

## 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](#).

### Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

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

Our web collection on [statistics for biologists](#) may be useful.

### Software and code

Policy information about [availability of computer code](#)

Data collection	No software was used for data collection
Data analysis	Samples from already published studies were processed using a close reference approach and rarefied at 1,000 sequences per sample, then we calculated PCoAs based on the unweighted UniFrac distances.

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 upon request. 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

Studies in this analysis:

<https://qiita.ucsd.edu/study/description/1928>

<https://qiita.ucsd.edu/study/description/1927>  
<https://qiita.ucsd.edu/study/description/1924>  
<https://qiita.ucsd.edu/study/description/1939>  
<https://qiita.ucsd.edu/study/description/1998>  
<https://qiita.ucsd.edu/study/description/11484>  
<https://qiita.ucsd.edu/study/description/1629>  
 Figure generation:  
<https://github.com/knightlab-analyses/qiita-paper>

## Field-specific reporting

Please select 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/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

## Life sciences study design

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

Sample size   
 Data exclusions   
 Replication   
 Randomization   
 Blinding

## Reporting for specific materials, systems and methods

### Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Unique biological materials
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants

### Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging