In harmony? A scoping review of methods to combine multiple 16S amplicon datasets

Zone Li, Yingan Chen, Yage Sun, Clifton McKee, Kristen McArthur, Megan Ulmer Carnes, Tiange Liu, Noel T Mueller, Grier P Page, Lori Rosman, Ekaterina Smirnova, Julie D White, Amii M Kress, Justine W Debelius

ORIGINAL ARTICLE

In harmony? A scoping review of methods to combine multiple 16S amplicon datasets

Zone Li[1]* (0009-0007-8893-114X), Yingan Chen[2]* (0000-0001-8591-8707), Yage Sun[3] (0009-0000-3564-0946), Clifton McKee[1] (0000-0002-6149-0598), Kristen McArthur[1](0000-0001-5601-1583), Megan Ulmer Carnes[4] (0000-0002-7270-132X), Tiange Liu[1,5] (0000-0003-2486-8691) , Noel T. Mueller[2,6] (0000-0002-7412-8352) , Grier P. Page[4] (0000-0003-2582-3786), Lori Rosman[7] (0000-0002-6935-1672), Ekaterina Smirnova[8] (0000-0001-9577-4102), Julie D. White[4] (0000-0002-2235-4817), Amii M. Kress[1,†] (0009-0003-1903-696X), Justine W Debelius[1,†] (0000-0002-8605-3546)

*These authors contributed equally to this work

†These authors contributed equally to this work and are co-corresponding authors

Affiliation(s)

1. Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health; Baltimore, MD, 21205. USA
2. Department of Epidemiology, Colorado School of Public Health, Lifecourse Epidemiology of Adiposity and Diabetes (LEAD) Center, University of Colorado, Anschutz Medical Campus, Aurora, CO, 80045. USA
3. Cutaneous Translational Research Program, Department of Dermatology, Johns Hopkins School of Medicine, Baltimore, MD, 21205. USA
4. GenOmics and Translational Research Center, RTI International Research Atlanta, GA, USA
5. Division of Women's Health, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School; Boston, MA, 02115. USA
6. Department of Pediatrics, Section of Nutrition, University of Colorado, Anschutz Medical Campus, Aurora, CO, 80045. USA
7. Welch Medical Library, Johns Hopkins School of Medicine, Baltimore, MD, 21205. USA
8. Department of Biostatistics, Virginia Commonwealth University, Richmond, VA

Corresponding authors

Justine Debelius (justine.debelius@jhu.edu)

Amii M Kress (akress1@jhu.edu)

Department of Epidemiology

Johns Hopkins Bloomberg School of Public Health

615 N. Wolfe Street

Baltimore, MD 21205. USA

Abstract

Robust evidence on relationships between the human microbiome and health is critical for understanding and improving the human condition. However, there is little information about methodological approaches to combine and analyze multiple microbiome (16S amplicon) datasets. To address this gap, we conducted a scoping review of studies that combined 16S sequencing data from multiple sources to understand the objectives, data sources and selection, bioinformatics, and analyses. References were identified through a systematic search of literature published between January 2011 and April 2026. Our final review included 136 articles. Despite the wide-spread use of the word “meta-analysis,” we found that only two-thirds of studies used a systematic process to select their dataset and 20% applied a statistical meta-analysis. Most studies (79%) combined datasets from multiple disjoint hypervariable regions. The number of hypervariable regions combined was not associated with the bioinformatic methods, but bioinformatics and the number of hypervariable regions influenced analytical resolution. Our results suggest the microbiome community needs to examine the terminology and analytic approaches for combining datasets; that additional work is needed to explore the impact of data source on bias in combined studies; and evaluation of methods used for feature table construction across disjoint regions is needed.

Keywords: meta research; microbiome; meta-analysis; scoping review; 16S sequencing; human microbiome

Introduction

The human microbiome is hypothesized to play a critical role in health, affecting development, metabolism, and the immune system. [1–3] Many of these conclusions are based on small, single center studies, limiting their generalizability. Understanding the role of the microbiome in health and moving forward interventions requires stronger evidence, such as that derived from evidence synthesis studies like systematic reviews and meta-analyses. However, such studies often fail to reproduce the findings of smaller studies, or they find taxonomic associations among such broad clades (i.e. class, phylum) that mechanistic inference is limited.[4–7] Synthesis studies rarely mention the effect of analytical choices of the included studies on the observed results, despite limited interoperability between commonly used taxonomic databases[8–11] and different differential abundance methods producing different results on the same dataset.[12] The inconsistencies noted in synthesis studies may therefore reflect analytical choices rather true biological variation.[8–11]

Re-analysis of previously published primary data (sequences and participant data) using a common bioinformatic pipeline and statistical approach is an ideal approach to evidence synthesis, however, this proposition is complicated by heterogeneity in extant datasets and published studies. Each step in sample collection, processing, amplification, sequencing, and bioinformatics influences the observed microbial community.[12,13] For example, microbiome characterization using the 16S rRNA gene often relies on short amplicons that target a fragment of the full gene.[14,15] Primer pairs

used to amplify different regions have different specificities and taxonomic resolutions; the "optimal" primer pair for a study depends on the research question.[14–17] While standardization is an oft-proposed goal, there are concerns that standardization may result in under-evaluated bias.[13] Thus, re-analysis of primary data necessitates harmonization across multiple non-overlapping fragments of the 16S rRNA gene. There are multiple possible bioinformatic strategies that might be used to address the challenge of harmonization, each with benefits and drawbacks (Table S1). While these approaches are common in microbiome analysis, little work has been done to characterize how researchers are applying these approaches in the context of combining multiple datasets.

To address this gap, we conducted a scoping review of studies that combined multiple primary 16S rRNA datasets and analyzed them together. Our goals were to describe the studies and how they selected and acquired specific datasets, how 16S data were harmonized across datasets, and what statistical methods were used to combine datasets.

Methods

Search Strategy and Screening.

We did not pre-register a study protocol.

Our search focused on three concepts: sequencing of the 16S rRNA gene; the human microbiome; and data combination (Supplemental File 1). References were limited to those published between January 2011 and April 2026. We considered Andersson et al, 2008[18] to be the first human microbiome study with next generation sequencing techniques; our 2011 start year therefore allowed time for studies combining multiple 16S sequencing datasets to be published. References were identified through a systematic search of PubMed, Embase, Scopus, Web of Science, and the Cochrane library, performed on June 9, 2022. A second search of PubMed and Web of Science was performed in April 2026. Embase and Scopus were excluded as the AI supplementation to their search substantially degraded the results.

References were screened through Covidence (Melbourne, Australia) with a two-step process. References were excluded at the title and abstract stage if they were a commentary, perspective, or erratum; did not include primary sequence analysis (e.g. systematic reviews); or represented a single 16S sequencing dataset. Title-abstract screening was performed independently by two reviewers (YC, TL, MC, KM, CM or JWD). In full text screening, eligible articles were peer reviewed studies, written in English, focused on the combination of 16S rRNA sequencing data of microbial communities from multiple sources (labs or published studies), and studies which included human samples. During extraction, articles suspected of being written with generative AI were also excluded.

Full text screening was performed independently by two reviewers (YC, MC, KM, JDW, CM, or JWD). In both title-abstract and full-text screening, conflicts were resolved by a

senior third-party reviewer (JWD, JDW, or MC) through group discussion and mutual consensus.

Data Extraction

We extracted information on the goal of combining 16S data, characteristics of the datasets included, feature table construction, and statistical analyses from each paper. Information was extracted using a structured form in Qualtrics (Qualtrics, Provo, Utah, USA). The full extraction form can be found on Zenodo.[19] Extractors used the term "meta-analysis" as short hand for the pooling of per dataset summary statistics.

