

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 Over the last century, the over-exploitation of natural resources through farming, mining, construction, and other industrial
55 practices has intensified pressure on the biosphere, hydrosphere, and atmosphere (Rockström et al. 2023). The resulting losses
56 of natural habitats have driven rapid shifts in the biogeochemical properties of ecosystems (Lewis and Maslin 2015; Waters et
57 al. 2016). These human-induced changes in global biogeochemical cycles alter energy fluxes that contribute to our current
58 climate and biodiversity crises. The resulting pollution and environmental degradation that have been measured and
59 communicated for decades by geoscientists and biologists necessitate continued biogeoscience research internationally (reports
60 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; Trowbridge et al. 2019;
66 Thibault et al. 2023). These local scale observatories are often equipped with instruments to measure the water, energy and
67 greenhouse gas (GHGs) fluxes (CO₂, CH₄, N₂O) 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).

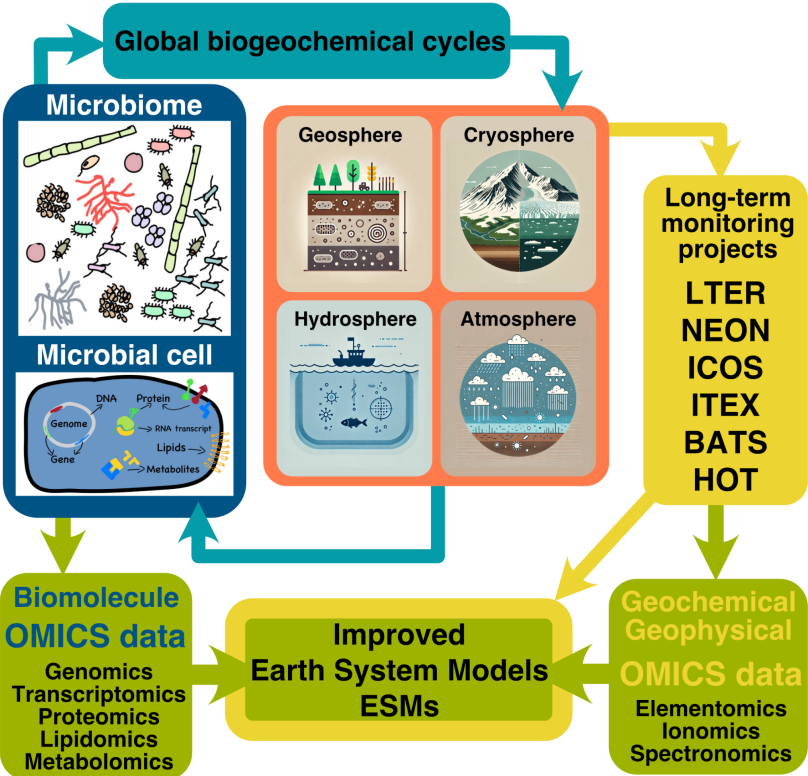
80 To date, most large-scale and long-term infrastructures have shown that representing a certain level of plant functional diversity
81 can improve the predictive accuracy of Earth System Models (ESMs), particularly in simulating interactions among carbon,
82 water, and nutrient cycles (Wullschlegel et al. 2014; Griffith et al., n.d.). However, ESMs still face considerable limitations in
83 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. 2024; He et al. 2026). The ecophysiology of these microbial communities has a
90 direct feedback on the chemistry of the atmosphere and consequently climate (Lovelock and Margulis 1974; Conrad 1994;
91 Monson and Holland 2001). Yet, integrating mechanistic representations of microbial abundance, diversity, and function into
92 ESMs remains incomplete, due to uncertainty about which processes are most critical, how to describe them mathematically,
93 and 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 metabolomics) 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., flux
99 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. The examples are deliberately presented as standalone Boxes, one per Earth
 105 system (air, cryosphere, soil, plant, ocean), each serving as a self-contained entry point for readers working in that system.

106



107

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

113

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 (Tara Ocean Foundation et al. 2022). 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

122 nutrient recycling, e.g., carbon and iron cycles (Twining and Baines 2013). Not only does much of the Earth's biodiversity
123 reside in microorganisms, microbes are highly abundant (est. 10^{30} cells on Earth) and, together, represent the second largest
124 contributor to Earth's biomass despite their microscopic size (Bar-On et al. 2018; Bar-On and Milo 2019). The immense
125 amount of microbial abundance and diversity reflects billions of years of metabolic innovation and evolution to exploit nearly
126 every resource on Earth. These key contributions of microbes as the engines of Earth's biogeochemical cycles and for assuring
127 the survival of Earth's life are grossly neglected by most people, including by the majority of scientists, although efforts are
128 now afoot to give a voice to this unseen majority (Falkowski and Knoll 2007; Doumeizel and Dolan 2024; Tara Ocean
129 Foundation et al. 2022).

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

138 **2.2 Using omics to link diversity to biogeochemical functions: a conceptual foundation**

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

141 Ecosystem ecology sits at the nexus of understanding the flow of energy and materials through the biosphere and other abiotic
142 pools. Many of the biological controls on these fluxes (e.g., reaction rates, substrate affinities, and temperature sensitivities)
143 are governed by functional traits that vary at finer genetic scales than are typically resolved in ecosystem studies, such that
144 aggregation into coarse taxonomic units obscures functionally relevant diversity and propagates uncertainty into predictions
145 of biogeochemical fluxes. Ecosystem ecologists have focused on the role of species variation across environmental gradients
146 to predict ecosystem functions; for example, correlating plant species diversity with an ecosystem's measure of productivity.
147 However, for microorganisms, we have limited knowledge of how microbial diversity translates to functional variation. For
148 one, most microorganisms are uncultured and, therefore, remain unknown and uncharacterized (Steen et al. 2019). Second, the
149 equivalent of the biological species concept does not apply to microorganisms, as there are ambiguities in what constitutes a
150 microbial population, or genetic units belonging to the same species (Shapiro et al. 2016; Chase et al. 2019). Therefore,
151 microbiome analyses typically rely on extracting all the genetic information from an environmental sample for microbial
152 classifications, as this circumvents the need to culture and identify individual microbes. Generally, microbiome scientists

153 cluster and collapse the immense genetic diversity within microbiomes into taxonomic units based on genetic relatedness, such
154 as operational taxonomic units (OTUs) or metagenome assembled genomes (MAGs), which can then be compared to reference
155 databases. However, these taxonomic groupings are defined largely by computational convenience and may not reflect
156 biological organization, as they can represent anything from a family or species to a mosaic genome composed of multiple
157 strains (Tyson et al. 2004; Meziti et al. 2021). Therefore, microbiome-focused ecosystem assessments are inherently forced to
158 move away from the rich literature and theory developed for species to a broader view centered on community-wide metrics
159 such as alpha- and beta-diversity metrics, community-level traits, or interaction metrics, e.g., with co-occurrence networks
160 (Chase and Martiny 2018).

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

176 Given that microbiome assessments rely on genetic material, we can also utilize that genomic information to infer functional
177 traits at multiple biological levels. All functional traits, to some degree, are phylogenetically conserved. Some, such as pH or
178 salinity responses, are broadly conserved at higher taxonomic levels while others, such as the responses to predicted
179 anthropogenic factors (i.e., added nitrogen or drought) (Chase et al. 2017), exhibit fine-scale genetic variation at the genus,
180 species, or even population levels (Martiny et al. 2015). Therefore, microbiome assessments of functional consequences need
181 to prioritize and identify their functional trait of interest and understand the system's diversity accordingly (McLaren and
182 Callahan 2018). For example, the globally distributed marine phototroph *Prochlorococcus* (Cyanobacteriota) is estimated to
183 produce ~20% of the oxygen generated each year by photosynthesis on the planet and is predicted to be highly sensitive to
184 changes in surface ocean temperatures (Flombaum et al. 2013). *Prochlorococcus* is composed of several ecotypes, or highly

185 clustered strains occupying similar ecological niches but differing in key functional traits (Cohan 2002). These ecotypes are
186 geographically and vertically distributed in the water column through their partitioning of resources, including light, iron, and
187 temperature, with closely related ecotypes differing in their thermal tolerance (ranging from ~15-30°C based on cultured
188 isolates) and persistence across surface and subsurface waters (Johnson et al. 2006). Therefore, an understanding of how
189 oxygen production may shift under future climate scenarios will be dependent on the individual responses of each ecotype and
190 their functional differences (Ustick et al. 2023). These patterns aren't unique to *Prochlorococcus* and have been increasingly
191 observed in abundant soil taxa driving leaf litter degradation and carbon turnover (Chase et al. 2018), glacier-fed streams
192 (Fodelianakis et al. 2022), marine sediments (Chase et al. 2021), and bogs (García-García et al. 2019).

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

205 **2.3 Using omics to predict biogeochemical functions at the site scale**

206 **2.3.1 Example from a subsurface system**

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

214 Omics approaches, such as metagenomics, are invaluable for unravelling the metabolic capacities of these microbial
215 communities, offering insights into the dynamic and complex biogeochemical processes that drive subsurface life.

216 Metagenomic studies have revealed a high abundance of chemolithoautotrophs in groundwater, accounting for 12% to 47% of
217 the microbial community, suggesting that dark carbon fixation is central to subsurface trophic webs (Overholt et al. 2022).
218 This process might also have global significance, as the subsurface and surface biospheres are tightly linked through the
219 exchange of energy and matter. Overholt et al. (2022) estimated that dark carbon fixation in deep subsurface environments
220 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 (Péguilhan et al. 2025; Amato et al. 2019). 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).

221

222 **2.3.2 Example from the global ocean**

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

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

241 layers of two basins in the Baltic Sea. Combining such contemporary meta-omic analyses with historical DNA data from
242 sedimentary archives offers a unique opportunity to explore the links between increasing anoxia in water columns and shifts
243 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) (Walter Anthony et al. 2024; Poppeliers et al. 2022).

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. 2021, 2018) 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., Woodcroft et al. 2018; Pessi et al. 2022) 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; Y. Wang et al. 2022). 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 is 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 (Defosse 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. 2025; Frugone-Álvarez et al. 2023; Dussarrat et al. 2022).

2.4 Using omics to predict biogeochemical functions at the global scale

2.4.1 Large-scale metagenomics

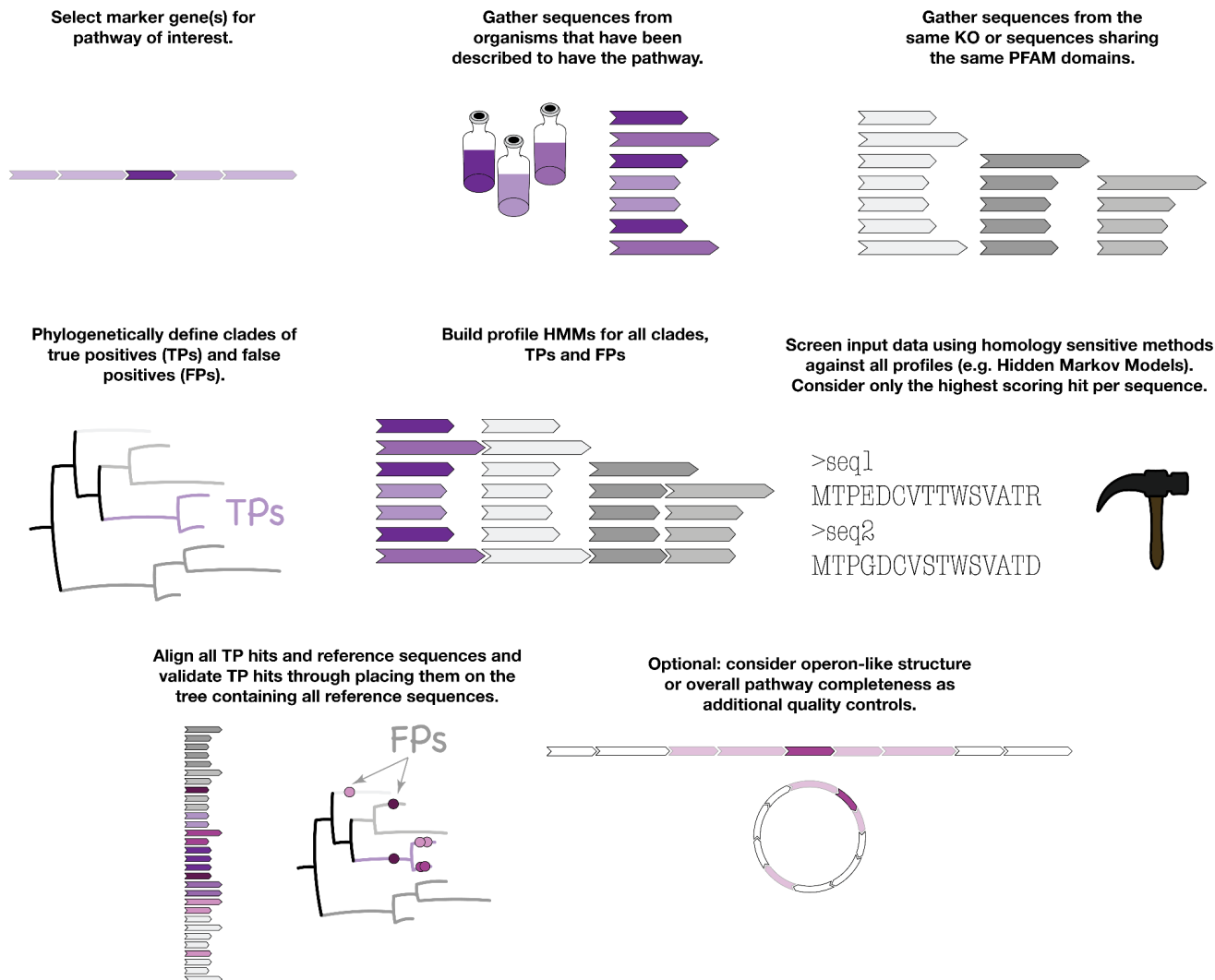
The examples in Sect. 2.2 illustrate how omics approaches can help identify the drivers of biogeochemical processes. Large-scale metagenomic analyses offer a powerful tool to extend findings from individual studies to a global scale. A variety of large-scale metagenomic datasets from diverse ecosystems are now available (Bahram et al. 2018; Sunagawa et al. 2020; Duarte 2015; Schmidt et al. 2024; Maeke et al. 2025), 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 operon-like organization of the genes involved in a pathway can also indicate its functionality, increasing annotation reliability (Bratlie et al. 2010). This property has been used to explore the biosynthetic potential of MAGs generated from *Tara* Oceans, leading to the discovery of a particularly talented group of formerly unknown bacteria able to generate multiple compounds of interest in biotechnology (Paoli et al. 2022).

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

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



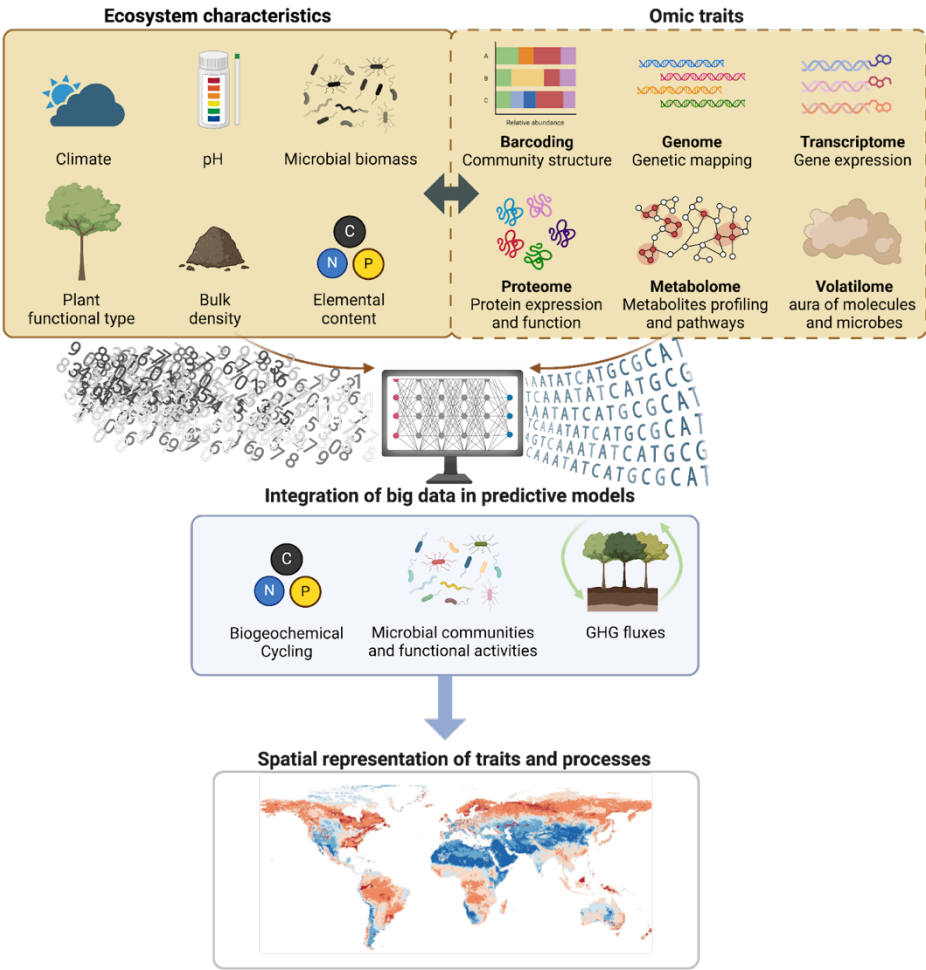
284

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

293
294
295
296
297
298
299
300
301
302

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).



303
304

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.

337 **Table 1: Databases and resources for linking omics data to ecosystem processes.** The table provides an illustrative, non-exhaustive
 338 overview of widely used resources across ecosystems, including data type, ecosystem coverage, strengths and limitations, access standards,
 339 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)
PhyCosm	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 transcriptome s initiative)	Transcriptome s	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	(Soudzilovskaia 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	(Sunagawa et al. 2020; Paoli et al. 2022)

				contextualize d				
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	(“The National Microbiome Data Collaborative Data Portal: An Integrated Multi-Omics Microbiome Data Resource” 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-Velásquez et al. 2019)
AERIS	Atmospheric dynamics and	Atmosphere (regional to	AERIS Parameter	Centralized access to	Focused on physico-	AERIS portal (https://www.	ISO 19115 / INSPIRE	(Ranjan Hatua and Madalli

	chemistry	global)	Vocabulary / GCMD Keywords	French/European atmospheric data	chemistry data	aeris- data.fr/catalog ue/)		2011)
--	-----------	---------	----------------------------------	--	-------------------	--	--	-------

340

341

2.4.3 Integrating omics data into biogeochemical models

342

343

344

345

346

347

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

348

349

350

351

352

353

354

355

356

357

358

359

360

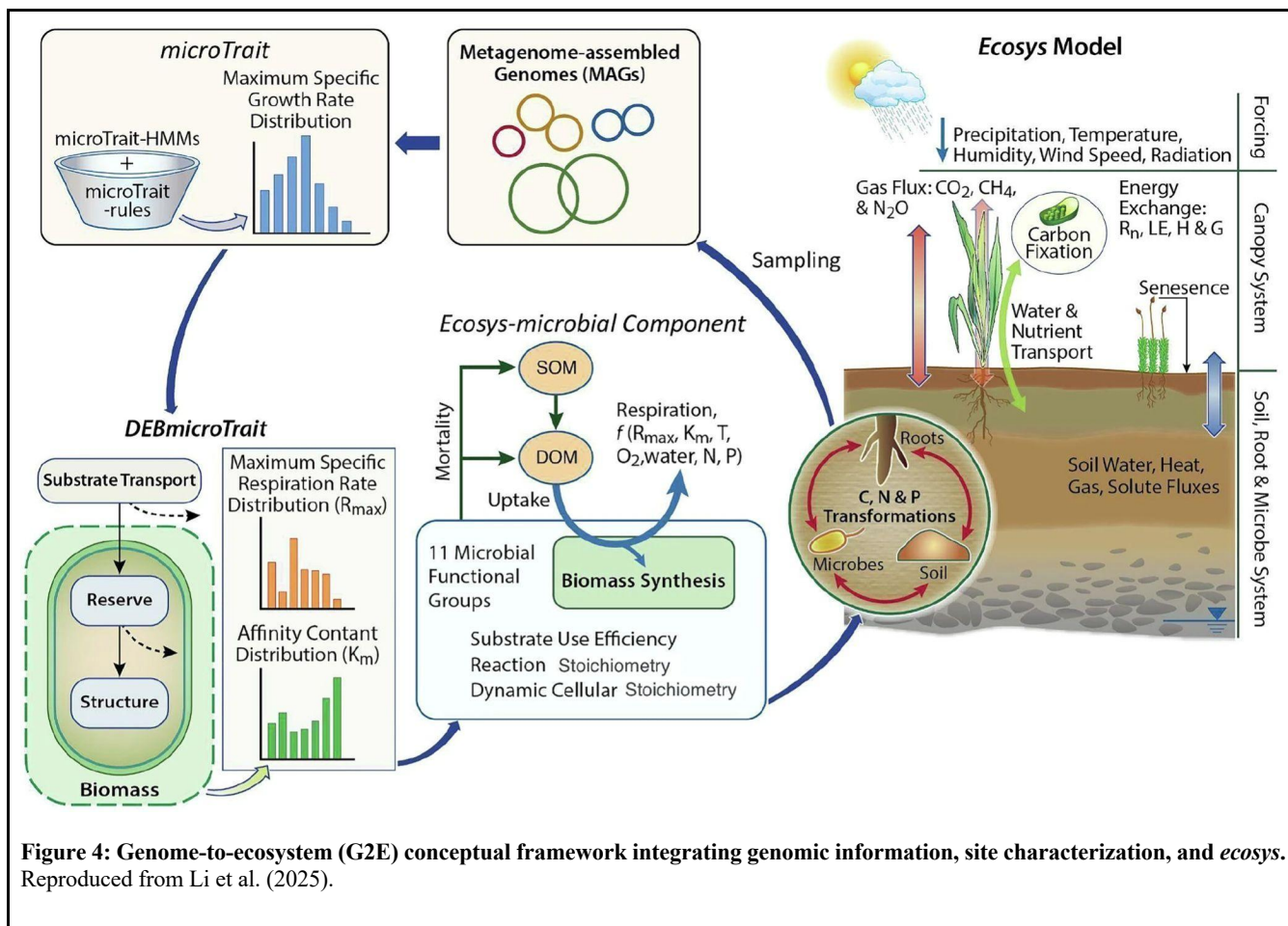
Incorporating genomic information into biogeochemical models requires a structured approach. Here, we review two groups of models: soil biogeochemical models, and full ecosystem-level models, each with distinct capabilities and approaches leveraging omics data. Soil biogeochemical models are designed to simulate detailed soil processes, and are typically driven by diagnostic inputs, such as prescribed litter, soil moisture, and temperature data. Some of these models have explicit representations of microbial processes, including the Microbial-Enzyme Decomposition (MEND) Model, Microbial-MIneral Carbon Stabilization (MIMICS), and CORPSE; see Chandel et al. (2023) for a full review. For example, MEND has been constrained and validated using GeoChip-based gene abundance measurements, which provides a basic platform for testing hypotheses about microbially-mediated biogeochemical processes (Wang et al. 2022; Gao et al. 2020). In contrast to soil biogeochemical models, full ecosystem models simulate the coupled dynamics of water, energy, plants, and soil processes driven by climate forcing. Such coupled models are valuable for projection or spatial extrapolation because many of the emergent properties of primary scientific and societal concern, such as net ecosystem exchanges of CO₂ and CH₄, are strongly regulated by interactions between these interconnected systems. One example of such a model that integrates microbial 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 successfully combined despite the variations in spatial scale of processes and taxonomic resolution. For example, omics have been utilized to differentiate past ecosystem drivers and expand assessments of vegetation changes in rangeland ecosystems (Box 4). However, combining traditional and omics evidence in palaeo-ecology is still in its infancy because there are few epistemically independent models (Jones and Bösl 2021; Edwards 2020).

