nature research

Corresponding author(s): Niranjan Nagarajan

Last updated by author(s): Apr 10, 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	Nanopore signal data was basecalled using Albacore (v2.3.1 to v2.3.3) or Guppy (v0.5.1 to v3.0.6) and fastq files were generated using Albacore and Porechop (v0.2.3).
Data analysis	Sequencing data was assembled using MEGAHIT (v 1.0.4-beta, Illumina) or Canu (v1.3/v1.7, Nanopore) or OPERA-MS (v0.8.3, Hybrid). The assemblies were further processed and analyzed using the following software: Pilon (v1.22), CheckM (v1.0.7), Mash (v1.1.1), sourmash (v3.0.1), blastn (v2.2.28), DIAMOND (v0.9.24.125) and MUMmer (v3.23), VirSorter (v1.0.5). Illumina shotgun metagenomic reads were analyzed with skewer (v0.2.2), bwa mem (v0.7.10-r789), MetaPhlAn2 (v2.6.0) and srst2 (v0.1.4). Nanopore metagenomes were profiled with kraken (v0.10.5-beta). Scripts for generation of figures are available on GitHub (https://github.com/csb5/hospital_microbiome).

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

All sequencing reads are available from the European Nucleotide Archive (ENA) under project PRJEB31632 (https://www.ebi.ac.uk/ena/data/view/PRJEB31632).

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	No statistical analysis was used to predetermine sample size. Sample size was determined based on available resources. Analysis of generated data via rarefaction showed that the sample size was largely adequate.
Data exclusions	Samples that did not generate enough DNA for sequencing had to be excluded as library preparation for them would inevitably fail.
Replication	No technical replication was attempted for environmental sampling and culture enrichment.
Randomization	No randomization was required as the study design was observational.
Blinding	No blinding was planned as the study design was observational.

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