

# Are short-read amplicons suitable for the prediction of microbiome functional potential? A critical perspective

Vitor Heidrich<sup>1,2</sup>  | Lukas Beule<sup>3</sup> 

<sup>1</sup>Centro de Oncologia Molecular, Hospital Sírio-Libanês, São Paulo, Brazil

<sup>2</sup>Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, São Paulo, Brazil

<sup>3</sup>Julius Kühn Institute (JKI)—Federal Research Centre for Cultivated Plants, Institute for Ecological Chemistry, Plant Analysis and Stored Product Protection, Berlin, Germany

## Correspondence

Vitor Heidrich, Centro de Oncologia Molecular, Hospital Sírio-Libanês, São Paulo 01308-060, Brazil.

Email: [vheidrich@mochsl.org.br](mailto:vheidrich@mochsl.org.br)

Lukas Beule, Julius Kühn Institute (JKI)—Federal Research Centre for Cultivated Plants, Institute for Ecological Chemistry, Plant Analysis and Stored Product Protection, Berlin 14195, Germany.

Email: [lukas.beule@julius-kuehn.de](mailto:lukas.beule@julius-kuehn.de)

## Funding information

Fundação de Amparo à Pesquisa do Estado de São Paulo, Grant/Award Number: 13996-0/2018

## Abstract

Taxonomic marker gene analysis allows uncovering taxonomic profiles of microbial communities at low cost, making it omnipresent in microbiome research. There is an ever-expanding set of tools to extract further biological information from this kind of data. In this perspective, we enunciate several concerns regarding the biological validity of predicting functional potential from taxonomic profiles, especially when they are generated by short-read sequencing. The taxonomic resolution of marker genes, intragenomic variability of marker genes, and the compositional nature of microbiome data are discussed. Combining actual measurements of microbiome functions with predicted functional potentials is proposed as a powerful approach to better understand microbiome functioning. In this context, the significance of predicted functional potentials for generating and testing hypotheses is highlighted. We argue that functions of microbiomes predicted from microbiome DNA read count data generated by short-read amplicon sequencing should not serve as the only basis to draw biological inferences.

## KEYWORDS

compositional data, functional potential profiles, intragenomic marker gene variability, microbial functions, microbiome data, short-read amplicon sequencing, taxonomic resolution

## Highlights

- Concerns regarding the biological validity of predicting functional potential from taxonomic profiles of microbiome data sets are enunciated.
- Taxonomic resolution of marker genes, intragenomic variability of marker genes, and the compositional nature of microbiome data are discussed.
- Combining measurements of actual functions and predicted functional potential profiles is a powerful approach to understanding microbial functioning.

Vitor Heidrich and Lukas Beule contributed equally to this study.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *iMeta* published by John Wiley & Sons Australia, Ltd on behalf of *iMeta* Science.

## INTRODUCTION

Over the last two decades, next-generation sequencing (NGS) technologies gained enormous popularity for the analysis of microbial communities. Modern NGS platforms are affordable, allowing users to multiplex large amounts of samples, and have high base calling accuracy. These technologies enabled researchers to decipher the composition of microbial communities in varied habitats (e.g., [1–3]). With the rapid rise in microbiome research, several tools emerged to predict functional potential profiles from the taxonomical profiles of microbiome data sets [4]. Functional potential profiles usually list the microbial functions predicted to be present in a microbiome as well as the relative importance of each function based on the proportion of microbes in which they are predicted to be present. As microbial functions are important for the basic understanding of microbial communities, these tools aim to complement taxonomical microbiome data sets. Despite all the additional information that such tools offer, we here communicate several concerns that we have regarding the use of taxonomic annotation of microbiome data sets generated by short-read amplicon sequencing to derive microbial functions. While the prediction algorithms themselves have been discussed intensively (e.g., [4]), this perspective focuses on the applicability of short-read amplicon sequencing data to the task of functional potential prediction, and not the efficiency of prediction algorithms.

