Received: 26 September 2023 | Revised: 12 November 2023 | Accepted: 15 November 2023

DOI: 10.1111/mec.17223

OPINION

MOLECULAR ECOLOGY WILEY

Molecular ecology of microbiomes in the wild: Common pitfalls, methodological advances and future directions

Nicholas M. Fountain-Jones[1] | **Tatiana Giraud**[2] | **Lucie Zinger**[3,4] | **Holly Bik**[5] | **Simon Creer**[6] | **Elin Videvall**[7,8,9]

[1]School of Natural Sciences, University of Tasmania, Hobart, Tasmania, Australia

[2]Laboratoire Ecologie Systématique et Evolution, UMR 8079, Bâtiment 680, Université Paris-Saclay, CNRS, AgroParisTech, Gif-sur-Yvette, France

[3]Institut de Biologie de l'École Normale Supérieure (IBENS), École Normale Supérieure, CNRS, INSERM, PSL Université Paris, Paris, France

[4]Laboratoire Evolution et Diversité Biologique (EDB), UMR5174, CNRS, Institut de Recherche pour le Développement (IRD), Université Toulouse 3 Paul Sabatier, Toulouse, France

[5]Department of Marine Sciences and Institute of Bioinformatics, University of Georgia, Athens, Georgia, USA

[6]School of Environmental and Natural Sciences, Bangor University, Bangor, UK

[7]Department of Ecology, Evolution and Organismal Biology, Brown University, Providence, Rhode Island, USA

[8]Institute at Brown for Environment and Society, Brown University, Providence, Rhode Island, USA

[9]Department of Ecology and Genetics, Uppsala University, Uppsala, Sweden

Correspondence
Nicholas M. Fountain-Jones, School of Natural Sciences, University of Tasmania, Hobart, Tas. 7001, Australia.
Email: nfountainjones@gmail.com

Handling Editor: Loren Rieseberg

Abstract

The study of microbiomes across organisms and environments has become a prominent focus in molecular ecology. This perspective article explores common challenges, methodological advancements, and future directions in the field. Key research areas include understanding the drivers of microbiome community assembly, linking microbiome composition to host genetics, exploring microbial functions, transience and spatial partitioning, and disentangling non-bacterial components of the microbiome. Methodological advancements, such as quantifying absolute abundances, sequencing complete genomes, and utilizing novel statistical approaches, are also useful tools for understanding complex microbial diversity patterns. Our aims are to encourage robust practices in microbiome studies and inspire researchers to explore the next frontier of this rapidly changing field.

KEYWORDS
bioinformatics, eDNA, host-parasite interactions, microbial communities, statistics

1 | INTRODUCTION

The study of microbial communities across host species and environments (hereafter the ‘microbiome’) is a major focus of research in the field of molecular ecology. As such, *Molecular Ecology* and *Molecular Ecology Resources* regularly publish papers and special issues in the field. *Molecular Ecology* is at the forefront of publishing research on the importance of microbiomes across ecosystems, from increasing our understanding of host–pathogen (Bergner et al., 2020; Wille et al., 2018) and host–symbiont interactions (Rubin et al., 2019) to investigating the impact of climate and other environmental factors on microbial populations (Santos-Júnior et al., 2022; Wu et al., 2022). Complementing these efforts, *Molecular Ecology Resources* publishes significant methodological advances that continue to shape the field

(Harrison et al., 2020; Schnell et al., 2015; Stothart et al., 2021). Based on our collective experience as subject editors at *Molecular Ecology* and *Molecular Ecology Resources*, here we discuss some of the best practices and advances across the discipline for analysing microbiomes, from study design to data analysis, and highlight anticipated future directions in the field. With this article, we hope to inspire and encourage researchers to obtain more robust insights from microbiome data, which will enable the field to advance and tackle the new horizons enabled by recent advances in technology.

2 | STUDY DESIGN

As in any study, the sampling or experimental design of microbiome studies should include sufficient independent replicates, avoiding confounding effects as much as possible, with the samples representing appropriate ecological scales given the processes investigated. Microbiome sampling design must also be well-planned and appropriate to the specific hypothesis that is being tested. When testing hypotheses pertaining to the impact of outlier external drivers (e.g. fire, pollution events, natural disasters), studies would ideally feature samples that were collected both before and after the event, that are not confounded by habitat type, geography or physicochemistry. Before embarking on microbiome studies in the wild, particularly those of which are opportunistic (i.e. with samples originally collected for other purposes), researchers should carefully consider if autocorrelation of factors beyond their control could impede the interpretation of results. In other words, researchers must be realistic about what can be accomplished with limited sample sets, since rigorous hypothesis testing requires equally rigorous sampling protocols and study design.

In addition, the sampling of microbial communities should take into account their high heterogeneity at small spatial scales due to micro/mesoscale heterogeneity of their environment (Vos et al., 2013; Zhang et al., 2014) or neutral assembly dynamics (Woodcock et al., 2007). For example, composite samples (i.e. pooled individual samples) can be combined prior to homogenization and sub-sampling, in order to reduce the local, micro-scale heterogeneity if it is irrelevant to the questions being studied (George et al., 2019). Here, knowledge of how, and at what scale, the target community responds to external drivers will inform adequate sampling design. For example, a composite 0.2 mg sediment sample is likely to be representative of the bacterial, archaeal and microbial eukaryotic biospheres, but not for metazoan communities, due to issues of biological scale and heterogeneous distribution (Nascimento et al., 2018). Over 14 g of homogenized sediment was needed to achieve adequate beta diversity measurements for meiofauna species (Nascimento et al., 2018). Smaller samples will contain some microscopic taxa and trace environmental DNA but they are inadequate at representing the underlying meio- and macro-faunal communities. As the target organisms grow in size, the sample volume and spatial extent of the studied area should be correspondingly expanded, although we note that extraction protocols and tissue type are also important factors. We encourage readers to refer to taxon- or study-specific guides for more granular information about study design (e.g. diet: Johnson et al., 2020; host-associated microbiomes: Kohl, 2017; fungi: Tedersoo et al., 2022).

3 | SAMPLE PREPARATION AND WORKFLOW

In addition to the above considerations, the study design needs to account for the sensitivity and error-prone nature of many molecular-based approaches. Both shotgun metagenomics and DNA metabarcoding (i.e. amplicon sequencing of marker genes) include numerous opportunities for introducing false negatives and positives during the data generation process, starting from sample collection to the laboratory, sequencing, bioinformatics and data analyses. Details of these issues are already largely covered in a *Molecular Ecology* editorial (Zinger et al., 2019), but to summarize briefly, some possible pitfalls include sample contamination stemming from the field or laboratory environment (de Goffau et al., 2018; Salter et al., 2014), extraction/PCR amplification biases and errors generated during PCR and sequencing. Technical considerations, such as sample volume and choice of laboratory reagents, are, in many cases, the result of a compromise between the research question, logistical feasibility, time and available funds (Taberlet et al., 2018). However, any compromise of the protocol should still allow one to appropriately address the research question. In addition, we want to re-emphasize the importance of adequately describing the whole data production workflow in the methods section of manuscripts (e.g. primer sequences, polymerase, molecular labelling/indexing strategy).

Environmental or laboratory contamination is a particularly large problem for samples with low microbial biomass (Eisenhofer et al., 2019), and the collection of such samples cannot be avoided in many study designs (e.g. host-associated microbiomes of small organisms or depauperate environmental habitats). The sequencing of negative controls including DNA extraction blanks and PCR negative controls (and potentially also positive controls and technical replicates) alongside experimental samples is important for quantifying errors and artefacts (Davis et al., 2018), and can improve data curation procedures through tuned, experiment-specific criteria, including for samples with low microbial biomass. While there is more than one way to implement such efforts, a thorough description of the controls, a rationale for including them, and the ways they are integrated into data analysis, are essential practices of good microbiome science (Hakimzadeh et al., 2023).

