Supplementary Material

Community composition, not diversity, governs soil multifunctionality along an ant-nesting gradient in a revegetated desert biocrust system

Peng Zhang, Yanxia Pan, Zengru Wang

This file contains Supplementary Methods S1 and S2, Supplementary Figures S1–S7, and Supplementary Tables S1–S7.

Supplementary Methods

The two sections below give the full protocols for the general soil and molecular determinations, which were moved here from Sections 2.3 and 2.4 of the main text to save space. The references cited in these two sections are listed at the end of this section.

Methods S1. Soil physicochemical measurements

Measurements were performed on the soil dataset described in Section 2.2 of the main text. Variables spanned soil physical structure, acid-base and ionic properties, macronutrients, and multiple organic carbon fractions.

Physical properties. Soil water content (SWC, %) was measured *in situ* concurrently with field sampling using a portable time-domain reflectometer (TDR 300, Spectrum Technologies, Plainfield, IL, USA) at a depth of 0–5 cm, with three readings averaged per sampling point (Topp et al., 1980). Soil particle-size composition (sand, 0.05–2 mm; silt, 0.002–0.05 mm; clay, < 0.002 mm) was determined by the wet-sieving method (Klute, 1986). Soil aggregate stability was characterized by the dry-sieving mean weight diameter (MWD_dry, mm) and the proportion of water-stable macroaggregates (Macro_agg_pct, %), determined using a series of standard soil sieves with progressively decreasing mesh sizes (2, 1, 0.5, 0.25, and 0.053 mm) (Dane and Topp, 2002).

Acid-base and ionic properties. Soil pH and electrical conductivity (EC, μS cm^{-1}) were determined in a 1:5 (w/v) soil-to-water suspension using glass-electrode potentiometry and conductivity measurement, respectively, on a Mettler Toledo (Greifensee, Switzerland) multiparameter analyzer.

Macronutrients. Total nitrogen (TN, g kg^{-1}) was measured directly using an elemental analyzer (Vario TOC Cube; Elementar, Hanau, Germany). Total phosphorus (TP, g kg^{-1}) was determined using the molybdenum-antimony-ascorbic acid colorimetric method following H_2SO_4 and $HClO_4$ digestion (Sparks et al., 1996).

Organic carbon fractions. A suite of organic carbon fractions was quantified to represent soil carbon storage and cycling: total organic carbon (TOC), dissolved organic carbon (DOC), microbial biomass carbon (MBC), easily oxidizable carbon (EOC), light- and heavy-fraction organic carbon (LFOC, HFOC), particulate and mineral-associated organic carbon (POC, MAOC), and active and recalcitrant carbon (AOC, RC). Briefly, DOC was extracted with deionized water (1:5 w/v), filtered through a 0.45-μm membrane, and quantified on a TOC analyzer (TOC-L, Shimadzu, Kyoto, Japan)

(Jones and Willett, 2006); MBC was determined by chloroform fumigation and K_2SO_4 (0.5 M) extraction (24-h fumigation, k_EC = 0.45) (Vance et al., 1987); and POC and MAOC were obtained by size fractionation at 53 µm following Lavallee et al. (2020). TOC was determined by potassium dichromate oxidation with external heating (Sparks et al., 1996), and EOC by 333 mmol L^{-1} $KMnO_4$ oxidation with colorimetric detection (Blair et al., 1995). AOC was determined by permanganate oxidation, and RC was calculated as the difference between TOC and AOC. LFOC and HFOC were separated by density fractionation in NaI solution (ρ = 1.8 g cm^{-3}).

Methods S2. DNA extraction, amplicon library preparation and high-throughput sequencing

Total microbial DNA extraction. Fresh soil (approximately 0.5 g) from each cryopreserved sample was used for total microbial genomic DNA extraction using the FastDNA™ SPIN Kit for Soil (MP Biomedicals, Solon, OH, USA), following the manufacturer's instructions. DNA concentration and purity were assessed using a NanoDrop 2000 UV-visible spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), and DNA integrity was verified by 1% agarose gel electrophoresis. All samples exhibited an OD_{260}/OD_{280} ratio between 1.8 and 2.0 and a DNA concentration of at least 20 ng Ml^{-1}, satisfying the requirements for downstream PCR amplification. DNA samples that passed quality control were stored at −20 °C until use.

PCR amplification of the 16S and 18S rRNA genes. The bacterial community was characterized by amplifying the V4 hypervariable region of the 16S rRNA gene using the primer pair 515F (5′-GTGYCAGCMGCCGCGGTAA-3′) and 806R (5′-GGACTACNVGGGTWTCTAAT-3′), carrying sample-specific 8-bp barcodes (Caporaso et al., 2011). The eukaryotic microbial community was characterized by amplifying the V4 hypervariable region of the 18S rRNA gene using the primer pair 528F (5′-GCGGTAATTCCAGCTCCAA-3′) and 706R (5′-AATCCRAGAATTTCACCTCT-3′) (Cheung et al., 2010; Stoeck et al., 2010).

The PCR reaction mixture (30 µL) contained 15 µL of Phusion® High-Fidelity PCR Master Mix (New England Biolabs, Ipswich, MA, USA), 1 µL each of 0.2 µM forward and reverse primers, 10 ng of template DNA, and ddH_2O to volume. The thermocycling program consisted of an initial denaturation step at 98 °C for 1 min, followed by 30 cycles of denaturation at 98 °C for 10 s, annealing at 50 °C for 30 s, and extension at 72 °C for 30 s, with a final extension step at 72 °C for 5 min.

