

REVIEW

Open Access



# Identifying biases and their potential solutions in human microbiome studies

Jacob T. Nearing<sup>1</sup>, André M. Comeau<sup>2</sup> and Morgan G. I. Langille<sup>2,3\*</sup>

## Abstract

Advances in DNA sequencing technology have vastly improved the ability of researchers to explore the microbial inhabitants of the human body. Unfortunately, while these studies have uncovered the importance of these microbial communities to our health, they often do not result in similar findings. One possible reason for the disagreement in these results is due to the multitude of systemic biases that are introduced during sequence-based microbiome studies. These biases begin with sample collection and continue to be introduced throughout the entire experiment leading to an observed community that is significantly altered from the true underlying microbial composition. In this review, we will highlight the various steps in typical sequence-based human microbiome studies where significant bias can be introduced, and we will review the current efforts within the field that aim to reduce the impact of these biases.

**Keywords:** Microbiome, Methodology, Bias, Metagenomics, Amplicon, Contamination

## Introduction

Recent advances in DNA sequencing technology have allowed community-wide investigation into the microbes that live on and within the human body. These communities and their cellular functions are known as the human microbiome, which has now been associated with multiple different influences on human health [1, 2]. Investigation into these microbes has led to the development of both novel therapeutics [3] and diagnostic tools [4]. However, results from different studies often do not match with previous findings [5, 6]. One possible reason for inconsistent results across studies is the high level of random and systemic bias that is introduced throughout sequenced-based human microbiome studies.

Microbiome studies are particularly at fault for elevated levels of these biases due to the high sensitivity of DNA sequencing instruments and the relatively

unknown underlying microbial compositions within a sample. Throughout a typical microbiome study, there are numerous areas where biases are introduced [7]. The introduction of these biases often results in distorted observations of the true underlying microbial composition contained within a sample [8]. While sometimes these biases can be subtle, they often result in significant impacts on biological conclusions. In this report, we will review each step in a typical marker gene and metagenomic shotgun sequencing microbiome study, starting with sample collection and ending with downstream bioinformatic analysis (Fig. 1). Throughout each section, we will highlight the various biases that are introduced during each step, and we will then review current approaches that have been taken to help alleviate those biases to improve our understanding of the human microbiome (Table 1).

\* Correspondence: [morgan.langille@dal.ca](mailto:morgan.langille@dal.ca)

<sup>2</sup>Integrated Microbiome Resource, Dalhousie University, Halifax, Nova Scotia, Canada

<sup>3</sup>Department of Pharmacology, Dalhousie University, Halifax, Nova Scotia, Canada

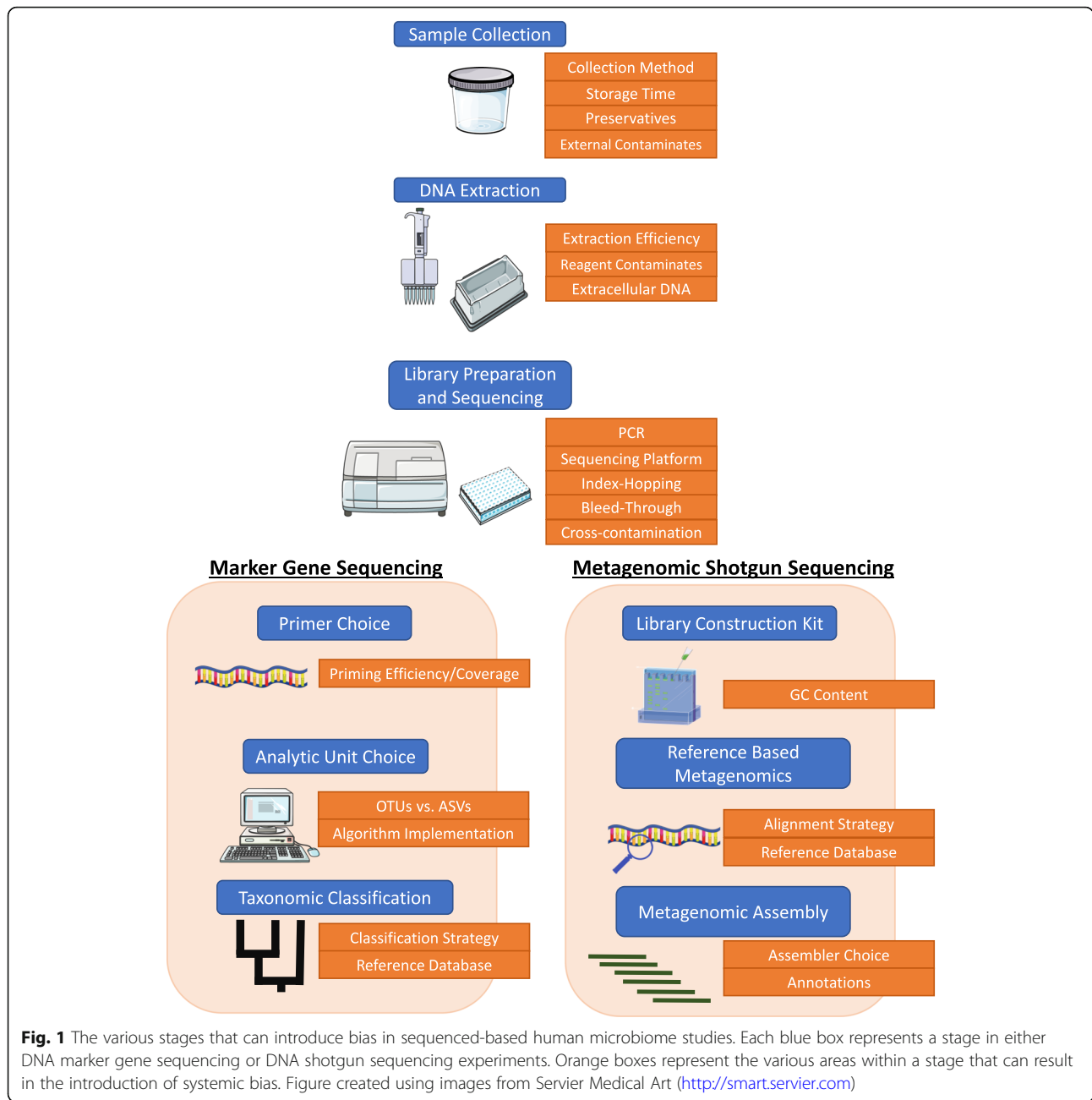
Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Sample collection and storage

Sample collection and storage are a vital part of an experiment and if not carefully planned can often introduce unaccounted bias. The way in which the sample is



collected, how long the sample is stored, and what the sample is stored in can all impact the underlying observed microbial composition. All these details should be noted during sample collection and accounted for during downstream analysis to ensure consistent results across studies.

**Sample collection**

The method of sample collection is particularly important to consider as different types of systematic bias can be introduced depending on the sampling environment. This has been highlighted in studies of the various areas of the human microbiome. Work on the gut microbiome

has compared colon biopsy samples to both stool and rectal swab samples. In both cases, it was found that collection of samples by biopsy introduced strong biases toward the identification of specific microbes such as those that adhere to the mucosal wall of the colon [9–12]. This is not surprising given the drastically different collection methods. For example, biopsies often have higher levels of human DNA and do not account for impacts on community composition during the transit of stool in the colon [13, 14].

In contrast, multiple studies have compared the observed microbial communities from rectal swabs and

**Table 1** Recommended best practices during sequenced-based microbiome studies

Processing step	Recommended best practices
Sample collection	<ul style="list-style-type: none"> <li>- Collect biological and technical replicates when possible.</li> <li>- Use the same collection device and manufacturer.</li> <li>- When possible, use the same collection personnel.</li> <li>- Use aseptic techniques during sample collection.</li> <li>- Make note of any variations during sample collection and include them in downstream analysis.</li> </ul>
Sample storage	<ul style="list-style-type: none"> <li>- When possible freeze samples at <math>-80^{\circ}\text{C}</math> immediately upon collection.</li> <li>- Exposure of samples to room temperature conditions should be minimized.</li> <li>- Preservatives should be used only when freezing of samples is infeasible (e.g., self-collected human samples).</li> <li>- If used, all samples should be stored in the same preservative.</li> <li>- Length of sample storage should be noted and included in downstream analysis.</li> </ul>
DNA extraction	<ul style="list-style-type: none"> <li>- Extractions should be done using validated extraction kits or validated protocols such as those presented by the Earth Microbiome project or International Human Microbiome Standards.</li> <li>- All samples must be extracted with the same protocol.</li> <li>- Extraction batches should be noted and used as covariates in downstream analysis.</li> <li>- Extraction should include a mechanical lysis step (e.g., bead-beating).</li> <li>- Extraction should be done using aseptic techniques and a biological safety cabinet to reduce the amount of possible contamination.</li> <li>- A small pool of samples should be extracted during each extraction batch and sequenced to determine technical variation.</li> <li>- Blanks should be carried through extraction to sequencing (critical for low-biomass samples).</li> </ul>
PCR amplification	<ul style="list-style-type: none"> <li>- Use high-fidelity polymerases.</li> <li>- Minimize the number of required PCR cycles (preferably max. 20–25).</li> <li>- Obtain as uniform amplification as possible for all samples.</li> <li>- Primers should be chosen based on the microbes of interest and whether the work is to be compared against previous literature.</li> </ul>
Metagenomic library construction	<ul style="list-style-type: none"> <li>- Use equal amounts of template for library construction.</li> <li>- Avoid usage of discontinued Illumina Nextera XT kit.</li> <li>- Note mechanical sonication can cause minor biases toward high-GC content sequences.</li> </ul>
Marker gene bioinformatics	<ul style="list-style-type: none"> <li>- Use denoising algorithms such as Deblur, DADA2, or UNOISE.</li> <li>- Use validated taxonomic classification methods such as QIIME2 feature classifiers or Kraken2.</li> <li>- Taxonomic classifiers should be consistent between comparison studies.</li> <li>- Use well-curated, up-to-date, comprehensive taxonomic databases such as SILVA, RDP, or NCBI.</li> </ul>
Metagenomic shotgun bioinformatics	<ul style="list-style-type: none"> <li>- Referenced-based analysis: <ul style="list-style-type: none"> <li>Use DNA-based KMER taxonomic assignment such as Kraken2 or taxMaps.</li> <li>Removal/filtering of low abundance taxa is recommended.</li> <li>Employ phylogenetic marker-gene-based strategies when examining low abundance taxa.</li> </ul> </li> <li>- Metagenomic assembly: <ul style="list-style-type: none"> <li>Use well-validated assembly methods such as MEGAHIT or metaSPAdes.</li> <li>Use validated binning methods such as DASTool, MetaWrap, or MetaBat2.</li> </ul> </li> <li>- For all methods, reference databases should be well-curated and up-to-date such as: <ul style="list-style-type: none"> <li>NCBI RefSeq</li> <li>Genome Taxonomy Database</li> </ul> </li> </ul>

fecal samples and have found similar profiles. Bassis et al. [15] compared stool samples and rectal swabs from the same patients and found no significant difference in microbial abundances between these collection methods. A further report by Reyman et al. [16] has shown similar results; however, a recent study by Jones et al. reported significant differences in both taxonomic makeup and functional pathway abundances between rectal swab and stool samples. They found that 24 of 48 families were significantly different in relative abundance between swab and stool samples through 16S rRNA gene sequencing, and a higher proportion of aerobic bacteria could be found in rectal swabs. Furthermore, they found numerous differences in levels 1 and 2 KEGG pathways through metagenomic shotgun sequencing (MGS) analysis. However, it should be noted that they did find

rectal swab microbial profiles, matched stool profiles closer than biopsies [17]. Overall, these results indicate that while stool and biopsy samples reveal different microbial communities in both membership and diversity, rectal swabs may provide a viable easy-to-collect alternative to stool while keeping in mind that the proportions of aerobic genera will be elevated.

Another body site of interest that has been studied comprehensively is the human oral microbiome. Within this environment, it has been shown that while different sites of the oral cavity such as the tongue, dental plaques, and saliva accommodate unique microbial communities [18, 19], the sampling of these communities using different techniques has been shown to be comparable. Work by Fan et al. [20] and Jo et al. [21] compared human oral microbiome compositions collected through

several approaches, including mouthwashes, water rinses, and unstimulated or stimulated saliva. They found no significant differences in community composition at the genus level due to the sampling method. Furthermore, work on comparing dental plaque collection methods found no differences in total DNA extracted, alpha diversity, or overall community structure between using a scaler or CytoSoft brush [22].

Several other body sites contain microbial communities that have been examined using several different collection methods. Studies that have compared different sampling techniques used in these areas have shown variable results. For example, reports have shown that skin swabs and tape-stripping result in similar family level relative abundances (Rho 0.792-0.999) [23] and no differences in alpha diversity [23, 24]. Interestingly, work by Bjerre et al. [25] comparing skin swabs and skin scrapping showed significant differences in alpha diversity between sampling techniques despite having an overlap of ~90% between the OTUs identified through each method. A similar conclusion was found between lung brushings and exhaled breath condensates where lung brushings were found to have significantly higher DNA levels and cluster separately from breath condensates when comparing overall community structure [26]. Overall, these cases highlight the fact that new collection methods should be compared with previous work to determine whether they will introduce, or conversely reduce, a significant amount of bias during downstream analysis.

Overall, the most important thing to consider during sample collection is the use of consistent collection methods within any given analysis. The same aseptic techniques should be employed during the collection of all samples. Furthermore, due to the high sensitivity of sequencing instruments, the use of the same type and manufacturer of collection devices is heavily advised, as DNA from collection devices can be introduced into samples [27]. When possible, all samples should be collected by the same individual as batch effects may be introduced due to differences in the individual collectors. This could be due to several issues including slight variations in collection technique between individuals or the individual's own microbial flora contaminating the sample. This principle also holds true for sample preparation and sequencing. Any variation in sample collection methodology, such as those highlighted above, should be noted during collection and be included as confounded effects during data analysis.