One overlooked problem in microbiome studies is cross-contamination between samples during library preparation procedures (Kim et al., 2017; Zinger et al., 2019), which can result in an artificial reduction in beta diversity (i.e. compositional differences between samples) and an increase in alpha diversity. Such cross-contamination can occur during the PCR plate preparation process through pipetting errors or aerosol production. Random positioning of samples in the PCR plates provides a relatively simple approach to reduce this problem (Minich et al., 2019; Taberlet et al., 2018). More often – and insidiously – cross-contamination can occur during the PCR cycles, a bias referred to as tag-switches (Carlsen et al., 2012; Esling et al., 2015), tag-jumps (Schnell et al., 2015), or, more recently in the microbiome literature, cross-talks (Edgar, 2018; Minich et al., 2019). During this laboratory step, amplicon molecules from different samples can recombine within conserved primer sequences, resulting in the production of new molecules containing the genuine DNA sequence, but the wrong sample/barcode label. As a consequence, the most abundant taxa will be detected in many samples, including the negative controls (Esling et al., 2015; Minich et al., 2019; Taberlet et al., 2018), preventing simple removal of all taxa occurring in controls as a fix for field and laboratory contamination. Alleviating the problem of tag-switches can be achieved with modified library preparation protocols (Carøe & Bohmann, 2020), appropriate sample labelling strategies and a posteriori using the information contained in both samples and negative controls (Bohmann et al., 2022; Hakimzadeh et al., 2023).

One limitation in microbiome studies, using either DNA metabarcoding or metagenomics, is the compositional nature of the sequence data (i.e. described as proportions, not absolute number of molecules). Like other count-based sequencing approaches, this limitation arises because the total number of reads in high-throughput sequencing platforms is uninformative (Gloor et al., 2017). As a result, we are unable to obtain data on absolute abundances and biomass for the different microbial members of the community. However, exciting new developments are emerging to overcome these limitations, relying on known reference values of DNA sequence abundances, allowing simple conversion of relative abundance into absolute values. Two broad classes of methods based on this approach have emerged: (i) quantification of target markers using q/ddPCR (quantitative/digital droplet PCR) prior to metabarcoding (Barlow et al., 2020; Callahan et al., 2019; Ji et al., 2019) and (ii) introduction of exogenous DNA spike-ins (i.e. DNA molecules of known sequence and quantity to calibrate measurements) (Harrison et al., 2020). Related to the latter, a host-associated microbiome PCR approach (HamPCR, Lundberg et al., 2021) represents a promising method to assess the ratio of the microbial population size relative to the amount of host tissue (i.e. microbial load). If these approaches are unavailable, log-ratio transformations of the data may be a useful method to reveal abundance changes in microbial composition (see Greenacre et al., 2021; Morton et al., 2019).

Despite obtaining a better estimate of the absolute number of molecules in a sample, it is still challenging to convert this number into the actual number of microbial cells. The calculation is often difficult because most gold-standard barcoding genes used for bacteria, fungi, and protists have multiple copies in the genome, with precise numbers varying across taxa and in unpredictable ways (Louca et al., 2018). Another problem is that some of the retrieved molecules can be derived from extracellular DNA or DNA adsorbed on cell debris or particles, that is, correspond to non-living organisms (Torti et al., 2015). The proportion of extracellular DNA in the environment is often not known but can be estimated with different approaches (reviewed by Nagler et al., 2022). Further, because extracellular DNA is often degraded, approaches including long-read sequencing targeting larger genomic regions will likely overcome this issue. Approaches that are able to quantify or eliminate extracellular DNA can prove useful when having correct snapshots of the microbial community is crucial (e.g. when studying short-term processes with repeated observations capturing microbiome variation within host individuals) but are likely less relevant when studying processes operating at larger temporal scales (e.g. microbiome response to climate change).

Lastly, incorporating site-occupancy modelling in microbiome studies presents an exciting avenue to quantify measurement uncertainty and to account for imperfect detection (Ficetola et al., 2015; McClenaghan et al., 2020; Willoughby et al., 2016). Site-occupancy models use data collected over multiple visits to sites (or across multiple technical/biological replicates) to quantify how likely it is to detect a taxon when it is present. For microbiome studies, including both biological and technical PCR replicates can enable rigorous statistical predictions regarding the true or false positive detection of microbial species within the community. Further, these predictions can be utilized to improve study design and provide increased confidence in metabarcoding datasets used for biodiversity assessments and monitoring (Fukaya et al., 2022; McClenaghan et al., 2020). How many replicates of each type are required is an open question, although biological replicates may improve detection probabilities (Willoughby et al., 2016).

4 | BEYOND ESTIMATING DIVERSITY: EXCITING ADVANCES IN STATISTICS

Modelling advances in community ecology offer exciting opportunities to understand the complex patterns in microbial diversity and complement robust sampling designs (Grantham et al., 2020; Trego et al., 2022). In addition, novel methods for analysing amplicon sequencing data are continuously emerging, primarily focused on the human gut microbiome but adaptable to other microbial ecology fields with suitable study designs and datasets (Trego et al., 2022). These tools tackle a broad range of ecological and evolutionary questions from quantifying community assembly processes, mapping occurrence networks, capturing spatial/temporal dynamics, integrating multi-omics, identifying differentially abundant taxa, finding species-environment associations and predicting functional patterns (Trego et al., 2022). However, despite the frequent use of high-throughput sequencing, there has been a slow uptake of these new analytical techniques, and many studies do not go much beyond basic comparisons of alpha and beta diversity estimates across samples. While important inferences can be made by examining overall patterns of composition and diversity (Grosser et al., 2019; Motta et al., 2018), they offer only a starting point towards having a more mechanistic understanding of the ecological drivers of microbiome variation (Shade, 2017).