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

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

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

399

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, Giguet-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).

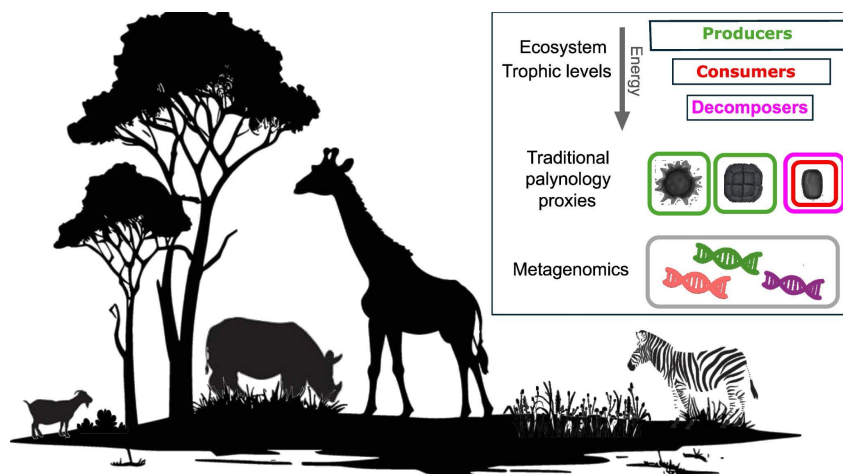


Figure 5: A graphical representation of trophic energy flow in African savanna rangelands featuring producers, consumers, and decomposers. The gray arrow indicates the flow of energy while the colour-coded trophic level boxes represent trophic biomass (not to scale) across the chain. Vegetation dynamics 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).

2.5.2 Future

Experimental approaches to studying microbial evolution and its role in biogeochemistry

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

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

424 However, microbes are distinct in that their generation times, and thus capacity for adaptation, operate on much faster
425 timelines. Combined with large population sizes and their capacity to exchange DNA within and across lineages, microbes are
426 exposed to a vast reservoir of genetic diversity within microbial communities that may contribute to rapid adaptive responses
427 to environmental change. Certainly, lab-based experimental evolutionary studies that manipulate a selective pressure (e.g.,
428 heat) have repeatedly demonstrated that microbes can rapidly evolve and adapt over weeks to months (Tenaillon et al. 2012;
429 Kent et al. 2018), although it's unclear how these artificial laboratory conditions translate to natural populations. In contrast to
430 species sorting, which alters the relative abundances of existing trait variants, evolutionary responses modify trait values
431 themselves, potentially changing microbial reaction norms and longer-term sensitivities to environmental change. Recently,
432 natural experiments using a reciprocal transplant approach across a regional climate gradient demonstrated that microbes, as
433 observed in lab-based experiments, can rapidly adapt through *de novo* mutations in response to predicted climate scenarios
434 (Chase et al. 2021). 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. 2021). 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 (Martiny et al. 2023; Abs et al. 2024; Abs et al. 2024).

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. 2024; Abs et al. 2022; 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.

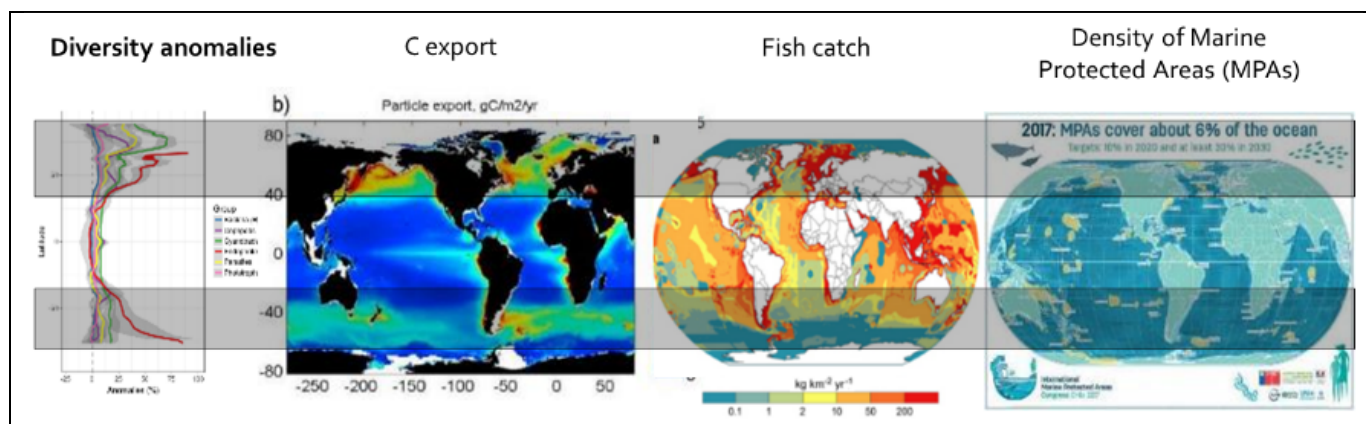


Figure 6: Projected effects of diversity changes within plankton communities on marine ecosystems, fisheries, and biogeochemical cycles (adapted from Ibarbalz et al. 2019). The panel in a) summarizes the biodiversity anomalies projected by end-century from *Tara* Oceans data for a range of different 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 (Feinberg and Irizarry 2010; Chatterjee 2024; 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 Earth System Models (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. 2024).

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) (“The National Microbiome Data Collaborative Data Portal: An Integrated Multi-Omics Microbiome Data Resource” 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. The interoperability they promote relies on community-agreed reporting standards and controlled vocabularies. The most widely adopted is the Minimum Information about any (x) Sequence (MIxS) standard, developed by the Genomic Standards Consortium, which specifies the minimum contextual metadata (habitat, geolocation, and sequencing methodology) required to describe genomes, metagenomes, and marker gene sequences through its associated checklists (e.g., MIGS for genomes, MIMS for metagenomes) (Yilmaz et al. 2011). MIxS is frequently paired with the Environment Ontology (ENVO), which provides a controlled vocabulary for the consistent annotation of biomes, environmental features, and materials such as soil, water, and air (Buttigieg et al. 2013). Other domains rely on complementary standards, such as the ISA framework (ISA-Tab) for metabolomics and the ISO 19115 and Climate and Forecast (CF) conventions for geospatial and atmospheric data, reflecting the heterogeneity of standards that a unified cross-ecosystem framework would need to reconcile. 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 (Nguyen et al. 2023; Kjær et al. 2022; Fernandez-Guerra, Wörmer, 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.

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

532 **3.4 The necessity of interdisciplinary collaboration**

533 Unsurprisingly, integrating omics data into ESMs demands collaboration across disciplines like ecology, geochemistry,
534 bioinformatics, modelling, and data science. Each field brings crucial expertise: ecologists and geochemists link biological and
535 chemical processes, bioinformaticians manage complex omics data, modelers translate findings into model parameters, and
536 data scientists develop frameworks to address challenges in data integration and accessibility. Environmental scientists play a
537 key role in identifying biochemical pathways and traits essential for model parameterization, thereby improving the accuracy
538 of ESMs in simulating ecosystem functions.

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

544 **Code/data/sample availability**

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

546 **Author contribution**

547 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
548 subsection and provided input on the manuscript text, table, figures and discussion of scientific content.

549 **Competing interests**

550 The authors declare that they have no conflict of interest.

551 **Acknowledgements**

552 No acknowledgements.

553 **Financial support**

554 Funding to EA was provided by the CALIPSO project provided by Schmidt Sciences. LW has received funding from the
555 European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant
556 agreement No. 101003125) and the French Govt in the framework of the IdEX Bordeaux University "Investments for the
557 Future" program / GPR Bordeaux Plant Sciences. CB has received funding from the European Research Council (ERC) under
558 the European Union’s Horizon 2020 research and innovation program (Diatomic; grant agreement No. 835067) and Horizon
559 Europe BlueRemediomics (grant agreement No. 101082304) and Marco-Bolo (grant agreement No. 101082021). KK has
560 received funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number
561 218627073: CRC1076 AquaDiva. JH has received funding from the Research Council of Finland (grant 354462) and the Kone
562 Foundation. EC has received funding from the Swedish Research Council VR (VR starting grant 2023-03504). Z.L. was
563 supported at Lawrence Livermore National Laboratory (LLNL) by the LLNL-LDRD Program under Project No. 24-SI-002
564 and by the U.S. Department of Energy (DOE) Office of Science, Energy Earthshot Initiative, as part of the LLNL Terraforming
565 Soils EERC under Award # SCW1841. Work at LLNL was conducted under the auspices of DOE Contract DE-AC52-
566 07NA27344. W.J.R. was supported by the U.S. Department of Energy, Office of Science, Office of Biological and
567 Environmental Research, Terrestrial Ecosystem Science Program, in the Belowground Biogeochemistry Scientific Focus Area,
568 under Award Number DE-AC02-05CH11231. S.R.S. and Z.L. were supported by the EMERGE Biology Integration Institute
569 (US NSF Award #2022070). CK has received funding from the European Union’s Horizon 2020 research and innovation
570 programme under the Marie Skłodowska-Curie grant agreement No 101108340.

571 **References**

- 572 Abs, Elsa, Michaeline B. N. Albright, and Steven D. Allison. 2024. “Invasions Eliminate the Legacy Effects of
573 Substrate History on Microbial Nitrogen Cycling.” *Ecosphere* 15 (1). <https://doi.org/10.1002/ecs2.4754>.
- 574 Abs, Elsa, Alexander B. Chase, and Steven D. Allison. 2023. “How Do Soil Microbes Shape Ecosystem
575 Biogeochemistry in the Context of Global Change?” *Environmental Microbiology* 25 (4): 780–785.
- 576 Abs, Elsa, Alexander B. Chase, Stefano Manzoni, Philippe Ciais, and Steven D. Allison. 2024. “Microbial
577 Evolution-An under-Appreciated Driver of Soil Carbon Cycling.” *Global Change Biology* 30 (4): e17268.
- 578 Abs, Elsa, David Coulette, Philippe Ciais, and Steven D. Allison. 2024. “Microbial Evolution Drives Adaptation
579 of Substrate Degradation on Decadal to Centennial Time Scales Relevant to Global Change.” *Ecology*
580 *Letters* 27 (10): e14530.
- 581 Abs, Elsa, and Régis Ferrière. 2020. “Modeling Microbial Dynamics and Heterotrophic Soil Respiration: Effect
582 of Climate Change.” In *Biogeochemical Cycles*. John Wiley & Sons, Inc.
- 583 Abs, Elsa, Hélène Leman, and Régis Ferrière. 2020. “A Multi-Scale Eco-Evolutionary Model of Cooperation
584 Reveals How Microbial Adaptation Influences Soil Decomposition.” *Communications Biology* 3 (1): 520.

585 Abs, Elsa, Scott Saleska, and Regis Ferriere. 2022. "Microbial Eco-Evolutionary Responses Amplify Global Soil
586 Carbon Loss with Climate Warming." September 20. <https://doi.org/10.21203/rs.3.rs-1984500/v1>.

587 Acinas, Silvia G., Pablo Sánchez, Guillem Salazar, et al. 2021. "Deep Ocean Metagenomes Provide Insight into
588 the Metabolic Architecture of Bathypelagic Microbial Communities." *Communications Biology* 4 (1): 604.

589 Ågren, Göran I., J. Å. Martin Wetterstedt, and Magnus F. K. Billberger. 2012. "Nutrient Limitation on Terrestrial
590 Plant Growth--Modeling the Interaction between Nitrogen and Phosphorus." *The New Phytologist* 194 (4):
591 953–960.

592 Amato, Pierre, Ludovic Besaury, Muriel Joly, Benjamin Penaud, Laurent Deguillaume, and Anne-Marie Delort.
593 2019. "Metatranscriptomic Exploration of Microbial Functioning in Clouds." *Scientific Reports* 9 (1): 4383.

594 Amato, Pierre, Marius Parazols, Martine Sancelme, Paolo Laj, Gilles Mailhot, and Anne-Marie Delort. 2007.
595 "Microorganisms Isolated from the Water Phase of Tropospheric Clouds at the Puy de Dôme: Major
596 Groups and Growth Abilities at Low Temperatures: Microorganisms from the Water Phase of Tropospheric
597 Clouds." *FEMS Microbiology Ecology* 59 (2): 242–254.

598 Armbrecht, Linda, Michael E. Weber, Maureen E. Raymo, et al. 2022. "Ancient Marine Sediment DNA Reveals
599 Diatom Transition in Antarctica." *Nature Communications* 13 (1): 5787.

600 Bahram, Mohammad, Falk Hildebrand, Sofia K. Forslund, et al. 2018. "Structure and Function of the Global
601 Topsoil Microbiome." *Nature* 560 (7717): 233–237.

602 Barbour, Kristin M., Alberto Barrón-Sandoval, Kendra E. Walters, and Jennifer B. H. Martiny. 2023. "Towards
603 Quantifying Microbial Dispersal in the Environment." *Environmental Microbiology* 25 (1): 137–142.

604 Bar-On, Yinon M., and Ron Milo. 2019. "Towards a Quantitative View of the Global Ubiquity of Biofilms."
605 *Nature Reviews. Microbiology* 17 (4): 199–200.

606 Bar-On, Yinon M., Rob Phillips, and Ron Milo. 2018. "The Biomass Distribution on Earth." *Proceedings of the
607 National Academy of Sciences of the United States of America* 115 (25): 6506–6511.

608 Barouillet, Cécilia, Marie-Eve Monchamp, Stefan Bertilsson, et al. 2022. "Investigating the Effects of
609 Anthropogenic Stressors on Lake Biota Using Sedimentary DNA." *Freshwater Biology*, ahead of print,
610 December 5. <https://doi.org/10.1111/fwb.14027>.

611 Bedoya-Velásquez, Andrés Esteban, Gloria Titos, Juan Antonio Bravo-Aranda, et al. 2019. "Long-Term Aerosol
612 Optical Hygroscopicity Study at the ACTRIS SIRTa Observatory: Synergy between Ceilometer and in Situ
613 Measurements." *Atmospheric Chemistry and Physics* 19 (11): 7883–7896.

614 Biller, Steven J., Paul M. Berube, Keven Dooley, et al. 2018. "Marine Microbial Metagenomes Sampled across
615 Space and Time." *Scientific Data* 5 (1): 180176.

616 Björk, Robert G., Mats P. Björkman, Mats X. Andersson, and Leif Klemetsson. 2008. "Temporal Variation in
617 Soil Microbial Communities in Alpine Tundra." *Soil Biology & Biochemistry* 40 (1): 266–268.

618 Bork, P., C. Bowler, C. de Vargas, G. Gorsky, E. Karsenti, and P. Wincker. 2015. "Tara Oceans. Tara Oceans

- Studies Plankton at Planetary Scale. Introduction.” *Science (New York, N.Y.)* 348 (6237): 873.
- Bouskill, Nicholas J., William J. Riley, Qing Zhu, Zelalem A. Mekonnen, and Robert F. Grant. 2020. “Alaskan Carbon-Climate Feedbacks Will Be Weaker than Inferred from Short-Term Experiments.” *Nature Communications* 11 (1): 5798.
- Bouslimani, Amina, Laura M. Sanchez, Neha Garg, and Pieter C. Dorrestein. 2014. “Mass Spectrometry of Natural Products: Current, Emerging and Future Technologies.” *Natural Product Reports* 31 (6): 718–729.
- Bratlie, Marit S., Jostein Johansen, and Finn Drabløs. 2010. “Relationship between Operon Preference and Functional Properties of Persistent Genes in Bacterial Genomes.” *BMC Genomics* 11 (1): 71.
- Bravo, Andrea G., and Claudia Cosio. 2020. “Biotic Formation of Methylmercury: A Bio-Physico-Chemical Conundrum.” *Limnology and Oceanography* 65 (5): 1010–1027.
- Breitburg, Denise, Lisa A. Levin, Andreas Oschlies, et al. 2018. “Declining Oxygen in the Global Ocean and Coastal Waters.” *Science* 359 (6371). <https://doi.org/10.1126/science.aam7240>.
- Briggs, Adrian W., Udo Stenzel, Philip L. F. Johnson, et al. 2007. “Patterns of Damage in Genomic DNA Sequences from a Neandertal.” *Proceedings of the National Academy of Sciences of the United States of America* 104 (37): 14616–14621.
- Brown, Robert W., David R. Chadwick, Huadong Zang, and Davey L. Jones. 2021. “Use of Metabolomics to Quantify Changes in Soil Microbial Function in Response to Fertiliser Nitrogen Supply and Extreme Drought.” *Soil Biology & Biochemistry* 160 (108351): 108351.
- Burrows, S. M., T. Butler, P. Jöckel, et al. 2009. “Bacteria in the Global Atmosphere – Part 2: Modeling of Emissions and Transport between Different Ecosystems.” *Atmospheric Chemistry and Physics* 9 (23): 9281–9297.
- Buttigieg, Pier Luigi, Norman Morrison, Barry Smith, Christopher J. Mungall, Suzanna E. Lewis, and ENVO Consortium. 2013. “The Environment Ontology: Contextualising Biological and Biomedical Entities.” *Journal of Biomedical Semantics* 4 (1): 43.
- Calayag, Alyzza M., Taylor Priest, Ellen Oldenburg, et al. 2025. “Arctic Ocean Virus Communities and Their Seasonality, Bipolarity, and Prokaryotic Associations.” *Nature Communications* 16 (1): 6427.
- Canadell, J. G., P. M. Monteiro, M. H. Costa, et al. 2021. *Global Carbon and Other Biogeochemical Cycles and Feedbacks. IPCC AR6 WGI, Final Government Distribution*.
- Capo, E., and C. Barouillet. 2023. “Tracking Environmental Change Using Lake Sediments.” In *Sedimentary DNA*, vol. 6, 6. Springer Nature.
- Capo, Eric, Caiyan Feng, Andrea G. Bravo, et al. 2022. “Expression Levels of hgcAB Genes and Mercury Availability Jointly Explain Methylmercury Formation in Stratified Brackish Waters.” *Environmental Science & Technology* 56 (18): 13119–13130.
- Capo, Eric, Marie-Eve Monchamp, Marco J. L. Coolen, Isabelle Domaizon, Linda Armbrrecht, and Stefan

- Bertilsson. 2022. “Environmental Paleomicrobiology: Using DNA Preserved in Aquatic Sediments to Its Full Potential.” *Environmental Microbiology* 24 (5): 2201–2209.
- Carpenter, Eric J., Naim Matasci, Saravanaraj Ayyampalayam, et al. 2019. “Access to RNA-Sequencing Data from 1,173 Plant Species: The 1000 Plant Transcriptomes Initiative (1KP).” *GigaScience* 8 (10). <https://doi.org/10.1093/gigascience/giz126>.
- Cavicchioli, Ricardo, William J. Ripple, Kenneth N. Timmis, et al. 2019. “Scientists’ Warning to Humanity: Microorganisms and Climate Change.” *Nature Reviews. Microbiology* 17 (9): 569–586.
- Chaffron, Samuel, Erwan Delage, Marko Budinich, et al. 2021. “Environmental Vulnerability of the Global Ocean Epipelagic Plankton Community Interactome.” *Science Advances* 7 (35): eabg1921.
- Chandel, Aneesh Kumar, Lifen Jiang, and Yiqi Luo. 2023. “Microbial Models for Simulating Soil Carbon Dynamics: A Review.” *Journal of Geophysical Research. Biogeosciences*, ahead of print, July 28. <https://doi.org/10.1029/2023jg007436>.
- Chase, Alexander B., Philip Arevalo, Eoin L. Brodie, Martin F. Polz, Ulas Karaoz, and Jennifer B. H. Martiny. 2019. “Maintenance of Sympatric and Allopatric Populations in Free-Living Terrestrial Bacteria.” *mBio* 10 (5). <https://doi.org/10.1128/mBio.02361-19>.
- Chase, Alexander B., Zulema Gomez-Lunar, Alberto E. Lopez, et al. 2018. “Emergence of Soil Bacterial Ecotypes along a Climate Gradient.” *Environmental Microbiology* 20 (11): 4112–4126.
- Chase, Alexander B., Ulas Karaoz, Eoin L. Brodie, Zulema Gomez-Lunar, Adam C. Martiny, and Jennifer B. H. Martiny. 2017. “Microdiversity of an Abundant Terrestrial Bacterium Encompasses Extensive Variation in Ecologically Relevant Traits.” *mBio* 8 (6). <https://doi.org/10.1128/mBio.01809-17>.
- Chase, Alexander B., and Jennifer B. H. Martiny. 2018. “The Importance of Resolving Biogeographic Patterns of Microbial Microdiversity.” *Microbiology Australia* 39 (1): 5.
- Chase, Alexander B., Douglas Sweeney, Mitchell N. Muskat, Dulce G. Guillén-Matus, and Paul R. Jensen. 2021. “Vertical Inheritance Facilitates Interspecies Diversification in Biosynthetic Gene Clusters and Specialized Metabolites.” *mBio* 12 (6): e0270021.
- Chase, Alexander B., Claudia Weihe, and Jennifer B. H. Martiny. 2021. “Adaptive Differentiation and Rapid Evolution of a Soil Bacterium along a Climate Gradient.” *Proceedings of the National Academy of Sciences of the United States of America* 118 (18). <https://doi.org/10.1073/pnas.2101254118>.
- Chatterjee, Rabindra Nath. 2024. “The Role of Epigenetics in Phenotypic Evolution: Lessons from the Genomic Response to Stresses, Genetic Conflicts and Developmental Plasticity in *Drosophila*.” *The Nucleus; an International Journal of Cytology and Allied Topics*, ahead of print, November 16. <https://doi.org/10.1007/s13237-024-00520-3>.
- Chen, I-Min A., Ken Chu, Krishnaveni Palaniappan, et al. 2023. “The IMG/M Data Management and Analysis System v.7: Content Updates and New Features.” *Nucleic Acids Research* 51 (D1): D723–D732.

687 Chen, Jianwei, Yangyang Jia, Ying Sun, et al. 2024. "Global Marine Microbial Diversity and Its Potential in
688 Bioprospecting." *Nature* 633 (8029): 371–379.

689 Cohan, Frederick M. 2002. "What Are Bacterial Species?" *Annual Review of Microbiology* 56 (1): 457–487.

690 Conrad, Ralf. 1994. "Compensation Concentration as Critical Variable for Regulating the Flux of Trace Gases
691 between Soil and Atmosphere." *Biogeochemistry* 27 (3). <https://doi.org/10.1007/bf00000582>.

692 DeAngelis, Kristen M., Grace Pold, Begüm D. Topçuoğlu, et al. 2015. "Long-Term Forest Soil Warming Alters
693 Microbial Communities in Temperate Forest Soils." *Frontiers in Microbiology* 6 (February): 104.

694 Defosse, Emmanuel, Camille Pitteloud, Patrice Descombes, et al. 2021. "Spatial and Evolutionary Predictability
695 of Phytochemical Diversity." *Proceedings of the National Academy of Sciences of the United States of*
696 *America* 118 (3). <https://doi.org/10.1073/pnas.2013344118>.

697 Defrenne, Camille E., Elsa Abs, Amanda Longhi Cordeiro, et al. 2021. "The Ecology Underground Coalition:
698 Building a Collaborative Future of Belowground Ecology and Ecologists." *The New Phytologist* 229 (6):
699 3058–3064.

700 Delgado-Baquerizo, Manuel, Fernando T. Maestre, Peter B. Reich, et al. 2016. "Carbon Content and Climate
701 Variability Drive Global Soil Bacterial Diversity Patterns." *Ecological Monographs* 86 (3): 373–390.

702 Demina, Tatiana, Heli Marttila, Igor S. Pessi, et al. 2025. "Tunturi Virus Isolates and Metagenome-Assembled
703 Viral Genomes Provide Insights into the Virome of Acidobacteriota in Arctic Tundra Soils." *Microbiome* 13
704 (1): 79.

705 Doney, Scott C., Mary Ruckelshaus, J. Emmett Duffy, et al. 2012. "Climate Change Impacts on Marine
706 Ecosystems." *Annual Review of Marine Science* 4: 11–37.

707 Doumeizel, Vincent, and John R. Dolan. 2024. "The Launch of The Plankton Manifesto in September 2024."
708 *Journal of Plankton Research*, ahead of print, November 13. <https://doi.org/10.1093/plankt/fbae061>.

709 Duarte, Carlos M. 2015. "Seafaring in the 21st Century: The Malaspina 2010 Circumnavigation Expedition."
710 *Limnology and Oceanography Bulletin* 24 (1): 11–14.

711 Duckworth, Jennifer C., Martin Kent, and Paul M. Ramsay. 2000. "Plant Functional Types: An Alternative to
712 Taxonomic Plant Community Description in Biogeography?" *Progress in Physical Geography* 24 (4): 515–
713 542.

714 Dussarrat, Thomas, Claudio Latorre, Millena C. Barros Santos, et al. 2025. "Rhizochemistry and Soil Bacterial
715 Community Are Tailored to Natural Stress Gradients." *Soil Biology & Biochemistry* 202 (109662): 109662.

716 Dussarrat, Thomas, Sylvain Prigent, Claudio Latorre, et al. 2022. "Predictive Metabolomics of Multiple Atacama
717 Plant Species Unveils a Core Set of Generic Metabolites for Extreme Climate Resilience." *The New*
718 *Phytologist* 234 (5): 1614–1628.

719 Edwards, Mary E. 2020. "The Maturing Relationship between Quaternary Paleocology and Ancient
720 Sedimentary DNA." *Quaternary Research* 96 (July): 39–47.

- 721 Falkowski, Paul G., and Andrew H. Knoll, eds. 2007. *Evolution of Primary Producers in the Sea*. Academic
722 Press.
- 723 Feinberg, Andrew P., and Rafael A. Irizarry. 2010. “Evolution in Health and Medicine Sackler Colloquium:
724 Stochastic Epigenetic Variation as a Driving Force of Development, Evolutionary Adaptation, and Disease.”
725 *Proceedings of the National Academy of Sciences of the United States of America* 107 Suppl 1 (suppl_1):
726 1757–1764.
- 727 Felden, Janine, Lars Möller, Uwe Schindler, et al. 2023. “PANGAEA - Data Publisher for Earth &
728 Environmental Science.” *Scientific Data* 10 (1): 347.
- 729 Feng, Kai, Shang Wang, Qing He, et al. 2024. “CoBacFM: Core Bacteria Forecast Model for Global Grassland
730 pH Dynamics under Future Climate Warming Scenarios.” *One Earth (Cambridge, Mass.)* 7 (7): 1275–1287.
- 731 Fernandez-Guerra, Antonio, Guillaume Borrel, Tom O. Delmont, et al. 2023. “A 2-Million-Year-Old Microbial
732 and Viral Communities from the Kap København Formation in North Greenland.” In *bioRxiv*. June 10.
733 <https://doi.org/10.1101/2023.06.10.544454>.
- 734 Fernandez-Guerra, Antonio, Lars Wörmer, Guillaume Borrel, et al. 2023. “Two-Million-Year-Old Microbial
735 Communities from the Kap København Formation in North Greenland.” In *Microbiology*, No.
736 Biorxiv;2023.06.10.544454v3. BioRxiv, June 10.
737 <https://www.biorxiv.org/content/10.1101/2023.06.10.544454v3>.
- 738 Fierer, Noah. 2014. “Microbial Biogeography: Patterns in Microbial Diversity across Space and Time.” In
739 *Accessing Uncultivated Microorganisms*. ASM Press.
- 740 Finks, Sarai S., Claudia Weihe, Sarah Kimball, et al. 2021. “Microbial Community Response to a Decade of
741 Simulated Global Changes Depends on the Plant Community.” *Elementa (Washington, D.C.)* 9 (1).
742 <https://doi.org/10.1525/elementa.2021.00124>.
- 743 Flombaum, Pedro, José L. Gallegos, Rodolfo A. Gordillo, et al. 2013. “Present and Future Global Distributions of
744 the Marine Cyanobacteria Prochlorococcus and Synechococcus.” *Proceedings of the National Academy of
745 Sciences of the United States of America* 110 (24): 9824–9829.
- 746 Fodelianakis, Stilianos, Alex D. Washburne, Massimo Bourquin, et al. 2022. “Microdiversity Characterizes
747 Prevalent Phylogenetic Clades in the Glacier-Fed Stream Microbiome.” *The ISME Journal* 16 (3): 666–675.
- 748 Franz, Daniela, Manuel Acosta, Núria Altimir, et al. 2018. “Towards Long-Term Standardised Carbon and
749 Greenhouse Gas Observations for Monitoring Europe’s Terrestrial Ecosystems: A Review.” *International
750 Agrophysics* 32 (4): 439–455.
- 751 Fraser, Lauchlan H. 2020. “TRY-A Plant Trait Database of Databases.” *Global Change Biology* 26 (1): 189–190.
- 752 Frémont, Paul, Marion Gehlen, and Olivier Jaillon. 2023. “Plankton Biogeography in the 21st Century and
753 Impacts of Climate Change: Advances through Genomics.” *Comptes Rendus Biologies* 346 (G1): 13–24.
- 754 Frémont, Paul, Marion Gehlen, Mathieu Vrac, et al. 2022. “Restructuring of Plankton Genomic Biogeography in

the Surface Ocean under Climate Change.” *Nature Climate Change* 12 (4): 393–401.

Friend, Andrew D., Wolfgang Lucht, Tim T. Rademacher, et al. 2014. “Carbon Residence Time Dominates Uncertainty in Terrestrial Vegetation Responses to Future Climate and Atmospheric CO₂.” *Proceedings of the National Academy of Sciences of the United States of America* 111 (9): 3280–3285.

Frostegård, Åsa, Silas H. W. Vick, Natalie Y. N. Lim, Lars R. Bakken, and James P. Shapleigh. 2022. “Linking Meta-Omics to the Kinetics of Denitrification Intermediates Reveals pH-Dependent Causes of N₂O Emissions and Nitrite Accumulation in Soil.” *The ISME Journal* 16 (1): 26–37.

Frugone-Álvarez, Matías, Sergio Contreras, Oliver Meseguer-Ruiz, et al. 2023. “Hydroclimate Variations over the Last 17,000 Years as Estimated by Leaf Waxes in Rodent Middens from the South-Central Atacama Desert, Chile.” *Quaternary Science Reviews* 311 (108084): 108084.

Gao, Qun, Gangsheng Wang, Kai Xue, et al. 2020. “Stimulation of Soil Respiration by Elevated CO₂ Is Enhanced under Nitrogen Limitation in a Decade-Long Grassland Study.” *Proceedings of the National Academy of Sciences of the United States of America* 117 (52): 33317–33324.

García-García, Natalia, Javier Tamames, Alexandra M. Linz, Carlos Pedrós-Alió, and Fernando Puente-Sánchez. 2019. “Microdiversity Ensures the Maintenance of Functional Microbial Communities under Changing Environmental Conditions.” *The ISME Journal* 13 (12): 2969–2983.

Giguet-Covex, Charline, Johan Pansu, Fabien Arnaud, et al. 2014. “Long Livestock Farming History and Human Landscape Shaping Revealed by Lake Sediment DNA.” *Nature Communications* 5 (1): 3211.

Glassman, Sydney I., Claudia Weihe, Junhui Li, et al. 2018. “Decomposition Responses to Climate Depend on Microbial Community Composition.” *Proceedings of the National Academy of Sciences of the United States of America* 115 (47): 11994–11999.

Grant, R. F. 1998. “Simulation of Methanogenesis in the Mathematical Model Ecosys.” *Soil Biology & Biochemistry* 30 (7): 883–896.

Grasso, Gianluca, Valeria Bianciotto, and Roland Marmeisse. 2024. “Paleomicrobiology: Tracking the Past Microbial Life from Single Species to Entire Microbial Communities.” *Microbial Biotechnology* 17 (1): e14390.

Gregory-Eaves, Irene, and John P. Smol. 2024. “Paleolimnology: Approaches and Applications.” In *Wetzel’s Limnology*, edited by Ian D. Jones and John P. Smol. Elsevier.