### Taxonomic resolution of amplicon sequencing

At the time of writing, short-read amplicon sequencing using the Illumina platform is the most common sequencing technology. On Illumina's widely used MiSeq System, short paired-end reads of up to  $2 \times 300$  base pairs can be generated. For the analysis of microbiomes, the most popular targets of short-read amplicon sequencing are taxonomically informative parts of amplified 16S ribosomal RNA (rRNA) and internal transcribed spacer (ITS) loci. Sequencing of parts of these loci often lacks taxonomic resolution at species level. For example, almost two decades ago, Blackwood et al. [5] sequenced the hypervariable V1 to V3 regions of the 16S rRNA genes of *Bacillus* spp. and found that species of the clinically relevant *Bacillus cereus* group, *Bacillus anthracis* and *Bacillus cereus*, could not be distinguished. In fact, sequencing of full 16S rRNA genes of the *B. cereus* group revealed only some isolates of *B. anthracis* and *B. cereus* presented discriminable 16S rRNA sequences

[6, 7], pointing to the limited taxonomic resolution of this marker gene for certain taxa. Recently, the importance of strain-level identification of members of the human microbiome has been pointed out [8, 9]. Short-read amplicon sequencing, however, mostly fails to differentiate closely related strains [9].

Previous studies proved the self-evident fact that a full-length 16S assessment, which is achievable by employing a long-read sequencing technology such as PacBio, increases taxonomic resolution (e.g., [10]). As an alternative to long-read sequencing, Loop Genomics proposed adapting Illumina short-read sequencing technology to produce synthetic long reads by using unique molecular barcodes in each 16S amplicon, which also showed to improve taxonomic resolution at species level [11]. Additionally, by using computational tools such as SMURF [12], it is possible to infer taxonomic profiles with a greater taxonomic resolution by aggregating information regarding amplicon sequencing libraries targeting different 16S hypervariable regions. This alternative is currently more accessible because there is no need for replacing sequencing platforms or adapting library preparation workflows. However, this is still suboptimal since, in opposition to Loop Genomics technology, full-length 16S sequences cannot be reconstructed with this strategy.

Although the costs are still prohibitive for wide adoption, we predict that full-length 16S assessment either by long-read amplicon sequencing or bioinformatics reconstruction of the full-length 16S based on uniquely identified short reads will gradually replace short-read amplicon sequencing in the long term, leading to increased taxonomic resolution and, thus, more accurate prediction of functional potential profiles.

Apart from taxonomic distinction, genomic variability among isolates of the same species further impedes the accurate prediction of functions. In 2004, Jaspers and Overmann [13] found that 11 isolates of *Brevundimonas alba* had identical full 16S rRNA gene sequences but high genomic diversity [13]. The authors were able to show that despite identical 16S rRNA gene sequences, these isolates showed pronounced differences in their physiology. This illustrates the need to isolate microorganisms to precisely assess their functions. Although the lack of reference genomes in public databases undoubtedly limits functional potential prediction tools [4], this example further illustrates that reference genome availability does not guarantee reliable functional potential inferences.

### Intragenomic variability of marker genes

Another complication in exploring microbiome structure using amplicon sequencing data is the variable copy

number of marker genes in microbial genomes. Fungi can contain up to several hundred copies of rRNA genes (which are interspaced by ITS sequences), but this varies by order of magnitude across fungi [14]. Even among isolates of a single fungal species, copy numbers of 18S and 28S rRNA genes per genome can vary largely [14–17]. Lavrinienko et al. [18] speculate that nontranscribed regions such as ITS may have an even greater copy number variability. Similarly, as shown in Rainey et al. [19], bacteria may contain multiple copies of the 16S rRNA gene. Recent estimates indicate the median 16S rRNA gene copy number per bacterial species can vary between 1 and 19 [20]. Although tools to correct for the gene copy number variability of 16S rRNA genes in archaeal and bacterial genomes have been developed [21, 22], the reliability of gene copy number correction remains questionable [23, 24]. In addition, it is known for over two decades that the 16S rRNA genes in a single bacterial genome are not always identical (e.g., [25]). This may cause amplicon sequencing studies to identify 16S rRNA alleles within the same bacterial cell as pertaining to different species [26], which often leads to inflated diversity estimates [27]. Therefore, besides the obvious implications of variable marker gene copy numbers in attempts of estimating the relative contribution of each species in the environment under study (i.e., distortion of relative abundances), this variability sometimes translates into allele diversity, which can even confound microbiome membership. As functional potential profiles reverberate all biases in taxonomic profiles, intragenomic variability of marker genes confounds the relative importance estimation of potential functions and likely inflates the predicted diversity of functional potential profiles.

