



BRCore: an R package implementing flexible selection of core taxa using contribution to Bray-Curtis dissimilarity and neutral model fitting

Bolívar Aponte Rolón,^{1,2} Brandon Kristy,^{3,4} Ashley Shade,⁵ Nejc Stopnisek,^{6,7} Sarah L. Lebeis,^{4,8,9,10} Adina Howe,^{1,2} Gian Maria Niccolò Benucci^{4,8}

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT Identifying core taxa in microbial ecology highlights groups likely to participate in a broad range of potential ecological interactions. Here, we present BRCore, an R package to identify core taxa using abundance-occupancy distributions and beta-diversity contributions across ecological niches, and predict stochastic and deterministic taxa.

KEYWORDS microbiome statistics, community ecology, persistent taxa, ASV/OTU abundance-occupancy distributions, Bray-Curtis dissimilarity, Sloan neutral model

Researchers define persistent microbiomes to identify abundant taxa, specifically those consistently present across samples, or enriched under distinct conditions. Core taxa selection is based on relative abundance and consistent detection, tailored to the research question, data set characteristics, and study system.

We introduce BRCore, a cross-platform R package compatible with Linux, macOS, and Windows, which implements the workflow proposed by Shade and Stopnisek (1) and incorporates additional rarefaction steps. The proposed workflow aims to identify persistent microbiomes using abundance-occupancy distributions and neutral community model fitting. Neutral models have previously been used to evaluate potential community assembly processes based on the underlying distributions of persistent taxa (1). Poor goodness-of-fit in neutral models has been used as evidence for deterministic selection (1–3). BRCore defines persistent taxa based on their contribution to beta-diversity, specifically Bray-Curtis dissimilarity (4), and provides neutral model testing, while accounting for features such as crop, site, or time for detection. This approach recognizes that high occupancy or abundance alone may not identify ecologically relevant taxa, and that persistent microbiomes may be transient and context-dependent across temporal and spatial gradients (1). Additionally, BRCore implements rarefaction of ASV/OTUs (i.e., repeatedly subsampling to a fixed depth without replacement) (5) and provides several visualization functions. BRCore requires R version ≥ 4.4 and depends on *phyloseq* ($\geq 1.54.0$), *dplyr* ($\geq 1.1.4$), *ggplot2* ($\geq 4.0.1$), *vegan* ($\geq 2.7.2$), and *rlang* ($\geq 1.1.6$).

As a demonstration, we applied BRCore to the built-in “bcse” data set, which comprises leaf microbiome samples collected from 10 cropping systems at the Bioenergy Cropping System Experiment in Kellogg Biological Station (6) (Fig. 1). In the pre-processing step, we established the rarefaction depth cutoff using various diagnostics, including per-sample read depth plots and Good’s coverage (7) assessments (Fig. 1A), and determining noise propagation across rarefaction iterations (8). The data set was rarefied to 1,000 sequences per sample across three iterations. We computed Bray-Curtis dissimilarities for each iteration, and averaged these to identify core taxa. *Elbow*

Editor Patrick D. Schloss, University of Michigan, Ann Arbor, Michigan, USA

Address correspondence to Bolívar Aponte Rolón, bolaponte@pm.me.

The authors declare no conflict of interest.

See the funding table on p. 3.

