nature portfolio

Corresponding author(s): Hirokazu Toju
Mikihito Noguchi

Last updated by author(s): Nov 7, 2025

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 (*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	Sequencing data were obtained with the Illumina MiSeq platform. Reads were processed using bcl2fastq (version 1.8.4), Claident (version 0.9.2022.01.26), Cutadapt (version. 3.7), R (version 4.2.2), DADA2 (version 1.18.0), decontam (version 1.1.1) and VSEARCH (version 2.21.1). Taxonomy assignment was performed using the UNITE general FASTA release for eukaryotes 2. Version 18.07.2023 and the SILVA version 138 databases.
Data analysis	Analyses were conducted in R (version 4.4.1) using vegan (version 2.6-8), rELA (version 0.60), compositions (version 2.0.8), purr (version 1.0.4), stringdist (version 0.9.15), parallel (version 4.4.1), foreach (version 1.5.2), doParallell (version 1.0.17), dplyr (version 1.1.4), ComplexHeatmap (version 2.22.0), circlize (version 0.4.16), cluster (version 2.1.8), tidyverse (version 2.0.0), renv (version 1.1.4), patchwork (version 1.3.0), ggrepel (version 0.9.6), ggalluvial (version 0.12.5), Rcpp (version 1.0.14), RcppArmadillo (version 14.4.3.1), gtools (version 3.9.5), ggplot2 (version 3.5.2), ggtext (version 0.1.2) , ggstar (version 1.0.4), ggnetwork (version 0.5.13), igraph (version 2.1.4), bipartite (version 2.20) and RColorBrewer (version 1.1.3) packages. All scripts and processed data are available at https://github.com/mikihito-noguchi/Assembly_landscape_of_the_forest_root_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 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

The DNA sequence data have been deposited to the DNA Data Bank of Japan (DDBJ accession: PRJDB17520 and PRJDB37499) [to be made public after the acceptance of the paper]. The computer codes are available from the GitHub repository (https://github.com/ngchngch/Assembly_landscape_of_the_forest_root_microbiome).

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	This study did not involve human participants, their data, or biological materials.
Reporting on race, ethnicity, or other socially relevant groupings	This study did not involve human participants, their data, or biological materials.
Population characteristics	This study did not involve human participants, their data, or biological materials.
Recruitment	This study did not involve human participants, their data, or biological materials.
Ethics oversight	This study did not involve human participants, their data, or biological materials.

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

Ecological, evolutionary & environmental sciences study design

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

Study description	This study aimed to characterize the assembly landscapes of root-associated microbiomes in a cool-temperate forest in central Japan. By integrating fungal ITS and prokaryotic 16S rRNA community data from 1,270 root-tip samples collected from six tree species across 125 sampling points, we applied energy landscape analysis to infer the assembly landscapes of fungal and prokaryotic communities and to identify potential keystone taxa influencing their community assembly.
Research sample	Root-tip samples (1-cm terminal roots) were collected for six dominant tree genera (Acer, Betula, Juglans, Larix, Pinus, and Populus) in a natural cool-temperate forest at the Sugadaira Research Station, University of Tsukuba, Nagano, Japan (36.524°N, 138.349°E, 1,340 m a.s.l.). Both fungal and prokaryotic communities were analyzed using amplicon sequencing of the ITS and 16S rRNA regions. The fungal dataset was previously described in our earlier publication (https://doi.org/10.1186/s40793-024-00628-8).
Sampling strategy	Sampling points (n = 125) were established across the forest to capture variation in root-associated microbial communities among coexisting tree species. A total of 1,270 root-tip samples were collected to represent variation within each sampling point and host species (more than 50 samples per host). Root samples yielding fewer than 1,000 reads were excluded. Data coverage was standardized using a coverage-based rarefaction method implemented in the R package vegan.
Data collection	MN and HT collected root samples at 125 points in a natural cool-temperate forest at the Sugadaira Research Station, University of Tsukuba, Nagano, Japan (36.524°N, 138.349°E, 1,340 m a.s.l.). MN conducted DNA metabarcoding for the fungal ITS and prokaryotic 16S rRNA regions. One-centimeter root tips were collected, surface sterilized, and subjected to DNA extraction and purification using the CTAB method, followed by PCR amplification of target regions. MiSeq sequencing data were quality-filtered, and operational taxonomic units (OTUs; 97% similarity threshold) were assigned to taxonomic groups.
Timing and spatial scale	The study covered an approximately 8.5-hectare forest (125 sampling points). Sampling was conducted from July 27 to August 5, 2022.

Data exclusions	Root samples with fewer than 1,000 sequencing reads were excluded from both fungal and prokaryotic datasets. Plant genera with fewer than 50 qualified root samples were not included. In the energy landscape analysis, samples in which more than 50% of reads could not be annotated at the family level were excluded.
Reproducibility	The outputs of the energy landscape analysis and the sensitivity analyses for identifying keystone taxa were highly reproducible, showing minimal variation when simulations with randomly generated initial community states were repeated 30 times.
Randomization	In the energy landscape analysis, host plant identity was included as a covariate. For the z-standardization of the Δtopography and Δevenness indices, relative abundance values were randomly shuffled within host-plant categories (10,000 permutations) to generate null distributions.
Blinding	Blinding was not applicable, as the study involved no experimental manipulation or treatment groups. All analyses were based on computational modeling of observational microbial community data.

Did the study involve field work? ☒ Yes ☐ No

Field work, collection and transport

Field conditions	Fieldwork was conducted under natural environmental conditions without any experimental manipulation. For example, a previous study reported the soil nitrogen concentrations (Urakawa et al., 2015; measurements at 0–10 cm soil depth for five sampling points: NH_4^+, = 2.426 mgN per kg soil; NO_3^- = 0.439 mgN per kg soil). Root samples were collected from July 27 to August 5, 2022. Sampling was performed under non-rainy conditions to minimize variation in soil moisture and contamination risk.
Location	Field sampling was carried out in the research forest of Sugadaira Research Station, Mountain Science Center, University of Tsukuba, Sugadaira, Ueda, Nagano Prefecture, Japan (36.524 ºN; 138.349 ºE; 1340 m a.s.l.).
Access & import/export	All field collections were conducted with permission from Sugadaira Research Station, Mountain Science Center, University of Tsukuba. No international transport of biological materials occurred; all samples were collected and processed domestically in Japan.
Disturbance	Sampling was designed to minimize ecological disturbance. At each sampling point, ca. 30 cm2 soil was excavated from a depth of 0 to 20 cm in order to collect woody plant roots. All sampling sites were restored to their original state immediately after collection.

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
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	We did not use any seed stocks.
Novel plant genotypes	We did not produce any new plant genotypes.
Authentication	We did not use any seed stocks. We did not produce any new plant genotypes.