Ranking the Biases: The Choice of OTUs vs. ASVs in 16S rRNA Amplicon Data Analysis Has Stronger Effects On Diversity Measures Than Rarefaction and Similarity Threshold

Marlène Chiarello (marlene.chiarello@gmail.com)
University of Mississippi, University

Mark McCauley
University of Mississippi, University

Sébastien Villéger
MARBEC, University of Montpellier, CNRS, Ifremer, IRD

Colin R Jackson
University of Mississippi, University

Research Article

Keywords: phylogenetic diversity, microbiome, high-throughput sequencing, Operational Taxonomic Unit, Exact Sequence Variant

DOI: https://doi.org/10.21203/rs.3.rs-764430/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background

Advances in the analysis of amplicon sequence datasets have introduced a methodological shift in how research teams investigate microbial biodiversity, away from the classification and downstream analyses of traditional operational taxonomic units (OTUs), and towards the usage of amplicon sequence variants (ASVs). While ASVs have several inherent properties that make them desirable compared to OTUs, questions remain as to the influence that these pipelines have on the ecological patterns being assessed, especially when compared to other methodological choices made when processing data (e.g. rarefaction) and computing diversity indices.

Results

We compared the respective influences of using ASVs vs. OTU-based pipelines, rarefaction of the community table, and OTU similarity threshold (97% vs. 99%) on the ecological signals detected in freshwater invertebrate and environmental (sediment, seston) 16S rRNA data sets, determining the effects on alpha diversity, beta diversity and taxonomic composition. While the choice of OTU vs. ASV pipeline significantly influenced unweighted alpha and beta diversities and changed the ecological signal detected, weighted indices such as the Shannon index, Bray-Curtis dissimilarity, and weighted Unifrac scores were not impacted by the pipeline followed. By comparison, OTU threshold and rarefaction had a minimal impact effect on all measurements, although rarefaction improved overall signals, especially in OTU-based datasets. The identification of major classes and genera identified revealed significant discrepancies across methodologies.

Conclusion

We provide a list of recommendations for the analysis of 16S rRNA amplicon data. We notably recommend the use of ASVs when analyzing alpha-diversity patterns, especially in species-rich or environmental samples. Abundance weighted alpha- and beta-diversity indices should also be preferred compared to ones based on the presence-absence of biological units.

Background

Evaluating the microbial diversity of various environments, from host-associated microbiomes to free-living communities in water, soil and air, is essential for understanding biodiversity, mitigating the effects of global change and improving human health and agriculture [1–3]. The development and increased accessibility of high-throughput sequencing technologies [4], has supported the advancement of large-scale assessments of microbial diversity over the past decade [5]. As with improvements in sequencing technology, bioinformatics techniques for the analysis of high-throughput sequencing data have been concurrently improving [6]. For targeted amplicon sequencing, i.e. the analysis of bacterial diversity by sequencing a specific gene (or region of a gene), there are now at least 15 dedicated pipelines available,
which have been developed thanks to increasing computational capacities and algorithm efficiencies [7–10]. Some pipelines have been created for specific needs [9–10], and others were developed to adapt to conceptual changes in the analysis methods [11–12].

In order to investigate biodiversity within amplicon sequencing data, researchers usually attempt to aggregate the >50,000 paired reads into sequences (e.g. Illumina MiSeq delivers up to 15Gb using kits for 2x300bp reads), before grouping them into biological units, widely termed operational taxonomic units (OTUs). OTUs are often defined as groups of sequences that differ by less than 97% of their nucleotides [7]. This similarity threshold is preferred as it is considered a suitable operational definition for “bacterial species” [13]. Clustering based on 97% similarity reduces the size of the raw 16S rRNA dataset, and therefore decreases the computational requirements for analysis [14]. It also reduces the impact of sequencing errors in downstream diversity estimations, as erroneous sequences are likely to be merged with correct sequences [11, 15].

Recently, a methodological shift has occurred with the increased use of amplicon sequence variants (ASVs) (e.g. [16]), also known as exact sequence variants. ASVs were first introduced by Eren et al. [11], originally using the term ‘oligotype’ to define sequences that differed by only one nucleotide. ASV-producing algorithms utilize different techniques for sequence denoising than OTU-based pipelines, by generating an error model tailored to an individual sequencing run, and using this model to distinguish between the predicted “true” biological variation and that likely generated by sequencing error [8, 17]. Therefore, ASVs are not simply equivalents to “100%-OTUs” [18]. One advantage of ASVs over 97%-similarity defined OTUs is their higher sensitivity to biological variation, as a change in one nucleotide in the 16S rRNA gene of a bacterial strain can indicate large variations within the rest of the genome [11]. Nonetheless, ASVs should not be considered as perfect proxies for different bacterial ecotypes [19].

With the increasing popularity of ASVs, studies have compared the results from OTU and ASV approaches. Based on mock communities, ASV-based pipelines had a higher sensitivity in detecting bacterial strains present, sometimes at the expense of specificity [20–23]. However, studies utilizing soil, rhizosphere, and human microbiome datasets found similar overall biological signals [16, 24–25]. From these studies, the main weakness of OTU approaches appear to be in the accurate detection of alpha diversity, as OTUs often overestimated bacterial richness when compared to ASVs, whereas beta diversity estimates were more congruent between methodologies [17, 22, 24–25]. Further an analysis using various OTU and ASV-based programs and a wide range of natural and mock communities identified significant discrepancies in the taxonomic assignment and estimation of relative abundance of functionally important taxa [26].

However, the drivers of these discrepancies have not been fully disentangled. For instance, what is the influence of choosing ASVs vs. OTUs when compared to the influence of other methodological choices that are known to impact biological patterns, such as rarefaction, removal of singletons, and specific diversity indices [27–30]? Further, while the relative impact of each of these choices may depend on the characteristics (richness, evenness of relative abundances, phylogenetic structure) of the community
analyzed, many of the aforementioned studies focused on just one environment (but see [20, 26]). A bacterial community with a few closely related but dominant taxa may be more sensitive to the choice of analysis, or rarefaction, than a community with a greater phylogenetic diversity or evenness of abundances. Further, while phylogenetic diversity indices (e.g. UniFrac) are typically more efficient at detecting ecological patterns and evolutionary processes than indices based on just biological units (e.g. Bray-Curtis or Jaccard indices, hereafter termed ‘taxonomic diversity indices’) [30–32], only one of the aforementioned studies compared phylogenetic diversity metrics obtained from ASVs vs. OTUs datasets [26].

In this study, we compare the influence of two widely used sequence-processing methods, namely the ASV-based DADA2 and OTU-based MOTHUR pipeline, on the diversity, ecological and compositional patterns of the bacterial communities present in three different microbial assemblages (aquatic sediment, particle-associated freshwater communities ‘seston’, and the freshwater mussel gut microbiome). Notably, we compared the effects of ASV vs. OTU approaches using two different similarity thresholds, to those with varying levels of sequence rarefaction, and eight contrasting alpha and beta-diversity metrics.

