

# A multi-organoid platform identifies *CIART* as a key factor for SARS-CoV-2 infection

Received: 2 May 2022

Accepted: 25 January 2023

Published online: 13 March 2023

 Check for updates

Xuming Tang<sup>1,2,9</sup>, Dongxiang Xue<sup>1,2,9</sup>, Tuo Zhang<sup>3,9</sup>, Benjamin E. Nilsson-Payant<sup>4,8,9</sup>, Lucia Carrau<sup>4,9</sup>, Xiaohua Duan<sup>1,2</sup>, Miriam Gordillo<sup>1,2</sup>, Adrian Y. Tan<sup>3</sup>, Yunping Qiu<sup>5</sup>, Jenny Xiang<sup>3</sup>, Robert E. Schwartz<sup>6,7,10</sup>, Benjamin R. tenOever<sup>4,10</sup>, Todd Evans<sup>1,2,10</sup> & Shuibing Chen<sup>1,2,10</sup> ✉

COVID-19 is a systemic disease involving multiple organs. We previously established a platform to derive organoids and cells from human pluripotent stem cells to model SARS-CoV-2 infection and perform drug screens<sup>1,2</sup>. This provided insight into cellular tropism and the host response, yet the molecular mechanisms regulating SARS-CoV-2 infection remain poorly defined. Here we systematically examined changes in transcript profiles caused by SARS-CoV-2 infection at different multiplicities of infection for lung airway organoids, lung alveolar organoids and cardiomyocytes, and identified several genes that are generally implicated in controlling SARS-CoV-2 infection, including *CIART*, the circadian-associated repressor of transcription. Lung airway organoids, lung alveolar organoids and cardiomyocytes derived from isogenic *CIART*<sup>-/-</sup> human pluripotent stem cells were significantly resistant to SARS-CoV-2 infection, independently of viral entry. Single-cell RNA-sequencing analysis further validated the decreased levels of SARS-CoV-2 infection in ciliated-like cells of lung airway organoids. CUT&RUN, ATAC-seq and RNA-sequencing analyses showed that *CIART* controls SARS-CoV-2 infection at least in part through the regulation of *NR4A1*, a gene also identified from the multi-organoid analysis. Finally, transcriptional profiling and pharmacological inhibition led to the discovery that the Retinoid X Receptor pathway regulates SARS-CoV-2 infection downstream of *CIART* and *NR4A1*. The multi-organoid platform identified the role of circadian-clock regulation in SARS-CoV-2 infection, which provides potential therapeutic targets for protection against COVID-19 across organ systems.

Coronavirus disease 19 (COVID-19) is a multi-organ disease caused by infection of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Although SARS-CoV-2 primarily infects the respiratory tract, patients with COVID-19 present with a wide range of disease indications, including the gastrointestinal, cardiovascular and neurological systems. Organoid models, derived from human pluripotent

stem cells (hPSCs) or adult tissues, have proven to be powerful tools to study viral tropism and the host response, and have also been used for drug screens. A wide range of human cells are permissive for SARS-CoV-2 infection, including those in lung alveolar<sup>2-7</sup>, lung airway<sup>5,8</sup>, small intestine and colon<sup>2</sup>, brain<sup>9-12</sup>, choroid plexus<sup>13</sup>, heart<sup>1,14-21</sup>, liver, pancreas<sup>22</sup>, kidney<sup>23</sup>, blood vessels<sup>24</sup> and tonsil<sup>25</sup> organoids. Although

A full list of affiliations appears at the end of the paper. ✉ e-mail: [shc2034@med.cornell.edu](mailto:shc2034@med.cornell.edu)

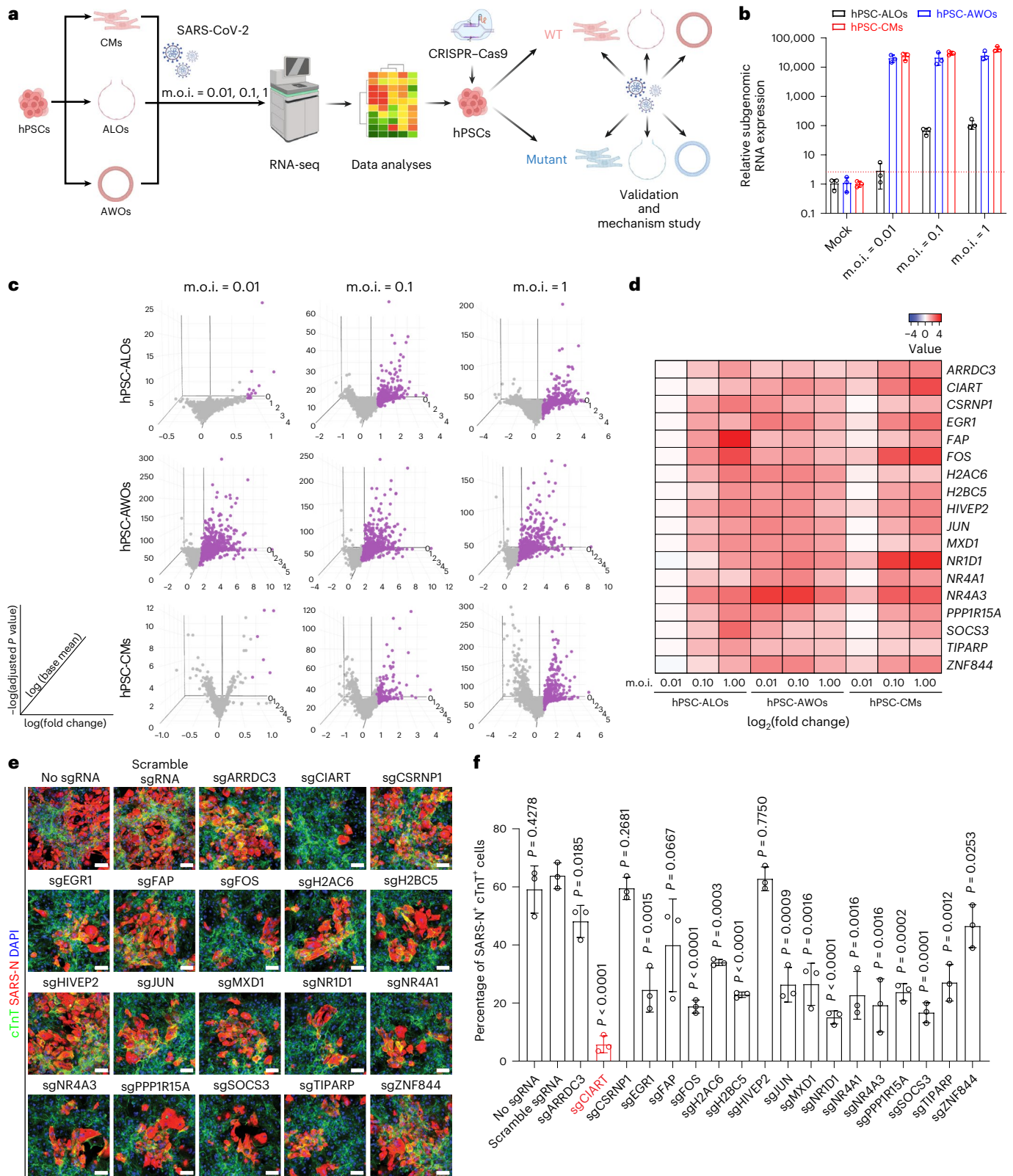
insight into cellular tropism and host responses has been gained, certain key aspects underlying the regulation of infection remain undetermined. In particular, it is unknown whether common factors could be targeted to block or resist SARS-CoV-2 infection across distinct tissue types, which would be a great advantage for the treatment of multi-systemic manifestations of the disease. To address this issue, we performed a systematic analysis of transcriptional changes caused by SARS-CoV-2 infection across several distinct hPSC-derived cell types and organoids. We identified *CIART*, a nuclear transcription factor, as a key regulator of SARS-CoV-2 infection. *CIART* (also known as *CHRONO*, *C1orf51* and *GMI29*) was originally identified as a regulator of a circadian-clock feedback loop<sup>26–28</sup>. Although *CIART* has not been associated with human disease, how the circadian rhythm of host cells may impact viral replication is an area of emerging interest. A recent study using the Calu-3 lung cancer cell line suggests that knockdown of *Bmal1* inhibits SARS-CoV-2 entry<sup>29</sup>. Here we systematically analysed the biological role and downstream mechanism of *CIART* regulation in SARS-CoV-2 infection.

To profile the relative changes in transcript patterns across cell types, hPSC-derived lung alveolar organoids (ALOs; Extended Data Fig. 1a,b), lung airway organoids (AWOs; Extended Data Fig. 1c,d) and cardiomyocytes (CMs; Extended Data Fig. 1e–g) were each exposed to SARS-CoV-2 at different multiplicities of infection (m.o.i. = 0.01, 0.10 and 1.00; Fig. 1a). To validate this range, we monitored the survival of SARS-CoV-2-infected hPSC-CMs (m.o.i. ranging from 0.01 to 4) at 48 h post infection (h.p.i.) and determined that the median lethal dose (LD<sub>50</sub>) is m.o.i. = 0.333 (Extended Data Fig. 2a). At 48 h.p.i., quantitative real-time PCR (qRT-PCR) determination of SARS-CoV-2 viral subgenomic RNA confirmed robust SARS-CoV-2 infection in hPSC-ALOs (m.o.i. = 0.1 and 1), hPSC-AWOs (all m.o.i.) and hPSC-CMs (all m.o.i.), with relatively higher expression levels in hPSC-AWOs and hPSC-CMs compared with hPSC-ALOs (Fig. 1b). RNA sequencing (RNA-seq) was used to systematically profile transcriptional changes of these hPSC-derived cells/organoids caused by SARS-CoV-2 infection. Three criteria were applied to choose the most highly expressed genes that were significantly changed in each condition: log<sub>2</sub>(fold change) > 0.75, base mean > 10 and adjusted *P* value < 0.05 (Fig. 1c). We first monitored the induction of genes involved in interferon I (IFN-I) pathways following SARS-CoV-2 infection. IFN-I-associated genes were mainly upregulated in hPSC-AWOs but not hPSC-ALOs or hPSC-CMs (Extended Data Fig. 2b). This is consistent with previous findings that the expression pattern of IFN-associated genes differs at distinct respiratory tract sites following SARS-CoV-2 infection<sup>30</sup>. SARS-CoV-2 infection can upregulate host genes, including *ACE2*, to facilitate its infection<sup>31</sup>. Given that we aimed to identify genes that are consistently upregulated following SARS-CoV-2 infection among different organoids/cell types, the IFN-I-associated genes do not meet this criterion. Eighteen genes were identified as significantly changed in seven of nine conditions, excepting the lowest m.o.i. for hPSC-ALOs and in some cases for the lowest m.o.i. for hPSC-CMs (Fig. 1d). These genes (*ARRDC3*, *CIART*, *CSRNPI*, *EGR1*, *FAP*, *FOS*, *H2AC6*, *H2BC5*, *HIVEP2*, *JUN*, *MXD1*, *NR1D1*, *NR4A1*, *NR4A3*, *PPP1R15A*, *SOCS3*, *TIPARP* and *ZNF844*) were individually tested to determine their relevance to infection efficiency. For this purpose, clustered regularly interspaced short palindromic repeats (CRISPR)-based knockout of each gene was performed in hPSC-CMs, which are relatively homogeneous. Briefly, hPSC-CMs were infected with lentivirus expressing Cas9 and two single-guide RNAs (sgRNAs) targeting each gene (Supplementary Table 1). After 3 d of puromycin selection, the knockout hPSC-CMs were subjected to infection with SARS-CoV-2. Knockout of *CIART* showed the greatest resistance to infection compared with the control hPSC-CMs (no sgRNA or scrambled sgRNA), although significantly reduced levels of SARS-CoV-2 infection were also observed by individually targeting *EGR1*, *FOS*, *H2AC6*, *H2BC5*, *JUN*, *MXD1*, *NR1D1*, *NR4A1*, *NR4A3*, *PPP1R15A*, *SOCS3* and *TIPARP* (Fig. 1e,f).

To validate the requirement for *CIART* in SARS-CoV-2 infections, we used CRISPR–Cas9-based gene targeting to create isogenic *CIART*<sup>−/−</sup> hPSCs. H1 human embryonic stem cells (H1-hESCs) were electroporated with a vector expressing Cas9 and a specific sgRNA targeting the first exon of *CIART* (Extended Data Fig. 3a). After subcloning, multiple independent clones with biallelic frameshift mutations were expanded. To account for possible variation between different clones, two wild-type (WT) clones (derived from the targeting process but without mutations in *CIART*) and two *CIART*<sup>−/−</sup> clones (no. 1 and no. 2) were chosen for further analysis. Biallelic indel mutations in each isogenic *CIART*<sup>−/−</sup> H1-hESC line were verified by DNA sequencing (Extended Data Fig. 3b). Mutant clone no. 1 has a single T insertion at both alleles, whereas mutant clone no. 2 has two distinct deletions. Both indel mutations create early frameshifts that are predicted to generate null alleles. All established clones displayed typical hPSC colony morphology and expressed pluripotency markers, including OCT4, NANOG, SOX2, SSEA4, TRA-1-60 and TRA-1-81 (Extended Data Fig. 3c). Western blotting further validated the knockout of *CIART* in the mutant H1-hESC clones as well as in hPSC-AWOs, hPSC-ALOs and hPSC-CMs (Extended Data Fig. 3d).

The isogenic WT and mutant *CIART*<sup>−/−</sup> hESC lines were differentiated to CMs, AWOs and ALOs to evaluate the impact of *CIART* in SARS-CoV-2 infection. Both WT and *CIART*<sup>−/−</sup> H1-hESCs showed comparable capacities to differentiate into CMs, identified by signature sarcomere structures and expression of the cardiac marker  $\alpha$ -actinin and cardiac Troponin-T (cTnT; Extended Data Fig. 4a). At 24 h.p.i., the *CIART*<sup>−/−</sup> hPSC-CMs were highly resistant to viral infection, as determined by qRT-PCR (Fig. 2a) and immunofluorescence staining analyses (Fig. 2b,c). The *CIART*<sup>−/−</sup> and WT hPSCs were also equally capable of differentiating to hPSC-AWOs, based on the generation of FOXJ1<sup>+</sup> ciliated cells (Extended Data Fig. 4b). Similar to mutant CMs, significantly decreased levels of SARS-CoV-2 infection were detected in *CIART*<sup>−/−</sup> hPSC-AWOs, as determined by qRT-PCR (Fig. 2d), whereas the percentages of SARS-N<sup>+</sup> cells were significantly lower in FOXJ1<sup>+</sup> ciliated cells of *CIART*<sup>−/−</sup> hPSC-AWOs compared with the WT hPSC-AWOs (Fig. 2e,f). Finally, *CIART*<sup>−/−</sup> hPSCs were differentiated to hPSC-ALOs. Both WT and *CIART*<sup>−/−</sup> cells generated hPSC-ALOs with abundant mature SP-B and mature SP-C double-positive (SP-B<sup>+</sup>SP-C<sup>+</sup>) alveolar type 2 cells (Extended Data Fig. 4c,d). As seen in hPSC-CMs and hPSC-AWOs, *CIART*<sup>−/−</sup> hPSC-ALOs were highly resistant to infection, as determined by qRT-PCR (Fig. 2g) and immunofluorescence staining in SP-B<sup>+</sup>SP-C<sup>+</sup> alveolar type 2 cells (Fig. 2h–k). Consistent with the results at 24 h.p.i., *CIART*<sup>−/−</sup> hPSC-CMs (Extended Data Fig. 5a–c), hPSC-AWOs (Extended Data Fig. 5d–f) and hPSC-ALOs (Extended Data Fig. 5g–k) also showed significantly reduced SARS-CoV-2 infection compared with their WT counterparts at 48 h.p.i. To determine whether loss of *CIART* affects viral entry, *CIART*<sup>−/−</sup> hPSC-AWOs and hPSC-ALOs were infected with a SARS-CoV-2 Spike protein pseudo-typed entry virus. No significant difference in viral entry was detected (Extended Data Fig. 5l,m). However, it is worth noting that the overall infection rate of pseudo-typed entry viruses in primary cells/organoids can be relatively low. Together, the data suggest that loss of *CIART* impairs SARS-CoV-2 infection through an entry-independent mechanism. We infected hPSC-ALOs, -AWOs and -CMs with influenza to test whether *CIART* upregulation is specific to SARS-CoV-2 infection. Robust influenza infection was determined by qRT-PCR (Extended Data Fig. 5n). *CIART* expression was also upregulated following influenza infection of ALOs, AWOs and CMs (Extended Data Fig. 5o), indicating that the induction of *CIART* is not specific to SARS-CoV-2 infection.