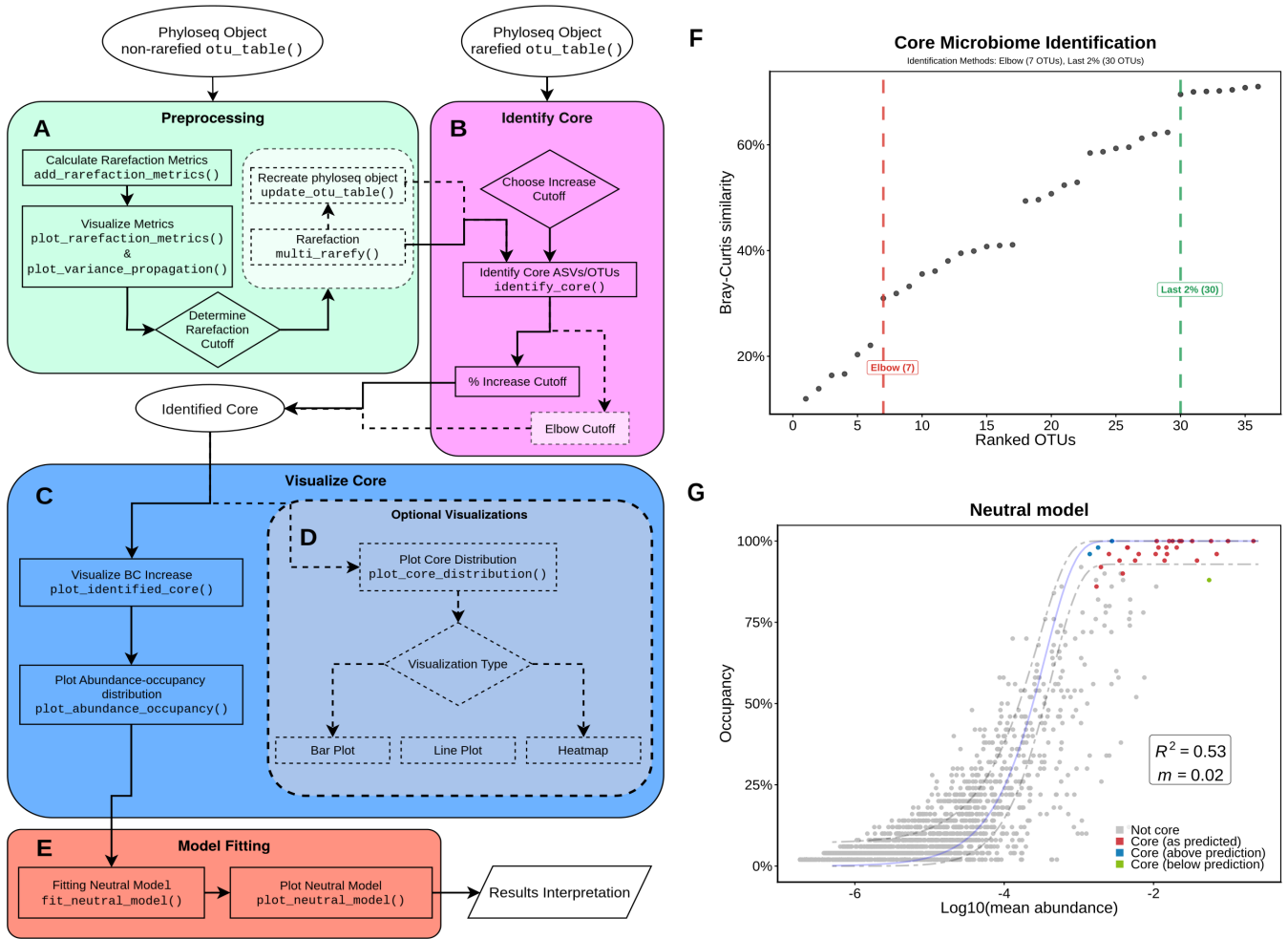
Ed. note: This article was handled by Pat Schloss, who acted as a guest editor in consultation with the editor in chief, Irene Newton. Dr. Schloss’s role as the guest editor of this article is outside of his duties as the chair of the Journals Committee.

Received 4 March 2026

Accepted 27 May 2026

Published 15 June 2026

Copyright © 2026 Aponte Rolón et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).



and 2% *increase* cutoffs were applied and plotted (Fig. 1B through F); see Shade and Stopnisek (1) for details. Briefly, the *elbow* method identifies the point (k threshold) at which incremental gains begin to plateau and additional retained taxa yield diminishing returns, while the 2% *increase* method maintains taxa until the additional gain from retaining one more taxon falls below a specified percentage (e.g., 2%). Neutral model goodness-of-fit statistics were calculated, including migration rate (m), coefficient of determination (R^2), root mean square error (RMSE), Akaike information criterion (AIC), and Bayesian information criterion (BIC) for both binomial and Poisson fits; R^2 and m are shown in the plots (Fig. 1E and G).

The BRCore R package offers a reproducible, user-friendly, and flexible pipeline for persistent microbiome identification. It enables users to emphasize study-specific conditions or treatments and to interpret ecological niches meaningfully.

The BRCore R package offers a reproducible, user-friendly, and flexible pipeline for persistent microbiome identification. It enables users to emphasize study-specific conditions or treatments and to interpret ecological niches meaningfully.

ACKNOWLEDGMENTS

This material is based upon work supported by the DOE Center for Advanced Bioenergy and Bioproducts Innovation (U.S. Department of Energy, Office of Science, Biological and Environmental Research Program under Award Number DE-SC0018420); and the Great Lakes Bioenergy Research Center (U.S. Department of Energy, Office of Science, Biological and Environmental Research Program under Award Numbers DE-SC0018409 and DE-FC02-07ER64494).

We thank Keara Grady, Kelsey Middlemiss, Kevin Myers, Michael Place, Sameer D'Costa, Phillip de Lorimier, and Hsun-Yi Hsieh for their assistance in testing the BRCore package.