### Study design

While the proper collection of samples is critical for any microbiome study, samples need to be collected in a manner that can be used to answer the hypothesis in

question. Common study designs for microbiome studies include longitudinal, cross-sectional, and cohort designs [28]. While a thorough description of study designs is beyond the scope of this review, we have highlighted a number of factors that should be of particular importance to reduce systemic bias or confounding information within your study.

All studies should randomize samples during extraction and sequencing to ensure that technical variation is equally spread between samples. Previous work has shown that batch effects due to non-random sample preparation can lead to spurious results [29]. Furthermore, when possible, samples should be extracted using the same extraction kit lot and within the smallest time frame as possible. When this is not possible, extraction kit lot numbers and extraction dates should be included as confounding variables during data analysis. During the identification and recruitment of participants in your study, we suggest matching them on basic dietary information along with their age and BMI, although it should be noted that these choices are sample specific, as recent work on the oral microbiome has shown that the effect size of these factors is relatively small with no single factor explaining a variation larger than 2% [30]. Finally, the most important thing researchers need to do is to be consistent in the way samples are handled, and when abnormalities occur, they should be noted and included during any data analysis.

### Timing of sample storage

It often is impractical to immediately analyze fresh samples in microbiome research. As a result, many samples are preserved through storage at  $-80^{\circ}\text{C}$ . It remains unclear whether freezing samples significantly affects the resulting community profiles. Work by Bahl et al. [31] compared frozen and fresh human stool samples from the same donors using quantitative PCR (qPCR) and found significant differences in the *Firmicutes* to *Bacteroidetes* ratio. However, subsequent work by Fouhy et al. [32] comparing frozen and fresh stool samples using both culture and DNA sequence-based approaches showed no significant differences in the relative abundances of bacteria at the phylum or family level. At the genus level, there was only minor evidence suggesting that the relative abundances of *Faecalibacterium* and *Leuconostoc* may be biased due to sample freezing. Taken together, these results have helped reinforce the current standard of immediately freezing samples at  $-80^{\circ}\text{C}$  upon collection [33].

Research on what happens to room temperature-stored stool, which can commonly occur when freezers are not readily available, has shown promising results. Work by both Shaw et al. [34] and Bokulich et al. [35] has shown that fresh and 2-day-old stool and stool

swabs have similar alpha diversity measurements and observed microbial relative abundances. However, Shaw et al. did note significant differences in weighted beta diversity measurements after 48 h of storage. Longer durations of room temperature storage, such as those sometimes seen during the shipping of self-collected stool samples, has indicated that Enterobacteriaceae and other aerobic bacteria can begin to bloom while alpha diversity measurements such as richness can decrease over time [35–37]. Fortunately, work by Amir et al. has characterized 20 different 16S rRNA gene sequences mainly belonging to Gammaproteobacteria that commonly bloom during room temperature storage [38]. Based on these identifications, they have developed bloom-filtering software that reduces taxonomic biases due to room temperature storage and have successfully implanted its use in the analysis of data from the American Gut project [36].

Not only does the length of time samples remain unpreserved need to be considered when looking at biases during microbiome studies, but so does the length of time samples are frozen. Work by Lauber et al. [39] compared multiple different short-length storage timings (3–14 days) of frozen stool and skin samples and found that samples clustered together based on their sampling source rather than the amount of time they were frozen, indicating minimal impact on overall microbial composition in the short term. Investigation into longer freezing times by Shaw et al. [34] found that all diversity metrics remained stable throughout storage with the exception of OTU richness, which slightly decreased over time. Comparing taxonomic abundances between samples stored from 2 months up to 2 years resulted in no significant differences. Overall, these findings indicated that long-term storage may introduce small biases in overall sample richness that should be taken note of during analysis, but the overall community structure and relative abundances of microbial DNA within samples remains similar for at least 2 years of storage time. Based on these results, when examining samples from longitudinal designs, researchers should be wary of any conclusions showing slight consistent reductions in microbial richness over long periods of time unless samples were sequenced upon or near the collection date.

### Sample preservatives

While the current gold standard of sample preservation is to immediately store samples at  $-80^{\circ}\text{C}$ , this is not always feasible. This has led to the development of multiple different preservatives that allow samples to be stored at room temperature for extended periods of time. This is accomplished by inhibiting microbial growth within the sample while preserving microbial DNA. Currently, there are several different preservatives

that can be used to store microbiome samples including RNAlater, OMNIgene preservatives, Tris-EDTA, FTA cards, and ethanol. While in many cases the use of preservatives has been shown to not impact DNA quality or yield [33], they often introduce biases toward the detection of specific microorganisms. In 2015, Choo et al. [40] compared commonly used microbiome preservatives for stool and found that each one was linked to the introduction of bias in observed community composition. Preservation in OMNIgene.GUT and RNAlater was linked with inflated microbial diversity when compared to frozen samples, and Tris-EDTA usage led to significant differences in the relative abundance of multiple taxa [40]. Additional work on biases introduced due to the use of preservatives has shown that the use of RNAlater leads to significant biases toward the detection of certain taxonomic groups [41, 42], and that Stool Nucleic Acid Collection and Preservation tubes (Norgen BioTek Corp.) can result in the over-identification of gram-negative bacteria [43]. Finally, work by Song et al. [37] has shown that the use of FTA cards results in higher levels of taxonomic diversity, but only small consistent differences in taxonomic proportions. Overall, these results indicate that the use of preservatives can have a substantial impact on the observed relative abundances of different taxonomic groups within a sample, but large scale overall community structure metrics may be comparable between some methods [44]. Since the type of preservative chosen can result in a biased observation of the true underlying microbial community, many may want to avoid their use. If freezing samples immediately is not possible, the same preservative should be used for all samples within a study and should be noted within the methodology section of the corresponding manuscripts.

### Benchmark sample/technical replicates

A possible solution that could aid in dealing with batch effects during both sample collection and downstream processing is the increased usage of technical replicates. Currently, the use of technical replicates during a microbiome study is uncommon. While sequencing facilities may initially benchmark their sequencing process on multiple mock communities, biological samples or samples reflecting a composition like those of the sample of interest are rarely sequenced twice. This is often due to budgetary constraints and the thought that a larger number of unique biological samples is a better use of resources than the creation of technical replicates, especially when they are pseudoreplicates and not true biological replicates. While there may be some truth to this argument, the use of technical replicates has numerous benefits and serves as an important tool to help identify batch effects during sample processing. It allows the

researchers to identify the amount of variance in their sample preparation procedures and, as such, allows one to determine the minimal effect size that can be reliably determined as being from biological signal vs. originating from technical variation.

One possible avenue researchers could take to help with the budgetary constraints of the creation of technical replicates is to pick a small group (or one pool) of samples that will be processed throughout each batch of sample processing. By comparing the same biological sample throughout numerous sample batches, researchers can benchmark the variance in composition that they can expect between sample sequencing and preparation rounds. While in the past, mock communities have generally served this purpose, they often do not reflect the composition of the samples of interest, and, as such, a biological sample that is similar in composition to those under examination should be chosen.

### DNA extraction

After sample collection and storage, the next step in most microbiome experiments is DNA extraction. During this step, there are three main areas that can result in significant biases within a study: the differing DNA extraction efficiencies of microorganisms, the introduction of contaminant DNA, and the introduction of DNA from non-viable microorganisms. Due to these issues, DNA extraction is one of the most biased steps in sequenced-based human microbiome studies [6, 45].

### Extraction efficiencies

It is well-known that many microbes exhibit different DNA extraction efficiencies that can result in lower or higher yields of DNA. For example, both bacterial endospores and gram-positive bacteria are known to produce less DNA upon extraction compared to gram-negative bacteria [46]. These large differences due to cell wall structure can be alleviated using mechanical cell lysis techniques such as bead-beating [47]. Nonetheless, even with the inclusion of bead-beating steps, it has been suggested that extraction biases remain even between species in the same genera [6, 48, 49]. For example, despite using the same number of starting cells and accounting for genome size, Morgan et al. [48] found that two strains within *Lactococcus lactis* can exhibit different DNA extraction efficiencies. While cell wall architecture could account for many differences, it is not understood why closely related species would have different DNA extraction efficiencies. Without a better understanding of the factors underlying differential extraction efficiency, this step will remain highly biased. As such, it is recommended that future researchers identify the

extraction efficiencies of health-related microbes starting from known cell counts for those organisms.

Further work in this area has also shown that the use of different DNA extraction kits can result in different DNA extraction efficiencies for taxa depending on the kit used [50]. This can result in systemic biases toward the detection of specific bacteria in microbiome studies. Early work by Carrigg et al. and Krsek et al. using denaturing gradient gel electrophoresis showed that different DNA extraction protocols used on soil and sediment samples significantly altered the observed microbial profiles [51, 52]. Furthermore, work by Costea et al. [53] compared 21 different extraction protocols and found that, while taxon abundances correlated strongly between protocols, there were significant differences between them based on various beta diversity measurements. Overall, their work along with others has shown that while technical biases due to the use of different DNA extraction kits are smaller than interindividual variation, they still posed significant issues for downstream analysis [54, 55]. Based on these studies, it is highly recommended that researchers include a mechanical lysis step during DNA extraction and either use a standardized kit [48] or well-validated protocols such as those presented by the International Human Microbiome Standards project [53].

In addition, the amount of systemic bias introduced due to differences in DNA extraction protocols has been recently linked to the level of biomass within a sample. A recent study by Davis et al. [56] has shown that biases due to choice of DNA extraction kit are stronger in low-biomass samples compared to high-biomass samples. This critically shows that comparisons of low-biomass samples should be done using the same DNA extraction protocols and that extreme caution should be taken when comparing results between studies with different DNA extraction methods.

In fact, almost all systematic bias that is introduced during sample collection and processing is amplified when working with low-biomass samples such as breast milk [57], lung samplings [58], tumor tissues [59], or placental tissue [60]. This is due to factors such as contamination or sampling procedures having a larger relative impact on community composition than they would in high-biomass samples such as human stool. As highlighted by Great-house et al. [50], it is important to reduce all possible sources of contaminants during sample preparation and to show the existence of microbial life beyond sequencing. This could include approaches such as fluorescence in situ hybridization, microscopy, or culturing. Furthermore, it is recommended to use technical replicates during the sequencing of low-biomass samples to ensure the reproducibility of the sample composition between extractions and sequencing runs.

### Contaminant DNA

Contamination is a common issue that can be introduced during multiple steps of sequencing experiments including during sample collection, DNA extraction, and DNA sequencing. Sequencing experiments are particularly vulnerable to the introduction of contaminants due to the high sensitivity of DNA sequencing platforms. This is highlighted in their ability to sequence DNA in samples that appear DNA-free based on gel electrophoresis [61]. In a recent review of 265 microbiome studies by Hornung et al. [62], only 30% of studies were found to report using any type of negative control during sequencing experiments. Furthermore, it was unclear in many cases whether the negative control was only analyzed by gel electrophoresis or also subjected to DNA sequencing. This highlights the fact that many microbiome studies could be potentially biased due to the introduction of contaminants.

Fortunately, in many cases, such as during the analysis of fecal samples or oral washes, contamination will only account for a very small proportion of total biomass and as a result only create a minor amount of bias [63]. However, as mentioned above, contamination can be a major issue in low-biomass samples such as those collected from the human airway, brain, or placenta [63]. In some cases, such as for placental samples, the identification of microbes has been argued to be purely due to contamination itself [64–66]. This has led to studies on the “kitome” which represents the microbial inhabitants of commonly used lab reagents. These studies have revealed that while lab reagents are DNase free, they are often not DNA free [61]. Multiple different strategies have been implemented to remove contaminant DNA from lab reagents, including the use of restriction endonucleases [67, 68], UV irradiation [69], and ultrafine filtration [70]. In a recent report, Stinson et al. [71] compared low-biomass samples that were prepared using standard PCR master mixes or PCR master mixes treated with dsDNase. They found that treatment with dsDNase resulted in a significant reduction of contaminant bacterial reads within their samples. This indicates that the simple pre-treatment of PCR master mixes with dsDNase may result in improved microbiome analysis for low biomass samples. However, it still remains unclear whether any of these potential solutions can reliably remove all contaminant DNA [71] and as such most microbiome samples will contain some amount of contamination. As this remains true, it is critical for low-biomass samples to include extraction blanks that have been carried through the entire extraction and sequencing process in order to determine the level of contaminants and whether your samples contain enough reads to determine true biological effects above the “background” of the blanks.

While DNA extraction is arguably the most susceptible step to contamination issues, there is evidence that sequencing and library preparation can also be affected [72]. Index hopping has been identified as a potential source of contamination between samples during library preparation, especially on Illumina instruments that use ExAmp chemistry on patterned flow cells [73, 74]. Index hopping refers to DNA sequence barcodes being swapped between samples due to recombination events, which leads to sample cross-contamination. Furthermore, there is evidence of sample bleed-through during sequencing, which refers to DNA from previous sequencing runs being incorporated into later runs. Although we are in the process of comprehensively evaluating the effect of bleed-through, on the Illumina MiSeq, it is currently thought to contribute to 0.1% of total DNA reads per run [75].

### Post hoc contamination removal

Due to the significance of this issue, various bioinformatic techniques have been suggested as ways to identify potential contaminant sequences. In 2017, Edmonds and Williams [76] suggested that complete removal of OTUs found in negative/blank sequencing controls could be used to eliminate potential contaminant sequences. However, this is not always advisable as complete removal can lead to the loss of biologically relevant signals due to cross-contamination between samples and negative controls [77]. A second common approach for contaminant removal is the removal of sequences below a specified relative abundance or read count threshold. Nonetheless, this also does not work well in practice on low-biomass samples and in some cases will lead to the removal of rare features that are truly present [78].

In 2018, Davis et al. [78] presented a novel method, “decontam,” which attempts to identify levels of contamination rather than completely remove suspected contaminants. In this method, the authors take advantage of two different patterns of contamination through the measurement of DNA concentration levels and the sequencing of negative controls. The first pattern is that the abundance of contaminant sequences tends to be inversely proportional to the total DNA concentration in a sample, whereas true sequence abundance is independent of DNA concentration levels. This allows the authors to fit models of contamination to various sequences and remove contaminant abundances according to significant model fits. The second feature they take advantage of is the greater presence of contaminant sequences in negative controls than in real samples. While the authors show their tool is effective at controlling contamination in multiple different types of samples, their models still cannot account for cross-contamination between biological samples, nor the more significant problem of sequence bleed-through on Illumina machines.

Currently, to the best of our knowledge there are no bioinformatic methods that can deal with cross-contamination between samples in microbiome data. However, Minich et al. [72] have suggested a few steps that can be used to reduce its effect on microbiome analysis. These steps include that samples should be randomized across plates, and, when possible, single-tube extraction methods should be employed. Furthermore, low-biomass samples should not be prepared on the same plates as samples with high biomass.

Overall, while the use of reagents to remove microbial DNA may be promising, they have yet to be consistently shown as a valid approach to reduce contamination. However, researchers might think about opting for the use of a similar strategy when working with low-biomass samples. Furthermore, the removal of contaminants post hoc using bioinformatic software such as “decontam” should be included as part of the standard analysis of low-biomass samples, but will only provide minimal improvements to high biomass samples such as stool [78]. Finally, as previously mentioned, sample collection and processing should be done using aseptic techniques and when possible completed within biosafety cabinets.

#### DNA from non-viable cells

One of the final bias that can be introduced during DNA extraction is the inclusion of extracellular DNA found in biofilms and dead cells. DNA included from these sources can often result in the obstruction of phylogenetic signals and the over-identification of specific microorganisms [79]. This has been shown to be particularly problematic in the study of the airway microbiome in cystic fibrosis patients due to the biofilms formed by resident bacteria [80]. However, recent work on the gut microbiome of rabbits has shown that the inclusion of dead cells in fecal samples can result in significant biases and interfere with identifying the true live microbial community within a sample [81].

Various methods have been introduced to identify the living portion of microbial communities [79]. One proposed method to distinguish viable and non-viable cells takes advantage of the fact that RNA is only actively transcribed in living cells and is degraded quickly when exposed to factors outside of the cell. Therefore, by sequencing RNA (cDNA) instead of DNA, only genetic information from the living fraction of cells within a community will be captured. Nonetheless, this method suffers from multiple weaknesses, including the lower stability of RNA prevent the use of some samples (ex: stool), a higher detection rate of non-dormant microorganisms [82], and that rRNA transcription levels and cellular growth rates are not perfectly associated [83].

A second promising approach takes advantage of cell membrane integrity and the ability to pass specific

chemicals through the membranes of non-viable cells. While these techniques are commonly used in microscopy and flow cytometry, it has only recently been applied to DNA sequencing experiments [79]. Recent work by Weinmaier et al. [84] has shown that the use of propidium monoazide (PMA), a chemical that binds extracellular DNA and inhibits PCR amplification, can significantly improve the ability to identify DNA from live cells prior to sequencing. Since this initial work, multiple other studies have successfully taken advantage of PMA to characterize differences between the “live” microbiome and the total microbiome [85–87]. While this method is promising, it cannot accurately distinguish between viable and non-viable spores and archaea [84]. Further caution is also required when using this method as certain microbes such as strict anaerobes that are alive within the sampling environment upon collection will die off during sample processing. This can result in the loss of information from these groups of microbes, despite them being alive and present in the original sample.

#### Library preparation and sequencing

##### DNA amplification

Multiple factors can lead to biases at the PCR amplification step of microbiome studies. One of these factors is the inclusion of both inorganic ions and organic material that can interfere with PCR amplification [88]. The inclusion of these materials is usually due to inefficient DNA extraction steps in poorly optimized sampling environments. Interference by the inclusion of these inhibitors can lead to reduced PCR amplification and the inability to identify low abundance organisms within a sample.

One countermeasure to the reduced PCR amplification due to inhibitors is to increase the number of PCR cycles. Increasing PCR cycle numbers has been shown to significantly increase bacterial richness while not significantly impacting overall community structure due to the stronger detection of low abundant organisms [89]. However, this does come at the cost of reduced specificity. Work by Qiu et al. [90] has shown that increasing PCR cycle numbers directly increases the number of PCR products that contain sequences from two different DNA templates. Accumulation of these sequences, known as chimeras, is due to the creation of incompletely extended primers during the later cycles of PCR amplification that can act as primers for new DNA templates [91]. This issue is further exacerbated in marker gene sequencing studies due to the higher likelihood of homology between template DNA strands from the same or highly related species within a PCR [92].

Various strategies have been developed to reduce chimeric sequences in multiple template PCRs, with the



use of high-fidelity polymerases with proof-reading capability being one of the most promising [93]. In 2019, Sze and Schloss [94] examined bias introduced through the use of different polymerases and found that significant differences in error rate and chimeric sequence generation could be found between them. Interestingly, they found that after 30 PCR cycles AccuPrime polymerases had the lowest chimera rate (0.9%), but the highest error rate (0.124%), indicating a strong trade-off between lowering error rate and chimera formation. They found that the Q5, KAPA, and Phusion polymerases had the lowest error rates (~0.06%), although the KAPA polymerase's chimera rates were also the second lowest (2.3%) among polymerases tested. Based on these results, they recommended the use of KAPA polymerase for 16S rRNA gene sequencing studies. However, despite these differences they concluded that the choice of polymerase has little impact on the biological interpretations of community-wide measurements of diversity when using appropriate chimera filtering software such as UCHIME [95]. Similarly, in 2016, Gohl et al. [93] examined multiple polymerases and found that the use of high-fidelity polymerases resulted in fewer chimeric sequences. In summary, these two reports indicate that the choice of polymerase can affect both the error rate of sequences and the abundance of chimeras. To reduce these issues during PCR amplification, it has been suggested that researchers should use the highest possible fidelity polymerases and minimize the number of PCR cycles used [94].

Finally, one potential source of bias in PCR is the amount of template added to reactions. Although the overall goal would be to standardize the input DNA amounts across all community samples, achieving this goal is currently not possible for most microbiome samples. Due to the abovementioned extraction differences, copy-number variations between organisms differing between samples and, most importantly, the unquantified amounts of background “non-target” DNA that can be present in samples, it is quite difficult to assess the true “on-target” template amounts being added to PCRs. Given this, it is usually more productive to generate similarly strong amplification, regardless of template input differences, from as many samples as possible so that post-library normalization will generate as consistent sequencing between all samples as possible (i.e., will generate similar “sequencing effort”). In conclusion, researchers should keep a consistent protocol throughout their microbiome study and should report all PCR parameters within the methodology section of their report.

### Sequencing platform

The majority of DNA sequencing projects are now done on Illumina sequencing platforms, which generally show preferable DNA sequencing performance when compared to

other short read technologies such as Roche 454 sequencing or Ion torrent sequencing [7]. In 2016, D'Amore et al. [96] compared several sequencing platforms using 16S rRNA gene sequencing to determine the amount of bias introduced by the choice of DNA sequencing platform. They found that each DNA sequencer had its own unique bias profile that caused significant differences in the observed community make up that explained small levels of variation between microbial profiles [96]. This confirms earlier work by Salipante et al. [97] who compared the Illumina MiSeq and Ion Torrent Personal Genome Machine and found that overall the disparities between the two platforms were minor. Accordingly, these studies show that, while the choice of DNA sequencing platform can impact the observed community, the systemic bias introduced tends to be smaller than sample-to-sample variation.

Recent advances in long-read sequencing technologies have now enabled the sequencing of large DNA fragments. These advancements have enabled full-length marker gene studies and more comprehensive evaluation of metagenomes [98, 99]. However, historically long-read technologies have been plagued by significantly higher error rates (5–15%) [100, 101], particularly when compared to short-read technologies such as Illumina sequencing (0.01–3%) [96]. Nonetheless, recent work by Schloss et al. [102] and Johnson et al. [98] has shown that advancements in PacBio circular consensus sequencing have led to observed error rates as low as 0.027%. Unfortunately, this comes at a significant cost to sequencing throughput and other long-read sequencing platforms such as the MinION still have significantly lower accuracies than its short-read competitors. While the throughput of long-read platforms are still significantly lower than Illumina short-read technologies, the use of full-length marker genes has been shown to provide considerable advantages [98]. These advantages include the elimination of systemic biases between studies that have sequenced different marker gene regions and improvement on our ability to resolve lower levels of taxonomy.

It should also be noted that advancements in DNA sequencing technology have introduced the use of linked reads and proximity ligation strategies to produce high-quality synthetic long reads [103]. Using these methods, it is possible to produce higher quality genome assemblies; however, users should be aware of the potential biases these methods introduce due to the various enzymes used during library preparation [104]. As such, researchers should be aware that data produced through other library preparation methods may not be directly comparable.

### Primer choice during marker gene studies

Despite the recent advances in long-read technology, marker gene sequencing is still predominately performed

with high-throughput short-read technologies. These technologies, such as the Illumina MiSeq and HiSeq, are limited to short fragments ranging from ~150 to 550 bp. This has resulted in the creation of multiple different primer sets to target different variable regions on marker genes, such as the bacterial 16S rRNA gene [105–107]. Unfortunately, these primer sets differentially amplify different bacterial taxa, and no 16S rRNA gene primer set has been shown to equally amplify all bacteria. This is due to multiple reasons, including differing hybridization rates between taxonomic groups [7] and the presence of inhibitory flanking DNA [108], often caused by secondary structure and/or GC content differences. These unequal priming events lead to over- and under-detection of various bacterial groups, depending on the primer set used, leading to obstruction of the true underlying community composition and large taxonomic differences between studies that use different primer sets [109]. Numerous studies have compared the detection biases of various primers in different environments, and primer choice remains a debated topic within the field [20, 75, 106, 110, 111]. While large projects such as the Earth Microbiome Project recommend the usage of the V4 region, this choice does suffer from the underrepresentation of several important taxonomic groups, including *Actinobacteria* and *Propionibacterium*, while over-representing groups such as *Streptococcus*, *Treponema*, and *Prevotella* [75, 107]. It is highly recommended that researchers use a primer set that is either directly comparable to previous literature or can detect the groups of bacteria they are interested in studying.

Fortunately, as mentioned previously, improvements in DNA sequencing technologies have now begun to enable higher throughput full-length marker gene sequencing. This has the promise to reduce biases between studies that use primer sets that target different regions on marker genes. However, the primers used in full length studies will still have different priming rates between taxonomic groups.

### Metagenomic library preparation

Currently, the majority of metagenomic shotgun sequencing projects rely on high-throughput short-read Illumina DNA sequencing [112]. Several different library construction methods have been developed for Illumina sequencers, which generally consist of three to four different stages: DNA fragmentation, optional repairing of DNA fragments, ligation of platform specific adaptors, and optional library PCR amplification [113]. Multiple different strategies have been used to accomplish these tasks including enzymatic digestion, mechanical shearing, and the use of transposons. In 2019, Sato et al. [113] compared the biases introduced by using different library preparation kits by examining both mock

communities and isolated bacteria. They found that library preparation using the now-discontinued Illumina Nextera XT kit suffered from strong biases toward high-GC content regions. Examination of mock communities prepared using the XT kit resulted in the under-observation of low-GC content species such as *Staphylococcus aureus*, *Brachyspira pilosicoli*, and *Streptobacillus moniliformis*. Further investigation also revealed that minor biases toward high-GC content were also present in library preparations using mechanical fragmentation, likely due to non-random DNA shearing during sonication methods [114]; however, these biases were smaller than in the case of XT kits. Interestingly, the other library preparation kits they tested, including the new Nextera Flex (replacement for XT) Illumina transposon-based method, showed relatively little bias when compared to one another. Overall, this indicates that the observed microbial community can be systemically biased when using library preparation methods such as the Illumina Nextera XT kit or sonication-based mechanical fragmentation.

### Marker gene bioinformatics

After sequencing has been accomplished, a significant amount of bias can be introduced during downstream bioinformatic analysis [55]. A typical analysis begins with quality filtering to remove reads with ambiguous or error-prone base calls. Despite this step removing a significant amount of poor-quality data, it has been shown to only introduce a relatively small amount of bias directed toward the detection of low-abundance microorganisms [8]. After quality filtering, reads must be assigned to an analytical unit using either error-correction denoising algorithms or operational taxonomic unit (OTU) picking.

### Taxonomic units

As both sequencing and PCR can introduce base errors during sequencing experiments, it is not recommended to analyze sequences directly. Instead, users can choose to either cluster sequences together into OTUs or correct sequencing errors using denoising algorithms. The choice to use denoising algorithms or OTU-picking strategies has been shown to create systematic bias in marker gene sequencing studies [8, 115]. Within each of these distinct categories, further systemic bias can be introduced based on the implementation of the chosen denoising or OTU-picking algorithm [7]. In the case of OTU picking, there are three major methods for defining OTUs: reference-based, de novo, and open-reference clustering. During reference-based clustering, sequences are compared to known marker genes within a database and clustered at a specific percent identity cutoff [116]. This differs from de novo clustering strategies as these

methods do not rely on reference databases, but rather cluster all sequences within a study based on pairwise nucleotide distances [117]. Finally, open-reference clustering uses both strategies by first performing reference-based OTU picking and then clustering the remaining sequences in a de novo manner [118]. Overall, the clustering of sequences in these manners allows researchers to mitigate the impact of sequencing errors on any one read.

Currently there is mixed evidence on which strategy is best when attempting to define OTUs [7]. While de novo methods have been shown to create higher quality OTU classifications [117], they often suffer from being unstable both within studies [119] and between studies [120]. Furthermore, while the use of different strategies can result in significantly different richness counts [117], they often resulted in similar beta-diversity metrics, indicating that, while biases are present, they are often smaller than sample-to-sample variation [121]. Overall, it currently remains unclear which clustering strategy reveals observations closest to the true community, although recent work by Edgar [122] has shown that, regardless of clustering strategy, the use of OTUs often inflates the diversity and richness present within a microbial community.

