MICROBIOLOGY

A global comparison of surface and subsurface microbiomes reveals large-scale biodiversity gradients, and a marine-terrestrial divide

S. Emil Ruff[1]*†, Isabella Hrabe de Angelis[2]†‡, Megan Mullis[3], Jérôme P. Payet[4], Cara Magnabosco[5], Karen G. Lloyd[6], Cody S. Sheik[7], Andrew D. Steen[6], Anna Shipunova[1], Aleksey Morozov[1], Brandi Kiel Reese[3,8], James A. Bradley[9,10], Clarisse Lemonnier[11], Matthew O. Schrenk[12], Samantha B. Joye[13], Julie A. Huber[14], Alexander J. Probst[15], Hilary G. Morrison[1], Mitchell L. Sogin[1], Joshua Ladau[16,17], Frederick Colwell[4]

Subsurface environments are among Earth's largest habitats for microbial life. Yet, until recently, we lacked adequate data to accurately differentiate between globally distributed marine and terrestrial surface and subsurface microbiomes. Here, we analyzed 478 archaeal and 964 bacterial metabarcoding datasets and 147 metagenomes from diverse and widely distributed environments. Microbial diversity is similar in marine and terrestrial microbiomes at local to global scales. However, community composition greatly differs between sea and land, corroborating a phylogenetic divide that mirrors patterns in plant and animal diversity. In contrast, community composition overlaps between surface to subsurface environments supporting a diversity continuum rather than a discrete subsurface biosphere. Differences in microbial life thus seem greater between land and sea than between surface and subsurface. Diversity of terrestrial microbiomes decreases with depth, while marine subsurface diversity and phylogenetic distance to cultured isolates rivals or exceeds that of surface environments. We identify distinct microbial community compositions but similar microbial diversity for Earth's subsurface and surface environments.

INTRODUCTION

Microbial life is pervasive in Earth's hugely varied habitats and environments. Microorganisms have adapted to grow in acidic and alkaline springs, salterns, deserts, environments with temperature extremes greater than 122°C or lower than −20°C, and from ambient pressures up to pressures greater than those of abyssal oceanic trenches (*1*–*4*). Although surface ecosystems are thought to harbor the majority of Earth's total biomass (*5*), most of Earth's bacterial and archaeal biomass carbon is stored in subsurface ecosystems from less than 1 m to many kilometers beneath Earth's surface. The local irregularity of boundaries between surface and subsurface ecosystems (*6*) due to the variability of physicochemical and biological parameters with depth supports only a loose definition of subsurface environments, which include soils, rocks, or sediments deeper than 1 to 8 m below Earth's land surface or deeper than 0.1 to 1 m below the seafloor (mbsf) (*7*–*12*). For the "subsurface" category in this study, we use samples from aquifers, rock fracture fluids, sediments, and rock cores that originated from 0.2 to 491 m below the seafloor or 15 to 4375 m below ground.

At continental margins, especially at the mouths of large rivers, sediment thickness can exceed 10 km (*13*). Similarly, the habitable continental crust can exceed 15 km in thickness in the Canadian, Fennoscandian, and Siberian shields (*7*). Archaea, bacteria, and eukaryotic microorganisms inhabit this subsurface biosphere (*12*). Marine subsurface ecosystems particularly have relatively high proportions and abundances of Archaea compared to most other ecosystems (*8*, *14*–*17*). Subsurface ecosystems may host more than half of all microbial cells (~5 to 12 × 10^{29}) on Earth (*7*, *18*–*21*) despite sometimes very low cell densities, for example, in subsurface sediments of oligotrophic South Pacific Gyre as low as approximately 100 cells per cubic centimeter of sediment (*22*).

Continuous deposition of organic matter from the overlying ocean and concomitant burial explains the introduction and presence of most microbes in deep sediment layers beneath the ocean floor. Life in marine deep subsurface sediments often bears a resemblance to the shallow subsurface life at a given location (*23*–*26*). Terrestrial subsurface ecosystems exhibit a similar connection to the surface world, with entrainment of shallow subsurface communities into the deeper realms via recharge and movement of fluids (*27*). This similarity exists despite the considerable environmental differences between surface and subsurface environments (i.e., differences in pressure, light, oxygen, energy, and nutrient availabilities, as well as available pore space within which cells might exist) that generally select for distinct microbial communities. Subsurface microbes can disperse via marine and terrestrial aquifers (*28*, *29*), fracture fluids (*30*), and porewaters (*31*). Hydrodynamic and hydrogeological processes such as eruptions of mud volcanoes (*32*, *33*) or fluid seepage (*34*, *35*) can reintroduce microbes from subsurface to surface environments.

[1]Marine Biological Laboratory, Woods Hole, MA, USA. [2]Multiphase Chemistry Department, Max Planck Institute for Chemistry, Mainz, Germany. [3]University of South Alabama, Mobile, AL, USA. [4]College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR, USA. [5]ETH Zürich, Zürich, Switzerland. [6]University of Tennessee, Knoxville, TN, USA. [7]Large Lakes Observatory and Department of Biology, University of Minnesota Duluth, Duluth, MN, USA. [8]Dauphin Island Sea Laboratory, Dauphin Island, AL, USA. [9]Aix Marseille University, University of Toulon, CNRS, IRD, MIO, Marseille, France. [10]Queen Mary University of London, London, UK. [11]UMR CARRTEL, INRAE, Université Savoie Mont-Blanc, Thonon-les-Bains, France. [12]Department of Earth and Environmental Sciences, Michigan State University, East Lansing, MI. USA. [13]Department of Marine Sciences, University of Georgia, Athens, GA, USA. [14]Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA, USA. [15]Environmental Metagenomics, Research Center One Health Ruhr of the University Alliance Ruhr, Faculty of Chemistry and Centre of Water and Environmental Research (ZWU), University of Duisburg-Essen, Essen, Germany. [16]Department of Computational Precision Health, University of California, San Francisco, CA, USA. [17]Arva Intelligence, Houston, TX, USA.
*Corresponding author. Email: eruff@mbl.edu
†These authors contributed equally to this work.
‡Present address: OceanX, 37 W 39th St., New York City, NY 10018, USA.

Intrinsic or acquired capacities to metabolize subsurface energy, carbon, and nutrient sources and/or improved survival and dormancy strategies allow microbes to survive and live in subsurface environments. The predominant occurrence of certain microbial lineages in subsurface ecosystems suggests that some organisms are better equipped for the subsurface than others. Coping strategies for survival may play a role, yet evidence also suggests that with increasing depth, e.g., in marine sediments, microbes change their expressed extracellular peptidases to those that specialize in highly degraded detrital proteins (*36*). Microbes can also increase cellular lifespan by slowing genome transcription (*37*) or actively expressing mRNA for DNA repair enzymes (*38*). Among many other processes acetogenesis (*39*, *40*), methanogenesis (*41*); hydrogen, methane, and sulfur oxidation (*42*, *43*); fermentation of microbial biomass and necromass (*44*–*46*); symbiosis (*47*, *48*); serpentinization (*49*); and even radiolysis (*50*, *51*) might contribute to the subsistence of deep life. While a portion of cells persist in a dormant state (*52*), many organisms actively metabolize (*22*, *53*, *54*) but often with generation times of decades to centuries (*24*, *55*–*58*).

While subsurface ecosystems harbor substantial biomass, diversity, and many lineages that are phylogenetically distant to cultured isolates, the degree to which the biodiversity of subsurface ecosystems directly compares with that of surface ecosystems remains unclear. Moreover, although global meta-analyses exist for either the marine (*17*) or the terrestrial subsurface (*59*), a standardized global dataset encompassing both marine and terrestrial subsurface environments was not available to date. Our study expands the insights into terrestrial subsurface diversity and composition beyond that of an earlier study that used data from two sequencing platforms (454 pyrosequencing and early Illumina) and multiple 16*S* ribosomal RNA (rRNA) primers, which thus allowed only a database-dependent operational taxonomic unit (OTU) approach (*59*). We also extend our observations beyond the bacterial domain (*59*) including a synthesis of archaeal communities, and we increase the total amount of samples and data, allowing a broader comparison of biomes and diverse environments.

Previous studies have compared marine and terrestrial surface environments suggesting distinct microbiomes between land and sea (*60*–*64*). Factors such as salinity, pH, and the availability of nutrients were identified as being among the major drivers of community variance (*60*, *62*, *65*–*67*). These works, however, focused on the principles structuring each habitat (*63*) on sea and land-derived bioaerosols (*64*) or on the ocean-land connectivity at specific sites (*60*–*62*). A standardized comparison of microbial metabolites across diverse marine and terrestrial environments is available (*68*), and numerous global surveys have studied either marine (*69*–*71*) or terrestrial microbiomes (*72*, *73*). Standardized comparisons of microbial community structure between sea and land, however, have not been published to our knowledge. Datasets that can compare marine and terrestrial microbiomes on a global scale are available (*67*, *74*, *75*) yet were analyzed with a different focus. To understand the differences and commonalities in the microbiomes of global surface and subsurface environments, we here provide a comparison between surface and subsurface as well as between marine and terrestrial environments. We investigate environments that span a broad range of depths from surface environments (under the influence of relatively fresh photosynthesis-derived organic matter) to very deep and isolated environments (cut-off from photosynthetic primary production for at least centuries). For this global comparison, we analyzed 1442 globally distributed 16*S* rRNA gene amplicon datasets and taxonomic marker genes of 147 metagenomes. We investigate (i) whether microbial communities of marine and terrestrial biomes and of surface and subsurface environments fundamentally differ, (ii) whether subsurface environments are generally less diverse than surface ecosystems, (iii) whether subsurface environments harbor distinct clades, and (iv) whether marine and terrestrial subsurface communities share a core microbiome.

RESULTS AND DISCUSSION

The extended Census of Deep Life atlas allows comparisons of surface and subsurface biodiversity

The Census of Deep Life (CoDL) under the auspices of the Deep Carbon Observatory organized a decadal effort (2010–2020) to characterize microbial diversity and function in subsurface ecosystems worldwide, carried out by more than two dozen groups and hundreds of researchers. After the completion of the CoDL, we compiled, re-analyzed, and compared 964 bacterial and 478 archaeal 16*S* rRNA gene amplicon datasets from 35 individual globally distributed CoDL projects together with 15 additional non-CoDL projects that originated mainly from surface ecosystems (Fig. 1, A and B, and datasets S1 and S2). Archaeal and bacterial 16*S* rRNA gene amplicon sequence variants (ASVs; quasi–strain-level microbial lineages) were amplified using domain-specific primers. We obtained 31,099 unique archaeal and 377,374 unique bacterial ASVs from a total of 4.9×10^7 archaeal and 9.6×10^7 bacterial reads. On average, we found 155 archaeal and 1183 bacterial ASVs per sample (datasets S3 and S4). Analyses of 147 metagenomes from 49 globally distributed CoDL subsurface projects complement the amplicon datasets. The metagenomes were used for taxonomic analyses based on two different taxonomic markers, 16*S* rRNA gene sequences retrieved by phyloFlash (pF16S; before read assembly) and ribosomal protein S3 (rpS3) genes (after read assembly). We obtained a total of 1552 unique archaeal, 23,206 unique bacterial, and 1387 unique eukaryotic pF16S genes, as well as 154 different archaeal and 1601 different bacterial rpS3 genes (datasets S5 and S6).

We analyzed the communities first by grouping the amplicon datasets based on the biome from which the samples originated: marine biome ($n_{\text{Archaea}} = 304$ datasets, $n_{\text{Bacteria}} = 503$ datasets) and terrestrial biome ($n_{\text{Archaea}} = 174$, $n_{\text{Bacteria}} = 461$). Within the marine and terrestrial biome, we grouped samples based on the depth realm they originated from: surface, interface, and subsurface.

Surface datasets ($n_{\text{Archaea}} = 183$, $n_{\text{Bacteria}} = 599$) include water samples from oceans and lakes (at various depths in the water column) and shallow (<0.1 mbsf) sediment samples from oceans, estuaries, and lakes. Surface ecosystems represent environments under the influence of relatively fresh photosynthesis-derived organic matter. Subsurface datasets ($n_{\text{Archaea}} = 85$, $n_{\text{Bacteria}} = 122$) originated from ecosystems that have likely been cut-off from photosynthetic primary production for decades to centuries. These ecosystems were accessed via boreholes or mines and include deep sediments, aquifers, and fracture fluids. Marine subsurface sediments were further subdivided based on depositional setting: shelf, slope, and abyssal domains. The location of each domain is defined by water depth (*76*–*78*). Shelf environments roughly correspond to water depths <200 m, except the Antarctic region where shelf area corresponds to water depths <500 m. The abyssal plain corresponds to areas of water depths >3500 m. Sediments under other water depths are referred to as slopes. Interface datasets ($n_{\text{Archaea}} = 210$, $n_{\text{Bacteria}} = 243$)

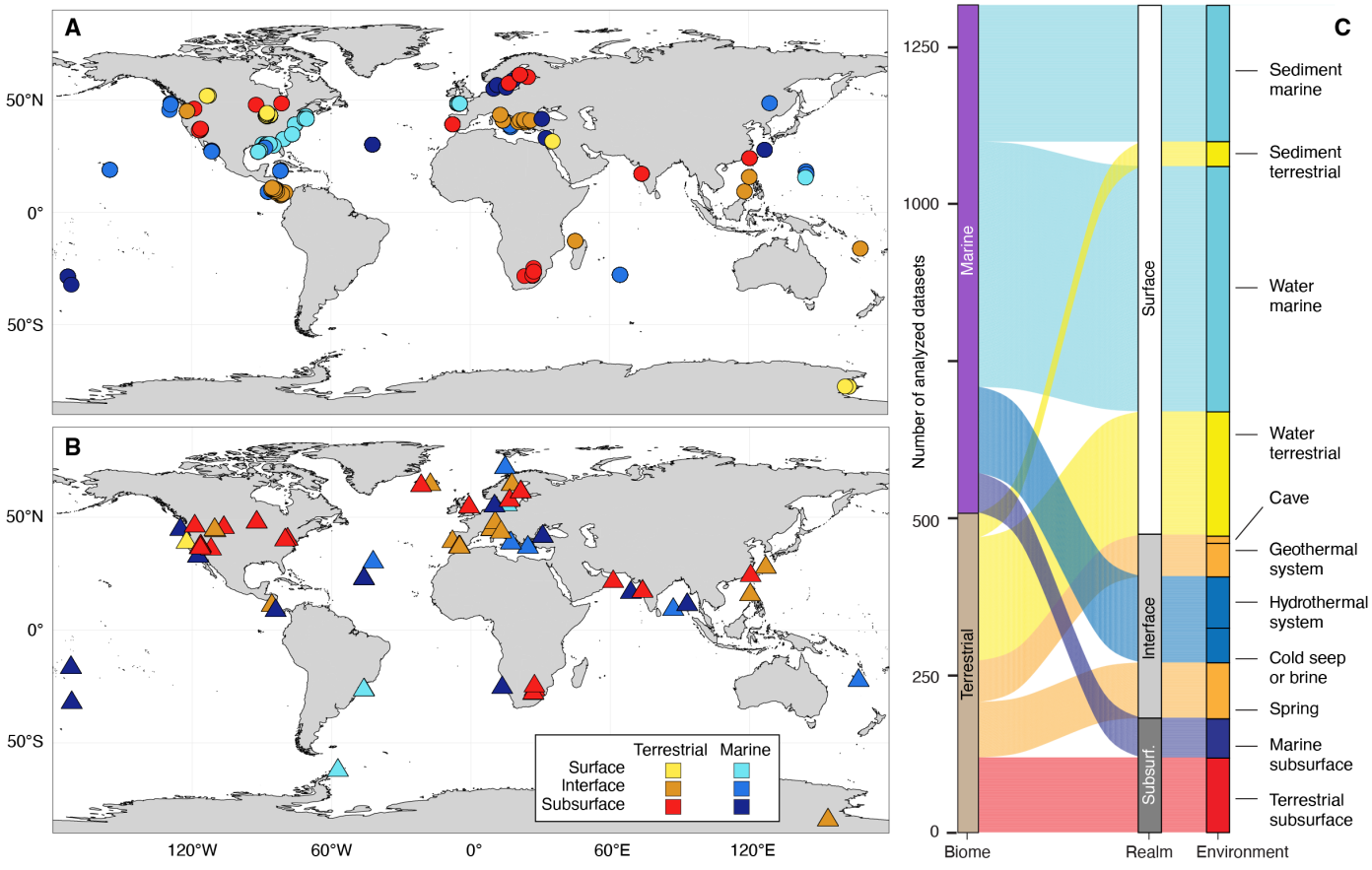

Fig. 1. Geographic location and origin of samples. Maps of samples used for metabarcoding of 16S rRNA gene amplicon taxonomic marker genes (**A**) and shotgun metagenomic analyses of unassembled and de novo assembled taxonomic marker genes (**B**). Each symbol represents one project, which comprises multiple individual samples. Both terrestrial and marine samples may contain rock, sediment, or water samples. Further maps showing sample material, pH, and temperature are included in fig. S1. (**C**) Overview of sample origins derived from marine and terrestrial biomes, depth realms, and environments.

originated from environments that are subjected to influences from surface and subsurface environments and processes. Interface environments can be surface sites that are driven by energy from the subsurface, or vice versa, and included water and sediment samples from caves, hot springs, hydrothermal vents, and cold seeps. We also grouped and analyzed the samples based on the 11 major environments they originated from, as listed in Fig. 1C.

