https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

Open Access Earth System Science Data Discussions

A global database of soil microbial communities and associated climate, soil and vegetation factors

Authors: Shutao Chen*, Qi Li, Zhenghua Hu

School of Ecology and Applied Meteorology, Nanjing University of Information Science and Technology, Nanjing 210044, China

* Corresponding author: Shutao Chen

E-mail: chenstyf@aliyun.com

Telephone: +86-25-58731539

Fax: +86-25-58731539

ORCID: 0000-0001-8513-1787

Submitted to Earth System Science Data

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

Open Access Earth System Science Data Discussions

Abstract. Few scholars have compiled databases of soil microbial communities and associated climate, soil and vegetation factors at the global scale. However, many studies involving high-throughput sequencing of soil bacteria and fungi have been published in the past decade. In this study, we constructed a global database of the soil microbial communities and the associated climate, soil and vegetation factors, with sites on each of the seven continents and eleven ecosystem types. There were 8490 sets of soil bacterial and fungal community data for the different treatments and study sites in the database. Soil bacterial and fungal diversities were highly variable across various ecosystems. There was a highly significant (R^2 = 0.4037, P < 0.001) linear regression relationship between the fungal and bacterial Shannon indices. Proteobacteria and Ascomycota were the most species-rich bacterial and fungal phyla, respectively, in most ecosystems. The median relative abundances of Proteobacteria and Ascomycota were 29.30 % and 57.49 %, respectively. The information (e.g., site names and ecosystem types) in the database enabled researchers to investigate where the most abundant bacterial or fungal phylum was located and whether the ecosystem type affected bacterial and fungal diversities and compositions at the global scale. We anticipated that this database could be further improved by adding more detailed information, such as bacterial and fungal compositions at the class, order, family, and genus levels. The database is available via Zenodo at https://doi.org/10.5281/zenodo.16195889 (Chen et al., 2025).

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

Open Access Earth System Science Data Discussions

Short summary

Soil microorganisms play important roles in carbon stock, nutrient cycling and vegetation production. We compiled a global database of soil microbial communities and potential driving factors, such as climate, soil and vegetation. Our database included soil bacterial and fungal diversity indices and compositions. Researchers can examine which ecosystems and locations have high bacterial and fungal diversities and where the dominant bacterial and fungal compositions exist.

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

1 Introduction

Soil microorganisms play key roles in soil functions such as carbon stock (Hanson et al. 2008; Jones and Lennon, 2010), nutrient cycling (Classen et al., 2015; Delgado-Baquerizo et al., 2020) and plant production (Bahram et al., 2018; Labouyrie et al., 2023). Ecosystem functions may be altered by altering the soil microbial community. It is necessary to explore the global distribution patterns and community compositions of soil microorganisms and associated driving factors to determine the function of soil. Bacteria and fungi are the two main microbial taxa that dominate soil habitats. Numerous diversity indices and compositions at various levels (e.g., phylum, class and order) are used to characterize bacterial and fungal communities.

Soil bacterial and fungal diversities and compositions are affected by a series of environmental driving factors, including climate (Bahram et al., 2018), soil (Lauber et al., 2008; Lu et al., 2023) and vegetation factors (van der Heijden et al., 2008; Delgado-Baquerizo et al., 2016). Climate factors, soil properties and vegetation characteristics influence bacterial and fungal diversities and compositions through the regulation of soil microorganism assemblages. These factors are tightly linked to microbial metabolism and thus affect soil bacterial and fungal diversity and composition (Fierer and Jackson, 2006; Green et al., 2008; Banerjee and Heijden, 2022). However, the key environmental variables that control spatial and temporal variations in the global soil microbial community are still unresolved because of the lack of a database and the high variability of microorganisms.

Moreover, scholars have rarely explained the relationship between fungal and

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

bacterial diversity at the global scale (Jiao et al., 2021; Zubek et al., 2024; Zhang et al., 2025). We do not know what kind of relationship (e.g., linear, nonlinear, or no relationship) exists between fungal and bacterial diversity among different ecosystem types. In addition, we do not know whether such relationships exist among different diversity indices that shape the complexity of soil microorganisms. Addressing these themes may aid in understanding the microbial coexistence relationship between relatively inferior prokaryotes of bacteria and more advanced eukaryotes of fungi.

In the past decade, high-throughput sequencing has been widely used to determine soil bacterial and fungal diversity and composition (Lozupone et al., 2011; Schmieder and Edwards, 2011). Second-generation sequencing has revolutionized genomics by overcoming the fundamental limitations of throughput, cost, and scalability inherent in first-generation sequencing, making large-scale genomic studies feasible and routine (Edgar, 2013; Rinke et al. 2013; Callahan et al., 2016). Although many studies have been published in the past decade on the high-throughput sequencing of soil bacteria and fungi, few researchers have compiled databases of soil microbial communities and associated climate, soil and vegetation factors at the global scale (Delgado-Baquerizo et al., 2018; Averill et al., 2019; Peng et al., 2024).

In this study, we constructed a global database of soil microbial communities using second-generation high-throughput sequencing publications and associated climate, soil and vegetation factors. The database included soil bacterial and fungal diversity indices and compositions at the phylum, class and order levels. Researchers can examine which ecosystem and location have high bacterial and fungal diversities and

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

where the dominant bacterial and fungal compositions exist. The relationships between fungal and bacterial diversities can be explored using the available data. Researchers can use databases to investigate the effects of experimental treatments on soil bacterial and fungal diversity and composition.

2 Materials and methods

2.1 Data sources

We searched the keywords “soil microbial community”, “soil bacterial community” and “soil fungal community” in the Web of Science from 2013 to 2024. The searched articles were subsequently found in the relevant literature databases, such as Elsevier ScienceDirect, Springer Link, Wiley Online Library, and Taylor and Francis. The literature featuring the second-generation high-throughput sequencing method (i.e., Illumina MiSeq, Illumina HiSeq and Illumina NovaSeq) was compiled, and manuscripts that used the first-generation sequencing method (e.g., Roche 454) were not included in the database. We found that no studies included a second-generation high-throughput sequencing method before 2016. Therefore, our earliest collection of target studies began in 2016. The spliced fasta sequences were clustered into operational taxonomic units (OTUs) with 97 % similarity (Edgar et al., 2013). The graphic digitizing software GetData Graph Digitizer 2.26 (GetData Inc., S. Federow, Moscow, Russia) was used to obtain data from figures if the data were unavailable in tables and text.

2.2 Structure of the database

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

In the database, 8490 sets of soil bacterial and fungal community data were obtained for the different treatments and study sites. The database included geographic information, sampling date, bacterial and fungal diversity indices, experimental treatments, climate factors, soil properties, vegetation characteristics, bacterial and fungal compositions, and site information (Table S1). The geographic information included latitude, longitude, altitude and site name. The sampling date was the 15th of the sampling month for convenience if only the year and month were reported in the text. The bacterial and fungal diversity indices included α and β diversities, although β diversity was available in only a few studies. The bacterial and fungal α diversity indices consisted of bacterial sequences, good coverage, OTU richness, the Chao1 index, the Shannon index, the Simpson index, the abundance-based coverage estimator (ACE) index, the PD whole-tree index, Pielou's evenness index, and 16S rRNA abundances. If one variable, such as the OTU richness, Chao1 index, Shannon index, Simpson index, ACE index, PD whole-tree index, Pielou's evenness index, or 16S rRNA abundance, appeared in an article, it was compiled in the database. The bacterial and fungal β diversity indices consisted of the Bray–Curtis distance, weighted UniFrac distance and unweighted UniFrac distance. The experimental treatments in the database included fertilization (e.g., nitrogen, phosphorus, and potassium application), warming, precipitation manipulation, CO_2 enrichment, manure, biochar, straw, and compost application, and drainage. The climate factors included precipitation, temperature, climate type, annual evaporation, and potential evapotranspiration. Many soil properties were collected in the database, as these properties concerning the soil

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

154 microbial community and potential soil drivers have been widely reported in the
155 literature. The vegetation characteristics included vegetation biomass, tree
156 physiological indicators, vegetation cover, plant diversity indices, and dominant
157 vegetation type. The bacterial and fungal compositions were evaluated considering their
158 relative abundances at the phylum, class and order levels. Eleven ecosystems were
159 included in our collection. The abbreviations for the bacterial and fungal diversity
160 indices, soil properties, vegetation characteristics and ecosystem types are shown in
161 Table 1.
162
163
164
165
166
167
168
169
170
171
172
173
174
175

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

Open Access Earth System Science Data Discussions

176 **Table 1.** Variable categories included in the dataset.