To help counteract these issues, various denoising algorithms have been created to resolve single-nucleotide accuracy while also generating sequences that can be compared between studies [120]. Currently, there are several denoising algorithms that generate amplicon sequence variants (ASVs), including DADA2 [123], Deblur [124], and UNOISE [125]. While the goal of these three tools is similar, they achieve it through different mechanisms. Deblur first filters reads for possible artifacts through the alignment of reads against sequence artifact databases and the removal of reads that do not map to the 88% OTU Greengenes database. It then aligns all remaining sequences against one another and uses information on their abundances and the error rate of Illumina sequencers to remove sequences derived from errors. UNOISE uses a similar approach that first clusters sequences together and then uses pre-set parameters to remove error-derived sequences. Unlike the other two methods, DADA2 generates a parametric error model for each sequencing run and then uses quality score information to remove sequencing errors.

These algorithms have been compared to standard OTU-clustering methods and in many cases have shown preferable results. Nearing et al. [8] found that while open reference-based OTU clustering resulted in the inflated richness of mock communities, all three denoising algorithms had richness scores comparable to the expected number of biological sequences. They also found that all three denoising algorithms resulted in

comparable biological conclusions when examining metrics that were weighted by abundance. Subsequently, Caruso et al. [115] have shown that denoising algorithms outperform conventional de novo OTU-clustering strategies on several different mock communities. It should be noted that denoising tools have been developed using mock communities and as such their pre-set parameters and modeling behaviors may be overfit toward simple microbial communities. This makes it extremely important to be consistent in your choice of algorithm when comparing data within and across studies. However, overall, the data does suggest that denoising algorithms may be a preferable option for analytic unit assignment in future marker gene studies.

Regardless of the advantages and disadvantages of each algorithm, it is important to record which strategy was used to assign analytical units during analysis. Due to the intrinsic nature of each algorithm, they all generate their own systemic biases which become especially apparent during the examination of rare taxonomic groups [8].

#### **Taxonomic classification**

Taxonomic classification of ASVs or OTUs poses a second challenge in the bioinformatic analysis of microbiome sequencing data. There have been multiple different strategies used to assign taxonomy to ASVs or OTUs throughout the years [126]. Common strategies include the use of kmer-based classifiers, such as naïve Bayes or Kraken2 [127, 128], local alignment search tools [129], or multiple global alignment strategies [130]. Each of these classifiers have been shown to have varying accuracy scores depending on the sample that is being analyzed [126, 128, 131] and introduce systemic biases in their ability to resolve the classification of novel taxonomic groups [131]. Furthermore, each of these methods is dependent on the use of reference databases that can result in additional differences between studies.

Currently, there are four commonly used marker gene databases for microbiome studies: Greengenes [132], SILVA [133], Ribosomal Database Project (RDP) [134], and NCBI [135, 136]. Interestingly, each of these databases use their own method for taxonomic classification assignment. While all four use seven main ranks: domain, phylum, class, order, family, and genus, both the SILVA and RDP databases also contain intermediate ranks. Furthermore, unlike the other three databases that use manually curated systematics to assign taxonomy, Greengenes taxonomy is assigned based on de novo 16S rRNA gene tree construction [137]. While work has shown that Greengenes, SILVA, and RDP can be mapped onto NCBI taxonomy with only low levels of dissimilarity, there are many instances where Greengenes, SILVA, and RDP cannot be mapped reliably to one another [137]. This distinction is important as slight

differences in taxonomic assignments between studies can lead to substantial differences in interpretation of the biological conclusions from studies.

### Copy number correction

Unfortunately, not all marker genes that are studied are present as single copies within microorganisms. The commonly sequenced 16S rRNA gene is present in variable numbers among both bacteria and archaea [138]. While it has been shown that 16S rRNA gene copy number rarely varies within species, it often increases in variation with taxonomic distance [139]. Furthermore, due to these differences, 16S rRNA gene sequencing experiments can be biased toward the detection of bacteria with higher 16S rRNA copy numbers [140]. While this bias will have minimal impact on between-study comparisons, they can severely distort the observed composition from the underlying truth [141].

Work by both Kembel et al. [140] and Angly et al. [142] has found that 16S rRNA gene copy number variation exhibits strong phylogenetic signals. This has led to the development of multiple different phylogenetic methods for correcting 16S rRNA gene copy number variation within microbiome studies. Louca et al. [141] recently compared three different tools (PICRUSt [143], CopyRighter [142], and PAPRICA [144]) in their ability to correctly predict the number of 16S rRNA gene copies within a wide range of both bacterial and archaeal clades. In their analysis, they found that independently of the tool used, 16S rRNA gene copy number could only be accurately predicted for a small number of tested genomes that had low levels of sequence divergence from reference 16S rRNA genes [141]. This indicates that while copy number correction methods could be useful in well-studied environments, they generally introduce more noise than they correct. Fortunately, like most bioinformatic methods that rely on reference databases, their accuracy may increase as new genomes are sequenced.

### Metagenomic shotgun bioinformatics

Unlike marker-gene sequencing, metagenomic shotgun sequencing involves the sequencing of all the DNA contained within a sample. This can significantly complicate bioinformatic workflows, and, as a result, two main approaches for data analysis are available: referenced-based metagenomics and de novo metagenomic assembly. Each of these methods have their own benefits and detriments and often analysis with one method can complement the other.

#### Reference based

Reference-based metagenomic shotgun sequencing strategies generally involve three different approaches by comparing sequenced DNA to databases containing reference genes, marker genes, or translated protein sequences [145].

Multiple tools exist for each strategy; however, they often are either comparing full-length sequences, marker gene sequences, or kmer sequence composition to reference databases to achieve taxonomic and functional assignment [146]. McIntyre et al. [147] compared 11 different metagenomic classifiers and found that the number of species identified within the same sample could differ by three orders of magnitude depending on tool choice. However, they found that the overlap between the true taxa within the sample and most abundant species identified was relatively high, indicating that filtering of low-proportion species could improve accuracy and reduce the number of false positives. Furthermore, they found that the most commonly identified false positives belong to the phyla Proteobacteria, Firmicutes, or Actinobacteria. More recent work by Ye et al. [145] compared 20 different metagenomic classifiers and found that all tools had relatively high area under the curve for precision and recall with the exception of the phylogenetic marker-gene-based tools MetaPhlan2 [148] and mOTUs2 [149]. The authors also found that DNA-based strategies, in particular Kraken [150] and taxMaps [151], tended to have abundance estimates that were closer to the underlying truth than other strategies. This is most likely due to protein classifiers lacking untranslated regions within their database and marker-gene strategies lacking marker genes for specific taxonomic lineages. However, it was also found that DNA strategies tended to suffer from large numbers of low-abundance false positives while marker gene strategies did not [145]. Overall, these studies indicate that the use of DNA-based strategies with an appropriate abundance filter, such as Kraken and its related tools Kraken2 [152] and Bracken [153], is recommended due to their favorable performance metrics. However, it should be noted that a combined approach using marker gene-based classifiers may be required when researchers are interested in studying low-abundance taxa.

Finally, it was shown that all classifiers underperformed as the number of poorly described taxonomic groups increased within a sample [145]. This issue has become especially apparent within the study of the human virome, where sequence databases cover only a small amount of total viral diversity [154, 155]. This issue highlights one of the largest biases in reference-based approaches as they are only able to classify functions and taxonomic groups within the target reference database. This limitation suggests that additional analysis on poorly described samples using de novo assembly could significantly improve insights into the observed underlying microbial composition.

#### Metagenomic assembly

Unlike reference-based metagenomics, de novo assembly methods do not rely on the use of existing reference

sequences. Instead, they take advantage of sequence overlaps between DNA reads to generate longer predicted sequences known as contigs and scaffolds [112]. This process is known as assembly and can be accomplished by multiple different programs including the popular tools MEGAHIT [156] and metaSPAdes [157]. The recovery of complete genomes from assembly is often not possible due to their inability to resolve highly repetitive regions [158] and regions of similarity between different genomes [112]. This has led to the development of a second step known as binning. Binning takes advantage of multiple features, including the co-abundance and coverage of contigs between samples [159] and the grouping of contigs with similar kmer frequencies and GC content [160, 161]. This allows contigs to be grouped into bins which represent groups of contigs likely co-occurring within the same genome.

This process has enabled the recent discovery of thousands of novel metagenome-assembled genomes in human microbiome samples [162–164]. However, while assembly methods have a lower dependency on reference databases, they are biased in multiple other aspects. Due to the complex composition of microbiome samples, assemblers are often not able to determine the relationship between a large portion of reads within a sample [112]. This can result in the under classification of functional genes such as rRNA genes that are often found in repetitive regions [158]. Vollmers et al. [165] recently compared various tools for assembling metagenomes and found significant biases in their ability to identify both high- and low-abundance organisms. Similarly, work by Yue et al. [166] compared various binning strategies and found that tools output differing results, with DASTool, MetaWrap, and MetaBat2 showing the strongest results. This indicates that, like reference-based approaches, the choice of bioinformatic software could have significant impacts on the observed results.

Finally, like reference-based metagenomics, the annotation of assembled bins and genomes is dependent on known reference databases. Several reference databases exist for this purpose including NCBI RefSeq [167] and the Genome Taxonomy Database [168]. Currently, there is debate within the field as to which databases provide the most accurate representation of the underlying data. This can result in biased annotations of genes or taxa that are overrepresented within a database which can considerably bias the information obtained during analysis [169]. Despite these drawbacks, assembly based methods offer a complimentary strategy to reference-based approaches that offer the ability to identify novel microorganisms within a sample.

## Abundance correction methods

### Bias adjustment using mock communities

In 2015, Brooks et al. [45] presented a novel protocol for identifying and correcting biases for specific taxa within a microbial community using mock community mixtures. Using 80 different mock communities of specific taxa, they were able to predict the true abundance of those taxa within sequenced samples of unknown composition. Based on their model, they found that biases could introduce abundance error rates as high as 85%, indicating the importance of incorporating them during downstream analysis. Furthermore, they found that the predicted true abundances of specific bacteria within clinical samples better reflected the physiology and diagnosis of patients. For example, they found that the predicted abundance of *G. vaginalis* using their modeling method correlated better with pH levels than the originally observed proportions.

Further work in later years by Krehenwinkel et al. [170] and Bell et al. [171] involved the use of simple linear regression models of observed proportions against actual abundances to estimate experimental biases. Like previous work, they found that each taxon within a sample had different levels of observed biases. While they were able to show strong model fits for many taxa, these studies suffered from their inability to consider the interactions between biases within a single sample, which is critical when modeling compositions using proportions [6].

Fortunately, recent work by McLaren et al. [6] has shown the promising ability to adjust biases introduced throughout microbiome sequencing experiments using a mathematical model based on the ratio of observed reads between different taxa. Within this model, they show that, unlike proportions, biases that impact ratios between taxa are independent of sample composition. This critical finding helps address previous issues with models that failed to consider interactions between various biases within samples. The authors then go on to show that, using a simple mathematical model based on taxon-specific multiplicative biases and the ratios between taxa, they could predict proportions that closely matched observed results from mock communities. Furthermore, they showed that the use of mock community calibration controls can help adjust observed proportions to their true values in samples with unknown composition.

Unfortunately, all these methods for bias adjustment rely on the ability to use mock communities as calibration controls. While this may be feasible for a small number of bacteria as suggested by Brooks et al. [45], it is currently not feasible for the majority of microbiome sequence-based experiments. With the inability to culture a significant proportion of microbes, many projects will suffer from the inability to create cell-based mock

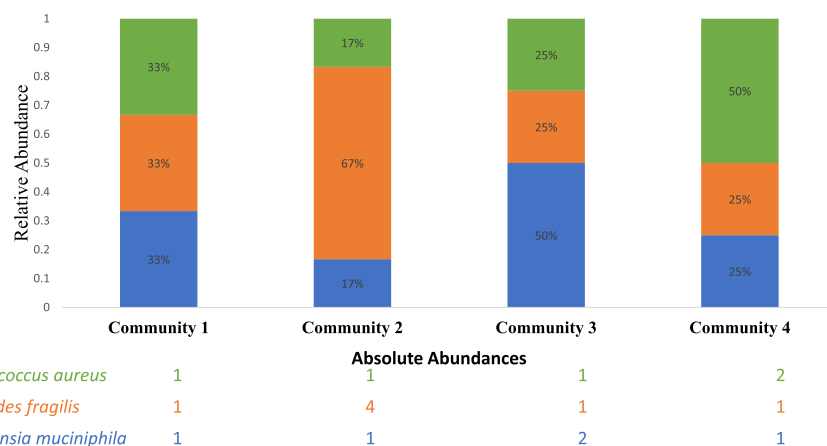
communities [172]. This will result in the inability to measure DNA extraction biases which have been shown to play a significant role in microbiome studies. Furthermore, due to the resolution of many 16S rRNA gene sequences, OTUs or ASVs of interest may not currently correspond with any known living organisms. This would again create barriers in the ability to collect and create mock communities. Despite these drawbacks, there are many instances where mock community construction is possible, as shown by Brooks et al. [45], and these methods should be considered when plausible. Furthermore, a publicly available database of extraction efficiencies of different microbes using different extraction protocols would substantially improve the usability of these methods in the future.