**Methods**

**Sampling of bacterial communities**

A total of 54 surface sediment, 54 seston and 121 mussel (host-associated) microbiome samples were collected from 18 sample sites located on six rivers in the Tennessee and Mobile River Basins, USA, in Summer 2019 (Table 1). At each site, three sediment and three seston samples were collected in the middle of the river channel. Sediment samples were collected using sterile 15 mL centrifuge tubes that were dragged through the top 5 cm of sediment. Seston samples were collected by filtering 100 mL of river water through sterile 47 mm, 1 µ pore size sterile glass fiber filters which were then placed in sterile 15 ml tubes. 4–5 mussel specimens belonging to three different species (*Lampsilis ornata, Amblema plicata, Cyclonaias asperata*) were manually collected (Supplementary Table 1). Mussels were collected under the authority of the permit (USFWS permit #TE68616B-1 and ALCDNR permit #2016077745468680) issued by the Alabama Department of Conservation and Natural Resources to Dr. Carla Atkinson, University of Alabama. After collection, all filters, sediment samples and mussels were stored on ice, and taken to the University of Alabama for storage at -80°C. Mussel gut tissue was excised using sterile dissecting equipment and all samples were transported to the University of Mississippi for microbiome analysis.

**DNA extraction and 16S rRNA gene high-throughput sequencing**

DNA was extracted from mussel gut tissue using a PowerSoil Pro extraction kit (Qiagen, Germantown, MD, USA) as described previously [36]. For seston and sediment microbiome analysis, filters or 30 mg of
sediment were added directly into the bead beating tube of the extraction kit, before performing the standard extraction procedure.

Dual-indexed barcoded primers were used to amplify the V4 region of the 16S rRNA gene of the extracted DNA following established techniques [7, 36]. The amplified 16S rRNA gene fragments were combined and spiked with 20% PhiX before being sequenced on an Illumina MiSeq at the University of Mississippi Medical Centre Molecular and Genomics Core facility. Sequence data was obtained as FASTQ files and two different approaches (OTU pipeline, ASV pipeline) used to assess patterns in microbiome diversity.

**OTU pipeline (Mothur)**

Mothur [44] is one of the most cited bioinformatic tools used to construct biological units by clustering similar bacterial sequences into OTUs. Mothur uses a percentage of similarity criterion for clustering, after several steps of sequence filtering. The most popular value of the threshold for percentage of similarity is 97%, but 99% similarity is also often reported. We compared 97%-similarity and 99%-similarity assembled OTUs, and generally followed the recommended procedures for Illumina MiSeq sequence data (https://mothur.org/wiki/miseq_sop/, December 2020). Briefly, after merging the forward and reverse reads, we screened the sequences and removed those of unusual length or that contained ambiguous bases. Unique sequences were then aligned to the Silva 16S rRNA gene database v. 138, and poorly aligned reads were removed. Sequences were classified using the default Wang approach in Mothur through the same database. All non-prokaryotic sequences, as well as sequences that were not classified to the Kingdom level, were removed. Chimeras were determined and removed using the chimera.vsearch command with default parameters. OTUs were assembled based on 97% - and 99% -similarity and OTU tables were constructed before retrieving the consensus classification for each OTU. The most abundant sequence within each OTU was then selected as the reference sequence for further phylogenetic analyses. A total of 31,453 97%-defined OTUs and 58,483 99%-defined OTUs were detected in the whole dataset of 217 samples, based on 3,420 – 143,827 valid sequence reads per sample.

**ASV pipeline (DADA2)**

DADA2 is a more recent pipeline that is implemented as an R-package to allow for the detection of ASVs [8]. DADA2 uses a sequence error model that aims to infer the biological sequences before the introduction of errors that are likely caused by amplification and sequencing. These sequences, that differ by as little as one nucleotide, are then considered as separate biological units. In DADA2, we followed the general tutorial available on the Github of the software (https://benjjneb.github.io/dada2/tutorial.html, November 2020). Forward and reverse reads with > 2 or > 5 estimated errors, respectively, were filtered, and were truncated at the 3’ end where read quality dropped below a quality score of 2 (TrunQ = 2). After separate estimations of error rates on forward and reverse reads, ASVs were predicted and merged using a minimal overlap of 12 nucleotides. Chimeras were removed using the consensus method with the removeBimeraDenovo() function. ASVs with < 243 or > 263 base pairs was removed. ASVs were classified using the naïve Bayesian classifier and the SILVA 16S rRNA gene database v. 138, using DADA2’s default parameters. There were a total of 21,606 ASVs in the unrarefied dataset of 217 samples, based on 3,164
−112,811 valid sequence reads per sample. Despite the same version of the SILVA database utilized for classification of both OTUs and ASVs, minor differences were observed, with Oxyphotobacteria considered as a class within the ASV dataset, and Planctomycetes in the OTU datasets named as Planctomycetacia in the ASV dataset. To simplify comparison across methodologies, we renamed Oxyphotobacteria as Cyanobacteria, and Planctomycetacia as Planctomycetes within the ASV dataset.

Relative abundance normalization and taxonomic diversity

ASV and OTU tables were imported into R using the phyloseq R-package [45]. All samples containing <3,000 sequences were removed, leaving a total of 217 samples, consisting of 45 sediment, 51 seston, and 121 host-associated microbiome communities (Supplementary Table 1). To assess the effect of sample rarefaction, random subsampling of 1,000, 2,000 and 3,000 sequences per sample was performed for each community table using phylosEq. Subsequent analyses were performed at each rarefaction level as well as on unrarefied OTU-97%, OTU-99% and ASVs tables, for a total of 12 community analysis tables (three ASV/OTU approaches x four levels of rarefaction). Sampling coverage was determined for all analysis tables using Chao's coverage index provided in the entropart R-package [46]. Indices of taxonomic alpha diversity (observed richness, iChao1 and ACE richness indices correcting for unsampled biological units, and Shannon alpha diversity) were assessed using vegan (for observed richness and Shannon), entropart (for iChao1) and fossil (for ACE) R-packages [46–47]. Following recommendations from Jost [48], all alpha diversity indices were expressed in equivalent number of biological units (i.e., equivalent numbers of species, also known as Hill numbers). Taxonomic beta diversity was assessed using complementary indices based on Jaccard and Bray-Curtis dissimilarities, using vegan. While Jaccard is solely based on the presence/absence of biological units, Bray-Curtis gives more weight to the most abundant biological units.

Phylogenetic diversity

ASV and OTU reference sequences were incorporated into the GreenGenes 99% phylogenetic tree version 13.8 [49], using SEPP software [50] implemented in QIIME2 [51] using default parameters. The obtained bacterial phylogenetic tree was then pruned using ape R-package [52] to remove all tree leaves absent from our dataset while keeping the structure of the tree. Phylogenetic beta diversity was then assessed using the weighted (W-) and unweighted (U-) versions of the Unifrac index, using GUniFrac R-package [40].

