

Using meta-omics to unravel biogeochemical changes from cell to planetary scales

Elsa Abs^{1*}, Christoph Keuschnig², Pierre Amato³, Chris Bowler^{4,5}, Eric Capo⁶, Alexander B. Chase⁷, Luciana Chavez Rodriguez⁸, Abraham N. Dabengwa^{9,10}, Thomas Dussarrat¹¹, Thomas Guzman¹², Linnea K. Hernandez¹³, Jenni Hultman¹⁴, Kirsten Küsel^{15,16}, Zhen Li^{13,17}, Anna Mankowski¹⁸, William J. Riley¹⁷, Scott R. Saleska¹⁹, Lisa Wingate¹²

¹Laboratoire des Sciences du Climat et de l'Environnement, IPSL-LSCE, CEA/CNRS/UVSQ, Orme des Merisiers, 91191, Gif sur Yvette, France

²GFZ Helmholtz Centre for Geosciences, 14473 Potsdam, Germany

³Laboratoire Microorganismes: Génome et Environnement, UMR 6023 CNRS-Université Clermont Auvergne, 63178, Aubière, France

⁴Institut de Biologie de l'Ecole Normale Supérieure (IBENS), Ecole Normale Supérieure, CNRS, INSERM, Université PSL, 75005 Paris, France

⁵CNRS Research Federation for the study of Global Ocean Systems Ecology and Evolution, FR2022/Tara Oceans GOSEE, Paris, France

⁶Department of Ecology and Environmental Science, Umeå University, 90187, Umeå, Sweden

⁷Department of Earth Sciences. Southern Methodist University, Dallas, TX, USA

⁸Soil Biology Group, Wageningen University & Research, Wageningen, Gelderland, The Netherlands

⁹Evolutionary Studies Institute, University of the Witwatersrand, Private Bag Wits 2050, Johannesburg, South Africa

¹⁰Centre de Recherche et d'Enseignement des Géosciences de l'Environnement (CEREGE), CNRS UMR7330, Europole de l'Arbois, Cedex 04, Aix -en-Provence 13545, France

¹¹Department of Chemical Ecology, Bielefeld University, Universitätsstr. 25, 33615 Bielefeld, Germany.

¹²INRAE UMR1391 ISPA, 33140 Villenave D'Ornon, France

¹³Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory; Livermore, CA 94550, USA

¹⁴Natural Resources Institute Finland, Helsinki, Finland

¹⁵Institute of Biodiversity, Ecology, and Evolution, Friedrich Schiller University Jena, Jena, Germany

¹⁶Cluster of Excellence Balance of the Microverse, Friedrich Schiller University Jena, Germany

¹⁷Climate and Ecosystem Sciences Division, Lawrence Berkeley National Laboratory; Berkeley, CA 94720, USA

¹⁸Molecular Systems Biology Unit, European Molecular Biology Laboratory, Heidelberg, Germany

¹⁹Department of Ecology and Evolutionary Biology. University of Arizona, Tucson, AZ 85721, USA

Correspondence to: Elsa Abs (elsa.abs@lsce.ipsl.fr)

Abstract. Increased human impacts on Earth systems are radically altering biogeochemical cycles. While long-term environmental observatories and Earth System Models (ESMs) provide valuable insights into the mechanisms of nutrient dynamics, their performance is limited at the fine spatial scales controlled by the functional diversity of plant and microbial communities. This gap in our understanding concerning the roles of microbial diversity and plant-microbial interactions in decomposition and nutrient dynamics extends across many global ecosystems. Recent advances in meta-omics technologies, including metagenomics, metatranscriptomics, metaproteomics, and metabolomics, offer a wide array of tools for assessing

39 metabolic to genetic to evolutionary drivers of ecosystem functioning. Here, we explore the integration of meta-omics with
40 traditional ecological approaches to examine responses to global environmental changes. We present case studies from diverse
41 environments—soils, aquatic systems, clouds, and paleoarchives—demonstrating how meta-omics can unravel the roles of
42 microbial diversity, metabolic pathways, and trait distributions critical to understanding greenhouse gas fluxes, nutrient
43 cycling, and biogeochemistry. Although meta-omics is still beset with challenges including data heterogeneity arising from
44 wide-ranging methods, omics-derived traits, kinetic parameters, and machine learning tools can be used to enhance ESM
45 predictive capability. For example, emerging applications of meta-omics to ancient environmental DNA are extending our
46 capacity to link historical patterns with future projections, offering a long-term perspective on ecosystem dynamics. This
47 review highlights the potential of integrating omics with experimental manipulations alongside existing monitoring and
48 modelling efforts to refine predictions of ecosystem responses to natural and anthropogenic-driven environmental changes.
49 Because omics approaches cross a range of scientific domains, they could be used to foster collaboration and even integration
50 within existing models, thus laying the foundation for informed conservation and ecosystem management strategies from local
51 to global scales.

52

53 **1 Introduction**

54 The rapid increase in human populations over the last century has intensified pressure on the biosphere, hydrosphere, and
55 atmosphere (Rockström et al., 2023). Losses of natural habitats and over-exploitation of resources (e.g., farming, mining,
56 construction, and other industrial practices) have led to rapid shifts in the biogeochemical properties of our ecosystems (Lewis
57 and Maslin, 2015; Waters et al., 2016). These human-induced changes in global biogeochemical cycles alter energy fluxes that
58 contribute to our current climate and biodiversity crises. The resulting pollution and environmental degradation that have been
59 measured and communicated for decades by geoscientists and biologists necessitate continued biogeoscience research
60 internationally (reports of the Intergovernmental Panel on Climate Change).

61 To understand the impacts on biogeochemical processes, long-term observatory studies with large-scale ground-based or
62 ocean-based infrastructures—such as LTER (terrestrial and aquatic systems across the USA), NEON (diverse U.S.
63 ecosystems), ICOS (European forests, wetlands, and grasslands), ITEX (Arctic tundra), BATS and SPOT (Atlantic and Pacific
64 oceanic waters), and HOT (subtropical North Pacific Ocean)—were established to document the dynamic responses of
65 ecosystems to changes in climate, species invasions, and management impacts (Franz et al., 2018; Thibault et al., 2023;
66 Trowbridge et al., 2019). These local scale observatories are often equipped with instruments to measure the water, energy and
67 greenhouse gas (GHGs) fluxes (CO_2 , CH_4 , N_2O) at the interface between surface ecosystems (soil, plant canopies, water
68 bodies) and the atmosphere, providing valuable insights on the response of ecosystem processes, that underpin GHG budgets,

69 to short (extreme and seasonal weather systems) and long term changes in climate and management. When these local scale
70 measurements are coordinated across networks, the spatial and temporal variations in GHG budgets and ecological processes
71 can be assessed over larger scales and used to generate predictive models, which in turn can be confronted at large scale by
72 satellite and atmospheric station observations.

73 In addition to modern-day monitoring, the information preserved in sedimentary archives provides knowledge about past
74 environmental changes and their potential consequences on biodiversity (Gregory-Eaves and Smol, 2024). The use of organic
75 climate proxies (contained in tree rings, pollen, charcoal, corals, diatoms and foraminifera) and inorganic proxies (contained
76 in varves, stalagmites, sediments and ice cores) provide a direct link to the effects of environmental change on ecosystem
77 properties, such as changes in functional diversity, species abundances, and biogeographic distributions. For example, by
78 identifying traits linked to present-day climate niches, researchers can extrapolate to past and future scenarios, using geological
79 climate parallels to anticipate ecosystem responses (McElwain et al., 2024).

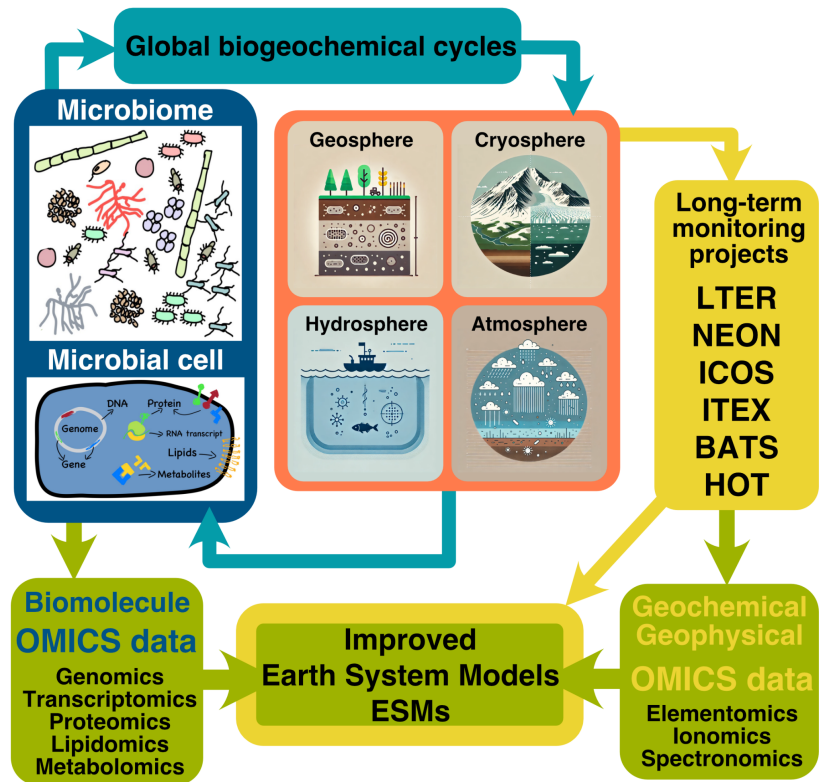
80 To date, most large-scale and long-term infrastructures have shown that representing a certain level of plant functional diversity
81 can improve the predictive accuracy of Earth System Models (ESMs), particularly in simulating interactions among carbon,
82 water, and nutrient cycles (Anderegg et al., 2021; Wullschleger et al., 2014). However, ESMs still face considerable limitations
83 in capturing the full complexity of biogeochemical processes and energy fluxes, especially those mediated by microbial
84 communities. These models often underrepresent key microbial functions that influence decomposition, nutrient cycling, and
85 GHG fluxes. While primary production by higher plants is relatively well parameterized in many land surface models, the
86 roles of microbial primary producers, heterotrophic decomposers, and organisms involved in secondary metabolism remain
87 poorly understood and largely absent from ESM formulations. This gap includes processes such as microbial biomass growth,
88 litter degradation, and the regulation of carbon use efficiency, which all influence the fate of organic matter and the balance
89 between carbon storage and release (He et al., 2024b, 2026). The ecophysiology of these microbial communities has a direct
90 feedback on the chemistry of the atmosphere and consequently climate (Conrad, 1994; Lovelock and Margulis, 1974; Monson
91 and Holland, 2001). Yet, integrating mechanistic representations of microbial abundance, diversity, and function into ESMs
92 remains incomplete, due to uncertainty about which processes are most critical, how to describe them mathematically, and
93 how to scale them appropriately across spatial and temporal dimensions.

94 Meta-omics data, which encompass the comprehensive analyses of genes, transcripts, proteins, and metabolites of
95 communities, provide powerful insights to bridge this knowledge gap. For example, molecular ecology techniques (e.g.,
96 metabarcoding, metagenomics, metatranscriptomics, metaproteomics, and metabonomics) used to study functional microbial
97 ecology provide new ways to quantify and identify GHG production hotspots with novel metrics based on key traits (Frostegård
98 et al., 2022). If implemented across networks that monitor large-scale and long-term changes in ecology and function (e.g.,
99 flux measurement sites) or through paleoecological archives, an omics framework could provide an important bridge between

100 disciplines (Fig. 1). For example, microbial ecologists could then help identify ‘keystone’ organisms that disproportionately
 101 impact ecosystem functioning, such as regulating GHG flux rates or those responding to climate variations.

102 In this perspective, we highlight the immense potential of omics data to advance our understanding of global biogeochemical
 103 cycles by showcasing key insights gained across diverse systems—from land to ocean to atmosphere—and addressing the
 104 remaining challenges and future opportunities.

105



106

107 **Figure 1: Conceptual framework for advancing our understanding and prediction of Earth’s biogeochemical cycles under**
 108 **environmental change, by integrating omics data into long-term monitoring and modelling efforts.** Red: Earth environmental spheres
 109 and ecosystems; blue: microbial cells and microbiomes interacting with biogeochemical processes critical in global biogeochemical cycles;
 110 yellow: long-term monitoring actions that supply key knowledge and response variables for biogeochemical models; green: integration of
 111 different levels of functional omics data to complement the environmental datasets used to advance Earth system models.

112

Glossary of omics vocabulary

Microbiome: the collective community of microorganisms (such as bacteria, archaea, fungi, viruses, and sometimes protists) that inhabit a specific environment.

Community: a group of populations found in a shared environment.

Population: a group of cells that have the same genome.

Molecular paleomicrobiology: The use of omics approaches (e.g., ancient DNA and metagenomics) to reconstruct past microbial communities and functions from environmental archives.

Amplicon sequencing

Amplicon sequencing: PCR amplification and sequencing of a portion of the genome found in multiple populations.

Amplicon Sequence Variant (ASV): similar to OTU - represents amplicon sequences where sequencing error has been taken into account and usually providing higher accuracy for taxonomic annotation.

Internal transcribed spacer (ITS): a sequence region between the small and large rRNA genes mostly used as a phylogenetic marker for fungi.

Metabarcoding: Method based on the high-throughput sequencing of DNA markers, typically specific gene or genomic regions, to simultaneously identify and characterize multiple species within complex environmental samples.

Operational taxonomic unit (OTU): a collection of amplicon sequences that are similar at a specific level of identity (i.e., 97 or 99 %) that is assumed to be phylogenetically closely related.

rRNA genes: ribosomal RNA coding genes - often used as a phylogenetic marker for prokaryotes (16S rRNA) and eukaryotes (18S rRNA).

Whole genome and metagenome sequencing

Assembler: bioinformatic software which creates contigs by various approaches.

Binning: grouping of contigs with similar features in a metagenome following the assumption they originate from the same genome.

Coding region: region of sequence/contig that encodes for a gene.

Contig: overlapping reads that have been assembled to a longer (contiguous) sequence. Resulting contigs are called assembly.

Gene: a segment of the region on a genome of sequence that encodes the information to produce a functional product such as proteins or RNA.

Genome: complete set of DNA from a single organism.

Genome sequencing: sequencing of the entire DNA of an organism.

Metagenome: the collection of genomes across multiple organisms in a given environment/sample.

Metagenome assembled genome (MAG): a genome that is constructed from a metagenome using sequence read assembly and binning.

Operon: cluster of functionally related genes found primarily in prokaryotes that are transcribed together. The encoded proteins often work in the same metabolic or regulatory pathway.

Protein: Amino acid sequence synthesized during translation of the RNA transcript by ribosomes, resulting in a functional molecule.

Read: a single sequence outputted by a DNA sequencer. This is the unit from which downstream analyses create assemblies from.

Transcript: RNA sequence produced from the transcription of a gene's DNA.

Meta-omics

Meta-omics: refer to the analysis of communities, as opposed to single cells or populations.

Metagenomics: method based on the sequencing of all DNA fragments present in a given DNA extract without the use of prior targeted approaches (e.g., PCR, hybridization capture). It is also referred to as shotgun sequencing or shotgun metagenomics.

Metatranscriptomics: method based on the reverse transcription of all RNA fragments and sequencing of complementary DNA (cDNA) from a given RNA extract without the use of prior targeted approaches.

Metaproteomics: method based on the direct interpretation of all proteins present in a sample, providing information on the active functions at the time of sampling.

Metabolomics: method based on the direct interpretation of all metabolites present in a sample, providing information regarding active functions and metabolic pathways at the time of sampling.

Metalipidomics: method based on the direct interpretation of all lipids present in a sample, providing information on i.e., cell membrane compositions, energy storage capacity and metabolic pathways of lipids.

Spectranomics: approach to link plant canopy species and their functional traits to their spectral-optical properties.

Volatilomics: subbranch of metabolomics, which studies the different types of volatile organic compounds that may be present in a certain biological matrix.

2 New insights on ecosystem dynamics and functional diversity

2.1 Global distribution of biodiversity

While around 1.5 million species have formally been catalogued to date, these estimates largely ignore microbes (both prokaryotes and eukaryotes, as well as viruses) that are estimated to encompass as many as a trillion species. The inability to characterize this dark matter of life is highly problematic, given microbes regulate and maintain all of Earth's biogeochemical processes including carbon (primary production in the ocean is performed almost entirely by microbes, and is comparable to all photosynthetic activity by plants on land), and nitrogen (e.g., nitrogen fixation and denitrification). Even viruses, generally not considered living organisms, are now recognized for playing major roles in nutrient recycling, e.g., carbon and iron cycles (Twining and Baines, 2013). Not only does much of the Earth's biodiversity reside in microorganisms, microbes are highly abundant (est. 10^{30} cells on Earth) and, together, represent the second largest contributor to Earth's biomass despite their

microscopic size (Bar-On et al., 2018; Bar-On and Milo, 2019). The immense amount of microbial abundance and diversity reflects billions of years of metabolic innovation and evolution to exploit nearly every resource on Earth. These key contributions of microbes as the engines of Earth's biogeochemical cycles and for assuring the survival of Earth's life are grossly neglected by most people, including by the majority of scientists, although efforts are now afoot to give a voice to this unseen majority (Doumeizel and Dolan, 2024; Falkowski and Knoll, 2007; Tara Ocean Foundation et al., 2022).

Thanks to multiple international efforts, a comprehensive picture of microbial diversity is emerging. Notable examples are the Earth Microbiome Project (Thompson et al., 2017), soil microbiome and fungal projects like SPUN (Sanchez-Tello and Corrales, 2024), and *Tara* Oceans (Sunagawa et al., 2020). These datasets allow us to build genomic catalogues (e.g., Ma et al., 2023; Maeke et al., 2025; Nayfach et al., 2021; Paoli et al., 2022; Schmidt et al., 2024) to characterize the patterns and processes governing microbial distributions. Gene atlases of these organisms could also be used to infer functions related to biogeochemical cycles and thus start to define who does what, how they interact with their environment, and how they impact ecosystem functioning (e.g., Guidi et al., 2016). As illustrated in the next section, such blueprints from multiple microbes provide multiple hypotheses to test the origins, current status, and future of life on Earth.

2.2 Using omics to link diversity to biogeochemical functions: a conceptual foundation

"We can say something about the community by giving a list of its species composition, but a community is poorly described by such a list alone." -Robert Whitaker (1975)

Ecosystem ecology sits at the nexus of understanding the flow of energy and materials through the biosphere and other abiotic pools. Many of the biological controls on these fluxes (e.g., reaction rates, substrate affinities, and temperature sensitivities) are governed by functional traits that vary at finer genetic scales than are typically resolved in ecosystem studies, such that aggregation into coarse taxonomic units obscures functionally relevant diversity and propagates uncertainty into predictions of biogeochemical fluxes. Ecosystem ecologists have focused on the role of species variation across environmental gradients to predict ecosystem functions; for example, correlating plant species diversity with an ecosystem's measure of productivity. However, for microorganisms, we have limited knowledge of how microbial diversity translates to functional variation. For one, most microorganisms are uncultured and, therefore, remain unknown and uncharacterized (Steen et al., 2019). Second, the equivalent of the biological species concept does not apply to microorganisms, as there are ambiguities in what constitutes a microbial population, or genetic units belonging to the same species (Chase et al., 2019; Shapiro et al., 2016). Therefore, microbiome analyses typically rely on extracting all the genetic information from an environmental sample for microbial classifications, as this circumvents the need to culture and identify individual microbes. Generally, microbiome scientists cluster and collapse the immense genetic diversity within microbiomes into taxonomic units based on genetic relatedness, such as operational taxonomic units (OTUs) or metagenome assembled genomes (MAGs), which can then be compared to reference databases. However, these taxonomic groupings are defined largely by computational convenience and may not reflect

154 biological organization, as they can represent anything from a family or species to a mosaic genome composed of multiple
155 strains (Meziti et al., 2021; Tyson et al., 2004). Therefore, microbiome-focused ecosystem assessments are inherently forced
156 to move away from the rich literature and theory developed for species to a broader view centered on community-wide metrics
157 such as alpha- and beta-diversity metrics, community-level traits, or interaction metrics, e.g., with co-occurrence networks
158 (Chase and Martiny, 2018).

159 This community-wide approach has several advantages, such as compiling global catalogs of microbial diversity across
160 geographic scales that reveal biogeographic processes structuring environmental microbiomes (Fierer, 2014) and
161 understanding potential metabolic interactions across various trophic levels encoded within the genetic makeup (Shaffer et al.,
162 2022). Yet, these broad approaches mask intra-community functional trait variation that contributes to ecosystem functioning
163 (Larkin and Martiny, 2017), especially in the context of environmental change (Scales et al., 2022). To illustrate this point, a
164 bacterial OTU is assigned by clustering related genetic sequences of the conserved 16S rRNA gene region at varying
165 thresholds. Depending on the genetic resolution used (ranging from amplicon sequence variants (ASVs) or 1-3% genetic
166 divergence), these OTUs are estimated to represent 50-150 MYA of evolutionary time, or roughly equivalent to when all
167 modern bird species evolved from a common ancestor (Martiny et al., 2023). Given the well-established trait and physiological
168 variation among bird species and their contributions to geographic distributions, no ornithologist would collapse all birds into
169 a single OTU to study ecosystem dynamics. From a biogeochemical perspective, doing so would implicitly assume uniform
170 kinetics and environmental sensitivities, an assumption rarely justified for processes such as decomposition, respiration, or
171 primary production. As such, microbial community analyses should understand the degree of trait variation within their
172 taxonomic designations and determine whether that functional variation is relevant to biogeochemical fluxes (Abs et al., 2023;
173 Defrenne et al., 2021).

174 Given that microbiome assessments rely on genetic material, we can also utilize that genomic information to infer functional
175 traits at multiple biological levels. All functional traits, to some degree, are phylogenetically conserved. Some, such as pH or
176 salinity responses, are broadly conserved at higher taxonomic levels while others, such as the responses to predicted
177 anthropogenic factors (i.e., added nitrogen or drought) (Chase et al., 2017), exhibit fine-scale genetic variation at the genus,
178 species, or even population levels (Martiny et al., 2015). Therefore, microbiome assessments of functional consequences need
179 to prioritize and identify their functional trait of interest and understand the system's diversity accordingly (McLaren and
180 Callahan, 2018). For example, the globally distributed marine phototroph *Prochlorococcus* (Cyanobacteriota) is estimated to
181 produce ~20% of the oxygen generated each year by photosynthesis on the planet and is predicted to be highly sensitive to
182 changes in surface ocean temperatures (Flombaum et al., 2013). *Prochlorococcus* is composed of several ecotypes, or highly
183 clustered strains occupying similar ecological niches but differing in key functional traits (Cohan, 2002). These ecotypes are
184 geographically and vertically distributed in the water column through their partitioning of resources, including light, iron, and
185 temperature, with closely related ecotypes differing in their thermal tolerance (ranging from ~15-30°C based on cultured

isolates) and persistence across surface and subsurface waters (Johnson et al., 2006). Therefore, an understanding of how oxygen production may shift under future climate scenarios will be dependent on the individual responses of each ecotype and their functional differences (Ustick et al., 2023). These patterns aren't unique to *Prochlorococcus* and have been increasingly observed in abundant soil taxa driving leaf litter degradation and carbon turnover (Chase et al., 2018), glacier-fed streams (Fodelianakis et al., 2022), marine sediments (Chase et al., 2021b), and bogs (García-García et al., 2019).

Given the current limitations in relating functional and taxonomic diversity, metagenomic (DNA-based) and/or metatranscriptomic (RNA-based) assessments of microbiomes can provide direct information on the functional genes present in a community. These functional approaches provide crucial insights into how the metabolic potential of communities may be used to predict responses to environmental change (Piton et al., 2023). These data reflect genomic potential rather than realized activity unless paired with proteomics or metabolomics; however, even with such paired datasets, sparse reference databases (e.g., 90-99% of environmental genes (Nayfach et al., 2016) and >95% of metabolites are unannotated (Bouslimani et al., 2014)) can bias trait interpretations and their links to environmental variables (Osburn et al., 2024). Ultimately, scaling from traits to ecosystem processes requires understanding how variation in microbial traits - and the genetic scales at which they are conserved - links to environmental gradients and biogeochemical rates (Abs et al., 2020, 2024a), as ecosystem ecologists have long done for macro-organisms. While physical state variables like temperature and pH are easily monitored, they represent the environment rather than the biological response to it. Meta-omics provides the link between these abiotic drivers and the internal biological “volume knobs”, such as V_{max} and K_m , that determine the actual rate of biogeochemical flux.

2.3 Using omics to predict biogeochemical functions at the site scale

2.3.1 Example from a subsurface system

We define the terrestrial subsurface to begin below the depth of greatest plant root density and extend into deeper aquifers. Life in the subsurface influences the composition of gases and liquids moving through it, acting as an open bioreactor driven by surface-derived organic matter, nutrient inputs, and a process known as dark carbon fixation (Küsel et al., 2016). In environments lacking sunlight, dark carbon fixation refers to the ability of chemolithoautotrophic microbes to use energy from reduced compounds (e.g., sulfur, nitrogen, hydrogen) to fix inorganic carbon into organic matter, thereby generating a self-sustaining energy source. This process is particularly significant in groundwater ecosystems, where reduced inorganic electron donors fuel primary production, replacing the role of photosynthesis at the surface.

Omics approaches, such as metagenomics, are invaluable for unravelling the metabolic capacities of these microbial communities, offering insights into the dynamic and complex biogeochemical processes that drive subsurface life. Metagenomic studies have revealed a high abundance of chemolithoautotrophs in groundwater, accounting for 12% to 47% of the microbial community, suggesting that dark carbon fixation is central to subsurface trophic webs (Overholt et al., 2022). This process might also have global significance, as the subsurface and surface biospheres are tightly linked through the

217 exchange of energy and matter. Overholt et al. (2022) estimated that dark carbon fixation in deep subsurface environments
218 could contribute between 1.4 and 2.5 Tg of CO₂ per year to the atmosphere globally.

Box 1: Example from air microbiology

The theory of spontaneous generation was disproved when atmospheric deposits were shown to inoculate environments with living microorganisms (Pasteur, 1861). Since then, it has been established that bacteria can remain suspended in the atmosphere for several days (Burrows et al., 2009) and be transported across continents (Griffin et al., 2017), enduring harsh conditions such as low water and nutrient availability, thermal and osmotic shocks, high UV exposure, and oxidants. Clouds are critical environments for complex multiphase chemical processes driven by light and oxidants, significantly impacting atmospheric composition and elemental cycles (Lelieveld and Crutzen, 1990). Additionally, they serve as relative hotspots of microbial activity, with biological processes contributing to chemical reactivity (Amato et al., 2019; Péguilhan et al., 2025). Clouds are also major components of the hydrological cycle, and play a key role in the deposition of airborne particles, including living microorganisms and genetic material, to surface ecosystems through precipitation (Péguilhan et al., 2021).

The low biomass in the atmosphere ($< 10^4$ cell m⁻³) and its high variability over short timescales pose significant challenges to study, but omics approaches offer decisive insights into a variety of scientific topics, including:

Early detection and monitoring of pathogens

The air contains a variety of potential pathogens affecting plants, animals, and humans, as well as microorganisms of sanitary concern (e.g., those harbouring antibiotic resistance genes). Omics-based methods offer rapid and accurate detection of these microorganisms at early stages, enabling powerful tools for epidemiological surveillance and outbreak prevention (Kuske et al., 2006).

Defining the role of biological processes in atmospheric chemistry

In the 2000s, clouds were identified as environments supporting microbial activity that interacts with atmospheric chemical processes. Laboratory experiments on natural cloud water incubated in bulk demonstrated active biological processes, where microorganisms metabolize dissolved organic compounds, produce biomass, and process dissolved chemicals (Amato et al., 2007). These findings suggested that microorganisms significantly influence atmospheric chemistry, which was further specified by modelling (Khaled et al., 2021; Nuñez López et al., 2024).

Defining the role of the atmosphere in microbial ecology

Microorganisms aerosolized from various surfaces enter the atmospheric environment, where they are exposed to extreme conditions. Some reach clouds and eventually return to ecosystems via precipitation. Certain microorganisms likely survive atmospheric processing better than others, enabling their long-distance dispersal. Investigating the biological functions involved in microbial survival—at the gene, transcript, protein, or metabolite level—can help us understand the selection pressures exerted by the atmosphere and its role as an environmental filter in microbial dispersion (Šantl-Temkiv et al., 2022).

Capturing environmental DNA (eDNA)

Airborne material from multiple surfaces carries genetic information, allowing the air to serve as a proxy for biodiversity within a given area. Airborne nucleic acids can be analysed through targeted or untargeted metagenomics to assess environmental biodiversity and track ecological changes (Lynggaard et al., 2022; Métris and Métris, 2023).

2.3.2 Example from the global ocean

Meta-omic data provide powerful tools for understanding the key role of microorganisms and in carbon sequestration, biogeochemical cycles of the global ocean including data generated by the GOS (Yooseph et al., 2010), GEOTRACES (Biller et al., 2018), Malaspina (Acinas et al., 2021) and Tara Oceans (Sunagawa et al., 2020) expeditions. Among other forcings, human-induced deoxygenation of marinel water columns is altering microbial metabolisms and their contributions to global biogeochemical cycles. Oxygen levels in marine waters are declining due to global warming and increased anthropogenic nutrient loading (Doney et al., 2012). This deoxygenation is expanding oxygen-minimum zones and creating so-called dead zones in near-shore ecosystems (Breitburg et al., 2018), with profound implications for ecosystem functioning. These changes reduce habitats and resources for organisms like fish and cause significant shifts in carbon, nitrogen, and sulfur fluxes between the sea and atmosphere (Wakeham, 2020). Future deoxygenation due to climate change may expand the extent of oxygen-minimum zones with unknown consequences such as potential increase of emissions of greenhouse gazes such as nitrous oxide (Canadell et al., 2021).

Microbial metabolisms are particularly affected, as anaerobic respiration processes become dominant in oxygen-depleted environments. These metabolisms include the use of nitrogen, sulfur, carbon compounds, and trace elements as alternative energy sources. Other processes, such as anoxygenic anaerobic photosynthesis (the use of light to produce organic matter in sunlit anoxic environments) and the microbial transformation of mercury into the neurotoxin methylmercury, are also induced (Bravo and Cosio, 2020). Thanks to meta-omic data, we can now study present-day microbial ecology and biogeochemical processes associated with oxygen depletion. For example, Capo et al. (2022b) demonstrated through environmental chemistry, metagenomics, and metatranscriptomics that microbial production of methylmercury is prevalent in oxygen-deficient water layers of two basins in the Baltic Sea. Combining such contemporary meta-omic analyses with historical DNA data from

240 sedimentary archives offers a unique opportunity to explore the links between increasing anoxia in water columns and shifts
241 in microbial metabolisms as recently done in the Black Sea (Zhong et al., 2025).

Box 2: Example from the cryosphere

Global warming is more pronounced in the Arctic regions than the rest of the globe, causing fundamental changes in ecosystem structure and functioning (Cavicchioli et al., 2019). Recent calculations showed that the Arctic is actually warming not two times but four times faster than the rest of the globe (Rantanen et al., 2022) and this warming has been observed to be stronger in the winter months (Post et al., 2019). Many omics studies, especially ones focusing on soil microbiomes, have been conducted in the summer period as the sites are accessible and measurements easier to conduct. In addition, safe transport of samples to the laboratory for analysis is simpler than in the winter seasons when sites are covered with up to meters of snow. Snow acts as an insulator and thus the soil may not freeze totally from the surface layers so that microbes may remain active and respire, producing both CO₂ and methane (CH₄) (Poppeliers et al., 2022; Walter Anthony et al., 2024).

Transition seasons with fluctuating conditions for microbes can also be of interest. In the springtime, the melt water brings nutrients for microbial metabolism. Besides, CH₄ and CO₂, produced by microbes, can be trapped in soil during winter and be released after thaw and for this reason these transition seasons should be studied in detail. In autumn, the length of the freeze-thaw events before the frozen season can vary depending on the timing of snowfall, resulting in an insulating snow cover which can fall on frozen or thawed soil—conditions that influence microbial processes during the winter months. With climate warming the autumns have changed in the Scandinavian subarctic while after plants have reduced photosynthesis with limited light sources, the weather stays warmer. Thus, in these seasons there is limited competition with the plants (Riley et al., 2018, 2021) for the resources and nutrients. Therefore, the need for arctic research is not only in the growing season but also during winter and transition seasons autumn and spring (Riley et al., 2021).

Omic approaches allow focus on metabolic function together with the taxonomic signatures of the microbial communities. Studies from the growing seasons have shown changes in microbial community composition and functions in sites with differing vegetation cover (e.g., Pessi et al., 2022; Woodcroft et al., 2018) and activity of the microbes has also been shown to vary (Viitamäki et al., 2022). Sub snow microbial communities in the arctic have shown diversity similar to summertime (Männistö et al., 2024). Intriguingly, the activity of functional genes has shown differences in CH₄, hydrogen and carbon metabolism between summer and winter (Viitamäki et al. in prep). In addition to soil, focus should also be in the snow as recent studies have demonstrated high microbial diversity and abundance of the same taxa as in soil in snowpacks (Keuschnig et al., 2023). Phospholipid fatty acid composition of tundra soils has shown temporal differences (Björk et al., 2008) and thus the possibilities of omics approaches in creating understanding on microbial processes and role during this long, forgotten season.

In addition to bacteria and archaea, viruses are abundant and diverse in the cryosphere (Heinrichs et al., 2024). These viruses control microbial communities through lytic infections as viral coverage can exceed cellular coverage from five up to eight-fold and therefore viruses indirectly affect biogeochemical cycling in (Calayag et al., 2025; Wang et al., 2022b). Recently metagenomic studies have also revealed widespread auxiliary metabolic genes (AMGs) for example in polar phages (Demina et al., 2025; Tian et al., 2024) and sea-ice viruses (Pettersson et al., 2025).

2.3.3 Example from a terrestrial plant system

Metabolomics are also a tool to better predict the influence of climate change on ecosystem dynamics. Chemical ecologists recently included phytochemical traits as niche dimensions based on their central functions in plant responses to their environment (Müller and Junker, 2022). For instance, the so-called “predictive metabolomics” approach can be used to predict patterns of phytochemical diversity in Alpine regions independently of plant lineage (Defossez et al., 2021). Similarly, modelling was used to predict the environment of multiple plant species from the Atacama Desert using their metabolism (Dussarrat et al., 2022). Metabolomics was also used to analyse changes in soil microbial function in response to drought (Brown et al., 2021). Overall, predictive metabolomics peaks as an ideal tool to predict and anticipate losses or modifications of chemical and functional diversity in response to climate change (Dussarrat et al., 2022, 2025; Frugone-Álvarez et al., 2023).

2.4 Using omics to predict biogeochemical functions at the global scale

2.4.1 Large-scale metagenomics

The examples in Sect. 2.2 illustrate how omics approaches can help identify the drivers of biogeochemical processes. Large-scale metagenomic analyses offer a powerful tool to extend findings from individual studies to a global scale. A variety of large-scale metagenomic datasets from diverse ecosystems are now available (Bahram et al., 2018; Duarte, 2015; Maeke et al., 2025; Schmidt et al., 2024; Sunagawa et al., 2020), allowing researchers to screen for the presence of pathways of interest at genome resolution thanks to the large numbers of MAGs that have been generated. These resources also enable us to explore global distributions, and assess community susceptibility to environmental conditions (Chaffron et al., 2021). This, in turn, helps to predict their potential impact on biogeochemical cycling at a planetary level (Fig. 2).

However, large-scale metagenomics presents several challenges, foremost among them being the accuracy of pathway annotations. Standard workflows rely on automated annotation pipelines, followed by manual curation to correct issues like spurious annotations. As data volume increases, however, manual curation becomes impractical. To overcome this, we propose several steps to enhance the accuracy of pathway predictions from large-scale metagenomic data. These steps include defining marker genes uniquely associated with specific pathways to serve as reliable indicators of their presence. Annotation of these genes should utilize methods sensitive to remote homology and be validated through phylogenetic approaches. Furthermore, pathway completeness should be assessed by identifying other related genes within the MAG of interest. In some cases, the

267 operon-like organization of the genes involved in a pathway can also indicate its functionality, increasing annotation reliability
268 (Bratlie et al., 2010). This property has been used to explore the biosynthetic potential of MAGs generated from *Tara* Oceans,
269 leading to the discovery of a particularly talented group of formerly unknown bacteria able to generate multiple compounds of
270 interest in biotechnology (Paoli et al., 2022).

271 Large-scale metagenomic screenings provide a comprehensive view of microbial processes across diverse ecosystems and
272 under varying environmental conditions, thus enhancing our understanding of the phylogenetic and ecological distribution of
273 biogeochemical cycles. Combining these with contextual data (i.e., physical, chemical and biological parameters) can help to
274 predict the link between environmental conditions and a respective biogeochemical pathway. Furthermore, understanding the
275 link between pathways and the environmental conditions in which they occur allows predicting the sensitivity of
276 biogeochemical cycles in a changing environment.

277 In summary, we can leverage large-scale metagenomics to study the distribution of biogeochemical pathways in diverse
278 environments and make predictions of how their distribution might change under changing environmental conditions. The
279 results from such analyses provide the foundation for new hypotheses and experimental research (e.g., isolation of organisms
280 or studying pathway activity in different conditions).

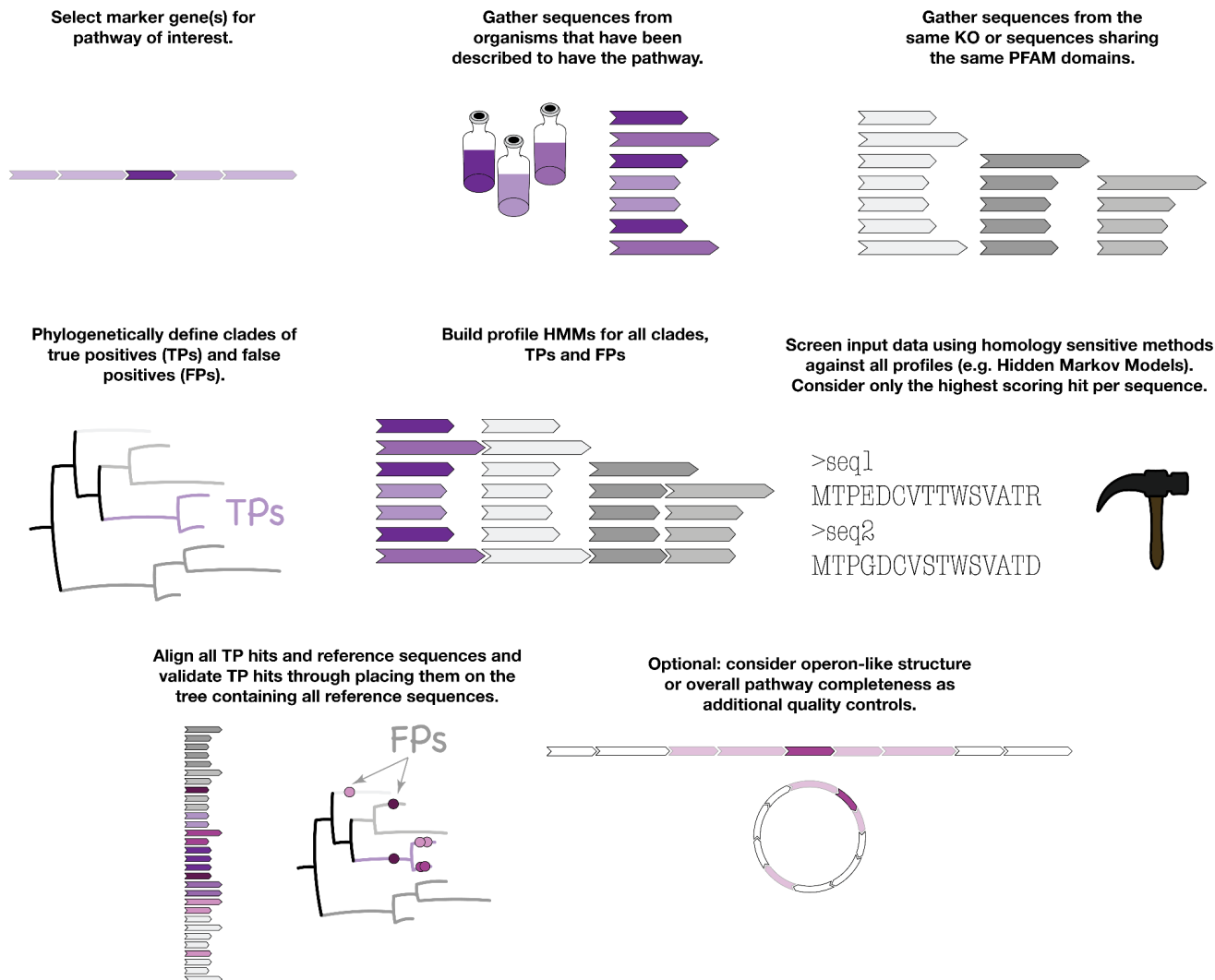
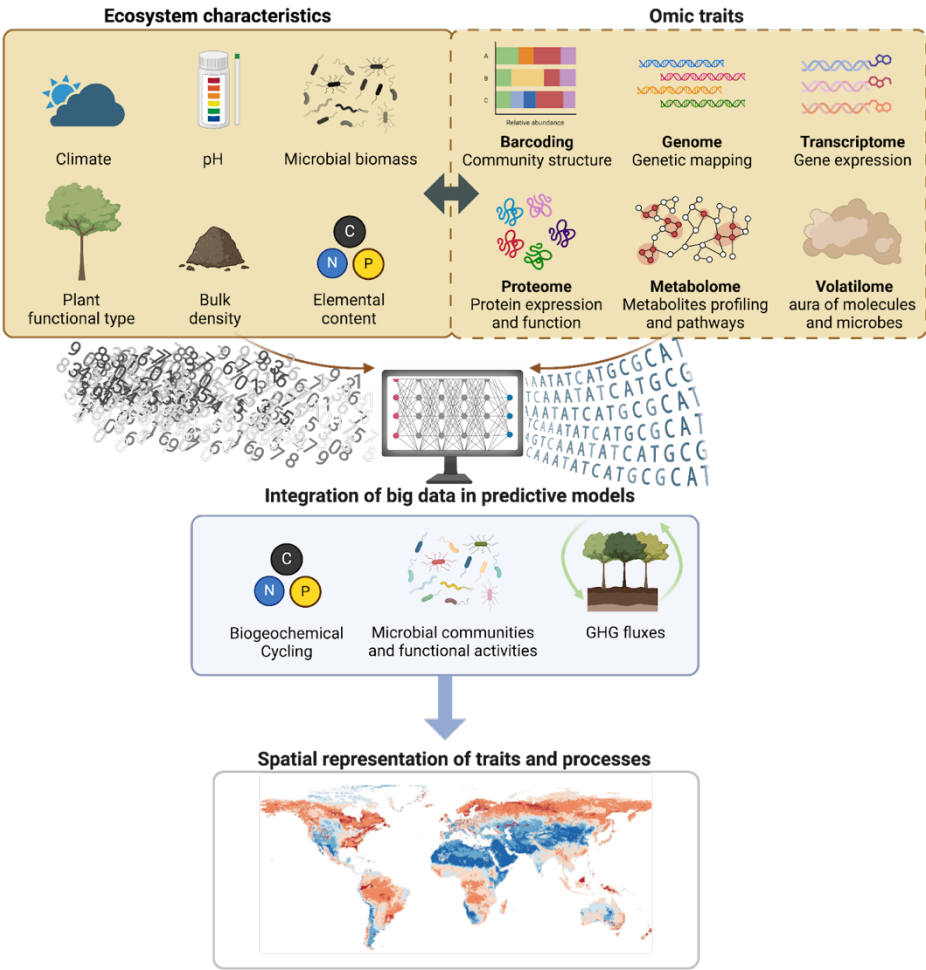


Figure 2: Methodological workflow for the robust screening and functional annotation of biogeochemically relevant pathways in large-scale metagenomic data. The strategy utilizes a multi-stage verification process to ensure high-specificity annotations, starting with the selection of specific marker genes and the gathering of reference sequences from known organisms and shared protein domains, such as PFAM. These reference sequences are used to phylogenetically define clades of true positives (TPs) and false positives (FPs) to build profile Hidden Markov Models (HMMs) for precise screening. Input metagenomic sequences are then screened using homology-sensitive methods, with potential hits validated by placement onto the reference phylogenetic tree to confirm their functional identity. To further handle technical uncertainty and ensure metabolic functionality, an optional final step evaluates the genomic context for operon-like organization or overall pathway completeness within the reconstructed Metagenome-Assembled Genome (MAG).