**Determining absolute microbial abundances**

Unlike the use of culturing, qPCR, or flow cytometry, DNA sequence-based microbiome studies currently lack the ability to determine absolute microbial abundances. This is due to the compositional nature of DNA sequencing data where one read does not correspond to one cellular unit [173]. This can often obscure the analysis of microbiome data, as the increased relative abundance of one taxon will always result in the decreased relative abundance of all other taxa in a sample, even if they have not actually changed (Fig. 2). In 2017, Vandeputte et al. [174] combined 16S rRNA gene sequencing data with flow cytometry and qPCR data to generate quantitative microbiome profiles (QMP). While this method allowed the authors to generate new insights into the microbial etiology of Crohn’s disease, it suffers from multiple different downfalls. To generate QMP data

from 16S rRNA gene sequencing data, the number of 16S rRNA gene copies in each cell must be determined. Unfortunately, this process is unreliable for most bacterial and archaeal genomes [141]. Furthermore, QMPs cannot consider multiple different biases that are introduced throughout sequencing experiments that disrupt the relationship between cellular counts and sequence abundance [6, 45]. This includes differing DNA extraction efficiencies and PCR amplification efficiencies between different taxonomic groups. Decades of work in “classical” environmental microbiology have also shown that it is not a trivial matter to obtain accurate cell counts for many samples. There are also the considerations of live vs. dead cells included in counts (or the accuracy of live-dead stains, if used) and the potentially variable biases (depending on sample type) introduced by the sample preparations required to enable microscopy or flow-cytometry counting. Even if accurate counts could finally be obtained, the end-point nature of current high-throughput sequencing PCRs means that resolved abundances are “semi-quantitative” at best, as they are not based upon the same technique as qPCR.

Tkacz et al. [175] presented a second method in 2018 for determining absolute microbial abundances from DNA sequencing data. They proposed the use of synthetic chimeric DNA with known quantities as controls during marker gene sequencing experiments. Furthermore, they suggest that these chimeric DNA sequences can be modeled after different taxonomic sequence signatures to control for amplification biases. This provides a significant improvement over the previous attempts; however, it still suffers from



**Fig. 2** Microbiome data is compositional and does not represent absolute counts. As a consequence of the arbitrary total amount of sequencing reads a DNA sequencing platform can generate, the resulting data is compositional. This means that one DNA sequence does not correspond to one cellular unit, and an increase in relative abundance from one group can result in the decrease of relative abundance of another. This can occur even when the absolute abundance of a group does not change

differences in marker gene copy numbers and differing DNA extraction efficiencies.

## Methodology standardization and reporting

### Consistent study protocols

A significant portion of the bias between microbiome studies is introduced through the use of different protocols [53, 176]. This has led to attempts to standardize protocols within the microbiome field to allow for improved inter-study comparisons. In 2010, the Earth Microbiome Project [176] was initiated as an attempt to collect and sequence microbiome samples from around the world. With the initiation of this project, several standardized protocols were developed including the acquisition of metadata, DNA extraction, and both marker-gene and shotgun-sequencing library preparation. Following this, the International Human Microbiome Standards group compared a number of different protocols on human fecal samples and presented recommended protocols based on reproducibility and accuracy of community diversity [53]. While following these recommended protocols would reduce biases between studies, not all protocols are appropriate for the diverse number of biological questions at hand [6]. For example, some protocols may need to be altered if the researchers are interested in either the live or total microbial community contained within a sample. Furthermore, the use of standardized protocols still does not solve the systemic biases introduced during sequencing experiments that led to distorted observations of the true underlying microbial community.

Despite these issues, it is still recommended that whenever possible, protocols between closely-related studies should be replicated to reduce the amount of bias introduced between each study. As previously mentioned, when possible, manufactured kits or the use of well-validated protocols such as those outlined by the International Human Microbiome Standards project or Earth Microbiome project should be used. Furthermore, the use of classification tools such as random-forest modeling or neural networks should be validated on data from more than one study group to ensure the reliability of their model on real world data. Overall, when interpreting results, researchers should remember that the observed communities are only one representation of the true underlying abundances contained within the sample.

### Methodology reporting

While there are several promising approaches to help reduce the amount of systematic bias within microbiome studies, the best way to help between study comparisons is the creation of a comprehensive methodology section. Not only will this allow researchers to replicate your

results but will also allow them to help identify why results may not be reproducible from one study to another. A number of studies including projects such as the Earth Microbiome project and more recently the STORMS microbiome project [177] have addressed this issue. As such, we will only briefly outline in the section below as well as in Table 2 key information that should be present in the methodology of all microbiome studies.

All sample collection procedures should be included within the methods section. This includes the collection method and sterile techniques used during sample collection, as well as the media used to store the samples. Furthermore, authors are also recommended to give the general length of time samples remained in storage before processing began. One of the most critical pieces of information included in the methods section should be the DNA extraction method used during sample preparation. As highlighted, above DNA extraction is one of the largest factors that introduces systematic bias within a study. If a DNA extraction kit is used, author must report the manufacturer and product name, along with any modification or optional steps that the authors used, such as the addition of proteinase K. Within this section, it is also recommended that the authors report all steps taken to reduce any possible sources of contamination, such as the use of biological safety hoods and inclusion of blanks (critical for low-biomass samples) during DNA extraction.

Depending on the type of DNA sequencing used to investigate the samples of interest, the exact reporting parameters will differ between marker gene or metagenomic shotgun sequencing. Researchers should report all parameters used during PCR amplification, including the number of cycles and the type of polymerase used. The authors should also report the primer names and sequences if custom primers are used, during marker gene sequencing. If a commercial metagenomic library preparation kit was used, the authors should report the manufacturer as well as the product name and any modifications or optional steps used during sample processing. The amounts of template and the method of library normalization should be reported within the methods for both marker gene and metagenomic shotgun sequencing studies. The DNA sequencing instruments, along with which consumable versions used during sequencing, should be reported.

No matter what approach is used to investigate your samples, the bioinformatics method section of a manuscript should include the names and versions of all bioinformatic tools used within the workflow. This includes any databases that were used for classification throughout the study and all non-default parameters used. Any programs used for statistical analysis should be reported,

**Table 2** Key information to report within sequenced-based microbiome manuscripts

Processing step	Parameters
Sample collection	<ul style="list-style-type: none"> <li>- Description of samples of interest</li> <li>- Collection location</li> <li>- Collection device</li> <li>- Number of different collection personnel</li> <li>- Sterile techniques used</li> <li>- Use of technical replicates</li> </ul>
Sample storage	<ul style="list-style-type: none"> <li>- Media used to store samples</li> <li>- Length of storage and storage conditions</li> <li>- Length of time spent unpreserved</li> </ul>
DNA extraction	<ul style="list-style-type: none"> <li>- DNA extraction method</li> <li>- If using manufacturer kit:               <ul style="list-style-type: none"> <li>Product name</li> <li>Optional steps used</li> </ul> </li> <li>- Any methods used to reduce or remove possible contaminants</li> </ul>
Library preparation	<ul style="list-style-type: none"> <li>- PCR parameters including:               <ul style="list-style-type: none"> <li>Polymerase</li> <li>Cycle number</li> <li>Thermal profile</li> </ul> </li> <li>- Marker gene studies:               <ul style="list-style-type: none"> <li>Primer names and gene target(s)</li> </ul> </li> <li>- Metagenomic shotgun studies:               <ul style="list-style-type: none"> <li>Library preparation kit name/protocol</li> <li>Detailed information about:                   <ul style="list-style-type: none"> <li>▪ DNA fragmentation</li> <li>▪ DNA fragment repair</li> <li>▪ Ligation</li> <li>▪ PCR amplification</li> </ul> </li> </ul> </li> <li>- Normalization method</li> <li>- Amount of starting template used</li> </ul>
Sequencing	<ul style="list-style-type: none"> <li>- Sequencing platform</li> <li>- Consumable product names</li> <li>- Whether demultiplexing or any other sequence pre-processing on-instrument was done</li> </ul>
Bioinformatics	<ul style="list-style-type: none"> <li>- All tools used to process raw data along with:               <ul style="list-style-type: none"> <li>Goal of tool</li> <li>Version number</li> <li>Non-default parameters</li> </ul> </li> <li>- Any database names along with their version numbers</li> <li>- Any statistical tests used to analysis data</li> </ul>

along with any hypothesis tests that were used to investigate the samples of interest. Finally, all code along with its documentation and comments should be deposited in a public repository (e.g., GitHub) and linked within the methodology section of its accompanying study.

### Closing remarks

The study of the human microbiome has clearly created a better understanding of the relationship between humans and microbes. However, multiple different experimental steps have led to the inclusion of significant biases within the human microbiome literature. In many cases, this has led to the inability to compare results from one study to another. The use of consistent protocols on similar sample types could help address issues of between-study bias [7]. This will, however, require the development of further environment-specific protocols, as not every sample is amendable to currently recommended solutions. Furthermore, authors should report the key parameters outlined within this review in all

microbiome-based sequence studies to help ensure the reproducibility of their results.

While reporting these key parameters will help with addressing between-study biases, it will not help address within-study biases that can lead to distorted observations of the true underlying microbial composition. These issues generally stem from multiple areas, including the inability to quantify extraction efficiencies between different microbes and different DNA amplification rates. Without solving these issues, the true view of the absolute abundances of microbes within the human microbiome will not be possible. This is critically important to understand when evaluating methods that rely on qPCR, flow cytometry, or DNA spike-in methods. Although recent work in the field of mathematical modeling shows the promising ability to adjust most biases within a microbiome experiment, they currently rely on the ability to create mock communities [6]. Unfortunately, as stated previously, this is currently not possible for many microbes. Despite this, we suggest



that future work should attempt to quantify DNA extraction efficiencies for various protocols and health-related microbes as it is currently one of the largest areas of bias within microbiome studies.

Overall, it is currently recommended that all methods in a study should be kept consistent and, when comparing to previous work, the same protocols should be followed. If this is not possible, protocols should be benchmarked between studies to identify whether consistent patterns of bias can be found between them. This will give researchers additional information that can be used when comparing results between studies. Finally, all biological conclusions generated from sequenced-based microbiome experiments should be validated through other molecular techniques to ensure that the underlying results are robust to biases introduced during microbiome sequencing experiments.

#### Abbreviations

MGS: Metagenomic shotgun sequencing; OTU: Operational taxonomic unit; ASVs: Amplicon sequence variants; PMA: Propidium monoazide; RDP: Ribosomal database project; QMP: Quantitative microbiome profile; PCR: Polymerase chain reaction; qPCR: Quantitative polymerase chain reaction

#### Acknowledgements

This manuscript was original written as part of J.T.N.'s PhD comprehensive exam. We would like to thank all members of the examining committee for their discussions and suggestions around this article: Joseph Bielawski, Brent Johnston, John Rhode, Ketan Kulkarni, Zhenyu Cheng, and Nikhil Thomas. We would also like to thank Gavin Douglas and Vanessa DeClercq for providing feedback on the manuscript.

#### Authors' contributions

J.T.N. and M.G.I.L. contributed to the conception and design of the review. J.T.N. performed literature review for the manuscript, drafted the manuscript, and created the figures within the manuscript. A.M.C. and M.G.I.L. provided critical feedback and suggestions to the manuscript. All authors approved the final version of the manuscript.

#### Funding

J.T.N. is supported by a researchNS Scotia Scholars award and a Nova Scotia Graduate Student Scholarship. M.G.I.L. is supported by an NSERC Discovery Grant.

#### Availability of data and materials

Not applicable.

#### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Department of Microbiology and Immunology, Dalhousie University, Halifax, Nova Scotia, Canada. <sup>2</sup>Integrated Microbiome Resource, Dalhousie University, Halifax, Nova Scotia, Canada. <sup>3</sup>Department of Pharmacology, Dalhousie University, Halifax, Nova Scotia, Canada.

Received: 11 July 2020 Accepted: 24 March 2021

Published online: 18 May 2021

#### References

- Pflughoeft KJ, Versalovic J. Human microbiome in health and disease. *Annu Rev Pathol Mech Dis.* 2012;7:99–122 Annual Reviews.
- Valdes AM, Walter J, Segal E, Spector TD. Role of the gut microbiota in nutrition and health. *BMJ.* 2018;361:36–44 BMJ Publishing Group.
- Wong AC, Levy M. New approaches to microbiome-based therapies. *mSystems.* American Society for Microbiology. 2019;4:e00122–19.
- Schlager R. Microbiome diagnostics. *Clin Chem.* 2020;66:68–76 American Association for Clinical Chemistry Inc. [cited 2020 Feb 18]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/31843867>.
- Schloss PD. Identifying and overcoming threats to reproducibility, replicability, robustness, and generalizability in microbiome research. *MBio.* American Society for Microbiology. 2018;9:e00525–18.
- McLaren MR, Willis AD, Callahan BJ. Consistent and correctable bias in metagenomic sequencing experiments. *Elife.* eLife Sciences Publications Ltd. 2019;8:e46923.
- Pollock J, Glendinning L, Wisedchanwet T, Watson M. The madness of microbiome: attempting to find consensus “best practice” for 16S microbiome studies. *Appl Environ Microbiol.* 2018; American Society for Microbiology. [cited 2020 Feb 18]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29427429>.
- Nearing JT, Douglas GM, Comeau AM, Langille MGI. Denoising the denoisers: an independent evaluation of microbiome sequence error-correction approaches. *PeerJ.* PeerJ Inc. 2018;2018:e5364.
- Araújo-Pérez F, McCoy AN, Okechukwu C, Carroll IM, Smith KM, Jeremiah K, et al. Differences in microbial signatures between rectal mucosal biopsies and rectal swabs. *Gut Microbes.* 2012;3:530–5 Taylor & Francis. [cited 2020 Feb 18]. Available from: <http://www.tandfonline.com/doi/abs/10.4161/gmic.22157>.
- Durbán A, Abellán JJ, Jiménez-Hernández N, Ponce M, Ponce J, Sala T, et al. Assessing gut microbial diversity from feces and rectal mucosa. *Microb Ecol.* Springer. 2011;61:123–33.
- Zoetendal EG, von Wright A, Vilponen-Salmela T, Ben-Amor K, Akkermans ADL, de Vos WM. Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. *Appl Environ Microbiol.* 2002;68:3401 LP–3407 Available from: <http://aem.asm.org/content/68/7/3401.abstract>.
- Codling C, O'Mahony L, Shanahan F, Quigley EMM, Marchesi JR. A molecular analysis of fecal and mucosal bacterial communities in irritable bowel syndrome. *Dig Dis Sci.* Available from. 2010;55:392–7 <https://doi.org/10.1007/s10620-009-0934-x>.
- Tottey W, Feria-Gervasio D, Gaci N, Laillet B, Pujos E, Martin JF, et al. Colonic transit time is a driven force of the gut microbiota composition and metabolism: in vitro evidence. *J Neurogastroenterol Motil.* 2017;23(1):124–34. <https://doi.org/10.5056/jnm16042>.
- Vandeputte D, Falony G, Vieira-Silva S, Tito RY, Joossens M, Raes J. Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. *Gut.* BMJ Publishing Group. 2016;65:57–62.
- Bassis CM, Moore NM, Lolans K, Seekatz AM, Weinstein RA, Young VB, et al. Comparison of stool versus rectal swab samples and storage conditions on bacterial community profiles. *BMC Microbiol.* 2017;17:78 BioMed Central Ltd. [cited 2020 Feb 18]. Available from: <http://bmcmicrobiol.biomedcentral.com/articles/10.1186/s12866-017-0983-9>.
- Reyman M, van Houten MA, Arp K, Sanders EAM, Bogaert D. Rectal swabs are a reliable proxy for faecal samples in infant gut microbiota research based on 16S-rRNA sequencing. *Sci Rep.* 2019;9:16072 Available from: <https://doi.org/10.1038/s41598-019-52549-z>.
- Jones RB, Zhu X, Moan E, Murff HJ, Ness RM, Seidner DL, et al. Inter-niche and inter-individual variation in gut microbial community assessment using stool, rectal swab, and mucosal samples. *Sci Rep.* 2018;8:4139 Available from: <https://doi.org/10.1038/s41598-018-22408-4>.
- Hall MW, Singh N, Ng KF, Lam DK, Goldberg MB, Tenenbaum HC, et al. Inter-personal diversity and temporal dynamics of dental, tongue, and salivary microbiota in the healthy oral cavity. *npj Biofilms Microbiomes.* 2017;3:2 Available from: <https://doi.org/10.1038/s41522-016-0011-0>.
- Proctor DM, Fukuyama JA, Loomer PM, Armitage GC, Lee SA, Davis NM, et al. A spatial gradient of bacterial diversity in the human oral cavity