In the first round, data extraction was performed by two independent reviewers (ZL, YC, YG, JWD). Reviewers trained on a common set of four articles and discrepancies were discussed. Inter-reviewer correlations were checked regularly, and all four reviewers met to discuss questions about the papers. The agreement between independent extractors in the first round was greater than 82.5%. Conflicts were resolved by a third reviewer who had access to the responses of the previous two reviewers. Compiled results were reviewed by an expert reviewer (JWD) with at least six months of separation from the original extraction. In the second round of extractions, two reviewers (JWD, CM) extracted all articles. Conflicts were resolved by consensus with specific textual documentation for the observation. The results of the data extraction were assembled using the pandas library (v1.2.5) in python 3.8.11.[20,21]

Synthesis of Results

All analyses were conducted in R (v4.2.2) with tidyverse (v2.0.0); plots were made with the ComplexUpset library (v1.3.3); other figures were made using ggplot2 (v3.5.1)[22–26]. Analyses were primarily descriptive. Associations between the number of hypervariable regions assessed, feature table construction method, and taxonomic level for analysis were tested with Fisher's exact tests. An alpha value of <0.05 was considered statistically significant.

Results

Our review focused on articles which combined primary data from multiple 16S rRNA sequencing datasets. Using a systematic search strategy, we identified 1524 references for evaluation, of which 996 were excluded at the title and abstract stage and an additional 400 were excluded at full text screening. We extracted and summarized 136 articles, although not all articles were relevant to all topics analyzed here (Figure 1, Table 1, Table S2).[19] The rate of publication increased over time, with the majority of studies published after 2017 (Figure S1). Studies were designed with several scientific goals, although most focused on a microbiome-outcome (n=102, 75%) or microbiome-exposure (n=31, 23%) association (Table 1). Eighteen studies included validation datasets (range 1-6). Three included only one dataset in their primary analysis also included at least one validation dataset.[27–29]

Dataset Selection and Description

We also abstracted information about how the studies identified datasets to include in their analysis, and how they described these datasets in the text. Of our original 136 studies, we chose to focus on cases where this was most relevant: research focused on a microbiome-exposure or microbiome-outcome relationship analyses of humans only or humans and non-human hosts. This left us with a total of 125 relevant studies. Ninety percent (113/125) of articles had the words “meta anal*” or “meta-anal*” in either the title or abstract. Most studies (n= 81, 65%) reported a systematic approach for dataset identification and inclusion; 50 included a PRISMA diagram (Table S3). While all 125 articles included at least some description of the datasets they included, reporting varied greatly, with 85% of studies (n=106) identifying the hypervariable used in their analysis, but only 24 (19%) describing sample collection.

Dataset Sources

The microbiome community has emphasized the importance of publicly available data to support robust and reproducible research.[30,31] Therefore, we were interested in where included studies found their datasets (Figure 2, Table S4). Data access is not related to study design or research question, so we included all 136 studies. The majority (n=108; 79%) used at least one publicly available data source, such as the European Nucleotide Archive (ENA), or Sequence Read Archive (SRA). There were 62 (46%) studies that used closed data sources, including datasets from the authors of the combined study or requests to the authors of the original dataset. Overall, we found 54 (40%) studies that

used only open data sources, while 36 (26%) studies required the authors to request additional information about a previously published dataset.

Sequencing Platforms and Hypervariable Regions

The technical parameters used to sequence the data were heterogeneous, which likely influences how 16S data were harmonized to produce a common feature table. While factors like sample collection and extraction are known to create biases in observed microbial communities,[15] they should not substantially influence bioinformatic processing. We anticipated sequencing platform and hypervariable regions would influence dataset processing and harmonization.[14] Among the 136 total studies reviewed, 28 (21%) did not specify their sequencing platforms. A total of 53 (39%) included datasets sequenced using the Roche 454 pyrosequencer; 102 (75%) included datasets amplified with Illumina and 20 studies (15%) had datasets amplified with Ion Torrent. Table S5 summarizes the observed combinations of sequencing platforms.

Previous research has shown that amplicons from different hypervariable regions within the same organism are not directly comparable, due to both the underlying biological functions of the 16S rRNA gene and technical biases introduced during DNA amplification, sequencing, and bioinformatic processing.[14] Additionally, the hypervariable region can have a large impact on the observed community, often exceeding the effect of biologically relevant factors.[32,33] Over 80% (n=115/136) of the total included studies reported the hypervariable regions amplified in their datasets

(Figure 3). We found 24 (21%) of studies included data from only one hypervariable region; 41(36%) used datasets which targeted two or three hypervariable regions, and 50 (43%) included datasets from four or more regions. The most common hypervariable region included was V4 which was included in 94 of the studies (82% of the 115 studies with described hypervariable regions). Primers targeting the V3V4 and V1V2 regions were also common, present in 88 and 44 studies, respectively.

Feature Table Construction Methods

Combining 16S sequencing data across multiple, disjoint hypervariable regions is challenging. One major goal in feature table construction is to harmonize the regional sequence identifiers in such a way that the same organism shares a common identifier across disjoint regions. This decision involves a balance between retaining high sequence-based resolution, which provides more precision and allows the use of phylogenetic techniques [34]; or working with data collapsed to a common taxonomic annotation, a choice which may decrease dataset-specific technical biases but may also obfuscate true ecological differences between microbial communities (Table S1).[33] Six out of our total 136 studies, 6 (4%) did not process their data with consistent pipelines, two studies did not include a taxonomic analysis, and the table processing was unclear for nine studies. Of the remaining 119 studies, closed reference operational taxonomic units (OTU) clustering (n=26, 22%), de novo OTU clustering (n=20, 17%), and denoising to amplicon sequence variants (ASVs; n=55, 46%) were the most common methods, although this varied over time (Table S6, Figure S2).

We specifically wanted to understand how the feature table construction method and number of hypervariable regions would influence the taxonomic resolution used in analysis. Table S1 highlights the relationship between these factors. We hypothesized studies that used feature table construction techniques that generated common feature identifiers across multiple disjoint hypervariable regions (e.g. closed reference OTU clustering) would be more likely to analyze with feature-level precision compared to those harmonizing using a technique that does not provide feature-level identifiers across multiple regions (e.g. ASVs, de novo OTUs), as this approach requires the data to be collapsed to a higher taxonomic level, sacrificing precision. We also hypothesized this might depend on the type of analysis being performed and its cross regional validity.

We identified 84 studies which used either closed reference OTU clustering (n=19; common features across multiple regions) or generated region-specific identifiers (n=65; de novo OTUs and ASVs), provided a defined number of hypervariable regions, and conducted alpha diversity (n=71; feature identifier agnostic), beta diversity (n=66; common identifier needed), and/or differential abundance (n=73; common identifier needed) analyses. We did not identify a statistically signification association between the number of hypervariable regions combined and the table construction method (Table S7). Studies which constructed an ASV table or did de novo OTU clustering were less likely to include beta diversity , (Table S6, $p<0.01$). Comparing the number of regions and feature table construction method by analysis, we found analyses conducted with ASVs or de novo OTUs were significantly more likely to conduct multi-region analyses

with collapsed data, but the same was not true for analyzed using closed reference clustering to scaffold the regions (Table 2).

Analyses

We explored the analyses conducted and how multiple datasets were combined statistically. In describing analyses, we used the 125 articles which were focused on humans only or humans and another non-human host and which were interested in microbiome-outcome or microbiome-exposure relationships. The most common analyses were alpha diversity (n=101, 81%), beta diversity (n=102, 82%), differential abundance testing for individual organisms (n=105, 84%), and sample classification (n=62, 50%; Figure 4, Table S8). Studies performed a median of four types of analyses (IQR 3 - 5 analyses) with 98 studies using three or more analytical techniques. For the three most common inference-based analyses (alpha diversity, beta diversity, differential abundance), we looked at the statistical techniques used for combining datasets across our analyses (Table S9). We classified inference approaches into five categories: (i) no acknowledgement of dataset; (ii) parallel analysis without pooling; (iii) adjustment for dataset or pooling; (iv) other analyses, for example, adjustment for technical effects, and (v) analyses not described (Table 3; Figure S3; Table S10-S13). We did not consider an adjustment for batch effects (e.g. ComBat) to be a statistical method to correct dataset, as the efficacy of these methods depend on the metric being characterized. Among the statistical analyses, 38% of alpha diversity, 27% of beta diversity, and 27% of differential abundance studies did not account for the dataset

effect . Interestingly, 9 studies showed ordinations by study and 15 studies performed beta diversity testing by dataset, which may not be appropriate given ordination and measures from commonly used tests are not directly comparable.