290
291
292
293
294
295
296
297
298
299

2.4.2 Combining meta-omics to infer large scale biogeochemistry

In most Land Surface Models (LSMs), vegetation and soil attributes are often prescribed in parameter lists and used to describe how biogeochemical processes vary over the land surface. For example, the representation of vegetation is distilled in a variety of plant functional types (PFTs), with each representing a prescribed range of anatomical and physiological parameter values such as leaf mass area, carboxylation rates, or leaf respiration (Duckworth et al., 2000; Woodward and Cramer, 1996). These PFTs can co-vary with certain soil chemical and physical characteristics that are often prescribed, assimilated and/or predicted after some model initialisation and predetermined spin up procedure that drives the vegetation, soil, and atmosphere into some quasi-steady-state to explore feedbacks on biogeochemical cycles (Friend et al., 2014; Matthaeus et al., 2021; Taylor et al., 2011).



300
301

Figure 3: Conceptual figure of a bottom-up system approach using meta-omics. The integration and combination of omic datasets with key vegetation and soil attributes through machine learning models present a great opportunity to enhance our understanding of the genes, pathways and mechanisms underlying biogeochemical cycling and functional activity of soil microbial communities and associated processes such as soil gas exchange for their representation at larger scales. Created with BioRender (<https://BioRender.com/9p03o4f>).

A growing number of databases and maps detailing key soil characteristics—such as soil bulk density, pH, carbon (C), nitrogen (N) and phosphorus (P) content—combined with climate forcing, are now openly available. These resources can support LSM numerical simulations, machine learning models, or hybrid approaches, to predict how soil properties, such as microbial C, N and P biomass, vary across the land surface. These soil characteristics are increasingly being linked to variations in soil microbial community structure and diversity. For example, Delgado-Baquerizo et al. (2016) found that certain microbial groups thrive in high-pH soils, itself shaped by vegetation and climate. Similarly, fungal communities can be mapped to how soil mycorrhizal fungal guilds will vary across the land surface, with implications for soil C storage and nutrient acquisition by vegetation (Kivlin, 2020).

More recently, these databases have been extended to include metagenomic data to show how genes linked to C storage vary across the global land surface (Piton et al., 2023) (Table 1). Linking metagenomic data with measurements of soil GHG fluxes or profiles of soil volatiles could provide mechanistic insight into the metabolic characteristics of soil communities that control GHG and VOC fluxes with the atmosphere. These, in turn, could offer predictive tools for understanding how GHG and VOC fluxes might change with climate or disturbance (Fig. 3). Likewise, other omic datasets, such as soil proteomes or metabolomes, may provide additional insights into the realized and functional activity of soil microbial communities. For example, examining how soil GHG fluxes and gridded soil characteristics vary with soil metabolomes, proteomes, and community structure could pave the way for predicting soil C storage and GHG variability across large scales and over time, even from discrete soil samples.

Beyond direct parameterization, emerging machine-learning (ML) and AI-based models provide a powerful bridge between microbial community structure and ecosystem functions. For example, ML frameworks have been successfully used to map bacterial communities to physicochemical variables and soil quality indicators, allowing for the prediction of ecosystem-level states from sequence data (Hermans et al., 2020). These predictive models can be further expanded to forecast environmental shifts, such as changes in soil pH in response to global warming, by training models on large-scale longitudinal datasets (Feng et al., 2024). Complementary to ML, genome-informed approaches allow for the inference of specific environmental preferences, such as microbial pH optima, directly from genomic traits (Ramoneda et al., 2023). By linking these inferred traits to functional outcomes, researchers can move beyond taxonomic lists to mechanistic representations of how microbial 'niche breadths' influence biogeochemical stability.

334 **Table 1: Databases and resources for linking omics data to ecosystem processes.** The table provides an illustrative, non-exhaustive
 335 overview of widely used resources across ecosystems, including data type, ecosystem coverage, strengths and limitations, access standards,
 336 and references where applicable.

Database	Data type	Ecosystem coverage	Functional annotation resource	Strengths	Limitations	Access	Metadata standard	Reference
Genomic and metagenomic reference resources								
GTDB	MAGs	Global	eggNOG-mapper, KEGG	>700,000 MAGs	MAGs created from sequences not from the same environment	GTDB portal (https://gtdb.cogonomic.org/)	MIxS (MIGS/MIMS)	Parks et al. (2022)
SPIRE	MAGs	Global (multi-ecosystem)	eggNOG-mapper, KEGG	> 1,000,000 MAGs from > 700 studies, covering diverse ecosystems, microntology based habitat annotation	Limited metadata annotations, no functional annotations	http://spire.embl.de/	Custom microntology	Schmidt et al. (2024)
Earth Microbiome Project (EMP)	Amplicons, metagenomes	Global (multi-ecosystem)	PICRUSt2 / QIIME 2 / KEGG	Unprecedented spatial coverage; harmonized protocols	Limited functional resolution for many processes	Qiita (https://qiita.ucsd.edu/emp/)	MIxS / ENVO	Thompson et al. (2017)
GOMC	MAGs, metagenomes	Marine (global)	KEGG	Global coverage of marine metagenomes		https://db.cngb.org/maya/datasets/MDB000002		Chen et al. (2024)
MycCosm	Genomes / MAGs	Fungi (global)	FUNGuild / Fun ^{Fun}	Highly curated fungal genomes; strong taxonomic coverage	Limited functional context	JGI portal (https://genome.jgi.doe.gov/portal/)	MIxS	Grigoriev et al. (2014)
PhycoCosm	Genomes	Algae (marine & freshwater)	Phytozome / KEGG	Broad phylogenetic coverage of algal lineages	Sparse ecological metadata	JGI portal (https://genome.jgi.doe.gov/portal/)	MIxS	Grigoriev et al. (2021)
1KP (1000 Plant transcriptomes initiative)	Transcriptomes	Plants (global)	FASTA sequence files	Large-scale effort to collect next-generation	Gaps in streptophyte algae, almost all of the	1KP portal (https://sites.google.com/a/ualberta.ca/one)	MIxS	Carpenter et al. (2019)

				sequencing (NGS) data for over a thousand phylogenetically diverse species across the tree of life for green plants	nonflowering plant lineages, angiosperm lineages outside of the core eudicot and grass clades, and many other lineages within these clades	kp/)		
Trait-based and metabolic specialized resources								
TRY	Functional traits	Plants (global)	Plant Trait Ontology (PTO)	Standardized plant trait database; widely used in ESMs	No microbial traits	TRY portal (https://www.try-db.org/TryWeb/Home.php)	TRY standard / Controlled vocabularies	Fraser (2020)
GRooT	Root traits	Plants (global)	Plant Trait Ontology (PTO)	Standardized root trait database; more than 1,000 species with a wide geographical range	Phylogenetic gaps are observed for traits related to key categories, such as anatomy, architecture, dynamics and physiology.	GRooT portal (https://groot-database.github.io/GRooT/)	MIxS	Guerrero-Ramírez et al. (2021)
SPUN Mapping Consortium	Fungal richness maps	Mycorrhizal fungi (global)	Virtual taxa (VT) and operational taxonomic units (OTUs); Fungal richness	Largest dataset of fungal internal transcribed spacer (ITS) and small subunit (SSU) rRNA amplicon sequences assembled so far.	Insufficient data on ericoid and orchid mycorrhizal fungal richness; relatively coarse spatial scale (1 km ²)	Data archived on Zenodo (https://doi.org/10.5281/zenodo.14871588) and full map predictions (1-km ² global GeoTIFFs) available (https://www.spun.earth/?r=0)	MIxS	Van Nuland et al. (2025)
FungalRoot	Fungal - Plant Mycorrhizal	Plants (global)	Plant mycorrhizal status	Large database including 14,870 plant species.	Records are unequally distributed among plant species; Observations unequally distributed,	PlutoF platform (https://www.gbif.org/dataset/744edc21-8dd2-474e-8a0b-b8c3d56a3c2)	MIxS	Soudzilovskai a et al. (2020)

					with greatest density in North Europe and North America	d)		
GSTI (The Global Spectra-Trait Initiative)	Leaf hyperspectral and gas exchange measurements	Plants (global)	Reflectance	>7,500 paired observations from around 400 species from a broad range of biomes	Overrepresentation of certain biomes and species	ESS-DIVE repository (doi:10.15485/2530733)	MIxS	Lamour et al. (2026)
MassIVE and GNPS	Web-based mass spectrometry analysis platform	Global	Multiple spectral libraries (e.g. GNPS library spectra, Massbank, HMDB). Built on top of MassIVE (MS data repository)	Community spectral sharing, include metadata. Provides additional tools such as feature-based molecular networking.	Metabolomics data must be pre-processed externally.	https://gnps2.org/homepage https://massiv.e.ucsd.edu/	No controlled vocabulary. DOI-linked citation.	Wang et al. (2016)
MetaboLights	Database for Metabolomics	Global	Primarily a metabolomics data repository, but it does link to annotation resources (e.g. CheBI, HMDB)	Curated. Standardised deposit procedure. Includes metadata. Classification system by organism, organism part, technology.	Metabolomics data must be pre-processed externally.	https://www.ebi.ac.uk/metabolights/	ISA framework (ISA-Tab format). Specific ontology. DOI-linked citation.	Yurekten et al. (2024)
Integrated contextual observatories								
IMG/M	Genomes, MAGs, metagenomes	Global (multi-ecosystem)	KEGG / microTrait	Integrated analysis platform; standardized annotations	Dependent on user-submitted metadata quality	JGI (https://genome.jgi.doe.gov/portal/)	MIxS	Chen et al. (2023)
Tara Oceans	MAGs, metagenomics, metatranscriptomics	Marine (global)	OM-RGC / eggNOG	Planetary-scale standardized sampling; deep functional resolution; highly	Limited function. No temporal resolution	ENA / PANGAEA (https://doi.pangaea.de/10.1594/PANGAEA.842237)	MIxS	Paoli et al. (2022); Sunagawa et al. (2020)

				contextualized				
NMDC	Multi-omics (genomes, transcriptomes, organic matter (FT-ICR MS), proteomes, lipidomes, metabolomes)	Global (multi-ecosystem)	KEGG / microTrait	Strong emphasis on metadata standards and interoperability	Adoption still limited; uneven ecosystem coverage	NMDC portal (https://data.microbiomedata.org/)	MIxS / NMDC Schema	Eloe-Fadrosch et al. (2022)
LUCAS	Soil chemical and physical traits	Soils (Europe)	GlobalSoilMap / WoSIS standards	Harmonized continental-scale soil dataset	Limited geographic coverage (EU only)	EU standard (https://esdac.jrc.ec.europa.eu/projects/lucas)	ISO 28258 (Soil quality) or INSPIRE compliant	Tóth et al. (2013)
Global data infrastructures								
PANGEA	Data from Earth system research including all Earth's spheres	Global (multi-ecosystem)		Contextual georeferenced metadata Editorial curation Interoperability + citability	Not limited to Omics data	https://www.pangaea.de	ISO 19115 Machine-readable exports (schema.org, OAI-PMH) DOI-linked citation metadata	Felden et al. (2023)
GEO BON (Group on Earth Observations Biodiversity Observation Network)	Essential Biodiversity Variables (EBVs)	Terrestrial, marine, and freshwater	Standardised indicators to quantify biodiversity	Central data repository	Limited function	https://geobon.org/	MIxS	Griffith et al. (2026)
ACTRIS	Aerosol, Clouds and Trace Gases	Atmosphere (mainly Europe)	Climate and Forecast (CF) Metadata Conventions	High-quality long-term physical/chemical contextual data	No biological and omics component implemented yet	ACTRIS portal (https://www.actris.fr/catalogue/)	ISO 19115; API - catalogue (machine-actionable); NetCDF + CF + ACDD	Bedoya-Velázquez et al. (2019)
AERIS	Atmospheric dynamics and chemistry	Atmosphere (regional to global)	AERIS Parameter Vocabulary / GCMD	Centralized access to French/European	Focused on physico-chemistry data	AERIS portal (https://www.aeris-data.fr/catalog)	ISO 19115 / INSPIRE	Ranjan Hatua and Madalli (2011)

			Keywords	atmospheric data		ue/)		
--	--	--	----------	------------------	--	------	--	--

337

338 **2.4.3 Integrate omics data into biogeochemical models**

339 Traditional biogeochemical models often apply simplified representation of microbial processes, relying on parameters derived
340 from cultivated microorganisms, which may not accurately capture the complexity and diversity of natural ecosystems (Steen
341 et al., 2019). The increasing availability of metagenomics data presents an opportunity to integrate those microbial traits from
342 genomes into mechanistic ecosystem models, potentially enhancing their predictive capabilities (Bahram et al., 2018;
343 Woodcroft et al., 2018). By inferring microbial traits from genomes obtained in-situ, the metabolic potential of the microbial
344 communities would be better represented, thereby providing more accurate and robust prediction of ecosystem functioning.

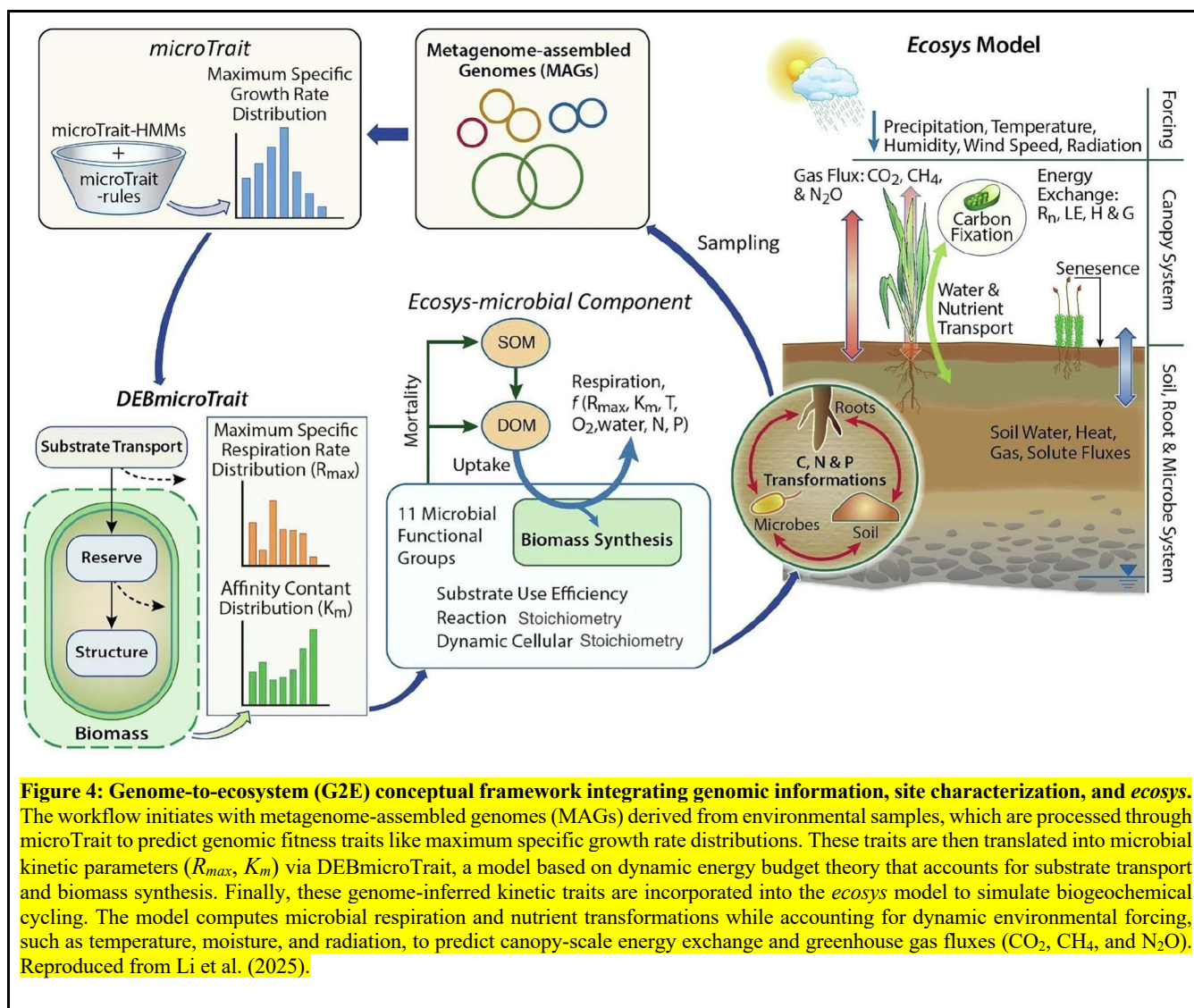
345 Incorporating genomic information into biogeochemical models requires a structured approach. Here, we review two groups
346 of models: soil biogeochemical models, and full ecosystem-level models, each with distinct capabilities and approaches
347 leveraging omics data. Soil biogeochemical models are designed to simulate detailed soil processes, and are typically driven
348 by diagnostic inputs, such as prescribed litter, soil moisture, and temperature data. Some of these models have explicit
349 representations of microbial processes, including the Microbial-Enzyme Decomposition (MEND) Model, MIcrobial-MIneral
350 Carbon Stabilization (MIMICS), and CORPSE; see Chandel et al. (2023) for a full review. For example, MEND has been
351 constrained and validated using GeoChip-based gene abundance measurements, which provides a basic platform for testing
352 hypotheses about microbially-mediated biogeochemical processes (Gao et al., 2020; Wang et al., 2022a). In contrast to soil
353 biogeochemical models, full ecosystem models simulate the coupled dynamics of water, energy, plants, and soil processes
354 driven by climate forcing. Such coupled models are valuable for projection or spatial extrapolation because many of the
355 emergent properties of primary scientific and societal concern, such as net ecosystem exchanges of CO2 and CH4, are strongly
356 regulated by interactions between these interconnected systems. One example of such a model that integrates microbial
357 genomics data is the genome-to-ecosystem (G2E) framework proposed in Li et al. (2025) (Box 3).

Box 3: A worked example: genome-to-ecosystem (G2E) framework

The G2E framework is a three-step process: (1) genomic trait prediction from genomes, 2) translation of genomic traits to kinetic traits, and 3) incorporation of kinetic traits into ecosystem models (Fig. 4). The first step involves predicting traits from metagenome-assembled genomes which are obtained from environmental samples. For this step, microTrait, which extracts fitness traits from genomic sequences (Karaoz and Brodie, 2022), was applied, but analogous tools are available (Shaffer et al., 2020). In the second step, the predicted fitness traits are translated to microbial kinetic trait parameters commonly used in ecosystem models. For this step, DEBmicroTrait, a model based on dynamic energy budget (DEB) theory

(Marschmann et al., 2024), was applied, but other approaches have also been used (Kumar et al., 2019). DEBmicroTrait translates fitness traits to Michaelis-Menten kinetic parameters, e.g., maximum specific respiration rate ($R_{\max}$) and half-saturation constant (K_m). In the third step, *ecosys*, a mechanistic ecosystem model, is used to compute microbial respiration rates based on the Michaelis-Menten rate law, while accounting for dynamic environmental constraints (water, temperature, oxygen, nutrients) (Grant, 1998).

The G2E framework was applied to study methane (CH_4) emissions from a thawing permafrost system (Li et al., 2025). The study found that variations in genome-inferred microbial traits lead to large differences in CH_4 emissions, demonstrating the critical role of genome-inferred microbial traits in regulating ecosystem functioning. In addition, the findings highlight the importance of including relative abundance of genomes in trait inference and integration with ecosystem models. While the study successfully linked genomic data to ecosystem processes and demonstrated the value of incorporating microbial relative abundance for improved model predictions, several caveats and needed future work are identified. For example, although supported by past research at the site, using relative abundance is a simplification of microbial activity (McCalley et al., 2014). Future work could enhance the trait-inference approach by including absolute abundance or other microbial activity indices. In addition, the current *ecosys* structure allows for 11 microbial functional groups but does not account for microbial diversity within each functional group, which could impact emergent carbon cycling rates. This caveat is also highly relevant for methane sink processes. Microbial methane oxidation represents the dominant methane filter in groundwater, removing an estimated ~66% of methane globally (Heinze et al., 2025). However, oxidation efficiency varies substantially with hydrogeochemical conditions and methanotrophic community composition, with different lineages associated with high- versus low-rate environments. These findings underscore that accounting for trait variation and microbial diversity within functional groups may be critical for improving model representations of methane cycling. Future work that expands the current *ecosys* structure to allow microbial diversity could potentially reduce model uncertainty. Furthermore, the study focused on two sub-habitats within an arctic wetland. Applying the G2E framework to other locations and ecosystems could extend the use of genomic data across larger landscapes, enhancing our understanding and prediction of carbon-climate feedback.