## The compositional nature of microbiome data

Because there is no relationship between the number of reads generated by NGS from a sample and the number of bacterial cells in that sample [28], bacterial reads do not translate into bacterial abundances. The number of reads generated for each taxon during NGS informs solely on the relative sizes of parts of the community, making NGS microbiome data sets compositional [29]. In other words, this means they unlock the relative sequencing read abundance of the taxa (i.e., proportions or frequencies) present in a microbial community, but because the size (microbial biomass) of the whole community remains unknown, they do not reveal absolute abundances of taxa [28–31]. Therefore, even if the predicted functional potential profile of a community

matches the profile of its actually realized functional potential, the total population size of the community is still ignored which disables estimation of the magnitude of the functional potential. For example, given two compositionally identical microbiomes (microbiome A and B) that differ in their overall population size by factor two (microbiome A has double the population size of B), we would like to compare their predicted to their realized functional potential. Although their predicted functional potentials based on taxonomic profiles will be identical, microbiome A has double the realized functional potential due to the double population size. Great efforts have been undertaken to overcome this limitation of microbiome analysis through various ways of additional quantification of microorganisms (e.g., [28, 30, 32–34]); however, different quantification methods may introduce additional data variability [35].

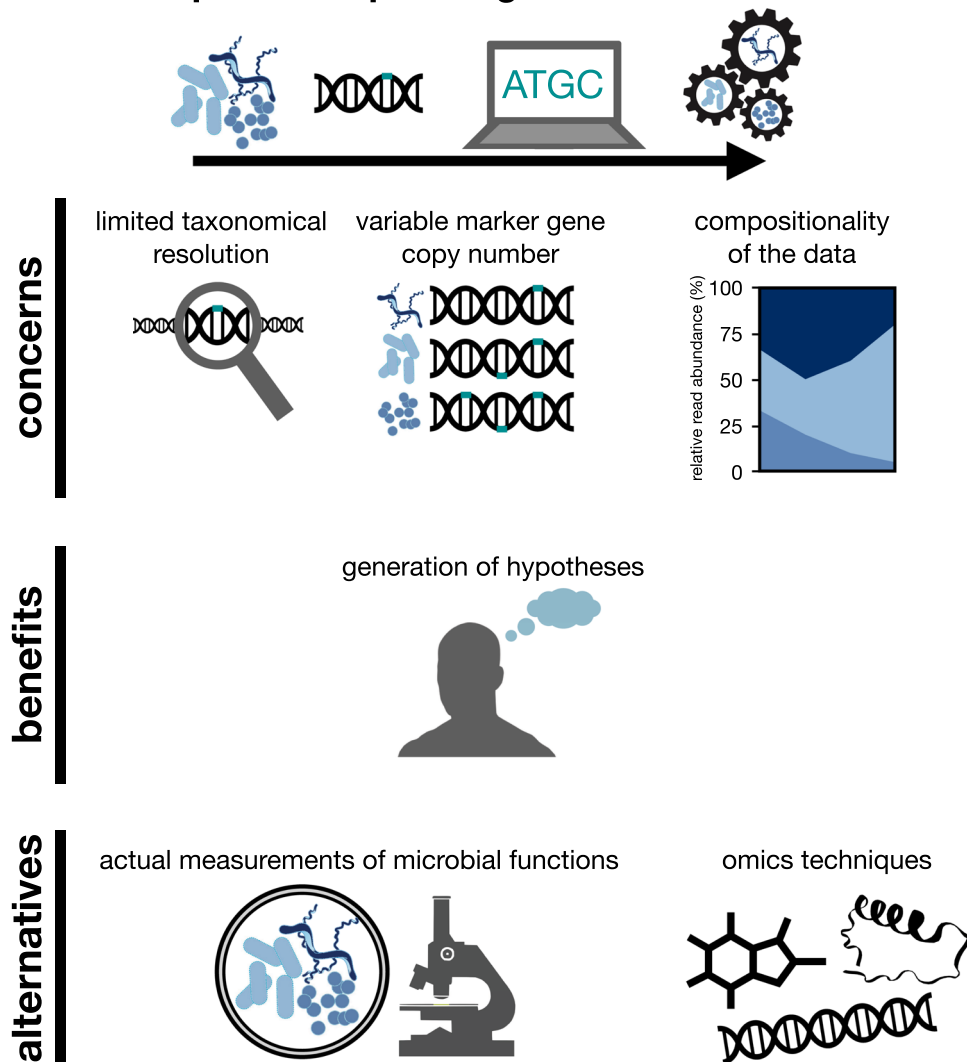
A further strategy to circumvent the caveats of compositional data analysis is the use of ratios (division between elements of the composition, i.e., proportions) [36]. This is because the microbial load bias vanishes in ratios [37]. So if the microbiome function of interest can be expressed in terms of ratios between molecules/functions/processes, as is often the case in biological systems (e.g., carbon/nitrogen, albumin/globulin, neutrophil/lymphocyte), relative abundances of potential functions turn into valuable information and quantification of the microbial load becomes less critical. Although not always practical, such that our concerns regarding the compositional nature of the data still hold, we suggest the use of ratios whenever possible when analyzing functional potential profiles. Regardless, compositionally-aware statistical methods, which employ ratio-based transformations to make the data less contingent on community sizes, are also useful and should be favored in this setting [29].