Methods used to understand microbiome variation can be broadly classified as distance-based (i.e. model pair-wise distances between samples), model-based (i.e. combining univariate models of each taxon into a community-wide model) and differential abundance models (i.e. models estimating which taxa are significantly more abundant than others). For brevity, we will not discuss differential abundance methods but see, for example, Nearing et al. (2022) for a comparison of methods. Common distance-based approaches to quantify differences in beta diversity across microbiome samples, such as the permutational multivariate analysis of variance (PERMANOVA), are algorithmic (i.e. not based on a statistical model) and do not explicitly account for uncertainty in ecological data (Björk et al., 2018; Warton et al., 2012, 2015). Importantly, making inferences about microbiome variation is often difficult using algorithmic distance-based approaches (Björk et al., 2018; Warton et al., 2012). Model-based approaches such as joint species distribution models (JSDMs) or stacked models (Powell-Romero et al., 2023) are multi-response extensions of generalized linear mixed models (GLMMs) that can overcome some of the limitations of the algorithmic distance-based methods to elucidate patterns of microbiome variation (Björk et al., 2018; Grantham et al., 2020). Often using a Bayesian framework, JSDMs simultaneously analyse multiple species and environmental variables, allowing for the assessment of community-level responses to environmental change and host effects (Björk et al., 2018; Ovaskainen et al., 2017; Pollock et al., 2014). JSDMs can (i) incorporate information on species traits and phylogenetic relatedness, improving estimation accuracy and power when there is a phylogenetic signal (Ovaskainen et al., 2017), and (ii) analyse patterns of taxon covariance to infer microbial co-occurrence networks (Björk et al., 2018; Fountain-Jones et al., 2020, 2023). Microbial co-occurrence networks are valuable tools in microbiome science, as they offer insights (but see *Current gaps and future directions* below) into the associations among microbial taxa, enhancing our understanding of microbial community dynamics and functioning. JSDM-based co-occurrence networks have an added advantage of interpretation as the major environmental and host effects shaping microbial presences are controlled for (i.e. an inferred association between microbes is then not likely a mere product of a shared environmental response). However, GLMM-based JSDM co-occurrence networks cannot untangle the relative roles of taxa associations, and environmental or host effects (Clark et al., 2018; Fountain-Jones et al., 2020) and tend to not scale well with large datasets (Pichler & Hartig, 2021). Approaches such as conditional random fields (CRF, Clark et al., 2018), multi-response interpretable machine learning (mrIML, Fountain-Jones et al., 2021), MIMIX (Microbiome MIXed Model, Grantham et al., 2020) and scalable JSDMs (sjSDM, Pichler & Hartig, 2021) can overcome these limitations. Importantly, approaches such as MrIML and MiMiX allow for predictions and treatment effects to be extracted for individual taxa, which can be useful if researchers have a set of focal taxa. We note that these methods are not appropriate in all situations. For particularly large datasets (thousands of samples), new distance-based methods such as D-MANOVA (Chen & Zhang, 2021) or multivariate distance matrix regression (MDMR, Zapala & Schork, 2012) may be better options. Boshuizen and te Beest (2023) provided a useful guide of the pitfalls in analysing amplicon data. While the tools mentioned here represent only a tiny fraction of the potential methods available, we encourage readers to go beyond diversity metrics and differentially abundant taxa to gain more mechanistic insights into microbiome data from wild species.

Incorporating some of the methodological advances in bioinformatics and statistics, coupled with robust study design and rigorous laboratory techniques, will improve current research efforts in the field (see Figure 1 for a summary). Moreover, taking into consideration both the limitations and opportunities of these various approaches allow us to open up new exciting avenues in the field of microbiome ecology research.

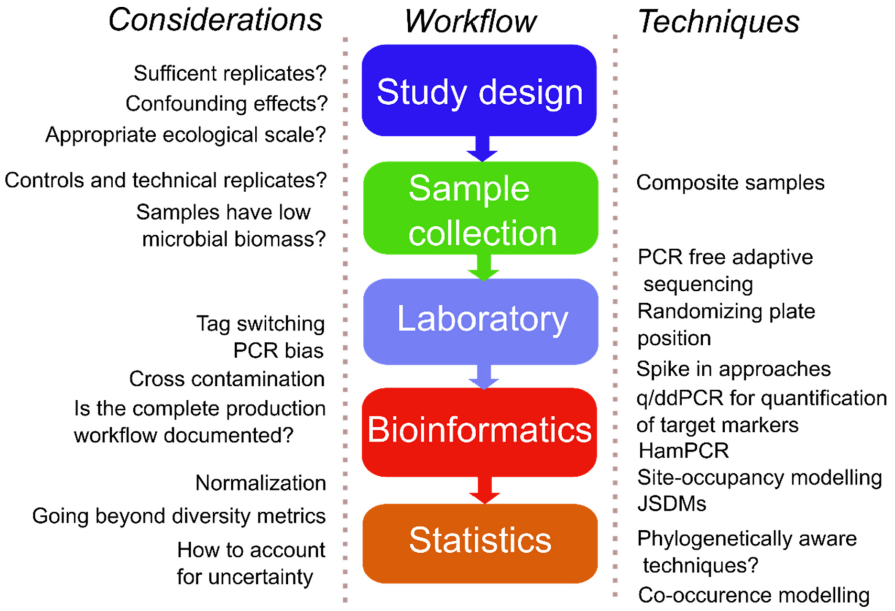

FIGURE 1 Overview of some of the considerations and techniques that can be employed across a general DNA-based microbiome research workflow. JSDMs, Joint species distribution models; q/dd PCR, quantitative/droplet digital PCR. [Colour figure can be viewed at wileyonlinelibrary.com]

5 | CURRENT GAPS AND FUTURE DIRECTIONS IN MICROBIOME ECOLOGY RESEARCH

Given the above considerations, how do we move the field forward? Here we outline some applicable research directions that will generate significant impact, by helping to close some of the most pressing knowledge gaps in the near future.

1. **Obtaining a better understanding of the ecological, evolutionary and mechanistic drivers of microbiome community assembly**. A key research gap in microbiome ecology is the need for a comprehensive understanding of the drivers of community assembly. While significant progress has been made regarding microbiomes associated with humans and model organisms (e.g. Drosophila or Arabidopsis), further study on non-model organisms and natural environments such as soil is required. Included within this goal is the investigation of phylosymbiosis – the topological congruence between host phylogenetic distance and the compositional similarity patterns of their associated microbiota (Brown et al., 2023); a pattern that can arise from both ecological and evolutionary processes. Mechanistic studies elucidating the specific processes that govern microbial transmission, colonization, competition and succession will help explain the presence or absence of phylosymbiosis signals across host species or populations, and are more broadly essential for a deeper understanding of microbiome assembly (Coyte et al., 2021). For both free-living and host-associated microbiomes, the debate on the relative role of stochastic (e.g. dispersal) and deterministic processes (e.g. environment) in shaping community assembly across habitats remains a fundamental research gap (Nemergut et al., 2013; Zhou & Ning, 2017). The model-based methods we discuss above provide a useful starting point to quantifying the importance of each community assembly mechanism (Ovaskainen et al., 2017).
2. **Linking microbiome composition and host genetics.** To gain insights into the ability of hosts to select specific microbes that may benefit their health or reproduction, future research should aim to link host-associated microbiome composition and diversity with genes and genomic regions of hosts. This approach will enable the identification of host genetic factors, such as immune genes, that are key players in shaping the host microbiome and ultimately the resulting host phenotypes. By integrating both host genome and microbiome data, researchers can make progress at uncovering host–microbiome interactions (Sutherland et al., 2022; West et al., 2023). Incorporation of long-read sequencing and hybrid assembly approaches which utilize both short and long reads now offer exciting opportunities for advancing this research area.
3. **Linking microbiomes to host traits and phenotypes**. Understanding the connection between host-associated microbiomes and host phenotypes is a related and important research avenue. Investigating the influence of the microbiome on host development, behaviour, metabolism and life-history traits can therefore provide valuable insights (Bestion et al., 2017; Wood et al., 2022). For example, some nematode species are known to demonstrate extreme phenotypic plasticity in response to environmental cues (e.g. chemical or bacterial stimuli [Hauquier et al., 2017; Sommer et al., 2017]), and one open question is whether microbial taxa play an integral role in initiating such host developmental switch genes (which are themselves under epigenetic control in the case of *Pristionchus* spp. fig nematodes). By integrating microbiota data with detailed trait measurements of diverse hosts, researchers will be able to identify host phenotypes associated with certain microbiome compositions and start to unravel the underlying mechanisms of how certain microbes can influence the phenotypes of hosts, and vice versa.
4. **Exploring microbial functions within microbiomes**. While microbial community composition and diversity have been extensively studied in microbiome ecology, there is a significant need to explore the functional attributes of whole communities, localized populations, and individual microorganisms (genes and pangenomes) within a microbiome. Investigating microbial functions, such as metabolic pathways and molecular interactions between members of the microbiome and with the host, can provide important insights into the contributions of specific microbial taxa/consortia and their functional roles in host and ecosystem health (Béchade et al., 2023; e.g. Hicks et al., 2018; Karmacharya et al., 2019). Furthermore, isolation and culturing of microbial strains can provide complementary information not otherwise accessible through community omics alone (e.g. physiological profiling of microbial growth rates and chemical/antibiotic sensitivity), while also paving the way for future experimental work using such microbial isolates.
5. **Disentangle the role of the non-bacterial components of the microbiome**. Although bacteria have been the primary focus of microbiome ecology research due to their overwhelming abundance, other components such as viruses, fungi or protists can play crucial roles in host–microbe interactions and ecosystem functioning (Jervis et al., 2021; Raghwani et al., 2023). Future investigations that include these non-bacterial components will allow us to more comprehensively understand the dynamics of the microbiome as a whole community, its interaction with hosts, and its role in the ecosystem.
6. **Elucidate the role of host-associated microbiomes in wildlife disease and conservation ecology**. Understanding the role of microbiomes in biological conservation, such as wildlife disease susceptibility and resistance, is an emerging and timely field of research within microbiome ecology. Investigating the interactions between host genetics, environmental factors and microbial communities can shed light on disease dynamics and the mechanisms through which microbiomes modulate host immune responses in wildlife populations (Bozzi et al., 2021; Gao et al., 2021; Jervis et al., 2021). Likewise, studies linking environmental microbiomes with land use, habitat fragmentation and climate change can provide important information on how to address ecosystem challenges in a changing world.