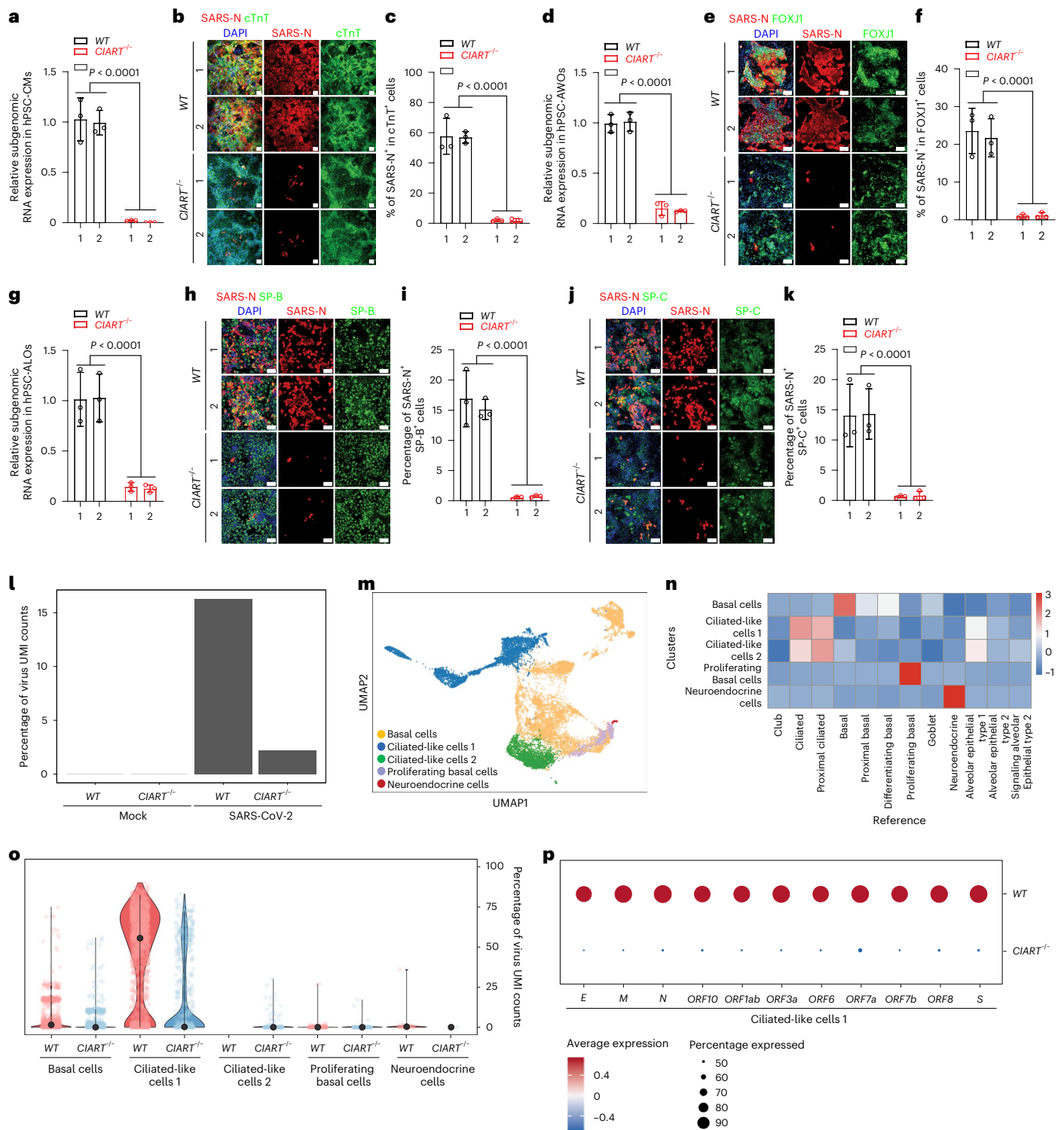
Lung organoids contain multiple cell types, so single-cell RNA-seq (scRNA-seq) was applied to analyse SARS-CoV-2 infection in defined lineages of hPSC-AWOs. Both WT and *CIART*<sup>−/−</sup> AWOs were infected with SARS-CoV-2 (m.o.i. = 0.1) and dissociated into single cells for scRNA-seq at 24 h.p.i. Consistent with previous observations (Fig. 2d), the percentage of viral reads in infected *CIART*<sup>−/−</sup> AWOs was significantly lower compared with WT AWOs (Fig. 2l). Five clusters were



**Fig. 1 | A multi-organoid platform to identify genes involved in SARS-CoV-2 infection.** **a**, Schematic of the experimental design. **b**, Levels of subgenomic viral transcripts, determined by qRT-PCR, in hPSC-derived AWOs, ALOs and CMs at 48 h.p.i. with SARS-CoV-2 at different m.o.i. (m.o.i. = 0.01, 0.10 and 1.00). The dashed red line indicates the detection limit. **c**, Three-dimensional analysis of transcriptional changes in hPSC-derived AWOs, ALOs and CMs infected at 48 h.p.i. (m.o.i. = 0.01, 0.10 and 1.00). The genes that were significantly changed ( $\log_2(\text{fold change}) > 0.75$ , base mean  $> 10$  and adjusted  $P < 0.05$ ) in each condition are highlighted in purple. **d**, Heatmap of the protein-coding genes that were

increased for at least seven of nine conditions in hPSC-derived AWOs, ALOs and CMs at 48 h.p.i. (m.o.i. = 0.01, 0.10 and 1.00). **d**, Data are presented as an integration of all biological replicates. **e, f**, Representative confocal images (**e**) and the calculated percentage of SARS-N<sup>+</sup> cells in the cTnT<sup>+</sup> subpopulation (**f**) of hPSC-CMs infected with lentivirus carrying Cas9 and sgRNAs targeting hit genes. P-values were calculated using an unpaired two-tailed Student's *t*-test. The red text and red bar highlight that knockout of CIART showed the greatest resistance to SARS-CoV-2 infection. **b, e, f**, Data are the mean  $\pm$  s.d. **b, c, f**,  $n = 3$  independent biological replicates.

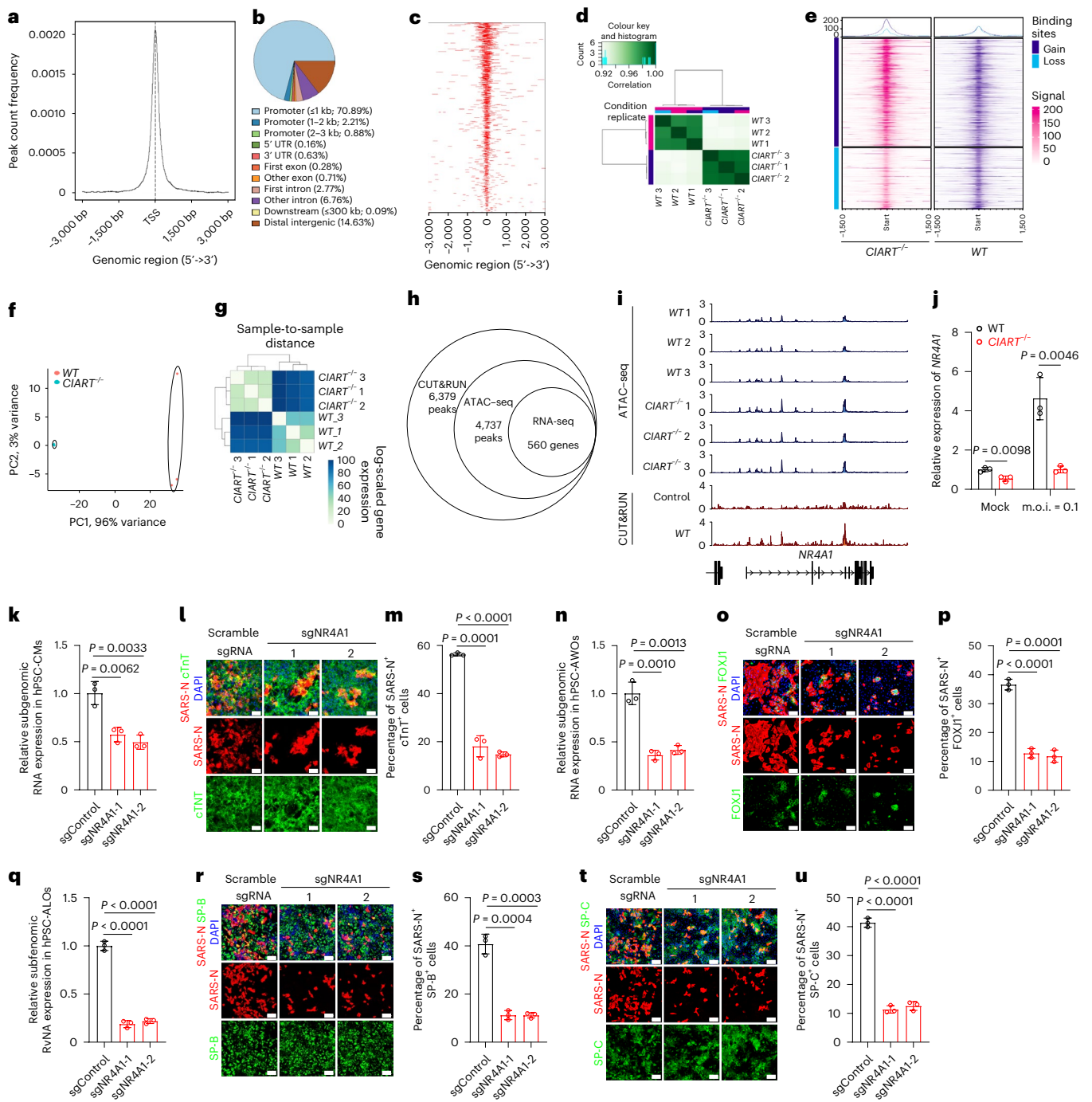




**Fig. 2 | Loss of CIART decreases SARS-CoV-2 infection.** **a**, Relative expression levels of viral RNA in WT and *CIART*<sup>-/-</sup> hPSC-CMs at 24 h.p.i. with SARS-CoV-2 (m.o.i. = 0.3). **b, c**, Representative confocal images (**b**) and calculated percentages (**c**) of SARS-N<sup>+</sup> cells within the cTnT<sup>+</sup> cell populations of WT and *CIART*<sup>-/-</sup> hPSC-CMs at 24 h.p.i. with SARS-CoV-2 (m.o.i. = 0.3). **d**, Relative expression levels of viral RNA in WT and *CIART*<sup>-/-</sup> hPSC-AWOs at 24 h.p.i. with SARS-CoV-2 (m.o.i. = 0.3). **e, f**, Representative confocal images (**e**) and calculated percentages (**f**) of SARS-N<sup>+</sup> cells within the FOXJ1<sup>+</sup> cell populations of WT and *CIART*<sup>-/-</sup> hPSC-AWOs at 24 h.p.i. with SARS-CoV-2 (m.o.i. = 0.3). **g**, Relative expression levels of viral RNA in WT and *CIART*<sup>-/-</sup> hPSC-ALOs at 24 h.p.i. with SARS-CoV-2 (m.o.i. = 0.3). **h, i**, Representative confocal images (**h**) and calculated percentages (**i**) of SARS-N<sup>+</sup> cells in the mature SP-B<sup>+</sup> cell populations of WT and *CIART*<sup>-/-</sup> hPSC-ALOs at 24 h.p.i. with SARS-CoV-2 (m.o.i. = 0.3). **j, k**, Representative confocal

images (**j**) and calculated percentages (**k**) of SARS-N<sup>+</sup> cells in the mature SP-C<sup>+</sup> cell populations of WT and *CIART*<sup>-/-</sup> hPSC-ALOs at 24 h.p.i. with SARS-CoV-2 (m.o.i. = 0.3). **a–k**, *n* = 3 independent biological replicates. **l**, Percentage of viral UMI counts in mock- and SARS-CoV-2-infected WT and *CIART*<sup>-/-</sup> hPSC-AWOs (m.o.i. = 0.1; 24 h.p.i.). **m**, Uniform manifold approximation and projection (UMAP) plot illustrating five cell clusters in the hPSC-AWOs. **n**, Correlation analysis of cell clusters in hPSC-AWOs and adult human lung. **o**, Percentage of viral UMI counts in each cluster of WT and *CIART*<sup>-/-</sup> hPSC-AWOs following SARS-CoV-2 infection (m.o.i. = 0.1; 24 h.p.i.). **p**, Levels of SARS-CoV-2 viral transcripts in ciliated-like cells 1 of WT and *CIART*<sup>-/-</sup> hPSC-AWOs following SARS-CoV-2 infection (m.o.i. = 0.1; 24 h.p.i.). Data are the mean ± s.d. *P* values were calculated using an unpaired two-tailed Student's *t*-test. Scale bars, 100 μm.





**Fig. 3 | CIART promotes SARS-CoV-2 infection through NR4A1 regulation.**

**a**, Average profile of *CIART* CUT&RUN peaks in WT hPSC-CMs around the TSS. **b**, Distribution of the genomic locations of CUT&RUN peaks in WT hPSC-CMs. **c**, Profile heatmap showing the distribution of CUT&RUN peaks in WT hPSC-CMs around the TSS. **a**–**c**, Data are presented as an integration of all biological replicates;  $n = 2$  independent biological replicates. **d**, Clustering analysis of ATAC-seq data of WT and *CIART*<sup>-/-</sup> hPSC-CMs. **e**, Profile ATAC-seq heatmap showing the enrichment of gain and loss sites in WT and *CIART*<sup>-/-</sup> hPSC-CMs. **f**, **g**, RNA-sequencing PCA (**f**) and sample clustering (**g**) analysis of WT and *CIART*<sup>-/-</sup> hPSC-AWOs under mock infection conditions. **d**–**g**, Data are presented as the individual biological replicates (**d**, **f**, **g**) or an integration of all biological replicates (**e**);  $n = 3$  independent biological replicates. **h**, Summary of peaks in the CUT&RUN, ATAC-seq and RNA-seq assays. **i**, Peaks associated with the *NR4A1* gene in WT and *CIART*<sup>-/-</sup> hPSC-CMs in the ATAC-seq and CUT&RUN assays. Data are presented as individual biological replicates ( $n = 3$  for ATAC-seq) and 2 for CUT&RUN). The schematic below the plots illustrates the exon and intron regions of gene *NR4A1* in the genome. **j**, Expression levels, measured by qRT-PCR, of *NR4A1* in WT

and *CIART*<sup>-/-</sup> hPSC-CMs following mock or SARS-CoV-2 infection (m.o.i. = 0.1). **k**–**m**, Relative expression levels of SARS-CoV-2 RNA in hPSC-CMs (**k**) as well as representative confocal images (**l**) and calculated percentages (**m**) of SARS-N<sup>+</sup> cells in the cTnT<sup>+</sup> subsets of hPSC-CMs expressing scramble sgRNA or sgNR4A1 at 24 h.p.i. (m.o.i. = 0.1). **n**–**p**, Relative expression levels of SARS-CoV-2 viral RNA in hPSC-AWOs (**n**) as well as representative confocal images (**o**) and calculated percentages (**p**) of SARS-N<sup>+</sup> cells within the FOXJ1<sup>+</sup> cell populations of hPSC-AWOs expressing scramble sgRNA or sgNR4A1 at 24 h.p.i. (m.o.i. = 0.1). **q**–**s**, Relative expression levels of SARS-CoV-2 RNA in hPSC-ALOs (**q**) as well as representative confocal images (**r**) and calculated percentages (**s**) of SARS-N<sup>+</sup> cells within the mature SP-B<sup>+</sup> cell populations of hPSC-ALOs expressing scramble sgRNA or sgNR4A1 at 24 h.p.i. (m.o.i. = 0.1). **t**, **u**, Representative confocal images (**t**) and calculated percentages (**u**) of SARS-N<sup>+</sup> cells within mature SP-C<sup>+</sup> cell populations of hPSC-ALOs expressing scramble sgRNA or sgNR4A1 at 24 h.p.i. (m.o.i. = 0.1). **j**–**u**, Data are presented as the mean  $\pm$  s.d. (**j**, **k**, **m**, **p**, **q**, **s**, **u**);  $n = 3$  independent biological replicates. *P* values were calculated by unpaired two-tailed Student's *t*-test. Scale bars, 100  $\mu$ m.

identified, including basal cells (cluster 0), ciliated-like cells (clusters 1 and 2), proliferating basal cells (cluster 3) and neuroendocrine cells (cluster 4; Fig. 2m and Extended Data Fig. 6a,b). Correlation analysis of signature genes further validated the identity of the hPSC-derived ciliated-like cell population, showing high similarity to adult human ciliated cells<sup>32</sup> (Fig. 2n). We examined the percentage of viral unique molecular identifier (UMI) counts in each cluster of WT and *CIART*<sup>-/-</sup> AWOs following SARS-CoV-2 infection. Viral UMIs were mainly detected in ciliated-like cell cluster 1 (Fig. 2o). The percentage of viral UMIs was significantly lower in each cluster of *CIART*<sup>-/-</sup> AWOs compared with their WT AWO counterparts (Fig. 2o). We further examined individual viral transcripts in ciliated-like cells and found that all viral transcripts were significantly decreased in *CIART*<sup>-/-</sup> ciliated-like cells (Fig. 2p). Together, these data further confirm that loss of *CIART* significantly impairs SARS-CoV-2 infection.