Statistical analyses

Statistical analyses were computed using R software v. 3.6.3 and data visualization was made using ‘ggplot2’ R-package. Differences in alpha and beta diversity values across different methodologies (ASVs vs. 97%-OTUs vs. 99%-OTUs), rarefaction level (no rarefaction vs. rarefaction to 1,000 vs. 2,000 vs. 3,000 sequences), and OTU threshold (99% vs. 97% similarity) were assessed using pairwise Wilcoxon signed-rank tests. Correlation in ranks of alpha and beta diversity metrics, and relative abundance of taxonomic groups across the different analysis treatments were made using Spearman’s signed rank tests. Similarly,
correlations of abundance of the top bacterial classes and genera were made using separate Spearman’s signed rank tests and plotted using ‘corrplot’ R-package [53].

Detection of ecological signals (effect of community type, river, and sample site) in alpha diversity was conducted with Kruskal-Wallis tests and associated post-hoc pairwise comparisons using ‘pgirmess’ R-package. Detection of the effects of these signals on microbial structure was conducted with PERMANOVAs performed on beta-diversity matrices using the ‘vegan’ R-package (900 permutations). R-square values from the PERMANOVAs were extracted and compared across treatments to assess the respective impact of ASVs vs. OTUs, rarefaction, and the beta diversity metric used, using random forest models provided in ‘randomForest’ R-package [54]. These models are based on successive decision trees to allow ranking of the effects of several correlated variables on a given response [55], here on the quality of detection of an ecological signal.

Results

We analyzed the microbiome of 54 surface sediment, 54 seston and 121 mussel gut samples collected from 18 sites located in the Tennessee and Mobile River Basins, USA (Table S1) through amplicon-based sequencing of the v4 region of the 16S rRNA gene. Sequence processing was performed through Mothur and DADA2, before performing rarefaction, if any, and computing diversity indices.

Structure and coverage of 16S rRNA datasets ASV and OTU based analyses resulted in significantly different numbers of biological units in non-rarefied datasets, with fewer ASVs (mean 303 ± 215 per sample) than 97% similarity defined OTUs (1,386 ± 1,026), and 99% similarity defined OTUs (1,708 ± 1,319) (Wilcoxon signed-rank tests, p < 0.001 in all comparisons). The abundance distribution of biological units was more even in the ASV dataset than in 99%-OTU dataset and 97%-OTU datasets (Fig. 1; Wilcoxon signed-rank tests p < 0.001 in all comparisons). This pattern was apparent across all sample types, although evenness was highest in the sediment, intermediate in seston and lowest in host-associated communities (Fig. 1; Kruskal-Wallis and post-hoc associated pairwise tests, p < 0.001 for all comparisons). Chao’s coverage, which assesses the completeness of community sampling, was approximately 100% in all ASV-based datasets, and was significantly lower in OTU-based datasets (Wilcoxon signed-rank tests, p < 0.001 in all comparisons, Figure-S1). For the OTU based datasets, Chao’s coverage index was the highest in the mussel gut microbiome samples (mean of 97 ± 3% in mussel microbiome), and lower in the seston (81 ± 7%) and sediment communities (72 ± 13%).

Figure 1: Effect of sequence processing methodology on the rank abundance of biological units. The relative abundance of ASVs (obtained from DADA2) and OTUs defined by 97% or 99% similarity (obtained from MOTHUR). Units were sorted according to their rank of abundance and represented separately for each community type (a: mussel gut microbiome, b: sediment, c: seston). Inserts focus on the 500 most abundant biological units. The evenness of the relative abundance of the biological units were computed using Bulla’s O and are displayed on each plot.
Alpha-diversity metric

All alpha-diversity metrics were distinct between sequence processing methods (ASVs vs. OTUs vs.), at every rarefaction level (Wilcoxon signed-rank tests p < 0.001 in all comparisons, Fig. 2, Figure-S1). ASV-derived richness was roughly twice as low as 99%-OTU and 97%-OTU richness, which were closer to each other (99%-OTU richness was 1.1 times higher than 97%-OTU richness). The difference between ASV- and OTU-derived richness was even higher using iChao1 (4.8–5.3 higher in 97%- and 99%-OTUs datasets than in ASVs, respectively, Figure-S1) and ACE (3.5–4.1 times respectively, Figure-S1). However, differences between the approaches were lower in the case of the Shannon alpha-diversity index (1.4 and 1.7 times respectively, Fig. 2).

Rarefaction decreased all alpha-diversity estimates (Fig. 2 & Figure-S1; Wilcoxon signed-rank tests p < 0.001 in all comparisons). This decrease was lower for the ASV-based datasets than for those derived from OTUs (Figure-S2). For example, after rarefaction to 2,000 sequences per sample, the decrease in observed richness compared to before rarefaction was 63 ± 17% for 99%-OTUs, 59 ± 17% for 97%-OTUs, but only 22 ± 15% for ASVs (Figure-S2). The post rarefaction decline varied across indices, being the greatest for observed richness, ACE and iChao1 (respectively, 48 ± 25%, 43 ± 24% and 40 ± 24% across all datasets), and lowest for Shannon alpha-diversity (21 ± 17%).

Figure 2: Effect of sequence processing methodology on alpha-diversity metrics for microbiome analyses. Measure of taxonomic richness (expressed as number of OTUs or ASVs and Shannon alpha-diversity depending on the methodology used, namely ASV, 97%-OTU and 99%-OTU, at 3 different levels of rarefaction (1,000; 2,000 and 3,000 sequences per sample) within each sample type studied (a-c). To ease the visualization of differences across methods and rarefaction levels, the y-axis of plots has been log transformed. The effect of methodological choices on other alpha-diversity metrics is available in Figure S1.