Library construction and Illumina sequencing. PCR products that passed verification by 2% agarose gel electrophoresis were purified using the GeneJET Gel Extraction Kit (Thermo Fisher Scientific) and quantified using a Qubit 3.0 fluorometer. PCR products from individual samples were pooled in equimolar amounts, and Illumina paired-end sequencing libraries were constructed using the standard sequence of end-repair, A-tailing, adapter ligation, and library amplification. Libraries that passed dual quantification by Qubit and qPCR were sequenced on the Illumina NovaSeq 6000 platform in the PE250 paired-end mode (Novogene Bioinformatics Technology Co., Ltd., Beijing, China).

References cited in the Supplementary Methods

Blair, G.J., Lefroy, R.D.B., Lisle, L., 1995. Soil carbon fractions based on their degree of oxidation, and the development of a carbon management index for agricultural systems. Australian Journal of Agricultural Research 46, 1459–1466. https://doi.org/10.1071/AR9951459

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., Fierer, N., Knight, R., 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proceedings of the National Academy of Sciences 108, 4516–4522. https://doi.org/10.1073/PNAS.1000080107

Cheung, M.K., Au, C.H., Chu, K.H., Kwan, H.S., Wong, C.K., 2010. Composition and genetic diversity of picoeukaryotes in subtropical coastal waters as revealed by 454 pyrosequencing. The ISME Journal 4, 1053–1059. https://doi.org/10.1038/ISMEJ.2010.26

Dane, J.H., Topp, G.C. (Eds.), 2002. Methods of Soil Analysis. Part 4. Physical Methods. SSSA Book Series 5. https://doi.org/10.2136/SSSABOOKSER5.4

Jones, D.L., Willett, V.B., 2006. Experimental evaluation of methods to quantify dissolved organic nitrogen (DON) and dissolved organic carbon (DOC) in soil. Soil Biology and Biochemistry 38, 991–999. https://doi.org/10.1016/J.SOILBIO.2005.08.012

Klute, A. (Ed.), 1986. Methods of Soil Analysis. Part 1. Physical and Mineralogical Methods, 2nd ed. SSSA Book Series 5. https://doi.org/10.2136/SSSABOOKSER5.1.2ED

Lavallee, J.M., Soong, J.L., Cotrufo, M.F., 2020. Conceptualizing soil organic matter into particulate and mineral-associated forms to address global change in the 21st century. Global Change Biology 26, 261–273. https://doi.org/10.1111/GCB.14859

Sparks, D.L., Page, A.L., Helmke, P.A., Loeppert, R.H., Soltanpour, P.N., Tabatabai, M.A., Johnston, C.T., Sumner, M.E. (Eds.), 1996. Methods of Soil Analysis. SSSA Book Series 5. https://doi.org/10.2136/SSSABOOKSER5.3

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M.D.M., Breiner, H.W., Richards, T.A., 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. Molecular Ecology 19, 21–31. https://doi.org/10.1111/J.1365-294X.2009.04480.X

Topp, G.C., Davis, J.L., Annan, A.P., 1980. Electromagnetic determination of soil water content: Measurements in coaxial transmission lines. Water Resources Research 16, 574–582. https://doi.org/10.1029/WR016I003P00574

Vance, E.D., Brookes, P.C., Jenkinson, D.S., 1987. An extraction method for measuring soil microbial biomass C. Soil Biology and Biochemistry 19, 703–707. https://doi.org/10.1016/0038-0717(87)90052-6

Supplementary Figures

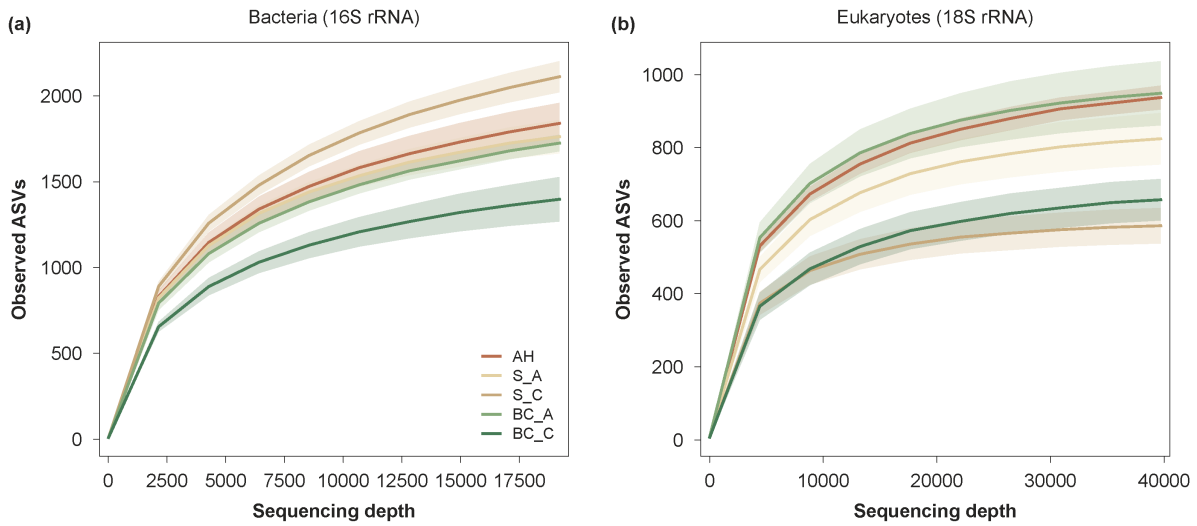


Fig. S1. Rarefaction curves of amplicon sequence variant (ASV) richness against sequencing depth for the bacterial (16S rRNA) and eukaryotic (18S rRNA) datasets across five microhabitats. Tables were rarefied to 19,210 (16S) and 39,703 (18S) reads per sample. Curves approaching an asymptote indicate that the sequencing depth was sufficient to capture community diversity. Microhabitat abbreviations: AH, ant hill soil; S_A, subsurface soil beneath ant hill-covered biocrust; S_C, subsurface soil beneath control biocrust; BC_A, ant hill-covered biocrust; BC_C, control biocrust.