We included samples from four different types of environmental materials: water or brine (724 samples), biofilm or mat (35 samples), sediment or soil (458 samples), and rock (102 samples). Samples from different materials were dispersed relatively homogeneous across the different biomes (figs. S2 and S3) and depth realms (figs. S4 and S5) to minimize diversity effects caused by the sample material. Sediments had the highest alpha diversity, while rocks had the lowest, in both the marine and terrestrial biome. To minimize batch effects and biases, we sequenced and analyzed all surface, interface, and subsurface samples using the same archaeal or bacterial primers, the same chemistry, the same Illumina instrument, and the same bioinformatic pipeline. We minimized the potential impact of contamination by including blanks and controls and by removing notorious contaminants listed in the Supplementary Materials (*79*, *80*). We minimized composition bias by normalizing template concentration and by standardizing workflows (*81*–*83*). Our dataset can deeply sample microbial richness and evenness, compare microbiomes at high taxonomic resolution, and study community similarities and shifts across large scales. Future work could expand our analyses to explore changes in the subsurface microbiome over time, perform long read 16S rRNA gene sequencing to achieve higher phylogenetic resolution, or use metagenomics to assess metabolic capabilities. These analyses have been done at individual sites (*37*, *84*, *85*) and promise—at a global scale—to greatly expand our understanding of the roles and nature of subsurface life.

Substantial differences exist between marine and terrestrial microbiomes

We found substantial differences in archaeal and bacterial community structure between marine and terrestrial biomes. For this analysis, we grouped all datasets solely based on whether they originated from a marine or terrestrial environment, regardless of their sample material or association with a surface, interface, or subsurface environment. Alpha diversity (community diversity per locality/sample) and beta diversity (community diversity between localities/samples) of marine and terrestrial microbiomes showed the same trends across ASV, rpS3, and pF16S-based analyses. Richness, estimated richness (Chao1), and evenness (Shannon entropy and inverse Simpson diversity) were similar or higher for archaea and bacteria in marine relative to terrestrial microbiomes (Fig. 2, A to F, and

Fig. 2. Microbial diversity in marine and terrestrial biome. Archaeal (**A** to **C**) and bacterial (**D** to **F**) alpha diversity (per sample community richness and evenness) in marine and terrestrial biomes using 16*S* rRNA gene ASVs [(A) and (D)], as well as metagenome-derived ribosomal protein S3 genes [rpS3; (B) and (E)] and 16*S* rRNA gene sequences detected by phyloFlash (pF16S). To allow comparison, the datasets were subsampled to the same number of reads. Pairwise comparisons were performed using a Wilcoxon rank sum test. Significance: $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$. The number of analyzed datasets is shown below the boxplots. Community dissimilarity between marine and terrestrial communities was shown by nonmetric multidimensional scaling ordinations of ASV-based dissimilarity matrices using 478 archaeal (**G**) and 964 bacterial datasets (**H**). Each dot represents the community structure of a dataset and is connected to the group centroid (weighted average mean of within-group distances); ellipses depict 1 SD of the centroid. Statistical testing using ANOSIM showed that the groups are overlapping but significantly different ($R \sim 0.5$, $P < 0.001$). (**I**) Differential sequence abundance analyses of marine versus terrestrial bacterial phyla. The phyla are ordered from top to bottom based on increasing phylum level MaAsLin2 coefficient, i.e., likeliness of their occurrence in terrestrial-derived samples ("terrestrialness"). Boxplots summarize order level MaAsLin2 coefficients, i.e., terrestrialness, within the listed phyla. Note that due to ease of visualization, boxplots are also shown for very small number (*n*) of orders. Significance levels are: $^{*}P < 0.01$, $^{**}P < 0.001$, $^{***}P < 0.0001$, and $^{****}P < 0.00001$. Additional phyla, particularly those that lack cultured representatives, are shown in fig. S6.

fig. S6). The high diversity in the marine biome was caused in both domains by sediments and rocks, which are more diverse in the sea than on land. Marine waters were similarly diverse as terrestrial waters for bacteria and slightly higher for archaea, whereas biofilms and microbial mats tended to be higher in diversity on land for both domains (figs. S2 and S3). In the subsampled datasets, which was done to account for the unequal sampling effort, i.e., different sequencing depth, we found, on average, 75 and 79 archaeal and 512 and 349 bacterial ASVs in marine and terrestrial samples, respectively (dataset S7). Microbial community structure and composition were overlapping but significantly different between the marine and terrestrial biomes (Fig. 2, G and H; R_{Arc}: 0.58, P_{Arc}: 0.001; R_{Bac}: 0.49, P_{Bac}: 0.001). On average, only 2 to 6% of ASVs, 2% of rpS3 genes, and 8 to 9% of pF16S genes were shared between any marine and terrestrial dataset (Fig. 2, G and H, and figs. S7 and S8). ASV-based gamma diversity (total community diversity per biome/environment) tended to be higher for archaea in the terrestrial biome and for the bacteria in the marine biome (figs. S6 and S7).

A multivariate association analysis (*86*) of our dataset suggests that most bacterial phyla are significantly more sequence abundant and prevalent in the marine biome (fig. S9A). Phyla that are more common in marine environments include well-known Proteobacteria, Planctomycetota, *and* Verrucomicrobiota, but also less prominent phyla such as Marinimicrobia, Nitrospinota, Fermentibacterota, Caldatribacteriota, Latescibacterota, Aerophobota, and Gemmatimonadota (Fig. 2I and fig. S9A). In terrestrial environments, we found more commonly Firmicutes, Chloroflexi, Acidobacteriota, Actinobacteriota, Nitrospirota, and Patescibacteria, as well as Methylomirabilota, Armatimonadota, and Elusimicrobiota among others. Both the marine- and terrestrial-associated phyla comprise members of the two major monophyletic branches of the bacterial tree Gracilicutes (mostly diderm lineages) and Terrabacteria (monoderm and atypical diderm lineages), respectively (*87*). We also find a versatile third group of bacterial phyla that occupy both biomes at comparable relative sequence abundances (RSAs) and prevalences, also regardless of being mono- or diderms. These include Desulfobacteria, Bacteroidota, Cyanobacteria, Spirochaetota, Gemmatimonadota, Campilobacterota, and Thermotogota. Hence, despite the marked differences between sea and land at ASV, species, and genus level, there are cosmopolitan lineages but at higher phylogenetic levels (e.g., phylum and class). Most archaeal phyla were more sequence abundant and prevalent in the terrestrial biome (fig. S9B). Only Thermoplasmatota were significantly more common in marine environments, and Crenarchaeota and Asgardarchaeota were found averaged across each biome at similar RSAs and prevalence. All other archaeal phyla including Euryarchaeota, Halobacterota, Hadarchaeota, Nanoarchaeota, Aenigmarchaeota, and Altiarchaeota were more sequence abundant and prevalent in the terrestrial biome. Relative sequence abundances of archaeal and bacterial phyla (fig. S9, C and D) support the trend revealed by the association analyses; however, association analyses or differential abundance analyses can change with every sample and lineage added or removed and thus must be interpreted with caution. Robust affiliations to a certain biome may only be universally true for the lineages with highest significances, those at the very top and bottom of the graph. Some phyla contribute disproportionally to ASV-level diversity including Nitrososphaeria, Bathyarchaeia, Chloroflexi, and Planctomycetota (fig. S9, E and F), being responsible for a larger part of the ASV-level diversity than their RSA suggests.

Subsurface microbiomes can be as diverse as surface microbiomes

Within each biome, differences in alpha diversity between surface, interface, and subsurface environments, i.e., depth realms, were overall less pronounced than between the biomes (figs. S10 and S11). Across depth realms, communities of either biome—marine or terrestrial—shared many ASVs (figs. S12 and S13, and movies S1 and S2). Archaeal and bacterial communities in the three depth realms were distinct but with considerable overlaps. The larger differences in community structure between marine and terrestrial biomes than between depth realms indicate that the divide between microbial life on land and in the sea is more pronounced than the divide between surface and subsurface communities. We found that species richness and evenness in many subsurface environments rival those in surface environments (fig. S14). This finding was consistent across the 11 investigated environment types, where levels of microbial diversity are often comparable across the surface, interface, and subsurface (Fig. 3, A to D), i.e., being within the same order of magnitude. Archaeal alpha diversity was overall highest in samples from cold seeps and brines, caves, springs, and the deep marine subsurface (Fig. 3A). Bacterial alpha diversity was, on average, highest in the cave samples and marine sediments, followed by terrestrial water and sediment, seeps and vents. Marine subsurface bacterial diversity rivaled the diversity found in springs and marine surface water (Fig. 3B). Total archaeal diversity (gamma diversity) was significantly higher in marine interface and subsurface environments than in the marine surface, even after correcting for the different numbers of samples per group using a subsampling approach (Fig. 3C and fig. S10, A to D). Similarly, bacterial gamma diversity was highest in marine interface environments (Fig. 3D). In the terrestrial biome, archaeal gamma diversity was comparable across all depth realms (Fig. 3C), whereas bacterial diversity was highest in the surface (Fig. 3D). Overall, archaeal and bacterial diversity was only sampled exhaustively in marine waters based on species accumulation curves (fig. S15), corroborating previous findings that vast numbers of microbial species remain to be found across global biomes (*11*, *74*, *88*–*90*).

Thermoplasmata, Nitrososphaeria, and Bathyarchaeia together comprised more than half of the archaeal RSAs in any investigated environment, except the terrestrial subsurface (Fig. 4A and fig. S16). The contribution of these three classes to global archaeal ASV richness was similarly outsized (Fig. 4B), while other clades that can occur at relatively high RSAs contributed very little to overall archaeal richness, including Hadarchaea, ANME-1, Methanobacteria, Thermoprotei, and Methanococci (Fig. 4, A and B, and fig. S16). This suggests that a large part—if not most—of the global archaeal abundance and richness can be attributed to three archaeal classes Thermoplasmata, Nitrososphaeria, and Bathyarchaeia. Bacterial communities were largely dominated by Proteobacteria and Bacteroidota, except those of caves and the subseafloor where Chloroflexi played an important role (Fig. 4C). The contribution of the lineages to overall bacterial richness was often not proportional to their RSA. Proteobacteria, Bacteroidota, Firmicutes, Cyanobacteria, and Desulfobacterota contributed less, while Chloroflexi, Planctomycetota, and Verrucomicrobia contributed more richness than expected based on their sequence abundance (Fig. 4D). We found increasing archaeal richness and evenness with depth, supporting the notion that archaea are well suited for life in subsurface and subsurface-influenced, i.e., interface, environments. The importance of archaea in marine subsurface

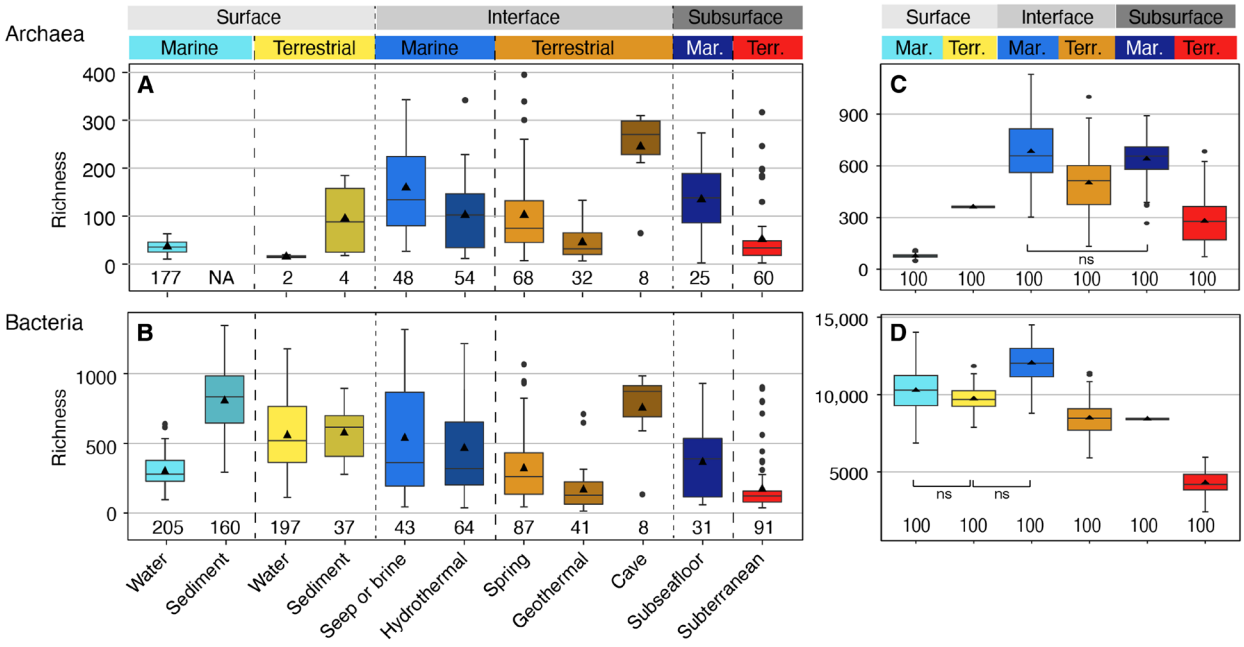

Fig. 3. Observed archaeal and bacterial richness across environments based on 16*S* rRNA gene ASVs. Archaeal (**A**) and bacterial (**B**) richness found in the 11 studied environments (*n*: number of included samples). Environments are grouped on the basis of biome (marine, terrestrial) and depth realm (surface, interface, subsurface). Note that Archaeal terrestrial water and sediment samples contain too few datapoints for robust visualization using boxplots, yet the plots were retained for completeness. NA, no value available. Total archaeal (**C**) and bacterial richness (**D**) using a subsampling approach to account for different group sizes (100 iterations using 1142 archaeal or 2271 bacterial reads, respectively). Normalized gamma diversity corroborates that archaeal diversity is highest in marine interface and subsurface ecosystems even when different group sizes are considered. All pairwise comparisons were significantly different ($P < 0.001$) except for those indicated with ns (not significant). Shannon entropy values and a subsampling approach using 50,000 reads show the same trends (fig. S11). The diversity found in individual subsurface environments and their comparison to surface and interface environments is shown in the fig. S14.

ecosystems has previously been demonstrated by studies assessing community structure using lipid analyses (*8*), cell abundances (*16*), and meta-omics (*14*, *91*). A high diversity of archaea in the terrestrial subsurface, rivaling that of terrestrial surface ecosystems, is less well documented. This work robustly shows that average archaeal diversity in the terrestrial subsurface is often equal to and can even exceed the diversity found at terrestrial surface environments, a finding that contextualizes the high archaeal diversity described in previous studies of the terrestrial subsurface (*59*, *92*, *93*). Similarly unexpected is the high diversity of bacteria in the marine subsurface in comparison to surface ecosystems, supporting previous findings that focused on marine sediments (*17*). This work collectively improves our understanding of the contribution of archaea and bacteria to global microbial diversity and ecosystem function. To further investigate the global trend and corroborate the high diversity of archaeal and bacterial diversity in marine subsurface ecosystems, future studies should include datasets of additional surface environments such as soils and freshwater sediments, as those were underrepresented in our datasets.

Interface and subsurface microbiomes are phylogenetically more diverse in marine than in terrestrial biomes

Average microbial species-level diversity was significantly higher in marine interface and subsurface environments than in terrestrial interface and subsurface environments (Fig. 3, C and D, and fig. S10). This was the case for archaea and bacteria and for all tested alpha diversity indices concerning richness, evenness, and estimated richness, as well as gamma diversity. We corroborated this ASV-based trend through comparisons of the diversity of rpS3 and pF16S genes (fig. S10), for which marine interface and subsurface microbiomes were also either significantly more diverse or statistically indistinguishable. Beta diversity between the marine and terrestrial subsurface was also high based on ASVs (fig. S12, A and B), showing that analogous to surface ecosystems marine and terrestrial subsurface environments harbor many unique microbial species and comprise fundamentally different habitats. Generally, the average community dissimilarity between terrestrial samples was greater than between marine samples. The processes that lead to lower richness and greater dissimilarity in the terrestrial subsurface cannot be explained with our dataset but could be due to fundamental properties of the environment, including high habitat heterogeneity, high disturbance, or low dispersal. The diversity trends were also supported by rpS3 and pF16S gene-based analyses (fig. S12, C and D).