The choice of ASV or OTU-based analysis also influenced how samples ranked in terms of alpha diversity. Correlations in sample ranking between ASVs and OTUs approaches were the lowest for richness indices that correct for sampling bias (iChao1 and ACE; Spearman’s r = 0.61–0.84 depending on index, rarefaction level and OTU threshold), intermediate for observed richness (0.80–0.85), and highest for Shannon diversity (0.90–0.97) (Table 1). Rarefaction increased the strength of these correlations, but only minimally (Table 1).
Table 1

**Correlation between alpha-diversity metrics between OTU- and ASV-based datasets accross every rarefaction level tested.** Correlations were tested using Spearman’s signed rank tests (all P values were < 0.001). Correlations showing poorer agreement between OTUs and ASVs are noted, i.e. <0.9 **(bold)** or < 0.8 **(bold and underlined)**.

| Diversity index | Rarefaction | ASVs vs. 97%-OTUs | ASVs vs. 99%-OTUs |
|-----------------|-------------|-------------------|-------------------|
| **Richness**    | No rarefaction | 0.80              | 0.81              |
|                 | 3,000 sequences | 0.81              | 0.81              |
|                 | 2,000 sequences | 0.85              | 0.85              |
|                 | 1,000 sequences | 0.90              | 0.91              |
| **Shannon**     | No rarefaction | 0.95              | 0.96              |
|                 | 3,000 sequences | 0.96              | 0.96              |
|                 | 2,000 sequences | 0.96              | 0.96              |
|                 | 1,000 sequences | 0.96              | 0.97              |
| **iChao1**      | No rarefaction | 0.62              | 0.61              |
|                 | 3,000 sequences | 0.65              | 0.66              |
|                 | 2,000 sequences | 0.70              | 0.69              |
|                 | 1,000 sequences | 0.76              | 0.78              |
| **ACE**         | No rarefaction | 0.70              | 0.69              |
|                 | 3,000 sequences | 0.73              | 0.73              |
|                 | 2,000 sequences | 0.77              | 0.77              |
|                 | 1,000 sequences | **0.84**          | **0.84**          |

In terms of sample type, observed richness and Shannon diversity were highest in the sediment, intermediate in the seston, and lowest in mussel host-associated communities, regardless of sequence processing method or rarefaction level (Kruskal-Wallis and associated pairwise post-hoc tests, p < 0.05 in all cases and for all comparisons, Fig. 2). However, more subtle biological effects, like that of river or site, varied with the sequence processing method and the alpha-diversity metric, with an intensity that was dependent on the type of the community (Fig. 3). While differences in Shannon diversity between rivers were consistent regardless of sequence processing method and sample type, patterns in observed richness were more variable. For sediment samples, richness patterns were generally stable to shifting from ASV to OTU-based approaches, with no significant differences in sediment microbiome richness between the six rivers. However, for seston and mussel host communities, the selection of an ASV or OTU-
approach influenced the outcome as to whether any rivers were significantly different in richness or not (Fig. 3).

**Figure 3: Comparison of observed richness and Shannon alpha-diversity across sampling rivers and sequence processing methods.** For each community type, (a) mussel microbiome, (b) sediment and (c) Seston, the average alpha-diversity value per river were represented by the size and color intensity of the corresponding circle. Letters representing whether diversity values of different rivers were significantly different or not, were displayed (Kruskal-Wallis' pairwise post-hoc tests, \(P < 0.05\)). For instance, ASV richness of mussel microbiome sampled on Bear Creek ('c') was significantly lower than the one of mussels collected on river Buttahatchee ('a'), but did not differ from the ones collected on Paint Rock ('abc'). To facilitate comparison across panels, all metrics were scaled between 0 and 1.

**Beta-diversity indices**

Beta-diversity metrics computed with ASVs and OTUs were highly correlated (Spearman's correlation values > 0.9), with the exception of U-Unifrac (< 0.81, Table 2). Rarefaction slightly increased agreement between ASV and OTU based beta-diversity metrics, for both OTU thresholds (Table 2). Rarefaction only had a low effect on Bray-Curtis and Jaccard indices, with a correlation > 0.9 between rarefied and unrarefied data and no effect on W-Unifrac (r = 1; Fig. 4 and Figure-S3). Rarefaction did, however, have a great influence on U-Unifrac values computed on OTU datasets (r = 0.82–0.87, Figure-S3).
Table 2

Correlation between beta-diversity metrics between OTU- and ASV-based datasets across every rarefaction level tested. Correlations were assessed using Mantel tests based on Spearman’s correlation coefficient, which is indicated for each correlation (all P values were < 0.001). Correlations < 0.9 are highlighted in bold, showing poorer agreement between OTUs and ASVs.

| Beta-diversity index | Rarefaction level | 97%-OTUs vs. ASVs | 99%-OTUs vs. ASV |
|----------------------|-------------------|-------------------|------------------|
| Jaccard              | No rarefaction    | 0.91              | 0.96             |
|                      | 3,000             | 0.93              | 0.96             |
|                      | 2,000             | 0.94              | 0.97             |
|                      | 1,000             | 0.94              | 0.97             |
| Bray-Curtis          | No rarefaction    | 0.92              | 0.96             |
|                      | 3,000             | 0.93              | 0.96             |
|                      | 2,000             | 0.94              | 0.97             |
|                      | 1,000             | 0.94              | 0.97             |
| U-Unifrac            | No rarefaction    | 0.80              | 0.82             |
|                      | 3,000             | 0.83              | 0.89             |
|                      | 2,000             | 0.84              | 0.90             |
|                      | 1,000             | 0.86              | 0.91             |
| W-Unifrac            | No rarefaction    | 0.93              | 0.95             |
|                      | 3,000             | 0.94              | 0.95             |
|                      | 2,000             | 0.94              | 0.95             |
|                      | 1,000             | 0.93              | 0.95             |

Figure 4: Change in Bray-Curtis and W-Unifrac beta-diversity after rarefaction to 2,000 sequences per sample within the ASV, 97%- and 99%-OTUs datasets across all sediment, seston and mussel microbiomes. The agreement of ranks of beta-diversities before and after rarefaction were assessed using a Mantel test based on Spearman’s coefficient of correlation, which result is indicated over each plot. The diagonal is marked by the red line. The same correlations on other beta-diversity metrics (Jaccard and U-Unifrac) is available in Figure-S3.

The selection of beta-diversity index had the strongest impact on detecting a biological signal (i.e. type of community, sampling river and sampling site). Random Forests tests estimated that the choice of beta-diversity index had a 7-15x higher contribution in detecting a biological signal than the choice of ASV vs. OTU pipeline, and both had greater contributions than rarefaction (Figure-S4). For instance, PERMANOVA R-square values varied by a six-fold factor across different beta-diversity indices within the same dataset.
and biological factor (Figure-S5), while the choice of ASVs vs. OTUs dataset only varied up to 3-fold, with all else being equal.

W-Unifrac was the most reliable index for detecting biological signals across different community types (PERMANOVA's $R^2 = 0.42 \pm 0.0$), for detecting the river effect ($R^2 = 0.32 \pm 0.14$), and for the site effect ($R^2 = 0.59 \pm 0.24$) (Figure-S5). The second most efficient index was Bray-Curtis, with significant signals in the same factors ($R^2 = 0.24 \pm 0.03$, $0.34 \pm 0.14$, and $0.57 \pm 0.21$, respectively). Bray-Curtis surpassed W-Unifrac for detecting the effect of river in ASV and 99%-OTU datasets (Figure-S5).

97%-OTUs were slightly better at detecting a difference of structure across community types (PERMANOVA's $R^2 = 0.26 \pm 0.10$ vs. ASVs $R^2 = 0.24 \pm 0.11$ across all indices and rarefaction levels, Wilcoxon signed-rank test, $p < 0.05$, Fig. 5). In contrast, when detecting river and site effects, while ASVs, 97%-OTUs and 99%-OTUs gave similar detection ranges in sediment and mussel microbiomes (River effect $R^2 = 0.24 \pm 0.04$ and $0.16 \pm 0.04$, respectively), ASVs performed better within seston communities (River effect $R^2 = 0.40 \pm 0.12$ in OTUs vs. $0.47 \pm 0.12$ in ASVs, Fig. 5). Usage of 99%-OTUs did not improve the detection of any biological effect compared to 97%-OTUs (Wilcoxon signed-rank test between 97%- and 99%-OTU datasets, $p > 0.05$).

**Figure 5: Quality of detection of biological signal across sequence processing methods and rarefaction levels.** The intensity of the effect of (a) community type, (b) sampling river and (c) collection site on community structure was assessed using separated PERMANOVAs for each factor, and each combination of sequence processing method, rarefaction level and index of dissimilarity. Here results are aggregated for all dissimilarity indices computed for this study (Jaccard, Bray-Curtis, and U- and W-Unifrac); comparison of the efficiency of the different indices for detecting the same biological effects is provided in Figure-S4.

**Composition**

While Gammaproteobacteria (12.0 ± 9.0%), Cyanobacteria (11.1 ± 12.4%), Alphaproteobacteria (10.0 ± 8.0%) and Planctomycetes (15.7 ± 15.3%) were the main classes identified in all datasets and rarefaction levels used (Fig. 6), several major classes of Bacteria had their relative abundance computed on OTUs poorly correlated with abundance computed on ASVs. These classes included Bacilli (Spearman's $r = 0.34 \pm 0.21$), Clostridia ($r = 0.66 \pm 0.28$), Bacteroidia (0.75 ± 0.13), and Actinobacteria (0.81 ± 0.17). Other major classes presented a higher reproducibility across methods (0.83 ± 0.09). The class Verrucomicrobiae, which was dominant in OTU datasets (8.7 ± 14% of relative abundance), was detected at much lower level within ASVs, where it wasn't among the most common classes (1.7 ± 0.6%) (Fig. 6). Conversely, the ASV pipeline characterized Mollicutes within the mussel gut microbiome (13.4 ± 13%), which remained completely undetected in the OTU datasets (Fig. 6).

**Figure 6: Composition of mussel gut microbiome, sediment and seston communities found using OTU-based and ASV-based methodologies.** (a) The relative abundance of the 11 most common bacterial classes in datasets rarefied to 2,000 sequences/sample. (b) Heatmap representing agreement of the 20
most detected bacterial genera across methodologies and rarefaction levels. Numbers represent the percentage of those 20 genera in common across the compared treatments ('shared'); color intensity represents overall Spearman's correlation coefficient of the shared genera between treatments.

Differences across methods were more pronounced at the genus level, with only ~ 50% of the most common 30 genera shared between ASV and OTU-based datasets (excluding genera belonging to unclassified families), and correlations of the relative abundance of such shared genera generally below 0.9 (Spearman's correlation Fig. 6; Supplementary Table 2). The 30 most detected genera unique to ASV datasets included *Methyloglobus, Escherichia/Shigella* and Clostridium sensu stricto (Supplementary Table 2). Compared to the differences between ASV and OTU methodology, rarefaction and OTU similarity threshold had only a minor impact on overall class- or genus-level classification, with, for instance, a correlation of > 0.9 of abundance of the major classes and genera between unrarefied and 2,000 sequence datasets, and between 99%- and 97% OTUs. As previously observed, with alpha- and beta-diversity indices, rarefaction had a lower impact on ASV- than OTU-based datasets, with all most detected genera shared before and after rarefaction (Fig. 6).

**Discussion**

Increasing popularity of ASV based analyses of microbiomes has raised questions as to whether (i) ASVs are accurately capturing the microbial diversity in a community [24], (ii) diversity measures based on newer ASV would still be comparable to former OTU-based ones, even though OTUs are difficult to compare across studies due to their method of construction [18].

Using a dataset that contained three different microbial community types (mussel gut tissue, sediment, and seston) we tested whether alpha- and beta-diversity metrics and community composition were consistent across ASVs vs. OTUs, in addition to comparing the influence of this methodological choice to others (rarefaction, OTU threshold, choice of diversity metrics). As others have suggested [17, 22, 26], we detected significant influences of ASVs vs. OTUs on richness estimates; both on their values and the ranking of samples, inducing discrepancies in terms of their biological signals. Importantly, our study revealed that these discrepancies were accentuated by sampling bias corrected richness estimators like iChao1 and ACE, which, of all those tested, were the least correlated across the ASV and OTU datasets. This results from the significant influence of the pipeline choice on the structure of abundance of biological units, revealing lower evenness of relative abundances together with a much higher number of rare OTUs compared to ASVs. Rare OTUs significantly impact the bias-corrected richness estimators, which in certain cases cannot work correctly [33–34]. The impact of ASVs vs. OTUs on richness measures was also greater for environmental samples than for host-associated (mussel) microbiome. In a previous study, Glassman and Martiny [16] did not find a major effect of ASV vs. OTU-based pipelines on richness for bacterial communities sampled on leaf litter. Such leaf litter environments may induce a narrower selection filter compared to environments such as bulk soil [26], and thus have lower intrinsic diversity (e.g. Purahong et al. [35] found maximum richness in leaf litter was ~ 500 OTUs), facilitating a more consistent measure of richness. Here, mussel gut microbiomes may be comparable to leaf litter in
that being under the influence of its host, bacterial diversity within mussels is lower than in the surrounding environment, and is typically dominated by a low number of biological units [36–37]. While this could potentially explain a relatively higher robustness of richness ranks within such samples, a correlation of 0.70 in richness measures across pipelines still highlight discrepancy between estimates for many communities as, for instance, 48% of such samples presented a very high variation (> 70%) in their richness estimate between ASVs and OTUs before rarefaction.

Although OTUs cluster more dissimilar sequences into the same unit than ASVs, they can also generate higher number of spurious, low-abundance OTUs that often inflate richness measures [17, 22]. Prodan et al [22] reported that applying a cutoff within mothur on sequence abundance before constructing the OTUs was necessary for reducing the erroneous inflation of observed richness. The removal of singletons (OTUs represented by only one sequence) can drastically reduce the variation across sample replicates within the same sequencing run [38], underlining the sensitivity of OTU-based methods to sequencing errors. Adding this step may be important for OTU-based analyses in order to pre-filter potentially spurious, low abundance sequences that could induce the construction of numerous erroneous OTUs. Interestingly, in our study, a simple rarefaction of the OTU community table increased the correlation between ASV- and OTU-based richness estimates, and decreased OTU-based richness to a level comparable to the estimate from ASVs. Rarefaction is more likely to remove the rare biological units, which are themselves more likely to originate from sequencing errors [39]. While rarefaction is a poor way of normalizing sequence data counts that can sometimes lead to false conclusions [29], in some cases, and especially in OTU-based environmental datasets, rarefaction may reduce the overestimation of richness that arises from spurious OTUs, and have comparable effects to an abundance cutoff on OTUs or sequences.

