https://editor.copernicus.org/egusphere-2025-1716/final-response

EGU Biogeosciences: egusphere-2025-1716

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

Elsa Abs, Christoph Keuschnig, Pierre Amato, Chris Bowler, Eric Capo, Alexander B. Chase, Luciana Chavez Rodriguez, Abraham Dabengwa, Thomas Dussarrat, Thomas Guzman, Linnea K. Hernandez, Jenni Hultman, Kirsten Küsel, Zhen Li, Anna Mankowski, William J. Riley, Scott S. Saleska, Lisa Wingate

RESPONSE TO REVIEWERS

Reviewer 1 (Anonymous, 27/09/2025)

Major comments:

Comment 1.1: This perspective manuscript highlighted the microbial datasets based on meta-omics technologies, and explored the integration of such datasets into Earth System Models to examine the microbial responses to environmental changes. Some key points and future efforts are discussed in the current version, and this would raise some concerns of ecologists to consider the important effects of microbiomes. The manuscript was written well to enjoy the reading, I have only some open concerns for specific issues.

Response 1.1:
We thank the reviewer for their thoughtful and constructive comments, which helped clarify the scope and strengthen the perspective.

Comment 1.2. The inclusion of meta-omics based microbial datasets in models was not well discussed. Because some of the biogeochemical features such as pH or organic matter are easy to measure with physical equipment, why is it important to model these parameters with microbial datasets? Actually, sequencing microbiomes will also consume additional time instead of reflecting simultaneous dynamics. Direct observation may provide more accurate results.

Response 1.2:
We thank the reviewer for this insightful comment regarding the trade-off between direct physical observation and meta-omics modeling. We have addressed this in the revised manuscript by clarifying that the value of meta-omics lies not in reconstructing easily measured state variables, but in parameterizing the underlying microbial “engine” that dictates ecosystem

sensitivity. In Section 2.2 (L. 200-202), we now explicitly contrast abiotic drivers with biological responses. We argue that while physical parameters like pH are easily monitored, they do not reveal the internal biological "volume knobs" (such as Vmax and Km) that determine how flux rates will shift under changing conditions. Furthermore, in Section 3.1 (L. 478-482), we address the reviewer's point about the time-intensive nature of sequencing. We clarify that while direct observation is essential for tracking current budgets, it often lags behind the microbial shifts that drive them. By characterizing functional potential, models can account for legacy effects and lagged responses, ensuring accurate projections even when historical environmental correlations break down.

Comment 1.3. Lack of summary of current well-established framework or models by integrating meta-omics microbial datasets into biogeochemical changes. Currently, the manuscript only provided one case study by using G2E framework to show this idea, but there was limited information on how to link the microbes with the biogeochemical features. Such as the Microbial-Enzyme Decomposition (MEND) model, which can indicate the soil functions based on soil enzyme and microbial GeoChip-based functional genes, already provides a basis for testing hypotheses about microbially mediated biogeochemical processes in response to environmental changes (Wang et al. 2022 https://doi.org/10.1111/gcb.16036 ; Gao et al., 2020 https://doi.org/10.1073/pnas.2002780117).

Response 1.3:

We agree that the manuscript would benefit from a clearer overview of existing frameworks that integrate microbial meta-omics information into biogeochemical modelling. In the revised version, we made a distinction between two modeling approaches that can couple omic-data: 1) soil biogeochemical models that use NPP, moisture as inputs, 2) full ecosystem models that simulate the plant, water, and energy processes. In discussion of soil biogeochemical models, we added a concise summary of well-established approaches, including the MEND framework and GeoChip-based model calibration, and positioned our perspective relative to these efforts (L. 338-357). Next, we discussed the value of fully integrated ecosystem models, and an example (G2E framework) of parameterizing such models with omic-derived microbial traits, in detail in a dedicated box (now Box 3). This helped clarify how genome-to-ecosystem approaches complement and extend existing microbial–biogeochemical-ecosystem model linkages.

Comment 1.4. Besides, the manuscript discussed some potential work to link biodiversity with ecosystem functions, but lacking detailed methods for this and details microbial traits or paths. For example, machine-learning or AI-based models were used to linking bacterial communities with physico-chemical variables and soil quality (Hermans et al., 2020 https://doi.org/10.1186/s40168-020-00858-1), further another study expanded the models to predict future grassland soil pH changes in response to global warming (Feng et al., 2024 https://doi.org/10.1016/j.oneear.2024.06.002). Also another pioneering work explored the bacterial pH preferences by linking microbial genomic traits to ecosystem functions, i.e., soil pH (Josep et al., 2023 https://doi.org/10.1126/sciadv.adf8998). More discussion about these possible methods or approaches would help readers to follow in their research.