*CIART* was identified as a key transcription factor involved in the regulation of the mammalian circadian clock<sup>26–28</sup>. We performed cleavage under targets and release using nuclease (CUT&RUN) chromatin profiling, assay for transposase-accessible chromatin using sequencing (ATAC-seq) and RNA-seq to identify downstream targets and signaling pathways regulated by *CIART* in hPSC-CMs. We identified 6,379 peaks using CUT&RUN, most of which were close to the transcription start site (TSS; Fig. 3a,b). More than 50% of the identified peaks were located in proximal promoter regions (Fig. 3c). The ATAC-seq peaks derived from WT and *CIART*<sup>-/-</sup> hPSC-CMs clustered separately (Fig. 3d) with 12,529 sites gained and 564 sites lost in *CIART*<sup>-/-</sup> hPSC-CMs compared with WT hPSC-CMs (Fig. 3e). From the RNA-seq data, clustering and principal component analyses (PCA), we confirmed that WT and *CIART*<sup>-/-</sup> hPSC-CMs are distinguishable at the transcript level (Fig. 3f,g). When we combined the CUT&RUN, ATAC-seq and RNA-seq assays, 671 peaks (associated with 560 genes) that are bound by *CIART* in WT hPSC-CMs and are significantly changed between WT and *CIART*<sup>-/-</sup> hPSC-CMs at both transcriptional and chromatin accessibility levels were identified (Fig. 3h). Interestingly, one of these genes, *NR4A1*, was also identified along with *CIART* as an upregulated hit gene following SARS-CoV-2 infection in the initial hPSC-derived multi-organoid platform (Fig. 1f). The peaks associated with *NR4A1* in the CUT&RUN assay (Fig. 3i) were correlated in the ATAC-seq data and significantly changed in *CIART*<sup>-/-</sup> hPSC-CMs compared with WT hPSC-CMs. Interestingly, some ATAC-seq peaks were enhanced in the mutant cells, suggesting an altered chromatin structure with loss of *CIART* binding. Finally, qRT-PCR assays confirmed that the transcript levels of *NR4A1* were significantly decreased in both mock- and SARS-CoV-2-infected *CIART*<sup>-/-</sup> hPSC-CMs (Fig. 3j), consistent with the RNA-seq data. To determine whether *NR4A1* impacts SARS-CoV-2 infection, hPSC-derived CMs, AWOs and ALOs were infected with lentivirus expressing Cas9 and one of two different sgRNAs targeting *NR4A1* (Supplementary Table 1). CMs, AWOs and ALOs expressing sgRNA targeting *NR4A1* (sgNR4A1) or scrambled sgRNA control were infected with SARS-CoV-2 (m.o.i. = 0.1). At 24 h.p.i., both subgenomic viral

RNAs and the percentage of SARS-N<sup>+</sup> cells in cTnT<sup>+</sup> CMs (Fig. 3k–m), FOXJ1<sup>+</sup> cells in AWOs (Fig. 3n–p) and SP-B<sup>+</sup>SP-C<sup>+</sup> cells in ALOs (Fig. 3q–u) were significantly decreased in cells expressing sgNR4A1 compared with those expressing scramble sgRNA.

To define the downstream signalling pathways regulated by *CIART*, RNA-seq was applied to analyse WT as well as *CIART*<sup>-/-</sup> hPSC-AWOs, hPSC-ALOs and hPSC-CMs. Clustering and PCA analyses showed that the transcript profiles of WT and *CIART*<sup>-/-</sup> hPSC-AWOs clustered separately both in mock conditions (PCA plot, Extended Data Fig. 7a; clustering, Extended Data Fig. 7b) and following SARS-CoV-2 infection (PCA plot, Extended Data Fig. 7c; clustering, Extended Data Fig. 7d). Similarly, the transcript profiles of WT and *CIART*<sup>-/-</sup> hPSC-ALOs clustered separately both in mock conditions (PCA plot, Extended Data Fig. 7e; clustering, Extended Data Fig. 7f) and 24 h.p.i. with SARS-CoV-2 (m.o.i. = 0.1; PCA plot, Extended Data Fig. 7g; clustering, Extended Data Fig. 7h). The transcript profiles of WT and *CIART*<sup>-/-</sup> hPSC-CMs clustered separately both in mock conditions (PCA plot, Extended Data Fig. 7i; clustering, Extended Data Fig. 7j) and 24 h.p.i. with SARS-CoV-2 (m.o.i. = 0.1; PCA plot, Extended Data Fig. 7k; clustering, Extended Data Fig. 7l). Ingenuity pathway analysis highlighted Retinoid X receptor (RXR) signalling pathways in *CIART*<sup>-/-</sup> hPSC-AWOs (Fig. 4a), *CIART*<sup>-/-</sup> hPSC-ALOs (Fig. 4b) and *CIART*<sup>-/-</sup> hPSC-CMs (Fig. 4c). Heatmaps showed the downregulation of RXR pathway-associated genes in *CIART*<sup>-/-</sup> hPSC-AWOs (Fig. 4d), hPSC-ALOs (Fig. 4e) and hPSC-CMs (Fig. 4f). In addition, RXR pathway-associated genes were also downregulated following SARS-CoV-2 infection in *CIART*<sup>-/-</sup> hPSC-AWOs (Extended Data Fig. 7m), hPSC-ALOs (Extended Data Fig. 7n) and hPSC-CMs (Extended Data Fig. 7o). Previous studies reported that NR4A1 could heterodimerize with RXR and increase the potential of RXR to modulate gene expression<sup>33</sup>. The downregulation of RXR pathway-associated genes in hPSC-AWOs (Extended Data Fig. 7p), hPSC-ALOs (Extended Data Fig. 7q) and hPSC-CMs (Extended Data Fig. 7r) expressing sgNR4A1 was confirmed using qRT-PCR assays. Finally, hPSC-AWOs, hPSC-ALOs and hPSC-CMs were treated with RXR inhibitors, followed by SARS-CoV-2 infection. Treatment with the RXR inhibitors HX531 and PA452 blocked SARS-CoV-2 infection in hPSC-CMs (subgenomic viral RNA, Fig. 4g; viral antigen, Fig. 4h,i), hPSC-AWOs (subgenomic viral RNA, Fig. 4j; viral antigen, Fig. 4k,l) and hPSC-ALOs (subgenomic viral RNA, Fig. 4m; viral antigen, Fig. 4n–q) at 24 h.p.i., as determined using qRT-PCR and immunostaining assays. Together, these data confirm that inhibition of the RXR pathway suppresses SARS-CoV-2 infection.

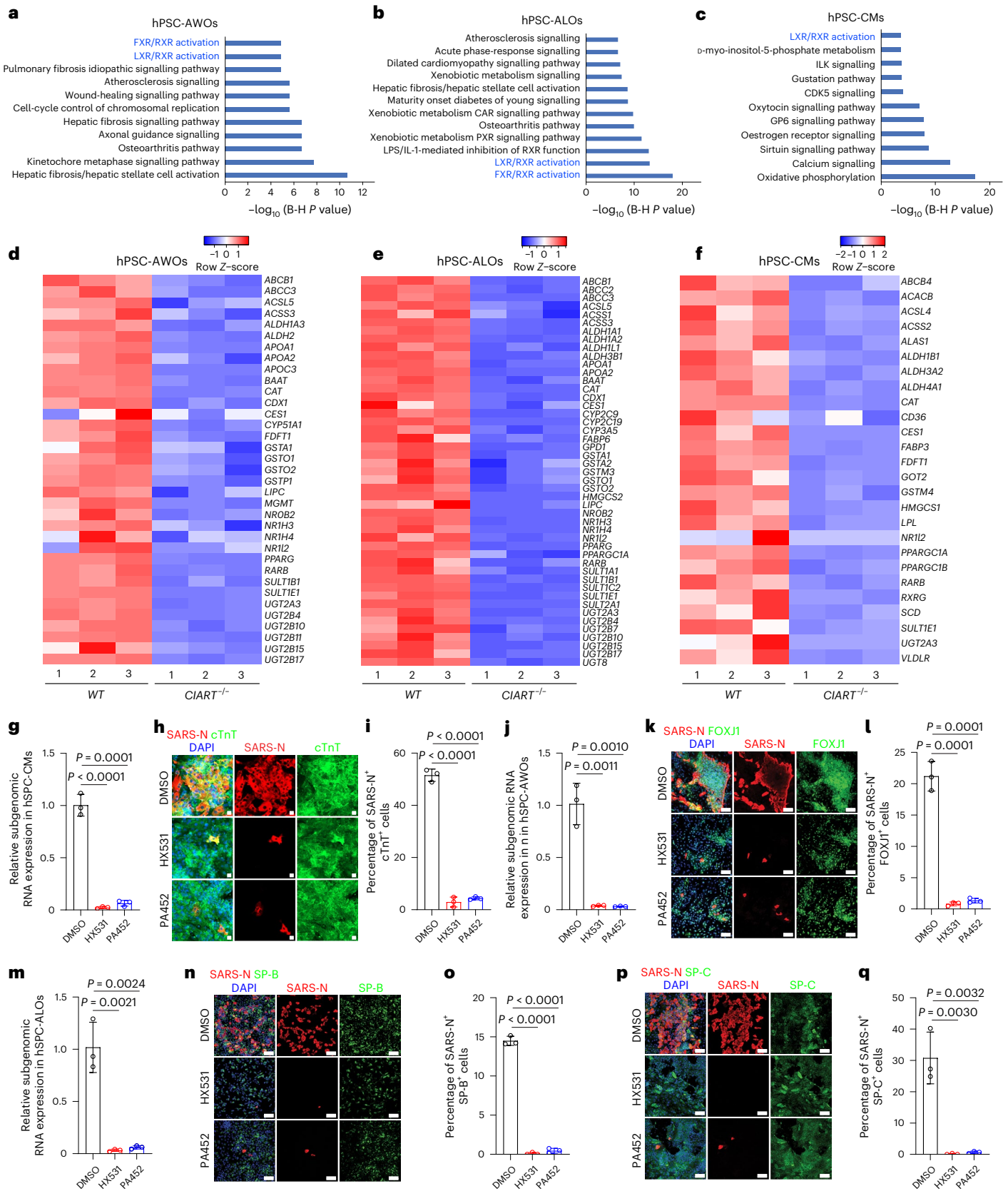
Previous studies demonstrated that the RXR signalling pathway regulates fatty-acid metabolism<sup>34,35</sup>. We analysed RNA-seq data of WT and *CIART*<sup>-/-</sup> hPSC-AWOs and found decreased expression levels in *CIART*<sup>-/-</sup> hPSC-AWOs for multiple genes involved in fatty-acid synthesis including *ACLY*, *ACOT1*, *ACSL4*, *CBR4*, *DLAT*, *FASN* and *MCAT* (Extended Data Fig. 8a). We performed RNA-seq analysis of HX531- or dimethylsulfoxide (DMSO)-treated AWOs. Consistent with the data for *CIART*<sup>-/-</sup> hPSC-AWOs, the expression of fatty-acid-synthesis-associated genes

**Fig. 4 | *CIART* regulates SARS-CoV-2 infection through the RXR pathway.** **a–c**, Enriched pathways in *CIART*<sup>-/-</sup> hPSC-AWOs (**a**), hPSC-ALOs (**b**) and hPSC-CMs (**c**) relative to their WT counterparts. The B–H p-value was corrected by the Benjamini–Hochberg method. The text in blue highlights that the RXR signaling pathway was enriched in *CIART*<sup>-/-</sup> AWOs, ALOs and CMs. **d–f**, Heatmap of RXR pathway-associated genes comparing WT and *CIART*<sup>-/-</sup> hPSC-AWOs (**d**), hPSC-ALOs (**e**) and hPSC-CMs (**f**). **g**, Relative expression levels of SARS-CoV-2 RNA in hPSC-CMs treated with DMSO, HX531 or PA452 (24 h.p.i.; m.o.i. = 0.3). **h,i**, Representative confocal images (**h**) and calculated percentages (**i**) of SARS-N<sup>+</sup> cells in the cTnT<sup>+</sup> subsets of hPSC-CMs treated with DMSO, HX531 or PA452 (24 h.p.i.; m.o.i. = 0.3). **j**, Relative expression levels of SARS-CoV-2 RNA in hPSC-AWOs treated with DMSO, HX531 or PA452 (24 h.p.i.; m.o.i. = 0.3). **k,l**, Representative confocal images (**k**) and calculated percentages (**l**) of

SARS-N<sup>+</sup> cells within the FOXJ1<sup>+</sup> cell populations of hPSC-AWOs treated with DMSO, HX531 or PA452 (24 h.p.i.; m.o.i. = 0.3). **m**, Relative expression levels of SARS-CoV-2 RNA in hPSC-ALOs treated with DMSO, HX531 or PA452 (24 h.p.i.; m.o.i. = 0.3). **n,o**, Representative confocal images (**n**) and calculated percentages (**o**) of SARS-N<sup>+</sup> cells within the mature SP-B<sup>+</sup> cell populations of hPSC-ALOs treated with DMSO, HX531 or PA452 (24 h.p.i.; m.o.i. = 0.3). **p,q**, Representative confocal images (**p**) and calculated percentages (**q**) of SARS-N<sup>+</sup> cells within the mature SP-C<sup>+</sup> cell populations of hPSC-ALOs treated with DMSO, HX531 or PA452 (24 h.p.i.; m.o.i. = 0.3). Data are presented as an integration of all biological replicates (**a–c**), individual biological replicates (**d–f**) or the mean ± s.d. (**g,i,j,l,m,o,q**); *n* = 3 independent biological replicates. *P* values were calculated using a paired or an unpaired two-tailed Student's *t*-test. Scale bars, 100 μm.

are downregulated in HX531-treated hPSC-AWOs (clustering, Extended Data Fig. 8b; PCA plot, Extended Data Fig. 8c; heatmap, Extended Data Fig. 8d). To further confirm the change in fatty-acid synthesis caused by *CIART* knockout or the RXR inhibitor HX531, we performed

metabolism profiling of WT, *CIART*<sup>-/-</sup> and WT + HX531-treated hPSC-AWOs. Several fatty acids, including palmitic acid, stearic acid, 11-eicosenoic acid, arachidic acid and myristic acid, were decreased in *CIART*<sup>-/-</sup> and WT + HX531 hPSC-AWOs compared with





the control WT without treatment (Extended Data Fig. 8e). Previous studies reported that inhibition of fatty-acid synthesis could suppress SARS-CoV-2 infection<sup>8,36</sup>. Together, these data suggest that loss of CIART decreases fatty-acid synthesis, which results in diminished SARS-CoV-2 infection.

Several host factors regulating SARS-CoV-2 infection have been identified by CRISPR-based screens using Vero E6 cells, cancer cell lines and patient-derived organoids<sup>37,38</sup>. Here we used the hPSC-derived multi-organoid system (including AWOs, ALOs and CMs). It is technically very challenging to isolate primary adult ALOs, AWOs and CMs from the same donor, and thus hPSCs were an optimal system for our purpose as they can be differentiated into multiple types of isogenic cells and organoids. However, a limitation of hPSC-derived organoids is that they are not as mature as adult tissue. Using the hPSC-derived multi-organoid system, we identified *CIART*, a nuclear transcription factor, as a key regulator of SARS-CoV-2 infection. Compared with control WT cells, isogenic derivatives from *CIART*<sup>-/-</sup> hPSCs poorly support SARS-CoV-2 infection across multiple lineages including in AWOs, ALOs and CMs. We found that *CIART* regulates SARS-CoV-2 infection through an entry-independent mechanism. Loss of *CIART* blocks SARS-CoV-2 infection by downregulating the RXR pathway, at least in part through NR4A1, leading to decreased fatty-acid synthesis, thereby impairing viral infection. Using the multi-organoid platform, we identified a previously unknown role for the *CIART*-RXR axis in supporting SARS-CoV-2 infection of human primary cells/organoids. *CIART* was originally identified as a transcriptional repressor that forms a negative regulatory component of the circadian clock<sup>26–28</sup>. Circadian rhythm has been shown to be involved in many biological processes. This study highlights the potential role of the host-cell circadian rhythm on viral infection. Interestingly, studies have been performed to develop RXR modulators as drug candidates<sup>39</sup>, and in that sense, this study could aid in the advancement of strategies for the development of antiviral drugs.

## Online content

Any methods, additional references, Nature Portfolio 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/s41556-023-01095-y>.