7. **Disentangling diet-microbiome associations**. Diet is a well-known driver of microbiome composition in animal hosts, but the mechanisms by which components of the diet promote certain microbial taxa and ultimately influence host health remain unclear. Therefore, the complexity of diet-microbiome associations requires further investigation (Kartzinel et al., 2019). Future research should aim to unravel the specific dietary components that shape microbial communities, how diet diversity is related to microbial diversity, and the mechanisms through which these interactions ultimately influence host health. Longitudinal studies and controlled dietary interventions can provide valuable insights into the specifics of diet-microbiome relationships (Couch et al., 2021).
8. **Unravelling microbial interactions within the microbiome.** Elucidating the nature of microbial interactions is crucial to understand the dynamics of microbial ecosystems. Patterns of co-occurrence are often used to evaluate microbial interactions (e.g. competition), yet doing so can be problematic (Blanchet et al., 2020). Interactions are highly scale-dependent, which poses unique challenges for microbial communities with fine spatial structuring (Goberna & Verdú, 2022; Peng et al., 2023). Future experimental and observational studies at relevant scales, with large numbers of samples across time and robust measures of abundance, will be able to better quantify microbial interactions (Blanchet et al., 2020). Statistical advances utilizing generalized Lotka–Volterra models across time-series (Stein et al., 2013), or employing conditional probabilities to more directly capture how taxa relate to each other will also help infer interactions (Blanchet et al., 2020).
9. **Determine microbial strain diversity and evolution within hosts**. Microbial communities are often highly heterogeneous and different strains of a single microbial species can exhibit significant genetic and functional variability (Anderson & Bisanz, 2023; Goyal et al., 2022). For example, a shift in *Escherichia coli* clones was documented in the gut microbiome of ageing mice, and these were characterized by an increase in bacterial mutations targeting stress-response genes (Barreto et al., 2020). Investigating microbial genetic diversity and the role of horizontal gene transfer are therefore crucial to better understand the adaptive processes and functional implications within a microbiome (Barreto & Gordo, 2023). Integrating high-resolution genomic techniques with longitudinal and repeated sampling schemes can capture key patterns in temporal variation among microbial communities and significantly improve our understanding of how microbes evolve within individual hosts or specific environments.
10. **The role of connectedness and dispersal in shaping environmental microbiomes.** Connectivity across habitats and microbial dispersal is increasingly recognized as important for free-living microbiomes in particular (Clarke et al., 2020; Dubey et al., 2021; Legeay & Hijri, 2023; Liu et al., 2023; Sessitsch et al., 2023). Dispersal limitation and level of connectivity can strongly shape microbial diversity. For example, experimental approaches have found that highly connected habitats (at a cellular level) can promote soil microbial diversity compared to less connected habitats (Dubey et al., 2021). The appropriate scale to measure connectivity can vary substantially; soil microbiomes can differ dramatically in the top 15 cm of substrate (Legeay & Hijri, 2023) or even within a single soil particle (Nemergut et al., 2013), yet in oceans, microbiomes can be highly connected from the surface to 1000 m in depth (Liu et al., 2023). However, the degree to which microbiomes are connected, and the mechanics of dispersal is unknown for most systems. Increased sampling across these heterogenous habitats coupled with increased genetic resolution (e.g. whole-genome shotgun sequencing [Zhou et al., 2022]) promises to garner new insights into how microbial populations are connected, move and evolve in time and space.

6 | CONCLUSIONS

We are currently at a major turning point in how we can derive valuable insights of the ecological processes shaping non-model and environmental microbiomes in the wild. The adoption of robust laboratory and bioinformatic techniques, together with sophisticated statistical approaches, enhance our ability to gain deeper insights into the factors influencing microbiome variation and the intricate relationships among microbial taxa. However, all of these techniques rely on robust study designs and appropriate sampling scales to address specific research questions. Ultimately, this perspective piece serves as a broad outline of some discussion points that ought to be considered in the field of microbiome ecology. Hopefully, with robust practices, we can, in turn, start to untangle the complex processes acting on these incredibly important but understudied communities.

AUTHOR CONTRIBUTIONS

NFJ and EV conceived the project. All authors contributed to writing the first draft of the manuscript and subsequent editing.

ACKNOWLEDGEMENTS

We thank Jacob A. Russell, Eric Coissac and Kayla King for their useful discussions and edits to the text. EV was funded by the Swedish Research Council (2020-00259).

CONFLICT OF INTEREST STATEMENT

None.

DATA AVAILABILITY STATEMENT

Not applicable.

ORCID

Nicholas M. Fountain-Jones https://orcid.org/0000-0001-9248-8493
Tatiana Giraud https://orcid.org/0000-0002-2685-6478
Lucie Zinger https://orcid.org/0000-0002-3400-5825
Holly Bik https://orcid.org/0000-0002-4356-3837
Elin Videvall https://orcid.org/0000-0002-9998-3689

REFERENCES

Anderson, B. D., & Bisanz, J. E. (2023). Challenges and opportunities of strain diversity in gut microbiome research. *Frontiers in Microbiology*, *14*, 1117122. https://www.frontiersin.org/articles/10.3389/fmicb.2023.1117122

Barlow, J. T., Bogatyrev, S. R., & Ismagilov, R. F. (2020). A quantitative sequencing framework for absolute abundance measurements of mucosal and lumenal microbial communities. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-16224-6

Barreto, H. C., & Gordo, I. (2023). Intrahost evolution of the gut microbiota. *Nature Reviews Microbiology*, *21*(9), Article 9. https://doi.org/10.1038/s41579-023-00890-6

Barreto, H. C., Sousa, A., & Gordo, I. (2020). The landscape of adaptive evolution of a gut commensal bacteria in aging mice. *Current Biology*, *30*(6), 1102–1109.e5. https://doi.org/10.1016/j.cub.2020.01.037

Béchade, B., Cabuslay, C. S., Hu, Y., Mendonca, C. M., Hassanpour, B., Lin, J. Y., Su, Y., Fiers, V. J., Anandarajan, D., Lu, R., Olson, C. J., Duplais, C., Rosen, G. L., Moreau, C. S., Aristilde, L., Wertz, J. T., & Russell, J. A. (2023). Physiological and evolutionary contexts of a new symbiotic species from the nitrogen-recycling gut community of turtle ants. *The ISME Journal*, *17*(10), 1751–1764. https://doi.org/10.1038/s41396-023-01490-1