Response 1.4:
We thank the reviewer for pointing out the need for more detailed methodological pathways. In the revised manuscript, we have added a new paragraph in Section 2.4.2 (L. 324–332) to explicitly describe emerging approaches for linking biodiversity with ecosystem function.

****

Reviewer 2 (Wang Minxiao, 01/12/2025)

Comment 2.1: The integration of genetic/omics data into Earth System Models (ESMs) is a promising direction to improve predictions of ecosystem evolution, particularly for element cycling and biodiversity–function relationships. This manuscript addresses a timely topic and provides a valuable forward-looking perspective. I recommend publication after minor-to-moderate revision, mainly to improve clarity and usefulness to readers. A key improvement would be to strengthen figure/box captions and add a few summary visuals/tables. Several sections contain important ideas (especially the model–omics linkage), but the presentation is sometimes too high-level for non-specialists to follow efficiently.

Response 2.1:
We thank the reviewer for their careful reading of the manuscript and for the detailed and constructive suggestions. We are pleased that the reviewer finds the perspective timely and valuable. We agree that several sections would benefit from clearer presentation and additional schematic support, especially for readers outside the immediate omics–modelling community. We address the comments point by point below.

Comment 2.2: General / presentation
Figures and boxes: Please provide more detailed captions for each figure/box so that they are informative as stand-alone items (what the panels show, what variables/axes represent, what the workflow inputs/outputs are, what assumptions/limitations apply).

Response 2.2:
We agree that figures and boxes should be more informative as stand-alone elements. In the revised manuscript, we have expanded all figure and box captions to clearly describe the panels, variables, workflow inputs and outputs, as well as key assumptions and limitations.

Comment 2.3: Add summary visuals: The manuscript proposes several frameworks/steps for future research; a small number of schematic summaries (one workflow diagram + one database table + one worked example box) would substantially improve readability.

Response 2.3:
We agree that additional summary visuals would substantially improve readability. As suggested, we added (i) a concise flow diagram outlining the proposed pathway annotation

strategy, from marker genes to pathway validation (see Response 2.10), (ii) a table compiling key databases and resources (see Response 2.11), and (iii) a boxed worked example illustrating one complete pathway from omics data to ESM-relevant parameters (see Response 2.12).

Specific comments (by line / element)

Comment 2.4: L119: Typo/formatting: please correct 10 ^ 30 to 10^30 (or "10^{30}"). The sentence currently reads "est. 1030 cells on Earth", which should be 10^30.

Response 2.4:
We corrected the formatting of the power notation to 10^{30} (L. 122).

Comment 2.5: L128: Please mention the latest microbial database/resource published in Nature (2024) that is relevant to large-scale ocean microbiome catalogues. For example, the 2024 Nature study that constructed a unified global ocean microbiome genome catalogue (GOMC) from large-scale marine metagenomes would fit well here and is directly aligned with the manuscript's message about global genomic catalogues.

Response 2.5:
We added a reference to Chen et al. (2024) L. 497 and in Table 1.

Comment 2.6: L177: Please add quantitative detail (exact numbers/ranges) to support the statement about trait variation across taxonomic/biological levels and its consequences. For instance, in the Prochlorococcus example (ecotypes/strain clusters), it would help to report at least one numeric range (e.g., contribution estimate, temperature niche range, or projected shift magnitude) rather than only qualitative wording.

Response 2.6:
We agree that quantitative context would strengthen this point. In the revised manuscript, we explicitly added the role of temperature-related trait variation in the *Procholococcus* example. Specifically, we clarified that closely related ecotypes differ in thermal tolerance (app. 15-30°C based on cultured isolates), linking fine-scale genetic variation directly to environmentally relevant temperature gradients and biogeochemical consequences (L. 183-186).

Comment 2.7: L215: I suggest revising the subsection title to "Example from an ocean system" and expanding the discussion with open-ocean examples. Many foundational concepts (e.g., mixotrophy, microbial loop) were developed in open-ocean systems and are now known to be important for carbon sequestration and broader element cycling. Adding 1–2 open-ocean examples would make this section more representative and historically accurate.

Response 2.7:
We revised the subsection title to "Example from the global ocean" (L. 220) and expanded the discussion to include open-ocean examples (L. 221–223), acknowledging the importance of

scientific expeditions (e.g., GOS, GEOTRACES, Malaspina, and Tara Oceans) over the past 25 years to unravel the role of microorganisms in carbon sequestration using genomic approaches.