Griffin, D. W., C. Gonzalez-Martin, C. Hoose, and D. J. Smith. 2017. “Global-Scale Atmospheric Dispersion of Microorganisms.” In *Microbiology of Aerosols*. John Wiley & Sons, Inc.

Griffith, D. M., J. Cavender-Bares, W. J. Riley, and J. A. Berry. n.d. *Representing Plant Diversity in Land Models: An Evolutionary Approach to make “Functional Types” More Functional Anderegg LDL*.

Griffith, Jory, Jean-Michel Lord, Michael D. Catchen, et al. 2026. “BON in a Box: An Open and Collaborative Platform for Biodiversity Monitoring, Indicator Calculation, and Reporting.” *Bioscience*, no. biaf189

(January). <https://doi.org/10.1093/biosci/biaf189>.

Grigoriev, Igor V., Richard D. Hayes, Sara Calhoun, et al. 2021. “PhycoCosm, a Comparative Algal Genomics Resource.” *Nucleic Acids Research* 49 (D1): D1004–D1011.

Grigoriev, Igor V., Roman Nikitin, Sajeet Haridas, et al. 2014. “MycoCosm Portal: Gearing up for 1000 Fungal Genomes.” *Nucleic Acids Research* 42 (Database issue): D699–704.

Guerrero-Ramírez, Nathaly R., Liesje Mommer, Grégoire T. Freschet, et al. 2021. “Global Root Traits (GRooT) Database.” *Global Ecology and Biogeography: A Journal of Macroecology* 30 (1): 25–37.

Guidi, Lionel, Samuel Chaffron, Lucie Bittner, et al. 2016. “Plankton Networks Driving Carbon Export in the Oligotrophic Ocean.” *Nature* 532 (7600): 465–470.

Heinrichs, Mara Elena, Gonçalo J. Piedade, Ovidiu Popa, et al. 2024. “Breaking the Ice: A Review of Phages in Polar Ecosystems.” *Methods in Molecular Biology (Clifton, N.J.)* 2738: 31–71.

Heinze, Beatrix M., Valérie F. Schwab, Kirsten Küsel, et al. 2025. “Microbial Oxidation Significantly Reduces Methane Export from Global Groundwaters.” *Proceedings of the National Academy of Sciences of the United States of America* 122 (42): e2508773122.

Hermans, Syrie M., Hannah L. Buckley, Bradley S. Case, Fiona Curran-Cournane, Matthew Taylor, and Gavin Lear. 2020. “Using Soil Bacterial Communities to Predict Physico-Chemical Variables and Soil Quality.” *Microbiome* 8 (1): 79.

He, Xianjin, Rose Z. Abramoff, Elsa Abs, Katerina Georgiou, Haicheng Zhang, and Daniel S. Goll. 2024. “Model Uncertainty Obscures Major Driver of Soil Carbon.” *Nature* 627 (8002): E1–E3.

He, Xianjin, Elsa Abs, Steven D. Allison, et al. 2024. “Emerging Multiscale Insights on Microbial Carbon Use Efficiency in the Land Carbon Cycle.” *Nature Communications* 15 (1): 8010.

He, Xianjin, Gaëlle Marmasse, Junxi Hu, et al. 2026. “Microbial Growth Rate Is a Stronger Predictor of Soil Organic Carbon than Carbon Use Efficiency.” *Nature Ecology & Evolution* 10 (2): 372–381.

Ibarbalz, Federico M., Nicolas Henry, Manoela C. Brandão, et al. 2019. “Global Trends in Marine Plankton Diversity across Kingdoms of Life.” *Cell* 179 (5): 1084–1097.e21.

Jia, Weihan, Sten Anslan, Fahu Chen, et al. 2022. “Sedimentary Ancient DNA Reveals Past Ecosystem and Biodiversity Changes on the Tibetan Plateau: Overview and Prospects.” *Quaternary Science Reviews* 293 (107703): 107703.

Johnson, Zackary I., Erik R. Zinser, Allison Coe, Nathan P. McNulty, E. Malcolm S. Woodward, and Sallie W. Chisholm. 2006. “Niche Partitioning among Prochlorococcus Ecotypes along Ocean-Scale Environmental Gradients.” *Science* 311 (5768): 1737–1740.

Jones, Elizabeth D., and Elsbeth Bösl. 2021. “Ancient Human DNA: A History of Hype (then and Now).” *Journal of Social Archaeology* 21 (2): 236–255.

- 822 Karaoz, Ulas, and Eoin L. Brodie. 2022. "MicroTrait: A Toolset for a Trait-Based Representation of Microbial
823 Genomes." *Frontiers in Bioinformatics* 2 (July): 918853.
- 824 Kent, Alyssa G., Catherine A. Garcia, and Adam C. Martiny. 2018. "Increased Biofilm Formation due to High-
825 Temperature Adaptation in Marine *Roseobacter*." In *Nature Microbiology*, vol. 3, 3. no. 9. Preprint.
826 <https://doi.org/10.1038/s41564-018-0213-8>.
- 827 Keuschnig, Christoph, Timothy M. Vogel, Elena Barbaro, et al. 2023. "Selection Processes of Arctic Seasonal
828 Glacier Snowpack Bacterial Communities." *Microbiome* 11 (1): 35.
- 829 Khaled, Amina, Minghui Zhang, Pierre Amato, Anne-Marie Delort, and Barbara Ervens. 2021. "Biodegradation
830 by Bacteria in Clouds: An Underestimated Sink for Some Organics in the Atmospheric Multiphase System."
831 *Atmospheric Chemistry and Physics* 21 (4): 3123–3141.
- 832 Kivlin, Stephanie N. 2020. "Global Mycorrhizal Fungal Range Sizes Vary within and among Mycorrhizal Guilds
833 but Are Not Correlated with Dispersal Traits." *Journal of Biogeography* 47 (9): 1994–2001.
- 834 Kjær, Kurt H., Mikkel Winther Pedersen, Bianca De Sanctis, et al. 2022. "A 2-Million-Year-Old Ecosystem in
835 Greenland Uncovered by Environmental DNA." *Nature* 612 (7939): 283–291.
- 836 Kumar, Manish, Boyang Ji, Karsten Zengler, and Jens Nielsen. 2019. "Modelling Approaches for Studying the
837 Microbiome." *Nature Microbiology* 4 (8): 1253–1267.
- 838 Küsel, Kirsten, Kai U. Totsche, Susan E. Trumbore, Robert Lehmann, Christine Steinhäuser, and Martina
839 Herrmann. 2016. "How Deep Can Surface Signals Be Traced in the Critical Zone? Merging Biodiversity
840 with Biogeochemistry Research in a Central German Muschelkalk Landscape." *Frontiers in Earth Science* 4
841 (April). <https://doi.org/10.3389/feart.2016.00032>.
- 842 Kuske, Cheryl R., Susan M. Barns, Christy C. Grow, Lori Merrill, and John Dunbar. 2006. "Environmental
843 Survey for Four Pathogenic Bacteria and Closely Related Species Using Phylogenetic and Functional
844 Genes." *Journal of Forensic Sciences* 51 (3): 548–558.
- 845 Lamour, Julien, Shawn P. Serbin, Alistair Rogers, et al. 2026. "The Global Spectra-Trait Initiative: A Database of
846 Paired Leaf Spectroscopy and Functional Traits Associated with Leaf Photosynthetic Capacity." *Earth
847 System Science Data* 18 (1): 245–265.
- 848 Larkin, Alyse A., and Adam C. Martiny. 2017. "Microdiversity Shapes the Traits, Niche Space, and
849 Biogeography of Microbial Taxa." *Environmental Microbiology Reports* 9 (2): 55–70.
- 850 Lelieveld, J., and P. J. Crutzen. 1990. "Influences of Cloud Photochemical Processes on Tropospheric Ozone."
851 *Nature* 343 (6255): 227–233.
- 852 Lewis, Simon L., and Mark A. Maslin. 2015. "A Transparent Framework for Defining the Anthropocene Epoch."
853 *The Anthropocene Review* 2 (2): 128–146.
- 854 Li, Jianwei, Gangsheng Wang, Melanie A. Mayes, et al. 2019. "Reduced Carbon Use Efficiency and Increased
855 Microbial Turnover with Soil Warming." *Global Change Biology* 25 (3): 900–910.

856 Li, Zhen, William J. Riley, Gianna L. Marschmann, et al. 2025. "A Framework for Integrating Genomics,
857 Microbial Traits, and Ecosystem Biogeochemistry." *Nature Communications* 16 (1): 2186.

858 Lovelock, James E., and Lynn Margulis. 1974. "Atmospheric Homeostasis by and for the Biosphere: The Gaia
859 Hypothesis." *Tellus A Dynamic Meteorology and Oceanography* 26 (1-2): 2.

860 Lynggaard, Christina, Mads Frost Bertelsen, Casper V. Jensen, et al. 2022. "Airborne Environmental DNA for
861 Terrestrial Vertebrate Community Monitoring." *Current Biology: CB* 32 (3): 701–707.e5.

862 Ma, Bin, Caiyu Lu, Yiling Wang, et al. 2023. "A Genomic Catalogue of Soil Microbiomes Boosts Mining of
863 Biodiversity and Genetic Resources." *Nature Communications* 14 (1): 7318.

864 Maeke, Mara D., Xiuran Yin, Lea C. Wunder, et al. 2025. "Extensive Data Mining Uncovers Novel Diversity
865 among Members of the Rare Biosphere within the Thermoplasmatota." *Microbiome* 13 (1): 155.

866 Männistö, Minna K., Saija H. K. Ahonen, Lars Ganzert, Marja Tirola, Sari Stark, and Max M. Häggblom. 2024.
867 "Bacterial and Fungal Communities in Sub-Arctic Tundra Heaths Are Shaped by Contrasting Snow
868 Accumulation and Nutrient Availability." *FEMS Microbiology Ecology* 100 (4).
869 <https://doi.org/10.1093/femsec/fiae036>.

870 Marschmann, Gianna L., Jinyun Tang, Kateryna Zhalnina, et al. 2024. "Predictions of Rhizosphere Microbiome
871 Dynamics with a Genome-Informed and Trait-Based Energy Budget Model." *Nature Microbiology* 9 (2):
872 421–433.

873 Martiny, Jennifer B. H., Stuart E. Jones, Jay T. Lennon, and Adam C. Martiny. 2015. "Microbiomes in Light of
874 Traits: A Phylogenetic Perspective." *Science (New York, N.Y.)* 350 (6261): aac9323.

875 Martiny, Jennifer B. H., Adam C. Martiny, Eoin Brodie, et al. 2023. "Investigating the Eco-Evolutionary
876 Response of Microbiomes to Environmental Change." *Ecology Letters*, ahead of print, March 25.
877 <https://doi.org/10.1111/ele.14209>.

878 Matthaeus, William J., Sophia I. Macarewich, Jon D. Richey, et al. 2021. "Freeze Tolerance Influenced Forest
879 Cover and Hydrology during the Pennsylvanian." *Proceedings of the National Academy of Sciences of the*
880 *United States of America* 118 (42): e2025227118.

881 McCalley, Carmody K., Ben J. Woodcroft, Suzanne B. Hodgkins, et al. 2014. "Methane Dynamics Regulated by
882 Microbial Community Response to Permafrost Thaw." *Nature* 514 (7523): 478–481.