According to the strictest definitions, a statistical meta-analysis is defined as the pooling of effect size estimates. [35,36] With this definition in mind, we consider statistical meta analysis only possible for alpha diversity and differential abundance, which produce generalizable point estimates. While 81 of 89 eligible papers (91%) included "meta-analysis" in the title or abstract, only 29 (33%) did a strict statistical meta-analysis (Table S14). A broader definition of "meta-analysis" allows the pooling of data or estimates with some adjustment for datasets; this can be applied to all three of our analyses. [37,38] Of 112 eligible studies, 102 (91%) were identified as meta-analyses, and 72 (64%) performed at least one analysis that adjusted for dataset differences (Table S14).

Predictive sample classification was performed in 63 studies. Unlike inferential statistics, sample classification does not provide an effect estimate that could be comparable through a meta-analytical framework; instead study effects are considered through either cross validation within the training set or independent validation of the classifier on a new dataset. We found the majority (n=34, 54%) of studies which built a classifier trained it on multiple datasets or used cross validation. Nineteen studies (30%) trained without considering dataset; one study contained two analyses which trained classifiers on a single dataset for one outcome and ignored dataset effects in the second;[39] one

study combined datasets only to validate their classifier;[28]five and five studies did not describe how the classifier was trained. However, we found only 14 studies validated their classifier on independent datasets, potentially limiting the generalizability of their results.[40]

Discussion

We performed a scoping review of English language peer-reviewed studies of human microbiome 16S rRNA datasets. While the intention of such studies is often to generate robust evidence relating the microbiome to an outcome or exposure, more than a quarter did not use systematic approach to dataset identification. This has the potential to create unobserved bias, limiting the strength of evidence generated by which can decrease confidence in the findings of these studies. About 40% of studies used open-source data alone, with the remaining studies functioning as either consortia or requiring the authors to request data from the original source. This suggests that while there may be a robust open data ecosystem in microbiome science, more work may be needed. We found denoising to ASVs to be the most common table construction method, followed by closed reference OTU clustering. While the number of hypervariable regions included was not associated with the feature table construction method, studies using ASVs or performing de novo OTU clustering – methods which were unable to provide common feature-level identifiers across hypervariable regions – were more likely to have performed taxonomically collapsed. Finally, we found that analyses often failed to account for the influence of datasets in their analyses.

One aim of our work was to summarize the re-use of publicly available data in studies that combined microbiome datasets. In the past decade, there have been increased calls for data sharing both among microbiome scientists;[30,31] from funding agencies;[41,42] and across scientific disciplines.[43–45] Proponents of publicly available microbiome data argue that it improves reproducibility, allows innovation, amortizes resources, facilitates better training, and can improve research equity.[31,46–48] However, there are concerns about data quality, participant privacy, data sovereignty, “scooping”, and feasibility.[48–50] Our scoping review provides insight into data re-use patterns across multiple human microbiome studies, with the caveat that human samples may come with more privacy concerns than environmental samples. While almost 80% of studies were able to use some open-source data, only 54 (40%) studies used only publicly available data. Thirty-six studies had to contact the original authors for additional data. However, it is possible that the studies which focused on open data sources also requested data from authors but were unable to include it due to a lack of response or time restrictions.[31] Both overall and in the context of absent or non-systematic study inclusion criteria, the process of requesting data from authors may also create an unobserved bias, since relevant datasets may have been lost because they could not be accessed. Our results suggest work is needed to explore the impact of using public data only versus requesting datasets from the original authors on both observed associations and the generalizability of findings.

One of the surprising themes in both our data extraction and conversations within our interdisciplinary team surrounded the meaning of the word “meta-analysis". There were

two major points of disagreement in this discussion: first, whether "meta-analysis" was restricted to a statistical combination of summary statistics or could include the re-analysis of primary data; and second, whether meta-analysis required a systematic approach to dataset selection.

Many researchers in fields like genetics, epidemiology, and medicine, have a strict concept of “meta-analysis.” The narrowest definitions define a “meta-analysis” as a statistical aggregation of dataset-specific estimates.[35] Some expanded definitions to allow for other statistical methods or results not presented in the original study, for example, within the individual participant data framework.[37,38] Geneticists, in contrast, disambiguate analyses of primary data as “mega-analyses”.[36] While not strictly necessary under the restricted statistical framework definition, in many fields the idea of a meta-analysis is intimately tied to a rigorous, systematic approach to dataset inclusion, description and analysis. The Cochrane guidelines[35,37] and others[51,52] caution researchers against conducting meta-analyses without a systematic approach,[35,37] which is designed to facilitate a more comprehensive understanding of the review topic, limit bias, enhance transparency, and improve replicability of the review process.[53] For many readers, the term “meta-analysis” may presupposes this process, with implications for the thoroughness and comprehensiveness of the evidence included and conclusions drawn.

In contrast, we found four prevailing definitions of “meta-analysis” in the microbiome literature. These ran the gamut from a re-analysis of primary datasets generated under a common protocol; to a systematic search, extraction of estimates, and statistical aggregation of results (Table 4). By the strictest definition, statistical meta-analyses were also rare, although a more relaxed was seen in almost two thirds or papers. Our results suggest the microbiome field needs to clarify the definition of "meta-analysis" and possibly align it more closely to what is seen in closely related disciplines. While we hesitate to suggest terms, microbiome literature might benefit from disambiguation between aggregate statistics and re-analysis of primary data. Choosing alternative terms for studies interested in combing data for other objectives like replication or contextualization might also benefit from alternative language. The development of a microbiome-specific PRISMA guideline, like the STROBE-metagenomics checklist,[54] STORMS checklist,[55] the PRISMA EcoEvo extension,[56] or PRISMA individual participant data extension[56] could provide a framework for robust evidence synthesis. These solutions will require community buy-in from funders, researchers, editors, and journals. While we, as a small group of researchers, can propose terms, a cross-community working group would be required to finalize guidelines. In the interim, we recommend readers carefully examine the study methods to understand how studies were gathered and to determine if a study presented as a meta-analysis is, in fact, systematic, whether a systematic approach matters for the application in question, and to consider how this may affect the quality of evidence relative to their own research question.

Our work has some limitations. The goal of this analysis was descriptive: we wanted to understand the ways in which microbiome datasets were combined. We did not evaluate the efficacy of the methods for minimizing study effects in data combination. As such, considerations like primer trimming, qualify filtering, algorithm-specific processing parameters, and normalization techniques like rarefaction, relative abundance scaling, and compositional transforms are not discussed. While a recently published evaluation of techniques to combine scaffold datasets may be of interest,[57] work based on both real data and simulations is needed to evaluate the best way to combine multiple 16S datasets into a single study. We acknowledge the potential for misclassification in our extraction. Microbiome research is a complex, interdisciplinary field with multiple analyses combinations and approaches.[12,58] There are linguistic differences around how techniques are described, for example phylotypes and OTUs. This was exacerbated in our analysis, where studies which included a primary dataset as well as a combination might conduct certain analyses on that primary dataset but not the combined data.[28,59] Additionally, our team found several methods sections to be sparse and occasionally disjointed, making it difficult to make a clear determination of the methods and analyses performed. Finally, our results are necessarily limited due to the scope of our work. We focused on the microbiome characterized using the 16S rRNA genes in samples from humans. Some of the issues we focused on, like the way multiple short read amplicons are combined, are specific to amplicon sequencing and will not generalize to other methods of microbiome characterization, like metagenomics. Future work is needed to evaluate language and methodology in metagenomics as well as the adoption of methods to combine 16S and metagenomic data.[60] Our results may also be less

relevant to non-human work, where there may be different concerns around issues like data sharing, metadata harmonization, or taxonomic database coverage of key microbes. Finally, our observations are limited by the scope of our work. Our structured approach ensured that we collected substantial information but may have inadvertently excluded uncommon protocols.

In conclusion, we conducted a scoping review of combined studies of the human microbiome. Our work suggests a need for a framework for high-quality evidence synthesis, which is essential for moving the microbiome field forward. We find that while open-source data are widely used, additional efforts could be made to further adoption. Our work re-enforces previous understanding of feature table construction methods to harmonize datasets with disjoint hypervariable regions and primers but begs additional work into their strengths and limitations.

DISCLOSURE STATEMENT

The authors declare that there are no conflicts of interest.

Research reported in this publication was supported by the Environmental influences on Child Health Outcomes (ECHO) Program, Office of the Director, National Institutes of Health, under Award Number U24OD023382 (Data Analysis Center). All analyses for scientific publication were performed by the study statisticians, independently of the sponsor. The study sponsor did not review or approve the manuscript for submission to the journal.

DATA AVAILABILITY STATEMENT

The data extraction form, codebook, raw, and cleaned data are available from Zenodo (https://zenodo.org/records/21072900; doi: 10.5281/zenodo.14767164). Analytical code can be found on GitHub (https://github.com/jwdebelius/MicroScoping; doi: 10.5281/zenodo.18202751).

ACKNOWLEDGEMENTS

The authors wish to thank Rob Wright at the Welch Medical library for his help in clarifying search terms. Thanks to Luisa Hugerth for her helpful discussion and meticulous copy editing. We also appreciate useful discussion of what a “meta-analysis” means in microbiome research with the attendees of the 2022 Lake Arrowhead Metagenomics meeting.

Author contributions

JWD, AK, NM, JDW, MC, KS, and GP designed the analysis. LR performed the systematic search with input from JWD and AK. YC, MC, KM, JDW, TL, CM, and JWD reviewed articles. YC, YS, ZL, CM, and JWD performed data extraction. JWD curated extracted data. YC, YS, ZL and JWD analyzed the results. ZL and JWD curated analytical code with help from YC and YS. YC, ZL, and JWD wrote the initial draft with help from YS and KM; AK, KM, and JDW provided critical edits. ZL and AK herded all the cats.

All authors reviewed and approved the final manuscript.

Tables

Table 1. Characteristics of studies included in this scoping review

	Total N (%)
Number of studies included	136
Analytical purpose of the combined analysis[a]	
Population synthesis	129 (95)
Describing a core microbiome	7 (5)
Replication and validation of previous findings	8 (6)
Other[b]	8 (6)
Scientific goal of the combined analysis[a]	
Describe the relationship between the microbiome and an outcome	102 (75)
Describe the relationship between the microbiome and an exposure	31 (23)
Studying community Assembly or Dynamics	9 (7)
To characterize a unique microbiome	8 (6)
Other[c]	8 (6)
Environment	
Human only	129 (95)
Human and other animals	3 (2)
Pan environmental surveys	4 (3%)
Human body site(s) analyzed[a]	
Gut	99 (73)
Oral	16 (12)
Airway	14 (10)
Urogenital	14 (10)
Skin	8 (6)
Other sites, including biopsies and gastric tissue	9 (7)
Pan environmental	4 (4)
Number of datasets included	
≤ 5	39 (29)
6-10	38 (28)
11-15	27 (20)
> 15	27 (20)
Not reported	5 (4)
Number of samples included	
< 500	24 (18)
500-1500	40 (29)
1501-2500	28 (21)
>2500	29 (21)
Not Reported	15 (11)
Included secondary dataset(s) for validation	18 (13)
Clear systematic approach for dataset gathering	83 (61)

[a]Multiple responses possible; number of papers with this response
[b]Other purposes included pan environmental comparisons; methodological assessments; cross species comparison; and microbiome-metabolite associations
[c]Other goals included characterization of a specific clade in a population; pooling randomized control trials; environmental comparison; and methodological assessment

Table 2. The relationship between the number of hypervariable regions combined, table construction method, and taxonomic collapse

	All			ASVs or de novo OTUs			Closed Reference OTUs		
	1 N(%)	2 or more N(%)	p	1 N(%)	2 or more N(%)	p	1 N(%)	2 or more N(%)	p
Alpha Diversity			0.03			0.05			1.00
feature	15 (21)	41 (57)		11 (20)	29 (54)		4 (22)	12 (67)	
collapse	0 (0)	16 (22)		0 (0)	14 (26)		0 (0)	2 (11)	
Beta Diversity			<0.01			<0.01			0.53
feature	16 (24)	30 (45)		12 (25)	19 (40)		4 (21)	11 (58)	
collapse	0 (0)	21 (31)		0 (0)	17 (35)		0 (0)	4 (21)	
Differential Abundance			<0.01			0.01			0.19
feature	10 (14)	12 (16)		7 (12)	8 (14)		3 (21)	4 (29)	
collapse	6 (8)	45 (62)		6 (10)	38 (64)		0 (0)	7 (50)	

Table 3. Dataset adjustment methods by analysis type

	All studies N (%)[a]	Meta Analysis in title or abstract N (%)[a]
Alpha diversity	(n=96)	(n=87)
Ignore dataset	36 (38)	32 (37)
Parallel analysis but no pooling	16 (17)	14 (16)
Adjust for dataset	32 (33)	30 (34)
Other method	6 (6)	6 (7)
Not described	6 (6)	5 (6)
Not analyzed	40	37
Beta Diversity Ordination	(n=87)	(n=80)
Ignore dataset	21 (27)	20 (29)
Parallel analysis but no pooling	15 (19)	13 (19)
Adjust for dataset	32 (42)	28 (41)
Other method	7 (9)	7 (10)
Not described	2 (3)	1 (1)
Not analyzed	59	55
Beta diversity statistics	(n=77)	(n=69)
Ignore dataset	21 (27)	20 (29)
Parallel analysis but no pooling	15 (19)	13 (19)
Adjust for dataset	32 (42)	28 (41)
Other method	7 (9)	7 (10)
Not described	2 (3)	1 (1)
Not analyzed	59	55
Differential Abundance	(n=102)	(n=93)
Ignore dataset	28 (27)	25 (27)
Parallel analysis but no pooling	13 (13)	12 (13)
Adjust for dataset	48 (47)	44 (47)
Other method	6 (6)	5 (5)
Not described	7 (7)	7 (8)
Not analyzed	34	31

[a]Percentage of studies with the analysis performed; denominator is given in first line of section

Table 4. Varying definitions of "meta-analysis" observed in microbiome studies

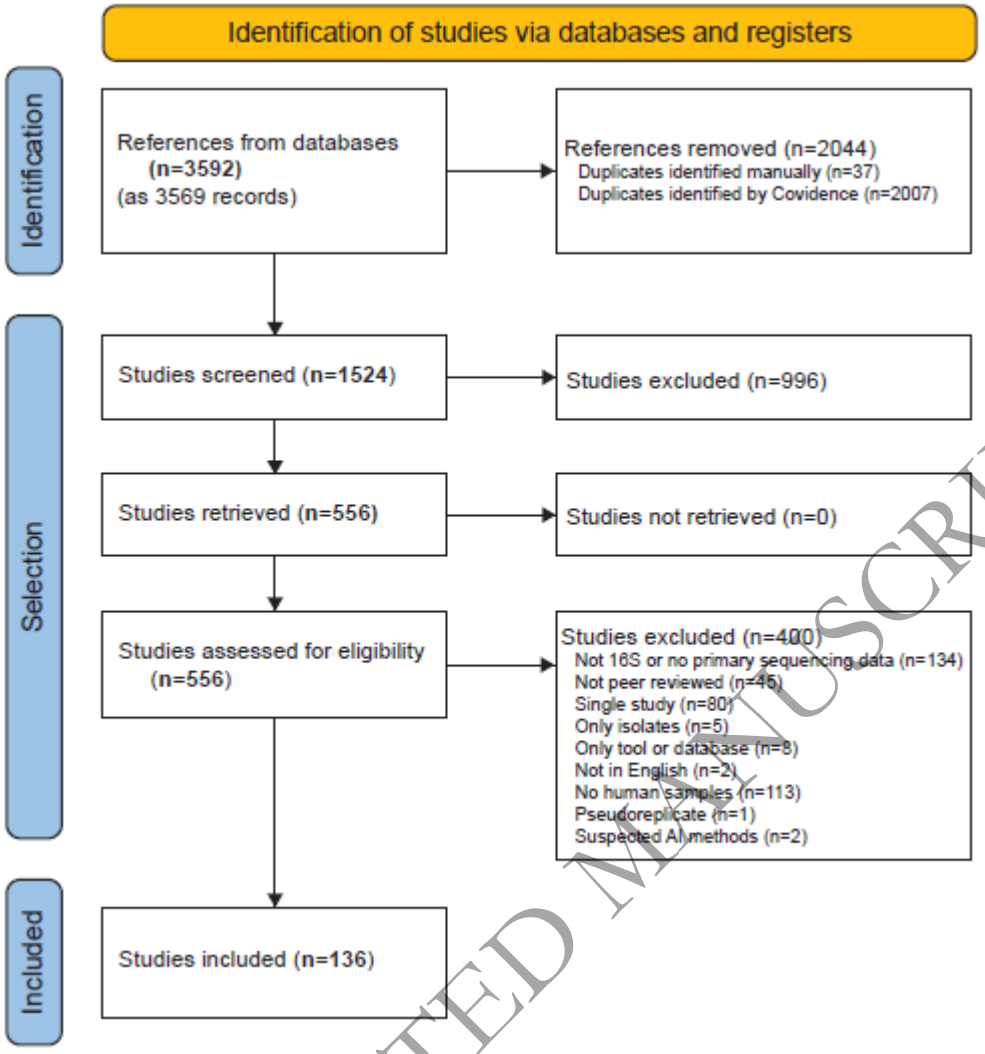
	Study protocol	Examples
1	Study processes and analyzes data using a common wet lab and bioinformatics protocol, where data were collected for a parent study (e.g., a cohort consortium) that encompasses multiple sub-studies with potentially different designs, study populations, or study sites.	61,62
2	Study combines two or more datasets identified without a systematic search process. Datasets are combined with a common bioinformatic process and analyzed together. Goal of the study is often to provide context or validate findings	27,28,63,64
3	Study identifies datasets through a systematic search process. Datasets are combined through a common bioinformatic process and analyzed together. While not universally true in the current literature, ideally the samples included in the dataset are described in a consistent manner, and analytical techniques are appropriate for dataset heterogeneity (e.g. adjustment for dataset specific effects through a fixed or random effect or the statistical pooling of estimates where appropriate for measurements).	65–67
4[a]	Study identifies datasets through a systematic search process. Summary statistics are extracted from published articles and combined using a statistical meta-analysis framework.	68–70

[a]This definition was observed during screening, but these articles were not included in our final scoping review since we required all included studies combine primary datasets from multiple sources.

Figure Captions

Figure 1. Selection of studies included in the scoping review

Alt text: A diagram showing the inclusion of studies in the analysis. The inital search identified 3592 references; 2044 were removed as duplicates. 1524 studies were screened at the title and abstract stage; 996 were discarded and 556 went through full text screening. There were 400 studies excluded: 134 without 16S sequencing data; 45 not peer reviewed; 80 did not combine data sets; 5 were sequencing isolates; 8 were tools or databases; 2 were not in English; 113 did not contain human samples; one study re-analyzed the same dataset; and the methods section of two studies was suspected of being AI. The final analysis included 136 studies.



11

12

Figure 2. Upset plot comparing the number of dataset sources used by study authors. The horizontal bar plot at the left gives the total number of studies which included a dataset from a given source; the vertical bar plot shows the number of studies with for each intersection of dataset sources. Closed data sources have a shaded background. SRA: Sequence Read Archive

Alt text: An upset plot showing the intersection between data sources used in manuscripts. There were two closed data sources (consortium, 28 articles total and request to original authors, 36 articles total) and three open sources (SRA, 105 articles total; Qiita 9 articles, total; and MG-Rast, 5 articles total). Common data source combinations included 64 articles with only SRA data; 17 articles with only consortium data; 24 with SRA data and request to authors; 6 with consortium and SRA data; and five that only requested data from authors.