Comment 2.8: L220: Please explicitly mention the expansion and intensification of OMZs (oxygen minimum zones) under climate change (spatial expansion/shallower boundaries) and clarify how this affects redox-sensitive element cycling and microbial functional shifts.

Response 2.8:
We now discuss the expansion and shoaling of oxygen minimum zones under climate change (L. 229–231) and clarify the implications for potential increases in greenhouse gas emissions.

Comment 2.9: Box 2: Omics studies also highlight the roles and risks of viruses in polar regions (e.g., viral impacts on microbial mortality, carbon shunting, lysogeny/reactivation risks in warming/ice melt contexts). Please add a short paragraph (and key citations) reflecting this.

Response 2.9:
We added a short paragraph at the end of Box 2 highlighting the abundance and diversity of viruses in the cryosphere. This addition discusses how viral infections control microbial communities and indirectly affect biogeochemical cycling, while also noting the recent discovery of widespread auxiliary metabolic genes (AMGs) in polar phages and sea-ice viruses that further influence metabolic potential.

Comment 2.10: ~L255 (pathway annotation paragraph): This part is important for future research. I strongly recommend adding a small schematic chart/flow diagram summarizing the proposed strategy (marker genes → remote homology methods → phylogenetic validation → pathway completeness checks → operon/context evidence).

Response 2.10:
We added a concise flow diagram outlining the proposed pathway annotation strategy, from marker genes to pathway validation (Figure 2).

Comment 2.11: ~L290: Please include a table listing recommended databases (genomes/MAG catalogues, functional annotation resources, metadata standards, etc.). Suggested columns: database name, data type (metagenome/MAG/metatranscriptome, viruses/euks), ecosystem coverage, key strengths/limitations, access/metadata standard, primary reference.

Response 2.11:
We included a table summarizing recommended databases and resources, indicating data types, ecosystem coverage, strengths and limitations, metadata standards, and key references (Table 1).

Comment 2.12: ~L320 (core section): This is the heart of the review. I suggest adding a boxed worked example that walks the reader through one complete pipeline from omics to ESM parameters (e.g., genome → inferred traits → kinetic parameters → model output), including what the inputs are, what assumptions are made, and where uncertainty enters.

Response 2.12:
We agree that a worked example would help readers navigate the conceptual pipeline. We will add a boxed example walking through a representative case from omics data to inferred traits and model parameters, explicitly noting assumptions and sources of uncertainty (Box 3).

Comment 2.13: ~L350: The text mentions the key complexity (typo in my notes: "kexy") that may confuse many readers, including researchers outside the subfield. Please add a short, concrete summary of how complexities are handled/discriminated in practice, e.g., a bullet list of strategy + toolkits (QC, assembly/binning, dereplication, contamination checks, trait inference tools, uncertainty quantification, validation pathways). (If you discuss ancient DNA here, also include standard authenticity/discrimination criteria.)

Response 2.13:
We appreciate the reviewer's request for clarity on how technical complexities and uncertainties are handled in practice. In the revised manuscript, we added a formal definition of molecular paleomicrobiology to the glossary and expanded the discussion in Box 3 to illustrate how trait variation and community composition are discriminated in functional models like methane cycling. To further assist readers outside the subfield, we have added a concrete summary of the standard 'omics' toolkit and validation pathways used to ensure data quality and handle uncertainties at the end of Section 3.1 (L. 483-493).

Comment 2.14: Fig. 4: Consider revising/expanding the figure to connect more directly to ocean microbiome projects (e.g., Tara Oceans), and to show at least one example where omics data have been used (or could be used) to predict carbon dynamics (carbon export, remineralization, plankton network effects). The current text already frames Tara Oceans as the archetypal global survey and notes that Fig. 4 is an example of projecting end-century biodiversity anomalies; strengthening the omics→carbon-dynamics linkage would better match the manuscript's thesis.

Response 2.14:
We appreciate this comment but propose to maintain the figure as is because it is embedded in a section covering long-term ecosystem responses, which is the topic of the figure. On the other hand, we agree that some reference should have been made in this box to how *Tara* Oceans data have been used to interrogate biogeochemical processes. This has now been done by describing results from the paper of Guidi et al. (2016), in which plankton networks were related to carbon export to the ocean interior (now Box 5).

Comment 2.15: L449: I agree that metatranscriptomics can provide essential information on ecosystem responses to climate change. Please consider explicitly integrating metatranscriptome data into your proposed data-to-model framework (where it fits, what it resolves beyond metagenomes, and how it helps constrain "potential vs realized" function).

Response 2.15:

We agree that metatranscriptomics deserves a clearer role in the proposed framework. In the revised version, we explicitly integrated metatranscriptome data into Figures 1 and 3 and Table 1.