Abbreviations	Variables	Abbreviations	Variables
MAP	Mean annual precipitation	AMN	Soil ammonium nitrogen
MAT	Mean annual temperature	NIN	Soil nitrate nitrogen
BAS	Bacterial sequences	STP	Soil total phosphorus
BGOC	Bacterial good coverage	SAP	Soil available phosphorus
BOTUr	Bacterial OTU richness	STK	Soil total potassium
BCHA	Bacterial Chao1 index	SAK	Soil available potassium
BSHA	Bacterial Shannon index	MBC	Microbial biomass carbon
BSIM	Bacterial Simpson index	MBN	Microbial biomass nitrogen
BACE	Bacterial abundance-based coverage estimator index	MBP	Microbial biomass phosphorus
BPDW	Bacterial PD whole-tree index	DOC	Dissolved organic carbon
BPIE	Bacterial Pielou's evenness index	DON	Dissolved organic nitrogen
BrRNA	16S rRNA abundance	CEC	Cation exchange capacity
BBCD	Bacterial Bray–Curtis distance	AGB	Above ground biomass
BWUD	Bacterial weighted UniFrac distance	BGB	Below ground biomass
BUUD	Bacterial unweighted UniFrac distance	TOB	Total biomass
FUS	Fungal sequences	TRA	Tree age
FGOC	Fungal good coverage	PLD	Plant (tree) density
FOTUr	Fungal OTU richness	DBH	Tree diameter at breast height
FCHA	Fungal Chao1 index	TRH	Tree height
FSHA	Fungal Shannon index	VEC	Vegetation cover
FSIM	Fungal Simpson index	PSR	Plant species richness
FACE	Fungal abundance-based coverage estimator index	PSH	Plant Shannon index
FPDW	Fungal PD whole-tree index	PMA	Plant Margalef index
FPIE	Fungal Pielou's evenness index	PPIE	Plant Pielou's evenness index
FrRNA	ITS rRNA abundance	PSIM	Plant Simpson index
FBCD	Fungal Bray–Curtis distance	CRY	Crop yield
FWUD	Fungal weighted UniFrac distance	EVA	Annual evaporation
FUUD	Fungal unweighted UniFrac distance	PET	Annual potential evapotranspiration
STE	Soil temperature	BNMF	Broad and needle-leaved mixed forest
SMO	Soil moisture	CL	Cropland
pH	Soil pH	DBF	Deciduous broad-leaved forest
SEC	Soil electric conductivity	DNF	Deciduous needle-leaved forest
CLA	Soil clay content	DS	Desert
SAN	Soil sand content	EBF	Evergreen broad-leaved forest
SIL	Soil silt content	ENF	Evergreen needle-leaved forest
SBD	Soil bulk density	GL	Grassland
SOC	Soil organic carbon	SL	Shrubland
STN	Soil total nitrogen	TD	Tundra
AVN	Soil available nitrogen	WL	Wetland

177

178

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

2.3 Data analysis

The spatial distributions of the soil bacterial and fungal measurement sites were plotted with ArcGis 10.2 software (Environmental Systems Research Institute Inc., Redlands, California, USA). Violin plots made by GraphPad Prism 9 (GraphPad Inc., San Diego, California, USA) were used to examine the distributions of the bacterial diversity indices of BOTUr, BCHA, BSHA, BSIM, BACE, and BGOC and the fungal diversity indices of FOTUr, FCHA, FSHA, FSIM, FACE, and FGOC. Linear regression was used to explore the relationships between the bacterial diversity indices of BOTUr, BCHA, BSHA, BSIM, BACE, and BGOC and the fungal diversity indices of FOTUr, FCHA, FSHA, FSIM, FACE, and FGOC. Pearson's correlations performed by SPSS Statistics 26 (SPSS Inc., Chicago, Illinois, USA) were used to investigate the relationships between the bacterial and fungal diversities and climate, soil and vegetation factors. Redundancy analysis (RDA) was performed by using Canoco 5 (Biometris Inc., Wageningen, Gelderland, The Netherlands) to analyze the effects of soil properties on the relative abundances of bacterial (top ten) and fungal (top eight) compositions at the phylum level.

3 Results

3.1 Bacterial and fungal diversities

The reviewed data points for the soil microbial community were distributed across each continent (Fig. 1). The northernmost and southernmost data points were located in Ny-Ålesund, Sweden and McMurdo Dry Valleys, Antarctica, respectively. Each of the

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

ecosystem types had data points for BOTUr, BCHA and BSHA (Figs. 2a, b and c). No BSIM data in the TD, no BACE data in the DNF and TD and no BGOC data in the DNF were reported (Figs. 2c, d and e). The CL had the most measurement data points for BOTUr, whereas the DS had the fewest. BOTUr varied from 176 to 19454 no. copies g^{-1} across all of the ecosystems, with the lowest and highest values appearing in the CL (Piauí, Brazil) and DBF (Cape Jervis, Australia), respectively. The median BOTUr, BCHA, BSHA, and BACE values in the different ecosystems ranged from 1348 to 3014 no. copies g^{-1}, 2061.00 to 4276.75, 5.600 to 7.708, and 2014 to 3333, respectively. The data distribution patterns of BCHA, BSHA and BACE were similar to those of BOTUr. The BSIM varied from 0.6000 to 1.0000. The lowest BSIM (0.6000) appeared in the CL, and the highest (1.0000) appeared in the CL, DBF, DS, EBF, ENF, GL, and SL. The lowest BGOC (0.8251) appeared in the EBF, and the highest (1.0000) appeared in the GL and SL.

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

Open Access Earth System Science Data Discussions

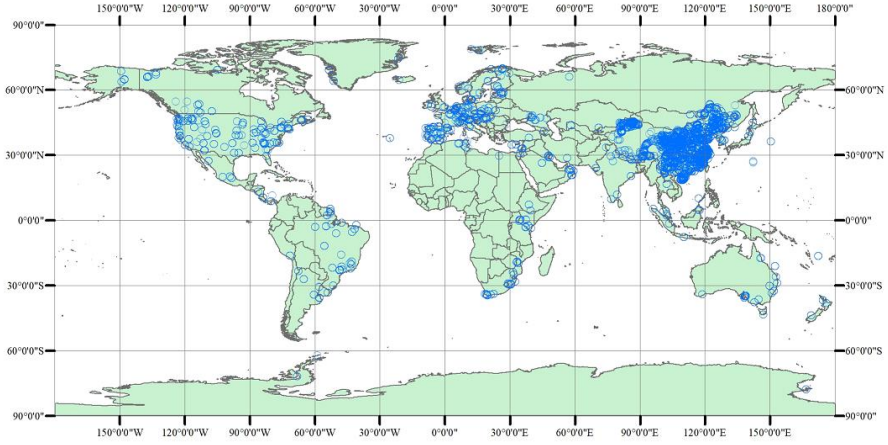
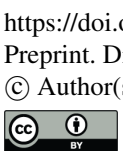


Figure 1. Spatial distributions of soil bacterial and fungal measurement sites.