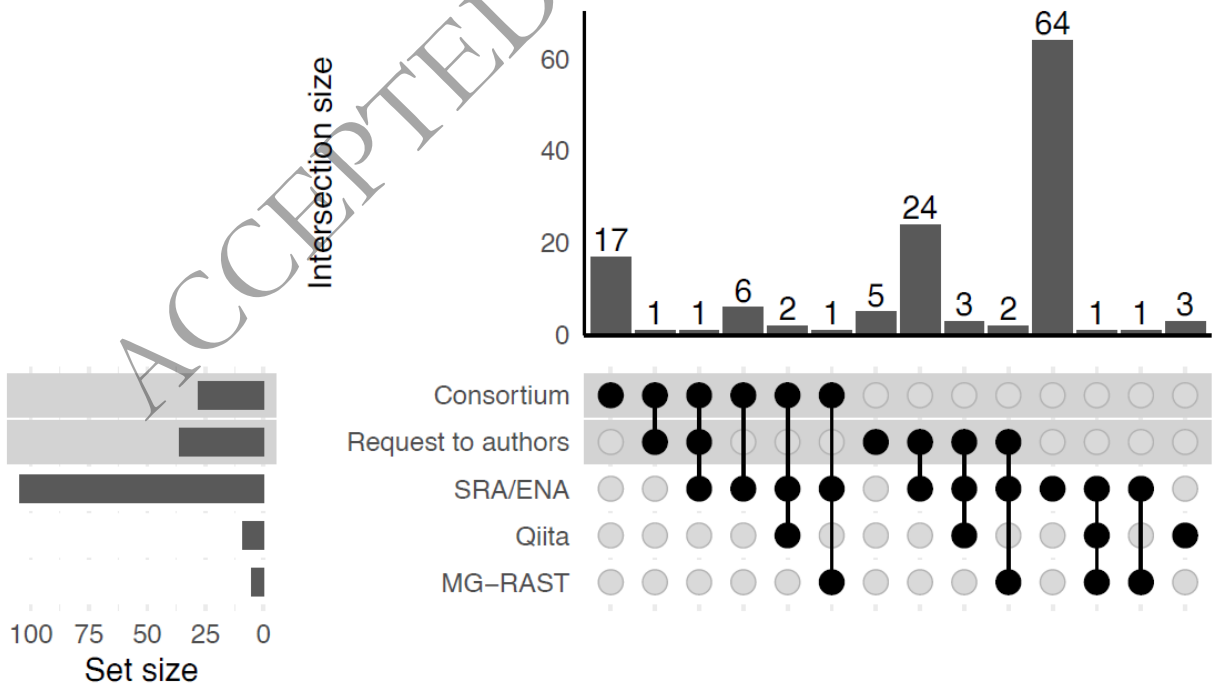
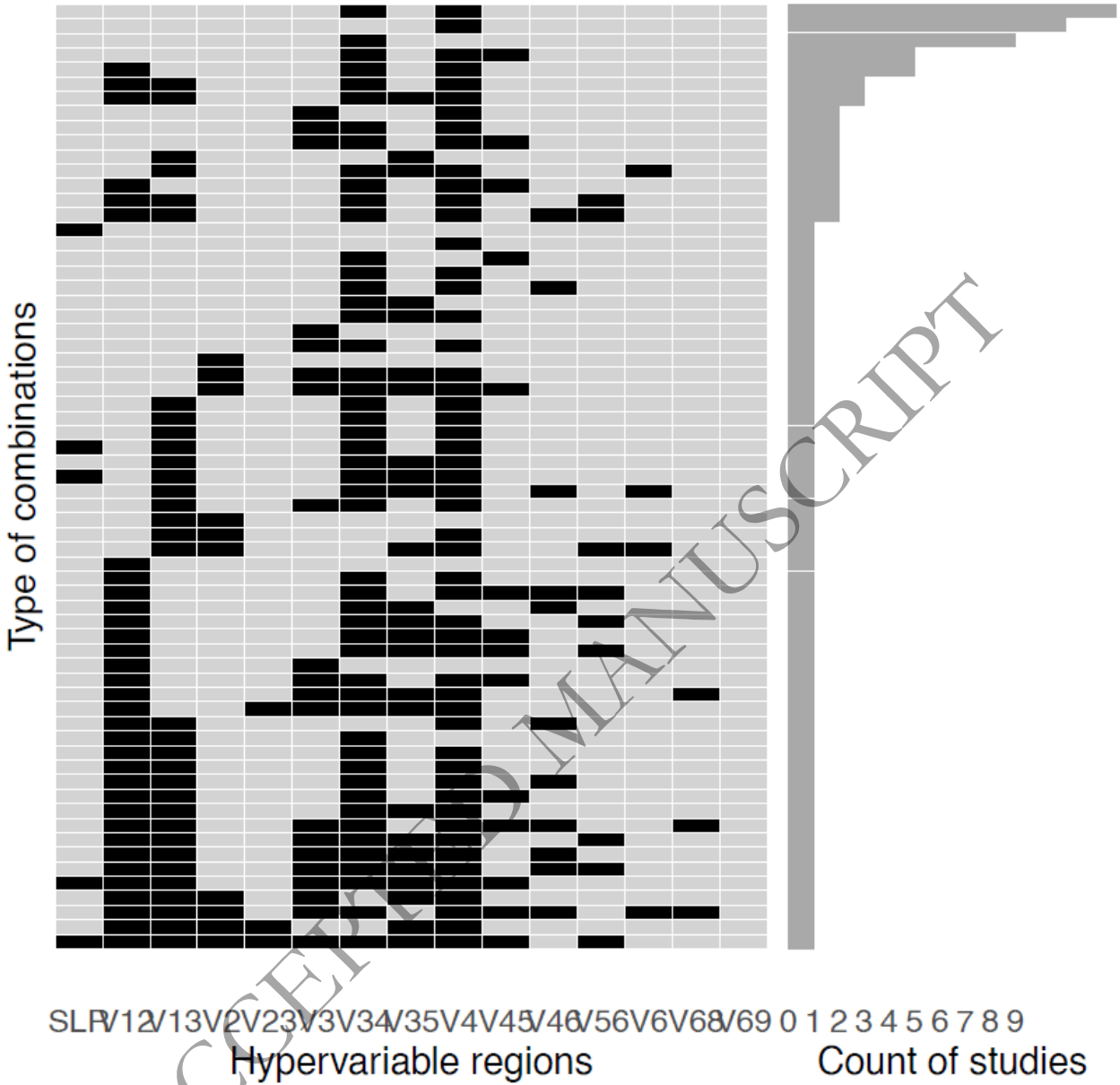


Figure 3. Hypervariable region combinations included in different studies. The matrix at the left shows the regions amplified by datasets included in each study; the bar chart at the right describes the number of studies which included that combination of hypervariable regions.

Alt text: Combinations of hypervariable regions and number of studies. Heatmap is arranged by hypervariable position along the 16S gene. The most common combinations were

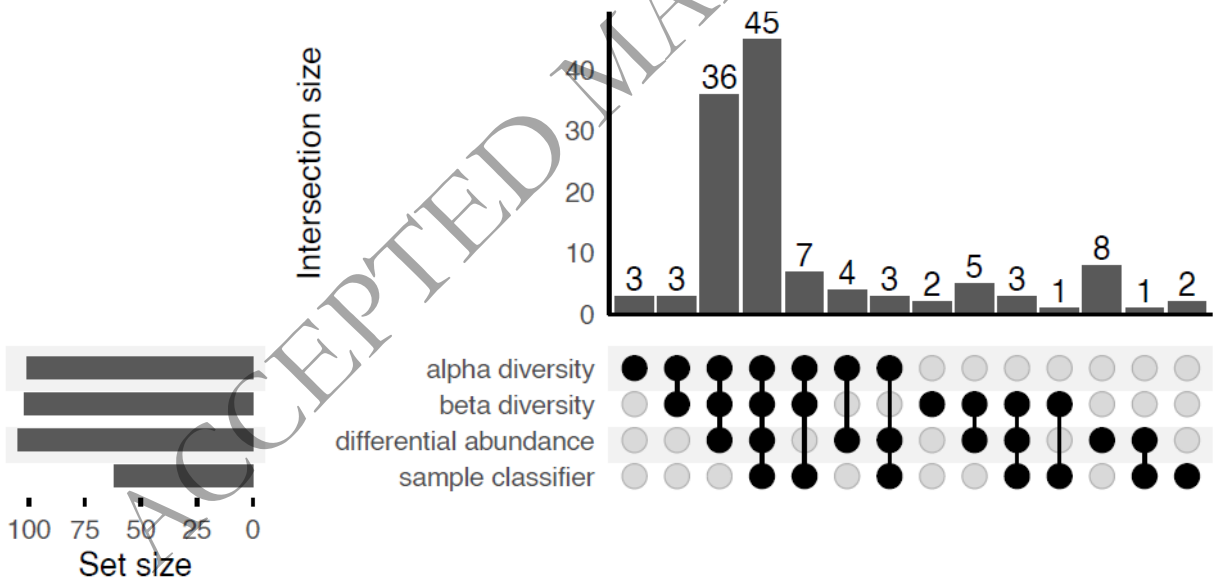


33

34

Figure 4. Analyses conducted by individual studies. Upset plot shows the number of studies which conducted each analysis, and the number of studies which conducted each combination.

Alt text: The intersection between different analyses. A total of 101 studies analyzed alpha diversity; 102 total analyzed beta diversity; 105 total performed differential abundance; and 62 total used a sample classifier. The most common combinations included 45 studies that did all four analyses; 36 studies that analyzed alpha diversity, beta diversity, and differential abundance; 8 studies which only performed differential abundance; and 7 studies which performed alpha diversity, beta diversity, and built a sample classifier.