- shaped by salivary flow. *Nat Commun.* 2018;9:681 Available from: <https://doi.org/10.1038/s41467-018-02900-1>.
20. Fan X, Peters BA, Min D, Ahn J, Hayes RB. Comparison of the oral microbiome in mouthwash and whole saliva samples. Arora PK, editor. *PLoS One.* 2018;13:e0194729 Public Library of Science. [cited 2020 Feb 18]. Available from: <https://dx.plos.org/10.1371/journal.pone.0194729>.
  21. Jo R, Nishimoto Y, Umezawa K, Yama K, Aita Y, Ichiba Y, et al. Comparison of oral microbiome profiles in stimulated and unstimulated saliva, tongue, and mouth-rinsed water. *Sci Rep. Nature Publishing Group.* 2019;9:1–7.
  22. Luo T, Srinivasan U, Ramadugu K, Shedden KA, Neiswanger K, Trumble E, et al. Effects of specimen collection methodologies and storage conditions on the short-term stability of oral microbiome taxonomy. Liu S-J, editor. *Appl Environ Microbiol.* 2016;82:5519 LP–5529 Available from: <http://aem.asm.org/content/82/18/5519.abstract>.
  23. Chng KR, Tay ASL, Li C, Ng AHQ, Wang J, Suri BK, et al. Whole metagenome profiling reveals skin microbiome-dependent susceptibility to atopic dermatitis flare. *Nat Microbiol.* 2016;1:16106 Available from: <https://doi.org/10.1038/nmicrobiol.2016.106>.
  24. Ogai K, Nagase S, Mukai K, Iuchi T, Mori Y, Matsue M, et al. A comparison of techniques for collecting skin microbiome samples: swabbing versus tape-stripping. *Front Microbiol.* 2018;2362 Available from: <https://www.frontiersin.org/article/10.3389/fmicb.2018.02362>.
  25. Bjerre RD, Hugerth LW, Boulund F, Seifert M, Johansen JD, Engstrand L. Effects of sampling strategy and DNA extraction on human skin microbiome investigations. *Sci Rep.* 2019;9:17287 Available from: <https://doi.org/10.1038/s41598-019-53599-z>.
  26. Glendinning L, Wright S, Tennant P, Gill AC, Collie D, McLachlan G. Microbiota in exhaled breath condensate and the lung. *Appl Environ Microbiol. American Society for Microbiology.* 2017;83:e00515–17.
  27. Willerslev E, Hansen AJ, Poinar HN. Isolation of nucleic acids and cultures from fossil ice and permafrost. *Trends Ecol Evol.* 2004;19:141–7 Available from: <http://www.sciencedirect.com/science/article/pii/S0169534703003574>.
  28. Knight R, Urbanac A, Taylor BC, Aksenov A, Callewaert C, Debelius J, et al. Best practices for analysing microbiomes. *Nat Rev Microbiol.* 2018;16:410–22 Available from: <https://doi.org/10.1038/s41579-018-0029-9>.
  29. Gilad Y, Mizrahi-Man O. A reanalysis of mouse ENCODE comparative gene expression data. *F1000Res.* 2015;4:121 Available from: <https://pubmed.ncbi.nlm.nih.gov/26236466>.
  30. Nearing JT, DeClercq V, Van Limbergen J, Langille MGI. Assessing the variation within the oral microbiome of healthy adults. Oh J, editor. *mSphere.* 2020;5 American Society for Microbiology Journals. Available from: <https://msphere.asm.org/content/5/5/e00451-20>.
  31. Bahl MI, Bergström A, Licht TR. Freezing fecal samples prior to DNA extraction affects the Firmicutes to Bacteroidetes ratio determined by downstream quantitative PCR analysis. *FEMS Microbiol Lett.* 2012;329:193–7 [cited 2020 Feb 18]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22325006>.
  32. Fouhy F, Deane J, Rea MC, O'Sullivan Ó, Ross RP, O'Callaghan G, et al. The effects of freezing on faecal microbiota as determined using MiSeq sequencing and culture-based investigations. Neu J, editor. *PLoS One.* 2015;10:e0119355 Public Library of Science. [cited 2020 Feb 18]. Available from: <https://dx.plos.org/10.1371/journal.pone.0119355>.
  33. Vandeputte D, Tito RY, Vanleeuwen R, Falony G, Raes J. Practical considerations for large-scale gut microbiome studies. *FEMS Microbiol Rev.* 2017;S154–67 [cited 2020 Feb 18]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28830090>.
  34. Shaw AG, Sim K, Powell E, Cornwell E, Cramer T, McClure ZE, et al. Latitude in sample handling and storage for infant faecal microbiota studies: the elephant in the room? *Microbiome.* 2016;4:40 BioMed Central Ltd. [cited 2020 Feb 18]. Available from: <http://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-016-0186-x>.
  35. Bokulich NA, Maldonado J, Kang D-W, Krajmalnik-Brown R, Caporaso JG. Rapidly processed stool swabs approximate stool microbiota profiles. Suen G, editor. *mSphere.* 2019;4:e00208–19 Available from: <http://msphere.asm.org/content/4/2/e00208-19.abstract>.
  36. McDonald D, Hyde E, Debelius JW, Morton JT, Gonzalez A, Ackermann G, et al. American Gut: an open platform for citizen science microbiome research. *mSystems. American Society for Microbiology.* 2018;3:e00031–18.
  37. Song SJ, Amir A, Metcalf JL, Amato KR, Xu ZZ, Humphrey G, et al. Preservation methods differ in fecal microbiome stability, affecting suitability for field studies. Dearing MD, editor. *mSystems.* 2016;1:e00021–16 Available from: <http://msystems.asm.org/content/1/3/e00021-16.abstract>.
  38. Amir A, McDonald D, Navas-Molina JA, Debelius J, Morton JT, Hyde E, et al. Correcting for microbial blooms in fecal samples during room-temperature shipping. *mSystems. American Society for Microbiology.* 2017;2:e00199–16.
  39. Lauber CL, Zhou N, Gordon JL, Knight R, Fierer N. Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples. *FEMS Microbiol Lett.* 2010;307:80–6 Oxford Academic. [cited 2020 Feb 18]. Available from: <https://academic.oup.com/femsle/article-lookup/doi/10.1111/j.1574-6968.2010.01965.x>.
  40. Choo JM, Leong LEX, Rogers GB. Sample storage conditions significantly influence faecal microbiome profiles. *Sci Rep. Nature Publishing Group.* 2015;5:1–10.
  41. Hale VL, Tan CL, Knight R, Amato KR. Effect of preservation method on spider monkey (*Ateles geoffroyi*) fecal microbiota over 8 weeks. *J Microbiol Methods. Elsevier.* 2015;113:16–26.
  42. Dominianni C, Wu J, Hayes RB, Ahn J. Comparison of methods for fecal microbiome biospecimen collection. *BMC Microbiol.* 2014;14:103 BioMed Central Ltd. [cited 2020 Feb 18]. Available from: <http://bmcmicrobiol.biomedcentral.com/articles/10.1186/1471-2180-14-103>.
  43. Watson EJ, Giles J, Scherer BL, Blatchford P. Human faecal collection methods demonstrate a bias in microbiome composition by cell wall structure. *Sci Rep; Nature Publishing Group.* 2019;9:1–8.
  44. Vogtmann E, Chen J, Amir A, Shi J, Abnet CC, Nelson H, et al. Comparison of collection methods for fecal samples in microbiome studies. *Am J Epidemiol.* 2017;185:115–23 Available from: <https://doi.org/10.1093/aje/kww177>.
  45. Brooks JP, Edwards DJ, Harwich MD, Rivera MC, Fettweis JM, Serrano MG, et al. The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies Ecological and evolutionary microbiology. *BMC Microbiol.* 2015;15:66 BioMed Central Ltd. [cited 2020 Feb 18]. Available from: <http://www.biomedcentral.com/1471-2180/15/66>.
  46. Kuske CR, Banton KL, Adorada DL, Stark PC, Hill KK, Jackson PJ. Small-scale DNA sample preparation method for field PCR detection of microbial cells and spores in soil. *Appl Environ Microbiol.* 1998;64:2463–72 American Society for Microbiology.
  47. Salonen A, Nikkilä J, Jalanka-Tuovinen J, Immonen O, Rajilčić-Stojanović M, Kekkonen RA, et al. Comparative analysis of fecal DNA extraction methods with phylogenetic microarray: effective recovery of bacterial and archaeal DNA using mechanical cell lysis. *J Microbiol Methods. Elsevier.* 2010;81:127–34.
  48. Morgan JL, Darling AE, Eisen JA. Metagenomic sequencing of an in vitro-simulated microbial community. *PLoS One.* 2010;5:e10209 Public Library of Science. Available from: <https://doi.org/10.1371/journal.pone.0010209>.
  49. Morita H, Kuwahara T, Ohshima K, Sasamoto H, Itoh K, Hattori M, et al. An improved DNA isolation method for metagenomic analysis of the microbial flora of the human intestine. *Microbes Environ.* 2007;22:214–22.
  50. Greathouse KL, Sinha R, Vogtmann E. DNA extraction for human microbiome studies: the issue of standardization. *Genome Biol.* 2019;20:212 Available from: <https://doi.org/10.1186/s13059-019-1843-8>.
  51. Carrigg C, Rice O, Kavanagh S, Collins G, O'Flaherty V. DNA extraction method affects microbial community profiles from soils and sediment. *Appl Microbiol Biotechnol.* 2007;77:955–64 Available from: <https://doi.org/10.1007/s00253-007-1219-y>.
  52. Krsek M, Wellington EMH. Comparison of different methods for the isolation and purification of total community DNA from soil. *J Microbiol Methods.* 1999;39:1–16 Available from: <http://www.sciencedirect.com/science/article/pii/S0167701299000937>.
  53. Costea PI, Zeller G, Sunagawa S, Pelletier E, Alberti A, Levenez F, et al. Towards standards for human fecal sample processing in metagenomic studies. *Nat Biotechnol. Nature Publishing Group.* 2017;35:1069–76.
  54. Mackenzie BW, Waite DW, Taylor MW. Evaluating variation in human gut microbiota profiles due to DNA extraction method and inter-subject differences. *Front Microbiol. Frontiers Media S.A.* 2015;6:130.
  55. Sinha R, Abu-Ali G, Vogtmann E, Fodor AA, Ren B, Amir A, et al. Assessment of variation in microbial community amplicon sequencing by the Microbiome Quality Control (MBQC) project consortium. *Nat Biotechnol.* 2017;35:1077–86 2017/10/02. Available from: <https://pubmed.ncbi.nlm.nih.gov/28967885>.
  56. Davis A, Kohler C, Alsallaq R, Hayden R, Maron G, Margolis E. Improved yield and accuracy for DNA extraction in microbiome studies with variation in