883 McElwain, Jennifer C., William J. Matthaeus, Catarina Barbosa, et al. 2024. "Functional Traits of Fossil Plants."
884 *The New Phytologist* 242 (2): 392–423.

885 McLaren, Michael R., and Benjamin J. Callahan. 2018. "In Nature, There Is Only Diversity." *mBio* 9 (1).
886 <https://doi.org/10.1128/mBio.02149-17>.

887 Métris, Kimberly L., and Jérémy Métris. 2023. "Aircraft Surveys for Air eDNA: Probing Biodiversity in the
888 Sky." *PeerJ* 11 (April): e15171.

889 Meziti, Alexandra, Luis M. Rodriguez-R, Janet K. Hatt, Angela Peña-Gonzalez, Karen Levy, and Konstantinos

- T. Konstantinidis. 2021. “The Reliability of Metagenome-Assembled Genomes (MAGs) in Representing Natural Populations: Insights from Comparing MAGs against Isolate Genomes Derived from the Same Fecal Sample.” *Applied and Environmental Microbiology* 87 (6). <https://doi.org/10.1128/AEM.02593-20>.
- Monson, Russell K., and Elisabeth A. Holland. 2001. “Biospheric Trace Gas Fluxes and Their Control over Tropospheric Chemistry.” *Annual Review of Ecology and Systematics* 32 (1): 547–576.
- Müller, Caroline, and Robert R. Junker. 2022. “Chemical Phenotype as Important and Dynamic Niche Dimension of Plants.” *The New Phytologist* 234 (4): 1168–1174.
- Nayfach, Stephen, Beltran Rodriguez-Mueller, Nandita Garud, and Katherine S. Pollard. 2016. “An Integrated Metagenomics Pipeline for Strain Profiling Reveals Novel Patterns of Bacterial Transmission and Biogeography.” *Genome Research* 26 (11): 1612–1625.
- Nayfach, Stephen, Simon Roux, Rekha Seshadri, et al. 2021. “A Genomic Catalog of Earth’s Microbiomes.” *Nature Biotechnology* 39 (4): 499–509.
- Nguyen, Ngoc-Loi, Dhanushka Devendra, Natalia Szymańska, et al. 2023. “Sedimentary Ancient DNA: A New Paleogenomic Tool for Reconstructing the History of Marine Ecosystems.” *Frontiers in Marine Science* 10 (June). <https://doi.org/10.3389/fmars.2023.1185435>.
- Nuñez López, Leslie, Pierre Amato, and Barbara Ervens. 2024. “Bacteria in Clouds Biodegrade Atmospheric Formic and Acetic Acids.” *Atmospheric Chemistry and Physics* 24 (9): 5181–5198.
- Osburn, Ernest D., Steven G. McBride, and Michael S. Strickland. 2024. “Microbial Dark Matter Could Add Uncertainties to Metagenomic Trait Estimations.” *Nature Microbiology* 9 (6): 1427–1430.
- Overholt, Will A., Susan Trumbore, Xiaomei Xu, et al. 2022. “Carbon Fixation Rates in Groundwater Similar to Those in Oligotrophic Marine Systems.” *Nature Geoscience* 15 (7): 561–567.
- Paoli, Lucas, Hans-Joachim Ruscheweyh, Clarissa C. Forneris, et al. 2022. “Biosynthetic Potential of the Global Ocean Microbiome.” *Nature* 607 (7917): 111–118.
- Parks, Donovan H., Maria Chuvochina, Christian Rinke, Aaron J. Mussig, Pierre-Alain Chaumeil, and Philip Hugenholtz. 2022. “GTDB: An Ongoing Census of Bacterial and Archaeal Diversity through a Phylogenetically Consistent, Rank Normalized and Complete Genome-Based Taxonomy.” *Nucleic Acids Research* 50 (D1): D785–D794.
- Pasteur, L. 1861. *Mémoire Sur Les Corpuscles Organisés Qui Existent Dans L’atmosphère, Examen de La Doctrine Des Générations Spontanées*.
- Péguilhan, Raphaëlle, Ludovic Besaury, Florent Rossi, et al. 2021. “Rainfalls Sprinkle Cloud Bacterial Diversity While Scavenging Biomass.” *FEMS Microbiology Ecology* 97 (11). <https://doi.org/10.1093/femsec/fiab144>.
- Péguilhan, Raphaëlle, Florent Rossi, Muriel Joly, et al. 2025. “Clouds Influence the Functioning of Airborne Microorganisms.” *Biogeosciences* 22 (5): 1257–1275.
- Pessi, Igor S., Sirja Viitamäki, Anna-Maria Virkkala, et al. 2022. “In-Depth Characterization of Denitrifier

- Communities across Different Soil Ecosystems in the Tundra.” *Environmental Microbiome* 17 (1): 30.
- Pettersson, Kalle J., Tatiana Demina, Eeva Eronen-Rasimus, et al. 2025. “Viral Genetic Diversity and Functional Potential in Polar and Subarctic Sea Ice.” In *bioRxiv*. BioRxiv, December 12.
<https://doi.org/10.64898/2025.12.12.693969>.
- Piton, Gabin, Steven D. Allison, Mohammad Bahram, et al. 2023. “Life History Strategies of Soil Bacterial Communities across Global Terrestrial Biomes.” *Nature Microbiology* 8 (11): 2093–2102.
- Poppeliers, Sanne W. M., Mariet Hefting, Ellen Dorrepaal, and James T. Weedon. 2022. “Functional Microbial Ecology in Arctic Soils: The Need for a Year-Round Perspective.” *FEMS Microbiology Ecology* 98 (12).
<https://doi.org/10.1093/femsec/fiac134>.
- Post, Eric, Richard B. Alley, Torben R. Christensen, et al. 2019. “The Polar Regions in a 2°C Warmer World.” *Science Advances* 5 (12): eaaw9883.
- Ramirez Flores, Ricardo Omar, Jan David Lanzer, Daniel Dimitrov, Britta Velten, and Julio Saez-Rodriguez. 2023. “Multicellular Factor Analysis of Single-Cell Data for a Tissue-Centric Understanding of Disease.” *eLife* 12 (November). <https://doi.org/10.7554/eLife.93161>.
- Ramonedá, Josep, Elias Stallard-Olivera, Michael Hoffert, et al. 2023. “Building a Genome-Based Understanding of Bacterial pH Preferences.” *Science Advances* 9 (17): eadf8998.
- Ranjan Hatua, Sudip, and Devika P. Madalli. 2011. “AERIS: An Integrated Domain Information System for Aerospace Science and Technology.” *Program* 45 (2): 199–212.
- Rantanen, Mika, Alexey Yu Karpechko, Antti Lipponen, et al. 2022. “The Arctic Has Warmed Nearly Four Times Faster than the Globe since 1979.” *Communications Earth & Environment* 3 (1).
<https://doi.org/10.1038/s43247-022-00498-3>.
- Riley, William J., Zelalem A. Mekonnen, Jinyun Tang, Qing Zhu, Nicholas J. Bouskill, and Robert F. Grant. 2021. “Non-Growing Season Plant Nutrient Uptake Controls Arctic Tundra Vegetation Composition under Future Climate.” *Environmental Research Letters* 16 (7): 074047.
- Riley, W. J., Q. Zhu, and J. Y. Tang. 2018. “Weaker Land–climate Feedbacks from Nutrient Uptake during Photosynthesis-Inactive Periods.” *Nature Climate Change* 8 (11): 1002–1006.
- Rockström, Johan, Joyeeta Gupta, Dahe Qin, et al. 2023. “Safe and Just Earth System Boundaries.” *Nature* 619 (7968): 102–111.
- Sanchez-Tello, Juan David, and Adriana Corrales. 2024. “Ectomycorrhizal Fungal Communities in Natural and Urban Ecosystems: *Quercus Humboldtii* as a Study Case in the Tropical Andes.” *Mycorrhiza* 34 (1-2): 45–55.
- Šantl-Temkiv, Tina, Pierre Amato, Emilio O. Casamayor, Patrick K. H. Lee, and Stephen B. Pointing. 2022. “Microbial Ecology of the Atmosphere.” *FEMS Microbiology Reviews* 46 (4).
<https://doi.org/10.1093/femsre/fuac009>.

958 Scales, N. C., A. B. Chase, S. S. Finks, et al. 2022. “Differential Response of Bacterial Microdiversity to
959 Simulated Global Change.” *Applied and Environmental Microbiology* 88 (6): e0242921.

960 Schmid, Marc W., Christian Heichinger, Diana Coman Schmid, et al. 2018. “Contribution of Epigenetic
961 Variation to Adaptation in Arabidopsis.” *Nature Communications* 9 (1): 4446.

962 Schmidt, Thomas S. B., Anthony Fullam, Pamela Ferretti, et al. 2024. “SPIRE: A Searchable, Planetary-Scale
963 mIcrobome REsource.” *Nucleic Acids Research* 52 (D1): D777–D783.

964 Schwarz, Erik, Elsa Abs, Arjun Chakrawal, Luciana Chavez Rodriguez, Pierre Quévreux, and Stefano Manzoni.
965 2025. “Eco-Evolutionary Optimality in Soil Organic Matter Models.” *Ecology Letters* 28 (12): e70278.

966 Shaffer, Justin P., Louis-Félix Nothias, Luke R. Thompson, et al. 2022. “Standardized Multi-Omics of Earth’s
967 Microbiomes Reveals Microbial and Metabolite Diversity.” *Nature Microbiology* 7 (12): 2128–2150.

968 Shaffer, Michael, Mikayla A. Borton, Bridget B. McGivern, et al. 2020. “DRAM for Distilling Microbial
969 Metabolism to Automate the Curation of Microbiome Function.” *Nucleic Acids Research* 48 (16): 8883–
970 8900.

971 Shaked, Yeala, Benjamin S. Twining, Alessandro Tagliabue, and Maria T. Maldonado. 2021. “Probing the
972 Bioavailability of Dissolved Iron to Marine Eukaryotic Phytoplankton Using in Situ Single Cell Iron
973 Quotas.” *Global Biogeochemical Cycles* 35 (8): e2021GB006979.

974 Shapiro, B. Jesse, Jean-Baptiste Leducq, and James Mallet. 2016. “What Is Speciation?” *PLoS Genetics* 12 (3):
975 e1005860.

976 Sokol, Noah W., Eric Slessarev, Gianna L. Marschmann, et al. 2022. “Life and Death in the Soil Microbiome:
977 How Ecological Processes Influence Biogeochemistry.” *Nature Reviews. Microbiology* 20 (7): 415–430.

978 Soudzilovskaia, Nadejda A., Stijn Vaessen, Milagros Barcelo, et al. 2020. “FungalRoot: Global Online Database
979 of Plant Mycorrhizal Associations.” *The New Phytologist* 227 (3): 955–966.

980 Steen, Andrew D., Alexander Crits-Christoph, Paul Carini, et al. 2019. “High Proportions of Bacteria and
981 Archaea across Most Biomes Remain Uncultured.” *The ISME Journal* 13 (12): 3126–3130.

982 Sunagawa, Shinichi, Silvia G. Acinas, Peer Bork, et al. 2020. “Tara Oceans: Towards Global Ocean Ecosystems
983 Biology.” *Nature Reviews. Microbiology* 18 (8): 428–445.

984 Tabares, Ximena, Heike Zimmermann, Elisabeth Dietze, et al. 2020. “Vegetation State Changes in the Course of
985 Shrub Encroachment in an African Savanna since about 1850 CE and Their Potential Drivers.” *Ecology and
986 Evolution* 10 (2): 962–979.