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025



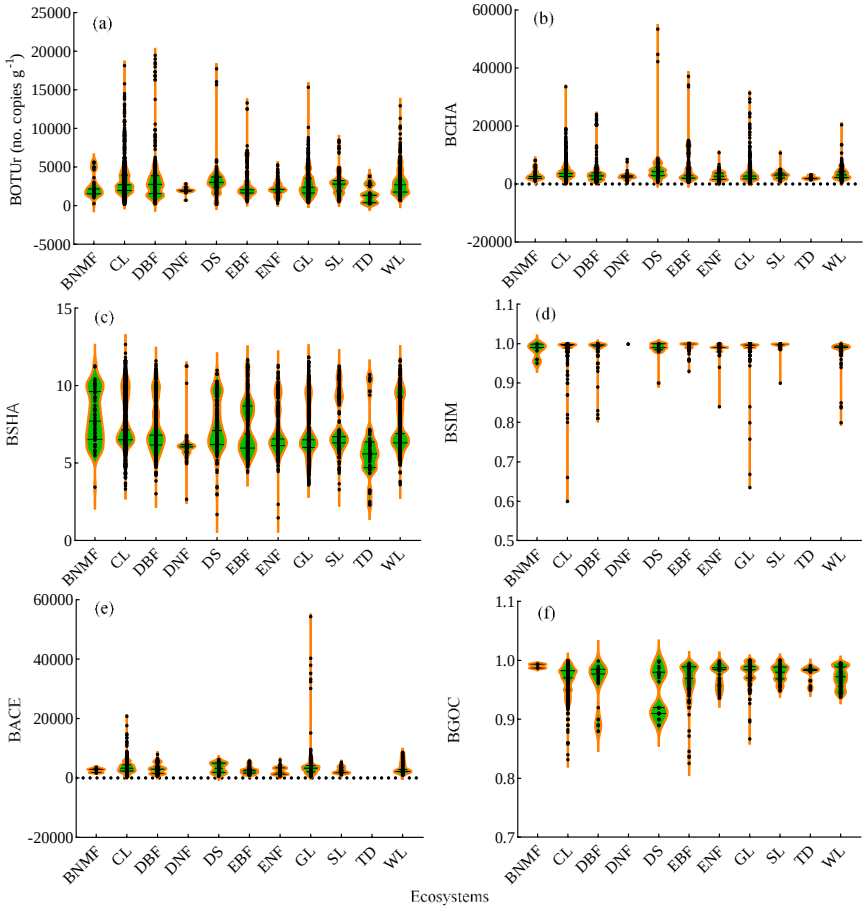


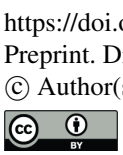
Figure 2. Violin plots of the distribution of bacterial diversity indices. Panels a, b, c, d, e, and f denote BOTUr, BCHA, BSHA, BSIM, BACE, and BGOC, respectively.

The FOTUr ranged from 24 to 3433 no. copies g^{-1}, with the lowest and highest values occurring in El Reno, USA, and Qingyuan, China, respectively. Most FOTUr data points were in the CL, with values ranging from 24 to 2719 no. copies g^{-1} (Fig. 3a). There were 788, 867, 1192, 178, 302, and 195 data points for FOTUr, FCHA, FSHA, FSIM, FACE, and FGOC, respectively, in the CL. The median FOTUr, FCHA, FSHA,

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

and FACE values across all of the ecosystems ranged from 206 to 639 no. copies g^{-1}, 215.90 to 835.50, 2.824 to 5.128 and 231.48 to 814.00, respectively (Figs. 3a, b, c, and e). Most of the FOTUr, FCHA, FSHA, and FACE data points were distributed close to the median values, whereas most of the FSIM and FGOC values were close to 1 (Figs. 3d and f). The FOTUr and FACE in the DBF exhibited the greatest variations across all of the ecosystems, with ranges from 57 to 3433 no. copies g^{-1} and 98 to 5768 for FOTUr and FACE, respectively.

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025



Open Access Earth System Science Data Discussions

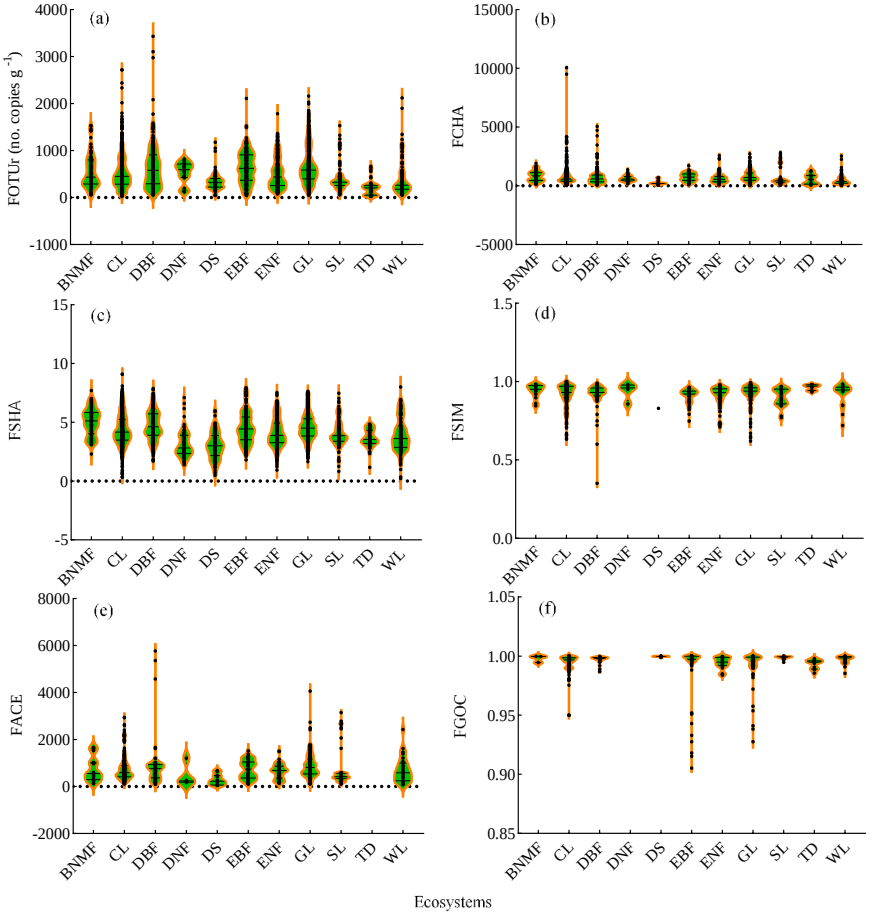


Figure 3. Violin plots of the distribution of fungal diversity indices. Panels a, b, c, d, e, and f denote FOTUr, FCHA, FSHA, FSIM, FACE, and FGOC, respectively.

There were fewer data points for the bacterial and fungal β diversities than for their α diversities (Table S1). The ranges of the BBCD and FBCD were from 0.050 to 0.600 and from 0.072 to 0.908, respectively. The lowest and highest BBCD values appeared in the EBF and GL, respectively, whereas both the lowest and highest FBCD values appeared in the ENF.

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

There was a significant linear regression relationship between FOTUr and BOTUr (Fig. 4a). The FOTUr in each of the ecosystems generally increased with increasing BOTUr. BOTUr explained 12.65 % ($R^2 = 0.1265$) of the temporal and spatial variations in FOTUr. The relationship between the FCHA and BCHA could be explained by a linear regression function, with an R^2 of 0.1805 and a P value less than 0.001 (Fig. 4b). There was a strong linear regression relationship between the FSHA and BSHA ($R^2 = 0.4037$; $P < 0.001$). BSHA explained 40.37 % of the variation in FSHA across all eleven ecosystems. In addition, the relationship between the FACE and BACE could be simulated by a linear regression model ($R^2 = 0.1294$; $P < 0.001$; Fig. 4e). No regression relationships between the FSIM and BSIM or between the FGOC and BGOC were found (Fig. 4d and f). The results of the regression analysis revealed that fungal diversity generally increased with increasing bacterial diversity, indicating a synchronous change in the two microbial communities.