Bergner, L. M., Orton, R. J., Benavides, J. A., Becker, D. J., Tello, C., Biek, R., & Streicker, D. G. (2020). Demographic and environmental drivers of metagenomic viral diversity in vampire bats. *Molecular Ecology*, *29*(1), 26–39. https://doi.org/10.1111/mec.15250

Bestion, E., Jacob, S., Zinger, L., Di Gesu, L., Richard, M., White, J., & Cote, J. (2017). Climate warming reduces gut microbiota diversity in a vertebrate ectotherm. *Nature Ecology & Evolution*, *1*(6), 161. https://doi.org/10.1038/s41559-017-0161

Björk, J. R., Hui, F. K. C., O'Hara, R. B., & Montoya, J. M. (2018). Uncovering the drivers of host-associated microbiota with joint species distribution modelling. *Molecular Ecology*, *27*(12), 2714–2724. https://doi.org/10.1111/mec.14718

Blanchet, F. G., Cazelles, K., & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions. *Ecology Letters*, *23*(7), 1050–1063. https://doi.org/10.1111/ele.13525

Bohmann, K., Elbrecht, V., Carøe, C., Bista, I., Leese, F., Bunce, M., Yu, D. W., Seymour, M., Dumbrell, A. J., & Creer, S. (2022). Strategies for sample labelling and library preparation in DNA metabarcoding studies. *Molecular Ecology Resources*, *22*(4), 1231–1246. https://doi.org/10.1111/1755-0998.13512

Boshuizen, H. C., & te Beest, D. E. (2023). Pitfalls in the statistical analysis of microbiome amplicon sequencing data. *Molecular Ecology Resources*, *23*(3), 539–548. https://doi.org/10.1111/1755-0998.13730

Bozzi, D., Rasmussen, J. A., Carøe, C., Sveier, H., Nordøy, K., Gilbert, M. T. P., & Limborg, M. T. (2021). Salmon gut microbiota correlates with disease infection status: Potential for monitoring health in farmed animals. *Animal Microbiome*, *3*(1), 30. https://doi.org/10.1186/s42523-021-00096-2

Brown, B. R. P., Goheen, J. R., Newsome, S. D., Pringle, R. M., Palmer, T. M., Khasoha, L. M., & Kartzinel, T. R. (2023). Host phylogeny and functional traits differentiate gut microbiomes in a diverse natural community of small mammals. *Molecular Ecology*, *32*(9), 2320–2334. https://doi.org/10.1111/mec.16874

Callahan, B. J., Wong, J., Heiner, C., Oh, S., Theriot, C. M., Gulati, A. S., McGill, S. K., & Dougherty, M. K. (2019). High-throughput amplicon sequencing of the full-length 16S rRNA gene with single-nucleotide resolution. *Nucleic Acids Research*, *47*(18), e103. https://doi.org/10.1093/nar/gkz569

Carlsen, T., Aas, A. B., Lindner, D., Vrålstad, T., Schumacher, T., & Kauserud, H. (2012). Don't make a mista(g)ke: Is tag switching an overlooked source of error in amplicon pyrosequencing studies? *Fungal Ecology*, *5*(6), 747–749. https://doi.org/10.1016/j.funeco.2012.06.003

Carøe, C., & Bohmann, K. (2020). Tagsteady: A metabarcoding library preparation protocol to avoid false assignment of sequences to samples. *Molecular Ecology Resources*, *20*(6), 1620–1631. https://doi.org/10.1111/1755-0998.13227

Chen, J., & Zhang, X. (2021). D-MANOVA: Fast distance-based multivariate analysis of variance for large-scale microbiome association studies. *Bioinformatics*, *38*(1), 286–288. https://doi.org/10.1093/bioinformatics/btab498

Clark, N. J., Wells, K., & Lindberg, O. (2018). Unravelling changing interspecific interactions across environmental gradients using Markov random fields. *Ecology*, *99*(6), 1277–1283. https://doi.org/10.1002/ecy.2221

Clarke, L. J., Jones, P. J., Ammitzboll, H., Barmuta, L. A., Breed, M. F., Chariton, A., Charleston, M., Dakwa, V., Dewi, F., Eri, R., Fountain-Jones, N. M., Freeman, J., Kendal, D., McDougal, R., Raes, E. J., Sow, S. L. S., Staples, T., Sutcliffe, B., Vemuri, R., … Flies, E. J. (2020). Mainstreaming microbes across biomes. *BioScience*, *70*(7), 589–596. https://doi.org/10.1093/biosci/biaa057

Couch, C. E., Stagaman, K., Spaan, R. S., Combrink, H. J., Sharpton, T. J., Beechler, B. R., & Jolles, A. E. (2021). Diet and gut microbiome enterotype are associated at the population level in African buffalo. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-22510-8

Coyte, K. Z., Rao, C., Rakoff-Nahoum, S., & Foster, K. R. (2021). Ecological rules for the assembly of microbiome communities. *PLoS Biology*, *19*(2), e3001116. https://doi.org/10.1371/journal.pbio.3001116

Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A., & Callahan, B. J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome*, *6*(1), 226. https://doi.org/10.1186/s40168-018-0605-2

de Goffau, M. C., Lager, S., Salter, S. J., Wagner, J., Kronbichler, A., Charnock-Jones, D. S., Peacock, S. J., Smith, G. C. S., & Parkhill, J. (2018). Recognizing the reagent microbiome. *Nature Microbiology*, *3*(8), Article 8. https://doi.org/10.1038/s41564-018-0202-y

Dubey, M., Hadadi, N., Pelet, S., Carraro, N., Johnson, D. R., & van der Meer, J. R. (2021). Environmental connectivity controls diversity in soil microbial communities. *Communications Biology*, *4*(1), Article 1. https://doi.org/10.1038/s42003-021-02023-2

Edgar, R. C. (2018). *UNCROSS2: Identification of cross-talk in 16S rRNA OTU tables* (p. 400762). bioRxiv. https://doi.org/10.1101/400762

Eisenhofer, R., Minich, J. J., Marotz, C., Cooper, A., Knight, R., & Weyrich, L. S. (2019). Contamination in low microbial biomass microbiome studies: Issues and recommendations. *Trends in Microbiology*, *27*(2), 105–117. https://doi.org/10.1016/j.tim.2018.11.003

Esling, P., Lejzerowicz, F., & Pawlowski, J. (2015). Accurate multiplexing and filtering for high-throughput amplicon-sequencing. *Nucleic Acids Research*, *43*(5), 2513–2524. https://doi.org/10.1093/nar/gkv107

Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., Gielly, L., Lopes, C. M., Boyer, F., Pompanon, F., Rayé, G., & Taberlet, P. (2015). Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources*, *15*(3), 543–556. https://doi.org/10.1111/1755-0998.12338

Fountain-Jones, N., Kozakiewicz, C., Forester, B., Landguth, E., Carver, S., Charleston, M., Gagne, R., Greenwell, B., Kraberger, S., Trumbo, D., Mayer, M., Clark, N., & Machado, G. (2021). MrIML: Multi-response interpretable machine learning to map genomic landscapes. *Molecular Ecology Resources*, *21*, 2766–2781. https://doi.org/10.22541/au.160855820.09604024/v1