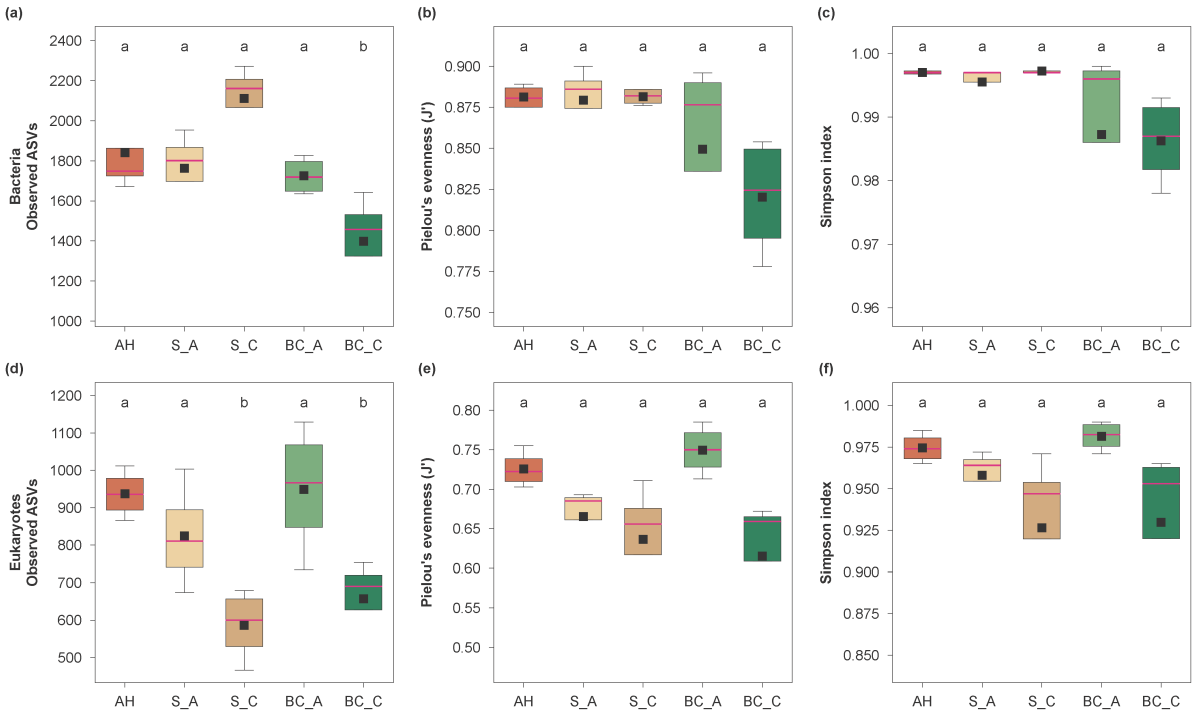


Fig. S2. Additional alpha-diversity indices (Observed ASVs, Pielou's evenness, Simpson index, and dominance) for the bacterial (16S) and eukaryotic (18S) communities across the five microhabitats (mean ± SD, n = 4). Microhabitat abbreviations as in Fig. S1.