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

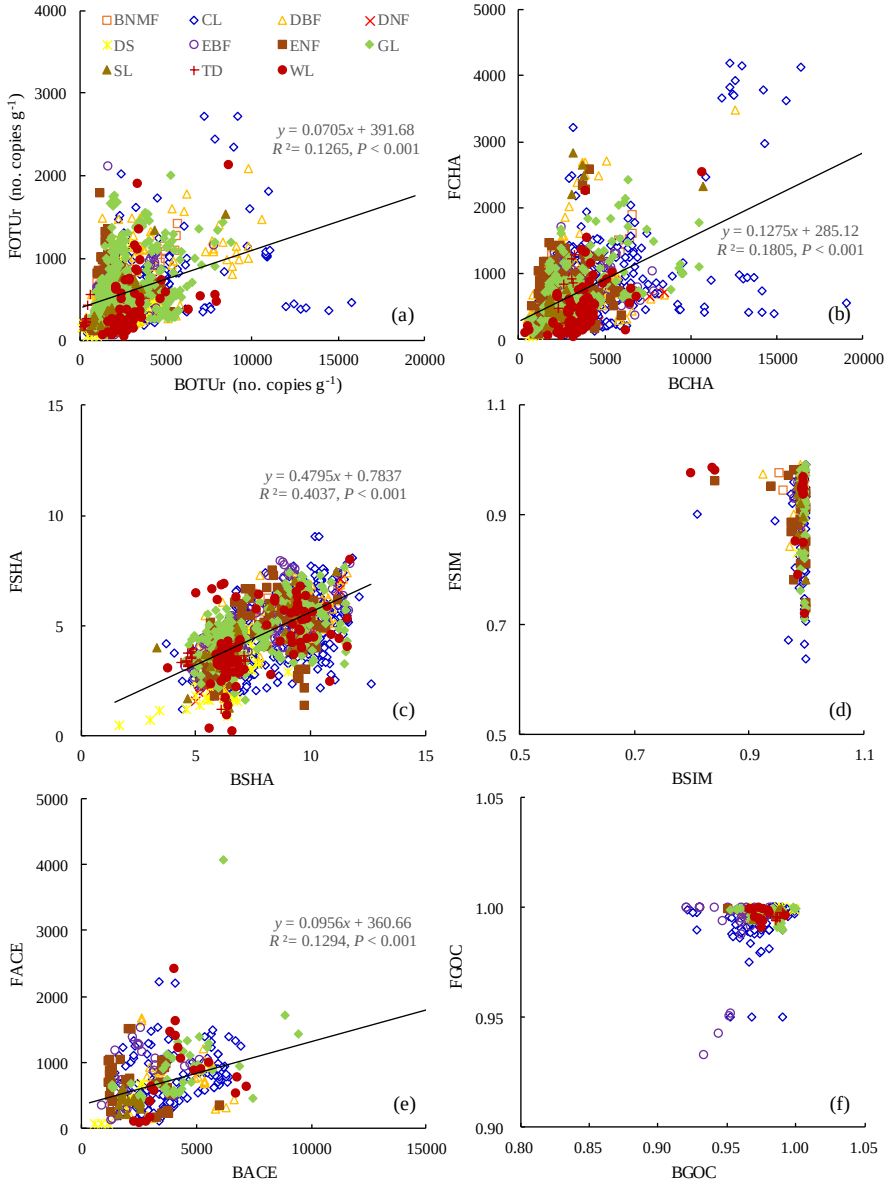


Figure 4. Regression relationships between the bacterial and fungal diversity indices. Panels a, b, c, d, e, and f denote the relationships between FOTUr and BOTUr, FCHA and BCHA, FSHA and BSHA, FSIM and BSIM, FACE and BACE, and FGOC and BGOC, respectively.

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

3.2 Bacterial and fungal compositions

The top ten dominant bacterial phyla were Proteobacteria, Acidobacteriota, Actinobacteriota, Chloroflexota, Verrucomicrobiota, Bacteroidota, Planctomycetota, Gemmatimonadota, Firmicutes, and Nitrospirota (Table S1); the median relative abundances were 29.30 %, 15.88 %, 14.40 %, 7.6 %, and 3.00 %, respectively. 4.60 %, 4.10 %, 3.70 %, 3.10 %, and 2.10 %, respectively. With respect to the simultaneously reported relative abundances of Proteobacteria and Acidobacteriota, 85.40 % of the data points presented higher relative abundances for the former. In addition to the ten dominant bacterial phyla, relatively few other phyla have been reported. The top eight dominant fungal phyla were Ascomycota, Basidiomycota, Mortierellomycota, Zygomycota, Mucoromycota, Rozellomycota, Chytridiomycota, and Glomeromycota (Table S1); the median relative abundances of the eight phyla were 57.49 %, 15.90 %, 6.94 %, 6.87 %, 3.39 %, 1.37 %, 0.95 %, and 1.00 %, respectively. The reported abundances of the other phyla were significantly less than those of the eight phyla. Ascomycota was clearly the most species-rich phylum in most ecosystems. With respect to the reported relative abundances of both Ascomycota and Basidiomycota, however, the relative abundance of the latter was higher than that of the former in 19.62 % of the ecosystems.

The classes (e.g., Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria, and Deltaproteobacteria) within the phylum Proteobacteria have been mostly reported in the literature. For instance, the reported relative abundance of Alphaproteobacteria varied from 2.13 % to 74.40 %. The classes (e.g., Acidobacteria_Gp1,

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

376 Acidobacteria_Gp2 and Acidobacteria_Gp3) within the phylum Acidobacteria have
377 also been reported in the literature. The most reported fungal class was Agaricomycetes,
378 whose relative abundance ranged from 0.30 to 95.70 %, with a median of 9.71 %. There
379 were 88 bacterial orders (e.g., Gemmatimonadales) and 52 fungal orders (e.g.,
380 Agaricales) in the database. In addition, there were fewer reported bacterial and fungal
381 orders than the number of reported phyla and classes.

382 **3.3 Relationships between the microbial community and potential driving factors**

383 Most of the bacterial and fungal diversity indices were significantly ($P < 0.05$)
384 correlated with the climate factors MAP and MAT, although the correlation coefficients
385 were relatively low (Table 2). The bacterial diversity indices of BOTUr, BCHA, BSHA,
386 and BGOC were significantly ($P < 0.001$) positively correlated with STE, whereas only
387 the fungal diversity index of FOTUr was significantly ($P = 0.027$) correlated with STE.
388 BOTUr, BCHA, BACE, and BGOC were significantly ($P < 0.001$) negatively
389 correlated with SMO. BOTUr was significantly ($P < 0.01$) correlated with numerous
390 soil and vegetation variables, such as pH, SBD, STK, MBN, BGB, VEC, and PSH. The
391 variables that were significantly correlated with BOTUr were similar to those that were
392 significantly correlated with other bacterial diversity indices. For instance, the SBD was
393 significantly correlated with BOTUr ($P < 0.001$), BCHA ($P < 0.001$), BSHA ($P < 0.001$),
394 BSIM ($P = 0.008$), BACE ($P < 0.001$), and BGOC ($P < 0.001$). SBD and PSH were the
395 key soil and vegetation variables, respectively, controlling the spatial and temporal
396 variations in bacterial diversity. FOTUr was significantly correlated with numerous soil
397 properties, such as AVN ($P = 0.001$), SAP ($P < 0.001$), SAK ($P < 0.001$), and MBN (P

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

398 = 0.003). The variables that were significantly correlated with FOTUr were similar to
399 those that were significantly correlated with the other fungal diversity indices.
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418

Table 2. Pearson's correlations between the bacterial and fungal diversities and climate, soil and vegetation factors.