2.5 Using omics to reconstruct past and forecast future environments

2.5.1 Past

Omics methodologies in paleo-ecological reconstructions have advanced from mere novelty to producing new ways of understanding ecosystem functions. Compared with traditional palynology techniques, such as pollen, fungal spores, and bacteria microfossils, metagenomics for instance captures changing trophic relationships among primary producers, consumers, and decomposers albeit at limited spatial scales (Jia et al., 2022). Still, omics and conventional proxies can be

365 successfully combined despite the variations in spatial scale of processes and taxonomic resolution. For example, omics have
366 been utilized to differentiate past ecosystem drivers and expand assessments of vegetation changes in rangeland ecosystems
367 (Box 4). However, combining traditional and omics evidence in paleo-ecology is still in its infancy because there are few
368 epistemically independent models (Edwards, 2020; Jones and Bösl, 2021).

369 Molecular paleomicrobiology is an emerging field that leverages ancient and/or historical DNA (aDNA) technologies to
370 uncover how microbial genomes and communities have evolved over time in response to environmental changes (Grasso et
371 al., 2024). By studying ancient microbial remains preserved in dental calculus, paleofeces, archaeological sediments, and even
372 environmental samples like permafrost, aquatic sediments and ice cores, scientists gain insights into historical human-microbe
373 interactions and environmental adaptations (Capo et al., 2022a; Warinner, 2022; Warinner et al., 2014, 2017). Remarkably
374 well-preserved microbial aDNA from stable environments, such as permafrost, allows us to track the evolutionary history of
375 microbial taxa and understand ecosystem responses to past climate shifts. For instance, DNA from two-million-year-old
376 Greenland permafrost has revealed how plant and animal communities from the Middle Pleistocene adapted to environmental
377 changes. However, the complex mixture of DNA from ancient and more recent microbial populations still complicates efforts
378 to disentangle the signals from pioneering microbes and those deposited at a later time (Fernandez-Guerra et al., 2023a; Kjær
379 et al., 2022).

380 While post-mortem DNA degradation complicates the analysis, predictable damage patterns, such as fragmentation and
381 specific base modifications, offer clues to the age of the samples (Briggs et al., 2007). This information is crucial as we work
382 to reconstruct microbial genomes from diverse environments, which might contain amounts of aDNA that may reflect past
383 ecosystems. Recently, DNA could be even extracted from low-biomass containing carbonate rocks with distinct DNA
384 fragmentation patterns allowing us to separate ancient from modern communities (Wegner et al., 2023). Thus, these rock
385 formations might act as potential archives of ancient microbial records. However, MAGs have to be reconstructed from these
386 challenging environments to deepen our understanding of subsurface microbial life and its long-term changes, ultimately
387 refining estimates of microbial populations in the Earth's crust.

388 Long-term records of aquatic biota are still scarce as monitoring using DNA approaches started about two decades ago.
389 Fortunately, the extraction and sequencing of environmental DNA preserved in sedimentary archives has shown its potential
390 to reconstruct past environmental changes and their impact on aquatic microbial diversity (Capo and Barouillet, 2023; Nguyen
391 et al., 2023). In a recent review, Barouillet et al. (2022) highlighted how molecular paleoecology research augments knowledge
392 about the effects of anthropogenic stressors on past aquatic biota. On a longer time scale, Armbrrecht et al. (Armbrrecht et al.,
393 2022) show that photosynthetic genes dated back to 540,000 thousand years ago can be detected from Southern Ocean
394 sedimentary archives. Overall, past DNA data is foreseen to be a valuable proxy for projecting the impacts of near-future
395 environmental changes on microbial diversity across both aquatic and terrestrial systems.

Box 4: Sedimentary ancient DNA in Rangelands

Rangelands provide a habitat for various native and domestic herbivores which interact with and impact these ecosystems throughout history. Wetland sedimentary environments, where pollen accumulates are important areas for herbivores by providing essential resources such as food and water. Herbivores deposit dung in these areas after feeding on various plants like grasses, forbs, shrubs, and trees from parts of landscapes.

Tabares et al. (2020) demonstrated that sedimentary aDNA provided a more detailed taxonomic resolution of grasses and forbs, revealing shifts in edible plant species due to fire management. Furthermore, Gigu et-Covex (2014) discovered that aDNA was able to trace the functional differences among herbivores, such as equids and bovids, which have varying feeding habits and effects on landscapes. aDNA from sedimentary sources not only enhances traditional pollen -based (palynology) techniques, but metagenomics—analysing all DNA in environments simultaneously—could be a stronger method for assessing trophic shifts in ecosystems thereby improving rangeland conservation (Fig. 5).

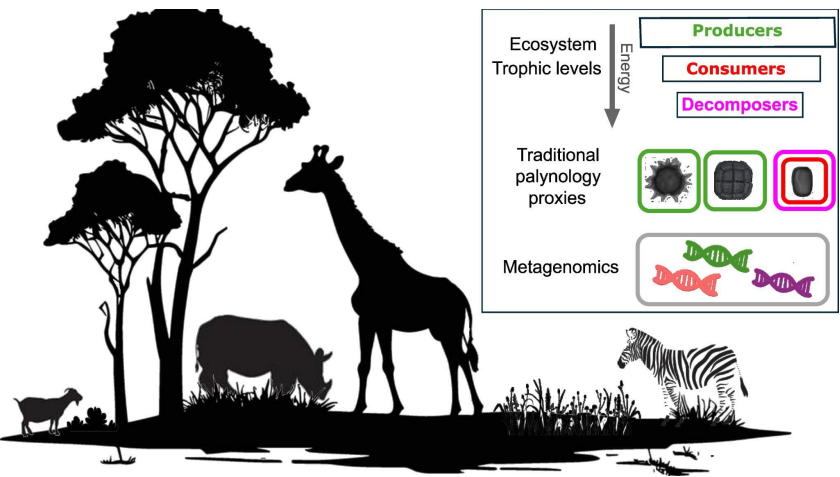


Figure 5: A graphical representation of trophic energy flow in African savanna rangelands featuring producers, consumers, and decomposers. The gray arrow indicates the flow of energy while the colour-coded trophic level boxes represent trophic biomass (not to scale) across the chain. Vegetation dynamics of this system can be reconstructed from fossil pollen, which can be related to herbivore grazing pressure and nutrient cycling via coprophilous fungal spores. Metagenomics, in contrast, is useful for resolving protein and genetic sequences, simultaneously tracing multiple trophic interactions. However, since preservation and strengths of trophic relationships may vary within ecosystems, combining palynology and omics produces better results (e.g., Tabares et al., 2020).

397

398 **2.5.2 Future**

399 *Experimental approaches to studying microbial evolution and its role in biogeochemistry*

400 Geological and paleoclimate records provide environmental analogs to how ecosystems may respond to future climate
401 scenarios. However, such studies are typically limited in their scope by generalizing at the ecosystem-level without insights
402 into the processes and mechanisms driving individual species responses. One such response is local adaptation to changing
403 environmental conditions, where species may disperse to new favourable conditions and expand their geographic ranges or
404 evolve adaptive traits over evolutionary time through gene flow with neighbouring populations and/or emergence of
405 advantageous de novo mutations. Given the long timescales to observe adaptive traits, reciprocal transplant experiments have
406 provided direct insights into the ecological and evolutionary responses to predicted climate conditions. For example, plant
407 communities, given their importance in terrestrial ecosystems, have long been studied demonstrating plant species are locally
408 adapted to current environmental conditions. Such studies highlight differential survival and fecundity (i.e., flowering time) to
409 changing conditions, highlighting phenological traits that may contribute to adaptive differentiation and possible evolutionary
410 responses (Ågren et al., 2012). In other words, past evolutionary divergence caused by local adaptation can restrict and
411 influence contemporary ecological patterns and processes (Urban et al., 2020). Extending these approaches to explicitly test
412 how microbial ecological and evolutionary responses alter biogeochemical rates represent a critical next step for predicting
413 ecosystem responses to future climate change (Box 5).

414 Numerous microbial studies have demonstrated, through reciprocal transplant experiments or environmental manipulations
415 (e.g., drought or temperature), that microbial communities can rapidly respond through the ecological process such as species
416 sorting, defined by the differential survival and reproduction of locally adapted taxa (e.g., DeAngelis et al., 2015; Finks et al.,
417 2021). These taxonomic shifts can have direct implications for ecosystem function, including changes in decomposition rates
418 (Glassman et al., 2018) and reductions in carbon use efficiency under long-term soil warming (Li et al., 2019). In addition,
419 dispersal can play a significant role in microbiome recovery and successional dynamics following large-scale disturbances
420 (i.e., wildfires; Barbour et al., 2023), which are expected to increase in frequency and intensity under future climate scenarios.

421 However, microbes are distinct in that their generation times, and thus capacity for adaptation, operate on much faster
422 timelines. Combined with large population sizes and their capacity to exchange DNA within and across lineages, microbes are
423 exposed to a vast reservoir of genetic diversity within microbial communities that may contribute to rapid adaptive responses
424 to environmental change. Certainly, lab-based experimental evolutionary studies that manipulate a selective pressure (e.g.,
425 heat) have repeatedly demonstrated that microbes can rapidly evolve and adapt over weeks to months (Kent et al., 2018;
426 Tenaillon et al., 2012), although it's unclear how these artificial laboratory conditions translate to natural populations. In
427 contrast to species sorting, which alters the relative abundances of existing trait variants, evolutionary responses modify trait

values themselves, potentially changing microbial reaction norms and longer-term sensitivities to environmental change. Recently, natural experiments using a reciprocal transplant approach across a regional climate gradient demonstrated that microbes, as observed in lab-based experiments, can rapidly adapt through *de novo* mutations in response to predicted climate scenarios (Chase et al., 2021a). However, such adaptive responses under natural conditions are likely operating under reduced selective pressures on monthly and annual timescales and, thus, may have a reduced role compared to rapid ecological responses (Chase et al., 2021a). Nonetheless, a shift in the microbiome under changing climate conditions is dependent on the continuum of ecological and evolutionary processes that operate at the same time. It is therefore the consideration of the relative influence of ecological and evolutionary processes - and the timescales over which they interact - that will be necessary to predict how microbial communities regulate biogeochemical rates and ecosystem functioning under environmental change (Abs et al., 2024b, c; Martiny et al., 2023).

Predicting long-term ecosystem responses using omics-parameterized models

Using microbial kinetic traits inferred from multi-year omics data, *ecosys* can be used to predict future ecosystem dynamics. Specifically, *ecosys* was employed to predict ecosystem dynamics in an Arctic wetland under the RCP8.5 scenario (RCP: Representative Concentration Pathway) through the year 2100 (Li et al., 2025). That *ecosys* simulation incorporated genome-inferred microbial traits (i.e., community-aggregated traits weighted by genome relative abundance) and baseline microbial traits. The model results show reasonable seasonal cycles of methane and net ecosystem exchange (NEE) during the late 21st century, and have about 18% and 21% difference in annual methane emissions and NEE, respectively. The differences were mediated by shifts in the overall carbon cycle, largely due to microbial activity, belowground nutrient transformations, and vegetation dynamics (Bouskill et al., 2020). We highlight some caveats to these results. First, the microbial traits inferred from omics that we applied are constants based on current life-history traits (e.g., growth rate, resources utilization, and stress tolerance); however, those traits may adapt to changing environmental conditions (Sokol et al., 2022). Work is underway to develop models of how adaptation may occur, and to integrate those approaches with ecosystem models (Abs et al., 2022, 2024b; Abs and Ferrière, 2020; Schwarz et al., 2025). Second, the genome-inferred microbial traits we determined span ranges within each microbial functional group, reflecting the multiple genomes that perform the same metabolic functions (Li et al., 2025). However, our application of these traits described here considered only a single value for each trait in each microbial group. Ongoing work in *ecosys* will allow for trait diversity within microbial functional groups, thereby allowing prediction of dynamic community assembly under dynamic and competitive environmental conditions.

Box 5: Predicting future ocean biogeochemistry

Predicting the fate of marine ecosystems in an ocean impacted by dramatic change is not possible without a base line. However, global surveys are rare because of the requisite commitment to standardize sampling and analysis methods on a huge scale.

Most attempts to predict future ocean biogeochemistry therefore tend to focus on one parameter, such as how concentrations of a critical nutrient may change as a function of altered temperature or ocean mixing (Shaked et al., 2021), or how the features of one specific region may change in the future (Tagliabue et al., 2017). Alternatively, efforts have been made to combine datasets from different sources, with the inherent problem of ensuring data consistency to avoid comparing apples with oranges. This generally means that only the most rudimentary data types can be combined from different sampling programs. A prominent example of a single global survey that has generated an internally consistent dataset is *Tara* Oceans (Bork et al., 2015). From 2009-2013, an international team of scientists onboard the research schooner *Tara* performed a comprehensive sampling program at more than 200 sites worldwide, aimed at exploring the entire size range of organisms within plankton communities in the upper 1,000 m of the ocean, from the smallest viruses of a few tens of nanometres to zooplankton several millimetres in size. The initiative generated tens of thousands of high quality samples, and their subsequent analysis using highly standardized protocols based on omics and microscopy has generated the first ever baseline of life in the ocean at global scale. As a consequence of the highly contextualized, comprehensive open access datasets that have been generated by *Tara* Oceans scientists that cover the entire community of organisms residing in the plankton (summarized in Sunagawa et al., 2020), many thousands of primary research papers have been generated by the international scientific community. These address multiple fundamental research questions that, in general, have been used for improving our understanding of: 1) Basic molecular and cellular processes in all domains of life, and beyond canonical model organisms, 2) Ecological and evolutionary principles, and 3) The interface between marine ecosystems and functioning of the Earth System, and the feedbacks between them. One prominent biogeochemical investigation that was focused on carbon export is that of Guidi et al. (2016). In this study, carbon export to the ocean interior (a proxy for which was measured during the expedition) was related to the microbial communities present in the upper photic zone using machine learning approaches. A network of plankton groups were found to correlate particularly strongly with carbon export, the most prominent and surprising of which were the smallest components of these ecosystems, such as *Synechococcus* and the viruses that infect them, known as *Synechophages*.

The *Tara* Oceans base line can also be used as the starting point to address how marine ecosystems and the geochemical pathways they manage may change in the future. There are now several examples of how this has been done (Frémont et al., 2022, 2023; Ibarbalz et al., 2019), with a typical approach being to project changes in the future ocean simulated in the IPCC climate models. An example is shown in Fig. 6, where anomalies in the biodiversity of major planktonic groups by end-century are predicted based on current patterns of latitudinal diversity gradients. The results predict major changes throughout the global ocean, and especially in high latitude regions that are of particular importance for carbon capture by the biological carbon pump and for the fishing industry. These regions also coincide with high densities of marine protected areas, highlighting the need to take heed of these results in future conservation efforts.

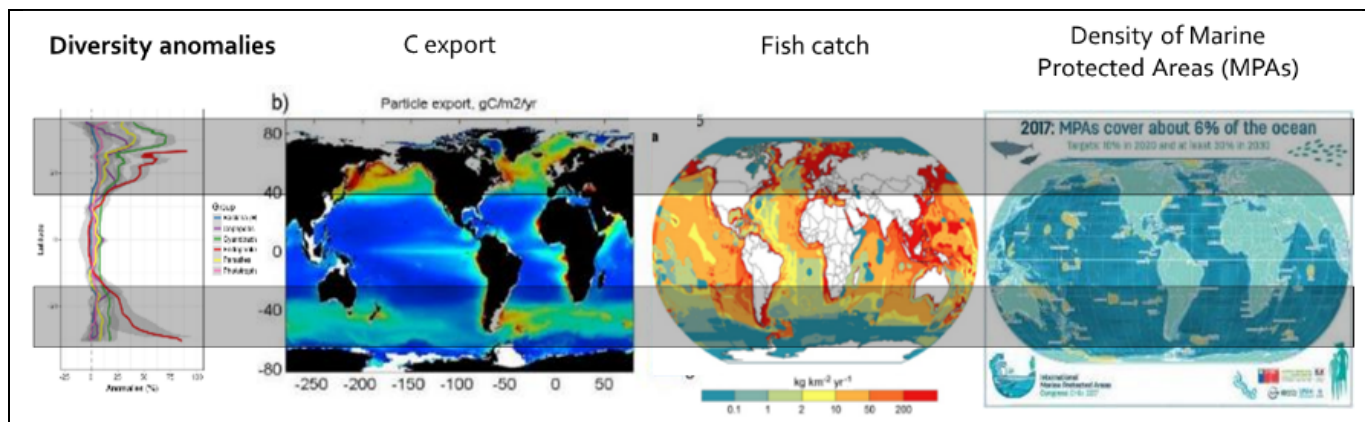


Figure 6: Projected effects of diversity changes within plankton communities on marine ecosystems, fisheries, and biogeochemical cycles. The panel in a) summarizes the biodiversity anomalies projected by end-century from *Tara* Oceans data for a range of different plankton groups (indicated by the different colours), and emphasizes that these projections indicate major changes at mid- to high latitudes (grey bars). The global maps in panel b) indicate estimates of carbon export in different oceanic regions (left), fish landings (centre), and the locations of Marine Protected Areas (right), all of which are likely to be impacted significantly by the projected biodiversity anomalies shown in a) (grey bars). Figure adapted from Ibarbalz et al., 2019; Tara Ocean Foundation et al., 2022).

While exercises such as these provide valuable information about how the ocean's biogeochemical cycles may be impacted by climate change in the future, a major caveat is that they do not include any information about how organisms may adapt to future conditions. Most likely, some organisms will migrate to new environments, others will go extinct, and yet others will adapt or evolve. Our lack of understanding about such processes is a major impediment to understanding what the future landscape of life on our planet will look like. Extensive studies in powerful experimental systems such as mouse, *Drosophila* and *Arabidopsis* have revealed much about the importance of both genetic and epigenetic drivers in evolution (Chatterjee, 2024; Feinberg and Irizarry, 2010; Schmid et al., 2018), but much remains to be done to understand adaptation in the natural world.

3 Challenges and perspectives

Long-term Earth system monitoring programs, along with preserved sample archives, present a valuable opportunity to enhance ESMs by providing insights into climate-driven ecosystem transformations. The recent surge in omics-based technologies has catalysed untargeted measurements across diverse environments, as highlighted in this paper with examples from deep-sea sediments, plant canopies, the atmosphere, cryospheric regions and marine ecosystems. However, the proposed implementation of omics driven data into modelling efforts faces challenges outlined in the following:

3.1 Data integration and interpretation

Integration of omics data into ESMs is complex due to the sheer diversity, scale, and volume of data these approaches generate. Multi-omics datasets capture different biological layers, often involving vast amounts of heterogeneous data. This heterogeneity arises because each of these layers represents different types of biological information, from genetic sequences to metabolite profiles and often vary in data structure, format, and the presence of missing information due to limitations inherent in omics technologies (Ramirez Flores et al., 2023).

Connecting these diverse datasets to critical ESM response variables—such as GHG emissions, soil carbon stocks, or other processes relevant to predicting human-induced climate change—is challenging. For instance, microbial-driven methane production is well characterized: its defined enzymatic pathways and environmental controls might enable more straightforward integration of omics data. In contrast, turnover of soil organic carbon is driven by a highly diverse array of microbes operating under variable conditions. This complexity might make it much harder to translate omics-based measurements into accurate process rates for ESM incorporation (He et al., 2024a).

Recent advances in deducing microbial traits from metagenomic data show great promise for integrating omics information into biogeochemical models by linking these data with kinetic parameters (Karaoz and Brodie, 2022). By identifying functional traits from metagenomes, researchers can begin to quantify the roles of various microbes in processes like nutrient cycling or carbon turnover, which holds the potential to enhance the predictive power of ESMs. Extending this approach to other types of omics data could lead to more comprehensive ESMs that incorporate detailed biotransformation kinetics, ultimately resulting in more accurate climate-related predictions. The trade-off between the speed of direct observation and the depth of omics sequencing is a matter of descriptive vs. predictive capacity. Direct measurements are essential for tracking current budgets, but physical state changes often lag behind the microbial shifts that drive them. In contrast, characterizing functional potential allows models to account for the 'legacy effects' and 'lagged responses' of communities, ensuring that projections remain accurate even as historical environmental correlations break down.

To navigate these complexities and ensure data reliability, researchers employ a standardized toolkit and validation strategy:

- **Quality Control & Assembly:** Raw sequences are processed to remove noise (QC) before being assembled into contiguous sequences (contigs).
- **Genome Reconstruction & Binning:** Contigs are grouped into Metagenome-Assembled Genomes (MAGs) and evaluated for completeness and contamination using tools like CheckM.
- **Trait Inference:** Bioinformatic pipelines like microTrait or DRAM are used to distill genomic potential into fitness traits and metabolic parameters (V_{max} , K_m).
- **Validation & Uncertainty:** Model results are confronted with in-situ flux measurements (e.g., from NEON or ICOS) to quantify the accuracy of microbial process representations.

- aDNA Authenticity: In molecular paleomicrobiology, specific damage patterns (e.g., fragmentation and base modifications) are used to discriminate ancient genetic signals from modern contamination.

3.2 Data accessibility and standardization

Ensuring accessibility to omics data and establishing robust community standards constitute essential steps to advance the integration of omics data into ESMs. Initiatives like the National Microbiome Data Collaborative (NMDC) (Eloe-Fadrosh et al., 2022) and Integrated Microbial Genomes & Microbiomes (IMG/M) (Chen et al., 2024) created frameworks that promote data sharing and enforce metadata standardization. These frameworks provide structured approaches for documenting and sharing omics data, ensuring that researchers follow uniform guidelines when collecting, annotating, and storing data. However, the adoption of these standards across the global scientific community has been limited so far. Such a globally implemented and executed framework would help address questions on large scales, as it would enable the aggregation and cross-referencing of omics data from diverse ecosystems worldwide. Additionally, standardized data formats would also facilitate data access to emerging AI-driven tools which hold the potential to play a key role in identifying yet unseen barriers in improvements of ESMs and might also help to overcome those.