Certain microbial lineages are globally abundant in the subsurface

Archaeal and bacterial communities differed between the surface and subsurface environments regarding composition, RSA and prevalence, i.e., the percentage of samples in which a lineage occurred (Figs. 5 and 6 and figs. S16 and S17). Certain archaeal and bacterial lineages were significantly more prevalent in samples from subsurface than surface ecosystems (fig. S18). In the marine biome, most archaeal phyla including Euryarchaeota and Asgardarchaeota predominantly occurred in the subsurface (Fig. 5A), except for Thermoplasmatota, which are very abundant in the ocean. Terrestrial archaeal phyla did not have a significant preferential occurrence in surface or subsurface environments (Fig. 5B). Among the bacteria, the phyla Cyanobacteria, Bdellovibrionota, Verrucomicrobiota, Planctomycetota, Bacteroidota, and Proteobacteria are statistically

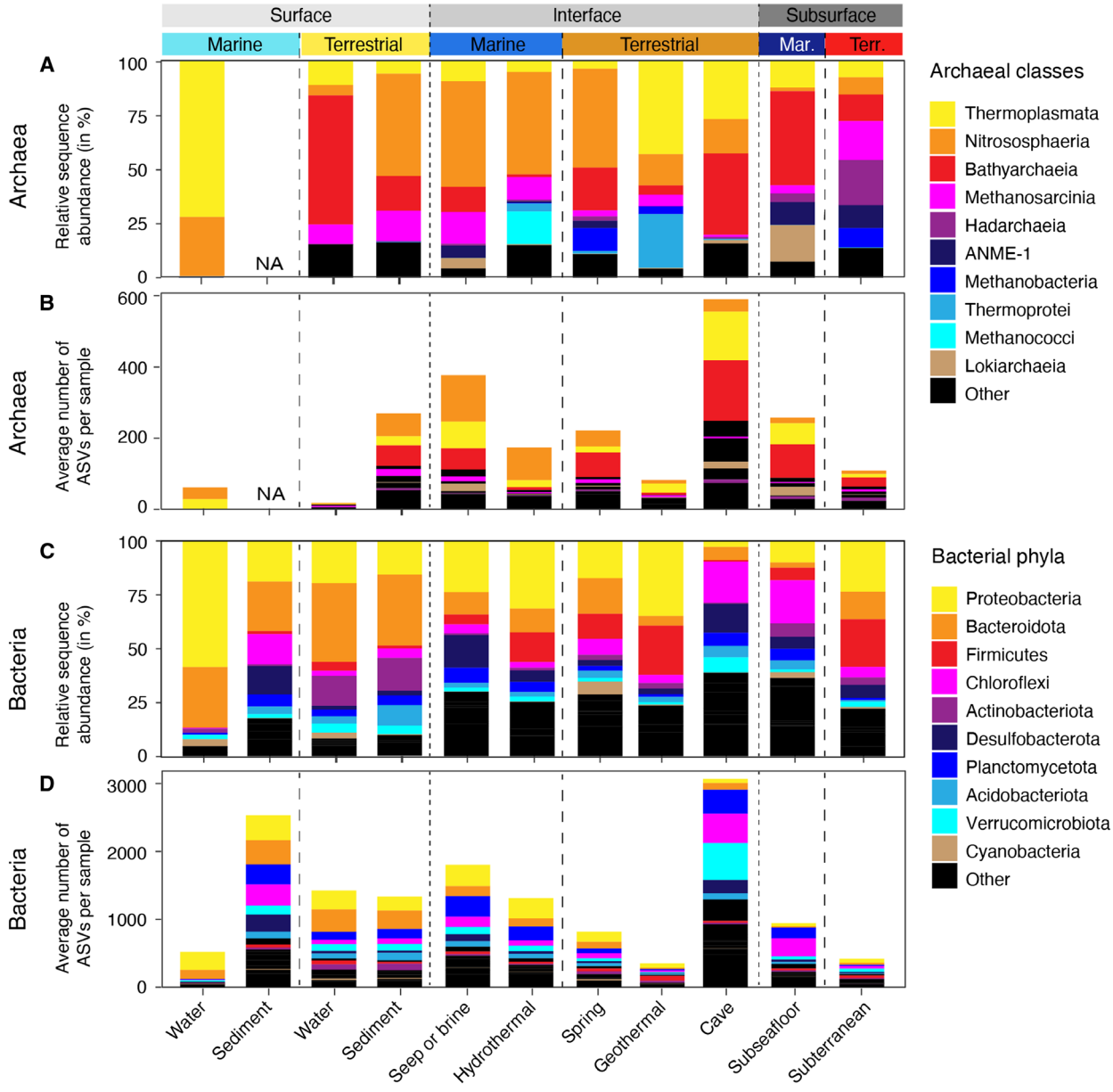

Fig. 4. Relative sequence abundance and richness of most important lineages across environments. Relative sequence abundance of top 10 most sequence abundant archaeal classes (**A**) and bacterial phyla (**C**) in the 11 studied environments. Contribution of most sequence abundant lineages to average number of archaeal (**B**) and bacterial (**D**) ASVs per sample (richness). Note that the most abundant clades do not necessarily show the highest richness, e.g., Planctomycetota are the seventh most sequence abundant clade [(C)] yet the third most diverse [(D)].

more common in surface ecosystems of the marine and terrestrial biome (Fig. 5, C and D, and figs. S19, S20, and S21). Firmicutes, Caldatribacteriota, Elusimicrobiota, and Patescibacteria are among those that are more likely found in the marine and terrestrial subsurface (Fig. 5, C and D, and figs. S17 and S18). In the marine subsurface, common phyla include Aerophobota, Chloroflexi, Desulfobacterota, and Methylomirabilota, while Nitrospirota was particularly common in the terrestrial subsurface. This finding supports previous reports of microbes that preferentially occur in deep subsurface environments including the sulfate-reducing firmicute *Candidatus* Desulforudis audaxviator (*94*–*96*), certain Spirochaeta (*97*), or organisms affiliating with (Cald)atribacteriota (*17*, *98*–*104*). In line with previous studies, we show that Proteobacteria are ubiquitous and often dominant in marine (*17*) and terrestrial subsurface (*59*) ecosystems analogous to their well-documented abundance in surface ecosystems (Fig. 5 and figs. S19 to S21) (*74*).

We have identified the most common subsurface class-level bacteria and order-level archaea based on their RSAs and prevalence and defined four categories. The first category comprises lineages with high RSAs and a high prevalence, e.g., Bathyarchaeia in the marine subsurface or Gammaproteobacteria in the terrestrial subsurface, agreeing with findings of earlier studies (*17*, *59*). Both lineages occur in all samples of the respective biome (prevalence = 1) at up to 2% RSA (Fig. 6, A and D). Bathyarchaeia still occur at more than half of all marine

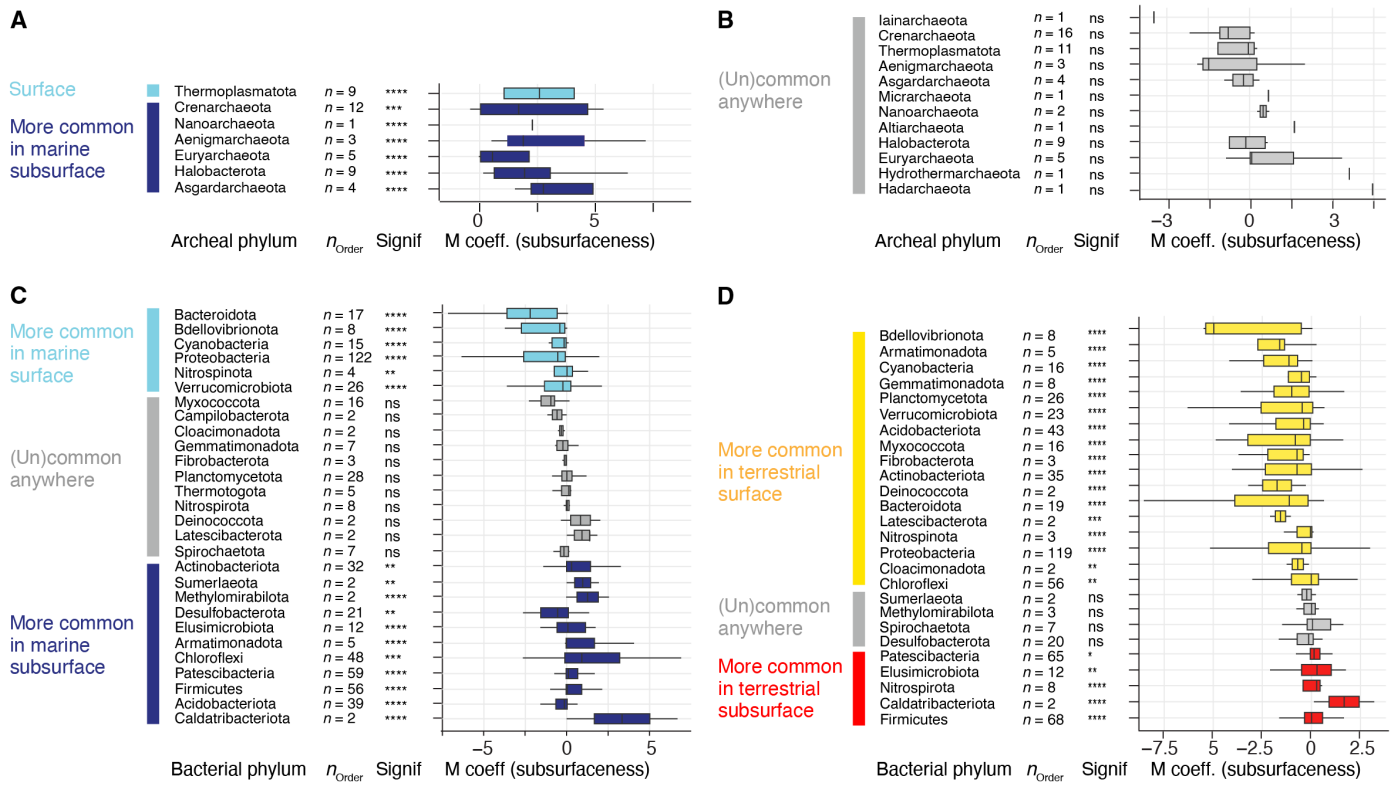

Fig. 5. Multivariate association analyses of microbial lineages. Analyses compare the occurrence of archaeal (**A** and **B**) bacterial (**C** and **D**) in marine [(A) and (C)] and terrestrial [(B) and (D)] surface versus subsurface realms. The phyla are ordered from top to bottom based on increasing likeliness of their occurrence in subsurface-derived samples (increasing MaAsLin2 coefficients; "subsurfaceness"). Boxplots summarize order level MaAsLin2 coefficients within the listed phyla. Note that due to ease of visualization, boxplots are even shown for very small number (*n*) of orders. The significance of the MaAsLin2 phylum level coefficient is shown in the column denoted "signif." Significance levels are: not significant (ns), *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, and ****$P < 0.0001$. An analysis based on all surface versus subsurface samples regardless of the biome of origin is shown in fig. S18, and additional bacterial phyla are shown in figs. S19 to S21.

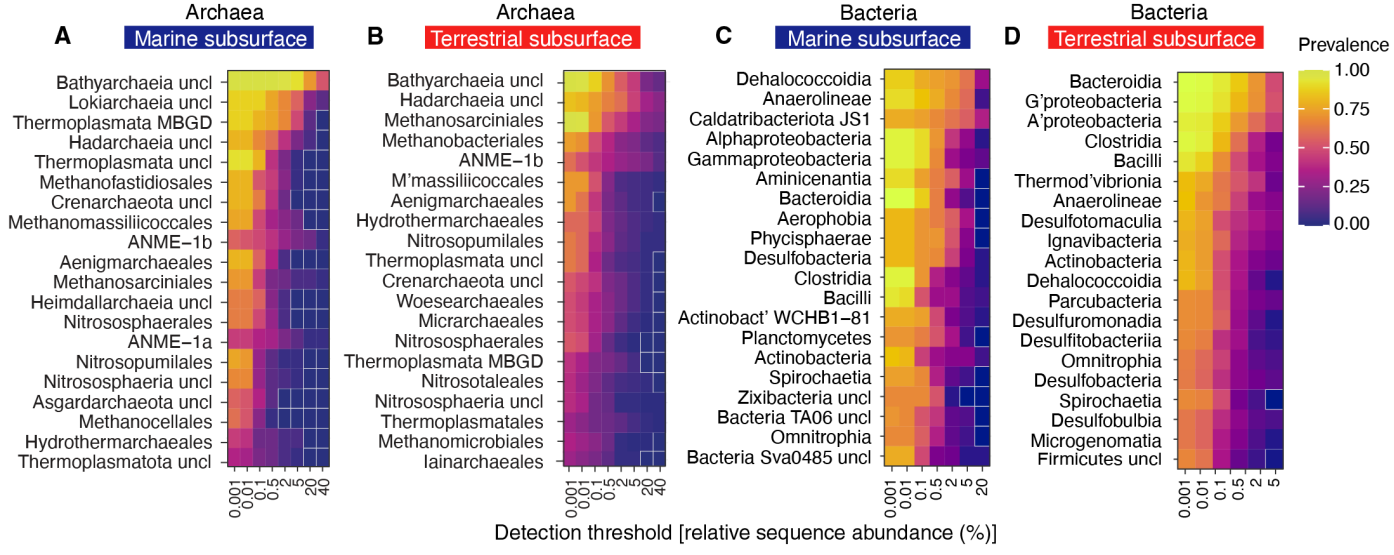

Fig. 6. Subsurface core microbiomes. The heatmaps show potential archaeal order level core microbiomes in the marine (**A**) and terrestrial subsurface (**B**), as well as bacterial class level core microbiomes in the marine (**C**) and terrestrial subsurface (**D**). The 20 most relevant lineages, regarding their prevalence and RSAs, are shown as rows, and each column color represents the lineages prevalence at a certain RSA threshold. Uncultured/unclassified lineages are denoted by "uncl," and the closest known phylogenetic level or the closest phylogenetic level with an isolated representative is shown. For example, Bathyarchaeia uncl are all uncultured/unclassified/order-level clades in the class Bathyarchaeia. Fields that are zero have a white outline.

subsurface samples (prevalence > 0.5) with RSAs of >40% (Fig. 6A), underlining their role as key players in the subsurface (*105*). The second category comprises lineages that have a high prevalence and low RSAs, i.e., they occur at almost all sites but are mostly rare. This category comprises terrestrial subsurface Bathyarchaeia and Methanosarciniales, as well as marine Gammaproteobacteria. Bathyarchaeia and Gammaproteobacteria were shown to be abundant and diverse by 16*S* rRNA gene amplicon studies in the marine (*17*, *105*, *106*) and terrestrial (*59*, *107*, *108*) subsurface biome, respectively. Our work finds that both lineages are in fact abundant and/or prevalent in all studied subsurface environments, further highlighting them as two of Earth's most widespread and sequence abundant microbial phyla. Remarkably, many other lineages also occur globally in marine and terrestrial subsurface environments at high prevalence and low RSA, including Bacilli, Clostridia, Bacteroidia, and Alphaproteobacteria (Fig. 6, C and D) and are thus truly cosmopolitan, albeit being mostly rare. Lineages of categories 1 and 2, being highly prevalent, likely comprise many generalists that have adapted to diverse habitats. Consistent with this idea, we find that, e.g., *Nitrosopumilus*, *Methanosaeta*, *Pseudomonas*, *Bacillus*, and *Desulfovibrio* are among the most common genera in the subsurface (figs. S22 and S23). The third category features those with low prevalence and high sequence abundance, which means that they do not occur at all sites, but when they occur, they are very sequence abundant or even dominant. This pattern fits to a specialist lifestyle in which organisms are very successful only in select habitats to which they are well adapted. This category includes marine *Lokiarchaeia* and ANME-1a, as well as Hadarchaeia and ANME-1b, and the latter too exhibit this distribution in both the marine and terrestrial subsurface. In the bacterial domain, examples include JS1 Caldatribacteriota, which constitute up to 20% of the community in about half of the marine subsurface samples (Fig. 6C), as well as marine Dehalococcoidia, Aminicenantia, Aerophophia, Phycisphaerae, and Desulfobacteria. The fourth category represents organisms with low prevalences and low sequence abundances. These lineages are generally rare, and their occurrence may be stochastic. Depending on an emphasis on prevalence or sequence abundance, the lineages in the first three categories, which are prevalent or abundant or both, can be interpreted as a core microbiome of the marine or terrestrial subsurface at the given phylogenetic level. Lineages that occur in our dataset widely, almost exclusively, in the marine subsurface include Lokiarchaeia, Heimdallarchaeia, Methanofastidiosales, Caldatribacteriota, Phycisphaerae, Aerophobia, and Aminicenantia. Lineages that occurred in our dataset almost exclusively in the terrestrial subsurface include Woesechaeales, Iainchaeales, Micrarchaeales, Thermodesulfovibrionia, Desulfotomaculia, and Ignavibacteria, a finding supported by previous studies (*42*, *92*). Although there is little community overlap between subsurface environments at species level, there are lineages at higher phylogenetic levels that occur in the marine and the terrestrial subsurface biome, including Bathyarchaeia, Hadarchaeia, ANME-1b, Alphaproteobacteria, Gammaproteobacteria, Anaerolineae, Bacteroidia, Bacilli, and Clostridia, which can be considered a global subsurface core microbiome at the given phylogenetic level.

Phylogenetic distance relative to cultured isolates is high in the subsurface

Many lineages that are abundant in Earth's environments belong to branches in the tree of life that have few or no cultured representatives (*74*, *109*). This is particularly true for lineages that are abundant in subsurface environments, e.g., Lokiarchaeia, Bathyarchaeia, Hadarchaeia, and Caldatribacteriota. Together, with these lineages often outsized contribution to richness (Fig. 4, B and D), it is likely that the subsurface holds a substantial degree of uncharted phylogenetic and likely metabolic diversity. We thus analyzed the phylogenetic distance of the detected lineages to their closest isolated relatives. For this, we mapped ASV against a SILVA reference database of cultivated isolates. Our analyses show that most archaeal ASVs in the marine subsurface share only 80 to 90% sequence identity to the closest isolated relative (Fig. 7A), on average 88% (fig. S24A), indicating that many uncharted family-, order-, and class-level clades can be expected (*110*). In the marine surface and interface and in terrestrial environments, archaeal ASVs are, on average, more closely related to a cultured isolate (95 to 97% sequence identity; Fig. 7A and fig. S24), nevertheless representing potential unseen lineages. In the case of bacteria, average sequence identity to the closest isolate is also lowest in the marine subsurface (88% on average), supporting the notion of extensive phylogenetically distant lineages in the seafloor (*17*). However, bacterial phylogenetic distance relative to cultured organisms is similarly high in all other environments, with most ASV being only ~90% sequentially identical to the closest isolate (Fig. 7B), ranging, on average, from 89 to 92% (fig. S24). Notably, many of the ASVs that are phylogenetically distant from their closest cultured relative have high RSAs (Fig. 7). This is especially true for Archaea, for which several phylogenetically distant ASV had RSAs between 1 and 10% (Fig. 7A). These findings support the notion of a largely untapped reservoir of uncultivated archaeal diversity in the subsurface (*111*). However, our findings also highlight the sequence abundance of lineages phylogenetically distant from the next isolate in most other environments, particularly in the bacterial domain. This supports findings of a previous global study in which most microbial population genomes were distantly related to reference genomes, suggesting that uncultivated organisms dominate the diversity within most phyla and environments (*74*).