Finally, although not discussed here, it is worth mentioning that PCR bias is a well-known source of error that can distort community composition (e.g., [38]) and consequently predicted functional potential profiles.

## Measuring the actual function of the microbiome

Another but more demanding approach is to measure the actual functions of interest of the microbiome if possible. For example, processes regulated by environmental microbiomes such as enzymatic activity [39], greenhouse gas fluxes [40], and nitrogen fixation [41] can often be measured in situ. Likewise, the function can be assessed from human/animal gut microbiomes by orthogonal measures of stool metabolites, which has been successfully used to investigate the relationship between the gut

## prediction of functional potential profiles from amplicon-sequencing microbiome data sets



**FIGURE 1** Schematic illustration of the concerns, benefits, and alternatives to the prediction of microbiome functional potential from amplicon sequencing microbiome data sets

microbiome and biotransformations that partly explain the interpatient variability in the efficacy and toxicity of several drugs [42], including immunosuppressants [43]. Measuring actual microbial functions is only an option if microbial communities are accessible to functional measurements, functionally active at the point of sampling, and have a sufficient population size and amount of sampling material to detect the functions. For microbiomes that do not meet these criteria, additional quantification of functional genes using techniques like real-time PCR is a valuable addition to quantifying their genetic potential for a given function. It has to be noted though that these genetic potentials do not necessarily translate into microbial activities and processes [44–46]. In this regard, omics techniques such as transcriptomics,

proteomics, and metabolomics could assist with the exploration of expressed genetic potentials. We argue that a combination of the measurement of selected actual functions and the predicted functional potential profiles is a powerful approach to understanding microbial functioning.

### Generating and testing hypotheses

Despite their limitations, we recognize the particular potential of predicted functional potential profiles for the generation of novel hypotheses. However, it is important not to neglect that microbiomes are complex. Accordingly, functional potential profiles are high-dimensional

and can be hard to analyze. This means that oftentimes functional potential prediction tools lead to too many research directions so that generating straightforward hypotheses becomes difficult. We strongly encourage researchers to select meaningful hypotheses and to independently test these hypotheses whenever possible. For example, Zhang et al. [47] predicted the metabolic functions of gut microbiota in mice and successfully validated their predictions by using nuclear magnetic resonance-based metabolomics. Likewise, Wu et al. [48] predicted altered biosynthesis pathways of the gut microbiome in patients with colorectal adenomas and colorectal cancer as compared to healthy individuals. The authors validated their predictions by quantifying genes from these pathways using real-time PCR. Even though studies validating predicted functional potential profiles are the exception rather than the rule, these examples illustrate the power of predicted functional potential profiles to find and explore new research directions.

## SUMMARY

An overview of the concerns, benefits, and alternatives to the prediction of microbiome functional potential discussed in this study is provided in Figure 1. Overall, we appreciate the efforts undertaken to enable the prediction of functional potential profiles from taxonomical microbiome data sets. We also believe that the prediction of functional potential profiles is useful to generate new ideas and explore new potential research directions. In conclusion, however, we argue that functions of microbiomes predicted from microbiome DNA read count data generated by short-read amplicon sequencing should not serve as the only basis to draw biological inferences. We believe that the transition from short- to long-read sequencing technologies will help to overcome some of the challenges presented here. Still, high-resolution taxonomic profiling does not resolve crucial concerns we raised (e.g., the compositional nature of microbiome data). Therefore, measuring microbiome activity using omics (e.g., metabolomics) and non-omics approaches (e.g., qPCR) will remain essential alongside taxonomic profiling to illuminate microbiome functioning.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

## ACKNOWLEDGMENTS

Vitor Heidrich was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, process no.

