nature research

Corresponding author(s): Leopoldo Segal

Last updated by author(s): Jul 9, 2021

Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection	All data collection instruments and analytical pipelines are detailed in the methods and in the description of the data analyses available below.
Data analysis	Available at https://github.com/segalmicrobiomelab/SARS_CoV2

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

Data

Policy information about availability of data

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

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

Sequencing data are available in NCBI's Sequence Read Archive under project numbers PRJNA688510 and PRJNA687506 (RNA and DNA sequencing, respectively).

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	The sample size was based on data available on a cohort of critically ill patients during the first surge of the COVID-19 pandemic where lower airway samples were collected. This is a convenience sample dataset and the largest cohort of patients with lower airway samples published with virome, metatranscriptome and metagenome data.
Data exclusions	There was no data or data point obtained that was excluded
Replication	In the paper we described multiple different methods used to replicate the results for key findings, such as the use of SARS-CoV-2 genomic RNA detection, subgenomic SARS-CoV-2 detection and RNA virome sequencing. Technical replicates were also used and indicated.
Randomization	All experiments were run using multiplexing approaches for what all samples were run at the same time
Blinding	None of the investigators performing the experiments and obtaining data were aware of the outcomes analyzed. Investigators were blinded to group allocation during data collection and/or upstream analysis. Downstream analysis including evaluation of variables associated with clinical outcome require unblinding that outcome

Reporting for specific materials, systems and methods

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

Materials & experimental systems

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

Methods

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

Antibodies

Antibodies used	Described in the methods in page 39: SAv (Streptavidin) (Sartorius 90792) that had been loaded with biotinylated Spike, biotinylated RBD or biotin (negative control) in wells of a 96 well HTS filter plate (MSHVN4550). As positive controls, we used CR3022 antibody, that recognizes SARS-CoV-2 Spike and RBD, in human IgG, IgA and IgM formats (Absolute Antibody; dilutions 1:1120, 1:1300 and 1:258, respectively). After washing the beads, bound antibodies were labeled with anti IgG-DyLight488, anti IgA-PE and anti IgM-PECy7, and the fluorescence intensities were measured in Intellicyt IQue3 (Sartorius)
Validation	Described in the methods

Human research participants

Policy information about studies involving human research participants

Population characteristics	The population characteristics collected and reported are extensive. These can be find in Table 1, Supplementary Tables 1-3
Recruitment	We included all subjects where lower airway samples were obtained during the first wave of the COVID-19 pandemic. As discussed in the manuscript, the presented data from lower airway samples are restricted to those subjects for whom bronchoscopy was performed as part of their clinical care. Thus, the culture independent data is biased towards patients that, while critically ill with COVID-19, may not be representative of the extremes in the spectrum of disease severity.

Ethics oversight	The study protocol was approved by the Institutional Review Board of New York University

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