	BOTUr		BCHA		BSHA		BSIM		BACE		BGOC		FOTUr		FCHA		FSHA		FSIM		FACE		FGOC	
	r	*P*	*r*	*P*	*r*	*P*	*r*	*P*	*r*	*P*	*r*	*P*	*r*	*P*	*r*	*P*	*r*	*P*	*r*	*P*	*r*	*P*	*r*	*P*
MAP	-0.046	0.003	-0.072	0.000	-0.053	0.000	-0.033	0.324	-0.109	0.000	0.148	0.000	0.112	0.000	0.087	0.000	0.015	0.398	-0.010	0.827	0.083	0.025	-0.209	0.000
MAT	0.081	0.000	0.132	0.000	0.003	0.834	0.072	0.033	-0.064	0.018	-0.136	0.000	0.052	0.009	0.019	0.388	-0.013	0.471	-0.083	0.056	-0.009	0.806	-0.143	0.004
STE	0.192	0.000	0.278	0.000	0.273	0.000	0.114	0.401	-0.076	0.490	0.405	0.000	0.154	0.027	-0.016	0.832	0.086	0.189	-0.201	0.221	-0.276	0.099	0.369	0.176
SMO	-0.163	0.000	-0.166	0.000	0.005	0.816	0.042	0.405	-0.192	0.000	0.188	0.000	-0.006	0.870	0.107	0.002	-0.004	0.901	0.015	0.829	0.128	0.020	0.011	0.873
pH	0.105	0.000	0.064	0.000	0.120	0.000	0.049	0.186	0.048	0.103	-0.100	0.003	-0.042	0.068	-0.038	0.119	0.010	0.629	0.027	0.589	-0.185	0.000	0.193	0.000
SEC	-0.135	0.003	-0.175	0.000	-0.194	0.000	-0.030	0.751	-0.140	0.038	0.187	0.023	-0.076	0.280	-0.173	0.023	0.065	0.249	0.276	0.050	-0.160	0.147	-0.075	0.641
SBD	0.227	0.000	0.198	0.000	0.219	0.000	0.208	0.008	0.342	0.000	-0.229	0.000	-0.039	0.518	-0.033	0.489	0.081	0.055	0.020	0.844	0.357	0.000	-0.214	0.022
SOC	-0.076	0.000	-0.069	0.000	-0.049	0.001	-0.012	0.750	-0.004	0.898	0.206	0.000	-0.069	0.004	-0.023	0.338	-0.021	0.298	0.086	0.081	0.010	0.789	-0.021	0.693
STN	-0.087	0.000	-0.063	0.001	-0.110	0.000	-0.081	0.054	-0.029	0.357	0.196	0.000	-0.003	0.908	0.012	0.656	-0.003	0.881	0.124	0.022	0.089	0.033	-0.069	0.211
AVN	0.077	0.056	-0.082	0.024	-0.008	0.804	-0.020	0.779	0.001	0.981	0.011	0.861	0.170	0.001	-0.016	0.733	-0.009	0.821	-0.390	0.001	-0.008	0.919	-0.164	0.031
AMN	-0.110	0.000	-0.085	0.002	-0.095	0.000	-0.010	0.858	-0.178	0.000	0.143	0.005	-0.043	0.233	-0.063	0.066	-0.100	0.001	-0.283	0.000	0.060	0.287	-0.061	0.446
NIN	0.024	0.383	-0.053	0.047	0.023	0.330	0.121	0.032	0.094	0.030	0.207	0.000	0.080	0.023	-0.020	0.556	-0.012	0.702	-0.165	0.024	0.181	0.001	0.066	0.399
STP	-0.002	0.953	0.004	0.871	0.049	0.030	0.017	0.779	0.101	0.015	0.016	0.716	0.005	0.889	0.052	0.134	0.058	0.053	0.072	0.357	0.105	0.065	-0.140	0.050
SAP	-0.047	0.049	-0.078	0.001	-0.029	0.152	0.040	0.413	0.023	0.555	-0.001	0.974	-0.121	0.000	-0.034	0.256	-0.092	0.000	-0.146	0.016	-0.072	0.137	0.016	0.781
STK	-0.208	0.000	-0.164	0.000	-0.004	0.896	0.039	0.671	-0.243	0.000	0.020	0.740	-0.028	0.648	-0.211	0.000	-0.074	0.130	-0.431	0.002	-0.131	0.150	0.206	0.039
SAK	-0.056	0.049	0.108	0.000	0.061	0.015	0.009	0.871	0.238	0.000	0.022	0.650	-0.189	0.000	0.023	0.532	0.023	0.485	-0.114	0.102	-0.014	0.815	-0.213	0.005
MBC	-0.105	0.008	-0.034	0.397	-0.074	0.023	0.204	0.017	-0.007	0.909	0.068	0.389	0.139	0.018	0.012	0.818	-0.102	0.027	0.387	0.002	0.082	0.361	0.511	0.000
MBN	-0.204	0.000	-0.072	0.121	-0.096	0.012	0.088	0.385	-0.081	0.261	0.128	0.259	0.207	0.003	0.000	0.994	0.000	0.998	0.273	0.034	0.141	0.125	0.159	0.309
MBP	-0.150	0.097	-0.029	0.795	-0.235	0.003	0.060	0.771	-0.445	0.005	-0.740	0.036	0.492	0.011	-0.127	0.263	-0.155	0.126	0.430	0.110	-0.566	0.044	-0.145	0.855
DOC	0.008	0.878	-0.122	0.013	-0.090	0.016	0.024	0.794	-0.008	0.934	-0.158	0.136	0.108	0.067	-0.067	0.239	-0.250	0.000	0.215	0.091	0.136	0.169	0.097	0.533
DON	0.151	0.018	0.136	0.109	-0.203	0.000	0.015	0.913	-0.181	0.356	0.230	0.117	-0.212	0.007	0.052	0.507	-0.152	0.033	-0.142	0.408	-0.139	0.386	0.525	0.000
CEC	0.121	0.160	0.009	0.920	-0.160	0.039	0.120	0.500	-0.256	0.076	-0.052	0.739	-0.039	0.684	0.046	0.645	-0.270	0.001	0.018	0.922	-0.050	0.733	0.726	0.000
AGB	-0.134	0.048	-0.087	0.217	-0.182	0.002	-0.142	0.293	-0.409	0.012	0.106	0.389	0.175	0.074	-0.054	0.557	-0.198	0.007	0.511	0.018	0.052	0.805	0.014	0.935
BGB	0.294	0.002	-0.012	0.897	-0.093	0.226	0.146	0.551	0.205	0.231	-0.935	0.000	-0.180	0.105	-0.251	0.034	-0.197	0.029	0.083	0.736	-0.194	0.363	0.349	0.121

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

421 **Table 2**. Continued.

	BOTUr		BCHA		BSHA		BSIM		BACE		BGOC		FOTUr		FCHA		FSHA		FSIM		FACE		FGOC	
	r	*P*	*r*	*P*	*r*	*P*	*r*	*P*	*r*	*P*	*r*	*P*	*r*	*P*	*r*	*P*	*r*	*P*	*r*	*P*	*r*	*P*	*r*	*P*
TOB	-0.122	0.171	-0.026	0.769	0.212	0.004	0.027	0.904	-0.185	0.318	-0.280	0.120	-0.303	0.016	-0.021	0.867	-0.119	0.231	-0.187	0.405	0.604	0.001	-0.054	0.830
TRA	-0.200	0.009	-0.102	0.164	0.182	0.004	-0.088	0.755	-0.141	0.311	0.513	0.010	-0.033	0.634	0.005	0.954	0.106	0.125	-0.201	0.297	0.062	0.713	0.358	0.086
PLD	-0.142	0.085	-0.060	0.394	-0.047	0.455	-0.310	0.011	0.156	0.336	0.340	0.018	-0.054	0.586	0.044	0.602	0.086	0.215	0.181	0.337	-0.100	0.604	-0.836	0.000
DBH	0.199	0.015	0.301	0.000	0.035	0.576	0.078	0.639	-0.162	0.294	0.169	0.490	-0.014	0.885	0.157	0.075	0.006	0.932	-0.416	0.016	-0.228	0.157	-0.639	0.361
TRH	-0.008	0.923	0.242	0.002	-0.182	0.009	-0.124	0.441	-0.643	0.000	-0.122	0.543	0.040	0.671	0.282	0.002	-0.165	0.034	-0.095	0.581	-0.154	0.302	-0.305	0.235
VEC	-0.244	0.000	-0.159	0.011	0.091	0.084	-0.318	0.018	0.045	0.664	-0.065	0.499	0.111	0.187	0.316	0.001	0.192	0.009	0.339	0.144	0.496	0.022	-0.282	0.153
PSR	0.134	0.045	-0.092	0.237	0.034	0.580	0.199	0.275	0.007	0.970	0.018	0.939	0.062	0.331	-0.163	0.099	0.147	0.034	0.243	0.348	-0.434	0.050	-0.507	0.093
PSH	-0.398	0.000	0.442	0.000	0.284	0.000	0.042	0.808	0.517	0.001	0.359	0.078	-0.109	0.179	-0.026	0.803	-0.240	0.001	0.288	0.193	-0.006	0.977	0.351	0.078

422

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

Figs. 5a and b show the soil properties that drove the variations in the relative abundances of bacteria and fungi, respectively. Although the RDA values were generally low for axes 1 and 2, several relative abundances of bacteria (e.g., Proteobacteria) showed nearly the same trends as those of soil properties (e.g., SEC). Therefore, SEC was a key factor influencing the variations in the relative abundance of Proteobacteria. The same trend was observed for the relative abundances of Verrucomicrobiota and CEC. Additional linear regression analysis indicated that the relative abundance of Acidobacteriota was significantly (R^2 = 0.1626, P < 0.001) negatively correlated with soil pH (Table S1), suggesting that low pH was more favorable for Acidobacteriota assemblage. As shown in Fig. 5b, the relative abundances of Ascomycota and Mortierellomycota were very close to those in the SBD, indicating that the SBD was a key soil property driving the variations in the relative abundances of Ascomycota and Mortierellomycota. In general, the directions of variation for all the relative abundances of bacteria and fungi were similar to the directions of variation for the soil properties.

