nature research

Corresponding author(s): Sheng Yang He

Last updated by author(s): Jan 8, 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	n/a
Data analysis	Description of published software used for data analysis (Fig. 6) is provided and cited in the Supplementary Methods section and figure legend. These include: FastQC (Version 0.11.5) https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ Ribosomal Database Project (RDP) preprocessing tools (Release 11) http://rdp.cme.msu.edu/ PANDAseq. https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-13-31 BBDuk (Version 36.77) http://jgi.doe.gov/data-and-tools/bb-tools/ USEARCH (Version 10) https://www.drive5.com/usearch/ UPARSE (Version 10) https://www.drive5.com/uparse/ UCHIME (Version 10) https://www.drive5.com/uchime/ References for each of these tools is included in the Supplementary Methods section.

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 source 16S rRNA gene sequences from this project are available in the Sequence Read Archive database under BioProject PRJNA689857, accession numbers SAMN17220890 to SAMN17220933.

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	This protocol does not involve statistical analysis. The sample size for Fig. 6 is described in the figure legend and was determined with consideration of previous publications on plant microbiota colonization experiments. No statistical methods were used to predetermine sample sizes.
Data exclusions	No data that pass quality control were excluded from analysis.
Replication	The GnotoPot and FlowPot experiments were repeated 3 times with similar results. Results were reproducible in all repeats with the same trend.
Randomization	The Microboxes placed in growth chambers randomly and rotated randomly to achieving consistent plant growth.
Blinding	Researchers were not blinded to allocation during experiments and outcome assessment. This is in part because different plant genotypes under study (Col-0 and the mbbc mutant) exhibit very distinct phenotypes visually; blinding was not possible. Routine practices included more than one author observing/assessing phenotypes, whenever possible.

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