## References

- 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).
- Han, Y. et al. Identification of SARS-CoV-2 inhibitors using lung and colonic organoids. *Nature* **589**, 270–275 (2021).
- Samuel, R. M. et al. Androgen signaling regulates SARS-CoV-2 receptor levels and is associated with severe COVID-19 symptoms in men. *Cell Stem Cell* **27**, 876–889 (2020).
- Tiwari, S. K., Wang, S., Smith, D., Carlin, A. F. & Rana, T. M. Revealing tissue-specific SARS-CoV-2 infection and host responses using human stem cell-derived lung and cerebral organoids. *Stem Cell Rep.* **16**, 437–445 (2021).
- Pei, R. et al. Host metabolism dysregulation and cell tropism identification in human airway and alveolar organoids upon SARS-CoV-2 infection. *Protein Cell* **12**, 717–733 (2021).
- Huff, S. et al. Discovery and Mechanism of SARS-CoV-2 Main Protease Inhibitors. *J. Med. Chem.* **65**, 2866–2879 (2021).
- Wang, S. et al. Cholesterol 25-hydroxylase inhibits SARS-CoV-2 and other coronaviruses by depleting membrane cholesterol. *EMBO J.* **39**, e106057 (2020).
- Duan, X. et al. An airway organoid-based screen identifies a role for the HIF1 $\alpha$ -glycolysis axis in SARS-CoV-2 infection. *Cell Rep.* **37**, 109920 (2021).
- Jacob, F. et al. Human pluripotent stem cell-derived neural cells and brain organoids reveal SARS-CoV-2 neurotropism predominates in choroid plexus epithelium. *Cell Stem Cell* **27**, 937–950 (2020).
- Wang, C. et al. ApoE-isoform-dependent SARS-CoV-2 neurotropism and cellular response. *Cell Stem Cell* **28**, 331–342 (2021).
- Song, E. et al. Neuroinvasion of SARS-CoV-2 in human and mouse brain. *J. Exp. Med.* **218**, e20202135 (2021).
- Poirier, E. Z. et al. An isoform of Dicer protects mammalian stem cells against multiple RNA viruses. *Science* **373**, 231–236 (2021).
- McMahon, C. L., Staples, H., Gazi, M., Carrion, R. & Hsieh, J. SARS-CoV-2 targets glial cells in human cortical organoids. *Stem Cell Rep.* **16**, 1156–1164 (2021).
- Sharma, A. et al. Human iPSC-derived cardiomyocytes are susceptible to SARS-CoV-2 infection. *Cell Rep. Med.* **1**, 100052 (2020).
- Wong, C. K. et al. Human-induced pluripotent stem cell-derived cardiomyocytes platform to study SARS-CoV-2 related myocardial injury. *Circ. J.* **84**, 2027–2031 (2020).
- Bojkova, D. et al. SARS-CoV-2 infects and induces cytotoxic effects in human cardiomyocytes. *Cardiovasc Res.* **116**, 2207–2215 (2020).
- Marchiano, S. et al. SARS-CoV-2 infects human pluripotent stem cell-derived cardiomyocytes, impairing electrical and mechanical function. *Stem Cell Rep.* **16**, 478–492 (2021).
- Perez-Bermejo, J. A. et al. SARS-CoV-2 infection of human iPSC-derived cardiac cells reflects cytopathic features in hearts of patients with COVID-19. *Sci. Transl. Med.* **13**, eabf7872 (2021).
- Li, Y. et al. SARS-CoV-2 induces double-stranded RNA-mediated innate immune responses in respiratory epithelial-derived cells and cardiomyocytes. *Proc. Natl Acad. Sci. USA* **118**, e2022643118 (2021).
- Mills, R. J. et al. BET inhibition blocks inflammation-induced cardiac dysfunction and SARS-CoV-2 infection. *Cell* **184**, 2167–2182 (2021).
- Yang, L. et al. Cardiomyocytes recruit monocytes upon SARS-CoV-2 infection by secreting CCL2. *Stem Cell Rep.* **16**, 2274–2288 (2021).
- Tang, X. et al. SARS-CoV-2 infection induces beta cell transdifferentiation. *Cell Metab.* **33**, 1577–1591 (2021).
- Calistri, A. et al. The new generation hDHODH inhibitor MEDS433 hinders the in vitro replication of SARS-CoV-2 and other human coronaviruses. *Microorganisms* **9**, 1731 (2021).
- Rauti, R. et al. Effect of SARS-CoV-2 proteins on vascular permeability. *eLife* **10**, e69314 (2021).
- Wagar, L. E. et al. Modeling human adaptive immune responses with tonsil organoids. *Nat. Med.* **27**, 125–135 (2021).
- Anafi, R. C. et al. Machine learning helps identify CHRONO as a circadian clock component. *PLoS Biol.* **12**, e1001840 (2014).
- Anayev, Y. et al. Gene model 129 (*Gm129*) encodes a novel transcriptional repressor that modulates circadian gene expression. *J. Biol. Chem.* **289**, 5013–5024 (2014).
- Yang, Y., Li, N., Qiu, J., Ge, H. & Qin, X. Identification of the repressive domain of the negative circadian clock component CHRONO. *Int. J. Mol. Sci.* **21**, 2469 (2020).
- Zhuang, X. et al. The circadian clock component BMAL1 regulates SARS-CoV-2 entry and replication in lung epithelial cells. *iScience* **24**, 103144 (2021).
- Sposito, B. et al. The interferon landscape along the respiratory tract impacts the severity of COVID-19. *Cell* **184**, 4953–4968 (2021).
- Ziegler, C. G. K. et al. SARS-CoV-2 receptor ACE2 is an interferon-stimulated gene in human airway epithelial cells and is detected in specific cell subsets across tissues. *Cell* **181**, 1016–1035 (2020).

32. Travaglini, K. J. et al. A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* **587**, 619–625 (2020).
33. Perlmann, T. & Jansson, L. A novel pathway for vitamin A signaling mediated by RXR heterodimerization with NGFI-B and NURR1. *Genes Dev.* **9**, 769–782 (1995).
34. Ide, T. et al. Cross-talk between peroxisome proliferator-activated receptor (PPAR)  $\alpha$  and liver X receptor (LXR) in nutritional regulation of fatty acid metabolism. II. LXRs suppress lipid degradation gene promoters through inhibition of PPAR signaling. *Mol. Endocrinol.* **17**, 1255–1267 (2003).
35. Ciaraldi, T. P. et al. Free fatty acid metabolism in human skeletal muscle is regulated by PPAR $\gamma$  and RXR agonists. *Ann. N. Y. Acad. Sci.* **967**, 66–70 (2002).
36. Chu, J. et al. Pharmacological inhibition of fatty acid synthesis blocks SARS-CoV-2 replication. *Nat. Metab.* **3**, 1466–1475 (2021).
37. Beumer, J. et al. A CRISPR/Cas9 genetically engineered organoid biobank reveals essential host factors for coronaviruses. *Nat. Commun.* **12**, 5498 (2021).
38. Hysenaj, L. et al. SARS-CoV-2 infection studies in lung organoids identify TSPAN8 as novel mediator. Preprint at *bioRxiv* <https://www.biorxiv.org/content/10.1101/2021.06.01.446640v1> (2021).
39. Schierle, S. & Merk, D. Therapeutic modulation of retinoid X receptors—SAR and therapeutic potential of RXR ligands and recent patents. *Expert Opin. Ther. Pat.* **29**, 605–621 (2019).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

---

<sup>1</sup>Department of Surgery, Weill Cornell Medicine, New York, NY, USA. <sup>2</sup>Center for Genomic Health, Weill Cornell Medicine, New York, NY, USA. <sup>3</sup>Genomics Resources Core Facility, Weill Cornell Medicine, New York, NY, USA. <sup>4</sup>Department of Microbiology, New York University, New York, NY, USA. <sup>5</sup>Stable Isotope and Metabolomics Core Facility, The Einstein-Mount Sinai Diabetes Research Center, Albert Einstein College of Medicine, Bronx, New York, USA. <sup>6</sup>Division of Gastroenterology and Hepatology, Department of Medicine, Weill Cornell Medicine, New York, NY, USA. <sup>7</sup>Department of Physiology, Biophysics and Systems Biology, Weill Cornell Medicine, New York, NY, USA. <sup>8</sup>Present address: TWINCORE Centre for Experimental and Clinical Infection Research, Hannover, Germany. <sup>9</sup>These authors contributed equally: Xuming Tang, Dongxiang Xue, Tuo Zhang, Benjamin E. Nilsson-Payant, Lucia Carrau. <sup>10</sup>These authors jointly supervised this work: Robert E. Schwartz, Benjamin R. tenOever, Todd Evans, Shuibing Chen. ✉e-mail: [shc2034@med.cornell.edu](mailto:shc2034@med.cornell.edu)

## Methods

All embryonic stem cell studies were approved by the Tri-Institutional ESCRO Committee (Weill Cornell Medicine, Memorial Sloan Kettering Cancer Center and Rockefeller University).

### Cell lines and culture conditions

H1-hESCs (WiCell, WA01) were cultured and maintained in StemFlex medium (Gibco) on 1% Matrigel-coated six-well plates. The medium was changed daily. When the H1-hESCs reached approximately 90% confluency, the cells were passaged at 1:6–1:10 with ReLeSR (Stem Cell Technology). Vero E6 cells (provided by the ATCC, CRL-1586) were cultured in DMEM medium supplemented with 2% fetal bovine serum (FBS), 4.5 g l<sup>-1</sup> D-glucose, 4 mM L-glutamine, 10 mM non-essential amino acids, 1 mM sodium pyruvate and 10 mM HEPES. HEK 293T cells (provided by the ATCC, CRL-3216) were cultured in DMEM supplemented with 10% FBS. All cell lines were cultured at 37 °C with 5% CO<sub>2</sub> and were tested for mycoplasma contamination every six months.

### Creation of isogenic H1-hESC lines

CRISPR sgRNA sequences were designed using the web resources available at <http://chopchop.cbu.uib.no/>. The target sequences are listed in Supplementary Table 1. Each target sequence was cloned into the pSpCas9(BB)-2A-Puro (PX459) V2.0 vector (Addgene, cat. no. 62988) to make the gene-targeting constructs. *CIART*-knockout lines were created based on H1-hESCs. Briefly, H1 cells were dissociated using Accutase (Innovative Cell Technologies) and electroporated ( $5 \times 10^5$  cells per sample) with 4 µg sgRNA-construct plasmids using Human Stem Cell Nucleofector solution (Lonza) following the manufacturer's instructions. The cells were then seeded into two wells of 24-well plates and cultured in StemFlex medium with 10 µM Y-27632. They were switched to StemFlex medium with 0.5 mg ml<sup>-1</sup> puromycin the following day and maintained for 2 d. After puromycin selection, hPSCs were dissociated into single cells using Accutase and re-plated at a density of 3 cells per well in 96-well plates. Y-27632 (10 µM) was added to the cells for the first 2 d. After 10 d, individual colonies were picked and re-plated into two wells of 96-well plates. One well of the cells from an individual clone was analysed by DNA sequencing. For biallelic frameshift mutants, we chose homozygous or compound heterozygous mutants. Wild-type lines from the same gene-editing experiment were included as controls to account for potential non-specific effects associated with the gene-targeting process.

### CM differentiation

To derive hPSC-CMs, we used a protocol from a previous study<sup>40</sup>, with slight modifications. The hPSCs were passaged at a density of  $3 \times 10^5$  cells per well in six-well plates and cultured in a humidified incubator with 5% CO<sub>2</sub> at 37 °C for 2 d. On day 0 of differentiation, hPSCs with around 90% confluency were treated with RPMI 1640 supplemented with B27 minus insulin (Thermo Fisher) and 6 µM CHIR99021 (Sigma) for 48 h. The medium was replaced with RPMI 1640 supplemented with B27 minus insulin on day 2. On day 3, the medium was changed to RPMI 1640 supplemented with B27 minus insulin and 5 µM XAV939 (Cayman Chemical). On day 5, the medium was switched back to RPMI 1640 supplemented with B27 minus insulin. From day 7, the medium was changed to RPMI 1640 plus normal B27 (Life Technologies); this medium was replaced every 2 d.

### ALO differentiation

The protocol used for the generation of hPSC-ALOs was slightly modified from a previous study<sup>2</sup>. Briefly, hPSCs were differentiated to definitive endoderm with RPMI 1640 supplemented with 100 ng ml<sup>-1</sup> Activin A and 3 µM CHIR99021 for 1 d, and then RPMI 1640 supplemented with 100 ng ml<sup>-1</sup> Activin A and 0.2% FBS for another 2 d. Definitive endoderm cells were further cultured in a complete serum-free differentiation medium (cSFDM) consisting of a base medium of 75% IMDM

(Thermo Fisher) and 25% Ham's F12 (Thermo Fisher) with B27 supplement plus retinoic acid (Invitrogen), N2 supplement (Invitrogen), 0.1% bovine serum albumin fraction V (Invitrogen), monothioglycerol (Sigma), GlutaMAX (Thermo Fisher) and ascorbic acid (Sigma). To induce definitive endoderm into anterior foregut endoderm, cells were cultured in cSFDM supplemented with 10 µM SB431542 (Tocris) and 2 µM dorsomorphin (Stemgent) for 3 d. The cells were then cultured for 9 d in cSFDM containing 3 µM CHIR99021 (Tocris), 10 ng ml<sup>-1</sup> recombinant human BMP4 (PeproTech) and 100 nM retinoid acid (Millipore-Sigma) to induce NKX2.1<sup>+</sup> lung progenitors. The NKX2.1<sup>+</sup> lung progenitors were enriched by sorting CD47<sup>hi</sup>CD26<sup>-</sup> cells and then resuspended in growth factor reduced Matrigel (Corning) at a density of 100,000 cells ml<sup>-1</sup> and pipetted (50 µl droplets) onto the base of tissue-culture plates. The three-dimensional culture was maintained in alveolar medium composed of cSFDM containing 3 µM CHIR99021, 10 ng ml<sup>-1</sup> recombinant human KGF (PeproTech), 50 nM dexamethasone, 100 nM 8-bromoadenosine 3',5'-cyclic monophosphate sodium salt (Millipore-Sigma) and 100 nM 3-isobutyl-1-methylxanthine (IBMX; Millipore-Sigma).

### AWO differentiation

First, hPSCs were induced to NKX2.1<sup>+</sup> lung progenitors as described earlier ('ALO differentiation' section). The NKX2.1<sup>+</sup> lung progenitors were then resuspended in growth factor reduced Matrigel (Corning) at a density of 100,000 cells ml<sup>-1</sup> and pipetted (50 µl droplets) onto the base of tissue-culture plates. The three-dimensional culture was maintained in alveolar medium composed of cSFDM containing 250 ng ml<sup>-1</sup> FGF2 (rhFGFbasic; R&D Systems), 100 ng ml<sup>-1</sup> FGF10, 50 nM dexamethasone, 100 nM 8-bromoadenosine 3',5'-cyclic monophosphate sodium salt (Millipore-Sigma), 100 nM 3-isobutyl-1-methylxanthine (Millipore-Sigma) and 10 µM Y-27632 (Tocris).

### Plasmid construction and lentivirus production

A lentiviral vector containing FLAG-tagged complementary DNA to the open reading frame of human *CIART*, pLV-Puro-EF1A >hCIART/FLAG, was purchased from VectorBuilder. For CRISPR-mediated gene knockout, sgRNA sequences (listed in Supplementary Table 1) were cloned into lentiCRISPRv2 vector (Addgene, cat. no. 52961) according to the instructions described by the Zhang laboratory ([https://media.addgene.org/cms/filer\\_public/53/09/53091cde-b1ee-47ee-97cf-9b3b05d290f2/lenticrisprv2-and-lentiguide-oligo-cloning-protocol.pdf](https://media.addgene.org/cms/filer_public/53/09/53091cde-b1ee-47ee-97cf-9b3b05d290f2/lenticrisprv2-and-lentiguide-oligo-cloning-protocol.pdf)). Lentivirus was produced in 10-cm-diameter petri dishes from HEK293T cells at 70–80% confluency that had been transfected with lentiviral plasmid and the packaging plasmids pMD2.G and psPAX2 (Addgene, cat. no. 12259 and cat. no. 12260).

### SARS-CoV-2 propagation and infection

SARS-CoV-2 isolate USA-WA1/2020 (NR-52281) was provided by the Center for Disease Control and Prevention and obtained through BEI Resources, NIAID, NIH. SARS-CoV-2 was propagated in Vero E6 cells (CRL-1586) in DMEM supplemented with 2% FBS, 4.5 g l<sup>-1</sup> D-glucose, 4 mM L-glutamine, 10 mM non-essential amino acids, 1 mM sodium pyruvate and 10 mM HEPES using a passage-2 stock of virus as described previously<sup>41</sup>. Three days after infection virus-containing supernatants were purified as described previously<sup>42</sup>. Briefly, supernatant containing propagated virus was filtered through an Amicon Ultra 15 (100 kDa) centrifugal filter (Millipore-Sigma) at approximately 3,000g for 20 min. The flow through was discarded and virus was resuspended in DMEM supplemented as described earlier. The infectious titres of SARS-CoV-2 were determined by plaque assay in Vero E6 cells in Minimum Essential Media supplemented with 2% FBS, 4 mM L-glutamine, 0.2% BSA, 10 mM HEPES, 0.12% NaHCO<sub>3</sub> and 0.7% agar. All m.o.i. values were based on the titre determined from plaque assays on Vero E6 cells. All work involving live SARS-CoV-2 was performed in the CDC and USDA-approved biosafety level 3 facilities of the Icahn School of Medicine at Mount



Sinai and NYU Langone in accordance with institutional biosafety requirements.

### scRNA-seq sample preparation

*CIART*<sup>-/-</sup> and wild-type hPSC-AWOs were infected with SARS-CoV-2 (m.o.i. = 0.1). The hPSC-AWOs were dissociated into a single-cell suspension using Accutase cell detachment solution (Innovative Cell Technologies) at 24 h.p.i. and incubated at 37 °C for approximately 5 min, followed by gentle pipetting to break apart groups of cells. The cells were then washed twice in 1×PBS and filtered using a 40-µm Flowmi cell strainer (Bel-Art Scienceware). Cell counts and viability were then determined using trypan blue staining and a Countess II automated cell counter (Thermo Fisher Scientific). Target cell inputs of 10,000 cells for each condition were then loaded into a Chromium Controller using Chromium Next GEM (Gel Bead-In Emulsion) single cell 5' library and gel bead kit v1.1 (10x Genomics) according to the manufacturer's instructions. After the generation of GEMs, cDNA synthesis and library preparation of all samples was completed using the Chromium single cell 5' library kit v1.1 (10x Genomics) according to the manufacturer's instructions.

### scRNA-seq data analysis

The 10x libraries were sequenced on an Illumina NovaSeq6000 sequencer with pair-end reads (28 base pairs (bp) for read 1 and 91 bp for read 2). The sequencing data were primarily analysed by the 10x cellranger pipeline (v5.0.0) in two steps. In the 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 UMI counts matrices. To measure both human and viral gene expression, we built a custom reference genome by integrating the SARS-CoV-2 virus genome into the 10x pre-built human reference (GRCh38 v3.0.0) using cellranger mkref. The SARS-CoV-2 virus genome (NC\_045512.2) was downloaded from NCBI.

Cell-free messenger RNA contamination in each scRNA-seq sample was estimated and removed using the R SoupX package (v1.6.1). Specifically, the cell-free mRNA expression profiles were estimated based on empty-droplet information stored in the 10x unfiltered feature barcode matrix and the contamination fraction in each cell was estimated after incorporating the clustering information produced by the cellranger pipeline using the autoEstCont function; the UMI counts matrix was then corrected to remove the contamination using the adjustCounts function.

Putative doublet cells were discarded using the R DoubletFinder package (v2.0.3), with an expected multiple rate of 0.8% per 1,000 cells according to the 10x Genomics guideline.

We filtered cells with fewer than 500 or more than 8,000 detected genes, cells with fewer than 1,000 or more than 70,000 detected UMIs as well as cells with mitochondria gene content greater than 15%, and used the remaining cells (2,851 cells for WT + mock sample, 3,337 cells for WT + SARS-CoV-2 sample, 6,408 cells for *CIART*<sup>-/-</sup> + mock sample and 4,503 cells for *CIART*<sup>-/-</sup> + SARS-CoV-2 sample) for downstream analysis. We filtered doublets using DoubletFinder.