AUTHOR AFFILIATIONS

¹Department of Agricultural and Biosystems Engineering, Iowa State University, Ames, Iowa, USA

²Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois Urbana-Champaign, Urbana, Illinois, USA

³Department of Integrative Biology, Michigan State University, East Lansing, Michigan, USA

⁴Great Lakes Bioenergy Research Center (GLBRC), East Lansing, Michigan, USA

⁵CNRS, INRAE, LEM, UMR 5557, UMR 1418, Université Lyon 1, Villeurbanne, France

⁶Department for Microbiological Research, National Laboratory of Health, Environment and Food, Maribor, Slovenia

⁷Department of Agricultural Ecology and Natural Resources, Agricultural Institute of Slovenia, Ljubljana, Slovenia

⁸Department of Plant, Soil, and Microbial Sciences, Michigan State University, East Lansing, Michigan, USA

⁹Department of Microbiology, Immunology and Genetics, Michigan State University, East Lansing, Michigan, USA

¹⁰Plant Resilience Institute, Michigan State University, East Lansing, Michigan, USA

AUTHOR ORCIDs

Bolívar Aponte Rolón  <http://orcid.org/0000-0002-2544-4551>

Brandon Kristy  <http://orcid.org/0000-0003-2421-7143>

Ashley Shade  <http://orcid.org/0000-0002-7189-3067>

Nejc Stopnisek  <http://orcid.org/0000-0002-1532-6826>

Sarah L. Lebeis  <http://orcid.org/0000-0001-9417-9678>

Adina Howe  <http://orcid.org/0000-0002-7705-343X>

Gian Maria Niccolò Benucci  <http://orcid.org/0000-0003-1589-947X>

FUNDING

Funder	Grant(s)	Author(s)
U.S. Department of Energy	SC0018420	Bolívar Aponte Rolón Adina Howe
U.S. Department of Energy	DE-SC0018409, DE-FC02-07ER64494	Brandon Kristy Sarah L. Lebeis Gian Maria Niccolò Benucci

AUTHOR CONTRIBUTIONS

Bolívar Aponte Rolón, Conceptualization, Data curation, Formal analysis, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review and editing | Brandon Kristy, Conceptualization, Data curation, Software,

Validation, Writing – review and editing | Ashley Shade, Data curation, Formal analysis, Software, Validation, Writing – review and editing | Nejc Stopnisek, Data curation, Formal analysis, Software, Validation, Writing – review and editing | Sarah L. Lebeis, Funding acquisition, Supervision, Validation, Writing – review and editing | Adina Howe, Funding acquisition, Supervision, Validation, Writing – review and editing | Gian Maria Niccolò Benucci, Conceptualization, Data curation, Formal analysis, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

BRCore is available on GitHub (<https://github.com/germs-lab/BRCore>), Zenodo (<https://doi.org/10.5281/zenodo.18866205>), and The Comprehensive R Archive Network (<https://cran.r-project.org/package=BRCore>).

REFERENCES

1. Shade A, Stopnisek N. 2019. Abundance-occupancy distributions to prioritize plant core microbiome membership. *Curr Opin Microbiol* 49:50–58. <https://doi.org/10.1016/j.mib.2019.09.008>
2. Sloan WT, Lunn M, Woodcock S, Head IM, Nee S, Curtis TP. 2006. Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ Microbiol* 8:732–740. <https://doi.org/10.1111/j.1462-2920.2005.00956.x>
3. Burns AR, Stephens WZ, Stagaman K, Wong S, Rawls JF, Guillemin K, Bohannan BJ. 2016. Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *ISME J* 10:655–664. <https://doi.org/10.1038/ismej.2015.142>
4. Shade A, Jones SE, Caporaso JG, Handelsman J, Knight R, Fierer N, Gilbert JA. 2014. Conditionally rare taxa disproportionately contribute to temporal changes in microbial diversity. *mBio* 5:e01371-14. <https://doi.org/10.1128/mBio.01371-14>
5. Schloss PD. 2024. Rarefaction is currently the best approach to control for uneven sequencing effort in amplicon sequence analyses. *mSphere* 9:e0035423. <https://doi.org/10.1128/msphere.00354-23>
6. Haan NL, Benucci GNM, Fiser CM, Bonito G, Landis DA. 2023. Contrasting effects of bioenergy crops on biodiversity. *Sci Adv* 9:eadh7960. <https://doi.org/10.1126/sciadv.adh7960>
7. Good IJ. 1953. The population frequencies of species and the estimation of population parameters. *Biometrika* 40:237–264. <https://doi.org/10.1093/biomet/40.3-4.237>
8. Cameron ES, Schmidt PJ, Tremblay BJ-M, Emelko MB, Müller KM. 2021. Enhancing diversity analysis by repeatedly rarefying next generation sequencing data describing microbial communities. *Sci Rep* 11:22302. <https://doi.org/10.1038/s41598-021-01636-1>