The CoDL atlas reveals global continua and great divides of microbial diversity

Overall, we demonstrate that communities in the marine and terrestrial subsurface can be as phylogenetically diverse as those on the surface, harboring a substantial part of Earth's biodiversity. This finding confirms previous insights from global marine sediments (*17*) and the terrestrial subsurface (*59*) yet compares the two distinct biomes in one global analysis. Our comparison of marine and terrestrial microbiomes and depth realms shows that microbial diversity is particularly high in interface environments that are influenced by surface and subsurface processes. The communities at interface ecosystems shared community membership with those of surface and subsurface environments, which reflects the location, as well as the exchange of fluids and other materials between the depth realms. Mud volcanoes and methane seeps have been shown to connect surface and subsurface environments (*32*–*34*), as have marine crustal fluids (*112*, *113*). Archaeal interface communities shared more genus-level clades with the subsurface (Fig. 8, A and B), whereas bacterial interface communities shared more genus-level clades with the surface (Fig. 8, C and D), suggesting that different ecological processes shape archaeal and bacterial communities at these sites.

Despite harboring many distinct microbial lineages, the microbiomes of surface, interface, and subsurface environments showed a relatively high overlap in community composition. In previous studies of individual sites and sediment cores, it was shown that

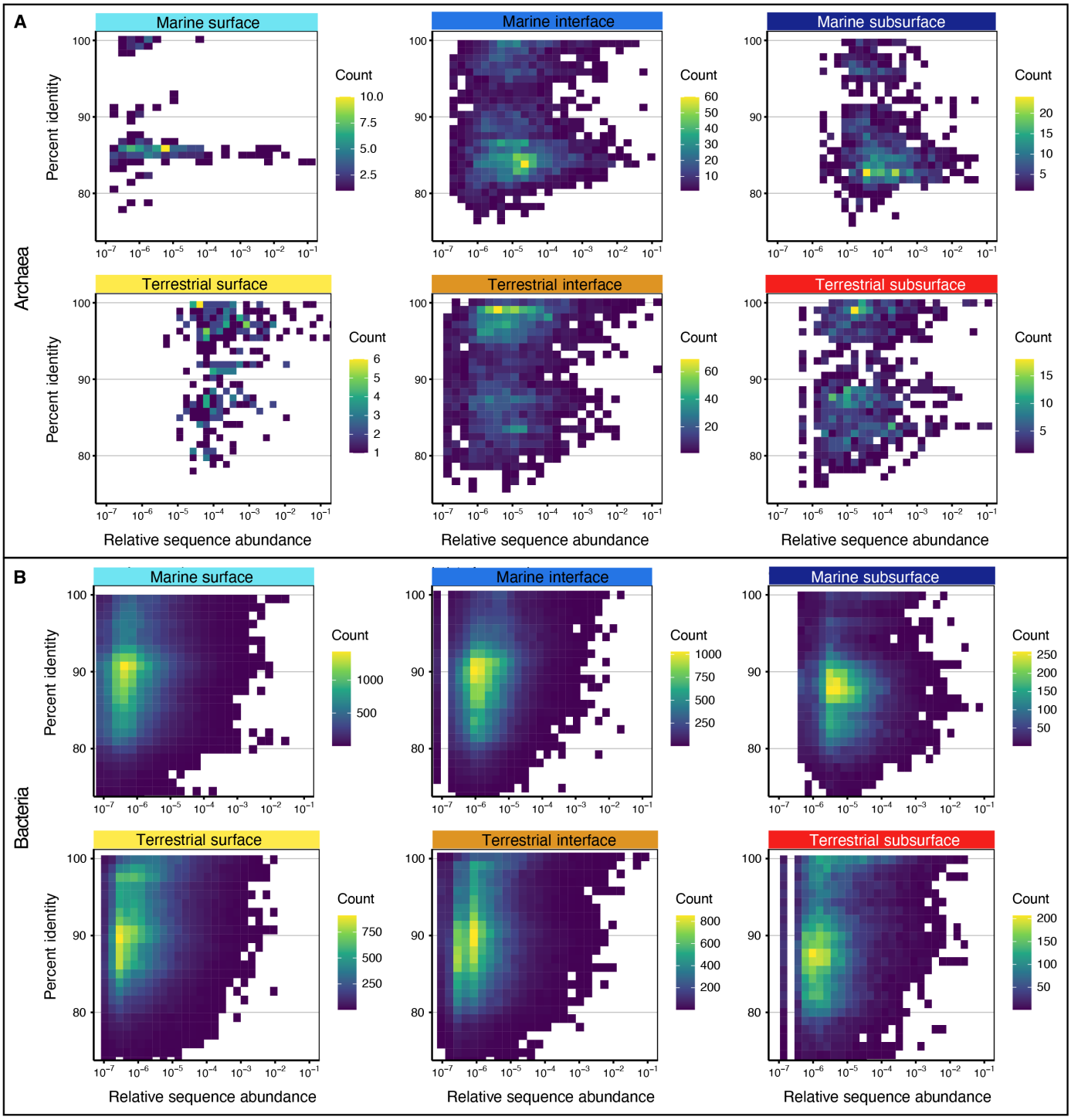

Fig. 7. Microbial phylogenetic distance relative to cultured organisms. Percent identity values (PIVs) of archaeal (**A**) and bacterial (**B**) ASVs relative to their closest cultured relative for the studied biomes and realms. Each square in the density plot represents one ASV and depicts its PIV (*y* axis) and RSA (*x* axis, logarithmic). The color gradient shows how many ASVs have identical PIV and RSA. The more yellow, the more ASVs are represented by the square.

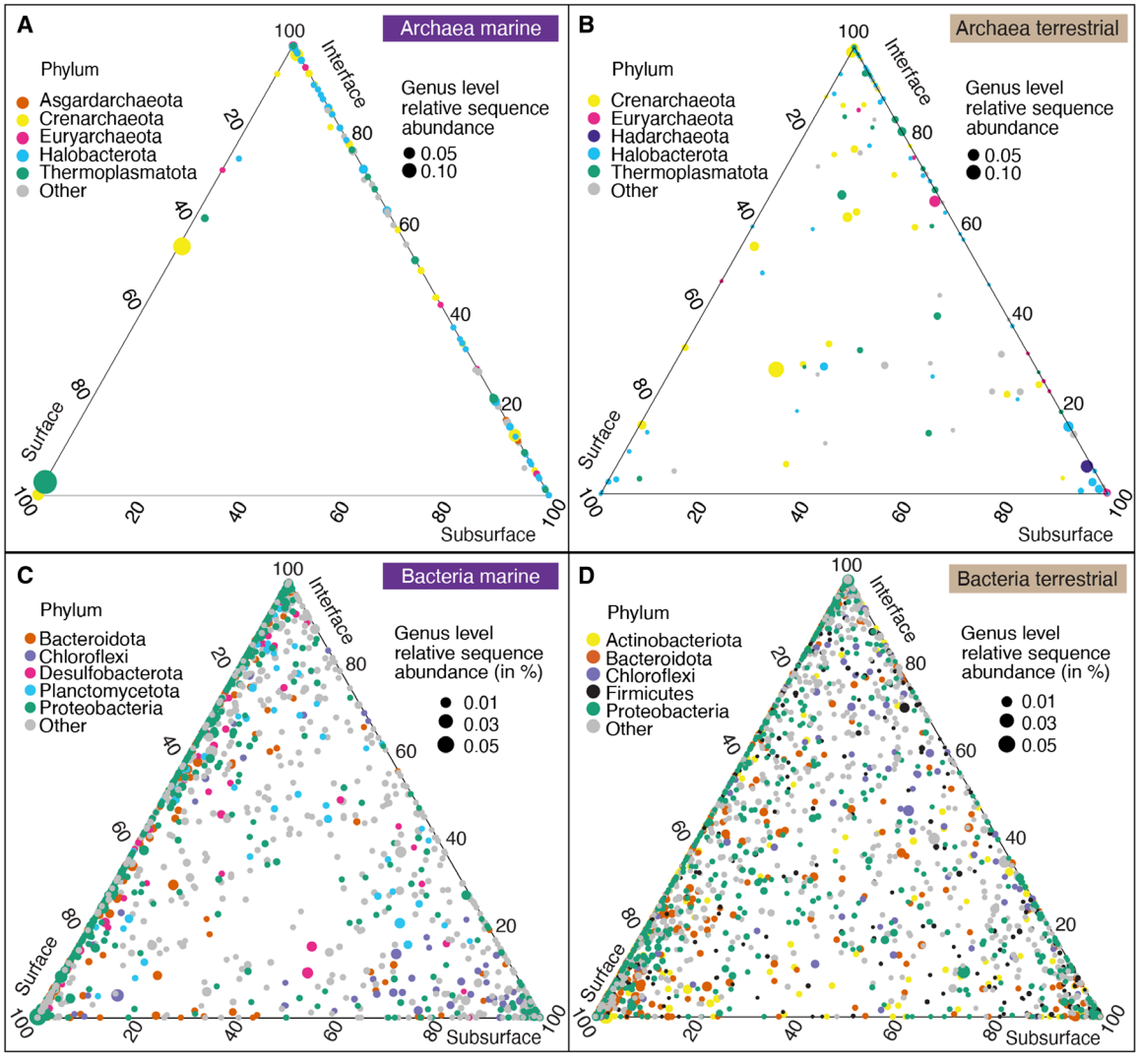

Fig. 8. Genus-level diversity within and between depth realms. Ternary plots show archaeal (**A** and **B**) and bacterial genus-level diversity (**C** and **D**) in the marine [(A) and (C)] and terrestrial biome [(B) and (D)]. Each circle is a genus-level clade, circle size is average RSA in the respective biome, and circle colors represent to which of the top 5 phyla the lineage belongs. The location of the circle shows the proportional average RSA in each depth realm (surface, interface and subsurface) scaled to sum 100%. For example, if a circle lies exactly in the center of the plot, then the respective lineage has an equal RSA in each of the three depths, if it lies at one of the corners (e.g., surface), then it means it occurs only in the surface, and circles that lie on the sides of the plot occur in only two of the three depths.

deep marine sediment layers contain communities similar to the surface seeded at the seafloor but getting more different with depth (*24*–*26*). Our standardized dataset provides further evidence at a large scale for such a diversity continuum with depth (fig. S12 and movies S1 and S2) that connects surface and subsurface communities across environments and deep time. Moreover, many of the environments we studied cannot be easily defined as surface or subsurface environment. Subsurface or subsurface-impacted ecosystems are not necessarily energy depleted and, in fact, can be relatively energy rich, such as hydrocarbon seeps (*114*); brines (*115*); hydrothermal vents (*116*); oil, coal, and shale deposits (*43*, *56*, *117*); serpentinizing systems (*118*–*120*); and caves (*121*). The subsurface is often replete with methane, hydrogen, and other energy and carbon sources (*50*, *122*–*125*). An absence of oxidants, nutrients, trace metals, or extreme physical or chemical conditions, however, can shape the specific community structures and hinder microbial growth or activity in relatively energy-rich ecosystems (*126*, *127*). Depth and oxygen content are also not always reliable proxies for subsurface conditions, e.g., in marine sediments with very low organic matter content, oxic conditions can prevail throughout the sediment column (*128*) as can oxygen be produced biotically or abiotically in subsurface ecosystems (*129*). Sediment depth and sediment age can be greatly different between locations (*77*), with past surface conditions or depositional environments (fig. S25) apparently being imprinted into a given subsurface sample (*25*, *26*). Defining environments and samples by a few select parameters such as depth, oxygen content, or age is useful, but to fully understand the subsurface geobiosphere, we may need to consider large-scale gradients and continua, analogous to the concept of the critical zone in soil science (*130*, *131*).

Our analyses further corroborate a major divide between marine and terrestrial microbiomes, showing significant differences in local and overall richness. In contrast to the studied depth realms, the differences between marine and terrestrial microbiomes were notable, with little to no overlap in community structure (Fig. 2 and fig. S6) between land- and sea-derived samples. This divide was also found on the basis of the analysis of metagenome-derived 16*S* rRNA and rpS3 genes (Fig. 2 and fig. S8). The taxonomic differentiation between marine and terrestrial microbial communities mirrors the differences between animal communities in these biomes (*132*) and is likely caused by the very different chemical and physical properties that shape these environments (*133*). The environmental drivers responsible for the gradients, continua, and divides remain undefined, but they likely include factors such as salinity (*65*, *66*), energy availability (*78*), geological activity (*134*, *135*), hydrologic conditions and recharge rates (*136*, *137*), and constrictions of void space within which microbial cells might exist (*127*). Improvements in global scale modeling and increasing availability of data can provide the tools to describe the surface and subsurface on large scales, refine our insights into microbial provinces in the subsurface (*138*), and potentially define regions of similar subsurface biogeochemical regimes analogous to Longhurst provinces for the surface ocean (*139*).

MATERIAL AND METHODS

Dataset specifications

To investigate the microbial diversity and composition of surface, interface, and subsurface microbiomes, we used 478 archaeal and 964 bacterial 16*S* rRNA gene amplicon datasets sequenced at the W. M. Keck Ecological and Evolutionary Genetics Facility of the Marine Biological Laboratory (Woods Hole, MA, USA). Most archaeal and bacterial datasets from subsurface and interface environments were part of the CoDL, and most datasets from surface ecosystems were not collected during the CoDL but were shared by numerous other investigators (see Acknowledgements). All sequencing datasets were prepared using identical primers, nearly identical library preparation protocols, identical sequencing chemistries, and bioinformatic analyses to minimize batch effects and biases and ensure comparability of communities based on ASVs. We removed datasets from enrichment samples, blanks, and controls, as well as datasets that we considered as failed runs (<2000 bacterial reads or <1000 archaeal reads) and subsurface datasets that contained lineages belonging to known contaminants of subsurface samples (*80*). Contextual data for each sample/sequence dataset is listed in dataset S2. Datasets S8 and S9, representing the archaeal and bacterial analysis logs using the "biome" grouping as a representative workflow). We are aware that certain environments are overrepresented (e.g., marine surface), while others are underrepresented (e.g., caves) or missing (e.g., soils). However, the highlighted global trends are likely reliable due to the size and breadth of the dataset, the comparability of environments, and the robustness of the results (e.g., toward the removal of reads; fig. S11).

16*S* rRNA gene library preparation and amplicon sequencing

Genomic DNA was extracted by the laboratories responsible for each project (dataset S1). DNA extraction could not be standardized due to commercial and logistical obstacles and laboratory-specific scientific practices and instrumentation. It was shown that DNA extraction kits can affect diversity indices and community structure (*140*); however, as this work comprises datasets from more than 50 laboratories worldwide, we can very likely exclude a systematic extraction bias, i.e., most/all samples of a certain tested category were extracted with the same method, which differs from the method used for samples of other categories. Sequencing was performed at the W. M. Keck Ecological and Evolutionary Genetics Facility at MBL. The bacterial 16*S* rRNA gene V4-V5 variable region was amplified using the forward primer 518F (5′-CCAGCAGCYGCGGTAAN-3′) and reverse primers 926R (5′-CCGTCAATTCNTTTRAGT-3′, 5′-CCGTCAATTTCTTTGAGT-3′, and 5′-CCGTCTATTCCTTTGANT-3′) (*141*). The archaeal 16*S* rRNA gene V4-V5 variable region was amplified using the forward primers 517F (5′-GCCTAAAGCATCCGTAGC-3′, 5′-GCCTAAARCGTYCGTAGC-3′, 5′-GTCTAAAGGGTCYGTAGC-3′, 5′-GCTTAAAGNGTYCGTAGC-3′, and 5′-GTCTAAARCGYYCGTAGC-3′) and reverse primer 958R (5′-CCGGCGTTGANTCCAATT-3′) (*142*). Amplicons were sequenced using Illumina's v3 600-cycle (paired-end) reagent kit on a MiSeq (Illumina Inc., San Diego, CA, USA). Reads were demultiplexed on the basis of the combination of index (CASAVA 1.8) and barcode (custom python scripts).