- microbial biomass. *Biotechniques*. 2019;66:285–9 Future Science. [cited 2020 Feb 18]. Available from: <https://www.future-science.com/doi/10.2144/btn-2019-0016>.
57. Kordy K, Gaufin T, Mwafu M, Li F, Cerini C, Lee DJ, et al. Contributions to human breast milk microbiome and enteromammary transfer of *Bifidobacterium breve*. *PLoS One*. 2020;15:e0219633 Public Library of Science. Available from: <https://doi.org/10.1371/journal.pone.0219633>.
  58. Lipinski JH, Moore BB, O'Dwyer DN. The evolving role of the lung microbiome in pulmonary fibrosis. *Am J Physiol Lung Cell Mol Physiol*. 2020;319(4):L675–82. <https://doi.org/10.1152/ajplung.00258.2020>.
  59. Nejman D, Livyatan I, Fuks G, Gavert N, Zwiang Y, Geller LT, et al. The human tumor microbiome is composed of tumor type-specific intracellular bacteria. *Science* (80- ) [Internet]. 2020;368:973 LP – 980. Available from: <http://science.sciencemag.org/content/368/6494/973.abstract>
  60. Lauder AP, Roche AM, Sherrill-Mix S, Bailey A, Laughlin AL, Bittinger K, et al. Comparison of placenta samples with contamination controls does not provide evidence for a distinct placenta microbiota. *Microbiome*. 2016;4:29 Available from: <https://doi.org/10.1186/s40168-016-0172-3>.
  61. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol*. BioMed Central Ltd. 2014;12:1–12.
  62. Hornung BVH, Zwiittink RD, Kuijper EJ. Issues and current standards of controls in microbiome research. *FEMS Microbiol Ecol*. 2019;95 Oxford University Press. [cited 2020 Feb 18]. Available from: <https://academic.oup.com/femsec/article/doi/10.1093/femsec/fiz045/5435435>.
  63. Karstens L, Asquith M, Davin S, Fair D, Gregory WT, Wolfe AJ, et al. Controlling for contaminants in low-biomass 16S rRNA gene sequencing experiments. *mSystems*. American Society for Microbiology. 2019;4:e00290–19.
  64. Leiby JS, McCormick K, Sherrill-Mix S, Clarke EL, Kessler LR, Taylor LJ, et al. Lack of detection of a human placenta microbiome in samples from preterm and term deliveries. *Microbiome*. NLM (Medline). 2018;6:196.
  65. de Goffau MC, Lager S, Sovio U, Gaccioli F, Cook E, Peacock SJ, et al. Human placenta has no microbiome but can contain potential pathogens. *Nature*. Nature Publishing Group. 2019;572:329–34.
  66. Segata N. No bacteria found in healthy placentas. *Nature*. NLM (Medline). 2019;317–8.
  67. Sharma JK, Gopalkrishna V, Das BC. A simple method for elimination of unspecific amplifications in polymerase chain reaction. *Nucleic Acids Res*. 1992;20(22):6117–8. Oxford University Press.
  68. Carroll NM, Adamson P, Okhravi N. Elimination of bacterial DNA from Taq DNA polymerases by restriction endonuclease digestion. *J Clin Microbiol*. American Society for Microbiology Journals. 1999;37:3402–4.
  69. Corless CE, Guiver M, Borrow R, Edwards-Jones V, Kaczmarek EB, Fox AJ. Contamination and sensitivity issues with a real-time universal 16S rRNA PCR. *J Clin Microbiol*. American Society for Microbiology Journals. 2000;38:1747–52.
  70. Wages JM, Decheng CAI, Fowler AK. Removal of contamination DNA from PCR reagents by ultrafiltration. *Biotechniques*. 1994;16(6):1014–1017.
  71. Stinson LF, Keelan JA, Payne MS. Identification and removal of contaminating microbial DNA from PCR reagents: impact on low-biomass microbiome analyses. *Lett Appl Microbiol*. 2019;68:2–8 Blackwell Publishing Ltd. [cited 2020 Feb 18]. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1111/lam.13091>.
  72. Minich JJ, Sanders JG, Amir A, Humphrey G, Gilbert JA, Knight R. Quantifying and understanding well-to-well contamination in microbiome research. *mSystems*. American Society for Microbiology. 2019;4:e00186–19.
  73. Costello M, Fleharty M, Abreu J, Farjoun Y, Ferreira S, Holmes L, et al. Characterization and remediation of sample index swaps by non-redundant dual indexing on massively parallel sequencing platforms. *BMC Genomics*. 2018;19:332 BioMed Central Ltd. [cited 2020 Feb 18]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29739332>.
  74. Illumina. Effects of index misassignment on multiplexing and downstream analysis. 2017.
  75. Comeau AM, Douglas GM, Langille MGI. Microbiome helper: a custom and streamlined workflow for microbiome research. *mSystems*. American Society for Microbiology. 2017;2.
  76. Edmonds K, Williams L. The role of the negative control in microbiome analyses. *FASEB J*. ; 2017;31:940.3. Federation of American Societies for Experimental Biology. Available from: [https://www.fasebj.org/doi/abs/10.1096/fasebj.31.1\\_supplement.940.3](https://www.fasebj.org/doi/abs/10.1096/fasebj.31.1_supplement.940.3)
  77. Larsson AJM, Stanley G, Sinha R, Weissman IL, Sandberg R. Computational correction of index switching in multiplexed sequencing libraries. *Nat Methods*. Nature Publishing Group. 2018;15:305–7.
  78. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome*. 2018;6:226 BioMed Central Ltd. [cited 2020 Feb 19]. Available from: <https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-018-0605-2>.
  79. Emerson JB, Adams RI, Román CMB, Brooks B, Coil DA, Dahlhausen K, et al. Schrödinger's microbes: tools for distinguishing the living from the dead in microbial ecosystems. *Microbiome*. BioMed Central. 2017;5:86.
  80. Costerton JW. Cystic fibrosis pathogenesis and the role of biofilms in persistent infection. *Trends Microbiol*. Elsevier Current Trends. 2001;9(2):50–2.
  81. Fu X, Zeng B, Wang P, Wang L, Wen B, Li Y, et al. Microbiome of total versus live bacteria in the gut of rex rabbits. *Front Microbiol*. 2018;9:733 Frontiers Media S.A. [cited 2020 Feb 19]. Available from: <http://journal.frontiersin.org/article/10.3389/fmicb.2018.00733/full>.
  82. Lennon JT, Jones SE. Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nat Rev Microbiol*. Nature Publishing Group. 2011;9:119–30.
  83. Blazewicz SJ, Barnard RL, Daly RA, Firestone MK. Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. *ISME J*. Nature Publishing Group. 2013;7:2061–8.
  84. Weinmaier T, Probst AJ, La Duc MT, Ciobanu D, Cheng J-F, Ivanova N, et al. A viability-linked metagenomic analysis of cleanroom environments: eukarya, prokaryotes, and viruses. *Microbiome*. 2015;3:62 BioMed Central. [cited 2020 Feb 19]. Available from: <http://www.microbiomejournal.com/content/3/1/62>.
  85. Stinson LF, Keelan JA, Payne MS. Characterization of the bacterial microbiome in first-pass meconium using propidium monoazide (PMA) to exclude nonviable bacterial DNA. *Lett Appl Microbiol*. 2019;68:378–85 Blackwell Publishing Ltd. [cited 2020 Feb 19]; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/30674082>.
  86. Marotz CA, Sanders JG, Zuniga C, Zaramela LS, Knight R, Zengler K. Improving saliva shotgun metagenomics by chemical host DNA depletion. *Microbiome*. 2018;6:42 BioMed Central Ltd. [cited 2020 Feb 19]. Available from: <https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-018-0426-3>.
  87. Checinska Sielaff A, Urbaniak C, Mohan GBM, Stepanov VG, Tran Q, Wood JM, et al. Characterization of the total and viable bacterial and fungal communities associated with the International Space Station surfaces. *Microbiome*. 2019;50–7 BioMed Central Ltd. [cited 2020 Feb 19]. Available from: <https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-019-0666-x>.
  88. Schrader C, Schielke A, Ellerbroek L, John R. PCR inhibitors - occurrence, properties and removal. *J Appl Microbiol*. 2012;113:1014–26 John Wiley & Sons, Ltd. [cited 2020 Feb 19]. Available from: <http://doi.wiley.com/10.1111/j.1365-2672.2012.05384.x>.
  89. Wu JY, Jiang XT, Jiang YX, Lu SY, Zou F, Zhou HW. Effects of polymerase, template dilution and cycle number on PCR based 16 S rRNA diversity analysis using the deep sequencing method. *BMC Microbiol*. BioMed Central. 2010;10:1–7.
  90. Qiu X, Wu L, Huang H, McDonel PE, Palumbo AV, Tiedje JM, et al. Evaluation of PCR-generated chimeras, mutations, and heteroduplexes with 16S rRNA gene-based cloning. *Appl Environ Microbiol*. American Society for Microbiology (ASM). 2001;67:880–7.
  91. Pääbo S, Irwin DM, Wilson AC. DNA damage promotes jumping between templates during enzymatic amplification. *J Biol Chem*. 1990;265:4718–21 [cited 2020 Feb 19]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/2307682>.
  92. Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G, et al. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res*. Cold Spring Harbor Laboratory Press. 2011;21:494–504.
  93. Gohl DM, Vangay P, Garbe J, MacLean A, Hauge A, Becker A, et al. Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. *Nat Biotechnol*. Nature Publishing Group. 2016;34:942–9.
  94. Sze MA, Schloss PD. The impact of DNA polymerase and number of rounds of amplification in PCR on 16S rRNA gene sequence data. *mSphere*. American Society for Microbiology. 2019;4:e00163–19.
  95. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*. 2011;27:2194–200 Oxford University Press. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3150044/>.

96. D'Amore R, Ijaz UZ, Schirmer M, Kenny JG, Gregory R, Darby AC, et al. A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling. *BMC Genomics*. 2016;17:55 BioMed Central Ltd. [cited 2020 Feb 19]. Available from: <http://www.biomedcentral.com/1471-2164/17/55>.
97. Salipante SJ, Kawashima T, Rosenthal C, Hoogestraat DR, Cummings LA, Sengupta DJ, et al. Performance comparison of Illumina and Ion Torrent next-generation sequencing platforms for 16S rRNA-based bacterial community profiling. *Appl Environ Microbiol*. 2014;80:7583–91 American Society for Microbiology. [cited 2020 Feb 19]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25261520>.
98. Johnson JS, Spakowicz DJ, Hong BY, Petersen LM, Demkowicz P, Chen L, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun*. Nature Publishing Group. 2019; 10:1–11.
99. Goldstein S, Beka L, Graf J, Klassen JL. Evaluation of strategies for the assembly of diverse bacterial genomes using MinION long-read sequencing. *BMC Genomics*. 2019;20:23 BioMed Central Ltd. [cited 2020 Feb 19]. Available from: <https://bmcbgenomics.biomedcentral.com/articles/10.1186/s12864-018-5381-7>.
100. Rhoads A, Au KF. PacBio sequencing and its applications. *Genomics Proteomics Bioinforma*. Beijing Genomics Institute. 2015;13(5):278–89.
101. Rang FJ, Kloosterman WP, de Ridder J. From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy [Internet]. *Genome Biol*. 2018;90 BioMed Central Ltd. [cited 2020 Feb 19]. Available from: <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-018-1462-9>.
102. Schloss PD, Jenior ML, Koumpouras CC, Westcott SL, Highlander SK. Sequencing 16S rRNA gene fragments using the PacBio SMRT DNA sequencing system. *PeerJ*. PeerJ Inc. 2016;2016:e1869.
103. Amarasinghe SL, Su S, Dong X, Zappia L, Ritchie ME, Gouli Q. Opportunities and challenges in long-read sequencing data analysis. *Genome Biol*. 2020; 21:30 Available from: <https://doi.org/10.1186/s13059-020-1935-5>.
104. Yaffe E, Tanay A. Probabilistic modeling of Hi-C contact maps eliminates systematic biases to characterize global chromosomal architecture. *Nat Genet*. 2011;43:1059–65 United States.
105. Huse SM, Dethlefsen L, Huber JA, Welch DM, Relman DA. Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genet*. 2008;4:1000255 [cited 2020 Feb 19]. Available from: <http://nihroadmap.nih.gov/hmp/>.
106. Kim M, Morrison M, Yu Z. Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. *J Microbiol Methods*. Elsevier. 2011;84:81–7.
107. Kumar PS, Brooker MR, Dowd SE, Camerlengo T. Target region selection is a critical determinant of community fingerprints generated by 16S pyrosequencing. *PLoS One*. 2011;6:e20956 Public Library of Science. Available from: <https://doi.org/10.1371/journal.pone.0020956>.
108. Hansen MC, Tolker-Nielsen T, Givskov M, Molin S. Biased 16S rDNA PCR amplification caused by interference from DNA flanking the template region. *FEMS Microbiol Ecol*. 1998;26:141–9 Oxford University Press (OUP). [cited 2020 Feb 19]. Available from: <https://academic.oup.com/femsec/article-lookup/doi/10.1111/j.1574-6941.1998.tb00500.x>.
109. Cai L, Ye L, Tong AHY, Lok S, Zhang T. Biased diversity metrics revealed by bacterial 16S pyrotags derived from different primer sets. Smidt H, editor. *PLoS One*. 2013;8:e53649 Public Library of Science. [cited 2020 Feb 19]. Available from: <https://dx.plos.org/10.1371/journal.pone.0053649>.
110. Willis C, Desai D, LaRoche J, et al. *FEMS Microbiol Lett*. 2019;366 Available from: <https://doi.org/10.1093/femsec/fnz1152>.
111. Zhang J, Ding X, Guan R, Zhu C, Xu C, Zhu B, et al. Evaluation of different 16S rRNA gene V regions for exploring bacterial diversity in a eutrophic freshwater lake. *Sci Total Environ*. 2018;618:1254–67 Available from: <http://www.sciencedirect.com/science/article/pii/S0048969717325792>.
112. Olson ND, Treangen TJ, Hill CM, Cepeda-Espinoza V, Ghurye J, Koren S, et al. Metagenomic assembly through the lens of validation: recent advances in assessing and improving the quality of genomes assembled from metagenomes. *Brief Bioinform*. 2019;20:1140–50 Oxford University Press. [cited 2020 Feb 19]. Available from: <https://academic.oup.com/bib/article/20/4/1140/4075034>.
113. Sato MP, Ogura Y, Nakamura K, Nishida R, Gotoh Y, Hayashi M, et al. Comparison of the sequencing bias of currently available library preparation kits for Illumina sequencing of bacterial genomes and metagenomes. *DNA Res*. 2019;26:391–8 NLM (Medline). [cited 2020 Feb 19]. Available from: <https://academic.oup.com/dnaresearch/article/26/5/391/5541856>.
114. Poptsova MS, Il'icheva IA, Nechipurenko DY, Panchenko LA, Khodikov MV, Oparina NY, et al. Non-random DNA fragmentation in next-generation sequencing. *Sci Rep*. 2014;4:1–6 Nature Publishing Groups.
115. Caruso V, Song X, Asquith M, Karstens L. Performance of microbiome sequence inference methods in environments with varying biomass. *mSystems*. American Society for Microbiology. 2019;4:e00163–18.
116. Schloss PD, Westcott SL. Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. *Appl Environ Microbiol*. American Society for Microbiology. 2011;77: 3219–26.
117. Westcott SL, Schloss PD. De novo clustering methods outperform reference-based methods for assigning 16S rRNA gene sequences to operational taxonomic units. *PeerJ*. PeerJ Inc. 2015;3:e1487.
118. Kopylova E, Navas-Molina JA, Mercier C, Xu ZZ, Mahé F, He Y, et al. Open-source sequence clustering methods improve the state of the art. *mSystems*. American Society for Microbiology. 2016;1:e00003–15.
119. He Y, Caporaso JG, Jiang X-T, Sheng H-F, Huse SM, Rideout JR, et al. Stability of operational taxonomic units: an important but neglected property for analyzing microbial diversity. *Microbiome*. 2015;3:20 Springer Science and Business Media LLC. [cited 2020 Feb 19]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25995836>.
120. Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J*. Nature Publishing Group. 2017;11:2639–43.
121. Sul WJ, Cole JR, Jesus EDC, Wang Q, Farris RJ, Fish JA, et al. Bacterial community comparisons by taxonomy-supervised analysis independent of sequence alignment and clustering. *Proc Natl Acad Sci U S A*. National Academy of Sciences. 2011;108:14637–42.
122. Edgar RC. Accuracy of microbial community diversity estimated by closed- and open-reference OTUs. *PeerJ*. PeerJ Inc. 2017;2017:e3889.
123. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*. Nature Publishing Group. 2016;13:581–3.
124. Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Zech Xu Z, et al. Deblur rapidly resolves single-nucleotide community sequence patterns. *mSystems*. American Society for Microbiology. 2017;2:e00191–16.
125. Edgar RC. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *bioRxiv*. 2016; Cold Spring Harbor Laboratory. Available from: <http://biorxiv.org/content/early/2016/10/15/081257.abstract>.
126. Edgar RC. Accuracy of taxonomy prediction for 16S rRNA and fungal ITS sequences. *PeerJ*. PeerJ Inc. 2018;2018:e4652.
127. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol*. American Society for Microbiology (ASM). 2007; 73:5261–7.
128. Lu J, Salzberg SL. Ultrafast and accurate 16S rRNA microbial community analysis using Kraken 2. *Microbiome*. 2020;8:124 Available from: <https://doi.org/10.1186/s40168-020-00900-2>.
129. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215(3):403–10. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
130. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ*. PeerJ Inc. 2016;2016:e2584.
131. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, et al. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome*. 2018;6:90 BioMed Central Ltd. [cited 2020 Feb 19]. Available from: <https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-018-0470-z>.
132. McDonald D, Price MN, Goodrich J, Nawrocki EP, Desantis TZ, Probst A, et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J*. Nature Publishing Group. 2012;6:610–8.
133. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yara P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2012;41:D590–6 Oxford Academic. [cited 2020 Feb 19]. Available from: <http://academic.oup.com/nar/article/41/D1/D590/1069277/The-SILVA-ribosomal-RNA-gene-database-project>.
134. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, et al. Ribosomal database project: data and tools for high throughput rRNA analysis. *Nucleic*