The impact of ASV or OTU selection and rarefaction was much lower on Shannon alpha-diversity (with a correlation > 0.9 at every rarefaction level studied), and was comparable to previous comparisons on natural and mock communities [26]. Giving less weight to the rarest biological units likely compensates for biases in the different methods in terms of the number of rare biological units, making the index of Shannon more robust regarding the choice of ASVs vs. OTUs than species richness metrics. Weighted alpha-diversity indices that account for the relative proportions of each biological unit should therefore be preferred to richness metrics when assessing alpha-diversity, especially in species-rich microbial communities such as those found in soils and waters.

In terms of beta-diversity, as in previous studies we reported more consistency across analysis methods [16, 25–26], although these findings were dependent on the particular beta-diversity metric. Indeed, in our study, we found that the choice of the beta diversity index was ten times more important than the choice of ASVs vs. OTUs in detecting a biological signal. The most powerful indices to detect the expected biological signal were W-Unifrac and Bray-Curtis, which performed 1.3–1.9 times better than U-Unifrac or Jaccard in detecting signals such as sample type, collection river and collection site. Surprisingly, the unweighted index of Unifrac (U-Unifrac) exhibited the lowest efficiency at detecting all patterns, for both the ASV and OTU datasets. This further supports the hypothesis that rare biological units may be
erroneous and obscure the signal produced by U-Unifrac, which is highly sensitive to rare biological units [40]. It also underlines the overall more suitable behavior of relative abundance weighted indices for the detection of ecological patterns [31]. Nearing et al. [17] determined Bray-Curtis, U-Unifrac and W-Unifrac dissimilarities on host-associated and soil datasets using several ASV based pipelines and one OTU based pipeline. They also observed that U-Unifrac correlated poorly across pipelines, while the weighted indices were highly consistent. We found that the most consistent index was W-Unifrac, with a near-perfect correlation between ASV and OTU datasets.

While 97% sequence similarity defined OTUs were better at detecting broader patterns in beta-diversity (e.g. differences between sample types), ASVs were more efficient at detecting more subtle effects (e.g. differences between rivers and sites within the seston communities). The use of a 99% threshold for defining the OTUs did not improve their efficiency at detecting subtle effects, suggesting that the better detection of such effects when using ASVs may not simply be because of a higher resolution compared to 97% OTUs.

In contrast to the low influence of analysis method on patterns in beta-diversity, there were major difference between ASVs and OTUs in terms of community composition, both at the genus and class level. Similar discrepancies were highlighted by Straub et al. [26], who found that the relative abundance of several taxa were poorly reproduced across both methodologies. We found that the OTU threshold and the level of rarefaction had only a marginal effect on the identity and the relative abundance of major classes and genera. ASVs, however, detected Mollicutes in the mussel gut microbiome, which were absent in the OTU datasets. In contrast, OTUs detected much higher amounts of Verrucomicrobiae in the seston and sediment samples than were detected by the ASV approach. These differences are surprising given that both OTU and ASV approaches classified sequences to the same database. In this specific case, while the cultivation of Mollicutes is difficult [41], they have been successfully isolated from the digestive gland of various hosts, including aquatic invertebrates [42], making their detection within the gut of the mussels studied here, plausible. On the other hand, Verrucomicrobia have been also identified within soil and water environments and they presence is not surprising here.

Conclusions - Recommendations

- We advise to compute alpha-diversity estimates using ASVs, especially in species-rich, environmental datasets, as OTUs tend to overestimate richness in these communities. This recommendation is also supported by previous studies [17, 26].
- Richness indices that correct for sampling bias are very sensitive to rare biological units, and should be avoided on OTU-based datasets if no rarefaction or filtering of erroneous or rare biological units has been performed.
- Relative abundance weighted alpha- and beta-diversity indices are more reliable, overall, than ones based on presence-absence only, and should be preferred. This has been highlighted previously [26].
If performing OTU-based analyses, rarefaction can in some cases remove the biological noise and therefore be desirable for alpha-diversity assessment. Other methods for the removal of erroneous OTUs may also be considered [43].

ASVs and OTUs show only a low taxonomic congruence within the same datasets, as shown previously [22]. Before switching from OTUs to ASVs, we therefore recommend to compare both methodologies to assess potential taxonomic bias.

**Declarations**

**Ethics approval and consent to participate**

Not applicable

**Consent for publication**

Not applicable

**Availability of data and materials**

Bacterial sequences for this paper will be available upon publication in the SRA database (http://www.ncbi.nlm.nih.gov/bioproject/740316), under the BioSample numbers SAMN19838587-SAMN19838803.

**Competing interest**

The authors declare that they have no competing interests.

**Funding**

Sample collection, sequencing and analysis were funded by the National Science Foundation (award #1831512).

**Authors’ contribution**

MC and CRJ conceptualized the study. MM and MC processed the samples and prepared the libraries. MC analyzed the data and wrote the manuscript. MM, CRJ and SV reviewed and edited the manuscript. All authors read and approved the final manuscript.

**Acknowledgments**

Carla Atkinson, Garrett Hopper, Jamie Bucholz, Irene Sanchez Gonzalez, and Megan Kubala at the University of Alabama coordinated and assisted with field collection of samples. Lauren Lawson assisted with DNA extractions of mussel samples. Funding was provided from National Science Foundation award DEB 1831531 to C.R.J.
References

1. Barea JM. Future challenges and perspectives for applying microbial biotechnology in sustainable agriculture based on a better understanding of plant-microbiome interactions. J soil sci plant nutr. 2015; ahead:0–0.

2. Bello MGD, Knight R, Gilbert JA, Blaser MJ. Preserving microbial diversity. Science. 2018;362:33–4.

3. McFall-Ngai M, Hadfield MG, Bosch TCG, Carey HV, Domazet-Lošo T, Douglas AE, et al. Animals in a bacterial world, a new imperative for the life sciences. Proc Natl Acad Sci U S A. 2013;110:3229–36.

4. Nikolaki S, Tsiamis G. Microbial diversity in the era of omic technologies. Biomed Res Int. 2013;2013:958719.

5. Gibbons SM, Gilbert JA. Microbial diversity–exploration of natural ecosystems and microbiomes. Curr Opin Genet Dev. 2015;35:66–72.

6. Gardner PP, Watson RJ, Morgan XC, Draper JL, Finn RD, Morales SE, et al. Identifying accurate metagenome and amplicon software via a meta-analysis of sequence to taxonomy benchmarking studies. PeerJ. 2019;7:e6160.

7. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol. 2013;79:5112–20.

8. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016;13:581–3.

9. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. 2014;15:R46.