16*S* rRNA gene amplicon-based community analyses

Raw sequences were analyzed using DADA2 (*143*) following the DADA2 Pipeline Tutorial v1.16 (https://benjjneb.github.io/dada2/tutorial.html). Each Illumina run was analyzed separately to use run-specific error profiles, as is recommended best practice for big data (https://benjjneb.github.io/dada2/bigdata.html). Briefly, forward and reverse reads were quality-trimmed to 275 and 205 base pairs (bp), respectively, and primer sequences (17-bp forward, 21-bp reverse) were removed. Reads with more than two expected errors were discarded, and paired reads were merged. All runs were combined, and chimeric sequences were removed. Species level taxonomy was assigned with the silva_nr_v138_train_set and silva_species_assignment_v138 based on the Silva small subunit reference database SSURef v138 [release date: 16 December 2019; (*144*)]. After sequence quality control and the removal of enrichments, blanks, and biological and technical replicates, we removed all sequences that affiliated with known contaminants (a list is given in the Supplementary Materials). Of the samples that remained after contaminant removal, we only analyzed the datasets comprising more than 1000 archaeal or 2000 bacterial reads because samples below these thresholds were either failed runs or originally highly contaminated. After all quality control steps, we obtained 478 archaeal and 964 bacterial amplicon datasets. Archaeal datasets contained a total of 4.9×10^7 sequence reads belonging to 31,099 unique ASVs. Archaeal samples had, on average, 1.02×10^5 reads and 155 ASVs (dataset S3). Bacterial datasets comprised a total of 9.6×10^7 sequence reads belonging to 377,374 unique ASVs. Bacterial samples had, on average, 0.99×10^5 reads and 1183 unique ASVs (dataset S4). Amplicon datasets may not quantitatively represent the sampled community yet are reliable to compare community structure and make ecological interpretations (*83*). Alpha diversity (richness, Shannon entropy, inverse Simpson diversity, and Chao1-estimated richness) was calculated from the ASV-by-sample table using a subsampling approach to account for unequal sampling effort. We used 1142 and 2271 randomly chosen reads from each archaeal and bacterial sample, respectively, calculated 10 iterations of the respective diversity index per sample, took the average value of the 10 iterations, and visualized these averages as boxplots. Total diversity within groups of samples (gamma diversity) was calculated similarly; 1142 archaeal or 2271

bacterial reads were used to subsample the ASV table; gamma diversity (total richness and Shannon) was calculated from *n* samples, where *n* is the sample number of the smallest category; this procedure was iterated 100 times, and the values were shown as boxplots. Even when using very stringent subsampling conditions of 50,000 randomly chosen archaeal and bacterial sequences, respectively, the trends in alpha and beta diversity did not change substantially; hence, we decided to include as many samples as possible. Bray-Curtis dissimilarities (*145*) between all samples were calculated and used for two-dimensional nonmetric multidimensional scaling (NMDS) ordinations with 20 random starts (*146*). All analyses were carried out with VisuaR (https://github.com/EmilRuff/VisuaR)—a workflow based on the R statistical environment, custom R scripts and several R packages including vegan (*147*) and ggplot2. Exemplary analysis logs (datasets S8 and S9), the used DADA2 script (dataset S10), the VisuaR v40 script (dataset S11), and an example VisuaR user input file (dataset S12) for this study are available as the Supplementary Materials.

Gene sequence identity analyses

To determine the relatedness of detected 16*S* rRNA gene sequences to those of the closest isolated strain, we performed a BLASTn search of each 16*S* rRNA gene amplicon sequence against a database of sequences from cultured archaeal and bacterial isolates, selected to yield only the single best hit for each amplicon. The isolate database was created with makeblastdb using sequences from cultured isolates in the SILVA NR reference database v138.2 available at (www.arb-silva.de/). We included recently cultured *(Cald)atribacteria* (*102*). Only alignment lengths >350 (97% of bacterial amplicons and 62% of archaeal amplicons) were analyzed.

Metagenomic DNA sequencing, contig assembly, gene annotation, and count

Paired-end libraries (2 × 151 bp) were prepared for a subset of samples using an Illumina TruSeq DNA library preparation kit (Illumina, San Diego, CA, USA) and sequenced on the Illumina NextSeq500 platform. Raw demultiplexed reads were quality processed using bbduk.sh in BBMap v.38.71 (Bushnell B; https://sourceforge.net/projects/bbmap/) in two successive steps by first removing Illumina adapters from read termini (with options: ktrim = r minlen = 40 mink = 11 tbo tpe k = 23 hdist = 1 hdist2 = 1 ftm = 5) before quality filtering against PhiX genome to an average quality threshold of 12 (with options: maq = 12 trimq = 12 qtrim = rl maxns = 3 minlen = 40 k = 31 hdist = 1). Quality trimmed reads mapping to the human HG19 genome with more than 95% identity were discarded using bbmap.sh in BBMap (with options: local minratio = 0.9 usemodulo maxindel = 3 bwr = 0.16 bw = 12 quickmatch fast minhits = 2 qtrim = rl trimq = 12 untrim idtag printunmappedcount kfilter = 25 maxsites = 1 k = 14). Cleaned reads were used to construct de novo assemblies for each sample using metaSPAdes (v3.13.0) with default parameters (*148*). Open reading frames (ORFs) were predicted on assembled contigs ≥1 kb using Prodigal (v2.6.3) (setting -p meta) (*149*). Predicted genes were clustered at ≥95% sequence identity and ≥ 80% overlap using dedupe.sh in BBMap (with option arc = t am = t ac = t mid = 95 mlp = 80) to produce a unique gene catalog of 15,663,206 nonredundant genes. The unigene was functionally assigned to a KEGG Orthology (KO) using kofamscan (v1.3) (*150*). rpS3 sequences were extracted from all predicted ORFs (*151*) using hmmsearch (*152*) against custom hidden Markov models for Archaea and Bacteria (https://github.com/AJProbst/rpS3_trckr), yielding a total of 8691 rpS3 sequences. RpS3 amino acids were dereplicated at ≥99% identity and ≥80% overlap with dedupe.sh (with options: arc = t am = t ac = t mid = 99 mlp = 80) to produce 4392 rpS3 single-gene sequences, equivalent to microbial species. Taxonomic classification of rpS3 genes was performed using Kaiju (v1.7.3, with option: greedy-5 mode) (*153*).

For taxonomic and gene count analyses, cleaned paired-end reads were subsampled to a depth of 2 M reads per sample using reformat.sh in BBMap (with options: samplereadstarget = 2,000,000, sampleseed = 121) and then mapped back to the annotated genes using BBMap with read alignment of ≥95% for the unigene (with options: minid = 0.95 idfilter = 0.95, ambiguous = random) and ≥ 99% for rpS3 genes (with options: minid = 0.99 idfilter = 0.99, ambiguous = random). Gene count tables were normalized to gene per million (e.g., equivalent to transcript per million), following (*154*). Metagenomic short reads were mapped to the SILVA SSU reference database (*144*) to assign nearest taxonomic units, and full-length 16*S*/18S rRNA gene sequences were reconstructed from metagenomes using phyloFlash v3.4 (*155*). Metagenome-derived rpS3 and 16*S* rRNA genes were subsampled (datasets S3 and S4), analyzed, and visualized analogous to the amplicon-derived 16*S* rRNA genes using the same VisuaR community analysis workflow. Overall, metagenome based diversity trends (figs. S6 to S8 and S10 to S12) and microbial composition (fig. S17, B and C) were similar to those detected by metabarcoding.

Statistical analyses

Differences in alpha diversity metrics between conditions were tested using the Wilcoxon signed rank test (ggsignif) as implemented in ggplot2 (*156*). *P* values were corrected stringently using the Bonferroni method. We used analysis of similarity (*157*) as implemented in vegan to test whether community structure between conditions was significantly different, as visualized in NMDS plots. Multivariate associations were tested as implemented in the MaAsLin2 method (*86*).

Supplementary Materials

The PDF file includes:

Supplementary Text
Supplementary Methods
Figs. S1 to S24
Legends for datasets S1 to S12
Legends for movies S1 and S2
References

Other Supplementary Material for this manuscript includes the following:

Datasets S1 to S12
Movies S1 and S2

REFERENCES AND NOTES

1. K. Takai, K. Nakamura, T. Toki, U. Tsunogai, M. Miyazaki, J. Miyazaki, H. Hirayama, S. Nakagawa, T. Nunoura, K. Horikoshi, Cell proliferation at 122°C and isotopically heavy CH_4 production by a hyperthermophilic methanogen under high-pressure cultivation. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 10949–10954 (2008).
2. N. Merino, H. S. Aronson, D. P. Bojanova, J. Feyhl-Buska, M. L. Wong, S. Zhang, D. Giovannelli, Living at the extremes: Extremophiles and the limits of life in a planetary context. *Front. Microbiol.* **10**, 780 (2019).
3. V. P. Edgcomb, S. J. Molyneaux, S. Böer, C. O. Wirsen, M. Saito, M. S. Atkins, K. Lloyd, A. Teske, Survival and growth of two heterotrophic hydrothermal vent archaea, Pyrococcus strain GB-D and Thermococcus fumicolans, under low pH and high sulfide concentrations in combination with high temperature and pressure regimes. *Extremophiles.* **11**, 329–342 (2007).

4. N. Ando, B. Barquera, D. H. Bartlett, E. Boyd, A. A. Burnim, A. S. Byer, D. Colman, R. E. Gillilan, M. Gruebele, G. Makhatadze, C. A. Royer, E. Shock, A. J. Wand, M. B. Watkins, The molecular basis for life in extreme environments. *Annu. Rev. Biophys.* **50**, 343–372 (2021).
5. Y. M. Bar-On, R. Phillips, R. Milo, The biomass distribution on Earth. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 6506–6511 (2018).
6. A. Teske, K. B. Sørensen, Uncultured archaea in deep marine subsurface sediments: Have we caught them all? *ISME J.* **2**, 3–18 (2008).
7. C. Magnabosco, L.-H. Lin, H. Dong, M. Bomberg, W. Ghiorse, H. Stan-Lotter, K. Pedersen, T. L. Kieft, E. van Heerden, T. C. Onstott, The biomass and biodiversity of the continental subsurface. *Nat. Geosci.* **11**, 707–717 (2018).
8. J. S. Lipp, Y. Morono, F. Inagaki, K.-U. Hinrichs, Significant contribution of Archaea to extant biomass in marine subsurface sediments. *Nature* **454**, 991–994 (2008).
9. T. W. Federle, D. C. Dobbins, J. R. Thornton-Manning, D. D. Jones, Microbial biomass, activity, and community structure in subsurface soils. *Groundwater* **24**, 365–374 (1986).
10. J. F. Biddle, J. B. Sylvan, W. J. Brazelton, B. J. Tully, K. J. Edwards, C. L. Moyer, J. F. Heidelberg, W. C. Nelson, Prospects for the study of evolution in the deep biosphere. *Front. Microbiol.* **2**, 285 (2012).
11. W. B. Whitman, D. C. Coleman, W. J. Wiebe, Prokaryotes: The unseen majority. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6578–6583 (1998).
12. R. C. Beaver, J. D. Neufeld, Microbial ecology of the deep terrestrial subsurface. *ISME J.* **18**, wrae091 (2024).
13. E. O. Straume, C. Gaina, S. Medvedev, K. Hochmuth, K. Gohl, J. M. Whittaker, R. Abdul Fattah, J. C. Doornenbal, J. R. Hopper, GlobSed: Updated total sediment thickness in the world's oceans. *Geochem. Geophys. Geosyst.* **20**, 1756–1772 (2019).
14. J. F. Biddle, J. S. Lipp, M. A. Lever, K. G. Lloyd, K. B. Sørensen, R. Anderson, H. F. Fredricks, M. Elvert, T. J. Kelly, D. P. Schrag, M. L. Sogin, J. E. Brenchley, A. Teske, C. H. House, K.-U. Hinrichs, Heterotrophic Archaea dominate sedimentary subsurface ecosystems off Peru. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 3846–3851 (2006).
15. S. Xie, J. S. Lipp, G. Wegener, T. G. Ferdelman, K.-U. Hinrichs, Turnover of microbial lipids in the deep biosphere and growth of benthic archaeal populations. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 6010–6014 (2013).
16. K. G. Lloyd, M. K. May, R. T. Kevorkian, A. D. Steen, Meta-analysis of quantification methods shows that archaea and bacteria have similar abundances in the subseafloor. *Appl. Environ. Microbiol.* **79**, 7790–7799 (2013).
17. T. Hoshino, H. Doi, G. Uramoto, L. Wörmer, R. R. Adhikari, N. Xiao, Y. Morono, S. D'Hondt, K.-U. Hinrichs, F. Inagaki, Global diversity of microbial communities in marine sediment. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 27587–27597 (2020).
18. J. Kallmeyer, R. Pockalny, R. R. Adhikari, D. C. Smith, S. D'Hondt, Global distribution of microbial abundance and biomass in subseafloor sediment. *Proc. Natl. Acad. Sci.* **109**, 16213–16216 (2012).
19. H.-C. Flemming, S. Wuertz, Bacteria and archaea on Earth and their abundance in biofilms. *Nat. Rev. Microbiol.* **17**, 247–260 (2019).
20. S. McMahon, J. Parnell, Weighing the deep continental biosphere. *FEMS Microbiol. Ecol.* **87**, 113–120 (2014).
21. R. J. Parkes, B. Cragg, E. Roussel, G. Webster, A. Weightman, H. Sass, A review of prokaryotic populations and processes in sub-seafloor sediments, including biosphere:Geosphere interactions. *Mar. Geol.* **352**, 409–425 (2014).
22. Y. Morono, M. Ito, T. Hoshino, T. Terada, T. Hori, M. Ikehara, S. D'Hondt, F. Inagaki, Aerobic microbial life persists in oxic marine sediment as old as 101.5 million years. *Nat. Commun.* **11**, 3626 (2020).
23. E. A. Walsh, J. B. Kirkpatrick, S. D. Rutherford, D. C. Smith, M. Sogin, S. D'Hondt, Bacterial diversity and community composition from seasurface to subseafloor. *ISME J.* **10**, 979–989 (2016).
24. P. Starnawski, T. Bataillon, T. J. G. Ettema, L. M. Jochum, L. Schreiber, X. Chen, M. A. Lever, M. F. Polz, B. B. Jørgensen, A. Schramm, K. U. Kjeldsen, Microbial community assembly and evolution in subseafloor sediment. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 2940–2945 (2017).
25. J. B. Kirkpatrick, E. A. Walsh, S. D'Hondt, Microbial selection and survival in subseafloor sediment. *Front. Microbiol.* **10**, 956 (2019).
26. C. Petro, B. Zäncker, P. Starnawski, L. M. Jochum, T. G. Ferdelman, B. B. Jørgensen, H. Røy, K. U. Kjeldsen, A. Schramm, Marine deep biosphere microbial communities assemble in near-Surface sediments in Aarhus Bay. *Front. Microbiol.* **10**, 758 (2019).
27. V. Hubalek, X. Wu, A. Eiler, M. Buck, C. Heim, M. Dopson, S. Bertilsson, D. Ionescu, Connectivity to the surface determines diversity patterns in subsurface aquifers of the Fennoscandian shield. *ISME J.* **10**, 2447–2458 (2016).
28. J. L. Meyer, U. Jaekel, B. J. Tully, B. T. Glazer, C. G. Wheat, H. Lin, C. Hsieh, J. P. Cowen, S. M. Hulme, P. R. Girguis, J. A. Huber, A distinct and active bacterial community in cold oxygenated fluids circulating beneath the western flank of the Mid-Atlantic ridge. *Sci. Rep.* **6**, 22541 (2016).
29. L. I. Putman, M. C. Sabuda, W. J. Brazelton, M. D. Kubo, T. M. Hoehler, T. M. McCollom, D. Cardace, M. O. Schrenk, Microbial communities in a serpentinizing aquifer are assembled through strong concurrent dispersal limitation and selection. *mSystems* **6**, e0030021 (2021).
30. Y. Zhang, A. E. Dekas, A. J. Hawkins, A. E. Parada, O. Gorbatenko, K. Li, R. N. Horne, *Water Resour. Res.*, 56, e2019WR025916 (2020).
31. D. Scheidweiler, F. Miele, H. Peter, T. J. Battin, P. de Anna, Trait-specific dispersal of bacteria in heterogeneous porous environments: From pore to porous medium scale. *J. R. Soc. Interface* **17**, 20200046 (2020).
32. T. Hoshino, T. Toki, A. Ijiri, Y. Morono, H. Machiyama, J. Ashi, K. Okamura, F. Inagaki, *Atribacteria* from the subseafloor sedimentary biosphere disperse to the hydrosphere through submarine mud volcanoes. *Front. Microbiol.* **8**, 1135 (2017).
33. S. E. Ruff, J. Felden, H. R. Gruber-Vodicka, Y. Marcon, K. Knittel, A. Ramette, A. Boetius, In situ development of a methanotrophic microbiome in deep-sea sediments. *ISME J.* **13**, 197–213 (2019).
34. A. Chakraborty, S. E. Ruff, X. Dong, E. D. Ellefson, C. Li, J. M. Brooks, J. McBee, B. B. Bernard, C. R. J. Hubert, Hydrocarbon seepage in the deep seabed links subsurface and seafloor biospheres. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 11029–11037 (2020).
35. D. A. Gittins, P.-A. Desiage, N. Morrison, J. E. Rattray, S. Bhatnagar, A. Chakraborty, J. Zorz, C. Li, O. Horanszky, M. A. Cramm, F. Bisiach, R. Bennett, J. Webb, A. MacDonald, M. Fowler, D. C. Campbell, C. R. J. Hubert, Geological processes mediate a microbial dispersal loop in the deep biosphere. *Sci. Adv.* **8**, eabn3485 (2022).
36. A. D. Steen, R. T. Kevorkian, J. T. Bird, N. Dombrowski, B. J. Baker, S. M. Hagen, K. H. Mulligan, J. M. Schmidt, A. T. Webber, T. M. Royalty, M. J. Alperin, Kinetics and identities of extracellular peptidases in subsurface sediments of the white Oak River Estuary, North Carolina. *Appl. Environ. Microbiol.* **85**, e00102-19 (2019).
37. J. T. Bird, E. D. Tague, L. Zinke, J. M. Schmidt, A. D. Steen, B. Reese, I. P. G. Marshall, G. Webster, A. Weightman, H. F. Castro, S. R. Campagna, K. G. Lloyd, Uncultured microbial phyla suggest mechanisms for multi-thousand-year subsistence in Baltic sea sediments. *MBio* **10**, e02376-18 (2019).
38. W. D. Orsi, V. P. Edgcomb, G. D. Christman, J. F. Biddle, Gene expression in the deep biosphere. *Nature.* **499**, 205–208 (2013).
39. M. A. Lever, Acetogenesis in the energy-starved deep biosphere—A paradox? *Front. Microbiol.* **2**, 284 (2012).
40. Y. He, M. Li, V. Perumal, X. Feng, J. Fang, J. Xie, S. M. Sievert, F. Wang, Genomic and enzymatic evidence for acetogenesis among multiple lineages of the archaeal phylum Bathyarchaeota widespread in marine sediments. *Nat. Microbiol.* **1**, 16035 (2016).
41. C. J. Newberry, G. Webster, B. A. Cragg, R. J. Parkes, A. J. Weightman, J. C. Fry, Diversity of prokaryotes and methanogenesis in deep subsurface sediments from the Nankai Trough, Ocean Drilling Program Leg 190. *Environ. Microbiol.* **6**, 274–287 (2004).
42. L. Momper, C. P. Casar, M. R. Osburn, A metagenomic view of novel microbial and metabolic diversity found within the deep terrestrial biosphere at DeMMO: A microbial observatory in South Dakota , USA. *Environ. Microbiol.* **25**, 3719–3737 (2023).
43. S. E. Ruff, P. Humez, I. H. de Angelis, M. Diao, M. Nightingale, S. Cho, L. Connors, O. O. Kuloyo, A. Seltzer, S. Bowman, S. D. Wankel, C. N. McClain, B. Mayer, M. Strous, Hydrogen and dark oxygen drive microbial productivity in diverse groundwater ecosystems. *Nat. Commun.* **14**, 3194 (2023).
44. M. H. Møller, C. Glombitza, M. A. Lever, L. Deng, Y. Morono, F. Inagaki, M. Doll, C. Su, B. A. Lomstein, D:L-amino acid modeling reveals fast microbial turnover of days to months in the subsurface hydrothermal sediment of Guaymas Basin. *Front. Microbiol.* **9**, 967 (2018).
45. J. A. Bradley, J. P. Amend, D. E. LaRowe, Necromass as a Limited Source of Energy for Microorganisms in Marine Sediments. *J. Geophys. Res. Biogeosci.* **123**, 577–590 (2018).
46. S. Pérez Castro, M. A. Borton, K. Regan, I. Hrabe de Angelis, K. C. Wrighton, A. P. Teske, M. Strous, S. E. Ruff, Degradation of biological macromolecules supports uncultured microbial populations in Guaymas Basin hydrothermal sediments. *ISME J.* **15**, 3480–3497 (2021).
47. N. M. Chaudhari, W. A. Overholt, P. A. Figueroa-Gonzalez, M. Taubert, T. L. V. Bornemann, A. J. Probst, M. Hölzer, M. Marz, K. Küsel, The economical lifestyle of CPR bacteria in groundwater allows little preference for environmental drivers. *Environ. Microbiome* **16**, 24 (2021).
48. K. Schwank, T. L. V. Bornemann, N. Dombrowski, A. Spang, J. F. Banfield, A. J. Probst, An archaeal symbiont-host association from the deep terrestrial subsurface. *ISME J.* **13**, 2135–2139 (2019).
49. M. O. Schrenk, W. J. Brazelton, S. Q. Lang, Serpentinization, Carbon, and Deep Life. *Rev. Mineral. Geochemistry.* **75**, 575–606 (2013).
50. J. F. Sauvage, A. Flinders, A. J. Spivack, R. Pockalny, A. G. Dunlea, C. H. Anderson, D. C. Smith, R. W. Murray, S. D'Hondt, The contribution of water radiolysis to marine sedimentary life. *Nat. Commun.* **12**, 1297 (2021).
51. L.-H. Lin, J. Hall, J. Lippmann-Pipke, J. A. Ward, B. Sherwood Lollar, M. DeFlaun, R. Rothmel, D. Moser, T. M. Gihring, B. Mislowack, T. C. Onstott, Radiolytic H_2 in continental crust: Nuclear power for deep subsurface microbial communities. *Geochem. Geophys. Geosyst.* **6**, Q07003 (2005).
52. J. T. Lennon, S. E. Jones, Microbial seed banks: The ecological and evolutionary implications of dormancy. *Nat. Rev. Microbiol.* **9**, 119–130 (2011).
53. A. Schippers, L. N. Neretin, J. Kallmeyer, T. G. Ferdelman, B. A. Cragg, R. John Parkes, B. B. Jørgensen, Prokaryotic cells of the deep sub-seafloor biosphere identified as living bacteria. *Nature* **433**, 861–864 (2005).