987 Tagliabue, Alessandro, Andrew R. Bowie, Philip W. Boyd, Kristen N. Buck, Kenneth S. Johnson, and Mak A.
988 Saito. 2017. “The Integral Role of Iron in Ocean Biogeochemistry.” *Nature* 543 (7643): 51–59.

989 Tara Ocean Foundation, Tara Oceans, European Molecular Biology Laboratory (EMBL), and European Marine
990 Biological Resource Centre - European Research Infrastructure Consortium (EMBRC-ERIC). 2022.
991 “Priorities for Ocean Microbiome Research.” *Nature Microbiology* 7 (7): 937–947.

992 Taylor, L., S. Banwart, J. Leake, and D. J. Beerling. 2011. “Modeling the Evolutionary Rise of Ectomycorrhiza
 993 on Sub-Surface Weathering Environments and the Geochemical Carbon Cycle.” *American Journal of*
 994 *Science* 311 (5): 369–403.

995 Tenaillon, Olivier, Alejandra Rodríguez-Verdugo, Rebecca L. Gaut, et al. 2012. “The Molecular Diversity of
 996 Adaptive Convergence.” *Science* 335 (6067): 457–461.

997 “The National Microbiome Data Collaborative Data Portal: An Integrated Multi-Omics Microbiome Data
 998 Resource.” 2022. *Nucleic Acids Research*, ahead of print, October 30. <https://doi.org/10.1093/nar/gkab990>.

999 Thibault, Katherine M., Christine M. Laney, Kelsey M. Yule, Nico M. Franz, and Paula M. Mabee. 2023. “The
 1000 US National Ecological Observatory Network and the Global Biodiversity Framework: National Research
 1001 Infrastructure with a Global Reach.” *Journal of Ecology and Environment* 47 (21).
 1002 <https://doi.org/10.5141/jee.23.076>.

1003 Thompson, Luke R., Jon G. Sanders, Daniel McDonald, et al. 2017. “A Communal Catalogue Reveals Earth’s
 1004 Multiscale Microbial Diversity.” *Nature* 551 (7681): 457–463.

1005 Tian, Funing, James M. Wainaina, Cristina Howard-Varona, et al. 2024. “Prokaryotic-Virus-Encoded Auxiliary
 1006 Metabolic Genes throughout the Global Oceans.” *Microbiome* 12 (1): 159.

1007 Tóth, Gergely, Arwyn Jones, and Luca Montanarella. 2013. “The LUCAS Topsoil Database and Derived
 1008 Information on the Regional Variability of Cropland Topsoil Properties in the European Union.”
 1009 *Environmental Monitoring and Assessment* 185 (9): 7409–7425.

1010 Trowbridge, John, Robert Weller, Deborah Kelley, et al. 2019. “The Ocean Observatories Initiative.” *Frontiers in*
 1011 *Marine Science* 6 (March). <https://doi.org/10.3389/fmars.2019.00074>.

1012 Twining, Benjamin S., and Stephen B. Baines. 2013. “The Trace Metal Composition of Marine Phytoplankton.”
 1013 *Annual Review of Marine Science* 5 (1): 191–215.

1014 Tyson, Gene W., Jarrod Chapman, Philip Hugenholtz, et al. 2004. “Community Structure and Metabolism
 1015 through Reconstruction of Microbial Genomes from the Environment.” *Nature* 428 (6978): 37–43.

1016 Urban, Mark C., Sharon Y. Strauss, Fanie Pelletier, et al. 2020. “Evolutionary Origins for Ecological Patterns in
 1017 Space.” *Proceedings of the National Academy of Sciences of the United States of America* 117 (30): 17482–
 1018 17490.

1019 Ustick, Lucas J., Alyse A. Larkin, and Adam C. Martiny. 2023. “Global Scale Phylogeography of Functional
 1020 Traits and Microdiversity in *Prochlorococcus*.” *The ISME Journal* 17 (10): 1671–1679.

1021 Van Nuland, Michael E., Colin Averill, Justin D. Stewart, et al. 2025. “Global Hotspots of Mycorrhizal Fungal
 1022 Richness Are Poorly Protected.” *Nature* 645 (8080): 414–422.

1023 Viitamäki, Sirja, Igor S. Pessi, Anna-Maria Virkkala, et al. 2022. “The Activity and Functions of Soil Microbial
 1024 Communities in the Finnish Sub-Arctic Vary across Vegetation Types.” *FEMS Microbiology Ecology* 98
 1025 (8). <https://doi.org/10.1093/femsec/fiac079>.

1026 Wakeham, Stuart G. 2020. "Organic Biogeochemistry in the Oxygen-Deficient Ocean: A Review." *Organic*
1027 *Geochemistry* 149 (November): 104096.

1028 Walter Anthony, K. M., P. Anthony, N. Hasson, et al. 2024. "Upland Yedoma Taliks Are an Unpredicted Source
1029 of Atmospheric Methane." *Nature Communications* 15 (1): 6056.

1030 Wang, Gangsheng, Qun Gao, Yunfeng Yang, Sarah E. Hobbie, Peter B. Reich, and Jizhong Zhou. 2022. "Soil
1031 Enzymes as Indicators of Soil Function: A Step toward Greater Realism in Microbial Ecological Modeling."
1032 *Global Change Biology* 28 (5): 1935–1950.

1033 Wang, Mingxun, Jeremy J. Carver, Vanessa V. Phelan, et al. 2016. "Sharing and Community Curation of Mass
1034 Spectrometry Data with Global Natural Products Social Molecular Networking." *Nature Biotechnology* 34
1035 (8): 828–837.

1036 Wang, Yajiao, Yu Liu, Yuxing Wu, Nan Wu, Wenwen Liu, and Xifeng Wang. 2022. "Heterogeneity of Soil
1037 Bacterial and Bacteriophage Communities in Three Rice Agroecosystems and Potential Impacts of
1038 Bacteriophage on Nutrient Cycling." *Environmental Microbiome* 17 (1): 17.

1039 Warinner, Christina. 2022. "An Archaeology of Microbes." *Journal of Anthropological Research*, December 7,
1040 000–000.

1041 Warinner, Christina, Alexander Herbig, Allison Mann, et al. 2017. "A Robust Framework for Microbial
1042 Archaeology." *Annual Review of Genomics and Human Genetics* 18 (August): 321–356.

1043 Warinner, Christina, João F. Matias Rodrigues, Rounak Vyas, et al. 2014. "Pathogens and Host Immunity in the
1044 Ancient Human Oral Cavity." *Nature Genetics* 46 (4): 336–344.

1045 Waters, Colin N., Jan Zalasiewicz, Colin Summerhayes, et al. 2016. "The Anthropocene Is Functionally and
1046 Stratigraphically Distinct from the Holocene." *Science (New York, N.Y.)* 351 (6269): aad2622.

1047 Wegner, Carl-Eric, Raphaela Stahl, Irina Velsko, et al. 2023. "A Glimpse of the Paleome in Endolithic Microbial
1048 Communities." *Microbiome* 11 (1): 210.

1049 Woodcroft, Ben J., Caitlin M. Singleton, Joel A. Boyd, et al. 2018. "Genome-Centric View of Carbon Processing
1050 in Thawing Permafrost." *Nature* 560 (7716): 49–54.

1051 Woodward, F. Ian, and Wolfgang Cramer. 1996. "Plant Functional Types and Climatic Change: Introduction."
1052 *Journal of Vegetation Science: Official Organ of the International Association for Vegetation Science* 7 (3):
1053 306–308.

1054 Wullschleger, Stan D., Howard E. Epstein, Elgene O. Box, et al. 2014. "Plant Functional Types in Earth System
1055 Models: Past Experiences and Future Directions for Application of Dynamic Vegetation Models in High-
1056 Latitude Ecosystems." *Annals of Botany* 114 (1): 1–16.

1057 Yilmaz, Pelin, Renzo Kottmann, Dawn Field, et al. 2011. "Minimum Information about a Marker Gene Sequence
1058 (MIMARKS) and Minimum Information about Any (x) Sequence (MIXS) Specifications." *Nature*
1059 *Biotechnology* 29 (5): 415–420.

1060 Yooseph, Shibu, Kenneth H. Nealson, Douglas B. Rusch, et al. 2010. "Genomic and Functional Adaptation in
1061 Surface Ocean Planktonic Prokaryotes." *Nature* 468 (7320): 60–66.

1062 Yurekten, Ozgur, Thomas Payne, Noemi Tejera, et al. 2024. "MetaboLights: Open Data Repository for
1063 Metabolomics." *Nucleic Acids Research* 52 (D1): D640–D646.

1064 Zhong, Meifang, Inés Barrenechea Angeles, Kuldeep D. More, et al. 2025. "Climate-Driven Deoxygenation
1065 Promoted Potential Mercury Methylators in the Past Black Sea Water Column." *Nature Water* 3 (12): 1389–
1066 1396.

1067 Zhong, Zhi-Ping, Funing Tian, Simon Roux, et al. 2021. "Glacier Ice Archives Nearly 15,000-Year-Old
1068 Microbes and Phages." *Microbiome* 9 (1): 160.

1069 Eric J Carpenter, Naim Matasci, Saravanaraj Ayyampalayam, Shuangxiu Wu, Jing Sun, Jun Yu, Fabio
1070 Rocha Jimenez Vieira, Chris Bowler, Richard G Dorrell, Matthew A Gitzendanner, Ling Li, Wensi Du,
1071 Kristian K. Ullrich, Norman J Wickett, Todd J Barkmann, Michael S Barker, James H Leebens-Mack, Gane
1072 Ka-Shu Wong, Access to RNA-sequencing data from 1,173 plant species: The 1000 Plant transcriptomes
1073 initiative (1KP), *GigaScience*, Volume 8, Issue 10, October 2019, giz126,
1074 <https://doi.org/10.1093/gigascience/giz126>

1075 Guerrero-Ramirez N, Mommer L, Freschet GT, Iversen CM, McCormack M.L, Kattge J, Poorter H, van der
1076 Plas F, Bergmann J, Kuyper TW, York LM, Bruelheide H, Laughlin DC, Meier IC, Roumet C, Semchenko
1077 M, Sweeney CJ, van Ruijven J, Valverde-Barrantes OJ, Aubin I, Catford JA, Manning P, Martin A, Milla R,
1078 Minden V, Pausas JG, Smith SW, Soudzilovskaia NA, Ammer C, Butterfield B, Craine J, Cornelissen JHC,
1079 de Vries FT, Isaac ME, Kramer K, König C, Lamb EG, Onipchenko VG, Peñuelas J, Reich PB, Rillig MC,
1080 Sack L, Shipley B, Tedersoo L, Valladares F, van Bodegom P, Weigelt P, Wright JP, Weigelt A. 2021.
1081 Global Root Traits (GRooT) Database. [Global Ecology and Biogeography 30\(1\): 25-37](https://doi.org/10.1111/geb.13179)
1082 <https://doi.org/10.1111/geb.13179>

1083 Van Nuland, M.E., Averill, C., Stewart, J.D. *et al.* Global hotspots of mycorrhizal fungal richness are poorly
1084 protected. *Nature* **645**, 414–422 (2025). <https://doi.org/10.1038/s41586-025-09277-4>

1085

1086

1087

1088

1089