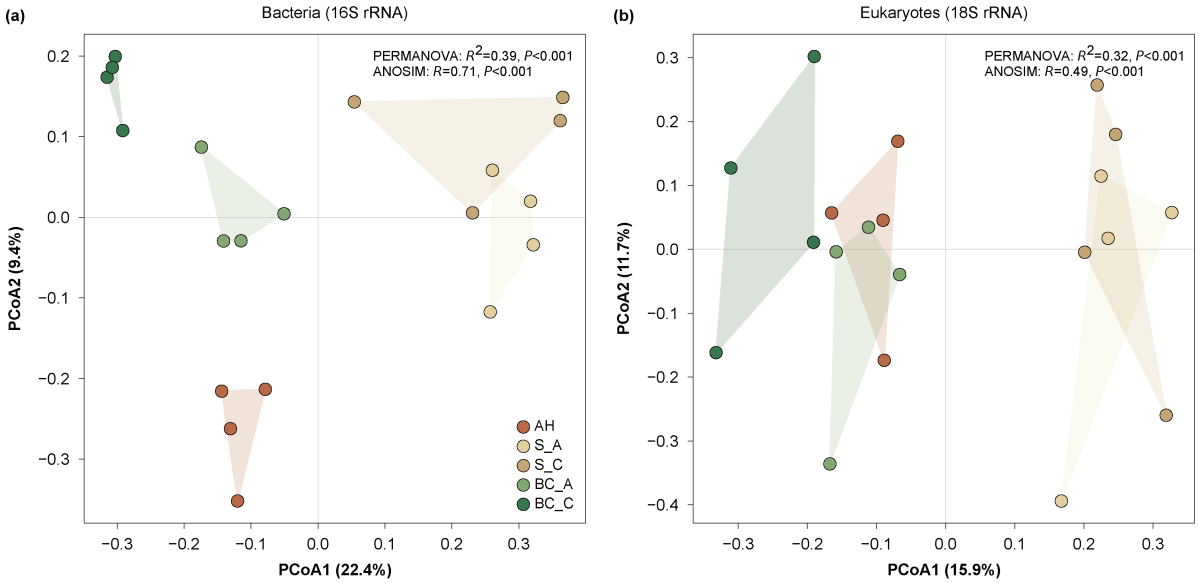


Fig. S3. Principal coordinate analysis (PCoA) based on Bray-Curtis dissimilarities of the bacterial (16S) and eukaryotic (18S) communities across the five microhabitats. Microhabitat abbreviations as in Fig. S1.

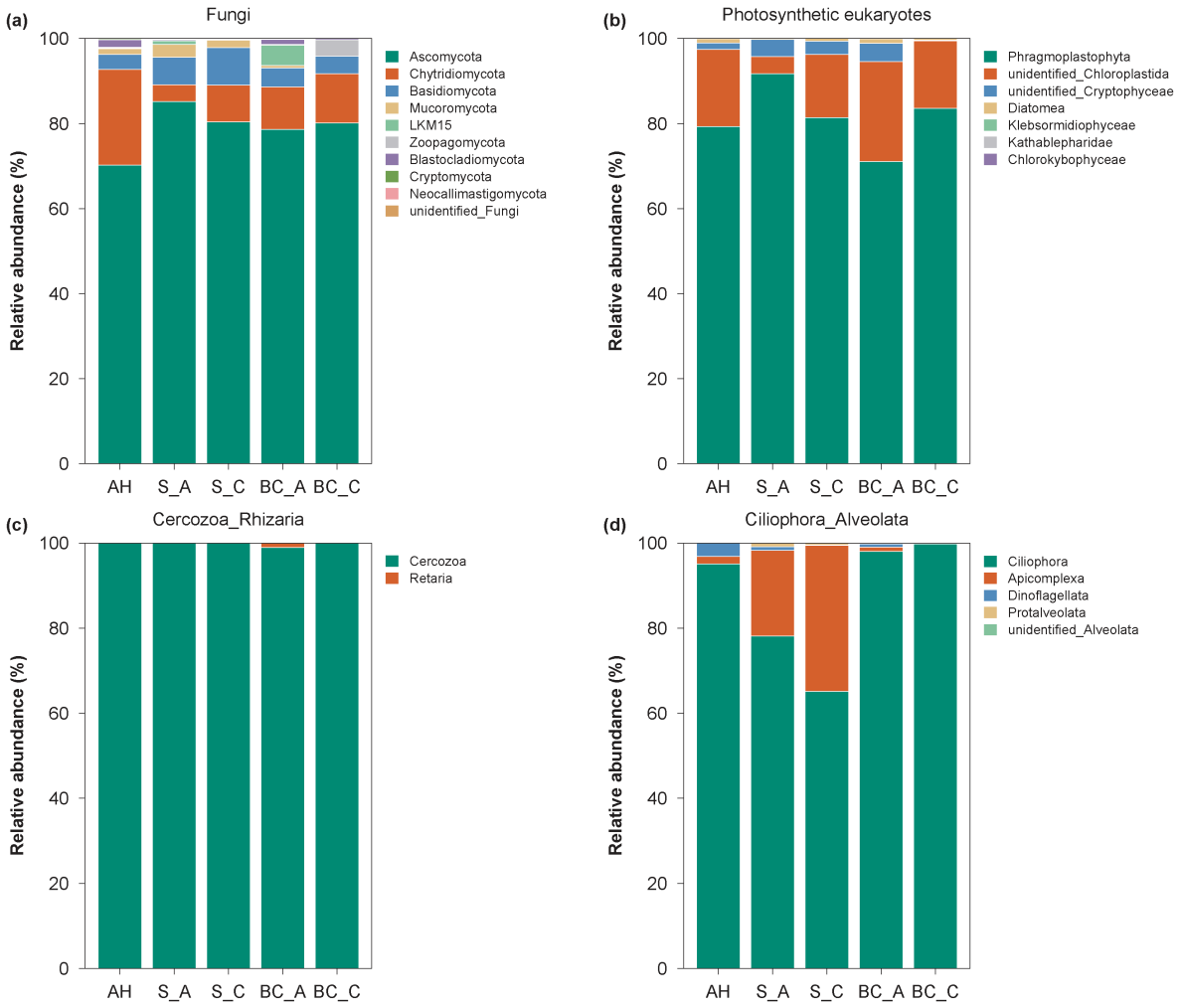


Fig. S4. Phylum-level composition of the main eukaryotic functional groups (for example, Fungi, photosynthetic eukaryotes, Rhizaria, and Alveolata) across the five microhabitats. Microhabitat abbreviations as in Fig. S1.