10. Ounit R, Wanamaker S, Close TJ, Lonardi S. CLARK: fast and accurate classification of metagenomic and genomic sequences using discriminative k-mers. BMC Genomics. 2015;16:236.

11. Eren AM, Maignien L, Sul WJ, Murphy LG, Grim SL, Morrison HG, et al. Oligotyping: Differentiating between closely related microbial taxa using 16S rRNA gene data. Methods Ecol Evol. 2013;4:1111–9.

12. Siegwald L, Touzet H, Lemoine Y, Hot D, Audebert C, Caboche S. Assessment of common and emerging bioinformatics pipelines for targeted metagenomics. PLoS One. 2017;12:e0169563.

13. Stackebrandt E, Goebel BM. Taxonomic note: A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. Int J Syst Evol Microbiol. 1994;44:846–9.

14. Schloss PD, Westcott SL. Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. Appl Environ Microbiol. 2011;77:3219–26.

15. Mysara M, Njima M, Leys N, Raes J, Monsieurs P. From reads to operational taxonomic units: an ensemble processing pipeline for MiSeq amplicon sequencing data. Gigascience. 2017;6:1–10.

16. Glassman SI, Martiny J BH. Broadscale ecological patterns are robust to use of exact sequence variants versus operational taxonomic units. mSphere. 2018;3. doi:10.1128/msphere.00148-18.
17. Nearing JT, Douglas GM, Comeau AM, Langille MGI. Denoising the Denoisers: an independent evaluation of microbiome sequence error-correction approaches. PeerJ. 2018;6:e5364.

18. Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. ISME J. 2017;11:2639–43.

19. Berry MA, White JD, Davis TW, Jain S, Johengen TH, Dick GJ, et al. Are oligotypes meaningful ecological and phylogenetic units? A case study of Microcystis in freshwater lakes. Front Microbiol. 2017;8:365.

20. Caruso V, Song X, Asquith M, Karstens L. Performance of microbiome sequence inference methods in environments with varying biomass. mSystems. 2019;4. doi:10.1128/mSystems.00163-18.

21. Needham DM, Sachdeva R, Fuhrman JA. Ecological dynamics and co-occurrence among marine phytoplankton, bacteria and myoviruses shows microdiversity matters. ISME J. 2017;11:1614–29.

22. Prodan A, Tremaroli V, Brolin H, Zwinderman AH, Nieuwdorp M, Levin E. Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. PLoS One. 2020;15:e0227434.

23. Xue Z, Kable ME, Marco ML. Impact of DNA sequencing and analysis methods on 16S rRNA gene bacterial community analysis of dairy products. mSphere. 2018;3. doi:10.1128/mSphere.00410-18.

24. Joos L, Beirinckx S, Haegeman A, Debode J, Vandecasteele B, Baeyen S, et al. Daring to be differential: metabarcoding analysis of soil and plant-related microbial communities using amplicon sequence variants and operational taxonomical units. BMC Genomics. 2020;21:733.

25. Moossavi S, Atakora F, Fehr K, Khafipour E. Biological observations in microbiota analysis are robust to the choice of 16S rRNA gene sequencing processing algorithm: case study on human milk microbiota. BMC Microbiol. 2020;20:290.

26. Straub D, Blackwell N, Langarica-Fuentes A, Peltzer A, Nahnsen S, Kleindienst S. Interpretations of environmental microbial community studies are biased by the selected 16S rRNA (gene) amplicon sequencing pipeline. Front Microbiol. 2020;11:550420.

27. Auer L, Mariadassou M, O’Donohue M, Klopp C, Hernandez-Raquet G. Analysis of large 16S rRNA Illumina data sets: Impact of singleton read filtering on microbial community description. Mol Ecol Resour. 2017;17:e122–32.

28. Chao A, Chiu C-H, Jost L. Unifying species diversity, phylogenetic diversity, functional diversity, and related similarity and differentiation measures through hill numbers. Annu Rev Ecol Evol Syst. 2014;45:297–324.

29. Willis AD. Rarefaction, alpha diversity, and statistics. Front Microbiol. 2019;10:2407.

30. Chiarello M, Auguet J-C, Bettarel Y, Bouvier C, Claverie T, Graham NAJ, et al. Skin microbiome of coral reef fish is highly variable and driven by host phylogeny and diet. Microbiome. 2018;6:147.

31. Fukuyama J, McMurdie PJ, Dethlefsen L, Relman DA, Holmes S. Comparisons of distance methods for combining covariates and abundances in microbiome studies. Pac Symp Biocomput. 2012;21:213–24.
32. Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. UniFrac: an effective distance metric for microbial community comparison. ISME J. 2011;5:169–72.

33. Chao A, Chiu C-H. Nonparametric estimation and comparison of species richness. eLS. 2016;:1–11.

34. Willie J, Petre C-A, Tagg N, Lens L. Evaluation of species richness estimators based on quantitative performance measures and sensitivity to patchiness and sample grain size. Acta Oecol (Montrouge). 2012;45:31–41.

35. Purahong W, Wubet T, Lentendu G, Schloter M, Pecyna MJ, Kapturska D, et al. Life in leaf litter: novel insights into community dynamics of bacteria and fungi during litter decomposition. Mol Ecol. 2016;25:4059–74.

36. McCauley M, Chiarello M, Atkinson CL, Jackson CR. Gut microbiomes of freshwater mussels (Unionidae) are taxonomically and phylogenetically variable across years but remain functionally stable. Microorganisms. 2021;9:411.

37. Weingarten EA, Atkinson CL, Jackson CR. The gut microbiome of freshwater Unionidae mussels is determined by host species and is selectively retained from filtered seston. PLoS One. 2019;14:e0224796.

38. Wen C, Wu L, Qin Y, Van Nostrand JD, Ning D, Sun B, et al. Evaluation of the reproducibility of amplicon sequencing with Illumina MiSeq platform. PLoS One. 2017;12:e0176716.

39. Brown SP, Veach AM, Rigdon-Huss AR, Grond K, Lickteig SK, Lothamer K, et al. Scraping the bottom of the barrel: are rare high throughput sequences artifacts? Fungal Ecol. 2015;13:221–5.

40. Chen J, Bittinger K, Charlson ES, Hoffmann C, Lewis J, Wu GD, et al. Associating microbiome composition with environmental covariates using generalized UniFrac distances. Bioinformatics. 2012;28:2106–13.

41. Lehmann D, Jouette S, Olivieri F, Laborde S, Rofel C, Simon E, et al. Novel sample preparation method for molecular detection of Mollicutes in cell culture samples. J Microbiol Methods. 2010;80:183–9.

42. Ramírez AS, Vega-Orellana OM, Viver T, Poveda JB, Rosales RS, Poveda CG, et al. First description of two moderately halophilic and psychrotolerant Mycoplasma species isolated from cephalopods and proposal of Mycoplasma marinum sp. nov. and Mycoplasma todarodis sp. nov. Syst Appl Microbiol. 2019;42:457–67.