54. Y. Morono, T. Terada, M. Nishizawa, M. Ito, F. Hillion, N. Takahata, Y. Sano, F. Inagaki, Carbon and nitrogen assimilation in deep subseafloor microbial cells. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 18295–18300 (2011).
55. B. B. Jørgensen, Deep subseafloor microbial cells on physiological standby. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 18193–18194 (2011).
56. E. Trembath-Reichert, Y. Morono, A. Ijiri, T. Hoshino, K. S. Dawson, F. Inagaki, V. J. Orphan, Methyl-compound use and slow growth characterize microbial life in 2-km-deep subseafloor coal and shale beds. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E9206–E9215 (2017).
57. P. B. Price, T. Sowers, Temperature dependence of metabolic rates for microbial growth, maintenance, and survival. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 4631–4636 (2004).
58. S. Braun, Y. Morono, S. Littmann, M. Kuypers, H. Aslan, M. Dong, B. B. Jørgensen, B. A. Lomstein, Size and carbon content of sub-seafloor microbial cells at landsort deep, Baltic Sea. *Front. Microbiol.* **7**, 1375 (2016).
59. A. Soares, A. Edwards, D. An, A. Bagnoud, J. Bradley, E. Barnhart, M. Bomberg, K. Budwill, S. M. Caffrey, M. Fields, J. Gralnick, V. Kadnikov, L. Momper, M. Osburn, A. Mu, J. W. Moreau, D. Moser, L. Purkamo, S. M. Rassner, C. S. Sheik, B. Sherwood Lollar, B. M. Toner, G. Voordouw, K. Wouters, A. C. Mitchell, A global perspective on bacterial diversity in the terrestrial deep subsurface. *Microbiology* **169**, 001172 (2023).
60. Y. Wang, H.-F. Sheng, Y. He, J.-Y. Wu, Y.-X. Jiang, N. F.-Y. Tam, H.-W. Zhou, Comparison of the levels of bacterial diversity in freshwater, Intertidal wetland, and marine sediments by using millions of illumina tags. *Appl. Environ. Microbiol.* **78**, 8264–8271 (2012).
61. E. R. Handler, S. D. J. Andersen, R. Gradinger, M. McGovern, A. Vader, A. E. Poste, Seasonality in land-ocean connectivity and local processes control sediment bacterial community structure and function in a High Arctic tidal flat. *FEMS Microbiol. Ecol.* **100**, fiad162 (2024).
62. H. S. Tee, D. Waite, G. Lear, K. M. Handley, Microbial river-to-sea continuum: gradients in benthic and planktonic diversity, osmoregulation and nutrient cycling. *Microbiome* **9**, 190 (2021).
63. T. Fenchel, M. Whitfield, P. Meadows, J. Huisman, Microbial ecology on land and sea [and discussion]. *Philos. Trans. Biol. Sci.* **343**, 51–56 (1994).
64. N. Lang-Yona, J. M. Flores, R. Haviv, A. Alberti, J. Poulain, C. Belser, M. Trainic, D. Gat, H. Ruscheweyh, P. Wincker, S. Sunagawa, Y. Rudich, I. Koren, A. Vardi, Terrestrial and marine influence on atmospheric bacterial diversity over the north Atlantic and Pacific Oceans. *Commun. Earth Environ.* **3**, 121 (2022).
65. C. A. Lozupone, R. Knight, Global patterns in bacterial diversity. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 11436–11440 (2007).
66. K. T. Jurdzinski, M. Mehrshad, L. F. Delgado, Z. Deng, S. Bertilsson, A. F. Andersson, Large-scale phylogenomics of aquatic bacteria reveal molecular mechanisms for adaptation to salinity. *Sci. Adv.* **9**, eadg2059 (2023).
67. L. R. Thompson, J. G. Sanders, D. McDonald, A. Amir, J. Ladau, K. J. Locey, R. J. Prill, A. Tripathi, S. M. Gibbons, G. Ackermann, J. A. Navas-Molina, S. Janssen, E. Kopylova, Y. Vázquez-Baeza, A. González, J. T. Morton, S. Mirarab, Z. Zech Xu, L. Jiang, M. F. Haroon, J. Kanbar, Q. Zhu, S. Jin Song, T. Kosciolek, N. A. Bokulich, J. Lefler, C. J. Brislawn, G. Humphrey, S. M. Owens, J. Hampton-Marcell, D. Berg-Lyons, V. McKenzie, N. Fierer, J. A. Fuhrman, A. Clauset, R. L. Stevens, A. Shade, K. S. Pollard, K. D. Goodwin, J. K. Jansson, J. A. Gilbert, R. Knight, Earth Microbiome Project Consortium, A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551**, 457–463 (2017).
68. T. M. Voser, M. D. Campbell, A. R. Carroll, How different are marine microbial natural products compared to their terrestrial counterparts? *Nat. Prod. Rep.* **39**, 7–19 (2022).
69. S. Sunagawa, L. P. Coelho, S. Chaffron, J. R. Kultima, K. Labadie, G. Salazar, B. Djahanschiri, G. Zeller, D. R. Mende, A. Alberti, F. M. Cornejo-Castillo, P. I. Costea, C. Cruaud, F. D'Ovidio, S. Engelen, I. Ferrera, J. M. Gasol, L. Guidi, F. Hildebrand, F. Kokoszka, C. Lepoivre, G. Lima-Mendez, J. Poulain, B. T. Poulos, M. Royo-Llonch, H. Sarmento, S. Vieira-Silva, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis, C. Bowler, C. de Vargas, G. Gorsky, N. Grimsley, P. Hingamp, D. Iudicone, O. Jaillon, F. Not, H. Ogata, S. Pesant, S. Speich, L. Stemmann, M. B. Sullivan, J. Weissenbach, P. Wincker, E. Karsenti, J. Raes, S. G. Acinas, P. Bork, E. Boss, C. Bowler, M. Follows, L. Karp-Boss, U. Krzic, E. G. Reynaud, C. Sardet, M. Sieracki, D. Velayoudon, Structure and function of the global ocean microbiome. *Science* **348**, 1261359 (2015).
70. M. Mestre, C. Ruiz-González, R. Logares, C. M. Duarte, J. M. Gasol, M. M. Sala, Sinking particles promote vertical connectivity in the ocean microbiome. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E6799–E6807 (2018).
71. S. G. Acinas, P. Sánchez, G. Salazar, F. M. Cornejo-Castillo, M. Sebastián, R. Logares, M. Royo-Llonch, L. Paoli, S. Sunagawa, P. Hingamp, H. Ogata, G. Lima-Mendez, S. Roux, J. M. González, J. M. Arrieta, I. S. Alam, A. Kamau, C. Bowler, J. Raes, S. Pesant, P. Bork, S. Agustí, T. Gojobori, D. Vaqué, M. B. Sullivan, C. Pedrós-Alió, R. Massana, C. M. Duarte, J. M. Gasol, Deep ocean metagenomes provide insight into the metabolic architecture of bathypelagic microbial communities. *Commun. Biol.* **4**, 604 (2021).
72. M. Delgado-Baquerizo, A. M. Oliverio, T. E. Brewer, A. Benavent-González, D. J. Eldridge, R. D. Bardgett, F. T. Maestre, B. K. Singh, N. Fierer, A global atlas of the dominant bacteria found in soil. *Science* **359**, 320–325 (2018).
73. E. B. Graham, A. P. Camargo, R. Wu, R. Y. Neches, M. Nolan, D. Paez-Espino, N. C. Kyrpides, J. K. Jansson, J. E. McDermott, K. S. Hofmockel, Soil Virosphere Consortium, A global atlas of soil viruses reveals unexplored biodiversity and potential biogeochemical impacts. *Nat. Microbiol.* , 1873–1883 (2024).
74. S. Nayfach, S. Roux, R. Seshadri, D. Udwary, N. Varghese, F. Schulz, D. Wu, D. Paez-Espino, I. Chen, M. Huntemann, K. Palaniappan, J. Ladau, S. Mukherjee, T. B. K. Reddy, T. Nielsen, E. Kirton, J. P. Faria, J. N. Edirisinghe, C. S. Henry, S. P. Jungbluth, D. Chivian, P. Dehal, E. M. Wood-Charlson, A. P. Arkin, S. G. Tringe, A. Visel, IMG/M Data Consortium, T. Woyke, N. J. Mouncey, N. N. Ivanova, N. C. Kyrpides, E. A. Eloe-Fadrosh, A genomic catalog of Earth's microbiomes. *Nat. Biotechnol.* **39**, 499–509 (2021).
75. J. P. Shaffer, L.-F. Nothias, L. R. Thompson, J. G. Sanders, R. A. Salido, S. P. Couvillion, A. D. Brejnrod, F. Lejzerowicz, N. Haiminen, S. Huang, H. L. Lutz, Q. Zhu, C. Martino, J. T. Morton, S. Karthikeyan, M. Nothias-Esposito, K. Dührkop, S. Böcker, H. W. Kim, A. A. Aksenov, W. Bittremieux, J. J. Minich, C. Marotz, M. M. Bryant, K. Sanders, T. Schwartz, G. Humphrey, Y. Vásquez-Baeza, A. Tripathi, L. Parida, A. P. Carrieri, K. L. Beck, P. Das, A. González, D. McDonald, J. Ladau, S. M. Karst, M. Albertsen, G. Ackermann, J. DeReus, T. Thomas, D. Petras, A. Shade, J. Stegen, S. J. Song, T. O. Metz, A. D. Swafford, P. C. Dorrestein, J. K. Jansson, J. A. Gilbert, R. Knight, the Earth Microbiome Project 500 (EMP500) Consortium, Standardized multi-omics of Earth's microbiomes reveals microbial and metabolite diversity. *Nat. Microbiol.* **7**, 2128–2150 (2022).
76. A. Vion, L. Menot, Continental margins between 140m and 3500m depth (IFREMER - COMARGE, 2009) (2009); www.marineregions.org/maps.php?album=3264&pic=64937).
77. D. E. LaRowe, S. Arndt, J. A. Bradley, E. Burwicz, A. W. Dale, J. P. Amend, Organic carbon and microbial activity in marine sediments on a global scale throughout the Quaternary. *Geochim. Cosmochim. Acta* **286**, 227–247 (2020).
78. J. A. Bradley, S. Arndt, J. P. Amend, E. Burwicz, A. W. Dale, M. Egger, D. E. LaRowe, Widespread energy limitation to life in global subseafloor sediments. *Sci. Adv.* **6**, eaba0697 (2020).
79. S. J. Salter, M. J. Cox, E. M. Turek, S. T. Calus, W. O. Cookson, M. F. Moffatt, P. Turner, J. Parkhill, N. J. Loman, A. W. Walker, Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **12**, 87 (2014).
80. C. S. Sheik, B. K. Reese, K. I. Twing, J. B. Sylvan, S. L. Grim, M. O. Schrenk, M. L. Sogin, F. S. Colwell, Identification and removal of contaminant sequences from ribosomal gene databases: Lessons from the census of deep life. *Front. Microbiol.* **9**, 840 (2018).
81. K. Kennedy, M. W. Hall, M. D. J. Lynch, G. Moreno-Hagelsieb, J. D. Neufeld, Evaluating bias of Illumina-based bacterial 16S rRNA gene profiles. *Appl. Environ. Microbiol.* **80**, 5717–5722 (2014).
82. M. Schirmer, U. Z. Ijaz, R. D'Amore, N. Hall, W. T. Sloan, C. Quince, Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Res.* **43**, e37 (2015).
83. K. Piwosz, T. Shabarova, J. Pernthaler, T. Posch, K. Šimek, P. Porcal, M. M. Salcher, Bacterial and eukaryotic small-subunit amplicon data do not provide a quantitative picture of microbial communities, but they are reliable in the context of ecological interpretations. *mSphere* **5**, e00052-20 (2020).
84. R. E. Anderson, E. D. Graham, J. A. Huber, B. J. Tully, Microbial populations are shaped by dispersal and recombination in a low biomass subseafloor habitat. *MBio*, e0035422 (2022).
85. J. E. Hinkle, J. P. Chanton, M. A. Moynihan, S. E. Ruff, A. Teske, Complex bacterial diversity of Guaymas Basin hydrothermal sediments revealed by synthetic long-read sequencing (LoopSeq). bioRxiv 589229 [Preprint] (2024). https://doi:10.1101/2024.04.12.589229.
86. H. Mallick, A. Rahnavard, L. J. McIver, S. Ma, Y. Zhang, L. H. Nguyen, T. L. Tickle, G. Weingart, B. Ren, E. H. Schwager, S. Chatterjee, K. N. Thompson, J. E. Wilkinson, A. Subramanian, Y. Lu, L. Waldron, J. N. Paulson, E. A. Franzosa, H. C. Bravo, C. Huttenhower, Multivariable association discovery in population-scale meta-omics studies. *PLOS Comput. Biol.* **17**, e1009442 (2021).
87. G. A. Coleman, A. A. Davín, T. A. Mahendrarajah, L. L. Szánthó, A. Spang, P. Hugenholtz, G. J. Szöllősi, T. A. Williams, A rooted phylogeny resolves early bacterial evolution. *Science* **372**, eabe0511 (2021).
88. J. T. Staley, A. Konopka, Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annu. Rev. Microbiol.* **39**, 321–346 (1985).
89. M. L. Sogin, H. G. Morrison, J. A. Huber, D. M. Welch, S. M. Huse, P. R. Neal, J. M. Arrieta, G. J. Herndl, Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc. Natl. Acad. Sci. U.S.A.* **103**, 12115–12120 (2006).
90. K. G. Lloyd, A. D. Steen, J. Ladau, J. Yin, L. Crosby, Phylogenetically novel uncultured microbial cells dominate earth microbiomes. *mSystems* **3**, e00055-18 (2018).
91. A. Vuillemin, S. D. Wankel, Ö. K. Coskun, T. Magritsch, S. Vargas, E. R. Estes, A. J. Spivack, D. C. Smith, R. Pockalny, R. W. Murray, S. D'Hondt, W. D. Orsi, Archaea dominate oxic subseafloor communities over multimillion-year time scales. *Sci. Adv.* **5**, eaaw4108 (2019).
92. A. J. Probst, B. Ladd, J. K. Jarett, D. E. Geller-McGrath, C. M. K. Sieber, J. B. Emerson, K. Anantharaman, B. C. Thomas, R. R. Malmstrom, M. Stieglmeier, A. Klingl, T. Woyke,