13996-0/2018). We thank Anamaria Camargo for critically revising the manuscript.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## DATA AVAILABILITY STATEMENT

No new data and script were used in this study. Supplementary materials (figures, tables, scripts, graphical abstract, slides, videos, Chinese translated version, and updated materials) may be found in the online DOI or iMeta Science <http://www.imeta.science>

## ORCID

Vitor Heidrich  <https://orcid.org/0000-0001-6617-9187>

Lukas Beule  <https://orcid.org/0000-0003-1107-7540>

## REFERENCES

1. Barberán, Albert, Joshua Ladau, Jonathan W. Leff, Katherine S. Pollard, Holly L. Menninger, Robert R. Dunn, and Noah Fierer. 2015. "Continental-Scale Distributions of Dust-Associated Bacteria and Fungi." *Proceedings of the National Academy of Sciences of the United States of America* 112: 5756–5761. <https://doi.org/10.1073/pnas.1420815112>
2. Siles, José A., and Rosa Margesin. 2016. "Abundance and Diversity of Bacterial, Archaeal, and Fungal Communities Along an Altitudinal Gradient in Alpine Forest Soils: What Are the Driving Factors." *Microbial Ecology* 72: 207–220. <https://doi.org/10.1007/s00248-016-0748-2>
3. Zhang, Likui, Manyu Kang, Jiajun Xu, Jian Xu, Yinjie Shuai, Xiaojian Zhou, Zhihui Yang, and Kesen Ma. 2016. "Bacterial and Archaeal Communities in the Deep-Sea Sediments of Inactive Hydrothermal Vents in the Southwest India Ridge." *Scientific Reports* 6: 25982. <https://doi.org/10.1038/srep25982>
4. Djemiel, Christophe, Pierre-Alain Maron, Sébastien Terrat, Samuel Dequiedt, Aurélien Cottin, and Lionel Ranjard. 2022. "Inferring Microbiota Functions from Taxonomic Genes: A Review." *GigaScience* 11: giab090. <https://doi.org/10.1093/gigascience/giab090>
5. Blackwood, Kym S., Christine Y. Turenne, Dag Harmsen, and Amin M. Kabani. 2004. "Reassessment of Sequence-Based Targets for Identification of Bacillus Species." *Journal of Clinical Microbiology* 42: 1626–1630. <https://doi.org/10.1128/JCM.42.4.1626-1630.2004>
6. Ash, Carol, John A. Farrow, Mathias Dorsch, Erko Stackebrandt, and Matthew D. Collins. 1991. "Comparative Analysis of Bacillus Anthracis, Bacillus Cereus, and Related Species on the Basis of Reverse Transcriptase Sequencing of 16S rRNA." *International Journal of Systematic Bacteriology* 41: 343–346. <https://doi.org/10.1099/00207713-41-3-343>
7. Bavykin, Sergei G., Yuri P. Lysov, Vladimir Zakhariyev, John J. Kelly, Joany Jackman, David A. Stahl, and Alexey Cherni. 2004. "Use of 16S rRNA, 23S rRNA, and GyrB Gene Sequence Analysis to Determine Phylogenetic Relationships of Bacillus Cereus Group Microorganisms." *Journal of Clinical Microbiology* 42: 3711–3730. <https://doi.org/10.1128/JCM.42.8.3711-3730.2004>