3.3 Emerging technologies

Technological advances, particularly in ancient/historical DNA (aDNA) analysis, have opened new pathways for studying long-term ecological responses to environmental changes. aDNA analysis enables scientists to retrieve genetic information from organisms preserved in ancient samples, such as ice cores (Zhong et al., 2021) and up to 2 million year old sediment layers (Fernandez-Guerra et al., 2023b; Kjær et al., 2022; Nguyen et al., 2023). By integrating biological data with chemical signatures found in these preserved samples, we can reconstruct ecosystem dynamics over extended historical periods, offering valuable insights into how biomes have responded to past climate variations. Such and other omics data will become increasingly available now as ongoing advancements in omics technology have gradually reduced the amount of analyte required for analysis, making it feasible to use small or degraded samples from ancient archives.

Incorporating these ancient biological signatures into ESMs has the potential to enhance model accuracy significantly. By cross-validating current ESM data with insights from historical records, scientists can better understand the relationships between abiotic factors and biotic responses over time. This integrative approach could improve the models' capacity to predict ecological responses to future climate changes by anchoring them in real historical data, providing a long-term perspective that modern data alone cannot offer.

3.4 The necessity of interdisciplinary collaboration

Unsurprisingly, integrating omics data into ESMs demands collaboration across disciplines like ecology, geochemistry, bioinformatics, modelling, and data science. Each field brings crucial expertise: ecologists and geochemists link biological and

chemical processes, bioinformaticians manage complex omics data, modelers translate findings into model parameters, and data scientists develop frameworks to address challenges in data integration and accessibility. Environmental scientists play a key role in identifying biochemical pathways and traits essential for model parameterization, thereby improving the accuracy of ESMs in simulating ecosystem functions.

Collaborative efforts like these would produce enhanced ESMs that incorporate omics data, rooted in comprehensive, relevant datasets and overcoming the barriers of data integration and accessibility. We believe that the EGU community offers an ideal platform for such interdisciplinary initiatives. Leveraging existing long-term ecological monitoring projects through collaborative projects generating omics data aimed for model integration, would be a logical first step in these efforts, and we encourage all stakeholders to use this opportunity and actively participate.

Code/data/sample availability

No code, data or sample were used for this perspective.

Author contribution

E.A. led the writing, E.A., C.K. and L.W. initiated the design and writing of the article. All co-authors wrote at least one subsection and provided input on the manuscript text, table, figures and discussion of scientific content.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

No acknowledgements.

Financial support

Funding to EA was provided by the CALIPSO project provided by Schmidt Sciences. LW has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 101003125) and the French Govt in the framework of the IdEX Bordeaux University "Investments for the Future" program / GPR Bordeaux Plant Sciences. CB has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Diatomic; grant agreement No. 835067) and Horizon Europe BlueRemediomics (grant agreement No. 101082304) and Marco-Bolo (grant agreement No. 101082021). KK has received funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number

218627073: CRC1076 AquaDiva. JH has received funding from the Research Council of Finland (grant 354462) and the Kone Foundation. EC has received funding from the Swedish Research Council VR (VR starting grant 2023-03504). Z.L. was supported at Lawrence Livermore National Laboratory (LLNL) by the LLNL-LDRD Program under Project No. 24-SI-002 and by the U.S. Department of Energy (DOE) Office of Science, Energy Earthshot Initiative, as part of the LLNL Terraforming Soils EERC under Award # SCW1841. Work at LLNL was conducted under the auspices of DOE Contract DE-AC52-07NA27344. W.J.R. was supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Terrestrial Ecosystem Science Program, in the Belowground Biogeochemistry Scientific Focus Area, under Award Number DE-AC02-05CH11231. S.R.S. and Z.L. were supported by the EMERGE Biology Integration Institute (US NSF Award #2022070). CK has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101108340.

References

- Abs, E. and Ferrière, R.: Modeling microbial dynamics and heterotrophic soil respiration: Effect of climate change, in: *Biogeochemical Cycles*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 103–129, 2020.
- Abs, E., Leman, H., and Ferrière, R.: A multi-scale eco-evolutionary model of cooperation reveals how microbial adaptation influences soil decomposition, *Commun. Biol.*, 3, 520, 2020.
- Abs, E., Saleska, S., and Ferriere, R.: Microbial eco-evolutionary responses amplify global soil carbon loss with climate warming, , <https://doi.org/10.21203/rs.3.rs-1984500/v1>, 2022.
- Abs, E., Chase, A. B., and Allison, S. D.: How do soil microbes shape ecosystem biogeochemistry in the context of global change?, *Environ. Microbiol.*, 25, 780–785, 2023.
- Abs, E., Albright, M. B. N., and Allison, S. D.: Invasions eliminate the legacy effects of substrate history on microbial nitrogen cycling, *Ecosphere*, 15, <https://doi.org/10.1002/ecs2.4754>, 2024a.
- Abs, E., Chase, A. B., Manzoni, S., Ciais, P., and Allison, S. D.: Microbial evolution-An under-appreciated driver of soil carbon cycling, *Glob. Chang. Biol.*, 30, e17268, 2024b.
- Abs, E., Coulette, D., Ciais, P., and Allison, S. D.: Microbial evolution drives adaptation of substrate degradation on decadal to centennial time scales relevant to global change, *Ecol. Lett.*, 27, e14530, 2024c.
- Anderegg, L. D., Griffith, D. M., Cavender-Bares, J., Riley, W. J., Berry, J. A., Dawson, T. E., & Still, C. J.: Representing plant diversity in land models: An evolutionary approach to make “Functional Types” more functional. *Glob. Chang. Biol.*, 28, 2541-2554, 2022.
- Acinas, S. G., Sánchez, P., Salazar, G., Cornejo-Castillo, F. M., Sebastián, M., Logares, R., Royo-Llonch, M., Paoli, L., Sunagawa, S., Hingamp, P., Ogata, H., Lima-Mendez, G., Roux, S., González, J. M., Arrieta, J. M., Alam, I. S., Kamau, A., Bowler, C., Raes, J., Pesant, S., Bork, P., Agustí, S., Gojobori, T., Vaqué, D., Sullivan, M. B., Pedrós-Alió, C., Massana, R., Duarte, C. M., and Gasol, J. M.: Deep ocean metagenomes provide insight into the metabolic architecture of bathypelagic microbial communities, *Commun. Biol.*, 4, 604, 2021.

581 Ågren, G. I., Wetterstedt, J. Å. M., and Billberger, M. F. K.: Nutrient limitation on terrestrial plant growth--
582 modeling the interaction between nitrogen and phosphorus, *New Phytol.*, 194, 953–960, 2012.

583 Amato, P., Parazols, M., Sancelme, M., Laj, P., Mailhot, G., and Delort, A.-M.: Microorganisms isolated from
584 the water phase of tropospheric clouds at the Puy de Dôme: major groups and growth abilities at low
585 temperatures: Microorganisms from the water phase of tropospheric clouds, *FEMS Microbiol. Ecol.*, 59, 242–
586 254, 2007.

587 Amato, P., Besaury, L., Joly, M., Penaud, B., Deguillaume, L., and Delort, A.-M.: Metatranscriptomic
588 exploration of microbial functioning in clouds, *Sci. Rep.*, 9, 4383, 2019.

589 Armbrrecht, L., Weber, M. E., Raymo, M. E., Peck, V. L., Williams, T., Warnock, J., Kato, Y., Hernández-
590 Almeida, I., Hoem, F., Reilly, B., Hemming, S., Bailey, I., Martos, Y. M., Gutjahr, M., Percuoco, V., Allen, C.,
591 Brachfeld, S., Cardillo, F. G., Du, Z., Fauth, G., Fogwill, C., Garcia, M., Glüder, A., Guitard, M., Hwang, J.-H.,
592 Iizuka, M., Kenlee, B., O’Connell, S., Pérez, L. F., Ronge, T. A., Seki, O., Tauxe, L., Tripathi, S., and Zheng, X.:
593 Ancient marine sediment DNA reveals diatom transition in Antarctica, *Nat. Commun.*, 13, 5787, 2022.

594 Bahram, M., Hildebrand, F., Forslund, S. K., Anderson, J. L., Soudzilovskaia, N. A., Bodegom, P. M.,
595 Bengtsson-Palme, J., Anslan, S., Coelho, L. P., Harend, H., Huerta-Cepas, J., Medema, M. H., Maltz, M. R.,
596 Mundra, S., Olsson, P. A., Pent, M., Pölme, S., Sunagawa, S., Ryberg, M., Tedersoo, L., and Bork, P.: Structure
597 and function of the global topsoil microbiome, *Nature*, 560, 233–237, 2018.

598 Barbour, K. M., Barrón-Sandoval, A., Walters, K. E., and Martiny, J. B. H.: Towards quantifying microbial
599 dispersal in the environment, *Environ. Microbiol.*, 25, 137–142, 2023.

600 Bar-On, Y. M. and Milo, R.: Towards a quantitative view of the global ubiquity of biofilms, *Nat. Rev.*
601 *Microbiol.*, 17, 199–200, 2019.

602 Bar-On, Y. M., Phillips, R., and Milo, R.: The biomass distribution on Earth, *Proc. Natl. Acad. Sci. U. S. A.*, 115,
603 6506–6511, 2018.

604 Barouillet, C., Monchamp, M.-E., Bertilsson, S., Brasell, K., Domaizon, I., Epp, L. S., Ibrahim, A., Mejbél, H.,
605 Nwosu, E. C., Pearman, J. K., Picard, M., Thomson-Laing, G., Tsugeki, N., Von Eggers, J., Gregory-Eaves, I.,
606 Pick, F., Wood, S. A., and Capo, E.: Investigating the effects of anthropogenic stressors on lake biota using
607 sedimentary DNA, *Freshw. Biol.*, <https://doi.org/10.1111/fwb.14027>, 2022.

608 Bedoya-Velásquez, A. E., Titos, G., Bravo-Aranda, J. A., Haefelin, M., Favez, O., Petit, J.-E., Casquero-Vera, J.
609 A., Olmo-Reyes, F. J., Montilla-Rosero, E., Hoyos, C. D., Alados-Arboledas, L., and Guerrero-Rascado, J. L.:
610 Long-term aerosol optical hygroscopicity study at the ACTRIS SIRTa observatory: synergy between ceilometer
611 and in situ measurements, *Atmos. Chem. Phys.*, 19, 7883–7896, 2019.

612 Biller, S. J., Berube, P. M., Dooley, K., Williams, M., Satinsky, B. M., Hackl, T., Hogle, S. L., Coe, A.,
613 Bergauer, K., Bouman, H. A., Browning, T. J., De Corte, D., Hassler, C., Hulston, D., Jacquot, J. E., Maas, E.
614 W., Reinthaler, T., Sintes, E., Yokokawa, T., and Chisholm, S. W.: Marine microbial metagenomes sampled
615 across space and time, *Sci. Data*, 5, 180176, 2018.

616 Björk, R. G., Björkman, M. P., Andersson, M. X., and Klemetsson, L.: Temporal variation in soil microbial

communities in Alpine tundra, *Soil Biol. Biochem.*, 40, 266–268, 2008.

Bork, P., Bowler, C., de Vargas, C., Gorsky, G., Karsenti, E., and Wincker, P.: Tara Oceans. Tara Oceans studies plankton at planetary scale. Introduction, *Science*, 348, 873, 2015.

Bouskill, N. J., Riley, W. J., Zhu, Q., Mekonnen, Z. A., and Grant, R. F.: Alaskan carbon-climate feedbacks will be weaker than inferred from short-term experiments, *Nat. Commun.*, 11, 5798, 2020.

Bouslimani, A., Sanchez, L. M., Garg, N., and Dorrestein, P. C.: Mass spectrometry of natural products: current, emerging and future technologies, *Nat. Prod. Rep.*, 31, 718–729, 2014.

Bratlie, M. S., Johansen, J., and Drabløs, F.: Relationship between operon preference and functional properties of persistent genes in bacterial genomes, *BMC Genomics*, 11, 71, 2010.

Bravo, A. G. and Cosio, C.: Biotic formation of methylmercury: A bio-physico-chemical conundrum, *Limnol. Oceanogr.*, 65, 1010–1027, 2020.

Breitburg, D., Levin, L. A., Oschlies, A., Grégoire, M., Chavez, F. P., Conley, D. J., Garçon, V., Gilbert, D., Gutiérrez, D., Isensee, K., Jacinto, G. S., Limburg, K. E., Montes, I., Naqvi, S. W. A., Pitcher, G. C., Rabalais, N. N., Roman, M. R., Rose, K. A., Seibel, B. A., Telszewski, M., Yasuhara, M., and Zhang, J.: Declining oxygen in the global ocean and coastal waters, *Science*, 359, <https://doi.org/10.1126/science.aam7240>, 2018.

Briggs, A. W., Stenzel, U., Johnson, P. L. F., Green, R. E., Kelso, J., Prüfer, K., Meyer, M., Krause, J., Ronan, M. T., Lachmann, M., and Pääbo, S.: Patterns of damage in genomic DNA sequences from a Neandertal, *Proc. Natl. Acad. Sci. U. S. A.*, 104, 14616–14621, 2007.

Brown, R. W., Chadwick, D. R., Zang, H., and Jones, D. L.: Use of metabolomics to quantify changes in soil microbial function in response to fertiliser nitrogen supply and extreme drought, *Soil Biol. Biochem.*, 160, 108351, 2021.

Burrows, S. M., Butler, T., Jöckel, P., Tost, H., Kerkweg, A., Pöschl, U., and Lawrence, M. G.: Bacteria in the global atmosphere – Part 2: Modeling of emissions and transport between different ecosystems, *Atmos. Chem. Phys.*, 9, 9281–9297, 2009.

Calayag, A. M., Priest, T., Oldenburg, E., Muschiol, J., Popa, O., Wietz, M., and Needham, D. M.: Arctic Ocean virus communities and their seasonality, bipolarity, and prokaryotic associations, *Nat. Commun.*, 16, 6427, 2025.

Canadell, J. G., Monteiro, P. M., Costa, M. H., Da Cunha, L. C., Cox, P. M., Eliseev, A. V., and Lebehent: Global carbon and other biogeochemical cycles and feedbacks. IPCC AR6 WGI, final government distribution, 2021.

Capo, E. and Barouillet, C.: Tracking environmental change using lake sediments, in: *Sedimentary DNA*, vol. 6, Springer Nature, 2023.

Capo, E., Monchamp, M.-E., Coolen, M. J. L., Domaizon, I., Armbricht, L., and Bertilsson, S.: Environmental paleomicrobiology: using DNA preserved in aquatic sediments to its full potential, *Environ. Microbiol.*, 24, 2201–2209, 2022a.

Capo, E., Feng, C., Bravo, A. G., Bertilsson, S., Soerensen, A. L., Pinhassi, J., Buck, M., Karlsson, C., Hawkes,

651 J., and Björn, E.: Expression Levels of hgcAB Genes and Mercury Availability Jointly Explain Methylmercury
652 Formation in Stratified Brackish Waters, *Environ. Sci. Technol.*, 56, 13119–13130, 2022b.

653 Carpenter, E. J., Matasci, N., Ayyampalayam, S., Wu, S., Sun, J., Yu, J., Jimenez Vieira, F. R., Bowler, C.,
654 Dorrell, R. G., Gitzendanner, M. A., Li, L., Du, W., K Ullrich, K., Wickett, N. J., Barkmann, T. J., Barker, M. S.,
655 Leebens-Mack, J. H., and Wong, G. K.-S.: Access to RNA-sequencing data from 1,173 plant species: The 1000
656 Plant transcriptomes initiative (1KP), *Gigascience*, 8, <https://doi.org/10.1093/gigascience/giz126>, 2019.

657 Cavicchioli, R., Ripple, W. J., Timmis, K. N., Azam, F., Bakken, L. R., Baylis, M., Behrenfeld, M. J., Boetius,
658 A., Boyd, P. W., Classen, A. T., Crowther, T. W., Danovaro, R., Foreman, C. M., Huisman, J., Hutchins, D. A.,
659 Jansson, J. K., Karl, D. M., Koskella, B., Mark Welch, D. B., Martiny, J. B. H., Moran, M. A., Orphan, V. J.,
660 Reay, D. S., Remais, J. V., Rich, V. I., Singh, B. K., Stein, L. Y., Stewart, F. J., Sullivan, M. B., van Oppen, M. J.
661 H., Weaver, S. C., Webb, E. A., and Webster, N. S.: Scientists’ warning to humanity: microorganisms and
662 climate change, *Nat. Rev. Microbiol.*, 17, 569–586, 2019.

663 Chaffron, S., Delage, E., Budinich, M., Vintache, D., Henry, N., Nef, C., Ardyna, M., Zayed, A. A., Junger, P. C.,
664 Galand, P. E., Lovejoy, C., Murray, A. E., Sarmiento, H., Tara Oceans coordinators, Acinas, S. G., Babin, M.,
665 Iudicone, D., Jaillon, O., Karsenti, E., Wincker, P., Karp-Boss, L., Sullivan, M. B., Bowler, C., de Vargas, C.,
666 and Eveillard, D.: Environmental vulnerability of the global ocean epipelagic plankton community interactome,
667 *Sci. Adv.*, 7, eabg1921, 2021.

668 Chandel, A. K., Jiang, L., and Luo, Y.: Microbial models for simulating soil carbon dynamics: A review, *J.*
669 *Geophys. Res. Biogeosci.*, <https://doi.org/10.1029/2023jg007436>, 2023.

670 Chase, A. B. and Martiny, J. B. H.: The importance of resolving biogeographic patterns of microbial
671 microdiversity, *Microbiol. Aust.*, 39, 5, 2018.

672 Chase, A. B., Karaoz, U., Brodie, E. L., Gomez-Lunar, Z., Martiny, A. C., and Martiny, J. B. H.: Microdiversity
673 of an Abundant Terrestrial Bacterium Encompasses Extensive Variation in Ecologically Relevant Traits, *MBio*,
674 8, <https://doi.org/10.1128/mBio.01809-17>, 2017.

675 Chase, A. B., Gomez-Lunar, Z., Lopez, A. E., Li, J., Allison, S. D., Martiny, A. C., and Martiny, J. B. H.:
676 Emergence of soil bacterial ecotypes along a climate gradient, *Environ. Microbiol.*, 20, 4112–4126, 2018.

677 Chase, A. B., Arevalo, P., Brodie, E. L., Polz, M. F., Karaoz, U., and Martiny, J. B. H.: Maintenance of
678 Sympatric and Allopatric Populations in Free-Living Terrestrial Bacteria, *MBio*, 10,
679 <https://doi.org/10.1128/mBio.02361-19>, 2019.

680 Chase, A. B., Weihe, C., and Martiny, J. B. H.: Adaptive differentiation and rapid evolution of a soil bacterium
681 along a climate gradient, *Proc. Natl. Acad. Sci. U. S. A.*, 118, <https://doi.org/10.1073/pnas.2101254118>, 2021a.

682 Chase, A. B., Sweeney, D., Muskat, M. N., Guillén-Matus, D. G., and Jensen, P. R.: Vertical inheritance
683 facilitates interspecies diversification in biosynthetic gene clusters and specialized metabolites, *MBio*, 12,
684 e0270021, 2021b.

685 Chatterjee, R. N.: The role of epigenetics in phenotypic evolution: lessons from the genomic response to stresses,
686 genetic conflicts and developmental plasticity in *Drosophila*, *Nucleus (Calcutta)*, <https://doi.org/10.1007/s13237->

687 [024-00520-3](#), 2024.

688 Chen, I.-M. A., Chu, K., Palaniappan, K., Ratner, A., Huang, J., Huntemann, M., Hajek, P., Ritter, S. J., Webb,
689 C., Wu, D., Varghese, N. J., Reddy, T. B. K., Mukherjee, S., Ovchinnikova, G., Nolan, M., Seshadri, R., Roux,
690 S., Visel, A., Woyke, T., Eloë-Fadrosh, E. A., Kyrpides, N. C., and Ivanova, N. N.: The IMG/M data
691 management and analysis system v.7: content updates and new features, *Nucleic Acids Res.*, 51, D723–D732,
692 2023.

693 Chen, J., Jia, Y., Sun, Y., Liu, K., Zhou, C., Liu, C., Li, D., Liu, G., Zhang, C., Yang, T., Huang, L., Zhuang, Y.,
694 Wang, D., Xu, D., Zhong, Q., Guo, Y., Li, A., Seim, I., Jiang, L., Wang, L., Lee, S. M. Y., Liu, Y., Wang, D.,
695 Zhang, G., Liu, S., Wei, X., Yue, Z., Zheng, S., Shen, X., Wang, S., Qi, C., Chen, J., Ye, C., Zhao, F., Wang, J.,
696 Fan, J., Li, B., Sun, J., Jia, X., Xia, Z., Zhang, H., Liu, J., Zheng, Y., Liu, X., Wang, J., Yang, H., Kristiansen, K.,
697 Xu, X., Mock, T., Li, S., Zhang, W., and Fan, G.: Global marine microbial diversity and its potential in
698 bioprospecting, *Nature*, 633, 371–379, 2024.

699 Cohan, F. M.: What are bacterial species?, *Annu. Rev. Microbiol.*, 56, 457–487, 2002.

700 Conrad, R.: Compensation concentration as critical variable for regulating the flux of trace gases between soil
701 and atmosphere, *Biogeochemistry*, 27, <https://doi.org/10.1007/bf00000582>, 1994.

702 DeAngelis, K. M., Pold, G., Topçuoğlu, B. D., van Diepen, L. T. A., Varney, R. M., Blanchard, J. L., Melillo, J.,
703 and Frey, S. D.: Long-term forest soil warming alters microbial communities in temperate forest soils, *Front.*
704 *Microbiol.*, 6, 104, 2015.

