nature research

Corresponding author(s): Sumit K. Chanda

Last updated by author(s): Jul 5, 2020

Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection	No software was used
Data analysis	The following softwares and websites were used for the analyses in this study: STAR 2.7.3a , R 3.6.3, DESeq2 1.26.0, fgsea 1.12.0, GSEA 4.0.3, Graph Pad Prism 8.4.1, CortellisTM (Clarivate Analytics) and drugs.com websites (analyses ran between April and June 2020), Metascape website (analysis ran in April 2020)

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/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The following data availability statement was added to the manuscript:

Data is available in the supplemental Tables S1, S2, S3 and S4, and through https://reframedb.org (assay A00440). Complete sequences of SARS-CoV-2 HKU-001a and SARS-CoV-2 USA-WA1/2020 are available through GenBank (accession number MT230904, and MT246667 and MN908947 respectively). RNAseq data in supplementary Table S2 were aligned with the genome of the African green monkey (Chlorocebus sabaeus, https://uswest.ensembl.org/Chlorocebus_sabaeus/Info/Annotation), and with the SARS-CoV-2 genome (https://www.ncbi.nlm.nih.gov/nuccore/NC_045512), selected as the reference genome. The dataset is available on GEO with accession number GSE153940. Figure ED3 derived from the analysis of publicly available single-cell RNA-seq dataset accessible at https://singlecell.broadinstitute.org/single_cell/study/SCP867/hca-lungmap-covid-19-barbry-lung?scpbr=hca-covid-19-integrated-analysis44. Gene expression analysis on human data shown in supplementary Table S2 (GSEA_Mason's paper) refers to the RNA-seq dataset available at https://www.biorxiv.org/

Field-specific reporting

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

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

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

Life sciences study design

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

Sample size	Vero E6 are a cell line. A Pearson's R-squared coefficient was calculated to assess the sample sizing was acceptable. The R-squared correlation coefficient >0.65, obtained from both the LOPAC and ReFRAME primary screens performed in duplicate, supports a good correlation for n=2. For dose-responses, Vero E6, HEK-293T and Huh-7 cell lines were used and n= at least 3 independent experiments were performed. In order to increase the statistical reliability for calculating EC50s, for some of the cell lines up to n=5 independent experiments were performed (specified in the figure legends). For hPSC-derived pneumocytes and ex vivo tissue lung samples, experiments were performed with n=3 (except for MTT assay in hPSC treated with the control remdesivir, where n=2 was used). The sample size for each experiment is specified in each corresponding figure legend.
Data exclusions	No data has been excluded from the analyses presented in this manuscript.
Replication	In order to verify the reproducibility of the experimental findings, the large-scale CPE screen was performed in duplicate and the R-squared correlation coefficient is indicated for both LOPAC and ReFRAME screens. The validation screens and all the confirmation studies in human cell models were performed at least in duplicate and the means +/- SEM or SD as well as the nature of 'n' are indicated in the figure legends. All the attempts were successful and no data was excluded from the analyses.
Randomization	The position of each compounds into the plates was randomly chosen for both the primary screen and the further validations.
Blinding	Data collection for the primary screen and the validation in Vero E6, 293T-ACE2 and Huh-7-ACE2 was run blinded, without knowing which compound was spot in the specific wells. Data analysis was also blinded, with the investigators only having internal IDs of compounds, without knowing their identity. For small-scale experiments (time-of-addition, VLPs, iPSC-derived cells, ex vivo tissues) the blinding was not relevant since the endpoint was determined independently from the final readout.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

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

Methods

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

Antibodies

Antibodies used	Rabbit-anti-SARS-CoV-1 nucleoprotein serum, which exhibits strong cross-reactivity with SARS-CoV-2 (A-G Sastre, unpublished data)(1:10,000 dilution). Alexa Fluor 488-conjugated goat-anti-rabbit IgG (Thermo Fisher Scientific, USA; Catalog # A27034) (1:2,000 dilution).
Validation	The antibody was tested for cross-reactivity with SARS-CoV2 in Vero E6 cells. The antibody showed specificity to SARS-CoV-2-infected cells and no background in non-infected cells.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Vero E6 and HEK-293T cells were obtained from ATCC (ATCC® CRL-1586 and ATCC®CRL-3216 TM respectively). Huh-7 were obtained from Apath LLC and BHK-21/WI-2 cells from Kerafast. hPSC1 cells (H9) were obtained from WiCell. hPSC2 (Lis38-derived) were kindly provided by Dr. Jacob Hanna(ISM ESCRO Project #14-005).
Authentication	All the cell lines were commercially available and have not been authenticated after receiving them.
Mycoplasma contamination	All cells were tested negative for mycoplasma contamination, except for Huh7-Ace2 cells which were tested positive.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in this study.

Human research participants

Policy information about studies involving human research participants

Population characteristics	This doesn't apply to this study, since tissues from only one donor were used.
Recruitment	No patients were recruited for this study. Biopsy samples that would have been otherwise discarded were used for experimental analyses.
Ethics oversight	The donors gave written consent as approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (UW13-364).

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

Flow Cytometry

Plots

Confirm that:

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

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

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

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

Methodology

Sample preparation	Cells were detached from the culture plate using 1ml of enzyme-free dissociation buffer (Sigma) and fixed by adding 1ml of 10% formaldehyde for 24h at room temperature. Cells were washed once by with Perm/Wash buffer (BD) and stained for the SARS N with in house produced mouse IgG2a monoclonal antibody conjugated to AF647. After 1h incubation at room temperature, cells were washed in phosphate buffered saline supplemented with 2mM EDTA once and resuspended in 200 microliter for analysis.
Instrument	Beckman Coulter Gallios
Software	Kaluza software version 1.0
Cell population abundance	flow cytometry to quantitate virus-infected cells but did not sort them. therefore the question for post sort abundance is actually irrelevant.
Gating strategy	Cells were first gated on singlets by plotting FSC-A versus FSC-H. All single cells were selected in a plot in which FSC-A was plotted versus SSC-A. Within this gate, virus-positive cells were quantified by plotting SSC-A versus SARS-N AF647.

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