Supplemental Materials

Supplementary File 1. Search Terms

Supplementary File 2. PRISMA scoping review checklist

Supplemental figures

Supplemental tables

Table S8. Percent of studies performing each analysis in outcome-exposure focused studies in humans

Table S9. Statistical considerations for dataset effects across common microbiome analyses

Table S10. Methods for handling dataset effects in alpha diversity analysis

Table S11.Methods for handling dataset effects in beta diversity ordination

Table S12. Methods for handling dataset effects in statistical analyses of beta diversity

Table S13. Methods for handling dataset effects in differential abundance analysis

Table S14. Meta-analysis reporting and application

References

1. Giles EM, Couper J. Microbiome in health and disease. *Journal of Paediatrics and Child Health*. 2020;56(11):1735-1738. doi:10.1111/jpc.14939

2. Frias-Lopez J, Duran-Pinedo AE. The Function of the Oral Microbiome in Health and Disease. In: Sahingur SE, ed. *Emerging Therapies in Periodontics*. Springer International Publishing; 2020:141-173. doi:10.1007/978-3-030-42990-4_10

3. Young VB. The role of the microbiome in human health and disease: an introduction for clinicians. *BMJ*. 2017;356:j831. doi:10.1136/bmj.j831

4. Castaner O, Goday A, Park YM, et al. The Gut Microbiome Profile in Obesity: A Systematic Review. *International Journal of Endocrinology*. 2018;2018(1):4095789. doi:10.1155/2018/4095789

5. Waterhouse M, Hope B, Krause L, et al. Vitamin D and the gut microbiome: a systematic review of in vivo studies. *Eur J Nutr*. 2019;58(7):2895-2910. doi:10.1007/s00394-018-1842-7

6. Grech A, Collins CE, Holmes A, et al. Maternal exposures and the infant gut microbiome: a systematic review with meta-analysis. *Gut Microbes*. 2021;13(1):1897210. doi:10.1080/19490976.2021.1897210

7. Sun J, Huang T, Debelius JW, Fang F. Gut microbiome and amyotrophic lateral sclerosis: A systematic review of current evidence. *Journal of Internal Medicine*. 2021;290(4):758-788. doi:10.1111/joim.13336

8. Robeson MS, O'Rourke DR, Kaehler BD, et al. RESCRIPt: Reproducible sequence taxonomy reference database management. *PLOS Computational Biology*. 2021;17(11):e1009581. doi:10.1371/journal.pcbi.1009581

9. Balvočiūtė M, Huson DH. SILVA, RDP, Greengenes, NCBI and OTT — how do these taxonomies compare? *BMC Genomics*. 2017;18(2):114. doi:10.1186/s12864-017-3501-4

10. Ceccarani C, Severgnini M. A comparison between Greengenes, SILVA, RDP, and NCBI reference databases in four published microbiota datasets. *bioRxiv*. Preprint posted online April 13, 2023:2023.04.12.535864. doi:10.1101/2023.04.12.535864

11. Parks DH, Chuvochina M, Waite DW, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol*. 2018;36(10):996-1004. doi:10.1038/nbt.4229

12. Nearing JT, Douglas GM, Hayes M, et al. Microbiome differential abundance methods produce disturbingly different results across 38 datasets. Published online May 10, 2021:2021.05.10.443486. doi:10.1101/2021.05.10.443486

13. Walker AW, Hoyles L. Human microbiome myths and misconceptions. *Nat Microbiol*. 2023;8(8):8. doi:10.1038/s41564-023-01426-7

14. Abellan-Schneyder I, Matchado MS, Reitmeier S, et al. Primer, Pipelines, Parameters: Issues in 16S rRNA Gene Sequencing. *mSphere*. 2021;6(1):10.1128/msphere.01202-20. doi:10.1128/msphere.01202-20

15. Nearing JT, Langille MGI. Identifying biases and their potential solutions in human microbiome studies. *Microbiome*. 2021;9(1):113. doi:10.1186/s40168-021-01059-0

16. Hugerth LW, Pereira M, Zha Y, et al. Assessment of In Vitro and In Silico Protocols for Sequence-Based Characterization of the Human Vaginal Microbiome. *mSphere*. 2020;5(6). doi:10.1128/mSphere.00448-20

17. Johnson JS, Spakowicz DJ, Hong BY, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun*. 2019;10(1):5029. doi:10.1038/s41467-019-13036-1

18. Andersson AF, Lindberg M, Jakobsson H, Bäckhed F, Nyrén P, Engstrand L. Comparative analysis of human gut microbiota by barcoded pyrosequencing. *PLoS One*. 2008;3(7):e2836. doi:10.1371/journal.pone.0002836

19. Sun Y, Chen Y, Li Z, Debelius J, Debelius J. Scoping review of microbiome harmonization techniques. Published online January 29, 2025. doi:10.5281/zenodo.14767165

20. Reback J, jbrockmendel, McKinney W, et al. pandas-dev/pandas: Pandas 1.2.5. Published online June 22, 2021. doi:10.5281/zenodo.5013202

21. McKinney W. Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference*. Published online 2010:56-61. doi:10.25080/Majora-92bf1922-00a

22. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2024. https://www.R-project.org/

23. Wickham H, Averick M, Bryan J, et al. Welcome to the Tidyverse. *Journal of Open Source Software*. 2019;4(43):1686. doi:10.21105/joss.01686

24. Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. UpSet: Visualization of Intersecting Sets. *IEEE Transactions on Visualization and Computer Graphics*. 2014;20(12):1983-1992. doi:10.1109/TVCG.2014.2346248

25. Krassowski M, Arts M, Lagger C, Max. krassowski/complex-upset: v1.3.5. Published online November 11, 2022. doi:10.5281/zenodo.7314197

26. Wickham H. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York; 2016. https://ggplot2.tidyverse.org

27. Beaumont M, Goodrich JK, Jackson MA, et al. Heritable components of the human fecal microbiome are associated with visceral fat. *Genome Biol*. 2016;17(1):189. doi:10.1186/s13059-016-1052-7

28. Dei-Cas I, Giliberto F, Luce L, Dopazo H, Penas-Steinhardt A. Metagenomic analysis of gut microbiota in non-treated plaque psoriasis patients stratified by disease severity: development of a new Psoriasis-Microbiome Index. *Sci Rep*. 2020;10(1):12754.

29. Demirkan A, van Dongen J, Finnicum CT, et al. Linking the gut microbiome to host DNA methylation by a discovery and replication epigenome-wide association study. *BMC Genomics*. 2024;25(1):1224. doi:10.1186/s12864-024-11136-x

30. Vangay P, Burgin J, Johnston A, et al. Microbiome Metadata Standards: Report of the National Microbiome Data Collaborative's Workshop and Follow-On Activities. *mSystems*. 2021;6(1):e01194-20. doi:10.1128/mSystems.01194-20

31. Langille MGI, Ravel J, Fricke WF. "Available upon request": not good enough for microbiome data! *Microbiome*. 2018;6(1):8. doi:10.1186/s40168-017-0394-z

32. Debelius J, Song SJ, Vazquez-Baeza Y, Xu ZZ, Gonzalez A, Knight R. Tiny microbes, enormous impacts: what matters in gut microbiome studies? *Genome Biol*. 2016;17(1):217. doi:10.1186/s13059-016-1086-x

33. Janssen S, McDonald D, Gonzalez A, et al. Phylogenetic Placement of Exact Amplicon Sequences Improves Associations with Clinical Information. *mSystems*. 2018;3(3). doi:10.1128/mSystems.00021-18

34. Washburne AD, Morton JT, Sanders J, et al. Methods for phylogenetic analysis of microbiome data. *Nat Microbiol*. 2018;3(6):652-661. doi:10.1038/s41564-018-0156-0

35. Deeks JJ, Higgins JP, Altman DG, McKenzie JE, Veroniki AA. Chapter 10: Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Page M, Welch V, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. 6.5. Cochrane; 2024. Accessed June 5, 2024. https://training.cochrane.org/handbook/current/chapter-10

36. Eisenhauer JG. Meta-analysis and mega-analysis: A simple introduction. *Teaching Statistics*. 2021;43(1):21-27. doi:10.1111/test.12242

37. Tierney JF, Stewart LA, Clarke M. Chapter 26: Individual participant data. In: Higgins JP, Thomas J, Chandler J, Li T, Page M, Welch V, eds. *Cochrane Handbook for*

Systematic Reviews of Interventions. 6.4. 2023. Accessed July 11, 2024. https://training.cochrane.org/handbook/current/chapter-26