8. Carrow, Hannah C., Lakshmi E. Batachari, and Hiutung Chu. 2020. "Strain Diversity in the Microbiome: Lessons from *Bacteroides fragilis*." *PLoS Pathogens* 16: e1009056. <https://doi.org/10.1371/journal.ppat.1009056>
9. Yan, Yan, Long H. Nguyen, Eric A. Franzosa, and Curtis Huttenhower. 2020. "Strain-Level Epidemiology of Microbial Communities and the Human Microbiome." *Genome Medicine* 12: 71. <https://doi.org/10.1186/s13073-020-00765-y>
10. Johnson, Jethro S., Daniel J. Spakowicz, Bo-Young Hong, Lauren M. Petersen, Patrick Demkowicz, Lei Chen, Shana R. Leopold, et al. 2019. "Evaluation of 16S rRNA Gene Sequencing for Species and Strain-Level Microbiome Analysis." *Nature Communications* 10: 5029. <https://doi.org/10.1038/s41467-019-13036-1>
11. Callahan, Benjamin J., Dmitry Grinevich, Siddhartha Thakur, Michael A. Balamotis, and Tuval Ben Yehzekel. 2021. "Ultra-Accurate Microbial Amplicon Sequencing with Synthetic Long Reads." *Microbiome* 9: 130. <https://doi.org/10.1186/s40168-021-01072-3>
12. Fuks, Garold, Michael Elgart, Amnon Amir, Amit Zeisel, Peter J. Turnbaugh, Yoav Soen, and Noam Shental. 2018. "Combining 16S rRNA Gene Variable Regions Enables High-Resolution Microbial Community Profiling." *Microbiome* 6: 17. <https://doi.org/10.1186/s40168-017-0396-x>
13. Jaspers, Elke, and Jörg Overmann. 2004. "Ecological Significance of Microdiversity: Identical 16S rRNA Gene Sequences Can be Found In Bacteria with Highly Divergent Genomes and Ecophysiologicals." *Applied and Environmental Microbiology* 70: 4831–4839. <https://doi.org/10.1128/AEM.70.8.4831-4839.2004>
14. Lofgren, Lotus A., Jessie K. Uehling, Sara Branco, Thomas D. Bruns, Francis Martin, and Peter G. Kennedy. 2019. "Genome-Based Estimates of Fungal rDNA Copy Number Variation Across Phylogenetic Scales and Ecological Lifestyles." *Molecular Ecology* 28: 721–730. <https://doi.org/10.1111/mec.14995>
15. Herrera, Monica L., Ana C. Vallor, Jonathan A. Gelfond, Thomas F. Patterson, and Brian L. Wickes. 2009. "Strain-Dependent Variation in 18S Ribosomal DNA Copy Numbers In *Aspergillus Fumigatus*." *Journal of Clinical Microbiology* 47: 1325–1332. <https://doi.org/10.1128/JCM.02073-08>
16. Alanio, Alexandre, Aude Sturny-Leclère, Marion Benabou, Nicolas Guigue, and Stéphane Bretagne. 2016. "Variation In Copy Number of the 28S rDNA of *Aspergillus Fumigatus* Measured by Droplet Digital PCR and Analog Quantitative Real-Time PCR." *Journal of Microbiological Methods* 127: 160–163. <https://doi.org/10.1016/j.mimet.2016.06.015>
17. Zhao, Shu, and John G. Gibbons. 2018. "A Population Genomic Characterization of Copy Number Variation in the Opportunistic Fungal Pathogen *Aspergillus Fumigatus*." *PLoS One* 13: e0201611. <https://doi.org/10.1371/journal.pone.0201611>
18. Lavrinienko, Anton, Toni Jernfors, Janne J. Koskimäki, Anna Maria Pirttilä, and Phillip C. Watts. 2021. "Does Intraspecific Variation In rDNA Copy Number Affect Analysis Of Microbial Communities? *Trends in Microbiology* 29: 19–27. <https://doi.org/10.1016/j.tim.2020.05.019>
