this study. This is a human research subject exempt protocol since it studies autopsy samples.
Ethics oversight	Tissue samples were provided by the Weill Cornell Medicine Department of Pathology. The Tissue Procurement Facility operates under Institutional Review Board (IRB) approved protocol and follows guidelines set by HIPAA. Experiments using samples from human subjects were conducted in accordance with local regulations and with the approval of the institutional review board at the Weill Cornell Medicine under protocol 20-04021814.

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

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Cells were differentiated from ESCs. Cells were dissociated with Accutase and resuspended in FACS buffer.
Instrument	Attune NxT
Software	FlowJo10
Cell population abundance	The PRO-SP-C cells are around 24% of the live cell population.
Gating strategy	FSC-A/SSC-A We have the isotype control for gating

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.