Fountain-Jones, N. M., Clark, N. J., Kinsley, A. C., Carstensen, M., Forester, J., Johnson, T. J., Miller, E. A., Moore, S., Wolf, T. M., & Craft, M. E. (2020). Microbial associations and spatial proximity predict north American moose (*Alces alces*) gastrointestinal community composition. *Journal of Animal Ecology*, *89*(3), 817–828. https://doi.org/10.1111/1365-2656.13154

Fountain-Jones, N. M., Khoo, B. S., Rau, A., Berman, J. D., Burton, E. N., & Oliver, J. D. (2023). Positive associations matter: Microbial relationships drive tick microbiome composition. *Molecular Ecology*, *32*(14), 4078–4092. https://doi.org/10.1111/mec.16985

Fukaya, K., Kondo, N. I., Matsuzaki, S. S., & Kadoya, T. (2022). Multispecies site occupancy modelling and study design for spatially replicated environmental DNA metabarcoding. *Methods in Ecology and Evolution*, *13*(1), 183–193. https://doi.org/10.1111/2041-210X.13732

Gao, M., Xiong, C., Gao, C., Tsui, C. K. M., Wang, M.-M., Zhou, X., Zhang, A.-M., & Cai, L. (2021). Disease-induced changes in plant microbiome assembly and functional adaptation. *Microbiome*, *9*(1), 187. https://doi.org/10.1186/s40168-021-01138-2

George, P. B. L., Lallias, D., Creer, S., Seaton, F. M., Kenny, J. G., Eccles, R. M., Griffiths, R. I., Lebron, I., Emmett, B. A., Robinson, D. A., & Jones, D. L. (2019). Divergent national-scale trends of microbial and animal biodiversity revealed across diverse temperate soil ecosystems. *Nature Communications*, *10*(1), 1107. https://doi.org/10.1038/s41467-019-09031-1

Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V., & Egozcue, J. J. (2017). Microbiome datasets are compositional: And this is not optional. *Frontiers in Microbiology*, *8*, 2224. https://www.frontiersin.org/articles/10.3389/fmicb.2017.02224

Goberna, M., & Verdú, M. (2022). Cautionary notes on the use of co-occurrence networks in soil ecology. *Soil Biology and Biochemistry*, *166*, 108534. https://doi.org/10.1016/j.soilbio.2021.108534

Goyal, A., Bittleston, L. S., Leventhal, G. E., Lu, L., & Cordero, O. X. (2022). Interactions between strains govern the eco-evolutionary dynamics of microbial communities. *eLife*, *11*, e74987. https://doi.org/10.7554/eLife.74987

Grantham, N. S., Guan, Y., Reich, B. J., Borer, E. T., & Gross, K. (2020). MIMIX: A Bayesian mixed-effects model for microbiome data from designed experiments. *Journal of the American Statistical Association*, *115*(530), 599–609. https://doi.org/10.1080/01621459.2019.1626242

Greenacre, M., Martínez-Álvaro, M., & Blasco, A. (2021). Compositional data analysis of microbiome and any-omics datasets: A validation of the additive Logratio transformation. *Frontiers in Microbiology*, *12*, 727398. https://www.frontiersin.org/articles/10.3389/fmicb.2021.727398

Grosser, S., Sauer, J., Paijmans, A. J., Caspers, B. A., Forcada, J., Wolf, J. B. W., & Hoffman, J. I. (2019). Fur seal microbiota are shaped by the social and physical environment, show mother–offspring similarities and are associated with host genetic quality. *Molecular Ecology*, *28*(9), 2406–2422. https://doi.org/10.1111/mec.15070

Hakimzadeh, A., Abdala Asbun, A., Albanese, D., Bernard, M., Buchner, D., Callahan, B., Caporaso, J. G., Curd, E., Djemiel, C., Brandström Durling, M., Elbrecht, V., Gold, Z., Gweon, H. S., Hajibabaei, M., Hildebrand, F., Mikryukov, V., Normandeau, E., Özkurt, E., Palmer, J., … Anslan, S. (2023). A pile of pipelines: An overview of the bioinformatics software for metabarcoding data analyses. *Molecular Ecology Resources*, 1–17. https://doi.org/10.1111/1755-0998.13847

Harrison, J. G., John Calder, W., Shuman, B., & Alex Buerkle, C. (2020). The quest for absolute abundance: The use of internal standards for DNA-based community ecology. *Molecular Ecology Resources*, *21*, 30–43. https://doi.org/10.1111/1755-0998.13247

Hauquier, F., Leliaert, F., Rigaux, A., Derycke, S., & Vanreusel, A. (2017). Distinct genetic differentiation and species diversification within two marine nematodes with different habitat preference in Antarctic sediments. *BMC Evolutionary Biology*, *17*(1), 120. https://doi.org/10.1186/s12862-017-0968-1

Hicks, A. L., Lee, K. J., Couto-Rodriguez, M., Patel, J., Sinha, R., Guo, C., Olson, S. H., Seimon, A., Seimon, T. A., Ondzie, A. U., Karesh, W. B., Reed, P., Cameron, K. N., Lipkin, W. I., & Williams, B. L. (2018). Gut microbiomes of wild great apes fluctuate seasonally in response to diet. *Nature Communications*, *9*, 1786. https://doi.org/10.1038/s41467-018-04204-w

Jervis, P., Pintanel, P., Hopkins, K., Wierzbicki, C., Shelton, J. M. G., Skelly, E., Rosa, G. M., Almeida-Reinoso, D., Eugenia-Ordoñez, M., Ron, S., Harrison, X., Merino-Viteri, A., & Fisher, M. C. (2021). Post-epizootic microbiome associations across communities of neotropical amphibians. *Molecular Ecology*, *30*(5), 1322–1335. https://doi.org/10.1111/mec.15789

Ji, B. W., Sheth, R. U., Dixit, P. D., Huang, Y., Kaufman, A., Wang, H. H., & Vitkup, D. (2019). Quantifying spatiotemporal variability and noise in absolute microbiota abundances using replicate sampling. *Nature Methods*, *16*(8), 731–736. https://doi.org/10.1038/s41592-019-0467-y

Johnson, A. J., Zheng, J. J., Kang, J. W., Saboe, A., Knights, D., & Zivkovic, A. M. (2020). A guide to diet-microbiome study design. *Frontiers in Nutrition*, *7*, 79. https://www.frontiersin.org/articles/10.3389/fnut.2020.00079

Karmacharya, D., Manandhar, P., Manandhar, S., Sherchan, A. M., Sharma, A. N., Joshi, J., Bista, M., Bajracharya, S., Awasthi, N. P., Sharma, N., Llewellyn, B., Waits, L. P., Thapa, K., Kelly, M. J., Vuyisich, M., Starkenburg, S. R., Hero, J.-M., Hughes, J., Wultsch, C., … Sinha, A. K. (2019). Gut microbiota and their putative metabolic functions in fragmented Bengal tiger population of Nepal. *PLoS One*, *14*(8), e0221868. https://doi.org/10.1371/journal.pone.0221868

Kartzinel, T. R., Hsing, J. C., Musili, P. M., Brown, B. R. P., & Pringle, R. M. (2019). Covariation of diet and gut microbiome in African megafauna. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(47), 23588–23593. https://doi.org/10.1073/pnas.1905666116

Kim, D., Hofstaedter, C. E., Zhao, C., Mattei, L., Tanes, C., Clarke, E., Lauder, A., Sherrill-Mix, S., Chehoud, C., Kelsen, J., Conrad, M., Collman, R. G., Baldassano, R., Bushman, F. D., & Bittinger, K. (2017). Optimizing methods and dodging pitfalls in microbiome research. *Microbiome*, *5*(1), 52. https://doi.org/10.1186/s40168-017-0267-5

