

¹ qdiv: A Python package for microbial ecology analysis ² using the Hill number framework

³ Oskar Modin  ¹

⁴ 1 Chalmers University of Technology, Department of Architecture and Civil Engineering, Division of
⁵ Water Environment Technology, Gothenburg, Sweden

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: Richard Liu  

Reviewers:

- [@erinyoung](#)
- [@Ruslan-Kurmashev](#)
- [@Shaun-Jiang](#)

Submitted: 05 January 2026

Published: unpublished

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).³⁰

⁶ Summary

⁷ Microorganisms are ubiquitous and play essential roles in biogeochemical cycles, human
⁸ health, and engineered systems. Microbial ecology research seeks to understand the structure
⁹ and function of microbial communities. Common approaches, such as amplicon sequencing
¹⁰ and metagenomics, generate tabular data that track the relative abundances of hundreds or
¹¹ thousands of species or genes across time or space. To explore how communities assemble and
¹² respond to environmental factors, microbial diversity is quantified using various metrics, some
¹³ incorporating phylogenetic and functional relationships.

¹⁴ The Hill number framework, also known as effective numbers, provides a unified and intuitive
¹⁵ methodology for assessing diversity and changes in community composition. However, few tools
¹⁶ systematically implement this framework across alpha and beta diversity, phylogenetic metrics,
¹⁷ multivariate statistics, and null models. The Python package qdiv fills this gap by applying
¹⁸ the Hill number framework to a broad range of ecological metrics, offering a streamlined and
¹⁹ versatile tool for investigating patterns of community assembly.

³⁰ Statement of need

²¹ qdiv is a Python package designed for microbial ecology. It was developed to facilitate the
²² management and analysis of data generated from amplicon sequencing and metagenomics
²³ within a Python programming environment. Specifically, qdiv implements the Hill number
²⁴ framework ([Chao et al., 2014](#)) for assessing diversity across a broad range of ecological metrics
²⁵ and analyses. This framework enables systematic evaluation of how relative abundances
²⁶ influence diversity measures.

²⁷ Traditionally, Hill numbers are applied to quantify alpha diversity, which is the diversity of
²⁸ taxa within a single community. However, the framework can also be applied to beta diversity,
²⁹ which measures differences in composition between communities, and extended to metrics that
³⁰ incorporate phylogenetic and functional relationships among taxa.

³¹ Many ecological statistical tests and null models, such as Raup-Crick ([Raup & Crick, 1979](#)),
³² the Net Relatedness Index, and the Nearest-Taxon Index ([Webb et al., 2002](#)), have used
³³ presence-absence data or relied on a single relative abundance-based index (typically Bray-
³⁴ Curtis). qdiv fills a critical gap by enabling the implementation of the Hill number framework
³⁵ for analysing diversity and microbial community assembly with a range of statistical methods.
³⁶ It provides a streamlined workflow for data management, diversity analysis, and visualization –
³⁷ all within a Python ecosystem.

³⁸ Software design

³⁹ qdiv has modular architecture centered around the core MicrobiomeData class. This class serves
⁴⁰ as a container for all essential components, including relative abundance tables, taxonomic
⁴¹ annotations, sequences, phylogenetic trees, and associated metadata. By loading these elements
⁴² into a single object, qdiv enables streamlined data handling and analysis. A comprehensive set of
⁴³ class methods supports convenient data management tasks such as merging datasets, subsetting
⁴⁴ samples, renaming features, and exporting files for downstream analyses. Functions for diversity
⁴⁵ calculations, statistical testing, and visualization are organized into dedicated subpackages that
⁴⁶ operate directly on MicrobiomeData instances. Results can be returned as publication-ready
⁴⁷ figures or as structured data tables for further processing in other software environments. The
⁴⁸ package leverages the widely used Python libraries pandas, numpy, and matplotlib, to ensure
⁴⁹ efficiency, flexibility, and compatibility with existing data science workflows.

⁵⁰ Research impact statement

⁵¹ qdiv was first introduced by (Modin et al., 2020) to enable Hill-based beta diversity calculations
⁵² and extend the Raup-Crick null model to such metrics. Since then, it has been adopted by
⁵³ other researchers, e.g., (Alberdi et al., 2021; Nikolova et al., 2021), and used by our own
⁵⁴ research group, e.g., (Abadikhah et al., 2024). It has been downloaded over 1800 times from
⁵⁵ PyPI in the past six months. Version 4 introduces the MicrobiomeData class for streamlined
⁵⁶ data management, along with new statistical and visualization capabilities.

⁵⁷ AI usage disclosure

⁵⁸ Microsoft Copilot was used to revise functions, to draft a new function for distance-based
⁵⁹ redundancy analysis, and to improve docstrings in version 4.0.0 of qdiv, as well as to refine the
⁶⁰ text of this manuscript.

⁶¹ Acknowledgements

⁶² Research projects that have supported the continued development of qdiv were funded by
⁶³ the Swedish Research Council VR (2023-03908), the Swedish Research Council FORMAS
⁶⁴ (2018-00622, 2024-01814), and the NovoNordisk Foundation (NNF24OC0093678).

⁶⁵ References

- ⁶⁶ Abadikhah, M., Persson, F., Farewell, A., Wilen, B.-M., & Modin, O. (2024). Viral diversity
⁶⁷ and host associations in microbial electrolysis cells. *ISME Communications*, 4. <https://doi.org/10.1093/ismeco/ycae143>
- ⁶⁹ Alberdi, A., Bideguren, G. M., & Aizpurua, O. (2021). Diversity and compositional changes in
⁷⁰ the gut microbiota of wild and captive vertebrates: A meta-analysis. *Scientific Reports*, 11.
⁷¹ <https://doi.org/10.1038/s41598-021-02015-6>
- ⁷² Chao, A., Chiu, C.-H., & Jost, L. (2014). Unifying species diversity, phylogenetic diversity,
⁷³ functional diversity, and related similarity and differentiation measures through hill numbers.
⁷⁴ *Annual Review of Ecology, Evolution, and Systematics*, 45. <https://doi.org/10.1146/annurev-ecolsys-120213-091540>
- ⁷⁶ Modin, O., Liebana, R., Saheb-Alam, S., Wilen, B.-M., Suarez, C., Hermansson, M., &
⁷⁷ Persson, F. (2020). Hill-based dissimilarity indices and null models for analysis of microbial
⁷⁸ community assembly. *Microbiome*, 8. <https://doi.org/10.1186/s40168-020-00909-7>

- 79 Nikolova, C., Ijaz, U. Z., & Gutierrez, T. (2021). Exploration of marine bacterioplankton
80 community assembly mechanisms during chemical dispersant and surfactant-assisted oil
81 biodegradation. *Ecology and Evolution*, 11. <https://doi.org/10.1002/ece3.8091>
- 82 Raup, D. M., & Crick, R. E. (1979). Measurement of faunal similarity in paleontology. *Journal*
83 of Paleontology, 53. <https://www.jstor.org/stable/1304099>
- 84 Webb, C. O., Ackerly, D. D., McPeek, M. A., & Donoghue, M. J. (2002). Phylogenies
85 and community ecology. *Annual Review of Ecology, Evolution, and Systematics*, 33.
86 <https://doi.org/10.1146/annurev.ecolsys.33.010802.150448>

DRAFT