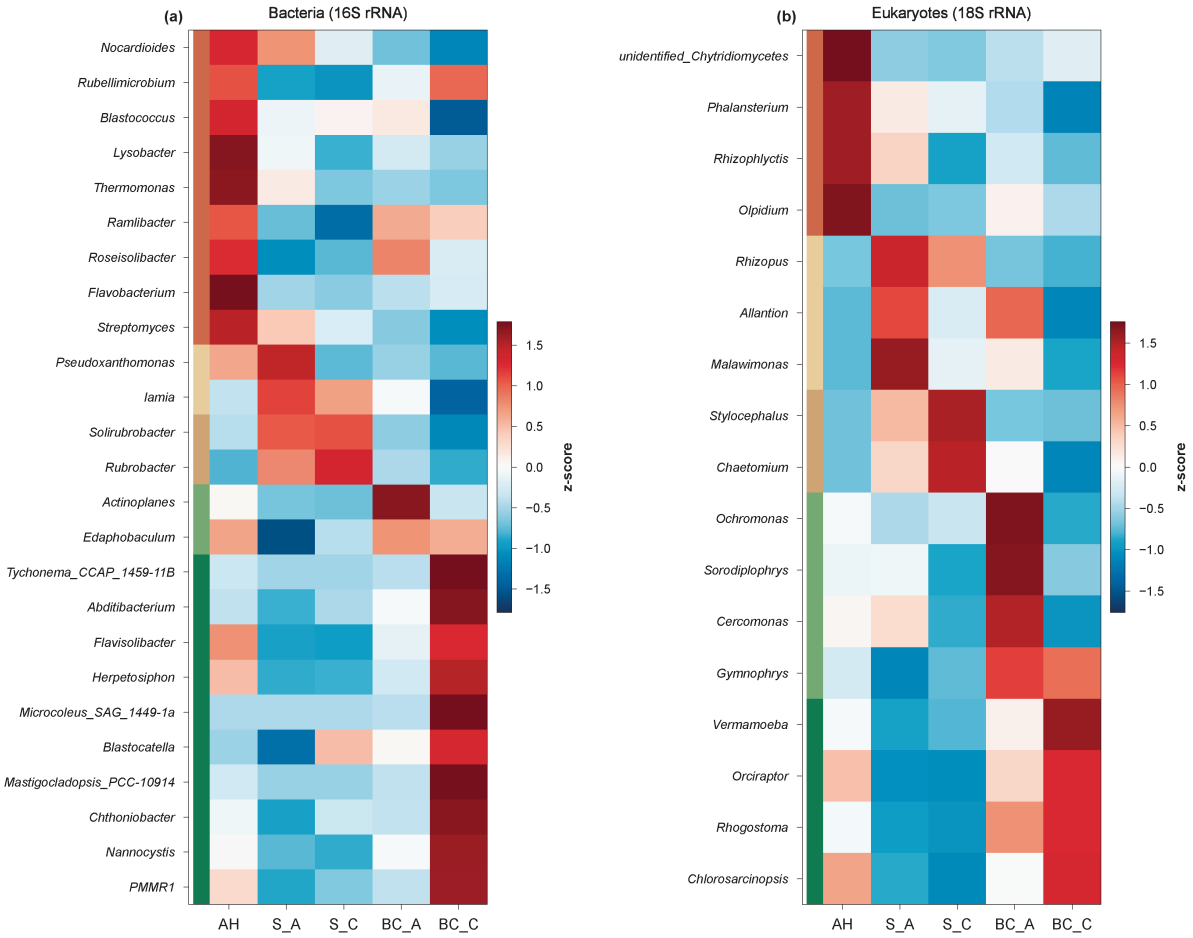


Fig. S5. Heatmap of genus-level indicator taxa for the bacterial (16S) and eukaryotic (18S) communities across the five microhabitats. The control biocrust was characterized by filamentous cyanobacterial genera such as *Microcoleus* and *Tychonema*, and the ant hill soil by oligotrophic actinobacterial and proteobacterial genera. Microhabitat abbreviations as in Fig. S1.

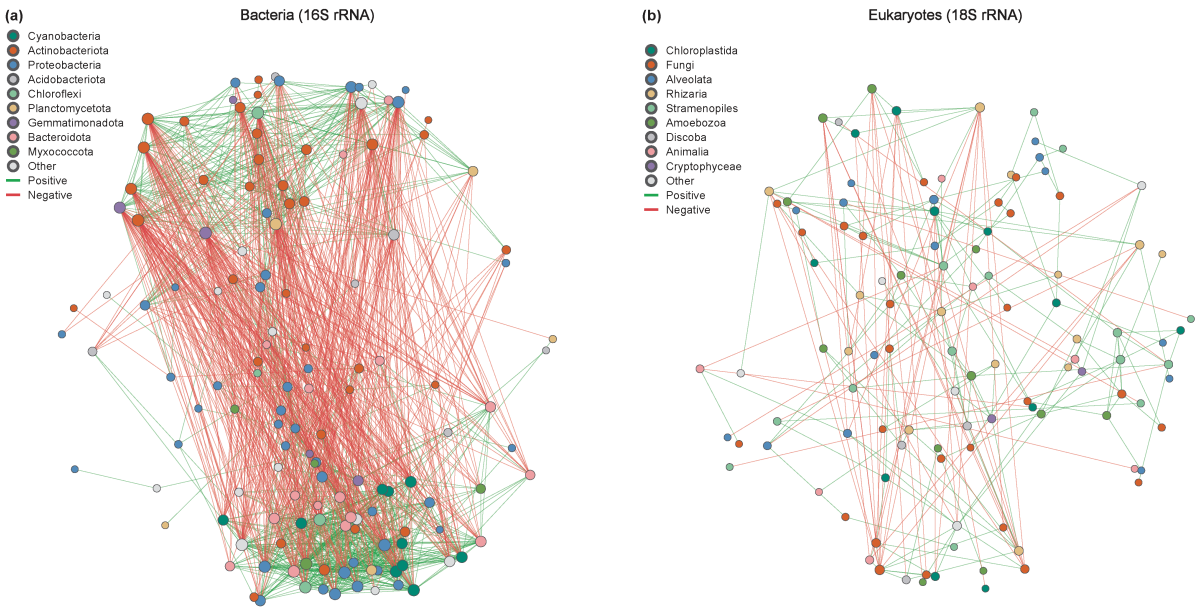


Fig. S6. Single-domain co-occurrence networks inferred separately for the bacterial (16S) and eukaryotic (18S) communities. The eukaryotic network is markedly sparser than the bacterial one. Microhabitat abbreviations as in Fig. S1.