19. Rainey, Frederick A., Naomi L. Ward-Rainey, Peter H. Janssen, Hans Hippe, and Erko Stackebrandt. 1996. "Clostridium Paradoxum DSM 7308T Contains Multiple 16S rRNA Genes with Heterogeneous Intervening Sequences." *Microbiology* 142(Pt 8): 2087–2095. <https://doi.org/10.1099/13500872-142-8-2087>
20. Schloss, Patrick D. 2021. "Amplicon Sequence Variants Artificially Split Bacterial Genomes into Separate Clusters." *mSphere* 6: e0019121. <https://doi.org/10.1128/mSphere.00191-21>
21. Angly, Florent E., Paul G. Dennis, Adam Skarszewski, Inka Vanwonterghem, Philip, Hugenholtz, and Gene W. Tyson. 2014. "CopyRighter: A Rapid Tool for Improving the Accuracy of Microbial Community Profiles Through Lineage-Specific Gene Copy Number Correction." *Microbiome* 2: 11. <https://doi.org/10.1186/2049-2618-2-11>
22. Stoddard, Steven F., Byron J. Smith, Robert Hein, Benjamin R. K. Roller, and Thomas M. Schmidt. 2015. "rrnDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development." *Nucleic Acids Research* 43: D593–D598. <https://doi.org/10.1093/nar/gku1201>
23. Louca, Stilianos, Michael Doebeli, and Laura Wegener Parfrey. 2018. "Correcting for 16S rRNA Gene Copy Numbers in Microbiome Surveys Remains an Unsolved Problem." *Microbiome* 6: 41. <https://doi.org/10.1186/s40168-018-0420-9>
24. Starke, Robert, Victor Satler Pylo, and Daniel Kumazawa Morais. 2021. "16S rRNA Gene Copy Number Normalization Does Not Provide More Reliable Conclusions in Metataxonomic Surveys." *Microbial Ecology* 81: 535–539. <https://doi.org/10.1007/s00248-020-01586-7>
25. Nübel, Ulrich, Bert Engelen, Andreas Felske, Jir Snaidr, A. Wieshuber, Rudolf I. Amann, Wolfgang Ludwig, and Horst Backhaus. 1996. "Sequence Heterogeneities of Genes Encoding 16S rRNAs in *Paenibacillus Polymyxa* Detected by Temperature Gradient Gel Electrophoresis." *Journal of Bacteriology* 178: 5636–5643. <https://doi.org/10.1128/jb.178.19.5636-5643.1996>
26. Strube, Mikael Lenz. 2021. "RibDif: Can Individual Species be Differentiated by 16S Sequencing? *Bioinformatics Advances* 1: vbab020. <https://doi.org/10.1093/bioadv/vbab020>
27. Sun, Dong-Lei, Xuan Jiang, Qinglong L. Wu, and Ning-Yi Zhou. 2013. "Intragenomic Heterogeneity of 16S rRNA Genes Causes Overestimation of Prokaryotic Diversity." *Applied and Environmental Microbiology* 79: 5962–5969. <https://doi.org/10.1128/AEM.01282-13>
28. Vandeputte, Doris, Gunter Kathagen, Kevin D'hoel, Sara Vieira-Silva, Mireia Valles-Colomer, João Sabino, Jun Wang, et al. 2017. "Quantitative Microbiome Profiling Links Gut Community Variation to Microbial Load." *Nature* 551: 507–511. <https://doi.org/10.1038/nature24460>
29. Gloor, Gregory B., Jean M. Macklaim, Vera Pawlowsky-Glahn, and Juan J. Egozcue. 2017. "Microbiome Datasets are Compositional: And this is not Optional." *Frontiers in Microbiology* 8: 2224. <https://doi.org/10.3389/fmicb.2017.02224>
30. Lou, Jun, Li Yang, Haizhen Wang, Laosheng Wu, and Jianming Xu. 2018. "Assessing Soil Bacterial Community and Dynamics by Integrated High-Throughput Absolute Abundance Quantification." *PeerJ* 6: e4514. <https://doi.org/10.7717/peerj.4514>
31. Beule, Lukas, Markus Arndt, and Petr Karlovsky. 2021. "Relative Abundances of Species or Sequence Variants Can Be Misleading: Soil Fungal Communities as an