Kohl, K. D. (2017). An introductory "how-to" guide for incorporating microbiome research into integrative and comparative biology. *Integrative and Comparative Biology*, *57*(4), 674–681. https://doi.org/10.1093/icb/icx013

Legeay, J., & Hijri, M. (2023). A comprehensive insight of current and future challenges in large-scale soil microbiome analyses. *Microbial Ecology*, *86*(1), 75–85. https://doi.org/10.1007/s00248-022-02060-2

Liu, S., Hu, R., Strong, P. J., Saleem, M., Zhou, Z., Luo, Z., Wu, Y., He, Z., & Wang, C. (2023). Vertical connectivity of microbiome and metabolome reveals depth-dependent variations across a deep cold-seep water column. *Environmental Research*, *239*, 117310. https://doi.org/10.1016/j.envres.2023.117310

Louca, S., Doebeli, M., & Parfrey, L. W. (2018). Correcting for 16S rRNA gene copy numbers in microbiome surveys remains an unsolved problem. *Microbiome*, *6*(1), 41. https://doi.org/10.1186/s40168-018-0420-9

Lundberg, D. S., Pramoj Na Ayutthaya, P., Strauß, A., Shirsekar, G., Lo, W.-S., Lahaye, T., & Weigel, D. (2021). Host-associated microbe PCR (hamPCR) enables convenient measurement of both microbial load and community composition. *eLife*, *10*, e66186. https://doi.org/10.7554/eLife.66186

McClenaghan, B., Compson, Z. G., & Hajibabaei, M. (2020). Validating metabarcoding-based biodiversity assessments with multi-species occupancy models: A case study using coastal marine eDNA. *PLoS One*, *15*(3), e0224119. https://doi.org/10.1371/journal.pone.0224119

Minich, J. J., Sanders, J. G., Amir, A., Humphrey, G., Gilbert, J. A., & Knight, R. (2019). Quantifying and understanding well-to-well contamination in microbiome research. *mSystems*, *4*, e00186-19. https://doi.org/10.1128/msystems.00186-19

Morton, J. T., Marotz, C., Washburne, A., Silverman, J., Zaramela, L. S., Edlund, A., Zengler, K., & Knight, R. (2019). Establishing microbial composition measurement standards with reference frames. *Nature Communications*, *10*(1), 2719. https://doi.org/10.1038/s41467-019-10656-5

Motta, E. V. S., Raymann, K., & Moran, N. A. (2018). Glyphosate perturbs the gut microbiota of honey bees. *Proceedings of the National Academy of Sciences*, *115*(41), 10305–10310. https://doi.org/10.1073/pnas.1803880115

Nagler, M., Podmirseg, S. M., Ascher-Jenull, J., Sint, D., & Traugott, M. (2022). Why eDNA fractions need consideration in biomonitoring. *Molecular Ecology Resources*, *22*(7), 2458–2470. https://doi.org/10.1111/1755-0998.13658

Nascimento, F. J. A., Lallias, D., Bik, H. M., & Creer, S. (2018). Sample size effects on the assessment of eukaryotic diversity and community structure in aquatic sediments using high-throughput sequencing. *Scientific Reports*, *8*(1), 11737. https://doi.org/10.1038/s41598-018-30179-1

Nearing, J. T., Douglas, G. M., Hayes, M. G., MacDonald, J., Desai, D. K., Allward, N., Jones, C. M. A., Wright, R. J., Dhanani, A. S., Comeau, A. M., & Langille, M. G. I. (2022). Microbiome differential abundance methods produce different results across 38 datasets. *Nature Communications*, *13*, 342. https://doi.org/10.1038/s41467-022-28034-z

Nemergut, D. R., Schmidt, S. K., Fukami, T., O'Neill, S. P., Bilinski, T. M., Stanish, L. F., Knelman, J. E., Darcy, J. L., Lynch, R. C., Wickey, P., & Ferrenberg, S. (2013). Patterns and processes of microbial community assembly. *Microbiology and Molecular Biology Reviews: MMBR*, *77*(3), 342–356. https://doi.org/10.1128/MMBR.00051-12

Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., Roslin, T., & Abrego, N. (2017). How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecology Letters*, *20*(5), 561–576. https://doi.org/10.1111/ele.12757

Peng, L., Hoban, J., Joffe, J., Smith, A. H., Carpenter, M., Marcelis, T., Patel, V., Lynn-Bell, N., Oliver, K. M., & Russell, J. A. (2023). Cryptic community structure and metabolic interactions among the heritable facultative symbionts of the pea aphid. *Journal of Evolutionary Biology*, 1–19. https://doi.org/10.1111/jeb.14216

Pichler, M., & Hartig, F. (2021). A new joint species distribution model for faster and more accurate inference of species associations from big community data. *Methods in Ecology and Evolution*, *12*(11), 2159–2173. https://doi.org/10.1111/2041-210X.13687

Pollock, L. J., Tingley, R., Morris, W. K., Golding, N., O'Hara, R. B., Parris, K. M., Vesk, P. A., & McCarthy, M. A. (2014). Understanding co-occurrence by modelling species simultaneously with a joint species distribution model (JSDM). *Methods in Ecology and Evolution*, *5*(5), 397–406. https://doi.org/10.1111/2041-210X.12180

Powell-Romero, F., Fountain-Jones, N. M., Norberg, A., & Clark, N. J. (2023). Improving the predictability and interpretability of co-occurrence modelling through feature-based joint species distribution ensembles. *Methods in Ecology and Evolution*, *14*(1), 146–161. https://doi.org/10.1111/2041-210X.13915

Raghwani, J., Faust, C. L., François, S., Nguyen, D., Marsh, K., Raulo, A., Hill, S. C., Parag, K. V., Simmonds, P., Knowles, S. C. L., & Pybus, O. G. (2023). Seasonal dynamics of the wild rodent faecal virome. *Molecular Ecology*, *32*, 4763–4776. https://doi.org/10.1111/mec.16778

Rubin, B. E. R., Kautz, S., Wray, B. D., & Moreau, C. S. (2019). Dietary specialization in mutualistic acacia-ants affects relative abundance but not identity of host-associated bacteria. *Molecular Ecology*, *28*(4), 900–916. https://doi.org/10.1111/mec.14834

Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., Turner, P., Parkhill, J., Loman, N. J., & Walker, A. W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology*, *12*, 87. https://doi.org/10.1186/s12915-014-0087-z

Santos-Júnior, C. D., Logares, R., & Henrique-Silva, F. (2022). Microbial population genomes from the Amazon River reveal possible modulation of the organic matter degradation process in tropical freshwaters. *Molecular Ecology*, *31*(1), 206–219. https://doi.org/10.1111/mec.16222

Schnell, I. B., Bohmann, K., & Gilbert, M. T. P. (2015). Tag jumps illuminated—Reducing sequence-to-sample misidentifications in metabarcoding studies. *Molecular Ecology Resources*, *15*(6), 1289–1303. https://doi.org/10.1111/1755-0998.12402

Sessitsch, A., Wakelin, S., Schloter, M., Maguin, E., Cernava, T., Champomier-Verges, M.-C., Charles, T. C., Cotter, P. D., Ferrocino, I., Kriaa, A., Lebre, P., Cowan, D., Lange, L., Kiran, S., Markiewicz, L., Meisner, A., Olivares, M., Sarand, I., Schelkle, B., … Kostic, T. (2023). Microbiome interconnectedness throughout environments with major consequences for healthy people and a healthy planet. *Microbiology and Molecular Biology Reviews*, *87*(3), e0021222. https://doi.org/10.1128/mmbr.00212-22