We normalized the gene-expression UMI counts using a deconvolution strategy implemented by the R scran package (v1.14.1). In particular, we pre-clustered cells using the quickCluster function, computed size factor per cell within each cluster and rescaled the size factors by normalization between clusters using the computeSumFactors function, 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 (v1.2.1).

We identified highly variable genes using the FindVariableFeatures function in the R Seurat package (v3.1.0)<sup>43</sup> and selected the top 3,000 variable genes after excluding mitochondrial, ribosomal, viral and dissociation-related genes. The list of dissociation-related genes was

originally built on mouse data; we converted them to human orthologue genes using Ensembl BioMart. We aligned the four samples based on their mutual nearest neighbours using the fastMNN function in the R batchelor package; this was done by performing a PCA analysis on the highly variable genes and then correcting the principal components according to their mutual nearest neighbours. We selected the corrected top 50 principal components for downstream visualization and clustering analysis.

We ran UMAP dimensional reduction using the RunUMAP function in the R Seurat package with the number of neighbouring points set to 35 and training epochs set to 2,500. We clustered cells into 17 clusters by constructing a shared nearest neighbour graph and then grouping cells of similar transcriptome profiles using the FindNeighbors and FindClusters functions (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. We compared the marker genes of each cluster against a reference human adult lung dataset and performed an unsupervised hierarchical clustering on these 17 clusters based on their marker gene similarity to the reference. We then merged the 17 clusters into five clusters representing basal cells, ciliated-like cells 1, ciliated-like cells 2, proliferating basal cells and neuroendocrine cells based on the hierarchical clustering and UMAP results, and used them for downstream analysis. We generated UMAP plots illustrating the five clusters as well as highlighting expressions of selected genes using the R ggplot2 package.

To validate the identity of the cell populations in hPSC-AWOs, we extracted the marker genes of the five clusters in hPSC-AWOs with adjusted  $P < 0.01$  and average  $\log_2(\text{fold change}) > 0$ , and compared them to the marker genes of various cell types in a human adult lung dataset<sup>32</sup>. We calculated the fraction of overlapping marker genes between hPSC-AWO and human adult lung cell clusters<sup>32</sup>.

We presented the expression difference in the SARS-CoV-2 genes between WT and *CIART*<sup>-/-</sup> samples by dot plot using the DotPlot function in the R Seurat package, where the size of a dot indicates the percentage of cells that express a gene and the colour represents the relative expression level of a gene.

To evaluate viral infections of WT and *CIART*<sup>-/-</sup> samples, we calculated the percentage of viral UMI counts of WT and *CIART*<sup>-/-</sup> hPSC-AWOs under mock- or SARS-CoV-2-infection conditions. We considered a cell SARS-CoV-2<sup>+</sup> if its viral UMI count was greater than 20 and calculated the percentage of SARS-CoV-2<sup>+</sup> cells in the five cell clusters in SARS-CoV-2-infected WT and *CIART*<sup>-/-</sup> samples. The virus infection results were presented as bar and violin plots using the R ggplot2 package.

### Bulk RNA-seq data analysis

The cDNA libraries were sequenced on a NovaSeq6000 sequencer with pair-end 51 bp. The sequencing reads were checked for quality using FastQC v0.10.1 and cleaned by trimming the adaptor sequences and low-quality bases using cutadapt v1.18. To measure both human and viral gene expression, the cleaned reads were aligned to the human reference genome (GRCh37) combined with the SARS-CoV-2 genome (NC\_045512.2) using STAR aligner v.2.5.2b. Raw gene counts were quantified using HTSeq-count v0.11.2. We performed differential expression analysis on the gene counts using the R DESeq2 package v1.26.0, applied regularized logarithm transformation to the count data and used the transformed data for sample clustering. In particular, we performed PCA using the plotPCA function and performed an unsupervised hierarchical clustering using Euclidean distance in the R pheatmap package.

We screened for changes in gene expression across multiple hPSC-derived organoids following SARS-CoV-2 infection at various m.o.i. We studied nine conditions (three organoids at three m.o.i.) and pre-filtered candidate genes for each condition with base mean > 10,

adjusted  $P$  value  $< 0.05$  and  $\log_2(\text{fold change}) > 0.75$ , and searched for genes shared by at least seven conditions.

We observed high viral reads for AWOs and ALOs following SARS-CoV-2 infection, resulting in a lower count of human genes than the mock samples. To make a fair comparison, we downsampled the raw gene counts in all AWOs and ALOs so that their human gene counts were comparable.

### ATAC-seq data analysis

The ATAC-seq libraries were sequenced on a NovaSeq6000 sequencer with pair-end 51 bp. The sequencing reads were checked for quality using FastQC v0.11.9 (ref. <sup>44</sup>) and trimmed to remove adaptor sequences and low-quality bases using cutadapt v3.4 (ref. <sup>45</sup>). The trimmed reads were aligned to the human GRCh37 reference genome using Bowtie2 v2.4.4 (ref. <sup>45</sup>) with the parameters `-X 2000 -very-sensitive -k 5`. Duplicate reads were discarded using Picard v2.26.2. Peaks were identified for each replicate sample using Genrich v0.6.1 with the parameters `-j -q 0.05 -a 200.0, -e` to remove mitochondrial genome and regions not assembled into chromosomes, and `-E` to exclude 'N' homopolymers or high mappability regions in the genome. The called peaks were loaded into the R DiffBind package v3.2.1 for downstream differential binding analysis. Briefly, consensus peaks were calculated for the WT and *CIART*<sup>-/-</sup> conditions by combining peaks that overlap in at least two replicate samples in each condition, and a consensus peak set was generated by taking a union of peaks from both conditions and filtering peaks in the ENCODE blacklisted regions. Reads overlapping the consensus peak set were counted for each sample and background normalization was applied to the read counts. Differential binding sites between WT and *CIART*<sup>-/-</sup> conditions were identified with false detection rate (FDR)  $< 0.05$ .

To view the peaks in the ATAC-seq data, reads in the binary alignment map (BAM) files were shifted +4 bp and -5 bp for the positive and negative strand, respectively, using alignmentSieve from the deepTools package<sup>46</sup> v3.5.1 to account for the 9-bp duplication created by DNA repair of the nick by Tn5 transposase<sup>47</sup>. Coverage tracks in bigWig format were generated based on the shifted BAM files using bamCoverage from the deepTools package v3.5.1 (ref. <sup>46</sup>) and visualized using the R karyoploteR package v1.18.0 (ref. <sup>48</sup>).

### CUT&RUN assay

CUT&RUN was performed using CUTANA ChIC/CUT&RUN Kit (Epiccypher) according to the user manual. Briefly, differentiation day 30 hPSC-CMs were transduced with lentiviruses encoding FLAG-tagged cDNA to the open reading frame of human *CIART*. Seven days after transduction,  $1 \times 10^6$  cells per sample were washed and bound to 11  $\mu\text{l}$  of activated Concanavalin A beads (Epiccypher). The bead-bound cells were incubated with monoclonal anti-FLAG M2 (Sigma Aldrich; 1:100) at 4 °C overnight. After washing, the cells were incubated with CUTANA pAG-MNase (Epiccypher) for 10 min, targeted chromatin tagmentation was initiated by the addition of 100 mM CaCl<sub>2</sub> and allowed to proceed for 2 h at 4 °C, and then stop buffer containing 0.5 ng *Escherichia coli* spike-in DNA was added to each sample. Released chromatin fragments were purified using DNA Clean & Concentrator-5 (Zymo Research). Libraries were generated using a NEBNext ultra II DNA library prep kit and sequenced on an Illumina NovaSeq (PE-00) at the Weill Cornell Medical College Genomics Core.

### CUT&RUN data analysis

The CUT&RUN libraries were sequenced on a NovaSeq6000 sequencer with pair-end 51 bp. The sequencing reads were processed using the CUT&RUNTools package<sup>49</sup> using the default settings. Briefly, reads were adaptor trimmed using Trimmomatic v0.36 (ref. <sup>50</sup>) and an additional trimming step was performed to remove up to 6 bp adaptor from each read. Next, the reads were aligned to the hg19 genome using Bowtie2 v2.2.9 (ref. <sup>51</sup>) with the 'dovetail' settings enabled. Alignments were

further divided into  $\leq 120$ -bp and  $> 120$ -bp fractions. Alignments from the  $\leq 120$ -bp fractions were used for peak calling with MACS2 v2.1.1 (ref. <sup>52</sup>). Candidate peaks were extracted by generating a consensus peak set based on the two WT replicate samples and then subtracting peaks overlapping with the negative control sample using the R DiffBind package v3.2.7. The candidate peaks were annotated and visualized using the R ChIPseeker<sup>53</sup> package v1.28.3. To identify the downstream targets regulated by CIART in hPSC-CMs, the candidate peaks were further filtered by incorporating the ATAC-seq and bulk RNA-seq data. In particular, we selected peaks for which chromatin accessibility changes were observed in promoter regions of the corresponding genes, with FDR  $< 0.05$ , and gene expression changes were observed in the corresponding genes, with adjusted  $P < 0.1$  and absolute  $\log_2(\text{fold change}) > 1$ . To view the peaks in the CUT&RUN and ATAC-seq data, coverage tracks in bigWig format were generated based on the BAM files using bamCoverage from the deepTools<sup>46</sup> package v3.5.1 and visualized using the R karyoploteR<sup>48</sup> package v1.18.0. Although there were very small peaks at the *NR4A1*-associated region in the control CUT&RUN track, these peaks were too weak to be considered real signals.

### Immunofluorescence staining and confocal microscopy

Cells or organoids were fixed with 4% paraformaldehyde at room temperature for 30 min. The samples were then blocked and permeabilized in PBS containing 5% horse serum and 0.1% Triton X-100 for 1 h at room temperature and incubated with primary antibodies at 4 °C overnight, followed by incubation with fluorescence-conjugated secondary antibodies at room temperature for 1 h. Nuclei were counterstained with 4,6-diamidino-2-phenylindole (DAPI). Information on the antibodies used for immunofluorescence staining is provided in Supplementary Table 2. Images were taken using a Zeiss LSM 800 confocal microscope and scored using the MetaMorph image analysis software (Molecular Devices).

### Flow cytometry analysis

Intracellular staining was performed following the instructions in the user manual for the Fixation and permeabilization solution kit (BD Biosciences). Briefly, cells were dissociated and resuspended in Fixation and Permeabilization solution for 20 min at 4 °C and washed twice in 1 $\times$ Perm and Wash buffer. The fixed cells were incubated with primary antibody at 4 °C overnight, washed three times with 1 $\times$ Perm and Wash buffer, and then incubated with fluorescence-conjugated secondary antibody for 1 h at 4 °C in the dark. The cells were washed three times before flow cytometry analysis using an Attune NxT instrument and the data were processed using the FlowJo v10 software. Information on the antibodies used for flow cytometry is provided in Supplementary Table 2.

### Western blotting

Total protein was extracted from WT and *CIART*<sup>-/-</sup> H1-hESCs, hPSC-ALOs, hPSC-AWOs and hPSC-CMs using RIPA buffer (Sigma) supplemented with protease and phosphatase inhibitor cocktail (Thermo Fisher). The protein samples were loaded onto NuPAGE 4-12% bis-Tris protein gels (Thermo Fisher), resolved by electrophoresis and transferred onto nitrocellulose membranes. The membranes were incubated with the following primary antibodies: GAPDH rabbit monoclonal antibody (Cell Signaling, cat. no. 5174S; 1:1,000) and CIART polyclonal antibody (Thermo Fisher, PA5-55643; 1:1,000). The primary antibodies were detected by fluorophore-conjugated secondary donkey anti-rabbit (IRDye 800CW, 926-32213; 1:15,000). Information on the antibodies used for western blotting is provided in Supplementary Table 2.

### qRT-PCR

Total RNA samples were prepared from cells/organoids using TRIzol and a 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 qRT-PCR was performed using a 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 described previously<sup>1</sup>. The qRT-PCR reactions were performed on an Applied Biosystems QuantStudio 6 flex real-time PCR instrument. The delta-delta-cycle threshold (DDC<sub>t</sub>) was determined relative to the *ACTB* and mock infected/treated samples. The sequences of primers/probes are provided in Supplementary Table 3.

### Metabolic profiling

Samples were extracted using 200 µl of 4:1 (vol/vol) methanol:water (containing internal standards) and quickly frozen in liquid nitrogen and thawed on ice. The thawed samples were sonicated for 2 min. The freeze–thaw–sonication procedure was repeated twice. Next, the samples were placed at –20 °C for 10 min and centrifuged at 13,523g for 10 min; the supernatant was transferred to a sampling vial. The samples were injected into a liquid chromatography-coupled mass spectrometry system (LC–MS; Waters UPLC coupled with ABSciex 6500+ QTrapMS) for acyl coenzyme A analysis. Following the LC–MS analysis, the samples were dried under gentle nitrogen flow, derivatized with a two-step derivatization procedure and analysed by gas chromatography–mass spectrometry for untargeted metabolomics. The derivatization was first methoximized with 50 µl methoxyamine hydrochloride (15 mg ml<sup>–1</sup> in pyridine) at 30 °C for 90 min. The silylation step was done with 50 µl *N,O*-bis(trimethylsilyl)trifluoroacetamide (containing 1% trimethylsilyl chloride) at 70 °C for 60 min. The samples were analysed by gas chromatography mass spectrometry (8890 GC with 5977B MS; Agilent). The LC–MS data were analysed using the Mutiquant (ABSciex) software and the gas chromatography–mass spectrometry data were analysed using the MS refiner (Genedata) software.

### Statistics and reproducibility

All statistical analyses were performed using the GraphPad Prism 6 software. Data are shown as the mean ± s.d. Data distribution was assumed to be normal but this was not formally tested. For two-group data, we used a two-tailed unpaired Student's *t*-test. For one-independent-variable data, we used a one-way analysis of variance. *P* < 0.05 was considered statistically significant. No data were excluded from the analyses. No statistical method was used to pre-determine sample size but our sample sizes are similar to those reported in previous publications<sup>2,54</sup>. Samples were assigned randomly. All experiments were performed at least three independent times with similar results, unless specified otherwise in the figure legends. Descriptions of each statistical test and the *n* and *P* values are included in each legend or experimental Source Data. All investigators analysing the data were blinded to the sample name.

### Reporting summary

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

### Data availability

The RNA-seq and scRNA-seq data that support the findings of this study have been deposited in the Gene Expression Omnibus under the accession code [GSE202967](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE202967) (including [GSE202963](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE202963), [GSE202964](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE202964) and [GSE202965](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE202965)). Source data are provided with this paper. All other relevant data are available from the corresponding author on reasonable request.

### Code availability

The code used for data analysis has been deposited to GitHub ([https://github.com/shuibingchen/COVID-19\\_Multi](https://github.com/shuibingchen/COVID-19_Multi)).

## References

- Lian, X. et al. Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/β-catenin signaling under fully defined conditions. *Nat. Protoc.* **8**, 162–175 (2013).
- Blanco-Melo, D. et al. Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell* **181**, 1036–1045 (2020).
- Nilsson-Payant, B. E. et al. The NF-κB transcriptional footprint is essential for SARS-CoV-2 replication. *J. Virol.* **95**, e0125721 (2021).
- Haghverdi, L., Lun, A. T. L., Morgan, M. D. & Marioni, J. C. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat. Biotechnol.* **36**, 421–427 (2018).
- de Sena Brandine, G. & Smith, A. D. Falco: high-speed FastQC emulation for quality control of sequencing data. *F1000Res* **8**, 1874 (2019).
- Martin, M. CUTADAPT removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* <https://doi.org/10.14806/ej.171.200> (2011).
- Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
- Yan, F., Powell, D. R., Curtis, D. J. & Wong, N. C. From reads to insight: a hitchhiker's guide to ATAC-seq data analysis. *Genome Biol.* **21**, 22 (2020).
- Gel, B. & Serra, E. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* **33**, 3088–3090 (2017).
- Zhu, Q., Liu, N., Orkin, S. H. & Yuan, G. C. CUT&RUNTools: a flexible pipeline for CUT&RUN processing and footprint analysis. *Genome Biol.* **20**, 192 (2019).
- Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
- Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
- Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
- Yu, G., Wang, L. G. & He, Q. Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).
- Huang, S. X. et al. Efficient generation of lung and airway epithelial cells from human pluripotent stem cells. *Nat. Biotechnol.* **32**, 84–91 (2014).

## Acknowledgements

Fig. 1a and Extended Data Fig. 1a,c,e were created with [BioRender.com](https://www.biorender.com) with publication licences. This work was supported by the National Institute of Diabetes, Digestive and Kidney Diseases (NIDDK; grant nos R01DK130454, R01DK119667, R01DK119667-02S1 and U01 DK127777 to S.C.); Tri-Institutional Stem Cell Initiative (grant no. Tri-Sci 202005), Department of Surgery, Weill Cornell Medicine (to T.E. and S.C.); Bill and Melinda Gates Foundation (grant no. INV-037420 to S.C., T.E., R.E.S. and B.t.O.); NCI R01CA234614, NIAID 2R01AI107301 and NIDDK R01DK121072 and 1R03DK117252; and the Defense Advanced Research Projects Agency (grant no. W911NF-16-C-0050 to B.R.t.). S.C. and R.E.S. are supported as Irma Hirschl Trust Research Award Scholars.

## Author contributions

X.T., S.C., T.E., B.R.t. and R.E.S. conceived the project. X.T., D.X. and S.C. designed the experiments. X.T. and D.X. performed gene editing, hPSC differentiation, scRNA-seq, CUT&RUN, RNA-seq and ATAC-seq. X.T., D.X., X.D. and M.G. performed confocal imaging and qRT-PCR. B.E.N.P.



and L.C. performed virus infections. T.Z., A.Y.T. and J.X. performed bioinformatics analysis. Y.Q. performed metabolic profiling. S.C., X.T., D.X., T.Z. and T.E. analysed the data and wrote the manuscript.