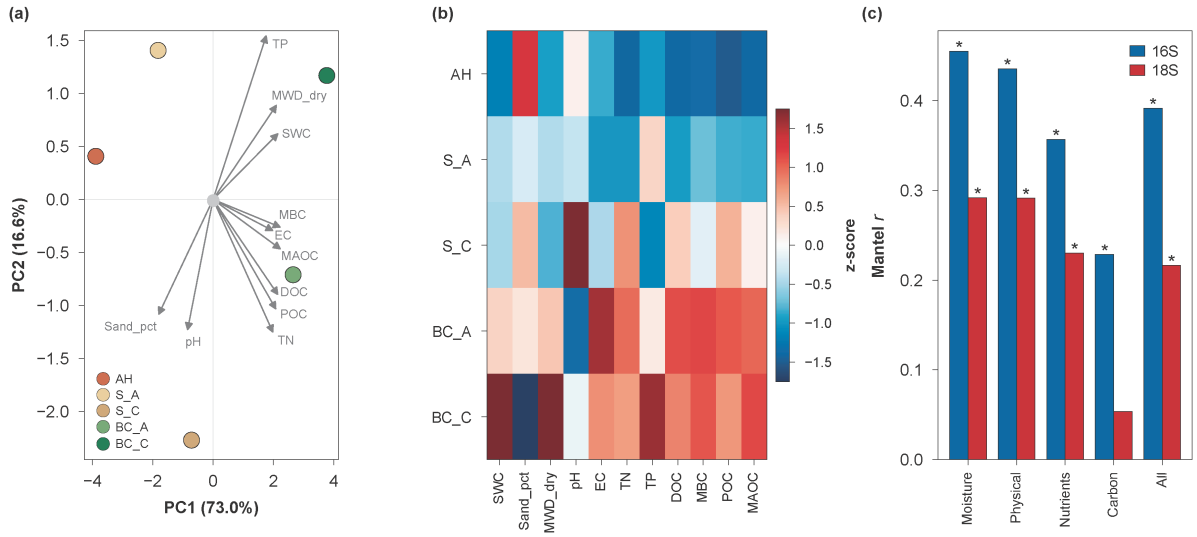


Fig. S7. Environmental drivers of community variation. Principal component analysis (PCA) and a z-score heatmap summarize the soil variables across microhabitats, and Mantel tests relate environmental subsets (Moisture, Physical, Nutrients, and Carbon fractions) to the bacterial (16S) and eukaryotic (18S) community distance matrices. Microhabitat abbreviations as in Fig. S1.

Supplementary Tables

Table S1. Alpha diversity indices across the five microhabitats (mean ± SD, n = 4).

Bacteria (16S rRNA)

Index	AH	S_A	S_C	BC_A	BC_C
Chao1	2173 ± 275 (a)	1980 ± 311 (a)	2549 ± 324 (a)	2064 ± 89 (a)	1689 ± 450 (b)
Observed ASVs	1840 ± 239 (a)	1762 ± 195 (a)	2112 ± 181 (a)	1725 ± 96 (a)	1398 ± 259 (b)
Shannon	9.55 ± 0.22 (a)	9.47 ± 0.23 (a)	9.73 ± 0.15 (a)	9.14 ± 0.77 (a)	8.55 ± 0.38 (b)
Pielou's J'	0.881 ± 0.007 (a)	0.879 ± 0.024 (a)	0.881 ± 0.005 (a)	0.850 ± 0.068 (a)	0.820 ± 0.037 (a)
Simpson	0.997 ± 0.001 (a)	0.996 ± 0.003 (a)	0.997 ± 0.001 (a)	0.987 ± 0.019 (a)	0.986 ± 0.007 (a)
Dominance	0.003 ± 0.001 (a)	0.004 ± 0.003 (a)	0.003 ± 0.001 (a)	0.013 ± 0.019 (a)	0.014 ± 0.007 (a)

Eukaryotes (18S rRNA)

Index	AH	S_A	S_C	BC_A	BC_C
Chao1	995 ± 78 (a)	851 ± 147 (a)	603 ± 109 (b)	986 ± 197 (a)	684 ± 118 (b)
Observed ASVs	937 ± 65 (a)	824 ± 141 (a)	586 ± 97 (b)	949 ± 176 (a)	658 ± 113 (b)
Shannon	7.16 ± 0.24 (a)	6.44 ± 0.57 (a)	5.86 ± 0.87 (b)	7.41 ± 0.50 (a)	5.77 ± 1.03 (b)
Pielou's J'	0.726 ± 0.023 (a)	0.665 ± 0.045 (a)	0.636 ± 0.080 (a)	0.750 ± 0.033 (a)	0.615 ± 0.096 (a)
Simpson	0.975 ± 0.009 (a)	0.958 ± 0.018 (a)	0.926 ± 0.058 (a)	0.982 ± 0.009 (a)	0.930 ± 0.055 (a)
Dominance	0.025 ± 0.009 (a)	0.042 ± 0.018 (a)	0.073 ± 0.058 (a)	0.018 ± 0.009 (a)	0.070 ± 0.055 (a)

Different lowercase letters denote significant differences among microhabitats (one-way ANOVA with Tukey's HSD, α = 0.05). Microhabitat abbreviations are shown in Fig. S1.