Shade, A. (2017). Diversity is the question, not the answer. *The ISME Journal*, *11*(1), 1–6. https://doi.org/10.1038/ismej.2016.118

Sommer, R. J., Dardiry, M., Lenuzzi, M., Namdeo, S., Renahan, T., Sieriebriennikov, B., & Werner, M. S. (2017). The genetics of phenotypic plasticity in nematode feeding structures. *Open Biology*, *7*(3), 160332. https://doi.org/10.1098/rsob.160332

Stein, R. R., Bucci, V., Toussaint, N. C., Buffie, C. G., Rätsch, G., Pamer, E. G., Sander, C., & Xavier, J. B. (2013). Ecological modeling from time-series inference: Insight into dynamics and stability of intestinal microbiota. *PLoS Computational Biology*, *9*(12), e1003388. https://doi.org/10.1371/journal.pcbi.1003388

Stothart, M. R., Greuel, R. J., Gavriliuc, S., Henry, A., Wilson, A. J., McLoughlin, P. D., & Poissant, J. (2021). Bacterial dispersal and drift drive microbiome diversity patterns within a population of feral hindgut fermenters. *Molecular Ecology*, *30*(2), 555–571. https://doi.org/10.1111/mec.15747

Sutherland, J., Bell, T., Trexler, R. V., Carlson, J. E., & Lasky, J. R. (2022). Host genomic influence on bacterial composition in the switchgrass rhizosphere. *Molecular Ecology*, *31*(14), 3934–3950. https://doi.org/10.1111/mec.16549

Taberlet, P., Bonin, A., Zinger, L., & Coissac, É. (2018). Environmental DNA: For biodiversity research and monitoring. In *Environmental DNA: For biodiversity research and monitoring*. Oxford University Press. https://doi.org/10.1093/oso/9780198767220.001.0001

Tedersoo, L., Bahram, M., Zinger, L., Nilsson, R. H., Kennedy, P. G., Yang, T., Anslan, S., & Mikryukov, V. (2022). Best practices in metabarcoding of fungi: From experimental design to results. *Molecular Ecology*, *31*(10), 2769–2795. https://doi.org/10.1111/mec.16460

Torti, A., Lever, M. A., & Jørgensen, B. B. (2015). Origin, dynamics, and implications of extracellular DNA pools in marine sediments. *Marine Genomics*, *24*(Pt 3), 185–196. https://doi.org/10.1016/j.margen.2015.08.007

Trego, A., Keating, C., Nzeteu, C., Graham, A., O'Flaherty, V., & Ijaz, U. Z. (2022). Beyond basic diversity estimates—Analytical tools for mechanistic interpretations of amplicon sequencing data. *Microorganisms*, *10*(10), Article 10. https://doi.org/10.3390/microorganisms10101961

Vos, M., Wolf, A. B., Jennings, S. J., & Kowalchuk, G. A. (2013). Micro-scale determinants of bacterial diversity in soil. *FEMS Microbiology Reviews*, *37*(6), 936–954. https://doi.org/10.1111/1574-6976.12023

Warton, D. I., Blanchet, F. G., O'Hara, R. B., Ovaskainen, O., Taskinen, S., Walker, S. C., & Hui, F. K. C. (2015). So many variables: Joint modeling in community ecology. *Trends in Ecology & Evolution*, *30*(12), 766–779. https://doi.org/10.1016/j.tree.2015.09.007

Warton, D. I., Wright, S. T., & Wang, Y. (2012). Distance-based multivariate analyses confound location and dispersion effects. *Methods in Ecology and Evolution*, *3*(1), 89–101. https://doi.org/10.1111/j.2041-210X.2011.00127.x

West, A. G., Digby, A., Santure, A. W., Guhlin, J. G., Dearden, P., Kākāpō Recovery Team, Taylor, M. W., & Urban, L. (2023). Capturing species-wide diversity of the gut microbiota and its relationship with genomic variation in the critically endangered kākāpō. *Molecular Ecology*, *32*(15), 4224–4241. https://doi.org/10.1111/mec.16999

Wille, M., Eden, J.-S., Shi, M., Klaassen, M., Hurt, A. C., & Holmes, E. C. (2018). Virus–virus interactions and host ecology are associated with RNA virome structure in wild birds. *Molecular Ecology*, *27*(24), 5263–5278. https://doi.org/10.1111/mec.14918

Willoughby, J. R., Wijayawardena, B. K., Sundaram, M., Swihart, R. K., & DeWoody, J. A. (2016). The importance of including imperfect detection models in eDNA experimental design. *Molecular Ecology Resources*, *16*(4), 837–844. https://doi.org/10.1111/1755-0998.12531

Wood, G., Steinberg, P. D., Campbell, A. H., Vergés, A., Coleman, M. A., & Marzinelli, E. M. (2022). Host genetics, phenotype and geography structure the microbiome of a foundational seaweed. *Molecular Ecology*, *31*(7), 2189–2206. https://doi.org/10.1111/mec.16378

Woodcock, S., van der Gast, C. J., Bell, T., Lunn, M., Curtis, T. P., Head, I. M., & Sloan, W. T. (2007). Neutral assembly of bacterial communities. *FEMS Microbiology Ecology*, *62*(2), 171–180. https://doi.org/10.1111/j.1574-6941.2007.00379.x

Wu, L., Yang, F., Feng, J., Tao, X., Qi, Q., Wang, C., Schuur, E. A. G., Bracho, R., Huang, Y., Cole, J. R., Tiedje, J. M., & Zhou, J. (2022). Permafrost thaw with warming reduces microbial metabolic capacities in subsurface soils. *Molecular Ecology*, *31*(5), 1403–1415. https://doi.org/10.1111/mec.16319

Zapala, M. A., & Schork, N. J. (2012). Statistical properties of multivariate distance matrix regression for high-dimensional data analysis. *Frontiers in Genetics*, *3*, 190. https://doi.org/10.3389/fgene.2012.00190

Zhang, Z., Geng, J., Tang, X., Fan, H., Xu, J., Wen, X., Ma, Z. (Sam), & Shi, P. (2014). Spatial heterogeneity and co-occurrence patterns of human mucosal-associated intestinal microbiota. *The ISME Journal*, *8*(4), Article 4. https://doi.org/10.1038/ismej.2013.185

Zhou, J., & Ning, D. (2017). Stochastic community assembly: Does it matter in microbial ecology? *Microbiology and Molecular Biology Reviews: MMBR*, *81*(4), e00002-17. https://doi.org/10.1128/MMBR.00002-17

Zhou, Y., Liu, M., & Yang, J. (2022). Recovering metagenome-assembled genomes from shotgun metagenomic sequencing data: Methods, applications, challenges, and opportunities. *Microbiological Research*, *260*, 127023. https://doi.org/10.1016/j.micres.2022.127023

Zinger, L., Bonin, A., Alsos, I. G., Bálint, M., Bik, H., Boyer, F., Chariton, A. A., Creer, S., Coissac, E., Deagle, B. E., De Barba, M., Dickie, I. A., Dumbrell, A. J., Ficetola, G. F., Fierer, N., Fumagalli, L., Gilbert, M. T. P., Jarman, S., Jumpponen, A., … Taberlet, P. (2019). DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. *Molecular Ecology*, *28*(8), 1857–1862. https://doi.org/10.1111/mec.15060

How to cite this article: Fountain-Jones, N. M., Giraud, T., Zinger, L., Bik, H., Creer, S., & Videvall, E. (2024). Molecular ecology of microbiomes in the wild: Common pitfalls, methodological advances and future directions. *Molecular Ecology*, *33*, e17223. https://doi.org/10.1111/mec.17223