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

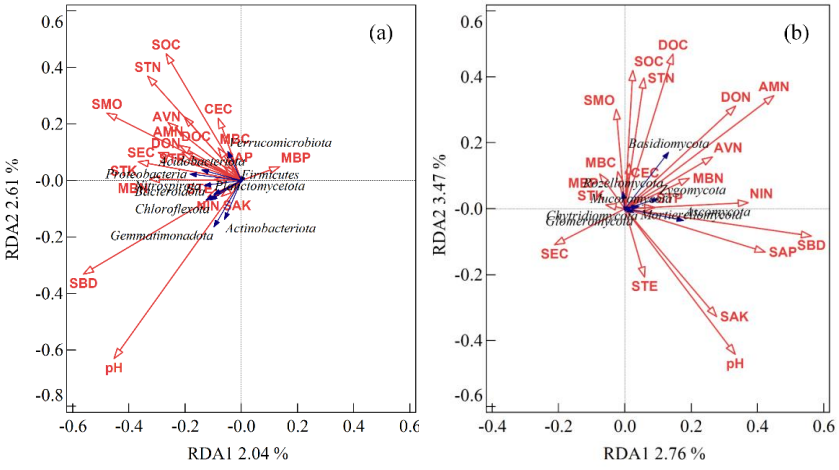


Figure 5. RDA of the relative abundances of bacterial (top ten) and fungal (top eight) components and soil properties. Panels a and b denote bacteria and fungi, respectively.

4 Discussion

4.1 Bacterial and fungal diversities and compositions

Our compilation of soil microbial community data enables researchers to analyze the effects of experimental treatments on the soil microbial community and its spatial and temporal variations. We listed various treatments, such as nitrogen application, warming, precipitation manipulation, and elevated CO_2, in the database, which may increase the data complexity because the treatments may interact with the environmental variables. Treatment effects may induce substantial variations in soil microorganisms and, in some cases, may be greater than site and seasonality effects, which can reduce the correlation coefficients between the soil microbial community and climate, soil and vegetation factors. However, the inclusion of environmental

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

treatments enables researchers to perform meta-analyses concerning soil microbial diversity and composition under different treatments. We added information (e.g., site names and ecosystem types) to the database to enable researchers to investigate where the most abundant bacterial or fungal phylum was located and whether the ecosystem type affected bacterial and fungal diversities and compositions at the global scale.

Fewer microbial taxa were reported at the class and order levels in the database than at the phylum level, despite the former being more interesting to researchers with various backgrounds. Soil function (e.g., carbon and nitrogen cycling) may be induced by certain soil microorganisms at the class or order level. For instance, Gammaproteobacteria are involved in soil nitrogen fixation (Wang et al., 2024). Alphaproteobacteria and Gammaproteobacteria play important roles in soil carbon use efficiency (Allison et al., 2010; Butler et al., 2023; Ding et al., 2025). In the database, the highest and lowest relative abundances of Gammaproteobacteria appeared in the DBF at Sanming, China, and in the GL at Alborz mountains, Iran, respectively. With the future increase in the collected literature, more relative abundance data at the class and order levels can be added to the database and be more conveniently consulted.

Much of the spatial and temporal variations in fungal diversity can be accounted for by bacterial diversity, indicating a tight link between the two microbial taxa. Previous studies have shown that interactions occur between bacteria and fungi at the site and regional scales (George et al. 2019; Baudy et al., 2021; Jiao et al., 2021; Pierce et al., 2021; Zhang et al., 2025). Positive correlations between bacterial and fungal richness widely exist across diverse temperate soil ecosystems (George et al., 2019). Bacterial

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

richness plays important roles in fungal community assembly and functional complementarity (Baudy et al., 2021; Pierce et al., 2021). A recent study by Zhang et al. (2025) revealed positive coupling between bacterial and fungal Shannon diversity in global grassland regions. The positive relationships between bacterial and fungal diversities at the global scale indicate that the two microbial groups are closely connected across various ecosystems. The inferior prokaryotes of bacteria may determine the composition and assembly of more advanced eukaryotes of fungi, or vice versa.

4.2 Drivers of the variations in bacterial and fungal communities

Numerous scholars have investigated the effects of climate, soil and vegetation on microbial communities at the regional scale (Labouyrie et al., 2023; Domeignoz-Horta et al., 2020; Xue et al., 2024). Temperature and precipitation have been shown to be the dominant drivers of microbial diversity at the regional scale (Borowik and Wyszkowska, 2016; Nottingham et al. 2018; Chen et al., 2020). Xue et al. (2024) reported that soil pH, clay content and organic carbon directly affect the microbial community. Acidic conditions (3.0–6.5 pH) are favorable for most Acidobacteria members to maintain optimum growth (Sait et al., 2006; Ward et al., 2009; Kalam et al., 2020). Nutrient (i.e., carbon, nitrogen, phosphorus and potassium) availability affects the microbial community by changing microbial growth and composition (Averill et al., 2019; Dai et al., 2020; Lauber et al., 2008). Labouyrie et al. (2023) reported that fungal α diversity across Europe is driven by vegetation covering the soil rather than by climate and soil properties. In this study, the spatial and temporal changes in bacterial and fungal

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

diversities and compositions tend to occur partially due to alterations in some climate, soil and vegetation variables, although the correlations between bacterial and fungal diversities and compositions and environmental predictors are weak (Table 2 and Fig. 5). Compared with MAT, STE measured at sampling time explains more of the variation in soil bacterial diversity (Table 2), indicating that seasonality plays an important role in structuring bacterial communities. On the basis of our global data analyses, bulk density is a critical soil property that drives spatial and temporal changes in soil microbial communities, particularly for bacteria. In addition, there is a significant correlation between bacteria and vegetation characteristics (e.g., PSH), indicating a close connection between aboveground vegetation and belowground microorganisms.

Our database provides evidence that the combination of climate, soil and vegetation variables is necessary for determining soil microbial communities at the global scale. Similar to the findings of previous studies, our findings demonstrate the complexity of the spatial and temporal variations in soil microbial communities. For instance, environmental predictors explain less of the variation in soil fungal diversity than in soil bacterial diversity (Table 2), increasing the difficulty in accurately simulating soil microbial communities.

4.3 Future improvements

Future improvements will concentrate on the following several aspects.

First, our present version of the database does not include the microbial groups at the family and genus levels. The bacterial and fungal compositions at the family and genus levels in different ecosystems may be more interesting to researchers with varied

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

backgrounds. Including the microbial composition data at the family and genus levels in the database may considerably improve the data quality, and more published literature in the future may contribute to achieving this goal.

Second, although numerous soil property data were collected in this study, soil enzyme data are not included in our database. The soil microbial community may be influenced by soil enzyme activity, particularly that of extracellular enzymes (Barbi et al., 2016; Dollete et al. 2024). In the future, more soil enzyme data associated with the soil microbial community should be collected.

Third, our present database includes literature in which the resulting high-quality sequences were clustered into OTUs at 97 % similarity. Studies (particularly in the year 2024) involving amplicon sequence variant (ASV) methods are not included. Numerous studies that have used the ASV method have been published since 2023 (e.g., Rodríguez et al., 2024; Shulman et al., 2024), and these studies can be compiled to enrich and update our database. A subdatabase including ASV data may be affiliated with the present main database using OTU clustering.

5 Data availability

The database is available via Zenodo at https://doi.org/10.5281/zenodo.16195889 (Chen et al., 2025).

6 Conclusions

We constructed a global database of soil microbial communities and potential influencing factors. The database included geographic information, sampling dates,

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

553 bacterial and fungal diversity indices, experimental treatments, climate factors, soil
554 properties, vegetation characteristics, bacterial and fungal compositions, and site
555 information. The database showed great variability in bacterial and fungal diversities
556 and compositions, but consistent linear regression relationships among the bacterial and
557 fungal diversity indices of OTU richness, Chao1, Shannon, and ACE existed. Potential
558 researchers could use databases to characterize spatial and temporal variations in soil
559 bacterial and fungal diversities and compositions and to analyze their key controlling
560 factors and experimental treatment effects.

561 **Supplement.** Table S1.

562 **Author contributions.** SC and QL designed the structure of the global soil microbial
563 community database. SC compiled the literature used in the database. SC, QL and ZH
564 performed the data analyses. SC wrote the article. QL and ZH reviewed and edited the
565 article.

566 **Competing interests.** The contact author has declared that none of the authors have
567 any competing interests.

568 **Disclaimer.** Publisher's note: Copernicus Publications remains neutral with regard to
569 jurisdictional claims made in the text, published maps, institutional affiliations, or any
570 other geographical representation in this paper. While Copernicus Publications makes
571 every effort to include appropriate place names, the final responsibility lies with the
572 authors.

573 **Acknowledgments.** We gratefully thank the scientists worldwide who contributed their

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

studies to our database.

Financial support. This research was financially supported by the National Natural Science Foundation of China (grant no. 42375114 and 41775151).