Table S2. Pairwise PERMANOVA of community composition (Bray-Curtis dissimilarity, 9999 permutations, FDR-corrected).

Dataset	Pair	F	R^2	P	P (FDR)	Sig.
16S Bacteria	AH-BC_A	1.742	0.225	0.027	0.034	*
16S Bacteria	AH-S_A	2.789	0.317	0.027	0.034	*
16S Bacteria	AH-BC_C	2.609	0.303	0.030	0.034	*
16S Bacteria	AH-S_C	3.016	0.335	0.029	0.034	*
16S Bacteria	BC_A-S_A	2.410	0.287	0.031	0.034	*
16S Bacteria	BC_A-BC_C	1.793	0.230	0.027	0.034	*
16S Bacteria	BC_A-S_C	2.016	0.252	0.028	0.034	*
16S Bacteria	S_A-BC_C	3.838	0.390	0.029	0.034	*
16S Bacteria	S_A-S_C	0.935	0.135	0.649	0.649	ns
16S Bacteria	BC_C-S_C	3.368	0.359	0.028	0.034	*
18S Eukaryotes	AH-BC_A	1.240	0.171	0.175	0.195	ns
18S Eukaryotes	AH-S_A	1.753	0.226	0.029	0.039	*
18S Eukaryotes	AH-BC_C	1.699	0.221	0.030	0.039	*
18S Eukaryotes	AH-S_C	1.971	0.247	0.026	0.039	*
18S Eukaryotes	BC_A-S_A	1.622	0.213	0.031	0.039	*
18S Eukaryotes	BC_A-BC_C	1.650	0.216	0.031	0.039	*
18S Eukaryotes	BC_A-S_C	2.060	0.256	0.029	0.039	*
18S Eukaryotes	S_A-BC_C	2.462	0.291	0.028	0.039	*
18S Eukaryotes	S_A-S_C	0.811	0.119	0.747	0.747	ns
18S Eukaryotes	BC_C-S_C	2.240	0.272	0.029	0.039	*

Microhabitat abbreviations as in Fig. S1.

Table S3. Topological properties of the five microhabitat-specific cross-domain sub-networks

Microhabitat	Nodes	Edges	Density	Avg. degree	Modularity	Avg. clustering	Avg. path length	Pos:neg ratio	Cross-domain edges
AH	141	273	0.028	3.872	0.383	0.141	2.870	1.874	69
S_A	123	291	0.039	4.732	0.311	0.149	2.879	2.464	59
S_C	115	273	0.042	4.748	0.295	0.162	2.557	2.689	52
BC_A	129	195	0.024	3.023	0.409	0.117	2.977	1.910	50
BC_C	111	169	0.028	3.045	0.355	0.130	2.839	4.281	41

Microhabitat abbreviations as in Fig. S1.

Table S4. Soil multifunctionality index (EMF): variable set and microhabitat values

(a) Variable set (15 soil functional variables assigned a priori to four categories; no within-category selection).

Functional category	Variables
Carbon storage and cycling	TOC, DOC, MBC, EOC, LFOC, HFOC, POC, MAOC, AOC, RC
Nutrient supply	TN, TP
Water regulation	SWC
Structural stability and erosion resistance	MWD_dry, Macro_agg_pct

Excluded: pH, EC (state or driver, not function), total carbon (includes a substantial carbonate fraction in these alkaline soils), particle size fractions, stoichiometric ratios (C:N, C:P, N:P), MWD_wet (not measured across all samples).

(b) Microhabitat values (mean ± SE, n = 3) for the primary index and two robustness formulations.

Formulation	AH	S_A	S_C	BC_A	BC_C	ANOVA
EMF (category-weighted; primary)	−0.72 ± 0.12 [b]	−0.35 ± 0.24 [a]	−0.25 ± 0.25 [a]	+0.46 ± 0.46 [a]	+0.86 ± 0.17 [a]	F = 5.56, P = 0.013
EMF (all-variable equal weight)	−0.80 ± 0.08	−0.44 ± 0.27	−0.04 ± 0.53	+0.58 ± 0.53	+0.69 ± 0.31	F = 2.79, P = 0.086
PC1 (74.4% of variance)	−3.21 ± 0.14	−1.79 ± 1.10	+0.19 ± 2.53	+2.20 ± 1.94	+2.60 ± 1.43	F = 2.34, P = 0.126

Construction: z-score each variable (N = 15), average within category, then equal weight across the four categories (primary). Robustness: The three formulations produced an identical microhabitat rank order (Spearman ρ = 1.00). Tukey HSD letters are shown for the primary index (significant ANOVA); lowercase letters denote groups that differ at α = 0.05. Microhabitat abbreviations are shown in Fig. S1.

Table S5. Community function coupling coefficients at the microhabitat level (n = 5).

