nature portfolio

Corresponding author(s): Chao Jiang

Last updated by author(s): Sep 6, 2024

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	no custom software was used
Data analysis	The raw sequencing reads were processed with Fastp to remove adaptor sequences and trim low-quality bases. Human reads were removed using a high-performance two-stage bioinformatics approach, in which reads were first aligned to human genome GRCh38 using Bowtie2 and underwent a secondary alignment using HISAT2. Taxonomic profiling and relative abundance of species were determined based on the exact alignment of unique clade-specific markers with MetaPhlAn3 v.3.0.13 on human-removed sequences. Functional profiling was performed using HUMAnN3 v.3.7, with the abundance of functional pathways classified against UniRef90 and ChocoPhlAn databases.

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

All human-removed metagenomic sequencing data and cultured isolate sequencing data generated in this study have been deposited in the European Nucleotide

Archive (https://www.ebi.ac.uk/ena/browser/home) under the project accession number PRJEB64676(https://www.ebi.ac.uk/ena/browser/view/PRJEB64676). The processed data and analytical code are available at the Github repository: https://github.com/processbys/LRT_microbiomes_critical_illness [https://zenodo.org/doi/10.5281/zenodo.13627161]93.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Although information on sex was collected for the participants of the cohort, it was not used in this study. The aim of our study was to explore the lower respiratory tract microbiome in critically ill patients.
Reporting on race, ethnicity, or other socially relevant groupings	Not relevant in this study.
Population characteristics	All host and demographic data, including age, gender, smoking status, pneumonia diagnosis, illness severity(APACHE II score, SOFA score, and oxygenation index), antibiotic usage, and the length of ICU stay were recorded and retrieved for our participants. These characteristics are summarized in Supplementary table S1.
Recruitment	Patients admitted to the ICU were screened based on the following criteria: (1) adult (>18 years of age) with an index admission to ICU; (2) requirement for mechanical ventilation, with an anticipated need for continuous mechanical ventilation (> 72 hours) as judged by the treating ICU specialist; and (3) absence of any pre-existing and transmissible diseases.
Ethics oversight	The study was approved by the Clinical Research Ethics Committee of the First Affiliated Hospital (reference: 1T20221227A), Peking University Third Hospital Medical Science Research Ethics Committee (reference: IRB00006761-M2022419), and the Medical Ethics Committee of the Second Xiangya Hospital (reference: LYG2020093). The written informed consent was obtained from all the participants for the collection of respiratory samples and the use of their relevant clinical data.

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

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	From the 157 enrolled patients, 453 endotracheal aspirates (ETA) were initially collected for downstream analysis, of which 11 samples were excluded due to insufficient DNA yielding, making the final sample size 442.
Data exclusions	No data were excluded from the analyses.
Replication	The analysis can be reproduced using the data and software described in the Methods sections.
Randomization	Samples were processed in random order.
Blinding	Participants were blinded to group allocation and all demographic information during sample processing and sequencing.

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
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

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

Plants

Seed stocks

Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.