705 Defosse, E., Pitteloud, C., Descombes, P., Glauser, G., Allard, P.-M., Walker, T. W. N., Fernandez-Conradi, P.,
706 Wolfender, J.-L., Pellissier, L., and Rasmann, S.: Spatial and evolutionary predictability of phytochemical
707 diversity, *Proc. Natl. Acad. Sci. U. S. A.*, 118, <https://doi.org/10.1073/pnas.2013344118>, 2021.

708 Defrenne, C. E., Abs, E., Longhi Cordeiro, A., Dietterich, L., Hough, M., Jones, J. M., Kivlin, S. N., Chen, W.,
709 Cusack, D., Franco, A. L. C., Khasanova, A., Stover, D., and Romero-Olivares, A. L.: The Ecology Underground
710 coalition: building a collaborative future of belowground ecology and ecologists, *New Phytol.*, 229, 3058–3064,
711 2021.

712 Delgado-Baquerizo, M., Maestre, F. T., Reich, P. B., Trivedi, P., Osanai, Y., Liu, Y.-R., Hamonts, K., Jeffries, T.
713 C., and Singh, B. K.: Carbon content and climate variability drive global soil bacterial diversity patterns, *Ecol.*
714 *Monogr.*, 86, 373–390, 2016.

715 Demina, T., Marttila, H., Pessi, I. S., Männistö, M. K., Dutilh, B. E., Roux, S., and Hultman, J.: Tunturi virus
716 isolates and metagenome-assembled viral genomes provide insights into the virome of Acidobacteriota in Arctic
717 tundra soils, *Microbiome*, 13, 79, 2025.

718 Doney, S. C., Ruckelshaus, M., Duffy, J. E., Barry, J. P., Chan, F., English, C. A., Galindo, H. M., Grebmeier, J.
719 M., Hollowed, A. B., Knowlton, N., Polovina, J., Rabalais, N. N., Sydeman, W. J., and Talley, L. D.: Climate
720 change impacts on marine ecosystems, *Ann. Rev. Mar. Sci.*, 4, 11–37, 2012.

721 Doumeizel, V. and Dolan, J. R.: The launch of The Plankton Manifesto in September 2024, *J. Plankton Res.*,
722 <https://doi.org/10.1093/plankt/fbae061>, 2024.

723 Duarte, C. M.: Seafaring in the 21st century: The Malaspina 2010 circumnavigation expedition, *Limnol.*
724 *Oceanog. Bull.*, 24, 11–14, 2015.

725 Duckworth, J. C., Kent, M., and Ramsay, P. M.: Plant functional types: an alternative to taxonomic plant
726 community description in biogeography?, *Prog. Phys. Geogr.*, 24, 515–542, 2000.

727 Dussarrat, T., Prigent, S., Latorre, C., Bernillon, S., Flandin, A., Díaz, F. P., Cassan, C., Van Delft, P., Jacob, D.,
728 Varala, K., Joubes, J., Gibon, Y., Rolin, D., Gutiérrez, R. A., and Pétriacq, P.: Predictive metabolomics of
729 multiple Atacama plant species unveils a core set of generic metabolites for extreme climate resilience, *New*
730 *Phytol.*, 234, 1614–1628, 2022.

731 Dussarrat, T., Latorre, C., Barros Santos, M. C., Aguado-Norese, C., Prigent, S., Díaz, F. P., Rolin, D., González,
732 M., Müller, C., Gutiérrez, R. A., and Pétriacq, P.: Rhizochemistry and soil bacterial community are tailored to
733 natural stress gradients, *Soil Biol. Biochem.*, 202, 109662, 2025.

734 Edwards, M. E.: The maturing relationship between Quaternary paleoecology and ancient sedimentary DNA,
735 *Quat. Res.*, 96, 39–47, 2020.

736 Eloë-Fadrosh, E. A., Ahmed, F., Babinski, M., Baumes, J., Borkum, M., Bramer, L., ... & Fagnan, K.: The
737 National Microbiome Data Collaborative Data Portal: an integrated multi-omics microbiome data resource.
738 *Nucleic Acids Research*, 50, D828-D836, 2022.

739 Falkowski, P. G. and Knoll, A. H. (Eds.): *Evolution of primary producers in the sea*, Academic Press, San Diego,
740 CA, 456 pp., 2007.

741 Feinberg, A. P. and Irizarry, R. A.: Evolution in health and medicine Sackler colloquium: Stochastic epigenetic
742 variation as a driving force of development, evolutionary adaptation, and disease, *Proc. Natl. Acad. Sci. U. S. A.*,
743 107 Suppl 1, 1757–1764, 2010.

744 Felden, J., Möller, L., Schindler, U., Huber, R., Schumacher, S., Koppe, R., Diepenbroek, M., and Glöckner, F.
745 O.: PANGAEA - data publisher for earth & environmental science, *Sci. Data*, 10, 347, 2023.

746 Feng, K., Wang, S., He, Q., Bonkowski, M., Bahram, M., Yergeau, E., Wang, Z., Peng, X., Wang, D., Li, S.,
747 Wang, Y., Ju, Z., Du, X., Yan, C., Gu, S., Li, T., Yang, X., Shen, W., Wei, Z., Hu, Q., Li, P., Zhu, Y., Lu, G.,
748 Qin, C., Zhang, G., Xiao, C., Yang, Y., Zhou, J., and Deng, Y.: CoBacFM: Core bacteria forecast model for
749 global grassland pH dynamics under future climate warming scenarios, *One Earth*, 7, 1275–1287, 2024.

750 Fernandez-Guerra, A., Borrel, G., Delmont, T. O., Elberling, B., Eren, A. M., Gribaldo, S., Jochheim, A.,
751 Henriksen, R. A., Hinrichs, K.-U., Korneliussen, T. S., Krupovic, M., Larsen, N. K., Laso-Pérez, R., Pedersen,
752 M. W., Pedersen, V. K., Sand, K. K., Sikora, M., Steinegger, M., Veseli, I., Wörmer, L., Zhao, L., Žure, M.,
753 Kjær, K., and Willerslev, E.: A 2-million-year-old microbial and viral communities from the Kap København
754 Formation in North Greenland, *bioRxiv*, <https://doi.org/10.1101/2023.06.10.544454>, 2023a.

755 Fernandez-Guerra, A., Wörmer, L., Borrel, G., Delmont, T. O., Elberling, B., Elvert, M., Eren, A. M., Gribaldo,
756 S., Henriksen, R. A., Hinrichs, K.-U., Jochheim, A., Korneliussen, T. S., Krupovic, M., Larsen, N. K., Perez-
757 Laso, R., Pedersen, M. W., Pedersen, V. K., Ruter, A. H., Sand, K. K., Sikora, M., Steinegger, M., Veseli, I.,
758 Wang, Y., Zhao, L., Žure, M., Kjær, K. H., and Willerslev, E.: Two-million-year-old microbial communities

759 from the Kap København Formation in North Greenland, *Microbiology*, 2023b.

760 Fierer, N.: Microbial biogeography: Patterns in microbial diversity across space and time, in: *Accessing*
761 *Uncultivated Microorganisms*, ASM Press, Washington, DC, USA, 95–115, 2014.

762 Finks, S. S., Weihe, C., Kimball, S., Allison, S. D., Martiny, A. C., Treseder, K. K., and Martiny, J. B. H.:
763 Microbial community response to a decade of simulated global changes depends on the plant community,
764 *Elementa* (Wash., DC), 9, <https://doi.org/10.1525/elementa.2021.00124>, 2021.

765 Flombaum, P., Gallegos, J. L., Gordillo, R. A., Rincón, J., Zabala, L. L., Jiao, N., Karl, D. M., Li, W. K. W.,
766 Lomas, M. W., Veneziano, D., Vera, C. S., Vrugt, J. A., and Martiny, A. C.: Present and future global
767 distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*, *Proc. Natl. Acad. Sci. U. S. A.*,
768 110, 9824–9829, 2013.

769 Fodelianakis, S., Washburne, A. D., Bourquin, M., Pramateftaki, P., Kohler, T. J., Styllas, M., Tolosano, M., De
770 Staercke, V., Schön, M., Busi, S. B., Brandani, J., Wilmes, P., Peter, H., and Battin, T. J.: Microdiversity
771 characterizes prevalent phylogenetic clades in the glacier-fed stream microbiome, *ISME J.*, 16, 666–675, 2022.

772 Franz, D., Acosta, M., Altimir, N., Arriga, N., Arrouays, D., Aubinet, M., Aurela, M., Ayres, E., López-
773 Ballesteros, A., Barbaste, M., Berveiller, D., Biraud, S., Boukir, H., Brown, T., Brümmer, C., Buchmann, N.,
774 Burba, G., Carrara, A., Cescatti, A., Ceschia, E., Clement, R., Cremonese, E., Crill, P., Darenova, E., Dengel, S.,
775 D’Odorico, P., Filippa, G., Fleck, S., Fratini, G., Fuß, R., Gielen, B., Gogo, S., Grace, J., Graf, A., Grelle, A.,
776 Gross, P., Grünwald, T., Haapanala, S., Hehn, M., Heinesch, B., Heiskanen, J., Herbst, M., Herschlein, C.,
777 Hörtnagl, L., Hufkens, K., Ibrom, A., Jolivet, C., Joly, L., Jones, M., Kiese, R., Klemetsson, L., Kljun, N.,
778 Klumpp, K., Kolari, P., Kolle, O., Kowalski, A., Kutsch, W., Laurila, T., de Ligne, A., Linder, S., Lindroth, A.,
779 Lohila, A., Longdoz, B., Mammarella, I., Manise, T., Jiménez, S. M., Matteucci, G., Mauder, M., Meier, P.,
780 Merbold, L., Mereu, S., Metzger, S., Migliavacca, M., Mölder, M., Montagnani, L., Moureaux, C., Nelson, D.,
781 Nemitz, E., Nicolini, G., Nilsson, M. B., de Beeck, M. O., Osborne, B., Löfvenius, M. O., Pavelka, M., Peichl,
782 M., Peltola, O., Pihlatie, M., Pitacco, A., Pokorný, R., Pumpanen, J., Ratié, C., Rebmann, C., Roland, M.,
783 Sabbatini, S., Saby, N. P. A., Saunders, M., Schmid, H. P., Schrumpf, M., Sedlák, P., et al.: Towards long-term
784 standardised carbon and greenhouse gas observations for monitoring Europe’s terrestrial ecosystems: a review,
785 *Int. Agrophys.*, 32, 439–455, 2018.

786 Fraser, L. H.: TRY-A plant trait database of databases, *Glob. Chang. Biol.*, 26, 189–190, 2020.

787 Frémont, P., Gehlen, M., Vrac, M., Leconte, J., Delmont, T. O., Wincker, P., Iudicone, D., and Jaillon, O.:
788 Restructuring of plankton genomic biogeography in the surface ocean under climate change, *Nat. Clim. Chang.*,
789 12, 393–401, 2022.

790 Frémont, P., Gehlen, M., and Jaillon, O.: Plankton biogeography in the 21st century and impacts of climate
791 change: advances through genomics, *C. R. Biol.*, 346, 13–24, 2023.

792 Friend, A. D., Lucht, W., Rademacher, T. T., Keribin, R., Betts, R., Cadule, P., Ciais, P., Clark, D. B., Dankers,
793 R., Falloon, P. D., Ito, A., Kahana, R., Kleidon, A., Lomas, M. R., Nishina, K., Ostberg, S., Pavlick, R., Peylin,
794 P., Schaphoff, S., Vuichard, N., Warszawski, L., Wiltshire, A., and Woodward, F. I.: Carbon residence time
795 dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO₂, *Proc. Natl.*
796 *Acad. Sci. U. S. A.*, 111, 3280–3285, 2014.

797 Frostegård, Å., Vick, S. H. W., Lim, N. Y. N., Bakken, L. R., and Shapleigh, J. P.: Linking meta-omics to the
798 kinetics of denitrification intermediates reveals pH-dependent causes of N₂O emissions and nitrite accumulation
799 in soil, *ISME J.*, 16, 26–37, 2022.

800 Frugone-Álvarez, M., Contreras, S., Meseguer-Ruiz, O., Tejos, E., Delgado-Huertas, A., Valero-Garcés, B., Díaz,
801 F. P., Briceño, M., Bustos-Morales, M., and Latorre, C.: Hydroclimate variations over the last 17,000 years as
802 estimated by leaf waxes in rodent middens from the south-central Atacama Desert, Chile, *Quat. Sci. Rev.*, 311,
803 108084, 2023.

804 Gao, Q., Wang, G., Xue, K., Yang, Y., Xie, J., Yu, H., Bai, S., Liu, F., He, Z., Ning, D., Hobbie, S. E., Reich, P.
805 B., and Zhou, J.: Stimulation of soil respiration by elevated CO₂ is enhanced under nitrogen limitation in a
806 decade-long grassland study, *Proc. Natl. Acad. Sci. U. S. A.*, 117, 33317–33324, 2020.

807 García-García, N., Tamames, J., Linz, A. M., Pedrós-Alió, C., and Puente-Sánchez, F.: Microdiversity ensures
808 the maintenance of functional microbial communities under changing environmental conditions, *ISME J.*, 13,
809 2969–2983, 2019.

810 Giguet-Covex, C., Pansu, J., Arnaud, F., Rey, P.-J., Griggo, C., Gielly, L., Domaizon, I., Coissac, E., David, F.,
811 Choler, P., Poulénard, J., and Taberlet, P.: Long livestock farming history and human landscape shaping revealed
812 by lake sediment DNA, *Nat. Commun.*, 5, 3211, 2014.

813 Glassman, S. I., Weihe, C., Li, J., Albright, M. B. N., Looby, C. I., Martiny, A. C., Treseder, K. K., Allison, S.
814 D., and Martiny, J. B. H.: Decomposition responses to climate depend on microbial community composition,
815 *Proc. Natl. Acad. Sci. U. S. A.*, 115, 11994–11999, 2018.

816 Grant, R. F.: Simulation of methanogenesis in the mathematical model ecosys, *Soil Biol. Biochem.*, 30, 883–896,
817 1998.

818 Grasso, G., Bianciotto, V., and Marmeisse, R.: Paleomicrobiology: Tracking the past microbial life from single
819 species to entire microbial communities, *Microb. Biotechnol.*, 17, e14390, 2024.

820 Gregory-Eaves, I. and Smol, J. P.: Paleolimnology: Approaches and applications, in: *Wetzel’s Limnology*, edited
821 by: Jones, I. D. and Smol, J. P., Elsevier, 1015–1043, 2024.

822 Griffin, D. W., Gonzalez-Martin, C., Hoose, C., and Smith, D. J.: Global-scale atmospheric dispersion of
823 microorganisms, in: *Microbiology of Aerosols*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 155–194, 2017.

824 Griffith, J., Lord, J.-M., Catchen, M. D., Arce-Plata, M. I., Blanchet, F. G., Chandramohan, M., Diaz-Corzo, M.
825 C., Gravel, D., Gutiérrez, C., Helfenstein, I. S., Hoban, S., Kass, J. M., Laikre, L., Larocque, G., Leigh, D. M.,
826 Leung, B., Mastretta-Yanes, A., Millette, K. L., Berbeo, M. A. M., Nguyen, D., Norman, K. E., Olaya-Rodríguez,
827 M. H., Pahls, S., Pereira, K., Peres-Neto, P. R., Poisot, T., Pollock, L. J., Rey-Velasco, J. C., Rincon-Parra, V. J.,
828 Roesli, C., Rousseu, F., Sánchez-Clavijo, L. M., Schuman, M. C., Selmoni, O., da Silva, J. M., Suarez-
829 Valencia, E., Surasinghe, T. D., Turak, E., Urbina, L. F., Valentin, S., Wightman, N., Zuloaga, J., Londoño, M.
830 C., and Gonzalez, A.: BON in a Box: An open and collaborative platform for biodiversity monitoring, indicator
831 calculation, and reporting, *Bioscience*, <https://doi.org/10.1093/biosci/biaf189>, 2026.

832 Grigoriev, I. V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otilar, R., Riley, R., Salamov, A., Zhao, X.,

833 Korzeniewski, F., Smirnova, T., Nordberg, H., Dubchak, I., and Shabalov, I.: MycoCosm portal: gearing up for
834 1000 fungal genomes, *Nucleic Acids Res.*, 42, D699–704, 2014.

835 Grigoriev, I. V., Hayes, R. D., Calhoun, S., Kamel, B., Wang, A., Ahrendt, S., Dusheyko, S., Nikitin, R., Mondo,
836 S. J., Salamov, A., Shabalov, I., and Kuo, A.: PhycoCosm, a comparative algal genomics resource, *Nucleic Acids*
837 *Res.*, 49, D1004–D1011, 2021.

838 Guerrero-Ramírez, N. R., Mommer, L., Freschet, G. T., Iversen, C. M., McCormack, M. L., Kattge, J., Poorter,
839 H., van der Plas, F., Bergmann, J., Kuyper, T. W., York, L. M., Bruehlheide, H., Laughlin, D. C., Meier, I. C.,
840 Roumet, C., Semchenko, M., Sweeney, C. J., van Ruijven, J., Valverde-Barrantes, O. J., Aubin, I., Catford, J. A.,
841 Manning, P., Martin, A., Milla, R., Minden, V., Pausas, J. G., Smith, S. W., Soudzilovskaia, N. A., Ammer, C.,
842 Butterfield, B., Craine, J., Cornelissen, J. H. C., de Vries, F. T., Isaac, M. E., Kramer, K., König, C., Lamb, E. G.,
843 Onipchenko, V. G., Peñuelas, J., Reich, P. B., Rillig, M. C., Sack, L., Shipley, B., Tedersoo, L., Valladares, F.,
844 van Bodegom, P., Weigelt, P., Wright, J. P., and Weigelt, A.: Global root traits (GRooT) database, *Glob. Ecol.*
845 *Biogeogr.*, 30, 25–37, 2021.

846 Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., Darzi, Y., Audic, S., Berline, L., Brum,
847 J., Coelho, L. P., Espinoza, J. C. I., Malviya, S., Sunagawa, S., Dimier, C., Kandels-Lewis, S., Picheral, M.,
848 Poulain, J., Searson, S., Tara Oceans coordinators, Stemann, L., Not, F., Hingamp, P., Speich, S., Follows, M.,
849 Karp-Boss, L., Boss, E., Ogata, H., Pesant, S., Weissenbach, J., Wincker, P., Acinas, S. G., Bork, P., de Vargas,
850 C., Iudicone, D., Sullivan, M. B., Raes, J., Karsenti, E., Bowler, C., and Gorsky, G.: Plankton networks driving
851 carbon export in the oligotrophic ocean, *Nature*, 532, 465–470, 2016.

852 Heinrichs, M. E., Piedade, G. J., Popa, O., Sommers, P., Trubl, G., Weissenbach, J., and Rahlff, J.: Breaking the
853 ice: A review of phages in polar ecosystems, *Methods Mol. Biol.*, 2738, 31–71, 2024.

854 Heinze, B. M., Schwab, V. F., Küsel, K., Schloemer, S., Roskam, A., Xu, X., and Trumbore, S.: Microbial
855 oxidation significantly reduces methane export from global groundwaters, *Proc. Natl. Acad. Sci. U. S. A.*, 122,
856 e2508773122, 2025.

857 Hermans, S. M., Buckley, H. L., Case, B. S., Curran-Cournane, F., Taylor, M., and Lear, G.: Using soil bacterial
858 communities to predict physico-chemical variables and soil quality, *Microbiome*, 8, 79, 2020.

859 He, X., Abs, E., Allison, S. D., Tao, F., Huang, Y., Manzoni, S., Abramoff, R., Bruni, E., Bowring, S. P. K.,
860 Chakrawal, A., Ciais, P., Elsgaard, L., Friedlingstein, P., Georgiou, K., Hugelius, G., Holm, L. B., Li, W., Luo,
861 Y., Marmasse, G., Nunan, N., Qiu, C., Sitch, S., Wang, Y.-P., and Goll, D. S.: Emerging multiscale insights on
862 microbial carbon use efficiency in the land carbon cycle, *Nat. Commun.*, 15, 8010, 2024a.

863 He, X., Abramoff, R. Z., Abs, E., Georgiou, K., Zhang, H., and Goll, D. S.: Model uncertainty obscures major
864 driver of soil carbon, *Nature*, 627, E1–E3, 2024b.

865 He, X., Marmasse, G., Hu, J., Varney, R. M., Manzoni, S., Ciais, P., Wang, Y.-P., Cui, Y., Bai, E., Abramoff, R.
866 Z., Abs, E., Schwarz, E., Zhang, H., and Goll, D. S.: Microbial growth rate is a stronger predictor of soil organic
867 carbon than carbon use efficiency, *Nat. Ecol. Evol.*, 10, 372–381, 2026.

868 Ibarbalz, F. M., Henry, N., Brandão, M. C., Martini, S., Busseni, G., Byrne, H., Coelho, L. P., Endo, H., Gasol, J.
869 M., Gregory, A. C., Mahé, F., Rigonato, J., Royo-Llonch, M., Salazar, G., Sanz-Sáez, I., Scalco, E., Soviadan, D.,

870 Zayed, A. A., Zingone, A., Labadie, K., Ferland, J., Marec, C., Kandels, S., Picheral, M., Dimier, C., Poulain, J.,
871 Pisarev, S., Carmichael, M., Pesant, S., Tara Oceans Coordinators, Babin, M., Boss, E., Iudicone, D., Jaillon, O.,
872 Acinas, S. G., Ogata, H., Pelletier, E., Stemmann, L., Sullivan, M. B., Sunagawa, S., Bopp, L., de Vargas, C.,
873 Karp-Boss, L., Wincker, P., Lombard, F., Bowler, C., and Zinger, L.: Global trends in marine plankton diversity
874 across kingdoms of life, *Cell*, 179, 1084–1097.e21, 2019.