- Acids Res. 2014;42:D633–42 Oxford Academic. [cited 2020 Feb 19]. Available from: <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkt1244>.
135. Sayers EW, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 2009;37:D10–7.
  136. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Res.* 2009;37:D26–D31.
  137. Balvočiute M, Huson DH, SILVA, RDP, Greengenes, NCBI and OTT - how do these taxonomies compare? *BMC Genomics.* 2017;18:114 BioMed Central Ltd. [cited 2020 Feb 19]. Available from: <http://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12864-017-3501-4>.
  138. Roux S, Enault F, Bronner G, Debroas D. Comparison of 16S rRNA and protein-coding genes as molecular markers for assessing microbial diversity (Bacteria and Archaea) in ecosystems. *FEMS Microbiol Ecol.* 2011;78:617–28 Oxford Academic. [cited 2020 Feb 19]. Available from: <https://academic.oup.com/femsec/article-lookup/doi/10.1111/j.1574-6941.2011.01190.x>.
  139. Větrovský T, Baldrian P. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. Neufeld J, editor. *PLoS One.* 2013;8:e57923 Public Library of Science. [cited 2020 Feb 19]. Available from: <https://dx.plos.org/10.1371/journal.pone.0057923>.
  140. Kembel SW, Wu M, Eisen JA, Green JL. Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance. *PLoS Comput Biol.* Public Library of Science. 2012;8:e1002743.
  141. Louca S, Doebeli M, Parfrey LW. Correcting for 16S rRNA gene copy numbers in microbiome surveys remains an unsolved problem. *Microbiome.* 2018;6:41 BioMed Central Ltd. [cited 2020 Feb 19]. Available from: <https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-018-0420-9>.
  142. Angly FE, Dennis PG, Skarshewski A, Vanwonterghem I, Hugenholtz P, Tyson GW. CopyRighter: a rapid tool for improving the accuracy of microbial community profiles through lineage-specific gene copy number correction. *Microbiome.* 2014;2:11 BioMed Central Ltd. [cited 2020 Feb 19]. Available from: <https://microbiomejournal.biomedcentral.com/articles/10.1186/2049-2618-2-11>.
  143. Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol.* Nature Publishing Group. 2013;31:814–21.
  144. Bowman JS, Ducklow HW. Microbial communities can be described by metabolic structure: a general framework and application to a seasonally variable, depth-stratified microbial community from the coastal West Antarctic Peninsula. *PLoS One.* Public Library of Science. 2015;10:e0135868.
  145. Ye SH, Siddle KJ, Park DJ, Sabeti PC. Benchmarking metagenomics tools for taxonomic classification. *Cell.* 2019;779–94 Cell Press. [cited 2020 Feb 19]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/31398336>.
  146. Breitwieser FP, Lu J, Salzberg SL. A review of methods and databases for metagenomic classification and assembly. [cited 2020 Feb 19]; Available from: <https://academic.oup.com/bib/article-abstract/doi/10.1093/bib/bbx120/4210288/A-review-of-methods-and-databases-for-metagenomic>.
  147. McIntyre ABR, Ounit R, Afshinnekoo E, Prill RJ, Hénaff E, Alexander N, et al. Comprehensive benchmarking and ensemble approaches for metagenomic classifiers. *Genome Biol.* 2017;18:182 Available from: <https://doi.org/10.1186/s13059-017-1299-7>.
  148. Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat Methods.* 2012;9:811 Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved. Available from: <https://doi.org/10.1038/nmeth.2066>.
  149. Milanese A, Mende DR, Paoli L, Salazar G, Ruscheweyh H-J, Cuenca M, et al. Microbial abundance, activity and population genomic profiling with mOTUs2. *Nat Commun.* 2019;10:1014 Available from: <https://doi.org/10.1038/s41467-019-08844-4>.
  150. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* 2014;15:R46 Available from: <https://doi.org/10.1186/gb-2014-15-3-r46>.
  151. Corvelo A, Clarke WE, Robine N, Zody MC. taxMaps: comprehensive and highly accurate taxonomic classification of short-read data in reasonable time. *Genome Res.* 2018;28:751–8 New York Genome Center, New York, New York 10013, USA. Available from: <http://europepmc.org/abstract/MED/29588360>.
  152. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* 2019;20:257 Available from: <https://doi.org/10.1186/s13059-019-1891-0>.
  153. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: estimating species abundance in metagenomics data. Procter J, editor. *PeerJ Comput Sci.* 2017; 3:e104 Available from: <https://doi.org/10.7717/peerj-cs.104>.
  154. Brister JR, Ako-adjei D, Bao Y, Blinkova O. NCBI viral genomes resource. *Nucleic Acids Res.* 2015;43:D571–7 Oxford University Press. [cited 2020 Feb 19]. Available from: <http://academic.oup.com/nar/article/43/D1/D571/2439171/NCBI-Viral-Genomes-Resource>.
  155. Garmaeva S, Sinha T, Kurilshikov A, Fu J, Wijmenga C, Zhernakova A. Studying the gut virome in the metagenomic era: challenges and perspectives. *BMC Biol.* BioMed Central Ltd. 2019;17(1):84.
  156. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics.* Oxford University Press. 2015;31:1674–6.
  157. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. MetaSPAdes: a new versatile metagenomic assembler. *Genome Res.* Cold Spring Harbor Laboratory Press. 2017;27:824–34.
  158. Kingsford C, Schatz MC, Pop M. Assembly complexity of prokaryotic genomes using short reads. *BMC Bioinformatics.* 2010;11:21 BioMed Central. [cited 2020 Feb 19]. Available from: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-11-21>.
  159. Alneberg J, Bjarnason BS, De Bruijn J, Schirmer M, Quick J, Ijaz UZ, et al. Binning metagenomic contigs by coverage and composition. *Nat Methods.* Nature Publishing Group. 2014;11:1144–6.
  160. Karlin S, Mrázek J, Campbell AM. Compositional biases of bacterial genomes and evolutionary implications. *J Bacteriol.* American Society for Microbiology. 1997;179:3899–913.
  161. Teeling H, Meyerdierks A, Bauer M, Amann R, Glockner FO. Application of tetranucleotide frequencies for the assignment of genomic fragments. *Environ Microbiol.* 2004;6:938–47 John Wiley & Sons, Ltd. [cited 2020 Feb 19]. Available from: <http://doi.wiley.com/10.1111/j.1462-2920.2004.00624.x>.
  162. Almeida A, Mitchell AL, Boland M, Forster SC, Gloor GB, Tarkowska A, et al. A new genomic blueprint of the human gut microbiota. *Nature.* Nature Publishing Group. 2019;568:499–504.
  163. Nayfach S, Shi ZJ, Seshadri R, Pollard KS, Kyrpides NC. New insights from uncultivated genomes of the global human gut microbiome. *Nature.* Nature Publishing Group. 2019;568:505–10.
  164. Pasolli E, Asnicar F, Manara S, Zolfo M, Karcher N, Armanini F, et al. Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. *Cell.* Cell Press. 2019;176:649–662.e20.
  165. Vollmers J, Wiegand S, Kaster AK. Comparing and evaluating metagenome assembly tools from a microbiologist's perspective - not only size matters! *PLoS One.* Public Library of Science. 2017;12(1):e0169662.
  166. Yue Y, Huang H, Qi Z, Dou H-M, Liu X-Y, Han T-F, et al. Evaluating metagenomics tools for genome binning with real metagenomic datasets and CAMI datasets. *BMC Bioinformatics.* 2020;21:334 Available from: <https://doi.org/10.1186/s12859-020-03667-3>.
  167. Pruitt KD, Tatusova T, Maglott DR. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* 2005;33:D501–4 Available from: <https://doi.org/10.1093/nar/gki025>.
  168. Parks DH, Chuvochina M, Chaumeil P-A, Rinke C, Mussig AJ, Hugenholtz P. A complete domain-to-species taxonomy for Bacteria and Archaea. *Nat Biotechnol.* 2020;38:1079–86 Available from: <https://doi.org/10.1038/s41587-020-0501-8>.
  169. Haynes WA, Tomczak A, Khatri P. Gene annotation bias impedes biomedical research. *Sci Rep.* Nature Publishing Group. 2018;8:1362.
  170. Krehenwinkel H, Wolf M, Lim JY, Rominger AJ, Simison WB, Gillespie RG. Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Sci Rep.* Nature Publishing Group. 2017;7:1–12.
  171. Bell KL, Burgess KS, Botsch JC, Dobbs EK, Read TD, Brosi BJ. Quantitative and qualitative assessment of pollen DNA metabarcoding using constructed species mixtures. *Mol Ecol.* 2019;28:431–55 Blackwell Publishing Ltd. [cited 2020 Feb 19]. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.14840>.
  172. Martiny AC. High proportions of bacteria are culturable across major biomes. *ISME J.* Nature Publishing Group. 2019;13:2125–8.
  173. Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ. Microbiome datasets are compositional: and this is not optional. *Front Microbiol.* 2017;8:2224 *Frontiers Media S.A.* [cited 2020 Feb 19]. Available from: <http://journal.frontiersin.org/article/10.3389/fmicb.2017.02224/full>.
  174. Vandeputte D, Kathagen G, D'Hoe K, Vieira-Silva S, Valles-Colomer M, Sabino J, et al. Quantitative microbiome profiling links gut community variation to microbial load. *Nature.* Nature Publishing Group. 2017;551:507–11.

175. Tkacz A, Hortala M, Poole PS. Absolute quantitation of microbiota abundance in environmental samples. *Microbiome*. 2018;6:110 BioMed Central Ltd. [cited 2020 Feb 19]. Available from: <https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-018-0491-7>.
176. Gilbert JA, Jansson JK, Knight R. The Earth Microbiome project: successes and aspirations. *BMC Biol*. 2014;12:69 BioMed Central Ltd. [cited 2020 Feb 19]. Available from: <http://bmcbiol.biomedcentral.com/articles/10.1186/s12915-014-0069-1>.
177. Mirzayi C, Renson A, Zohra F, Elsafoury S, Kasselmann L, van de Wijkert J, et al. Strengthening The Organizing and Reporting of Microbiome Studies (STORMS). *bioRxiv*. 2020:2020.06.24.167353 Available from: <http://biorxiv.org/content/early/2020/06/24/2020.06.24.167353.abstract>.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