M. C. Ryan, J. F. Banfield, Differential depth distribution of microbial function and putative symbionts through sediment-hosted aquifers in the deep terrestrial subsurface. *Nat. Microbiol.* **3**, 328–336 (2018).
93. C. He, R. Keren, M. L. Whittaker, I. F. Farag, J. A. Doudna, J. H. D. Cate, J. F. Banfield, Genome-resolved metagenomics reveals site-specific diversity of episymbiotic CPR bacteria and DPANN archaea in groundwater ecosystems. *Nat. Microbiol.* **6**, 354–365 (2021).
94. D. Chivian, E. L. Brodie, E. J. Alm, D. E. Culley, P. S. Dehal, T. Z. DeSantis, T. M. Gihring, A. Lapidus, L.-H. Lin, S. R. Lowry, D. P. Moser, P. M. Richardson, G. Southam, G. Wanger, L. M. Pratt, G. L. Andersen, T. C. Hazen, F. J. Brockman, A. P. Arkin, T. C. Onstott, Environmental genomics reveals a single-species Ecosystem deep within earth. *Science.* **322**, 275–278 (2008).
95. J. M. Labonté, E. K. Field, M. Lau, D. Chivian, E. Van Heerden, K. E. Wommack, T. L. Kieft, T. C. Onstott, R. Stepanauskas, Single cell genomics indicates horizontal gene transfer and viral infections in a deep subsurface Firmicutes population. *Front. Microbiol.* **6**, 349 (2015).
96. O. V. Karnachuk, Y. A. Frank, A. P. Lukina, V. V. Kadnikov, A. V. Beletsky, A. V. Mardanov, N. V. Ravin, Domestication of previously uncultivated Candidatus Desulforudis audaxviator from a deep aquifer in Siberia sheds light on its physiology and evolution. *ISME J.* **13**, 1947–1959 (2019).
97. O. V. Karnachuk, A. P. Lukina, V. V. Kadnikov, V. A. Sherbakova, A. V. Beletsky, A. V. Mardanov, N. V. Ravin, Targeted isolation based on metagenome-assembled genomes reveals a phylogenetically distinct group of thermophilic spirochetes from deep biosphere. *Environ. Microbiol.* **23**, 3585–3598 (2021).
98. S. A. Carr, B. N. Orcutt, K. W. Mandernack, J. R. Spear, Abundant Atribacteria in deep marine sediment from the Adélie Basin, Antarctica. *Front. Microbiol.* **6**, 872 (2015).
99. Y. M. Lee, K. Hwang, J. Il Lee, M. Kim, C. Y. Hwang, H.-J. Noh, H. Choi, H. K. Lee, J. Chun, S. G. Hong, S. C. Shin, Genomic insight into the predominance of candidate phylum Atribacteria JS1 lineage in marine sediments. *Front. Microbiol.* **9**, 2909 (2018).
100. A. Vuillemin, D. Ariztegui, F. Horn, J. Kallmeyer, W. D. Orsi, PASADO Science Team, Microbial community composition along a 50 000-year lacustrine sediment sequence. *FEMS Microbiol. Ecol.* **94**, fiy029 (2018).
101. A. Vuillemin, S. Vargas, Ö. K. Coskun, R. Pockalny, R. W. Murray, D. C. Smith, S. D'hondt, W. D. Orsi, Atribacteria reproducing over millions of years in the atlantic abyssal subseafloor. *MBio* **11**, e01937-20 (2020).
102. T. Katayama, M. K. Nobu, H. Kusada, X. Meng, N. Hosogi, K. Uematsu, H. Yoshioka, Y. Kamagata, H. Tamaki, Isolation of a member of the candidate phylum 'Atribacteria' reveals a unique cell membrane structure. *Nat. Commun.* **11**, 6381 (2020).
103. R. Z. Moger-Reischer, J. I. Glass, K. S. Wise, L. Sun, D. M. C. Bittencourt, B. K. Lehmkuhl, D. R. Schoolmaster, M. Lynch, J. T. Lennon, Evolution of a minimal cell. *Nature.* **620**, 122–127 (2023).
104. J. B. Glass, P. Ranjan, C. B. Kretz, B. L. Nunn, A. M. Johnson, M. Xu, J. McManus, F. J. Stewart, Microbial metabolism and adaptations in Atribacteria -dominated methane hydrate sediments. *Environ. Microbiol.* **23**, 4646–4660 (2021).
105. K. Kubo, K. G. Lloyd, J. F. Biddle, R. Amann, A. Teske, K. Knittel, Archaea of the Miscellaneous Crenarchaeotal Group are abundant, diverse and widespread in marine sediments. *ISME J.* **6**, 1949–1965 (2012).
106. J. F. Biddle, S. Fitz-Gibbon, S. C. Schuster, J. E. Brenchley, C. H. House, Metagenomic signatures of the Peru Margin subseafloor biosphere show a genetically distinct environment. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 10583–10588 (2008).
107. A. Dutta, S. Dutta Gupta, A. Gupta, J. Sarkar, S. Roy, A. Mukherjee, P. Sar, Exploration of deep terrestrial subsurface microbiome in late Cretaceous Deccan traps and underlying Archean basement, India. *Sci. Rep.* **8**, 17459 (2018).
108. L. Purkamo, R. Kietäväinen, M. Nuppunen-Puputti, M. Bomberg, C. Cousins, Ultradeep microbial communities at 4.4 km within crystalline bedrock: Implications for habitability in a planetary context. *Life.* **10**, 2 (2020).
109. C. Rinke, P. Schwientek, A. Sczyrba, N. N. Ivanova, I. J. Anderson, J.-F. Cheng, A. E. Darling, S. Malfatti, B. K. Swan, E. A. Gies, J. A. Dodsworth, B. P. Hedlund, G. Tsiamis, S. M. Sievert, W.-T. Liu, J. A. Eisen, S. J. Hallam, N. C. Kyrpides, R. Stepanauskas, E. M. Rubin, P. Hugenholtz, T. Woyke, Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**, 431–437 (2013).
110. P. Yarza, P. Yilmaz, E. Pruesse, F. O. Glöckner, W. Ludwig, K.-H. Schleifer, W. B. Whitman, J. Euzéby, R. Amann, R. Rosselló-Móra, Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat. Rev. Microbiol.* **12**, 635–645 (2014).
111. B. J. Baker, K. E. Appler, X. Gong, New microbial biodiversity in marine sediments. *Ann. Rev. Mar. Sci.* **13**, 161–175 (2021).
112. B. J. Tully, C. G. Wheat, B. T. Glazer, J. A. Huber, A dynamic microbial community with high functional redundancy inhabits the cold, oxic subseafloor aquifer. *ISME J.* **12**, 1–16 (2018).
113. K. J. Edwards, C. G. Wheat, J. B. Sylvan, Under the sea: Microbial life in volcanic oceanic crust. *Nat. Rev. Microbiol.* **9**, 703–712 (2011).
114. S. E. Ruff, in *Marine Hydrocarbon Seeps - Microbiology and Biogeochemistry of a Global Marine Habitat*, A. P. Teske, V. Carvalho, Eds. (Springer, 2020), pp. 1–19; http://link.springer.com/10.1007/978-3-030-34827-4_1).
115. M. Crespo-Medina, M. W. Bowles, V. A. Samarkin, K. S. Hunter, S. B. Joye, Microbial diversity and activity in seafloor brine lake sediments (Alaminos Canyon block 601, Gulf of Mexico). *Geobiology* **14**, 483–498 (2016).
116. N. Le Bris, M. Yücel, A. Das, S. M. Sievert, P. LokaBharathi, P. R. Girguis, Hydrothermal energy transfer and organic carbon production at the deep seafloor. *Front. Mar. Sci.* **5**, 531 (2019).
117. P. J. Mouser, M. Borton, T. H. Darrah, A. Hartsock, K. C. Wrighton, Hydraulic fracturing offers view of microbial life in the deep terrestrial subsurface. *FEMS Microbiol. Ecol.* **92**, fiw166 (2016).
118. O. Plümper, H. E. King, T. Geisler, Y. Liu, S. Pabst, I. P. Savov, D. Rost, T. Zack, Subduction zone forearc serpentinites as incubators for deep microbial life. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 4324–4329 (2017).
119. W. J. Brazelton, B. Nelson, M. O. Schrenk, Metagenomic evidence for H2 oxidation and H2 production by serpentinite-hosted subsurface microbial communities. *Front. Microbiol.* **2**, 268 (2012).
120. S. Motamedi, B. N. Orcutt, G. L. Früh-Green, K. I. Twing, H. L. Pendleton, W. J. Brazelton, Microbial residents of the Atlantis Massif's shallow serpentinite subsurface. *Appl. Environ. Microbiol.* **86**, e00356-20 (2020).
121. J. L. Macalady, E. H. Lyon, B. Koffman, L. K. Albertson, K. Meyer, S. Galdenzi, S. Mariani, Dominant microbial populations in limestone-corroding stream biofilms, Frasassi cave system, Italy. *Appl. Environ. Microbiol.* **72**, 5596–5609 (2006).
122. F. Inagaki, K. U. Hinrichs, Y. Kubo, M. W. Bowles, V. B. Heuer, W. L. Hong, T. Hoshino, A. Ijiri, H. Imachi, M. Ito, M. Kaneko, M. A. Lever, Y. S. Lin, B. A. Methé, S. Morita, Y. Morono, W. Tanikawa, M. Bihan, S. A. Bowden, M. Elvert, C. Glombitza, D. Gross, G. J. Harrington, T. Hori, K. Li, D. Limmer, C. H. Liu, M. Murayama, N. Ohkouchi, S. Ono, Y. S. Park, S. C. Phillips, X. Prieto-Mollar, M. Purkey, N. Riedinger, Y. Sanada, J. Sauvage, G. Snyder, R. Susilawati, Y. Takano, E. Tasumi, T. Terada, H. Tomaru, E. Trembath-Reichert, D. T. Wang, Y. Yamada, Exploring deep microbial life in coal-bearing sediment down to ~2.5 km below the ocean floor. *Science.* **349**, 420–424 (2015).
123. M. C. Y. Lau, T. L. Kieft, O. Kuloyo, B. Linage-Alvarez, E. van Heerden, M. R. Lindsay, C. Magnabosco, W. Wang, J. B. Wiggins, L. Guo, D. H. Perlman, S. Kyin, H. H. Shwe, R. L. Harris, Y. Oh, M. J. Yi, R. Purtschert, G. F. Slater, S. Ono, S. Wei, L. Li, B. Sherwood Lollar, T. C. Onstott, An oligotrophic deep-subsurface community dependent on syntrophy is dominated by sulfur-driven autotrophic denitrifiers. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E7927–E7936 (2016).
124. L.-H. Lin, G. F. Slater, B. Sherwood Lollar, G. Lacrampe-Couloume, T. C. Onstott, The yield and isotopic composition of radiolytic H_2, a potential energy source for the deep subsurface biosphere. *Geochim. Cosmochim. Acta* **69**, 893–903 (2005).
125. K. Pedersen, Exploration of deep intraterrestrial microbial life: Current perspectives. *FEMS Microbiol. Lett.* **185**, 9–16 (2000).
126. F. S. Colwell, R. P. Smith, in *Geophysical Monograph Series* (Wiley, 2004), vol. 144, pp. 355–367; https://agupubs.onlinelibrary.wiley.com/doi/10.1029/144GM22.
127. V. Rebata-Landa, J. C. Santamarina, Mechanical limits to microbial activity in deep sediments. *Geochem. Geophys. Geosyst.* **7**, Q11006 (2006).
128. S. D'Hondt, F. Inagaki, C. A. Zarikian, L. J. Abrams, N. Dubois, T. Engelhardt, H. Evans, T. Ferdelman, B. Gribsholt, R. N. Harris, B. W. Hoppie, J.-H. Hyun, J. Kallmeyer, J. Kim, J. E. Lynch, C. C. McKinley, S. Mitsunobu, Y. Morono, R. W. Murray, R. Pockalny, J. Sauvage, T. Shimono, F. Shiraishi, D. C. Smith, C. E. Smith-Duque, A. J. Spivack, B. O. Steinsbu, Y. Suzuki, M. Szpak, L. Toffin, G. Uramoto, Y. T. Yamaguchi, G. Zhang, X.-H. Zhang, W. Ziebis, Presence of oxygen and aerobic communities from sea floor to basement in deep-sea sediments. *Nat. Geosci.* **8**, 299–304 (2015).
129. S. E. Ruff, L. Schwab, E. Vidal, J. Hemingway, B. Kraft, R. Murali, Widespread occurrence of dissolved oxygen anomalies, aerobic microbes, and oxygen-producing metabolic pathways in apparently anoxic environments. *FEMS Microbiol. Ecol.* **100**, fiae132 (2024).
130. S. L. Brantley, M. B. Goldhaber, K. V. Ragnarsdottir, Crossing disciplines and scales to understand the critical zone. *Elements* **3**, 307–314 (2007).
131. K. Küsel, K. U. Totsche, S. E. Trumbore, R. Lehmann, C. Steinhäuser, M. Herrmann, How deep can surface signals be traced in the critical zone? Merging biodiversity with biogeochemistry research in a central German Muschelkalk landscape. *Front. Earth Sci.* **4**, 32 (2016).
132. J. H. Steele, K. H. Brink, B. E. Scott, Comparison of marine and terrestrial ecosystems: Suggestions of an evolutionary perspective influenced by environmental variation. *ICES J. Mar. Sci.* **76**, 50–59 (2019).
133. M. H. Carr, J. E. Neigel, J. A. Estes, S. Andelman, R. R. Warner, J. L. Largier, Comparing marine and terrestrial ecosystems: Implications for the design of coastal marine reserves. *Ecol. Appl.* **13**, 90–107 (2003).
134. Y. Zhang, R. N. Horne, A. J. Hawkins, J. C. Primo, O. Gorbatenko, A. E. Dekas, Geological activity shapes the microbiome in deep-subsurface aquifers by advection. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2113985119 (2022).