### Competing interests

R.E.S. is on the scientific advisory board of Miromatrix Inc, and is a paid consultant and speaker for Alnylam Inc. S.C. and T.E. are the co-founders of OncoBeat, LLC. S.C. is a consultant for Vesalius Therapeutics. The remaining authors declare no competing interests.

### Additional information

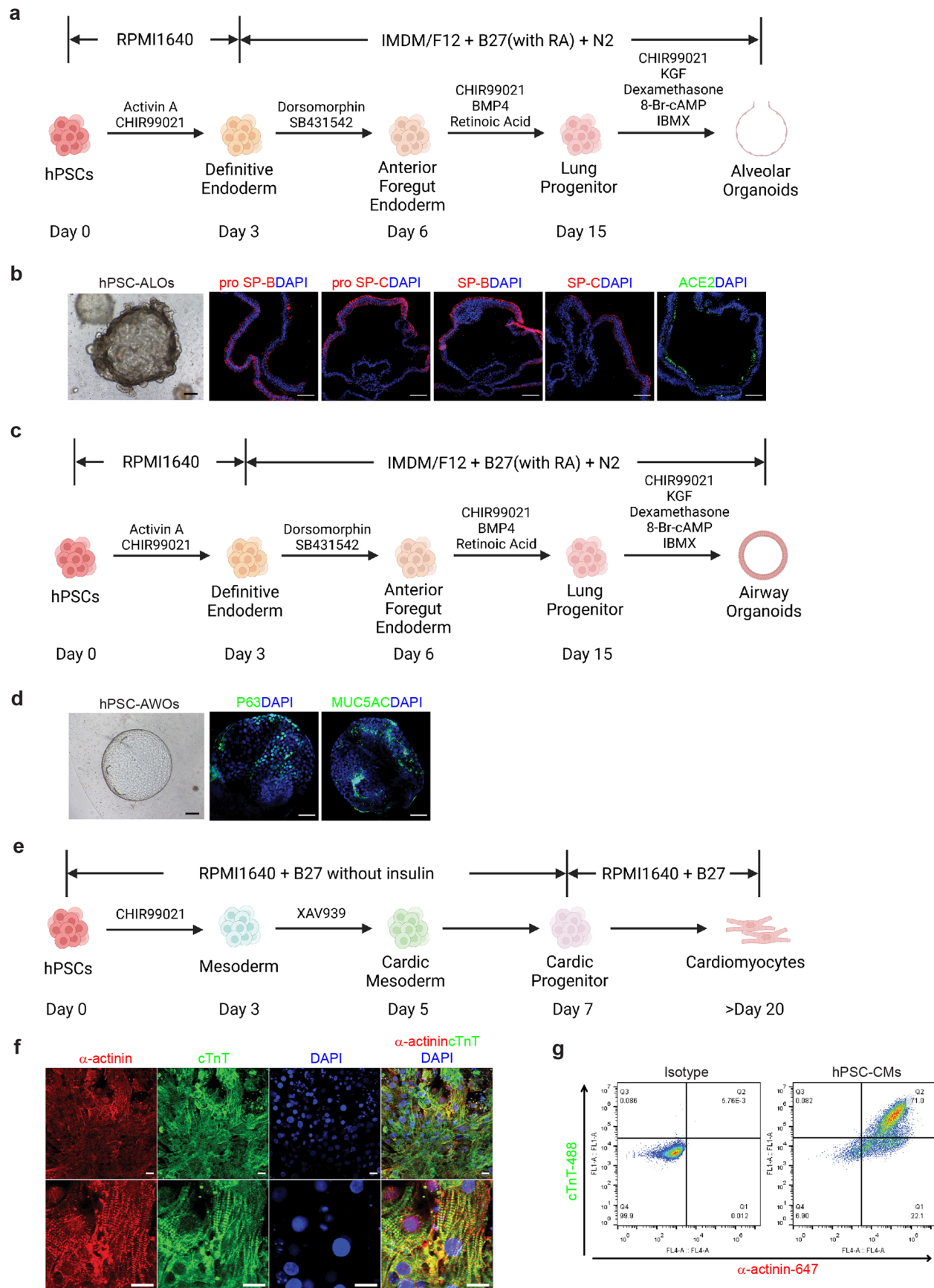
**Extended data** is available for this paper at <https://doi.org/10.1038/s41556-023-01095-y>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41556-023-01095-y>.

**Correspondence and requests for materials** should be addressed to Shuibing Chen.

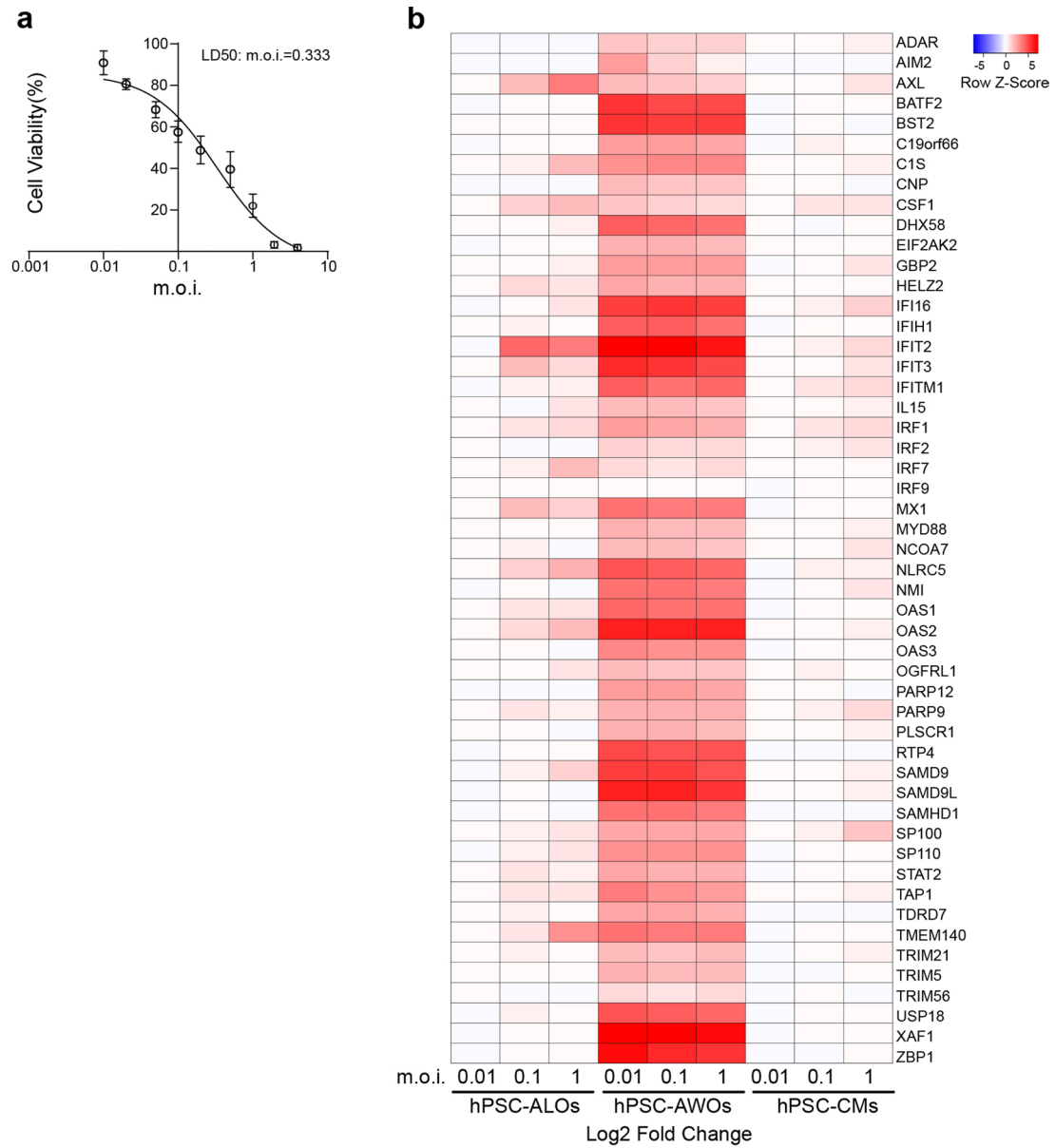
**Peer review information** *Nature Cell Biology* thanks Frank Kirchhoff, Soumita Das, and Rajarajan A. Thandavarayan for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).



**Extended Data Fig. 1 | Characterization of hPSC-ALOs, hPSC-AWOs and hPSC-CMs. a**, Scheme of hPSC-ALO differentiation. **b**, Representative confocal images of day 45 hPSC-ALOs. Scale bar, 100  $\mu$ m. **c**, Scheme of hPSC-AWO differentiation. **d**, Representative confocal images of day 40 hPSC-AWOs. Scale

bar, 100  $\mu$ m. **e**, Scheme of hPSC-CMs differentiation. **f**, Representative confocal images of day 30 hPSC-CMs. Scale bar, 100  $\mu$ m. **g**, Flow cytometry analysis of  $\alpha$ -actinin and cTnT expression in day 30 hPSC-CMs.

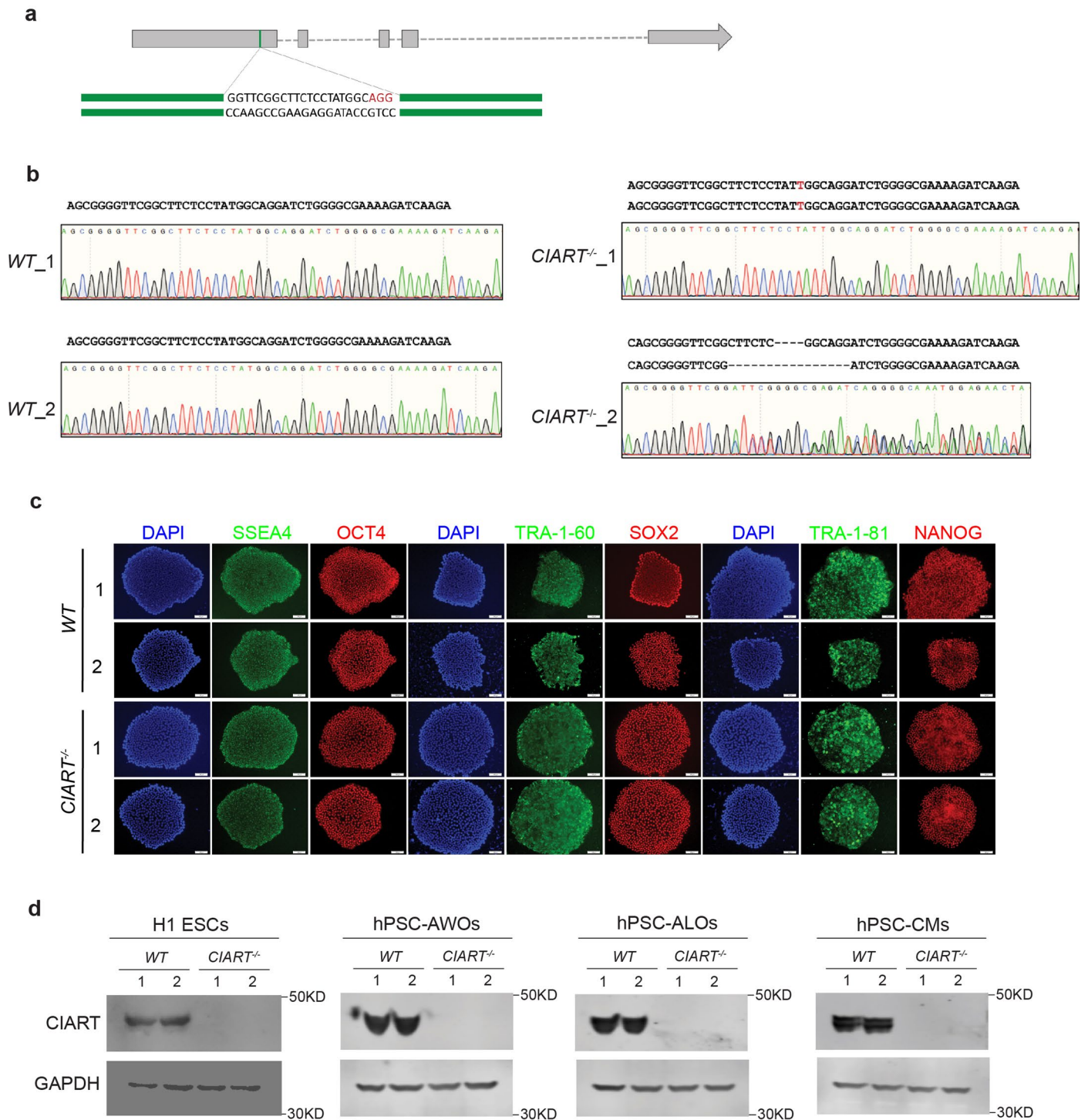


**Extended Data Fig. 2 | LD<sub>50</sub> of hPSC-CMs and heatmap of *IFN-1* genes.**

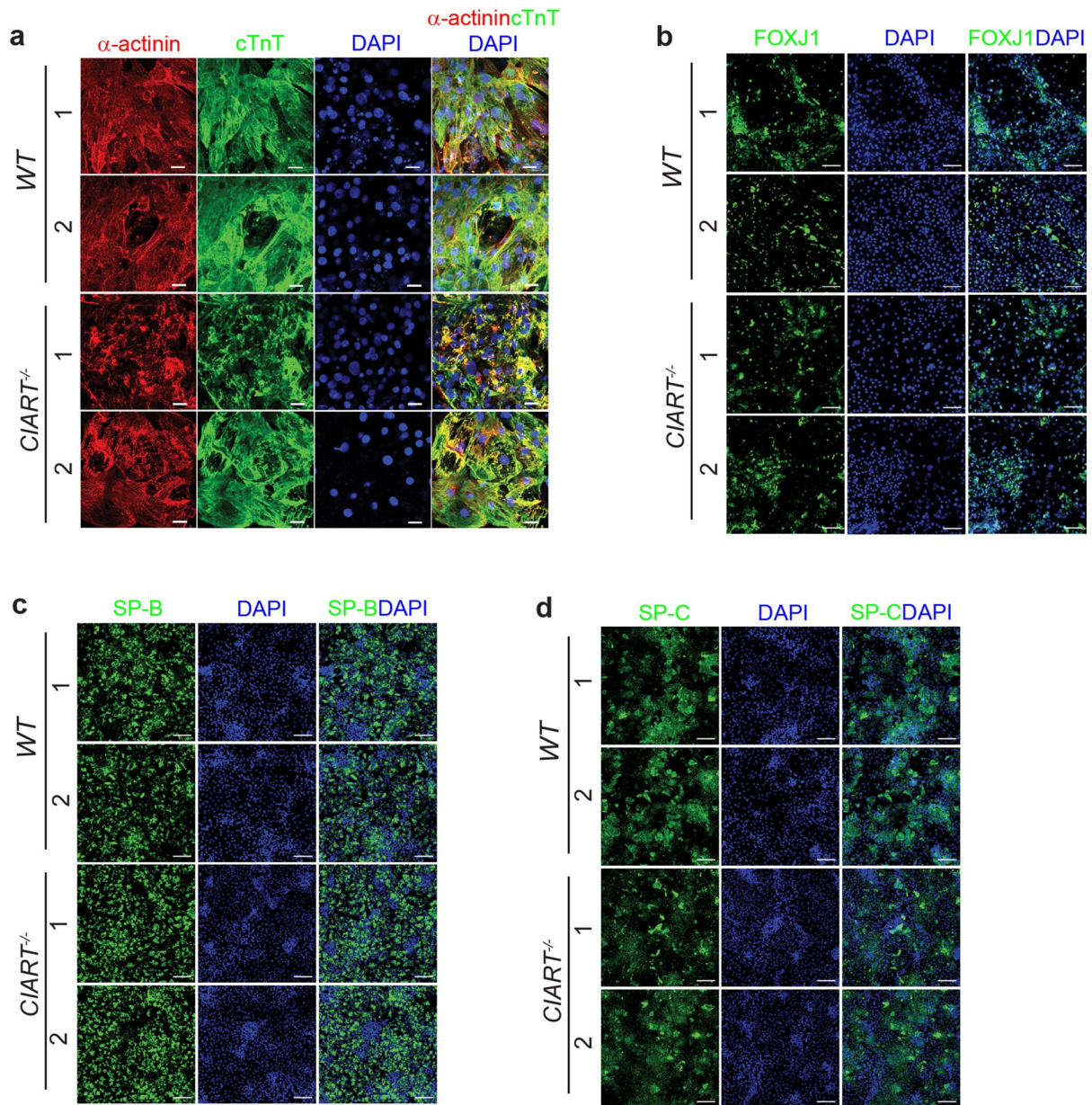
**a**, Cell viability of hPSC-CMs at 48 h.p.i. with SARS-CoV-2 at different m.o.i. LD<sub>50</sub>, m.o.i. = 0.333. *n* = 3 independent biological replicates. Data are presented as the mean ± s.d. **b**, Heatmap showing induction of genes involved in the type I

IFN signalling pathway in SARS-CoV-2-infected hPSC-ALOs, -AWOs and -CMs (m.o.i. = 0.01, 0.1 and 1). *n* = 3 independent biological replicates. Data are presented as an integration of all biological replicates.