Function	Community metric	Pearson r	P	Spearman ρ	P
Soil multifunctionality (EMF)	Bacterial Chao1	−0.61	0.270	−0.50	0.391
	Bacterial Shannon	**−0.89**	**0.042**	−0.70	0.188
	Cyanobacteria	0.74	0.155	0.40	0.505
	Actinobacteriota	**−0.89**	**0.042**	−0.70	0.188
	Proteobacteria	−0.43	0.470	−0.30	0.624
	Eukaryotic photoautotrophs (Phragmoplastophyta)	0.64	0.242	0.10	0.873
Cumulative CO_2 efflux	Bacterial Chao1	−0.76	0.140	−0.40	0.505
	Bacterial Shannon	**−0.95**	**0.013**	−0.70	0.188
	Cyanobacteria	**0.99**	**0.002**	**0.90**	**0.037**
	Actinobacteriota	**−0.98**	**0.003**	**−1.00**	**<0.001**
	Proteobacteria	0.05	0.933	0.30	0.624
	Eukaryotic photoautotrophs (Phragmoplastophyta)	**0.96**	**0.011**	0.50	0.391
Cumulative N_2O flux	Bacterial Chao1	−0.62	0.263	−0.60	0.285
	Bacterial Shannon	−0.64	0.242	−0.70	0.188
	Cyanobacteria	0.80	0.101	**0.90**	**0.037**
	Actinobacteriota	−0.69	0.197	−0.80	0.104
	Proteobacteria	0.68	0.211	0.80	0.104
	Eukaryotic photoautotrophs (Phragmoplastophyta)	0.86	0.064	**0.90**	**0.037**

Bold marks P < 0.05. Coefficients relate microhabitat means (n = 5) and are reported for both Pearson and Spearman because several taxa (notably cyanobacteria) have a bimodal microhabitat distribution. Community metrics are relative abundances, except for the diversity indices. Microhabitat abbreviations as in Fig. S1.

Table S6. Sensitivity of the soil multifunctionality index to the carbon fractions. Values are microhabitat means ± SD (n = 3).

Formulation	AH	S_A	S_C	BC_A	BC_C	F	P
EMF_cat (primary; four categories, 15 variables)	-0.72 ± 0.20	-0.34 ± 0.41	-0.25 ± 0.43	+0.46 ± 0.80	+0.86 ± 0.29	5.56	0.0128
EMF without carbon fractions (three categories, 5 variables)	-0.67 ± 0.25	-0.29 ± 0.38	-0.38 ± 0.13	+0.39 ± 0.74	+0.94 ± 0.28	7.58	0.0045
Unweighted mean of the five non-carbon variables	-0.60 ± 0.27	-0.27 ± 0.33	-0.37 ± 0.11	+0.41 ± 0.77	+0.83 ± 0.19	6.48	0.0077
PC1 of all 15 standardized variables	-3.21 ± 0.24	-1.79 ± 1.91	+0.19 ± 4.39	+2.20 ± 3.36	+2.60 ± 2.48	2.34	0.1261
PC1 of the five non-carbon variables	-1.61 ± 0.48	-0.47 ± 1.26	-0.89 ± 0.33	+0.54 ± 1.08	+2.43 ± 1.10	8.53	0.0029

Removing all ten carbon fractions preserved the ant hill minimum and the control biocrust maximum and left the microhabitat effect significant. Microhabitat means of the primary and carbon-free formulations were closely correlated (Pearson r = 0.99; r = 0.97 across the 15 samples), and pairwise Spearman ρ among formulations was 0.90–1.00. The only reordering occurred between S_A and S_C, which did not differ from one another in any analysis. Microhabitat abbreviations as in Fig. S1.

Table S7. Leave-one-out resampling of the cross-domain co-occurrence network. The inference was repeated 20 times, each time omitting one of the 20 samples.

Sample omitted	Pearson r vs full-sample solution	Edges	Reference edges recovered (%)	Intra-bacterial edges	Intra-eukaryotic edges	Cross-domain edges	Hub genera retained (of 9)
AH1	0.990	906	75.9	633	38	235	9
BC_A1	0.995	928	79.8	654	28	246	8
S_A1	0.985	878	73.5	621	34	223	8
BC_C1	0.972	743	65.7	540	23	180	7
S_C1	0.989	886	78.5	625	30	231	9
AH2	0.996	881	76.3	621	30	230	9
BC_A2	0.997	897	78.3	644	29	224	7
S_A2	0.986	819	72.7	588	29	202	8
BC_C2	0.987	801	69.1	556	33	212	9
S_C2	0.980	838	72.4	581	31	226	7
AH3	0.994	905	76.1	630	30	245	7
BC_A3	0.990	931	76.6	687	28	216	7
S_A3	0.992	873	75.2	615	33	225	8
BC_C3	0.988	802	70.0	571	27	204	9
S_C3	0.975	819	71.0	599	23	197	7
AH4	0.995	907	76.9	641	31	235	8
BC_A4	0.996	876	76.2	614	30	232	8
S_A4	0.986	848	72.4	611	31	206	8
BC_C4	0.989	780	71.8	550	24	206	8
S_C4	0.973	935	76.3	687	21	227	7

Across the 20 replicates the inferred correlation matrix remained highly concordant with the full-sample solution (mean Pearson r = 0.99, range 0.97–1.00), and intra-bacterial associations outnumbered intra-eukaryotic associations in every replicate. The same result was obtained when edges were defined by the correlation threshold alone. Sample codes denote microhabitat and plot; microhabitat abbreviations as in Fig. S1.