

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

Oskar Modin ¹

¹ Chalmers University of Technology, Department of Architecture and Civil Engineering, Division of 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Published: unpublished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 copyright¹⁹ and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/))³⁰.

Summary

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

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

Statement of need

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

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

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

Software design

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

Research impact statement

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

AI usage disclosure

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

Acknowledgements

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

References

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