**Extended Data Fig. 3 | Characterization of WT and CIART<sup>-/-</sup> hPSCs. a**, Scheme of gene targeting. **b**, DNA sequencing of WT and CIART<sup>-/-</sup> hPSCs. **c**, Immunostaining of pluripotency markers of WT and CIART<sup>-/-</sup> hPSCs. Scale bar, 100 μm. **d**, Western blotting for CIART in lysates from WT and CIART<sup>-/-</sup> H1 ESCs, hPSC-AWOs, hPSC-ALOs and hPSC-CMs.

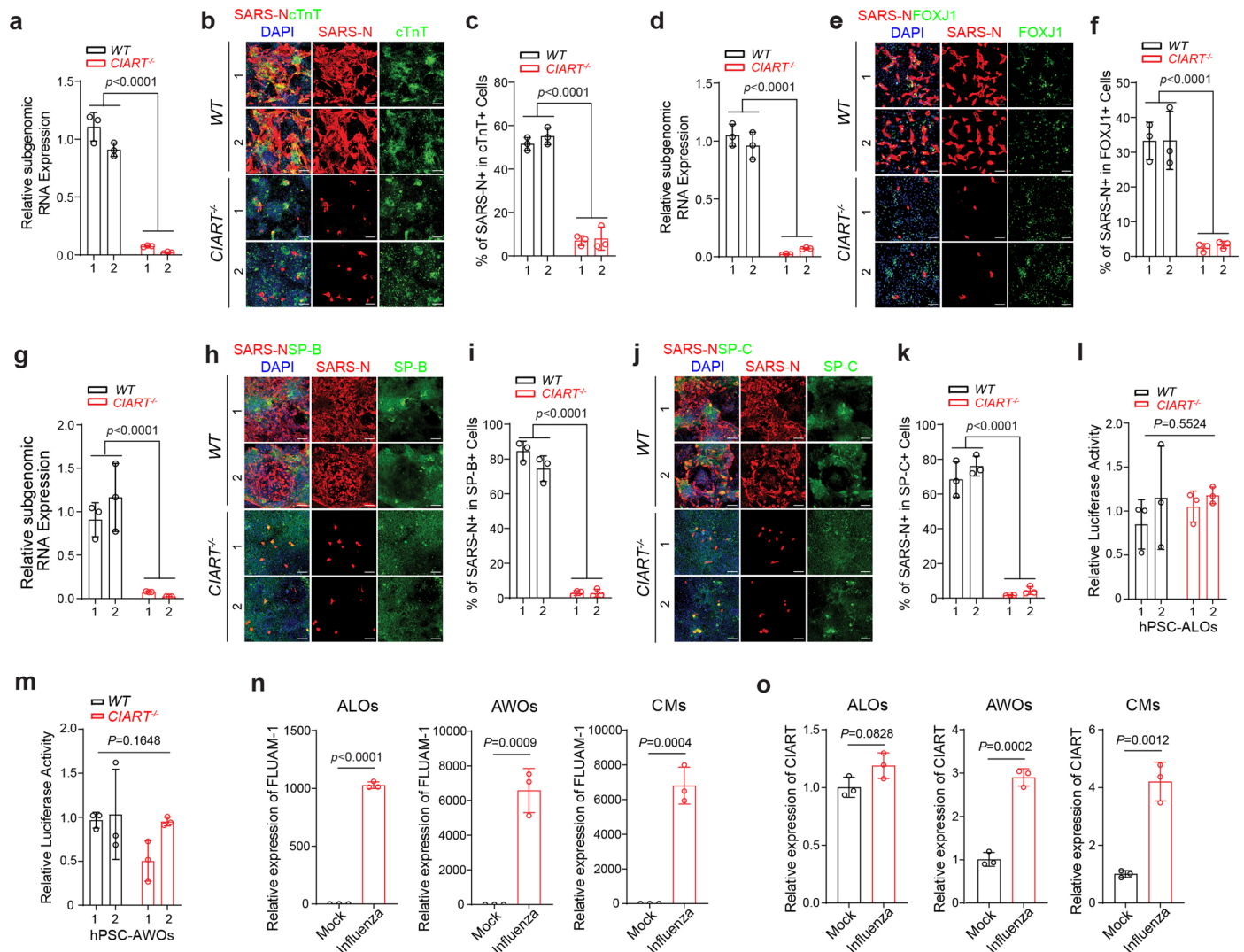


**Extended Data Fig. 4 | Differentiation of WT and *CIART*<sup>-/-</sup> hPSCs.**

**a**, Representative confocal images of  $\alpha$ -actinin<sup>+</sup>cTnT<sup>+</sup> cells of WT and *CIART*<sup>-/-</sup> hPSC-CMs. Scale bar, 100  $\mu$ m. **b**, Representative confocal images of FOXJ1<sup>+</sup> cells of WT and *CIART*<sup>-/-</sup> hPSC-AWOs. Scale bar, 100  $\mu$ m. **c**, Representative confocal

images of mature SP-B<sup>+</sup> cells of WT and *CIART*<sup>-/-</sup> hPSC-ALOs. Scale bar, 100  $\mu$ m. **d**, Representative confocal images of mature SP-C<sup>+</sup> cells of WT and *CIART*<sup>-/-</sup> hPSC-ALOs. Scale bar, 100  $\mu$ m.

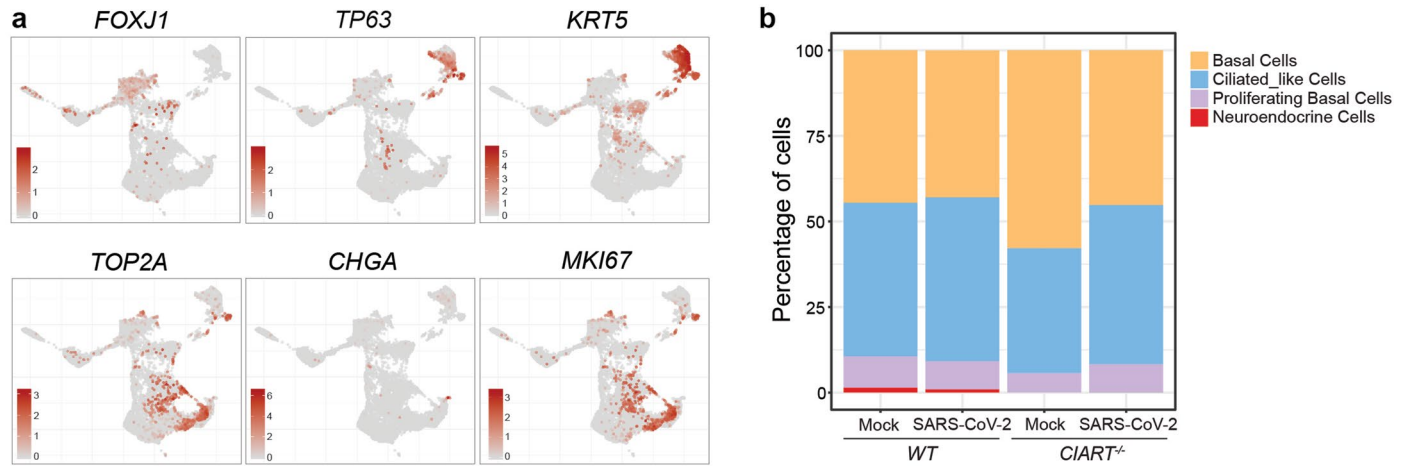




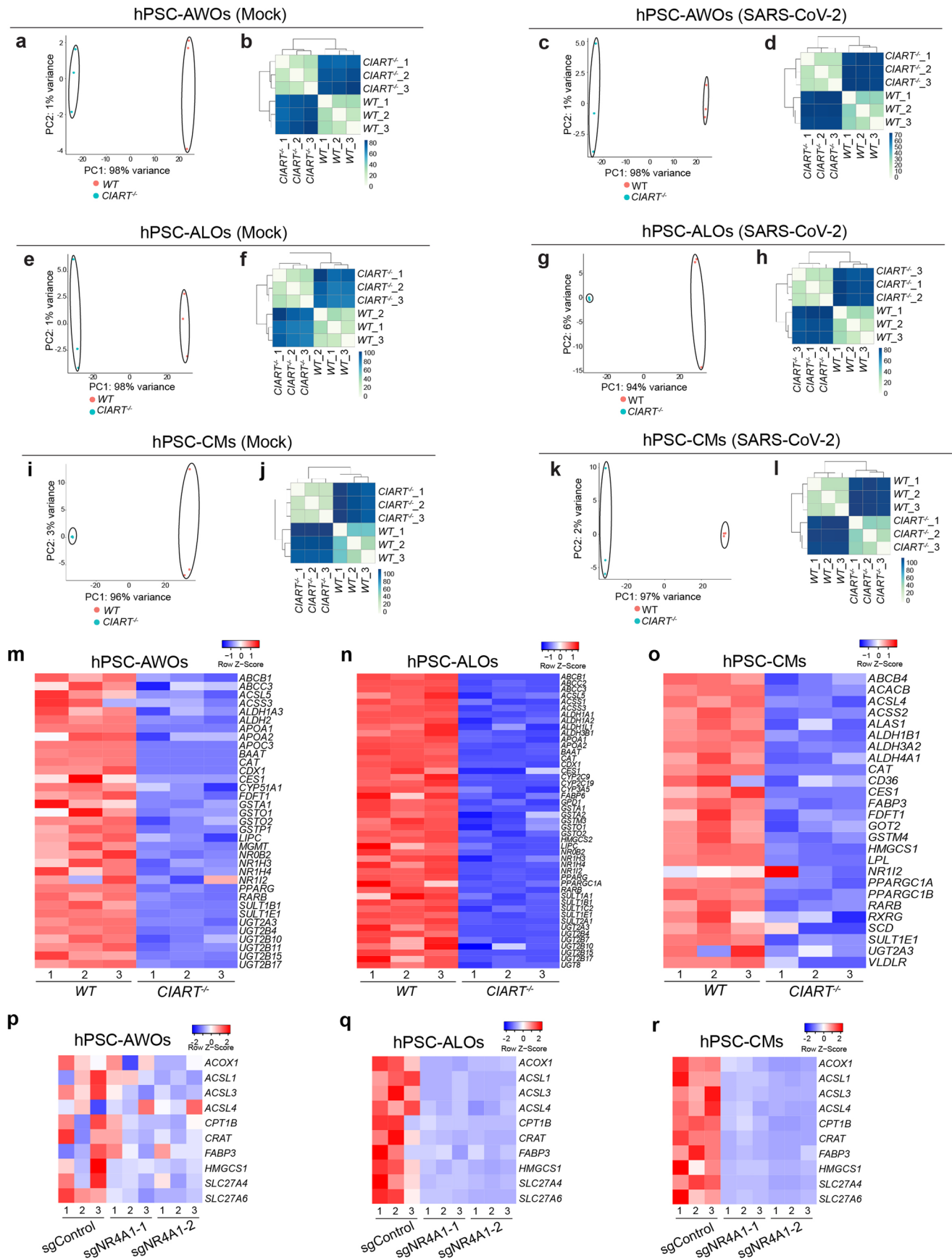
**Extended Data Fig. 5 | Loss of CIART decreases SARS-CoV-2 infection at 48 h.p.i.** **a**, Relative SARS-CoV-2 viral RNA expression levels (at 48 h.p.i.) of WT and *CIART*<sup>-/-</sup> hPSC-CMs infected with SARS-CoV-2 (m.o.i. = 0.3). *n* = 3 independent biological replicates. **b, c**, Representative confocal images (b) and quantification (c) of SARS-N<sup>+</sup> cells in cTnT<sup>+</sup> cells (at 48 h.p.i.) of WT and *CIART*<sup>-/-</sup> hPSC-CMs infected with SARS-CoV-2 virus (m.o.i. = 0.3). Scale bar, 100 μm. *n* = 3 independent biological replicates. **d**, Relative SARS-CoV-2 viral RNA expression levels at 48 h.p.i. in WT and *CIART*<sup>-/-</sup> hPSC-AWOs infected with SARS-CoV-2 (m.o.i. = 0.3). *n* = 3 independent biological replicates. **e, f**, Representative confocal images (e) and quantification (f) of SARS-N<sup>+</sup> cells in FOXJ1<sup>+</sup> cells (at 48 h.p.i.) of WT and *CIART*<sup>-/-</sup> hPSC-AWOs infected with SARS-CoV-2 (m.o.i. = 0.3). Scale bar, 100 μm. *n* = 3 independent biological replicates. **g**, Relative SARS-CoV-2 viral RNA expression levels (at 48 h.p.i.) of WT and *CIART*<sup>-/-</sup> hPSC-ALOs infected with SARS-CoV-2 (m.o.i. = 0.3). *n* = 3 independent biological replicates. **h, i**, Representative confocal images (h) and quantification (i) of SARS-N<sup>+</sup> cells in

mature SP-B<sup>+</sup> cells (at 48 h.p.i.) of WT and *CIART*<sup>-/-</sup> hPSC-ALOs infected with SARS-CoV-2 (m.o.i. = 0.3). Scale bar, 100 μm. *n* = 3 independent biological replicates. **j, k**, Representative confocal images (j) and quantification (k) of SARS-N<sup>+</sup> cells in mature SP-C<sup>+</sup> cells (at 48 h.p.i.) of WT and *CIART*<sup>-/-</sup> hPSC-ALOs infected with SARS-CoV-2 (m.o.i. = 0.3). Scale bar, 100 μm. *n* = 3 independent biological replicates. **l**, Relative luciferase activity (at 24 h.p.i.) for hPSC-ALOs infected with SARS-CoV-2 pseudo-typed entry reporter virus (m.o.i. = 0.01). *n* = 3 independent biological replicates. **m**, Relative luciferase activity (at 24 h.p.i.) for hPSC-AWOs infected with SARS-CoV-2 pseudo-typed entry reporter virus (m.o.i. = 0.01). *n* = 3 independent biological replicates. **n**, Relative influenza viral RNA expression level of mock- and influenza-infected hPSC-ALOs, -AWOs and -CMs at 48 h.p.i. (m.o.i. = 1). *n* = 3 independent biological replicates. **o**, Relative *CIART* expression level of mock- and influenza-infected hPSC-ALOs, -AWOs and -CMs at 48 h.p.i. (m.o.i. = 1). *n* = 3 independent biological replicates. Data are presented as the mean ± s.d. *P* values were calculated using an unpaired two-tailed Student's *t*-test.





**Extended Data Fig. 6 | scRNA-seq analysis of mock- and SARS-CoV-2-infected WT and *CIART*<sup>-/-</sup> hPSC-AWOs. a, UMAP of marker genes in hPSC-AWOs. b, Percentage of each cell type in mock- and SARS-CoV-2-infected WT and *CIART*<sup>-/-</sup> hPSC-AWOs.**

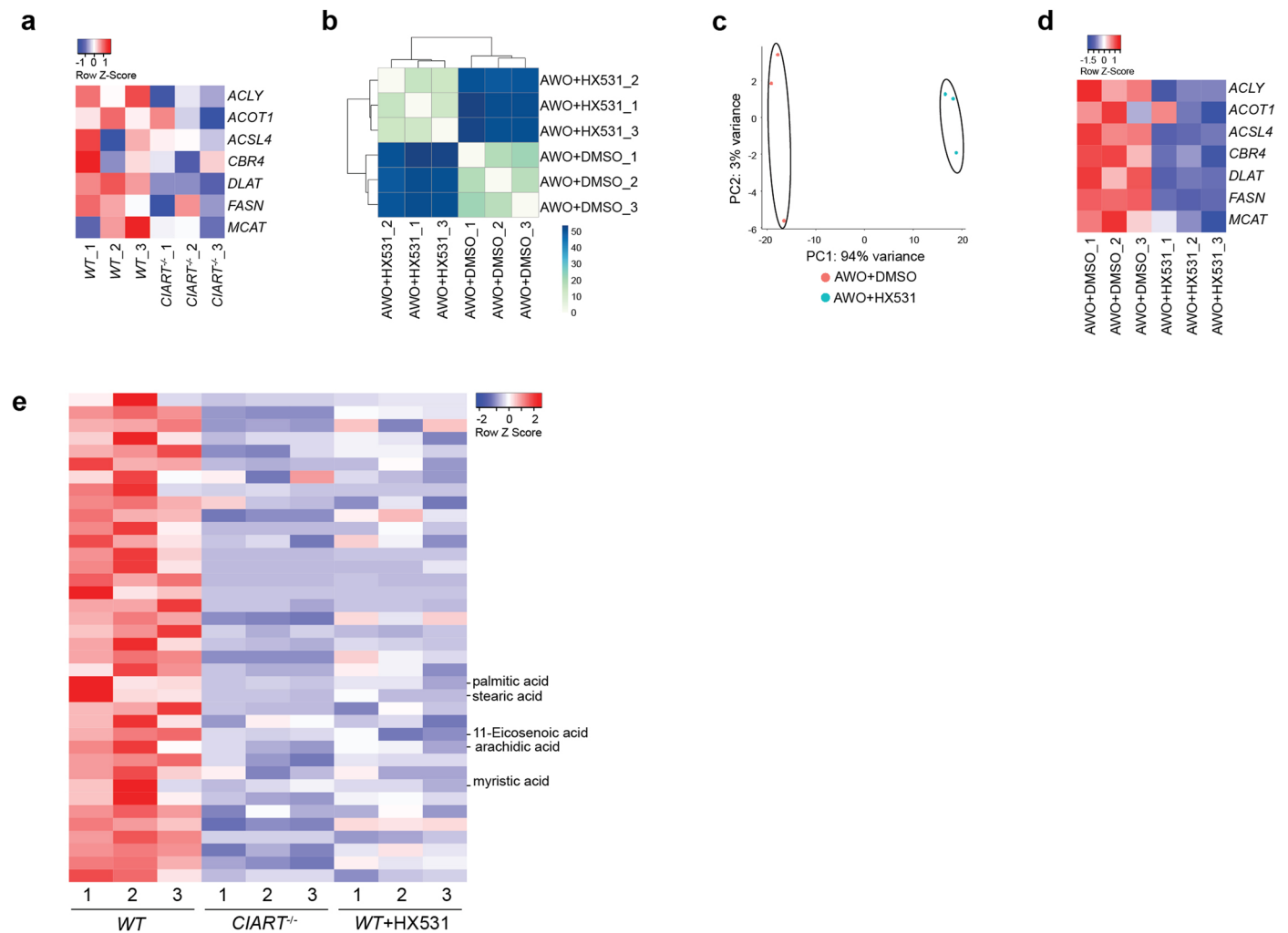


Extended Data Fig. 7 | See next page for caption.

