nature research

Corresponding author(s): Simona Radutoiu
Paul Schulze-Lefert
Ruben Garrido-Oter

Last updated by author(s): May 14, 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	16S amplicon data was collected using an Illumina Miseq sequencer. RNA-Seq data was generated by sequencing on the Illumina HiSeq3000 platform.
Data analysis	DADA2 v1.16 QIIME v1.9.1 USEARCH v8.0 UPARSE v8.0 R v4.0.0 A5 v20160825 METABAT v2.15 CheckM v.1.1.2

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

Raw 16S rRNA amplicon reads are deposited in the European Nucleotide Archive (ENA) under the accession number PRJEB37695. Similarly, genome and transcriptome sequencing reads are uploaded to the same database with the accession number PRJEB37696.

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 size was based on experience from previous experiments using the same gnotobiotic system to allow for confident statistical analyses. Sample size was indicated as "n" in the corresponding figures. No statistical methods were used to determine sample size.
Data exclusions	No data was excluded.
Replication	Experiments B and C included 3 independent preparations of the corresponding SynComs with at least 18 biological replicates of planted pots each. All other experiments were performed with one SynCom preparation and at least 18 biological replicates of planted pots. Experiment M was performed in 3 biological and 6 technical replicates.
Randomization	Position of boxes with planted pots was randomized in the growth chamber during the course of the experiments. Position of pots with different plant genotypes within a box was randomized.
Blinding	Samples in each experiment were harvested as well as processed by multiple researchers, and every researcher processed samples of all genotypes and compartment types. Experiments were independently validated or performed by different researchers to ensure reproducibility of the data.

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