135. T. L. V. Bornemann, P. S. Adam, V. Turzynski, U. Schreiber, P. A. Figueroa-Gonzalez, J. Rahlff, D. Köster, T. C. Schmidt, R. Schunk, B. Krauthausen, A. J. Probst, Genetic diversity in terrestrial subsurface ecosystems impacted by geological degassing. *Nat. Commun.* **13**, 284 (2022).
136. T. Gleeson, N. Moosdorf, J. Hartmann, L. P. H. van Beek, A glimpse beneath earth's surface: GLobal HYdrogeology MaPS (GLHYMPS) of permeability and porosity. *Geophys. Res. Lett.* **41**, 3891–3898 (2014).
137. K. J. Edwards, A. T. Fisher, C. G. Wheat, The deep subsurface biosphere in igneous ocean crust: Frontier habitats for microbiological exploration. *Front. Microbiol.* **3**, 8 (2012).
138. M. O. Schrenk, J. A. Huber, K. J. Edwards, Microbial provinces in the subseafloor. *Ann. Rev. Mar. Sci.* **2**, 279–304 (2010).
139. A. Longhurst, S. Sathyendranath, T. Platt, C. Caverhill, An estimate of global primary production in the ocean from satellite radiometer data. *J. Plankton Res.* **17**, 1245–1271 (1995).
140. G. Galla, N. Praeg, T. Rzehak, E. Sprecher, F. Colla, J. Seeber, P. Illmer, H. C. Hauffe, Comparison of DNA extraction methods on different sample matrices within the same terrestrial ecosystem. *Sci. Rep.* **14**, 8715 (2024).
141. M. C. Nelson, H. G. Morrison, J. Benjamino, S. L. Grim, J. Graf, Analysis, optimization and verification of illumina-generated 16S rRNA gene amplicon surveys. *PLOS ONE* **9**, e94249 (2014).
142. B. D. Topçuoğlu, L. C. Stewart, H. G. Morrison, D. A. Butterfield, J. A. Huber, J. F. Holden, Hydrogen limitation and syntrophic growth among natural assemblages of thermophilic methanogens at deep-sea hydrothermal vents. *Front. Microbiol.* **7**, 1240 (2016).
143. B. J. Callahan, P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, S. P. Holmes, DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods.* **13**, 581–583 (2016).
144. C. Quast, E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, F. O. Glöckner, The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
145. J. R. Bray, J. T. Curtis, An ordination of the upland forest communities of Southern Wisconsin. *Ecol. Monogr.* **27**, 325–349 (1957).
146. J. B. Kruskal, Nonmetric multidimensional scaling: A numerical method. *Psychometrika* **29**, 115–129 (1964).
147. J. Oksanen, F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, H. Wagner, vegan: Community Ecology Package (2012); http://cran.r-project.org/package=vegan).
148. S. Nurk, D. Meleshko, A. Korobeynikov, P. A. Pevzner, MetaSPAdes: A new versatile metagenomic assembler. *Genome Res.* **27**, 824–834 (2017).
149. D. Hyatt, G.-L. Chen, P. F. LoCascio, M. L. Land, F. W. Larimer, L. J. Hauser, Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119 (2010).
150. T. Aramaki, R. Blanc-Mathieu, H. Endo, K. Ohkubo, M. Kanehisa, S. Goto, H. Ogata, KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* **36**, 2251–2252 (2020).
151. A. J. Probst, F. J. Elling, C. J. Castelle, Q. Zhu, M. Elvert, G. Birarda, H.-Y. N. Holman, K. R. Lane, B. Ladd, M. C. Ryan, T. Woyke, K.-U. Hinrichs, J. F. Banfield, Lipid analysis of CO_2-rich subsurface aquifers suggests an autotrophy-based deep biosphere with lysolipids enriched in CPR bacteria. *ISME J.* **14**, 1547–1560 (2020).
152. S. R. Eddy, Accelerated profile HMM searches. *PLOS Comput. Biol.* **7**, e1002195 (2011).
153. P. Menzel, K. L. Ng, A. Krogh, Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* **7**, 11257 (2016).
154. G. P. Wagner, K. Kin, V. J. Lynch, Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.* **131**, 281–285 (2012).
155. H. R. Gruber-Vodicka, B. K. B. Seah, E. Pruesse, PhyloFlash: Rapid small-subunit rRNA profiling and targeted assembly from metagenomes. *mSystems* **5**, e00920-20 (2020).
156. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag, 2009); https://ggplot2.tidyverse.org.
157. K. R. Clarke, Non-parametric multivariate analyses of changes in community structure. *Austral. J. Ecol.* **18**, 117–143 (1993).
158. J. J. Marlow, I. Colocci, S. P. Jungbluth, N. M. Weber, A. Gartman, J. Kallmeyer, Mapping metabolic activity at single cell resolution in intact volcanic fumarole sediment. *FEMS Microbiol. Lett.* **367**, fnaa031 (2020).
159. S. Crognale, S. Venturi, F. Tassi, S. Rossetti, H. Rashed, J. Cabassi, F. Capecchiacci, B. Nisi, O. Vaselli, H. G. Morrison, M. L. Sogin, S. Fazi, Microbiome profiling in extremely acidic soils affected by hydrothermal fluids: The case of the Solfatara Crater (Campi Flegrei, southern Italy). *FEMS Microbiol. Ecol.* **94**, fiy190 (2018).
160. A. Dutta, P. Sar, J. Sarkar, S. Dutta Gupta, A. Gupta, H. Bose, A. Mukherjee, S. Roy, Archaeal communities in deep terrestrial subsurface underneath the Deccan Traps, India. *Front. Microbiol.* **10**, 1362 (2019).
161. D. V. Bekaert, E. Gazel, S. Turner, M. D. Behn, J. M. de Moor, S. Zahirovic, V. C. Manea, K. Hoernle, T. P. Fischer, A. Hammerstrom, A. M. Seltzer, J. T. Kulongoski, B. S. Patel, M. O. Schrenk, S. A. Halldórsson, M. Nakagawa, C. J. Ramírez, J. A. Krantz, M. Yücel, C. J. Ballentine, D. Giovannelli, K. G. Lloyd, P. H. Barry, High $^3He/^4He$ in central Panama reveals a distal connection to the Galápagos plume. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2110997118 (2021).
162. C. Magnabosco, P. H. A. Timmers, M. C. Y. Lau, G. Borgonie, B. Linage-Alvarez, O. Kuloyo, R. Alleva, T. L. Kieft, G. F. Slater, E. van Heerden, B. Sherwood Lollar, T. C. Onstott, Fluctuations in populations of subsurface methane oxidizers in coordination with changes in electron acceptor availability. *FEMS Microbiol. Ecol.* **94**, fiy089 (2018).
163. G. Etiope, I. Vadillo, M. J. Whiticar, J. M. Marques, P. M. Carreira, I. Tiago, J. Benavente, P. Jiménez, B. Urresti, Abiotic methane seepage in the Ronda peridotite massif, southern Spain. *Appl. Geochem.* **66**, 101–113 (2016).
164. D. R. Meyer-Dombard, M. R. Osburn, D. Cardace, C. A. Arcilla, The effect of a tropical climate on available nutrient resources to springs in ophiolite-hosted, deep biosphere ecosystems in the Philippines. *Front. Microbiol.* **10**, 761 (2019).
165. D. Cardace, D. R. Meyer-Dombard, K. M. Woycheese, C. A. Arcilla, Feasible metabolisms in high pH springs of the Philippines. *Front. Microbiol.* **6**, 10 (2015).
166. L. L. Vigni, K. Daskalopoulou, S. Calabrese, K. Kyriakopoulos, F. Parello, F. Brugnone, W. D'Alessandro, Geochemical characterisation of the thermo-mineral waters of Greece. *Environ. Geochem. Health* **44**, 2111–2133 (2022).
167. K. Daskalopoulou, S. Calabrese, F. Grassa, K. Kyriakopoulos, F. Parello, F. Tassi, W. D'Alessandro, Origin of methane and light hydrocarbons in natural fluid emissions: A key study from Greece. *Chem. Geol.* **479**, 286–301 (2018).
168. M. Bomberg, T. Lamminmäki, M. Itävaara, Microbial communities and their predicted metabolic characteristics in deep fracture groundwaters of the crystalline bedrock at Olkiluoto, Finland. *Biogeosciences* **13**, 6031–6047 (2016).
169. R. A. Edwards, B. Rodriguez-Brito, L. Wegley, M. Haynes, M. Breitbart, D. M. Peterson, M. O. Saar, S. Alexander, E. C. Alexander, F. Rohwer, Using pyrosequencing to shed light on deep mine microbial ecology. *BMC Genomics* **7**, 57 (2006).
170. C. Magnabosco, M. Tekere, M. C. Y. Lau, B. Linage, O. Kuloyo, M. Erasmus, E. Cason, E. van Heerden, G. Borgonie, T. L. Kieft, J. Olivier, T. C. Onstott, Comparisons of the composition and biogeographic distribution of the bacterial communities occupying South African thermal springs with those inhabiting deep subsurface fracture water. *Front. Microbiol.* **5**, 679 (2014).
171. H. J. Lavalleur, F. S. Colwell, Microbial characterization of basalt formation waters targeted for geological carbon sequestration. *FEMS Microbiol. Ecol.* **85**, 62–73 (2013).
172. A. A. P. Koppers, T. Yamazaki, J. Geldmacher, IODP Expedition 330: Drilling the louisville Seamount trail in the SW pacific. *Sci. Drill.* **15**, 11–22 (2013).
173. J. B. Sylvan, S. D. Wankel, D. E. LaRowe, C. N. Charoenpong, J. A. Huber, C. L. Moyer, K. J. Edwards, Evidence for microbial mediation of subseafloor nitrogen redox processes at Loihi Seamount, Hawaii. *Geochim. Cosmochim. Acta* **198**, 131–150 (2017).
174. L. A. Zinke, C. Glombitza, J. T. Bird, H. Røy, B. B. Jørgensen, K. G. Lloyd, J. P. Amend, B. K. Reese, Microbial organic matter degradation potential in Baltic Sea sediments is influenced by depositional conditions and *in situ* geochemistry. *Appl. Environ. Microbiol.* **85**, e02164-18 (2019).
175. L. A. Zinke, B. K. Reese, J. McManus, C. G. Wheat, B. N. Orcutt, J. P. Amend, Sediment microbial communities influenced by cool hydrothermal Fluid migration. *Front. Microbiol.* **9**, 1249 (2018).
176. J. Goordial, T. D'Angelo, J. M. Labonté, N. J. Poulton, J. M. Brown, R. Stepanauskas, G. L. Früh-Green, B. N. Orcutt, Microbial diversity and function in shallow subsurface sediment and oceanic lithosphere of the Atlantis Massif. *mBio* **12**, e0049021 (2021).
177. A. Teske, L. J. McKay, A. C. Ravelo, I. Aiello, C. Mortera, F. Núñez-Useche, C. Canet, J. P. Chanton, B. Brunner, C. Hensen, G. A. Ramírez, R. J. Sibert, T. Turner, D. White, C. R. Chambers, A. Buckley, S. B. Joye, S. A. Soule, D. Lizarralde, Characteristics and evolution of sill-driven off-axis hydrothermalism in Guaymas Basin—The Ringvent site. *Sci. Rep.* **9**, 13847 (2019).
178. L. D. Brandt, C. H. House, Marine subsurface microbial community shifts across a hydrothermal gradient in Okinawa trough sediments. *Archaea* **2016**, 2690329 (2016).
179. F. Dowell, Z. Cardman, S. Dasarathy, M. Kellermann, J. S. Lipp, S. E. Ruff, J. F. Biddle, L. McKay, B. J. MacGregor, K. G. Lloyd, D. B. Albert, H. Mendlovitz, K.-U. Hinrichs, A. Teske, Microbial communities in methane- and short chain alkane-rich hydrothermal sediments of Guaymas Basin. *Front. Microbiol.* **7**, 17 (2016).
180. A. Lecoeuvre, B. Ménez, M. Cannat, V. Chavagnac, E. Gérard, Microbial ecology of the newly discovered serpentinite-hosted Old City hydrothermal field (southwest Indian ridge). *ISME J.* **15**, 818–832 (2021).
181. S. Kleindienst, S. Grim, M. Sogin, A. Bracco, M. Crespo-Medina, S. B. Joye, Diverse, rare microbial taxa responded to the *Deepwater Horizon* deep-sea hydrocarbon plume. *ISME J.* **10**, 400–415 (2016).
182. S. A. Klasek, M. E. Torres, M. Loher, G. Bohrmann, T. Pape, F. S. Colwell, Deep-sourced fluids from a convergent margin host distinct Subseafloor microbial communities that change upon mud flow expulsion. *Front. Microbiol.* **10**, 1436 (2019).
183. A. Angermeyer, S. C. Crosby, J. A. Huber, Decoupled distance-decay patterns between dsrA and 16S rRNA genes among salt marsh sulfate-reducing bacteria. *Environ. Microbiol.* **18**, 75–86 (2016).

184. E. Trembath-Reichert, D. A. Butterfield, J. A. Huber, Active subseafloor microbial communities from Mariana back-arc venting fluids share metabolic strategies across different thermal niches and taxa. *ISME J.* **13**, 2264–2279 (2019).
185. A. Angermeyer, S. C. Crosby, J. A. Huber, Salt marsh sediment bacterial communities maintain original population structure after transplantation across a latitudinal gradient. *PeerJ.* **6**, e4735 (2018).
186. A. Roguet, A. M. Eren, R. J. Newton, S. L. McLellan, Fecal source identification using random forest. *Microbiome.* **6**, 185 (2018).
187. R. J. Newton, S. M. Huse, H. G. Morrison, C. S. Peake, M. L. Sogin, S. L. McLellan, Shifts in the microbial community composition of Gulf coast beaches following beach oiling. *PLOS ONE* **8**, e74265 (2013).
188. S. Feng, M. Bootsma, S. L. McLellan, Human-associated lachnospiraceae genetic markers improve Detection of fecal pollution sources in urban waters. *Appl. Environ. Microbiol.* **84**, e00309-18 (2018).
189. J. A. Bradley, S. Arndt, J. P. Amend, E. Burwicz-Galerne, D. E. LaRowe, Sources and fluxes of organic carbon and energy to microorganisms in global marine sediments. *Front. Microbiol.* **13**, 910694 (2022).

Acknowledgments: We are indebted to all our colleagues that shared their datasets with us: W. Brazelton (U of Utah), Z. Cardon (Marine Biological Laboratory), N. Fernandez-Gonzalez (U de Valladolid), M. Saxton (Miami U), L. Maignien (U of Western Brittany), K. Meyer (U Willamette), B. Osborne (Brown U), J. Rich (U of Maine), A. Roguet, R. Newton, D. Dila and S. McLellan (U of Wisconsin Milwaukee), I. Tiago (U of Coimbra), E. Trembath-Reichert (Arizona State U), F. Trigodet (U Chicago), and L. Amaral Zettler (Royal Netherlands Institute for Sea Research). Without including outgroup datasets generously provided by these researchers, the comparison between surface and subsurface biomes would not have been possible. We are also grateful to all colleagues participating in the CoDL, who organized field work and collected and processed samples, as well as the expeditions and programs that enabled the sample collection. For further details on specific CoDL projects, results, site descriptions, funding, and publications, refer to the Supplementary Materials and dataset S1. We wish to acknowledge the support of the Sloan Foundation and the Deep Carbon Observatory. Sequencing was performed at the Marine Biological Laboratory (Woods Hole, MA, United States), and we are grateful for the assistance of S. Huse, J. Vineis, and A. Voorhis. Among CoDL principal investigators and colleagues, we especially thank P. Adam, T. Bornemann, B. Briggs, S. D'Hondt, A. Eleish, E. Gaidos, A. Hoarfrost, F. Huang, T. Kieft, R. Leon-Zayas, K. Locey, M. Osburn, L. Purkamo, K. Rogers, and B. Sherwood Lollar for helpful discussions. Above all we want to thank our friend, colleague, and mentor J. Amend for being a continuous source of inspiration, an ally of early career scientists, and a driving force in geobiology and subsurface science. This study was conceived during the Deep Carbon Observatory funded "Origins and Movement of Life" workshop, which J. Amend co-organized and cohosted in 2016. **Funding:** This work was supported by the Simons Foundation (824763, to S.E.R.), the Human Frontier Science Program (RGEC34/2023, to S.E.R.), the Deep Carbon Observatory (Alfred P. Sloan Foundation G-2016-7206, to K.G.L.), the Office of Biological and Environmental Research of the US Department of Energy (DE-SC0020369, to K.G.L.), the National Science Foundation Division of Ocean Sciences (NSF OCE-2145434, to A.D.S.), the CNRS Chaires de Professeur Junior (to J.A.B.), the Agence Nationale de la Recherche (ANR23-CPJ1-0172-01, to J.A.B.), the Center for Dark Energy Biosphere Investigations (C-DEBI OCE-0939564, to J.A.H.), and the Deep Life I (2011-12-01, to J.A.H. and M.O.S.). This is C-DEBI contribution number 619. **Author contributions:** Conceptualization: S.E.R., A.J.P., M.L.S., J.L., and F.C. Methodology: S.E.R., I.H.d.A., K.G.L., C.S.S., A.D.S., J.A.B., J.A.H., A.J.P., H.G.M., J.L., and F.C. Investigation: S.E.R., I.H.d.A., M.M., J.P.P., C.M., K.G.L., C.S.S., A.D.S., A.S., A.M., B.K.R., J.A.B., C.L., M.O.S., S.B.J., H.G.M., J.L., and F.C. Visualization: S.E.R. and I.H.d.A. Supervision: S.E.R., B.K.R., H.G.M., and F.C. Writing—original draft: S.E.R., I.H.d.A., K.G.L., and J.A.H. Writing—review and editing: S.E.R., I.H.d.A., K.G.L., C.S.S., A.D.S., J.A.B., M.O.S., J.A.H., A.J.P., M.L.S., J.L., and F.C. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** 16*S* rRNA gene amplicon datasets and metagenomic datasets are available under the NCBI umbrella Bioproject PRJNA183206, and the individual project accessions are listed in dataset S1. Contextual data are available in dataset S2. The DADA2-based workflow to process 16*S* rRNA gene amplicon sequences is provided in dataset S10. The R-based workflow to analyze marker gene data, VisuaR, is available on GitHub (https://github.com/EmilRuff/VisuaR) and datasets S11 and S12 (exemplary user input and workflow for archaeal ASV analysis grouped by "biome"). All other data needed to evaluate the conclusions in this paper are present in the paper and/or the Supplementary Materials.

Submitted 26 April 2024
Accepted 13 November 2024
Published 18 December 2024
10.1126/sciadv.adq0645