References

Allison, S. D., Wallenstein, M. D., and Bradford, M. A.: Soil-carbon response to warming dependent on microbial physiology, Nat. Geosci., 3, 336–340, https://doi.org/10.1038/NGEO846, 2010.

Averill, C., Cates, L. L., Dietze, M. C., and Bhatnagar, J. M.: Spatial vs. temporal controls over soil fungal community similarity at continental and global scales, The ISME J., 13, 2082–2093, https://doi.org/10.1038/s41396-019-0420-1, 2019.

Bahram, M., Hildebrand, F., Forslund, S. K., Anderson, J. L., Soudzilovskaia, N. A., Bodegom, P. M., Bengtsson-Palme, J., Anslan, S., Coelho, L. P., Harend, H., Huerta-Cepas, J., Medema, M. H., Maltz, M. R., Mundra, S., Olsson, P. A., Pent, M., Põlme, S., Sunagawa, S., Ryberg, M., Tedersoo, L., and Bork, P.: Structure and function of the global topsoil microbiome, Nature, 560, 233–237, https://doi.org/10.1038/s41586-018-0386-6, 2018.

Banerjee, S. and Heijden, M. G. A. Soil microbiomes and one health, Nat. Rev. Microbiol., 21, 6–20, https://doi.org/10.1038/s41579-022-00779-w, 2022.

Barbi, F., Prudent, E., Vallon, L., Buée, M., Dubost, A., Legout, A., Marmeisse, R., Fraissinet-Tachet, L., and Luis, P.: Tree species select diverse soil fungal communities expressing different sets of lignocellulolytic enzyme-encoding genes, Soil Biol.

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

595 Biochem., 100, 149–159, https://doi.org/10.1016/j.soilbio.2016.06.008, 2016.

596 Baudy, P., Zubrod, J. P., Konschak, M., Kolbenschlag, S., Pollitt, A., Baschien, C.,
597 Schulz, R., and Bundschuh, M.: Fungal–fungal and fungal–bacterial interactions in
598 aquatic decomposer communities: bacteria promote fungal diversity, Ecology, 102,
599 e03471, https://doi.org/10.1002/ecy.3471, 2021.

600 Borowik, A. and Wyszkowska, J.: Soil moisture as a factor affecting the microbiological
601 and biochemical activity of soil, Plant Soil Environ., 62, 250–255,
602 https://doi.org/10.17221/158/2016-Pse, 2016.

603 Butler, O. M., Manzoni, S., and Warren, C. R.: Community composition and
604 physiological plasticity control microbial carbon storage across natural and
605 experimental soil fertility gradients, The ISME J., 17, 2259–2269,
606 https://doi.org/10.1038/s41396-023-01527-5, 2023.

607 Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and
608 Holmes, S. P.: DADA2: high-resolution sample inference from Illumina amplicon data,
609 Nat. Methods, 13, 581–583, https://doi.org/10.1101/024034, 2016.

610 Chen, Q., Niu, B., Hu, Y., Luo, T., and Zhang, G.: Warming and increased precipitation
611 indirectly affect the composition and turnover of labile-fraction soil organic matter by
612 directly affecting vegetation and microorganisms, Sci. Total Environ., 714, 136787,
613 https://doi.org/10.1016/j.scitotenv.2020.136787, 2020.

614 Chen, S., Li, Q., and Hu, Z.: A global database of soil microbial community and
615 associated climate, soil and vegetation factors, Zenodo [data set],
616 https://doi.org/10.5281/zenodo.16195889, 2025.

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

617 Classen, A. T., Sundqvist, M. K., Henning, J. A., Newman, G. S., Moore, J. A. M.,
618 Cregger, M. A., Moorhead, L. C., and Patterson, C. M.: Direct and indirect effects of
619 climate change on soil microbial and soil microbial-plant interactions: What lies ahead?,
620 Ecosphere, 6, art130, https://doi.org/10.1890/ES15-00217.1, 2015.
621 Dai, Z., Liu, G., Chen, H., Chen, C., Wang, J., Ai, S., Wei, D., Li, D., Ma, B., Tang, C.,
622 Brookes, P. C., and Xu, J.: Long-term nutrient inputs shift soil microbial functional
623 profiles of phosphorus cycling in diverse agroecosystems, The ISME J., 14, 757–770,
624 https://doi. org/10.1038/s4139 6-019-0567-9, 2020.
625 Delgado-Baquerizo, M., Maestre, F. T., Reich, P. B., Jeffries, T. C., Gaitan, J. J., Encinar,
626 D., Berdugo, M., Campbell, C. D., and Singh, B. K.: Microbial diversity drives
627 multifunctionality in terrestrial ecosystems, Nat. Commun., 7, 10541,
628 https://doi.org/10.1038/ncomms10541, 2016.
629 Delgado-Baquerizo, M. Oliverio, A. M.; Brewer, T. E., Benavent-González, A.,
630 Eldridge, D. J., Bardgett, R. D., Maestre, F. T., Singh, B. K., and Fierer, N.: A global
631 atlas of the dominant bacteria found in soil, Science, 359, 320–325,
632 https://doi.org/10.1126/science.aap9516, 2018.
633 Delgado-Baquerizo, M., Reich, P. B., Trivedi, C., El dridge, D. J., Abades, S., Alfaro,
634 F. D., Bastida, F., Berhe, A. A., Cutler, N. A., Gallardo, A., García-Velázquez, L.,
635 Hart, S. C., Hayes, P. E., He, J., Hseu, Z., Hu, H., Kirchmair, M., Neuhauser, S., Pérez,
636 C. A., Reed, S. C., Santos, F., Sullivan, B. W., Trivedi, P., Wang, J., Weber-Grullon, L.,
637 Williams, M. A., and Singh, B. K.: Multiple elements of soil biodiversity drive
638 ecosystem functions across biomes, Nat. Ecol. Evol., 4, 210–220,

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

https://doi.org/10.1038/s41559-019-1084-y, 2020.

Ding, J., Liu, F., Zeng, J. Gu, H., Huang, J., Wu, B., Shu, L., Yan, Q., He, Z., and Wang, C.: Depth heterogeneity of lignin-degrading microbiome and organic carbon processing in mangrove sediments, npj Biofilms Microbi., 11, 5, https://doi.org/10.1038/s41522-024-00638-x, 2025.

Dollete, D., Lumactud, R. A., Carlyle, C. N., Szczyglowski, K., and Thilakarathna, M. S.: Effects of defoliation on root traits, nitrogen fixation, soil nitrogen availability, soil enzyme activities and soil bacterial communities of forage legumes, Plant Soil, https://doi.org/10.1007/s11104-024-06802-8, 2024.

Domeignoz-Horta, L. A., Pold, G., Liu, X.-J. A., Frey, S. D., Melillo, J. M., and DeAngelis, K. M.: Microbial diversity drives carbon use efficiency in a model soil, Nat. Commun., 11, 3684, https://doi.org/10.1038/s41467-020-17502-z, 2020.

Edgar, R. C.: UPARSE, highly accurate OTU sequences from microbial amplicon reads, Nat. Methods., 10, 996–999, https://doi.org/10.1038/nmeth.2604, 2013.

Fierer, N. and Jackson, R. B.: The diversity and biogeography of soil bacterial communities, Proc. Natl Acad. Sci. U.S.A., 103, 626–631, https://doi.org/10.1073/pnas.0507535103, 2006.

George, P. B. L., Lallias, D., Creer, S., Seaton, F. M., Kenny, J. G., Eccles, R. M., Griffiths, R. I., Lebron, I., Emmett, B. A., Robinson, D. A., and Jones, D. L.: Divergent national-scale trends of microbial and animal biodiversity revealed across diverse temperate soil ecosystems, Nat. Commun., 10, 1107, https://doi.org/10.1038/s41467-019-09031-1, 2019.