875 Jia, W., Anslan, S., Chen, F., Cao, X., Dong, H., Dulias, K., Gu, Z., Heinecke, L., Jiang, H., Kruse, S., Kang, W.,
876 Li, K., Liu, S., Liu, X., Liu, Y., Ni, J., Schwalb, A., Stoof-Leichsenring, K. R., Shen, W., Tian, F., Wang, J.,
877 Wang, Y., Wang, Y., Xu, H., Yang, X., Zhang, D., and Herzsuh, U.: Sedimentary ancient DNA reveals past
878 ecosystem and biodiversity changes on the Tibetan Plateau: Overview and prospects, *Quat. Sci. Rev.*, 293,
879 107703, 2022.

880 Johnson, Z. I., Zinser, E. R., Coe, A., McNulty, N. P., Woodward, E. M. S., and Chisholm, S. W.: Niche
881 partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients, *Science*, 311, 1737–
882 1740, 2006.

883 Jones, E. D. and Bösl, E.: Ancient human DNA: A history of hype (then and now), *J. Soc. Archaeol.*, 21, 236–
884 255, 2021.

885 Karaoz, U. and Brodie, E. L.: MicroTrait: A toolset for a trait-based representation of microbial genomes, *Front.*
886 *Bioinform.*, 2, 918853, 2022.

887 Kent, A. G., Garcia, C. A., and Martiny, A. C.: Increased biofilm formation due to high-temperature adaptation in
888 marine *Roseobacter*, <https://doi.org/10.1038/s41564-018-0213-8>, 2018.

889 Keuschnig, C., Vogel, T. M., Barbaro, E., Spolaor, A., Koziol, K., Björkman, M. P., Zdanowicz, C., Gallet, J.-C.,
890 Luks, B., Layton, R., and Larose, C.: Selection processes of Arctic seasonal glacier snowpack bacterial
891 communities, *Microbiome*, 11, 35, 2023.

892 Khaled, A., Zhang, M., Amato, P., Delort, A.-M., and Ervens, B.: Biodegradation by bacteria in clouds: an
893 underestimated sink for some organics in the atmospheric multiphase system, *Atmos. Chem. Phys.*, 21, 3123–
894 3141, 2021.

895 Kivlin, S. N.: Global mycorrhizal fungal range sizes vary within and among mycorrhizal guilds but are not
896 correlated with dispersal traits, *J. Biogeogr.*, 47, 1994–2001, 2020.

897 Kjær, K. H., Winther Pedersen, M., De Sanctis, B., De Cahsan, B., Korneliussen, T. S., Michelsen, C. S., Sand,
898 K. K., Jelavić, S., Ruter, A. H., Schmidt, A. M. A., Kjeldsen, K. K., Tesakov, A. S., Snowball, I., Gosse, J. C.,
899 Alsos, I. G., Wang, Y., Dockter, C., Rasmussen, M., Jørgensen, M. E., Skadhauge, B., Prohaska, A., Kristensen,
900 J. Å., Bjerager, M., Allentoft, M. E., Coissac, E., PhyloNorway Consortium, Rouillard, A., Simakova, A.,
901 Fernandez-Guerra, A., Bowler, C., Macias-Fauria, M., Vinner, L., Welch, J. J., Hidy, A. J., Sikora, M., Collins,
902 M. J., Durbin, R., Larsen, N. K., and Willerslev, E.: A 2-million-year-old ecosystem in Greenland uncovered by
903 environmental DNA, *Nature*, 612, 283–291, 2022.

904 Kumar, M., Ji, B., Zengler, K., and Nielsen, J.: Modelling approaches for studying the microbiome, *Nat.*
905 *Microbiol.*, 4, 1253–1267, 2019.

906 Küsel, K., Totsche, K. U., Trumbore, S. E., Lehmann, R., Steinhäuser, C., and Herrmann, M.: How deep can
907 surface signals be traced in the critical zone? Merging biodiversity with biogeochemistry research in a central
908 German muschelkalk landscape, *Front. Earth Sci.*, 4, <https://doi.org/10.3389/feart.2016.00032>, 2016.

909 Kuske, C. R., Barns, S. M., Grow, C. C., Merrill, L., and Dunbar, J.: Environmental survey for four pathogenic
910 bacteria and closely related species using phylogenetic and functional genes, *J. Forensic Sci.*, 51, 548–558, 2006.

911 Lamour, J., Serbin, S. P., Rogers, A., Acebron, K. T., Ainsworth, E., Albert, L. P., Alonzo, M., Anderson, J.,
912 Atkin, O. K., Barbier, N., Barnes, M. L., Bernacchi, C. J., Besson, N., Burnett, A. C., Caplan, J. S., Chave, J.,
913 Cheesman, A. W., Clocher, I., Coast, O., Coste, S., Croft, H., Cui, B., Dauvissat, C., Davidson, K. J., Doughty,
914 C., Ely, K. S., Evans, J. R., Féret, J.-B., Filella, I., Fortunel, C., Fu, P., Furbank, R. T., Garcia, M., Gimenez, B.
915 O., Guan, K., Guo, Z., Heckmann, D., Heuret, P., Isaac, M., Kothari, S., Kumagai, E., Kyaw, T. Y., Liu, L., Liu,
916 L., Liu, S., Llusà, J., Magney, T., Maréchaux, I., Martin, A. R., Meacham-Hensold, K., Montes, C. M., Ogaya,
917 R., Ojo, J., Oliveira, R., Paquette, A., Peñuelas, J., Placido, A. D., Posada, J. M., Qian, X., Renninger, H. J.,
918 Rodriguez-Caton, M., Rojas-González, A., Schlüter, U., Sellan, G., Siegert, C. M., Silva-Perez, V., Song, G.,
919 Southwick, C. D., Souza, D. C., Stahl, C., Su, Y., Sujeeun, L., Ting, T.-C., Vasquez, V., Vijayakumar, A., Vilas-
920 Boas, M., Wang, D. R., Wang, S., Wang, H., Wang, J., Wang, X., Weber, A. P. M., Wong, C. Y. S., Wu, J., Wu,
921 F., Wu, S., Yan, Z., Yang, D., and Zhao, Y.: The Global Spectra-Trait Initiative: A database of paired leaf
922 spectroscopy and functional traits associated with leaf photosynthetic capacity, *Earth Syst. Sci. Data*, 18, 245–
923 265, 2026.

924 Larkin, A. A. and Martiny, A. C.: Microdiversity shapes the traits, niche space, and biogeography of microbial
925 taxa, *Environ. Microbiol. Rep.*, 9, 55–70, 2017.

926 Lelieveld, J. and Crutzen, P. J.: Influences of cloud photochemical processes on tropospheric ozone, *Nature*, 343,
927 227–233, 1990.

928 Lewis, S. L. and Maslin, M. A.: A transparent framework for defining the Anthropocene Epoch, *Anthr. Rev.*, 2,
929 128–146, 2015.

930 Li, J., Wang, G., Mayes, M. A., Allison, S. D., Frey, S. D., Shi, Z., Hu, X.-M., Luo, Y., and Melillo, J. M.:
931 Reduced carbon use efficiency and increased microbial turnover with soil warming, *Glob. Chang. Biol.*, 25, 900–
932 910, 2019.

933 Li, Z., Riley, W. J., Marschmann, G. L., Karaoz, U., Shirley, I. A., Wu, Q., Bouskill, N. J., Chang, K.-Y., Crill, P.
934 M., Grant, R. F., King, E., Saleska, S. R., Sullivan, M. B., Tang, J., Varner, R. K., Woodcroft, B. J., Wrighton, K.
935 C., EMERGE Biology Integration Institute Coordinators, and Brodie, E. L.: A framework for integrating
936 genomics, microbial traits, and ecosystem biogeochemistry, *Nat. Commun.*, 16, 2186, 2025.

937 Lovelock, J. E. and Margulis, L.: Atmospheric homeostasis by and for the biosphere: the gaia hypothesis, *Tellus*
938 A, 26, 2, 1974.

939 Lynggaard, C., Bertelsen, M. F., Jensen, C. V., Johnson, M. S., Frøslev, T. G., Olsen, M. T., and Bohmann, K.:
940 Airborne environmental DNA for terrestrial vertebrate community monitoring, *Curr. Biol.*, 32, 701–707.e5, 2022.

941 Ma, B., Lu, C., Wang, Y., Yu, J., Zhao, K., Xue, R., Ren, H., Lv, X., Pan, R., Zhang, J., Zhu, Y., and Xu, J.: A
942 genomic catalogue of soil microbiomes boosts mining of biodiversity and genetic resources, *Nat. Commun.*, 14,

7318, 2023.

Maeke, M. D., Yin, X., Wunder, L. C., Vanni, C., Richter-Heitmann, T., Miravet-Verde, S., Ruscheweyh, H.-J., Sunagawa, S., Fabian, J., Piontek, J., Friedrich, M. W., and Hassenrück, C.: Extensive data mining uncovers novel diversity among members of the rare biosphere within the Thermoplasmata, *Microbiome*, 13, 155, 2025.

Männistö, M. K., Ahonen, S. H. K., Ganzert, L., Tirola, M., Stark, S., and Häggblom, M. M.: Bacterial and fungal communities in sub-Arctic tundra heaths are shaped by contrasting snow accumulation and nutrient availability, *FEMS Microbiol. Ecol.*, 100, <https://doi.org/10.1093/femsec/fiae036>, 2024.

Marschmann, G. L., Tang, J., Zhalnina, K., Karaoz, U., Cho, H., Le, B., Pett-Ridge, J., and Brodie, E. L.: Predictions of rhizosphere microbiome dynamics with a genome-informed and trait-based energy budget model, *Nat. Microbiol.*, 9, 421–433, 2024.

Martiny, J. B. H., Jones, S. E., Lennon, J. T., and Martiny, A. C.: Microbiomes in light of traits: A phylogenetic perspective, *Science*, 350, aac9323, 2015.

Martiny, J. B. H., Martiny, A. C., Brodie, E., Chase, A. B., Rodríguez-Verdugo, A., Treseder, K. K., and Allison, S. D.: Investigating the eco-evolutionary response of microbiomes to environmental change, *Ecol. Lett.*, <https://doi.org/10.1111/ele.14209>, 2023.

Matthaeus, W. J., Macarewich, S. I., Richey, J. D., Wilson, J. P., McElwain, J. C., Montañez, I. P., DiMichele, W. A., Hren, M. T., Poulsen, C. J., and White, J. D.: Freeze tolerance influenced forest cover and hydrology during the Pennsylvanian, *Proc. Natl. Acad. Sci. U. S. A.*, 118, e2025227118, 2021.

McCalley, C. K., Woodcroft, B. J., Hodgkins, S. B., Wehr, R. A., Kim, E.-H., Mondav, R., Crill, P. M., Chanton, J. P., Rich, V. I., Tyson, G. W., and Saleska, S. R.: Methane dynamics regulated by microbial community response to permafrost thaw, *Nature*, 514, 478–481, 2014.

McElwain, J. C., Matthaeus, W. J., Barbosa, C., Chondrogiannis, C., O’Dea, K., Jackson, B., Knetge, A. B., Kwasniewska, K., Nair, R., White, J. D., Wilson, J. P., Montañez, I. P., Buckley, Y. M., Belcher, C. M., and Nogué, S.: Functional traits of fossil plants, *New Phytol.*, 242, 392–423, 2024.

McLaren, M. R. and Callahan, B. J.: In nature, there is only diversity, *MBio*, 9, <https://doi.org/10.1128/mBio.02149-17>, 2018.

Métris, K. L. and Métris, J.: Aircraft surveys for air eDNA: probing biodiversity in the sky, *PeerJ*, 11, e15171, 2023.

Meziti, A., Rodriguez-R, L. M., Hatt, J. K., Peña-Gonzalez, A., Levy, K., and Konstantinidis, K. T.: The Reliability of Metagenome-Assembled Genomes (MAGs) in Representing Natural Populations: Insights from Comparing MAGs against Isolate Genomes Derived from the Same Fecal Sample, *Appl. Environ. Microbiol.*, 87, <https://doi.org/10.1128/AEM.02593-20>, 2021.

Monson, R. K. and Holland, E. A.: Biospheric trace gas fluxes and their control over tropospheric chemistry, *Annu. Rev. Ecol. Syst.*, 32, 547–576, 2001.

- 977 Müller, C. and Junker, R. R.: Chemical phenotype as important and dynamic niche dimension of plants, *New*
978 *Phytol.*, 234, 1168–1174, 2022.
- 979 Nayfach, S., Rodriguez-Mueller, B., Garud, N., and Pollard, K. S.: An integrated metagenomics pipeline for
980 strain profiling reveals novel patterns of bacterial transmission and biogeography, *Genome Res.*, 26, 1612–1625,
981 2016.
- 982 Nayfach, S., Roux, S., Seshadri, R., Udworthy, D., Varghese, N., Schulz, F., Wu, D., Paez-Espino, D., Chen, I.-M.,
983 Huntemann, M., Palaniappan, K., Ladau, J., Mukherjee, S., Reddy, T. B. K., Nielsen, T., Kirton, E., Faria, J. P.,
984 Edirisinghe, J. N., Henry, C. S., Jungbluth, S. P., Chivian, D., Dehal, P., Wood-Charlson, E. M., Arkin, A. P.,
985 Tringe, S. G., Visel, A., IMG/M Data Consortium, Woyke, T., Mouncey, N. J., Ivanova, N. N., Kyrpides, N. C.,
986 and Elie-Fadrosh, E. A.: A genomic catalog of Earth’s microbiomes, *Nat. Biotechnol.*, 39, 499–509, 2021.
- 987 Nguyen, N.-L., Devendra, D., Szymańska, N., Greco, M., Angeles, I. B., Weiner, A. K. M., Ray, J. L., Cordier,
988 T., De Schepper, S., Pawłowski, J., and Pawłowska, J.: Sedimentary ancient DNA: a new paleogenomic tool for
989 reconstructing the history of marine ecosystems, *Front. Mar. Sci.*, 10,
990 <https://doi.org/10.3389/fmars.2023.1185435>, 2023.
- 991 Nuñez López, L., Amato, P., and Ervens, B.: Bacteria in clouds biodegrade atmospheric formic and acetic acids,
992 *Atmos. Chem. Phys.*, 24, 5181–5198, 2024.
- 993 Osburn, E. D., McBride, S. G., and Strickland, M. S.: Microbial dark matter could add uncertainties to
994 metagenomic trait estimations, *Nat. Microbiol.*, 9, 1427–1430, 2024.
- 995 Overholt, W. A., Trumbore, S., Xu, X., Bornemann, T. L. V., Probst, A. J., Krüger, M., Herrmann, M.,
996 Thamdrup, B., Bristow, L. A., Taubert, M., Schwab, V. F., Hölzer, M., Marz, M., and Küsel, K.: Carbon fixation
997 rates in groundwater similar to those in oligotrophic marine systems, *Nat. Geosci.*, 15, 561–567, 2022.
- 998 Paoli, L., Ruscheweyh, H.-J., Forneris, C. C., Hubrich, F., Kautsar, S., Bhushan, A., Lotti, A., Clayssen, Q.,
999 Salazar, G., Milanese, A., Carlström, C. I., Papadopoulou, C., Gehrig, D., Karasikov, M., Mustafa, H., Larralde,
1000 M., Carroll, L. M., Sánchez, P., Zayed, A. A., Cronin, D. R., Acinas, S. G., Bork, P., Bowler, C., Delmont, T. O.,
1001 Gasol, J. M., Gossert, A. D., Kahles, A., Sullivan, M. B., Wincker, P., Zeller, G., Robinson, S. L., Piel, J., and
1002 Sunagawa, S.: Biosynthetic potential of the global ocean microbiome, *Nature*, 607, 111–118, 2022.
- 1003 Parks, D. H., Chuvochina, M., Rinke, C., Mussig, A. J., Chaumeil, P.-A., and Hugenholtz, P.: GTDB: an ongoing
1004 census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete
1005 genome-based taxonomy, *Nucleic Acids Res.*, 50, D785–D794, 2022.
- 1006 Pasteur, L.: Mémoire sur les corpuscles organisés qui existent dans l’atmosphère, examen de la doctrine des
1007 générations spontanées, 1861.
- 1008 Péguilhan, R., Besaury, L., Rossi, F., Enault, F., Baray, J.-L., Deguillaume, L., and Amato, P.: Rainfalls sprinkle
1009 cloud bacterial diversity while scavenging biomass, *FEMS Microbiol. Ecol.*, 97,
1010 <https://doi.org/10.1093/femsec/fiab144>, 2021.
- 1011 Péguilhan, R., Rossi, F., Joly, M., Nasr, E., Batut, B., Enault, F., Ervens, B., and Amato, P.: Clouds influence the
1012 functioning of airborne microorganisms, *Biogeosciences*, 22, 1257–1275, 2025.

1013 Pessi, I. S., Viitamäki, S., Virkkala, A.-M., Eronen-Rasimus, E., Delmont, T. O., Marushchak, M. E., Luoto, M.,
1014 and Hultman, J.: In-depth characterization of denitrifier communities across different soil ecosystems in the
1015 tundra, *Environ. Microbiome*, 17, 30, 2022.

1016 Pettersson, K. J., Demina, T., Eronen-Rasimus, E., Roux, S., Viitamäki, S., Pessi, I. S., Oksanen, H. M., Assmy,
1017 P., Kaartokallio, H., and Hultman, J.: Viral genetic diversity and functional potential in polar and subarctic sea
1018 ice, *bioRxiv*, 2025.12.12.693969, <https://doi.org/10.64898/2025.12.12.693969>, 2025.

1019 Piton, G., Allison, S. D., Bahram, M., Hildebrand, F., Martiny, J. B. H., Treseder, K. K., and Martiny, A. C.: Life
1020 history strategies of soil bacterial communities across global terrestrial biomes, *Nat. Microbiol.*, 8, 2093–2102,
1021 2023.

1022 Poppeliers, S. W. M., Hefting, M., Dorrepaal, E., and Weedon, J. T.: Functional microbial ecology in arctic soils:
1023 the need for a year-round perspective, *FEMS Microbiol. Ecol.*, 98, <https://doi.org/10.1093/femsec/fiac134>, 2022.

1024 Post, E., Alley, R. B., Christensen, T. R., Macias-Fauria, M., Forbes, B. C., Gooseff, M. N., Iler, A., Kerby, J. T.,
1025 Laidre, K. L., Mann, M. E., Olofsson, J., Stroeve, J. C., Ulmer, F., Virginia, R. A., and Wang, M.: The polar
1026 regions in a 2°C warmer world, *Sci. Adv.*, 5, eaaw9883, 2019.

1027 Ramirez Flores, R. O., Lanzer, J. D., Dimitrov, D., Velten, B., and Saez-Rodriguez, J.: Multicellular factor
1028 analysis of single-cell data for a tissue-centric understanding of disease, *Elife*, 12,
1029 <https://doi.org/10.7554/eLife.93161>, 2023.

1030 Ramoneda, J., Stallard-Olivera, E., Hoffert, M., Winfrey, C. C., Stadler, M., Niño-García, J. P., and Fierer, N.:
1031 Building a genome-based understanding of bacterial pH preferences, *Sci. Adv.*, 9, eadf8998, 2023.

1032 Ranjan Hatua, S. and Madalli, D. P.: AERIS: an integrated domain information system for aerospace science and
1033 technology, *Program*, 45, 199–212, 2011.

1034 Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., Vihma, T., and
1035 Laaksonen, A.: The Arctic has warmed nearly four times faster than the globe since 1979, *Commun. Earth*
1036 *Environ.*, 3, <https://doi.org/10.1038/s43247-022-00498-3>, 2022.

1037 Riley, W. J., Zhu, Q., and Tang, J. Y.: Weaker land–climate feedbacks from nutrient uptake during
1038 photosynthesis-inactive periods, *Nat. Clim. Chang.*, 8, 1002–1006, 2018.

1039 Riley, W. J., Mekonnen, Z. A., Tang, J., Zhu, Q., Bouskill, N. J., and Grant, R. F.: Non-growing season plant
1040 nutrient uptake controls Arctic tundra vegetation composition under future climate, *Environ. Res. Lett.*, 16,
1041 074047, 2021.

1042 Rockström, J., Gupta, J., Qin, D., Lade, S. J., Abrams, J. F., Andersen, L. S., Armstrong McKay, D. I., Bai, X.,
1043 Bala, G., Bunn, S. E., Ciobanu, D., DeClerck, F., Ebi, K., Gifford, L., Gordon, C., Hasan, S., Kanie, N., Lenton,
1044 T. M., Loriani, S., Liverman, D. M., Mohamed, A., Nakicenovic, N., Obura, D., Ospina, D., Prodani, K.,
1045 Rammelt, C., Sakschewski, B., Scholtens, J., Stewart-Koster, B., Tharammal, T., van Vuuren, D., Verburg, P. H.,
1046 Winkelmann, R., Zimm, C., Bennett, E. M., Bringezu, S., Broadgate, W., Green, P. A., Huang, L., Jacobson, L.,
1047 Ndehedehe, C., Pedde, S., Rocha, J., Scheffer, M., Schulte-Uebbing, L., de Vries, W., Xiao, C., Xu, C., Xu, X.,
1048 Zafra-Calvo, N., and Zhang, X.: Safe and just Earth system boundaries, *Nature*, 619, 102–111, 2023.

1049 Sanchez-Tello, J. D. and Corrales, A.: Ectomycorrhizal fungal communities in natural and urban ecosystems:
1050 *Quercus humboldtii* as a study case in the tropical Andes, *Mycorrhiza*, 34, 45–55, 2024.

1051 Šantl-Temkiv, T., Amato, P., Casamayor, E. O., Lee, P. K. H., and Pointing, S. B.: Microbial ecology of the
1052 atmosphere, *FEMS Microbiol. Rev.*, 46, <https://doi.org/10.1093/femsre/fuac009>, 2022.

1053 Scales, N. C., Chase, A. B., Finks, S. S., Malik, A. A., Weihe, C., Allison, S. D., Martiny, A. C., and Martiny, J.
1054 B. H.: Differential response of bacterial microdiversity to simulated global change, *Appl. Environ. Microbiol.*,
1055 88, e0242921, 2022.

