

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 (<i>n</i>) 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 <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☐	☒ 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. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
☐	☒ 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 <i>d</i> , Pearson's <i>r</i>), 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	Data (taxonomic profiles) were provided by a commercial company (identities to be kept blinded) or generated at NIST. Data generated at NIST was done using R 4.1.0 and DADA2 (1.20.0) for 16S and Kraken2 with bracken using default parameters and the prebuilt GTDB-tk database (r89)
Data analysis	Provider reports varied in content and format necessitating manual curation of the taxonomic data from each report to produce a consistent format. In addition to creating a consistent file and format, some datasets provided taxonomic resolution down to the species or even strain level whereas other taxa were identified at a higher order, for example only down to the family level. A genus level taxonomic classification was chosen to facilitate comparisons between datasets; therefore, the relative abundance of all species and strain level designations were summed to the genus level. The relative abundances for unclassified taxa and taxa where a specific genus designation was not given were summed and given an unspecified designation. Taxonomic tables generated from the NIST 16S and WMS analysis were put in the same format. All resulting tables were used as the input files for future analysis conducted in R Studio along with the complete code for analysis and figure generation can be found at https://doi.org/10.18434/mds2-3348 . We are unable to share the actual reports in a blinded fashion so this version of the data will not be made publicly available.

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 Data (including raw sequencing data generated at NIST, taxonomic profiles, and code used for analysis) are publicly available through the National Institute of Standards and Technology MIDAS platform <https://doi.org/10.18434/mds2-3348>. Sequencing data generated at NIST has also been uploaded to the SRA (BioProject ID PRJNA1364055). The reports and sequencing data from the companies will not be made available as this can not be done while maintaining blinding.

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	Information on material source was collected but is not reported in this study since the focus is on a control material and not biological attributes.
Reporting on race, ethnicity, or other socially relevant groupings	NA
Population characteristics	NA
Recruitment	Material was collected from generous donors. We do not believe the donor source impact the results of this study since the focus was measurement reproducibility not biological conclusions from the reports.
Ethics oversight	All whole stool samples were collected after informed consent under approved IRB protocols at The BioCollective. subsequent work conducted at NIST was reviewed and approved by the U. S. National Institute of Standards and Technology Research Protections Office

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	No sample size calculation was done, statistical analysis was not conducted as part of this comparison - replication was done to compare between lab variation to within lab variation on an n of 3.
Data exclusions	1 data set was excluded from some analyses due to aberrant taxonomic assignment and low (less than half) the number of reads compared to the other 2 replicates.
Replication	This was an output of the study - an assessment of variability within a company compared to between companies on a common sample.
Randomization	This was not relevant for this study - a single common sample was needed for this study
Blinding	commercial companies were blinded to the sample site and replication. a mock account was created by the researchers.

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

Methods

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

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, mosaicism, off-target gene editing) were examined.