- Example.” *Microorganisms* 9: 589. <https://doi.org/10.3390/microorganisms9030589>
32. Tkacz, Andrzej, Marion Hortala, and Philip S. Poole. 2018. “Absolute Quantitation of Microbiota Abundance In Environmental Samples.” *Microbiome* 6: 110. <https://doi.org/10.1186/s40168-018-0491-7>
  33. Barlow, Jacob T., Said R. Bogatyrev, and Rustem F. Ismagilov. 2020. “A Quantitative Sequencing Framework for Absolute Abundance Measurements of Mucosal and Lumenal Microbial Communities.” *Nature Communications* 11: 2590. <https://doi.org/10.1038/s41467-020-16224-6>
  34. Jian, Ching, Panu Luukkonen, Hannele Yki-Järvinen, Anne Salonen, and Katri Korpela. 2020. “Quantitative PCR Provides a Simple and Accessible Method for Quantitative Microbiota Profiling.” *PLoS One* 15: e0227285. <https://doi.org/10.1371/journal.pone.0227285>
  35. Galazzo, Gianluca, Niels van Best, Birke J. Benedikter, Kevin Janssen, Liene Bervoets, Christel Driessen, Melissa Oomen, et al. 2020. “How to Count Our Microbes? The Effect of Different Quantitative Microbiome Profiling Approaches.” *Frontiers in Cellular and Infection Microbiology* 10: 403. <https://doi.org/10.3389/fcimb.2020.00403>
  36. Pawlowsky-Glahn, Vera, Juan J. Egozcue, and Raimon Tolosana-Delgado. 2015. *Modeling and Analysis of Compositional Data*. London: John Wiley & Sons.
  37. Morton, James T., Clarisse Marotz, Alex Washburne, Justin Silverman, Livia S. Zaramela, Anna Edlund, Karsten Zengler, and Rob Knight. 2019. “Establishing Microbial Composition Measurement Standards with Reference Frames.” *Nature Communications* 10: 2719. <https://doi.org/10.1038/s41467-019-10656-5>
  38. Silverman, Justin D., Rachael J. Bloom, Sharon Jiang, Heather K. Durand, Eric Dallow, Sayan Mukherjee, and Lawrence A. David. 2021. “Measuring and Mitigating PCR Bias In Microbiota Datasets.” *PLoS Computational Biology* 17: e1009113. <https://doi.org/10.1371/journal.pcbi.1009113>
  39. Dong, Shufu, Denise Brooks, Melanie D. Jones, and Susan J. Grayston. 2007. “A Method for Linking In Situ Activities of Hydrolytic Enzymes to Associated Organisms in Forest Soils.” *Soil Biology and Biochemistry* 39: 2414–2419. <https://doi.org/10.1016/j.soilbio.2007.03.030>
  40. Savage, Kathleen, Rebecca Phillips, and Eric Davidson. 2014. “High Temporal Frequency Measurements of Greenhouse Gas Emissions from Soils.” *Biogeosciences* 11: 2709–2720. <https://doi.org/10.5194/bg-11-2709-2014>
  41. Hardy, Ralph W., Richard D. Holsten, E. K. Jackson, and Richard C. Burns. 1968. “The Acetylene-Ethylene Assay for n(2) Fixation: Laboratory and Field Evaluation.” *Plant Physiology* 43: 1185–1207. <https://doi.org/10.1104/pp.43.8.1185>
  42. Tsunoda, Shirley M., Christopher Gonzales, Alan K. Jarmusch, Jeremiah D. Momper, and Joseph D. Ma. 2021. “Contribution of the Gut Microbiome to Drug Disposition, Pharmacokinetic and Pharmacodynamic Variability.” *Clinical Pharmacokinetics* 60: 971–984. <https://doi.org/10.1007/s40262-021-01032-y>
  43. Guo, Yukuang, Camila Manoel Crnkovic, Kyoung-Jae Won, Xiaotong Yang, John Richard Lee, Jimmy Orjala, Hyunwoo Lee, and Hyunyoung Jeong. 2019. “Commensal Gut Bacteria Convert the Immunosuppressant Tacrolimus to Less Potent Metabolites.” *Drug Metabolism and Disposition: The Biological Fate of Chemicals* 47: 194–202. <https://doi.org/10.1124/dmd.118.084772>
  44. Louca, Stilianos, Laura Wegener Parfrey, and Michael Doebeli. 2016. “Decoupling Function and Taxonomy in the Global Ocean Microbiome.” *Science* 353: 1272–1277. <https://doi.org/10.1126/science.aaf4507>
  45. Heintz-Buschart, Anna, and Paul Wilmes. 2018. “Human Gut Microbiome: Function Matters.” *Trends in Microbiology* 26: 563–574. <https://doi.org/10.1016/j.tim.2017.11.002>
  46. Jansson, Janet K., and Kirsten S. Hofmockel. 2018. “The Soil Microbiome—from Metagenomics to Metaphenomics.” *Current Opinion in Microbiology* 43: 162–168. <https://doi.org/10.1016/j.mib.2018.01.013>
  47. Zhang, Limin, Cen Xie, Robert G. Nichols, Siu H. J. Chan, Changtao Jiang, Ruixin Hao, Philip B. Smith, et al. 2016. “Farnesoid X Receptor Signaling Shapes the Gut Microbiota and Controls Hepatic Lipid Metabolism.” *mSystems* 1: e00070-16. <https://doi.org/10.1128/mSystems.00070-16>
  48. Wu, Yuanqi, Na Jiao, Ruixin Zhu, Yida Zhang, Dingfeng Wu, An-Jun Wang, Sa Fang, et al. 2021. “Identification of Microbial Markers Across Populations in Early Detection of Colorectal Cancer.” *Nature Communications* 12: 3063. <https://doi.org/10.1038/s41467-021-23265-y>

**How to cite this article:** Heidrich, Vitor, and Lukas Beule. 2022. “Are Short-Read Amplicons Suitable for the Prediction of Microbiome Functional Potential? A Critical Perspective.” *iMeta* 1, e38. <https://doi.org/10.1002/imt2.38>