43. Frøslev TG, Kjøller R, Bruun HH, Ejrnæs R, Brunbjerg AK, Pietroni C, et al. Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. Nat Commun. 2017;8:1188.

44. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol. 2009;75:7537–41.

45. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One. 2013;8:e61217.

46. Marcon E, Hérault B. entropart: An R Package to Measure and Partition Diversity. J Stat Softw. 2015;67. doi:10.18637/jss.v067.i08.
47. Dixon P. VEGAN, a package of R functions for community ecology. J Veg Sci. 2003;14:927–30.
48. Jost L. Partitioning diversity into independent alpha and beta components. Ecology. 2007;88:2427–39.
49. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol. 2006;72:5069–72.
50. Janssen S, McDonald D, Gonzalez A, Navas-Molina JA, Jiang L, Xu ZZ, et al. Phylogenetic placement of exact amplicon sequences improves associations with clinical information. mSystems. 2018;3. doi:10.1128/msystems.00021-18.
51. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Author Correction: Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37:1091.
52. Paradis E, Claude J, Strimmer K. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics. 2004;20:289–90.
53. Wei T, Simko V, Levy M, Xie Y, Jin Y, Zemla J. Package ‘corrplot.’ Statistician. 2017;56:24.
54. Liaw A, Wiener M. Classification and regression by randomForest. R news. 2002;2:18–22.
55. Boulesteix A-L, Janitza S, Kruppa J, König IR. Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics: Random forests in bioinformatics. Wiley Interdiscip Rev Data Min Knowl Discov. 2012;2:493–507.

Figures
Figure 1

Effect of sequence processing methodology on the rank abundance of biological units. The relative abundance of ASVs (obtained from DADA2) and OTUs defined by 97% or 99% similarity (obtained from MOTHUR). Units were sorted according to their rank of abundance and represented separately for each community type (a: mussel gut microbiome, b: sediment, c: seston). Inserts focus on the 500 most
abundant biological units. The evenness of the relative abundance of the biological units were computed using Bulla’s O and are displayed on each plot.

**Figure 2**

Effect of sequence processing methodology on alpha-diversity metrics for microbiome analyses. Measure of taxonomic richness (expressed as number of OTUs or ASVs and Shannon alpha-diversity depending on the methodology used, namely ASV, 97%-OTU and 99%-OTU, at 3 different levels of
rarefaction (1,000; 2,000 and 3,000 sequences per sample) within each sample type studied (a-c). To ease the visualization of differences across methods and rarefaction levels, the y-axis of plots has been log transformed. The effect of methodological choices on other alpha-diversity metrics is available in Figure S1.

**Figure 3**

(a) Mussel microbiome

|                  | Bogue Chitto | Sipsey | Butternutshere | Bear Creek | Paint Rock | Duck |
|------------------|--------------|--------|----------------|------------|------------|------|
| ASVs Richness    | ab abc a c abc bc |        |                |            |            |      |
| 97%-OTUs Richness| a a a a a a a a a a |        |                |            |            |      |
| 99%-OTUs Richness| a a a a a a a a a a |        |                |            |            |      |
| ASVs Shannon     | ab abc a c abc bc |        |                |            |            |      |
| 97%-OTUs Shannon | a a a a a a a a b ab b |        |                |            |            |      |
| 99%-OTUs Shannon | a a a a a a a a a a |        |                |            |            |      |

(b) Sediment

|                  |             |           |               |            |            |      |
|------------------|-------------|-----------|---------------|------------|------------|------|
| ASVs Richness    | a a a a a a a a |        |                |            |            |      |
| 97%-OTUs Richness| a a a a a a a a a a |        |                |            |            |      |
| 99%-OTUs Richness| ab ab a ab b ab ab |        |                |            |            |      |
| ASVs Shannon     | a a a a a a a a a a |        |                |            |            |      |
| 97%-OTUs Shannon | a a a a a a a a a a |        |                |            |            |      |
| 99%-OTUs Shannon | a a a a a a a a a a |        |                |            |            |      |

(c) Seston

|                  |             |           |               |            |            |      |
|------------------|-------------|-----------|---------------|------------|------------|------|
| ASVs Richness    | a a a a a a a a |        |                |            |            |      |
| 97%-OTUs Richness| ab a b b ab ab |        |                |            |            |      |
| 99%-OTUs Richness| ab a bc c ab abc |        |                |            |            |      |
| ASVs Shannon     | a a ab b a ab b |        |                |            |            |      |
| 97%-OTUs Shannon | ab a abc c ab bc |        |                |            |            |      |
| 99%-OTUs Shannon | ab a abc c ab bc |        |                |            |            |      |
Comparison of observed richness and Shannon alpha-diversity across sampling rivers and sequence processing methods. For each community type, (a) mussel microbiome, (b) sediment and (c) Seston, the average alpha-diversity value per river were represented by the size and color intensity of the corresponding circle. Letters representing whether diversity values of different rivers were significantly different or not, were displayed (Kruskal-Wallis’ pairwise post-hoc tests, P<0.05). For instance, ASV richness of mussel microbiome sampled on Bear Creek (‘c’) was significantly lower from than the one of mussels collected on river Buttahatchee (‘a’), but did not differ from the ones collected on Paint Rock (‘abc’). To facilitate comparison across panels, all metrics were scaled between 0 and 1.

**Figure 4**

Change in Bray-Curtis and W-Unifrac beta-diversity after rarefaction to 2,000 sequences per sample within the ASV-, 97%- and 99%-OTUs datasets across all sediment, seston and mussel microbiomes. The agreement of ranks of beta-diversities before and after rarefaction were assessed using a Mantel test based on Spearman’s coefficient of correlation, which result is indicated over each plot. The diagonal is marked by the red line. The same correlations on other beta-diversity metrics (Jaccard and U-Unifrac) is available in Figure-S3.
Figure 5

Quality of detection of biological signal across sequence processing methods and rarefaction levels. The intensity of the effect of (a) community type, (b) sampling river and (c) collection site on community structure was assessed using separated PERMANOVAs for each factor, and each combination of sequence processing method, rarefaction level and index of dissimilarity. Here results are aggregated for
all dissimilarity indices computed for this study (Jaccard, Bray-Curtis, and U- and W-Unifrac); comparison of the efficiency of the different indices for detecting the same biological effects is provided in Figure-S4.

**Figure 6**

Composition of mussel gut microbiome, sediment and seston communities found using OTU-based and ASV-based methodologies. (a) The relative abundance of the 11 most common bacterial classes in datasets rarefied to 2,000 sequences/sample. (b) Heatmap representing agreement of the 20 most detected bacterial genera across methodologies and rarefaction levels. Numbers represent the percentage of those 20 genera in common across the compared treatments (‘shared’); color intensity represents overall Spearman’s correlation coefficient of the shared genera between treatments.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryfigures.pdf
- Supplementarytables.pdf