1056 Schmid, M. W., Heichinger, C., Coman Schmid, D., Guthörl, D., Gagliardini, V., Bruggmann, R., Aluri, S.,
1057 Aquino, C., Schmid, B., Turnbull, L. A., and Grossniklaus, U.: Contribution of epigenetic variation to adaptation
1058 in *Arabidopsis*, *Nat. Commun.*, 9, 4446, 2018.

1059 Schmidt, T. S. B., Fullam, A., Ferretti, P., Orakov, A., Maistrenko, O. M., Ruscheweyh, H.-J., Letunic, I., Duan,
1060 Y., Van Rossum, T., Sunagawa, S., Mende, D. R., Finn, R. D., Kuhn, M., Pedro Coelho, L., and Bork, P.: SPIRE:
1061 A Searchable, Planetary-scale mIcrobome REsource, *Nucleic Acids Res.*, 52, D777–D783, 2024.

1062 Schwarz, E., Abs, E., Chakrawal, A., Chavez Rodriguez, L., Quévreux, P., and Manzoni, S.: Eco-evolutionary
1063 optimality in soil organic matter models, *Ecol. Lett.*, 28, e70278, 2025.

1064 Shaffer, J. P., Nothias, L.-F., Thompson, L. R., Sanders, J. G., Salido, R. A., Couvillion, S. P., Brejnrod, A. D.,
1065 Lejzerowicz, F., Haiminen, N., Huang, S., Lutz, H. L., Zhu, Q., Martino, C., Morton, J. T., Karthikeyan, S.,
1066 Nothias-Esposito, M., Dührkop, K., Böcker, S., Kim, H. W., Aksenov, A. A., Bittremieux, W., Minich, J. J.,
1067 Marotz, C., Bryant, M. M., Sanders, K., Schwartz, T., Humphrey, G., Vásquez-Baeza, Y., Tripathi, A., Parida, L.,
1068 Carrieri, A. P., Beck, K. L., Das, P., González, A., McDonald, D., Ladau, J., Karst, S. M., Albertsen, M.,
1069 Ackermann, G., DeReus, J., Thomas, T., Petras, D., Shade, A., Stegen, J., Song, S. J., Metz, T. O., Swafford, A.
1070 D., Dorrestein, P. C., Jansson, J. K., Gilbert, J. A., Knight, R., and Earth Microbiome Project 500 (EMP500)
1071 Consortium: Standardized multi-omics of Earth’s microbiomes reveals microbial and metabolite diversity, *Nat.*
1072 *Microbiol.*, 7, 2128–2150, 2022.

1073 Shaffer, M., Borton, M. A., McGivern, B. B., Zayed, A. A., La Rosa, S. L., Solden, L. M., Liu, P., Narrowe, A.
1074 B., Rodríguez-Ramos, J., Bolduc, B., Gazitúa, M. C., Daly, R. A., Smith, G. J., Vik, D. R., Pope, P. B., Sullivan,
1075 M. B., Roux, S., and Wrighton, K. C.: DRAM for distilling microbial metabolism to automate the curation of
1076 microbiome function, *Nucleic Acids Res.*, 48, 8883–8900, 2020.

1077 Shaked, Y., Twining, B. S., Tagliabue, A., and Maldonado, M. T.: Probing the bioavailability of dissolved iron to
1078 marine eukaryotic phytoplankton using in situ single cell iron quotas, *Global Biogeochem. Cycles*, 35,
1079 e2021GB006979, 2021.

1080 Shapiro, B. J., Leducq, J.-B., and Mallet, J.: What is speciation?, *PLoS Genet.*, 12, e1005860, 2016.

1081 Sokol, N. W., Slessarev, E., Marschmann, G. L., Nicolas, A., Blazewicz, S. J., Brodie, E. L., Firestone, M. K.,
1082 Foley, M. M., Hestrin, R., Hungate, B. A., Koch, B. J., Stone, B. W., Sullivan, M. B., Zablocki, O., LLNL Soil
1083 Microbiome Consortium, and Pett-Ridge, J.: Life and death in the soil microbiome: how ecological processes
1084 influence biogeochemistry, *Nat. Rev. Microbiol.*, 20, 415–430, 2022.

1085 Soudzilovskaia, N. A., Vaessen, S., Barcelo, M., He, J., Rahimlou, S., Abarenkov, K., Brundrett, M. C., Gomes,
1086 S. I. F., Merckx, V., and Tedersoo, L.: FungalRoot: global online database of plant mycorrhizal associations,
1087 New Phytol., 227, 955–966, 2020.

1088 Steen, A. D., Crits-Christoph, A., Carini, P., DeAngelis, K. M., Fierer, N., Lloyd, K. G., and Cameron Thrash, J.:
1089 High proportions of bacteria and archaea across most biomes remain uncultured, ISME J., 13, 3126–3130, 2019.

1090 Sunagawa, S., Acinas, S. G., Bork, P., Bowler, C., Tara Oceans Coordinators, Eveillard, D., Gorsky, G., Guidi,
1091 L., Iudicone, D., Karsenti, E., Lombard, F., Ogata, H., Pesant, S., Sullivan, M. B., Wincker, P., and de Vargas,
1092 C.: Tara Oceans: towards global ocean ecosystems biology, Nat. Rev. Microbiol., 18, 428–445, 2020.

1093 Tabares, X., Zimmermann, H., Dietze, E., Ratzmann, G., Belz, L., Vieth-Hillebrand, A., Dupont, L., Wilkes, H.,
1094 Mapani, B., and Herzsuh, U.: Vegetation state changes in the course of shrub encroachment in an African
1095 savanna since about 1850 CE and their potential drivers, Ecol. Evol., 10, 962–979, 2020.

1096 Tagliabue, A., Bowie, A. R., Boyd, P. W., Buck, K. N., Johnson, K. S., and Saito, M. A.: The integral role of iron
1097 in ocean biogeochemistry, Nature, 543, 51–59, 2017.

1098 Tara Ocean Foundation, Tara Oceans, European Molecular Biology Laboratory (EMBL), and European Marine
1099 Biological Resource Centre - European Research Infrastructure Consortium (EMBRC-ERIC): Priorities for ocean
1100 microbiome research, Nat. Microbiol., 7, 937–947, 2022.

1101 Taylor, L., Banwart, S., Leake, J., and Beerling, D. J.: Modeling the evolutionary rise of ectomycorrhiza on sub-
1102 surface weathering environments and the geochemical carbon cycle, Am. J. Sci., 311, 369–403, 2011.

1103 Tenaillon, O., Rodríguez-Verdugo, A., Gaut, R. L., McDonald, P., Bennett, A. F., Long, A. D., and Gaut, B. S.:
1104 The molecular diversity of adaptive convergence, Science, 335, 457–461, 2012.

1105 Thibault, K. M., Laney, C. M., Yule, K. M., Franz, N. M., and Mabee, P. M.: The US National Ecological
1106 Observatory Network and the Global Biodiversity Framework: national research infrastructure with a global
1107 reach, J. Ecol. Environ., 47, <https://doi.org/10.5141/jee.23.076>, 2023.

1108 Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J., Prill, R. J., Tripathi, A.,
1109 Gibbons, S. M., Ackermann, G., Navas-Molina, J. A., Janssen, S., Kopylova, E., Vázquez-Baeza, Y., González,
1110 A., Morton, J. T., Mirarab, S., Zech Xu, Z., Jiang, L., Haroon, M. F., Kanbar, J., Zhu, Q., Jin Song, S., Kosciolk,
1111 T., Bokulich, N. A., Lefler, J., Brislawn, C. J., Humphrey, G., Owens, S. M., Hampton-Marcell, J., Berg-Lyons,
1112 D., McKenzie, V., Fierer, N., Fuhrman, J. A., Clauet, A., Stevens, R. L., Shade, A., Pollard, K. S., Goodwin, K.
1113 D., Jansson, J. K., Gilbert, J. A., Knight, R., and Earth Microbiome Project Consortium: A communal catalogue
1114 reveals Earth’s multiscale microbial diversity, Nature, 551, 457–463, 2017.

1115 Tian, F., Wainaina, J. M., Howard-Varona, C., Domínguez-Huerta, G., Bolduc, B., Gazitúa, M. C., Smith, G.,
1116 Gittrich, M. R., Zablocki, O., Cronin, D. R., Eveillard, D., Hallam, S. J., and Sullivan, M. B.: Prokaryotic-virus-
1117 encoded auxiliary metabolic genes throughout the global oceans, Microbiome, 12, 159, 2024.

1118 Tóth, G., Jones, A., and Montanarella, L.: The LUCAS topsoil database and derived information on the regional
1119 variability of cropland topsoil properties in the European Union, Environ. Monit. Assess., 185, 7409–7425, 2013.

1120 Trowbridge, J., Weller, R., Kelley, D., Dever, E., Plueddemann, A., Barth, J. A., and Kawka, O.: The ocean
1121 observatories initiative, *Front. Mar. Sci.*, 6, <https://doi.org/10.3389/fmars.2019.00074>, 2019.

1122 Twining, B. S. and Baines, S. B.: The trace metal composition of marine phytoplankton, *Ann. Rev. Mar. Sci.*, 5,
1123 191–215, 2013.

1124 Tyson, G. W., Chapman, J., Hugenholtz, P., Allen, E. E., Ram, R. J., Richardson, P. M., Solovyev, V. V., Rubin,
1125 E. M., Rokhsar, D. S., and Banfield, J. F.: Community structure and metabolism through reconstruction of
1126 microbial genomes from the environment, *Nature*, 428, 37–43, 2004.

1127 Urban, M. C., Strauss, S. Y., Pelletier, F., Palkovacs, E. P., Leibold, M. A., Hendry, A. P., De Meester, L.,
1128 Carlson, S. M., Angert, A. L., and Giery, S. T.: Evolutionary origins for ecological patterns in space, *Proc. Natl.*
1129 *Acad. Sci. U. S. A.*, 117, 17482–17490, 2020.

1130 Ustick, L. J., Larkin, A. A., and Martiny, A. C.: Global scale phylogeography of functional traits and
1131 microdiversity in *Prochlorococcus*, *ISME J.*, 17, 1671–1679, 2023.

1132 Van Nuland, M. E., Averill, C., Stewart, J. D., Prylutskyi, O., Corrales, A., van Galen, L. G., Manley, B. F., Qin,
1133 C., Lauber, T., Mikryukov, V., Dulia, O., Furci, G., Marín, C., Sheldrake, M., Weedon, J. T., Peay, K. G.,
1134 Cornwallis, C. K., Větrovský, T., Kohout, P., Baldrian, P., Tedersoo, L., West, S. A., Crowther, T. W., Kiers, E.
1135 T., SPUN Mapping Consortium, and van den Hoogen, J.: Global hotspots of mycorrhizal fungal richness are
1136 poorly protected, *Nature*, 645, 414–422, 2025.

1137 Viitamäki, S., Pessi, I. S., Virkkala, A.-M., Niittynen, P., Kemppinen, J., Eronen-Rasimus, E., Luoto, M., and
1138 Hultman, J.: The activity and functions of soil microbial communities in the Finnish sub-Arctic vary across
1139 vegetation types, *FEMS Microbiol. Ecol.*, 98, <https://doi.org/10.1093/femsec/fiac079>, 2022.

1140 Wakeham, S. G.: Organic biogeochemistry in the oxygen-deficient ocean: A review, *Org. Geochem.*, 149,
1141 104096, 2020.

1142 Walter Anthony, K. M., Anthony, P., Hasson, N., Edgar, C., Sivan, O., Eliani-Russak, E., Bergman, O., Minsley,
1143 B. J., James, S. R., Pastick, N. J., Kholodov, A., Zimov, S., Euskirchen, E., Bret-Harte, M. S., Grosse, G., Langer,
1144 M., and Nitzbon, J.: Upland Yedoma taliks are an unpredicted source of atmospheric methane, *Nat. Commun.*,
1145 15, 6056, 2024.

1146 Wang, G., Gao, Q., Yang, Y., Hobbie, S. E., Reich, P. B., and Zhou, J.: Soil enzymes as indicators of soil
1147 function: A step toward greater realism in microbial ecological modeling, *Glob. Chang. Biol.*, 28, 1935–1950,
1148 2022a.

1149 Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., Nguyen, D. D., Watrous, J., Kapono,
1150 C. A., Luzzatto-Knaan, T., Porto, C., Bouslimani, A., Melnik, A. V., Meehan, M. J., Liu, W.-T., Crüsemann, M.,
1151 Boudreau, P. D., Esquenazi, E., Sandoval-Calderón, M., Kersten, R. D., Pace, L. A., Quinn, R. A., Duncan, K. R.,
1152 Hsu, C.-C., Floros, D. J., Gavilan, R. G., Kleigrew, K., Northen, T., Dutton, R. J., Parrot, D., Carlson, E. E.,
1153 Aigle, B., Michelsen, C. F., Jelsbak, L., Sohlenkamp, C., Pevzner, P., Edlund, A., McLean, J., Piel, J., Murphy,
1154 B. T., Gerwick, L., Liaw, C.-C., Yang, Y.-L., Humpf, H.-U., Maansson, M., Keyzers, R. A., Sims, A. C.,
1155 Johnson, A. R., Sidebottom, A. M., Sedio, B. E., Klitgaard, A., Larson, C. B., P, C. A. B., Torres-Mendoza, D.,
1156 Gonzalez, D. J., Silva, D. B., Marques, L. M., Demarque, D. P., Pociute, E., O'Neill, E. C., Briand, E., Helfrich,

1157 E. J. N., Granatosky, E. A., Glukhov, E., Ryffel, F., Houson, H., Mohimani, H., Kharbush, J. J., Zeng, Y.,
1158 Vorholt, J. A., Kurita, K. L., Charusanti, P., McPhail, K. L., Nielsen, K. F., Vuong, L., Elfeki, M., Traxler, M. F.,
1159 Engene, N., Koyama, N., Vining, O. B., Baric, R., Silva, R. R., Mascuch, S. J., Tomasi, S., Jenkins, S., Macherla,
1160 V., Hoffman, T., Agarwal, V., Williams, P. G., Dai, J., Neupane, R., Gurr, J., Rodríguez, A. M. C., Lamsa, A.,
1161 Zhang, C., Dorrestein, K., Duggan, B. M., Almaliti, J., Allard, P.-M., et al.: Sharing and community curation of
1162 mass spectrometry data with Global Natural Products Social Molecular Networking, *Nat. Biotechnol.*, 34, 828–
1163 837, 2016.

1164 Wang, Y., Liu, Y., Wu, Y., Wu, N., Liu, W., and Wang, X.: Heterogeneity of soil bacterial and bacteriophage
1165 communities in three rice agroecosystems and potential impacts of bacteriophage on nutrient cycling, *Environ.*
1166 *Microbiome*, 17, 17, 2022b.

1167 Warinner, C.: An archaeology of microbes, *J. Anthropol. Res.*, 000–000, 2022.

1168 Warinner, C., Rodrigues, J. F. M., Vyas, R., Trachsel, C., Shved, N., Grossmann, J., Radini, A., Hancock, Y.,
1169 Tito, R. Y., Fiddymment, S., Speller, C., Hendy, J., Charlton, S., Luder, H. U., Salazar-García, D. C., Eppler, E.,
1170 Seiler, R., Hansen, L. H., Castruita, J. A. S., Barkow-Oesterreicher, S., Teoh, K. Y., Kelstrup, C. D., Olsen, J. V.,
1171 Nanni, P., Kawai, T., Willerslev, E., von Mering, C., Lewis, C. M., Jr, Collins, M. J., Gilbert, M. T. P., Rühli, F.,
1172 and Cappellini, E.: Pathogens and host immunity in the ancient human oral cavity, *Nat. Genet.*, 46, 336–344,
1173 2014.

1174 Warinner, C., Herbig, A., Mann, A., Fellows Yates, J. A., Weiß, C. L., Burbano, H. A., Orlando, L., and Krause,
1175 J.: A robust framework for microbial archaeology, *Annu. Rev. Genomics Hum. Genet.*, 18, 321–356, 2017.

1176 Waters, C. N., Zalasiewicz, J., Summerhayes, C., Barnosky, A. D., Poirier, C., Gałuszka, A., Cearreta, A.,
1177 Edgeworth, M., Ellis, E. C., Ellis, M., Jeandel, C., Leinfelder, R., McNeill, J. R., Richter, D. D., Steffen, W.,
1178 Syvitski, J., Vidas, D., Waple, M., Williams, M., Zhisheng, A., Grinevald, J., Odada, E., Oreskes, N., and
1179 Wolfe, A. P.: The Anthropocene is functionally and stratigraphically distinct from the Holocene, *Science*, 351,
1180 aad2622, 2016.

1181 Wegner, C.-E., Stahl, R., Velsko, I., Hübner, A., Fagernäs, Z., Warinner, C., Lehmann, R., Ritschel, T., Totsche,
1182 K. U., and Küsel, K.: A glimpse of the paleome in endolithic microbial communities, *Microbiome*, 11, 210, 2023.

1183 Woodcroft, B. J., Singleton, C. M., Boyd, J. A., Evans, P. N., Emerson, J. B., Zayed, A. A. F., Hoelzle, R. D.,
1184 Lamberton, T. O., McCalley, C. K., Hodgkins, S. B., Wilson, R. M., Purvine, S. O., Nicora, C. D., Li, C.,
1185 Frolking, S., Chanton, J. P., Crill, P. M., Saleska, S. R., Rich, V. I., and Tyson, G. W.: Genome-centric view of
1186 carbon processing in thawing permafrost, *Nature*, 560, 49–54, 2018.

1187 Woodward, F. I. and Cramer, W.: Plant functional types and climatic change: Introduction, *J. Veg. Sci.*, 7, 306–
1188 308, 1996.

1189 Wullschleger, S. D., Epstein, H. E., Box, E. O., Euskirchen, E. S., Goswami, S., Iversen, C. M., Kattge, J., Norby,
1190 R. J., van Bodegom, P. M., and Xu, X.: Plant functional types in Earth system models: past experiences and
1191 future directions for application of dynamic vegetation models in high-latitude ecosystems, *Ann. Bot.*, 114, 1–16,
1192 2014.

1193 Yooseph, S., Nealson, K. H., Rusch, D. B., McCrow, J. P., Dupont, C. L., Kim, M., Johnson, J., Montgomery, R.,

1194 Ferriera, S., Beeson, K., Williamson, S. J., Tovchigrechko, A., Allen, A. E., Zeigler, L. A., Sutton, G., Eisenstadt,
1195 E., Rogers, Y.-H., Friedman, R., Frazier, M., and Venter, J. C.: Genomic and functional adaptation in surface
1196 ocean planktonic prokaryotes, *Nature*, 468, 60–66, 2010.

1197 Yurekten, O., Payne, T., Tejera, N., Amaladoss, F. X., Martin, C., Williams, M., and O’Donovan, C.:
1198 MetaboLights: open data repository for metabolomics, *Nucleic Acids Res.*, 52, D640–D646, 2024.

1199 Zhong, M., Barrenechea Angeles, I., More, K. D., Picard, M., Bertilsson, S., Bravo, A. G., Björn, E., Coolen, M.
1200 J. L., and Capo, E.: Climate-driven deoxygenation promoted potential mercury methylators in the past Black Sea
1201 water column, *Nat. Water*, 3, 1389–1396, 2025.

1202 Zhong, Z.-P., Tian, F., Roux, S., Gazitúa, M. C., Solonenko, N. E., Li, Y.-F., Davis, M. E., Van Etten, J. L.,
1203 Mosley-Thompson, E., Rich, V. I., Sullivan, M. B., and Thompson, L. G.: Glacier ice archives nearly 15,000-
1204 year-old microbes and phages, *Microbiome*, 9, 160, 2021.

1205 Eric J Carpenter, Naim Matasci, Saravananaray Ayyampalayam, Shuangxiu Wu, Jing Sun, Jun Yu, Fabio
1206 Rocha Jimenez Vieira, Chris Bowler, Richard G Dorrell, Matthew A Gitzendanner, Ling Li, Wensi Du,
1207 Kristian K. Ullrich, Norman J Wickett, Todd J Barkmann, Michael S Barker, James H Leebens-Mack, Gane
1208 Ka-Shu Wong, Access to RNA-sequencing data from 1,173 plant species: The 1000 Plant transcriptomes
1209 initiative (1KP), *GigaScience*, Volume 8, Issue 10, October 2019, giz126,
1210 <https://doi.org/10.1093/gigascience/giz126>

1211 Guerrero-Ramirez N, Mommer L, Freschet GT, Iversen CM, McCormack M.L, Kattge J, Poorter H, van der
1212 Plas F, Bergmann J, Kuyper TW, York LM, Bruelheide H, Laughlin DC, Meier IC, Roumet C, Semchenko
1213 M, Sweeney CJ, van Ruijven J, Valverde-Barrantes OJ, Aubin I, Catford JA, Manning P, Martin A, Milla R,
1214 Minden V, Pausas JG, Smith SW, Soudzilovskaia NA, Ammer C, Butterfield B, Craine J, Cornelissen JHC,
1215 de Vries FT, Isaac ME, Kramer K, König C, Lamb EG, Onipchenko VG, Peñuelas J, Reich PB, Rillig MC,
1216 Sack L, Shipley B, Tedersoo L, Valladares F, van Bodegom P, Weigelt P, Wright JP, Weigelt A. 2021.
1217 Global Root Traits (GRooT) Database. [Global Ecology and Biogeography 30\(1\): 25-37](https://doi.org/10.1111/geb.13179)
1218 <https://doi.org/10.1111/geb.13179>

1219 Van Nuland, M.E., Averill, C., Stewart, J.D. *et al.* Global hotspots of mycorrhizal fungal richness are poorly
1220 protected. *Nature* **645**, 414–422 (2025). <https://doi.org/10.1038/s41586-025-09277-4>

1221

1222

1223

1224

1225