nature research

Corresponding author(s): Carolina Lopez

Last updated by author(s): 02/09/2021

Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection	no software was used to collect the data.
Data analysis	Statistical analyses were performed as indicated in each figure using GraphPad Prism Version 9 or using the R environment (R package, version 3.6.0, tidyverse, tximport, biomaRt, reshape2, genefilter, edgeR, matrixStats, ggplot, limma, statmod, RColorBrewer, DEseq2). All R codes were summarized in R markdown file (supplementary information). VODKA: self-developed, deposit in GitHub at https://github.com/itmat/VODKA.

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

All data are available upon request to the corresponding author. Raw sequence data are deposited on SRA or GEO (Fig.1: PRJNA681672; Fig.2 and S1A: GSE146925; Fig 6: GSE166161).

Field-specific reporting

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

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

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

Life sciences study design

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

Sample size	Sample sizes of three cohorts were dependent on availability. No sample-size calculation were performed.
Data exclusions	For cohort 1, patients infected with RSV strain B were excluded from the study to be consistent with the other two cohorts. For cohort 3, three subjects that were negative for viral load, cbDVGs and did not show any symptoms were considered as uninfected and thus were excluded from the study.
Replication	Nasal washes were screened by PCR for the presence of cbDVGs. PCR was repeated blindly by two independent researchers for up to a total of 4 times for each clinical sample. Samples that never had detectable levels of cbDVGs were categorized as cbDVG-. Samples that had detectable levels of cbDVG between 1/4 to 4/4 repeats were categorized as cbDVG +. All PCR reactions were performed well.
Randomization	N/A: All patient samples were analyzed in the order they were received.
Blinding	DVG screening was repeated blindly by two independent researchers for up to a total of 4 times for each clinical sample. Once DVG groups were established researchers were given access to the clinical data related to these patients to compare groups.

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	Anti-human IgA-peroxidase antibody produced in goat (Sigma A0295-1ML) was used to detect IgA in nasal lavage at a 1:1000 dilution
Validation	According to the manufacturer's datasheet, antiserum is produced in goat using the Fc fragment of human IgG as the immunogen. The antibody is isolated from anti-human IgG antiserum by immunospecific purification to remove essentially all goat serum proteins, including immunoglobulins, which do not specifically bind to the Fc fragment of human IgG. Anti-Human IgG is conjugated to peroxidase by means of a two-step glutaraldehyde method. The product is purified to remove unconjugated material. Specificity of the Anti-Human IgG (Fc specific)-peroxidase is determined by ELISA. The conjugate is specific for human IgG (Fc fragment) when tested against human IgA, IgG (Fab and Fc fragments), IgM, Bence Jones kappa, and lambda myeloma proteins. Cross-reactivity of the antibody-conjugate is determined by ELISA. The conjugate shows no reactivity with mouse or rat IgG. Identity and purity of the antibody is established by immunoelectrophoresis (IEP), prior to conjugation. Electrophoresis of the antibody preparation followed by diffusion versus anti-goat IgG and anti-goat whole serum results in single arcs of precipitation.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Cohort 1: all patients were confirmed with RSV infection and were under 2 years age (range 30-714 days) at the time of sample collection and required hospitalization. 64 males and 58 females were included in this cohort. Median for the length of stay in hospital was 3 days (IQR 2-5 days) and median for the length staying in ICU was 0 (IQR 0-2 days).

	Cohort 2: all patients were age from 1.57 weeks to 33 weeks, including 73 non-hospitalized patients and 27 hospitalized patients. Among non-hospitalized patients, there are 42 males and 31 females. All patients were confirmed with RSV infection at the time of sample collection. Median for the symptom scores was 2 (IQR 1-4) in non-hospitalized patients. Cohort 3: all healthy adults with age range from 18 years to 50 years were infected with RSV A Memphis. There are 35 males and 21 females. Median for their clinical scores is 20 (IQR 7.3-36.8) and median for viral load is 643.2 (IQR 0-34,707).
Recruitment	Cohort 1: A total 122 banked nasopharyngeal swabs were obtained. All samples were from hospitalized pediatric patients less than 2 years old. Premature infants, patients with congenital heart disease and patients co-infected with other respiratory viruses were excluded from the study. If the same patient had more than two RSV positive nasopharyngeal swabs, only the first one was included in the current study. No death cases were included in this study. Cohort 2: Healthy, term, normal birth weight infants were enrolled in INSPIRE shortly following birth and were followed during infancy with biweekly surveillance during RSV season of each infant's first year of life. For this study, we evaluated a random sample of 100 nasal washes. Cohort 3: 61 healthy nonsmoking adults aged 18-50 were recruited and experimentally inoculated with 10^4 plaque-forming units of RSV A Memphis 37 (M37) by intranasal drops. Of the 61 individuals enrolled in this study we had access to nasal samples from 59 individuals and among them we eliminated from the analysis the 3 that did not show detectable virus nor symptoms at any time post infection. There is no self-selection bias in three cohorts' recruitments.
Ethics oversight	Cohort 1: approved by Children's Hospital of Philadelphia, Institutional Review Board. No consent was required as these were banked samples from a retrospective study. Cohort 2: approved by Vanderbilt University School of Medicine, Institutional Review Board. One parent of each participant provided written informed consent for participation in this study. Cohort 3: approved by UK National Ethics Service London-Fulham. Written informed consent was obtained from all volunteers prior to inclusion in the study.

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