38. Debray TPA, Moons KGM, van Valkenhoef G, et al. Get real in individual participant data (IPD) meta-analysis: a review of the methodology. *Res Synth Methods*. 2015;6(4):293-309. doi:10.1002/jrsm.1160

39. Walters WA, Xu Z, Knight R. Meta-analyses of human gut microbes associated with obesity and IBD. *FEBS Lett*. 2014;588(22):4223-4233. doi:10.1016/j.febslet.2014.09.039

40. Quinn TP. Stool Studies Don't Pass the Sniff Test: A Systematic Review of Human Gut Microbiome Research Suggests Widespread Misuse of Machine Learning. *arXiv*. Preprint posted online July 8, 2021:arXiv:2107.03611. doi:10.48550/arXiv.2107.03611

41. Rahaman T. Open Data and the 2023 NIH Data Management and Sharing Policy. *Med Ref Serv Q*. 2023;42(1):71-78. doi:10.1080/02763869.2023.2168103

42. Martone ME, Nakamura R. Changing the Culture on Data Management and Sharing: Getting Ready for the New NIH Data Sharing Policy. *Harvard Data Science Review*. 2022;4(3). doi:10.1162/99608f92.6650ce2b

43. Federer LM, Belter CW, Joubert DJ, et al. Data sharing in PLOS ONE: An analysis of Data Availability Statements. *PLoS One*. 2018;13(5):e0194768. doi:10.1371/journal.pone.0194768

44. Tedersoo L, Küngas R, Oras E, et al. Data sharing practices and data availability upon request differ across scientific disciplines. *Sci Data*. 2021;8(1):192. doi:10.1038/s41597-021-00981-0

45. Gorgolewski KJ, Margulies DS, Milham MP. Making data sharing count: a publication-based solution. *Front Neurosci*. 2013;7:9. doi:10.3389/fnins.2013.00009

46. Wilkinson MD, Dumontier M, Aalbersberg IJJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 2016;3:160018. doi:10.1038/sdata.2016.18

47. Huttenhower C, Finn RD, McHardy AC. Challenges and opportunities in sharing microbiome data and analyses. *Nat Microbiol*. 2023;8(11):1960-1970. doi:10.1038/s41564-023-01484-x

48. Perrier L, Blondal E, MacDonald H. The views, perspectives, and experiences of academic researchers with data sharing and reuse: A meta-synthesis. *PLoS One*. 2020;15(2):e0229182. doi:10.1371/journal.pone.0229182

49. Byrd JB, Greene AC, Prasad DV, Jiang X, Greene CS. Responsible, practical genomic data sharing that accelerates research. *Nat Rev Genet*. 2020;21(10):615-629. doi:10.1038/s41576-020-0257-5

50. Carroll SR, Garba I, Figueroa-Rodríguez OL, et al. The CARE Principles for Indigenous Data Governance. *Data Science Journal*. 2020;19(1). doi:10.5334/dsj-2020-043

51. Muka T, Glisic M, Milic J, et al. A 24-step guide on how to design, conduct, and successfully publish a systematic review and meta-analysis in medical research. *Eur J Epidemiol*. 2020;35(1):49-60. doi:10.1007/s10654-019-00576-5

52. Leonardi-Bee J, Flohr C, van Zuuren EJ, Le Cleach L, Hollestein LM. Common methodological pitfalls and new developments in systematic review meta-analyses. *Br J Dermatol*. 2019;181(4):649-651. doi:10.1111/bjd.18336

53. Cook DJ, Sackett DL, Spitzer WO. Methodologic guidelines for systematic reviews of randomized control trials in health care from the Potsdam Consultation on Meta-Analysis. *J Clin Epidemiol*. 1995;48(1):167-171. doi:10.1016/0895-4356(94)00172-m

54. Bharucha T, Oeser C, Balloux F, et al. STROBE-metagenomics: a STROBE extension statement to guide the reporting of metagenomics studies. *Lancet Infect Dis*. 2020;20(10):e251-e260. doi:10.1016/S1473-3099(20)30199-7

55. Mirzayi C, Renson A, Genomic Standards Consortium, et al. Reporting guidelines for human microbiome research: the STORMS checklist. *Nat Med*. 2021;27(11):1885-1892. doi:10.1038/s41591-021-01552-x

56. O'Dea RE, Lagisz M, Jennions MD, et al. Preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology: a PRISMA extension. *Biol Rev Camb Philos Soc*. 2021;96(5):1695-1722. doi:10.1111/brv.12721

57. Minot SS, Garb B, Roldan A, et al. MaLiAmPi enables generalizable and taxonomy-independent microbiome features from technically diverse 16S-based microbiome studies. *Cell Reports Methods*. 2023;3(11):100639. doi:10.1016/j.crmeth.2023.100639

58. Kleine Bardenhorst S, Berger T, Klawonn F, Vital M, Karch A, Rübsamen N. Data Analysis Strategies for Microbiome Studies in Human Populations—a Systematic Review of Current Practice. *mSystems*. 2021;6(1):10.1128/msystems.01154-20. doi:10.1128/msystems.01154-20

59. Vitali F, Colucci R, Di Paola M, et al. Early melanoma invasivity correlates with gut fungal and bacterial profiles. *Br J Dermatol*. 2022;186(1):106-116. doi:10.1111/bjd.20626

60. McDonald D, Jiang Y, Balaban M, et al. Greengenes2 unifies microbial data in a single reference tree. *Nat Biotechnol*. 2024;42(5):715-718. doi:10.1038/s41587-023-01845-1

61. Thompson LR, Sanders JG, McDonald D, et al. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature*. 2017;551(7681):457+. doi:10.1038/nature24621

62. McDonald D, Hyde E, Debelius JW, et al. American Gut: an Open Platform for Citizen Science Microbiome Research. *mSystems*. 2018;3(3). doi:10.1128/mSystems.00031-18

63. Lozupone CA, Stombaugh J, Gonzalez A, et al. Meta-analyses of studies of the human microbiota. *Genome Res*. 2013;23(10):1704-1714. doi:10.1101/gr.151803.112

64. Smits SA, Leach J, Sonnenburg ED, et al. Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. *Science*. 2017;357(6353):802-806. doi:10.1126/science.aan4834

65. Nishiwaki H, Ito M, Ishida T, et al. Meta-Analysis of Gut Dysbiosis in Parkinson's Disease. *Mov Disord*. 2020;35(9):1626-1635.

66. Gao C, Guo Y, Chen F. Cross-Cohort Microbiome Analysis of Salivary Biomarkers in Patients With Type 2 Diabetes Mellitus. *Front Cell Infect Microbiol*. 2022;12. doi:10.3389/fcimb.2022.816526

67. Wang Z, Yang Y, Yan Z, et al. Multi-omic meta-analysis identifies functional signatures of airway microbiome in chronic obstructive pulmonary disease. *ISME J*. 2020;14(11):2748-2765. doi:10.1038/s41396-020-0727-y

68. Bai J, Barandouzi ZA, Rowcliffe C, Meador R, Tsementzi D, Bruner DW. Gut Microbiome and Its Associations With Acute and Chronic Gastrointestinal Toxicities in Cancer Patients With Pelvic Radiation Therapy: A Systematic Review. *Frontiers in Oncology*. 2021;11. doi:10.3389/fonc.2021.745262

69. Gudnadottir U, Debelius JW, Du J, et al. The vaginal microbiome and the risk of preterm birth: a systematic review and network meta-analysis. *Sci Rep*. 2022;12(1):7926. doi:10.1038/s41598-022-12007-9

70. Pinart M, Dötsch A, Schlicht K, et al. Gut Microbiome Composition in Obese and Non-Obese Persons: A Systematic Review and Meta-Analysis. *Nutrients*. 2021;14(1):12. doi:10.3390/nu14010012