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

661 Green, J. L., Bohannan, B. J., and Whitaker, R. J.: Microbial biogeography: from
662 taxonomy to traits, Science, 320, 1039–1043, https://doi.org/10.1126/science.1153475,
663 2008.
664 Hanson, C. A., Allison, S. D., Bradford, M. A., Wallenstein, M. D., and Treseder, K. K.:
665 Fungal taxa target different carbon sources in forest soil, Ecosystems, 11, 1157–1167,
666 https://doi.org/10.1007/s10021-008-9186-4, 2008.
667 Jiao, S., Peng, Z., Qi, J., Gao, J., and Wei, G.: Linking bacterial-fungal relationships to
668 microbial diversity and soil nutrient cycling, mSystems, 6, e01052-20,
669 https://doi.org/10.1128/mSystems.01052-20, 2021.
670 Jones, S. E. and Lennon, J. T.: Dormancy contributes to the maintenance of microbial
671 diversity, Proc. Natl Acad. Sci. U.S.A., 107, 5881–5886,
672 https://doi.org/10.1073/PNAS.0912765107, 2010.
673 Kalam, S., Basu, A., Ahmad, I., Sayyed, R. Z., El-Enshasy, H. A., Dailin, D. J.,
674 and Suriani, N. L.: Recent understanding of soil Acidobacteria and their ecological
675 significance: A critical review, Front. Microbio., 11,
676 580024, https://doi.org/10.3389/fmicb.2020.580024, 2020.
677 Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J., and Knight, R.: UniFrac: an
678 effective distance metric for microbial community comparison, The ISME J., 5, 169–
679 172, https://doi.org/10.1038/ismej.2010.133, 2011.
680 Labouyrie, M., Ballabio, C., Romero, F., Panagos, P., Jones, A., Schmid, M. W.,
681 Mikryukov, V., Dulya, O., Tedersoo, L., Bahram, M., Lugato, E., van der Heijden, M.
682 G. A., and Orgiazzi, A.: Patterns in soil microbial diversity across Europe, Nat.

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

683 Commun., 14, 3311, https://doi.org/10.1038/s41467-023-37937–4, 2023.

684 Lauber, C. L., Strickland, M. S., Bradford, M. A., and Fierer, N.: The influence of soil
685 properties on the structure of bacterial and fungal communities across land-use types,
686 Soil Biol. Biochem., 40, 2407–2415, https://doi.org/10.1016/j.soilbio.2008.05.021,
687 2008.

688 Lu, T., Xu, N., Lei, C., Zhang, Q., Zhang, Z., Sun, L., He, F., Zhou, N., Peñuelas, J.,
689 Zhu, Y., and Qian, H.: Bacterial biogeography in China and its association to land use
690 and soil organic carbon, Soil Ecol. Lett., 5, 230172, https://doi.org/10.1007/s42832-
691 023-0172-8, 2023.

692 Nottingham, A. T., Fierer, N., Turner, B. L., Whitaker, J., Ostle, N. J., McNamara, N.
693 P., Bardgett, R. D., Leff, J. W., Salinas, N., Silman, M. R., Kruuk, L. E. B., and Meir,
694 P.: Microbes follow Humboldt: Temperature drives plant and soil microbial diversity
695 patterns from the Amazon to the Andes, Ecology, 99, 2455–2466,
696 https://doi.org/10.1002/ecy.2482, 2018.

697 Peng, Z., Qian, X., Liu, Y., Li, X., Gao, H., An, Y., Qi, J., Jiang, L., Zhang, Y., Chen, S.,
698 Pan, H., Chen, B., Liang, C., van der Heijden, M. G. A., Wei, G., and Jiao, S.: Land
699 conversion to agriculture induces taxonomic homogenization of soil microbial
700 communities globally, Nat. Commun., 15, 3624, https://doi.org/10.1038/s41467-024-
701 47348-8, 2024.

702 Pierce, E. C., Morin, M., Little, J. C., Liu, R. B., Tannous, J., Keller, N. P., Pogliano,
703 K., Wolfe, B. E., Sanchez, L. M., and Dutton, R. J.: Bacterial–fungal interactions
704 revealed by genome-wide analysis of bacterial mutant fitness, Nat. Microbiol., 6, 87–

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

705 102, https://doi.org/10.1038/s41564-020-00800-z, 2021.

706 Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N. N., Anderson, I. J., and Cheng, J.:
707 Insights into the phylogeny and coding potential of microbial dark matter, Nature, 499,
708 431–437, https://doi.org/10.1038/nature12352, 2013.

709 Rodríguez, V., Bartholomäus, A., Witzgall, K., Riveras-Muñoz, N., Oses R, Liebner, S.,
710 Kallmeyer, J., Rach, O., Mueller, C. W., Seguel, O., Scholten, T., and Wagner, D.:
711 Microbial impact on initial soil formation in arid and semiarid environments under
712 simulated climate change, Front. Microbiol. 15, 1319997,
713 https://doi.org/10.3389/fmicb.2024.1319997, 2024.

714 Sait, M., Davis, K. E., and Janssen, P. H.: Effect of pH on isolation and distribution of
715 members of subdivision 1 of the phylum Acidobacteria occurring in soil, Appl. Environ.
716 Microbiol., 72, 1852–1857, https://doi.org/10.1128/aem.72.3.1852-1857.2006, 2006.

717 Schmieder, R. and Edwards, R.: Quality control and preprocessing of metagenomics
718 datasets, Bioinformatics, 27, 863–864, https://doi.org/10.1093/bioinformatics/btr026,
719 2011.

720 Shulman, H. B., Aronson, E. L., Dierick, D., Pinto-Tomás, A. A., Botthoff, J. K.,
721 Artavia-León, A., and Allen, M. F.: Leafcutter ants enhance microbial drought
722 resilience in tropical forest soil, Env. Microbiol. Rep., 16, e13251,
723 https://doi.org/10.1111/1758-2229.13251, 2024.

724 van der Heijden, M. G. A., Bardgett, R. D., and van Straalen, N. M.: The unseen
725 majority: soil microbes as drivers of plant diversity and productivity in terrestrial
726 ecosystems, Ecol. Lett., 11, 296–310, https://doi.org/10.1111/j.1461-

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

0248.2008.01199.x, 2008.

Wang, S., Jiang, L., Zhao, Z., Zhen, C., Wang, J., Alain, K., Cui, L., Zhong, Y., Peng, Y., Lai, Q., Dong, X., and Shao, Z.: Chemolithoautotrophic diazotrophs dominate dark nitrogen fixation in mangrove sediments, The ISME J., 18, wrae119, https://doi.org/10.1093/ismejo/wrae119, 2024.

Ward, N. L., Challacombe, J. F., Janssen, P. H., Henrissat, B., Coutinho, P. M., Wu, M., Xie, G., Haft, D. H., Sait, M., Badger, J., Barabote, R. D., Bradley, B., Brettin, T. S., Brinkac, L. M., Bruce, D., Creasy, T., Daugherty, S. C., Davidsen, T. M., DeBoy, R. T., Detter, J. C., Dodson, R. J., Durkin, A. S., Ganapathy, A., Gwinn-Giglio, M., Han, C. S., Khouri, H., Kiss, H., Kothari, S. P., Madupu, R., Nelson, K. E., Nelson, W. C., Paulsen, I., Penn, K., Ren, Q., Rosovitz, M. J., Selengut, J. D., Shrivastava, S., Sullivan, S. A., Tapia, R., Thompson, L. S., Watkins, K. L., Yang, Q., Yu, C., Zafar, N., Zhou, L., and Kuske, C. R.: Three genomes from the phylum Acidobacteria provide insight into the lifestyles of these microorganisms in soils, Appl. Environ. Microbiol., 75, 2046–2056, https://doi.org/10.1128/aem.02294-08, 2009.

Xue, P., Minasny, B., Wadoux, A. M. J.-C., Dobarco, M. R., McBratney, A., Bissett, A., and de Caritat, P.: Drivers and human impacts on topsoil bacterial and fungal community biogeography across Australia, Global Change Biol., 30, e17216, https://doi.org/10.1111/gcb.17216, 2024.

Zhang, T., Li, S., Xing, T., Liu, J., Sun, Z., Yao, Z., Tang, S., and Jin, K.: The coupling of bacterial and fungal diversity under land-use conversion in global grassland regions is limited by perturbation intensity, Soil Till. Res., 254, 106751,

https://doi.org/10.5194/essd-2025-501
Preprint. Discussion started: 8 October 2025

749 https://doi.org/10.1016/j.still.2025.106751, 2025.

750 Zubek S., Rożek, K., Chmolowska, D., Odriozola, I., Větrovský, T., Skubała, K., Dobler,

751 P. T., Stefanowicz, A. M., Stanek, M., Orzechowska, A., Kohout, P., and Baldrian, P.:

752 Dominant herbaceous plants contribute to the spatial heterogeneity of beech and

753 riparian forest soils by influencing fungal and bacterial diversity, Soil Biol. Biochem.,

754 193, 109405, https://doi.org/10.1016/j.soilbio.2024.109405, 2024.