**Extended Data Fig. 7 | RNA-seq analysis of WT and *CIART*<sup>-/-</sup> hPSC-derived AWOs, ALOs and CMs. **a,b**, PCA (**a**) and sample clustering (**b**) analysis of WT and *CIART*<sup>-/-</sup> hPSC-AWOs under mock infection conditions. **c,d**, PCA (**c**) and sample clustering (**d**) analysis of WT and *CIART*<sup>-/-</sup> hPSC-AWOs at 24 h.p.i. (m.o.i. = 0.1). **e,f**, PCA (**e**) and sample clustering (**f**) analysis of WT and *CIART*<sup>-/-</sup> hPSC-ALOs under mock infection conditions. **g,h**, PCA (**g**) and sample clustering (**h**) analysis of WT and *CIART*<sup>-/-</sup> hPSC-ALOs at 24 h.p.i. (m.o.i. = 0.1). **i,j**, PCA (**i**) and sample clustering (**j**) analysis of WT and *CIART*<sup>-/-</sup> hPSC-CMs under mock infection**

conditions. **k,l**, PCA (**k**) and sample clustering (**l**) analysis of WT and *CIART*<sup>-/-</sup> hPSC-CMs at 24 h.p.i. (m.o.i. = 0.1). **m-o**, Heatmaps of RXR pathway-associated genes in WT and *CIART*<sup>-/-</sup> hPSC-derived AWOs (**m**), ALOs (**n**) and CMs (**o**) at 24 h.p.i. (m.o.i. = 0.1). *n* = 3 independent biological replicates. Data are presented as individual biological replicates. **p-r**, Heatmap of RXR pathway-associated genes in control and sgNR4A1-infected hPSC-derived AWOs (**p**), ALOs (**q**) and CMs (**r**). *n* = 3 independent biological replicates. Data are presented as individual biological replicates.





**Extended Data Fig. 8 | RNA-seq and metabolism profiling of WT and *CIART*<sup>-/-</sup> hPSC-AWOs. **a**, Heatmap of genes involved in fatty-acid synthesis for WT and *CIART*<sup>-/-</sup> hPSC-AWOs. *n* = 3 independent biological replicates. **b**, Clustering analysis of DMSO- and HX531-treated hPSC-AWOs. *n* = 3 independent biological replicates. **c**, PCA of DMSO- and HX531-treated hPSC-AWOs. *n* = 3 independent**

biological replicates. **d**, Heatmap of genes involved in fatty-acid synthesis for DMSO- and HX531-treated WT hPSC-AWOs. *n* = 3 independent biological replicates. **e**, Heatmap of metabolic profiles for WT, *CIART*<sup>-/-</sup> and HX531-treated hPSC-AWOs. *n* = 3 independent biological replicates.

## Reporting Summary

Nature Portfolio 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 Portfolio 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

Applied Biosystems QuantStudio 6 Flex Real-Time PCR, Zeiss LSM 800 confocal microscope, 10X Chromium system, Attune NxT, NovaSeq6000 sequencer

Data analysis

Graphpad Prism 6, MetaMorph image analysis software, Flowjo v10 software, 10X cellranger, R Seurat package, R SoupX package v1.6.1, R DoubletFinder package v2.0.3, R scran package v.1.14.1, R batchelor package v1.2.1, R batchelor package, R ggplot2 package, STAR aligner v.2.5.2b, R DESeq2 package v1.26.0, R pheatmap package, FastQC v0.11.9, cutadapt v3.4, bowtie2 v2.4.4, picard v2.26.2, Genrich v0.6.1, R DiffBind package v3.2.1, deepTools package v3.5.1, R karyoploteR package v1.18.0, CUT&RUNTools package, Trimmomatic v0.36, bowtie2 v2.2.9, MACS2 v2.1.1, R DiffBind package v3.2.7, R ChIPseeker package v1.28.3

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 Portfolio [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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

RNA-seq and scRNA-seq data associated with this study have been deposited in the Gene Expression Omnibus (GEO) under accession code GSE202967 (including GSE202963, GSE202964, GSE202965).

## 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](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

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

Sample size	Sample size was determined by previous studies. (Nature 589 (7841), 270-275; Cell metabolism 33 (8), 1577-1591; Cell stem cell 27 (1), 125-136)
Data exclusions	No exclusions
Replication	All attempts at replication are successful and have been included in the studies.
Randomization	Cells/organoids were randomly separated into different groups for DMSO vehicle or RXR inhibitors treatment.
Blinding	The investigators who performed SARS-CoV-2 infection were blinded to the group allocation. The investigators who prepared cells/organoids for SARS-CoV-2 infection were not blinded to group allocation because it was impossible.

## 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
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

## Antibodies

Antibodies used

Immunostaining SARS-CoV-2 Nucleocapsid Antibody 40143-R001 #R001 Rabbit Sino Biological 1:1,000  
 Immunostaining SARS-CoV-2 Nucleocapsid Antibody 40143-MM08 #MM08 Mouse Sino Biological 1:1,000  
 Immunostaining Anti-Cardiac Troponin T antibody ab8295 #1C11 Mouse ABCAM 1:1,000  
 Immunostaining Anti-Sarcomeric Alpha Actinin antibody ab137346 Polyclonal Rabbit ABCAM 1:500  
 Immunostaining Anti-Mucin 5AC antibody ab3649 #45M1 Mouse ABCAM 1:500  
 Immunostaining Anti-P63 CM163A #4A4 Mouse Biocare Medical 1:500  
 Immunostaining FOXJ1 Monoclonal Antibody 14-9965-82 #2A5 Mouse Thermo Fisher Scientific 1:500  
 Immunostaining Anti-Pro-SP-B WRAB-55522 Polyclonal Rabbit Seven Hills Bioreagents 1:500  
 Immunostaining Anti-Mature SP-B WRAB-48604 Polyclonal Rabbit Seven Hills Bioreagents 1:500  
 Immunostaining Anti-Pro-SP-C WRAB-9337 Polyclonal Rabbit Seven Hills Bioreagents 1:500  
 Immunostaining Anti-Mature-SP-C WRAB-76694 Polyclonal Rabbit Seven Hills Bioreagents 1:500  
 Immunostaining Human/Mouse/Rat/Hamster ACE-2 Antibody AF933 Polyclonal Goat R&D Systems 1:500  
 Immunostaining Oct-4A (C30A3) Rabbit mAb 2840 #C30A3 Rabbit Cell Signaling Technologies 1:500  
 Immunostaining Sox2 (D6D9) Rabbit mAb 3579 #D6D9 Rabbit Cell Signaling Technologies 1:500  
 Immunostaining Nanog (D73G4) XP® Rabbit mAb 4903 #D73G4 Rabbit Cell Signaling Technologies 1:500  
 Immunostaining SSEA4 (MC813) Mouse mAb 4755 #MC813 Mouse Cell Signaling Technologies 1:500  
 Immunostaining TRA-1-60(S) Mouse mAb 4746 #TRA-1-60(S) Mouse Cell Signaling Technologies 1:500  
 Immunostaining TRA-1-81 Mouse mAb 4745 #TRA-1-81 Mouse Cell Signaling Technologies 1:500  
 Immunostaining Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 A-21202 Polyclonal Donkey Thermo Fisher Scientific 1:1,000  
 Immunostaining Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 594 A-21203 Polyclonal Donkey Thermo Fisher Scientific 1:1,000  
 Immunostaining Donkey anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor™ Plus 488 A32790 Polyclonal



Donkey Thermo Fisher Scientific 1:1,000  
 Immunostaining Donkey anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor 594 conjugate A-21207 Polyclonal Donkey Thermo Fisher Scientific 1:1,000  
 Immunostaining Donkey anti-Goat IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 A-11055 Polyclonal Donkey Thermo Fisher Scientific 1:1,000  
 FACS FITC Mouse Anti-Human CD47 556045 #B6H12 Mouse BD Biosciences 1:200  
 FACS APC Mouse Anti-Human CD26 563670 #M-A261 Mouse BD Biosciences 1:200  
 Flow cytometry Anti-Cardiac Troponin T antibody ab8295 #1C11 Mouse ABCAM 1:1,000  
 Flow cytometry Anti-Sarcomeric Alpha Actinin antibody ab137346 Polyclonal Rabbit ABCAM 1:500  
 Flow cytometry Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 A-21202 Polyclonal Donkey Thermo Fisher Scientific 1:1,000  
 Flow cytometry Donkey anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor 647 conjugate A-31573 Polyclonal Donkey Thermo Fisher Scientific 1:1,000  
 CUT&RUN Anti-FLAG® M2 antibody F1804 #M2 Mouse Sigma Aldrich 1:100  
 CUT&RUN Recombinant mouse IgG1 monoclonal isotype control ab280974 #R312-MouseIgG1 Mouse ABCAM 1:100  
 Western Blot GAPDH (D16H11) XP® Rabbit mAb 5174S #D16H11 Rabbit Cell Signaling Technologies 1:1,000  
 Western Blot CIART Antibody NBP1-88685 Polyclonal Rabbit NOVUS Biologicals 1:500  
 Western Blot IRDye® 800CW Donkey anti-Rabbit IgG Secondary Antibody 926-32213 Polyclonal Donkey LI-COR 1:15,000

## Validation

The antibodies were validated by either western blotting or immunostaining by vendors or previously published paper. please refer to the vendor website or previously published paper as listed after each antibody below for the detailed validation conditions and results.

SARS-CoV-2 Nucleocapsid Antibody Rabbit 40143-R001 Sino Biological <https://www.sinobiological.com/antibodies/cov-nucleocapsid-40143-r001>  
 SARS-CoV-2 Nucleocapsid Antibody Mouse 40143-MM08 Sino Biological <https://www.sinobiological.com/antibodies/cov-nucleocapsid-40143-mm08>  
 Anti-Cardiac Troponin T antibody Mouse ab8295 ABCAM <https://www.abcam.com/cardiac-troponin-t-antibody-1c11-ab8295.html>  
 Anti-Sarcomeric Alpha Actinin antibody Rabbit ab137346 ABCAM <https://www.abcam.com/sarcomeric-alpha-actinin-antibody-ab137346.html>  
 Anti-Mucin 5AC antibody Mouse ab3649 ABCAM <https://www.abcam.com/mucin-5ac-antibody-45m1-ab3649.html>  
 Anti-P63 Mouse CM163A Biocare Medical <https://biocare.net/product/p63-antibody/>  
 FOXJ1 Monoclonal Antibody Mouse 14-9965-82 Thermo Fisher Scientific <https://www.thermofisher.com/antibody/product/FOXJ1-Antibody-clone-2A5-Monoclonal/14-9965-95>  
 Anti-Pro-SP-B Rabbit WRAB-55522 Seven Hills Bioreagents <https://www.sevenhillsbioreagents.com/products/anti-pro-sp-b-rabbit>  
 Anti-Mature SP-B Rabbit WRAB-48604 Seven Hills Bioreagents <https://www.nature.com/articles/s41586-020-2901-9#Sec34>  
 Anti-Pro-SP-C Rabbit WRAB-9337 Seven Hills Bioreagents <https://www.sevenhillsbioreagents.com/products/anti-pro-sp-c-rabbit-terminal>  
 Anti-Mature-SP-C Rabbit WRAB-76694 Seven Hills Bioreagents <https://www.sevenhillsbioreagents.com/products/anti-mature-sp-c-rabbit>  
 Human/Mouse/Rat/Hamster ACE-2 Antibody Goat AF933 R&D Systems [https://www.rndsystems.com/products/human-mouse-rat-hamster-ace-2-antibody\\_af933](https://www.rndsystems.com/products/human-mouse-rat-hamster-ace-2-antibody_af933)  
 Oct-4A (C30A3) Rabbit mAb Rabbit 2840 Cell Signaling Technologies <https://www.cellsignal.com/products/primary-antibodies/oct-4a-c30a3-rabbit-mab/2840>  
 Sox2 (D6D9) Rabbit mAb Rabbit 3579 Cell Signaling Technologies <https://www.cellsignal.com/products/primary-antibodies/sox2-d6d9-xp-rabbit-mab/3579>  
 Nanog (D73G4) XP® Rabbit mAb Rabbit 4903 Cell Signaling Technologies <https://www.cellsignal.com/products/primary-antibodies/nanog-d73g4-xp-rabbit-mab/4903>  
 SSEA4 (MC813) Mouse mAb Mouse 4755 Cell Signaling Technologies <https://www.cellsignal.com/products/primary-antibodies/ssea4-mc813-mouse-mab/4755>  
 TRA-1-60(S) Mouse mAb Mouse 4746 Cell Signaling Technologies <https://www.cellsignal.com/products/primary-antibodies/tra-1-60-s-tra-1-60-s-mouse-mab/4746>  
 TRA-1-81 Mouse mAb Mouse 4745 Cell Signaling Technologies <https://www.cellsignal.com/products/primary-antibodies/tra-1-81-tra-1-81-mouse-mab/4745>  
 Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 Donkey A-21202 Thermo Fisher Scientific <https://www.thermofisher.com/antibody/product/Donkey-anti-Mouse-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21202>  
 Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 594 Donkey A-21203 Thermo Fisher Scientific <https://www.thermofisher.com/antibody/product/Donkey-anti-Mouse-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21203>  
 Donkey anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor™ Plus 488 Donkey A32790 Thermo Fisher Scientific <https://www.thermofisher.com/antibody/product/Donkey-anti-Rabbit-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A32790>  
 Donkey anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor 594 conjugate Donkey A-21207 Thermo Fisher Scientific <https://www.thermofisher.com/antibody/product/Donkey-anti-Rabbit-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21207>  
 Donkey anti-Goat IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 Donkey A-11055 Thermo Fisher Scientific <https://www.thermofisher.com/antibody/product/Donkey-anti-Goat-IgG-H-L-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-11055>  
 FITC Mouse Anti-Human CD47 Mouse 556045 BD Biosciences <https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/fic-mouse-anti-human-cd47.556045>  
 APC Mouse Anti-Human CD26 Mouse 563670 BD Biosciences <https://www.bdbiosciences.com/en-us/products/reagents/flow-cytometry-reagents/research-reagents/single-color-antibodies-ruo/apc-mouse-anti-human-cd26.563670>  
 Anti-Sarcomeric Alpha Actinin antibody Rabbit ab137346 ABCAM <https://www.abcam.com/sarcomeric-alpha-actinin-antibody-ab137346.html>  
 Donkey anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor 647 conjugate Donkey A-31573 Thermo Fisher Scientific <https://www.thermofisher.com/antibody/product/Donkey-anti-Rabbit-IgG-H-L-Highly-Cross-Adsorbed-Secondary-Antibody-Polyclonal/>

A-31573

Anti-FLAG® M2 antibody Mouse F1804 Sigma Aldrich <https://www.sigmaaldrich.com/US/en/product/sigma/f1804>Recombinant mouse IgG1 monoclonal isotype control Mouse ab280974 ABCAM <https://www.abcam.com/mouse-igg1-monoclonal-r312-mouseigg1-isotype-control-ab280974.html>GAPDH (D16H11) XP® Rabbit mAb Rabbit 5174S Cell Signaling Technologies <https://www.cellsignal.com/products/primary-antibodies/gapdh-d16h11-xp-rabbit-mab/5174>CIART Antibody Rabbit NBP1-88685 NOVUS Biologicals [https://www.novusbio.com/products/ciart-antibody\\_nbp1-88685](https://www.novusbio.com/products/ciart-antibody_nbp1-88685)IRDye® 800CW Donkey anti-Rabbit IgG Secondary Antibody Donkey 926-32213 LI-COR <https://www.licor.com/bio/reagents/irdye-800cw-donkey-anti-rabbit-igg-secondary-antibody>

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)	hESC line H1 (WiCell, WA01); Vero E6 (ATCC #CRL-1586); 293T (ATCC #CRL3216)
Authentication	H1 cells were authenticated by morphology and immunostaining, Vero E6 cells and 293T cells were authenticated by morphology
Mycoplasma contamination	All cell lines tested negative for mycoplasma contamination
Commonly misidentified lines (See <a href="#">ICLAC</a> register)	No

## 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	Flow cytometry intracellular staining was performed following the instruction of user manual of Fixation and Permeabilization Solution Kit (BD Biosciences). Briefly, cells were dissociated and resuspended in Fixation and Permeabilization solution for 20 minutes at 4°C, then washed twice in 1x Perm and Wash buffer. Fixed cells were incubated with primary antibody at 4°C overnight, washed three times with 1x Perm and Wash buffer and then incubated with fluorescence-conjugated secondary antibody for 1 hour at 4°C in the dark. Cells were washed three times prior to flow cytometry analysis.
Instrument	Accuri C6 Flow Cytometer
Software	Flowjo v10 software
Cell population abundance	Live cell populations range from 50%~75%, cTnT and a-actinin double cells account for ~75% of the live cell population.
Gating strategy	Live cell population was determined by FSC/SSC. Positive stained cells were defined by using non-stained cells as reference. Isotype control was used to confirm the specificity of staining.

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