



# Primer, Pipelines, Parameters: Issues in 16S rRNA Gene Sequencing

Isabel Abellan-Schneyder,<sup>a</sup>  Monica S. Matchado,<sup>b</sup> Sandra Reitmeier,<sup>a</sup> Alina Sommer,<sup>a</sup> Zeno Sewald,<sup>a</sup> Jan Baumbach,<sup>b,c,d</sup>  Markus List,<sup>b</sup>  Klaus Neuhaus<sup>a</sup>

<sup>a</sup>Core Facility Microbiome, ZIEL—Institute for Food & Health, Technische Universität München, Freising, Germany

<sup>b</sup>Chair of Experimental Bioinformatics, TUM School of Life Sciences Weihenstephan, Technische Universität München, Freising, Germany

<sup>c</sup>Computational Biomedicine Lab, Department of Mathematics and Computer Science, University of Southern Denmark, Odense, Denmark

<sup>d</sup>Chair of Computational Systems Biology, University of Hamburg, Hamburg, Germany

**ABSTRACT** Short-amplicon 16S rRNA gene sequencing is currently the method of choice for studies investigating microbiomes. However, comparative studies on differences in procedures are scarce. We sequenced human stool samples and mock communities with increasing complexity using a variety of commonly used protocols. Short amplicons targeting different variable regions (V-regions) or ranges thereof (V1-V2, V1-V3, V3-V4, V4, V4-V5, V6-V8, and V7-V9) were investigated for differences in the composition outcome due to primer choices. Next, the influence of clustering (operational taxonomic units [OTUs], zero-radius OTUs [zOTUs], and amplicon sequence variants [ASVs]), different databases (GreenGenes, the Ribosomal Database Project, Silva, the genomic-based 16S rRNA Database, and The All-Species Living Tree), and bioinformatic settings on taxonomic assignment were also investigated. We present a systematic comparison across all typically used V-regions using well-established primers. While it is known that the primer choice has a significant influence on the resulting microbial composition, we show that microbial profiles generated using different primer pairs need independent validation of performance. Further, comparing data sets across V-regions using different databases might be misleading due to differences in nomenclature (e.g., *Enterorhabdus* versus *Adlercreutzia*) and varying precisions in classification down to genus level. Overall, specific but important taxa are not picked up by certain primer pairs (e.g., *Bacteroidetes* is missed using primers 515F-944R) or due to the database used (e.g., *Acetatifactor* in GreenGenes and the genomic-based 16S rRNA Database). We found that appropriate truncation of amplicons is essential and different truncated-length combinations should be tested for each study. Finally, specific mock communities of sufficient and adequate complexity are highly recommended.

**IMPORTANCE** In 16S rRNA gene sequencing, certain bacterial genera were found to be underrepresented or even missing in taxonomic profiles when using unsuitable primer combinations, outdated reference databases, or inadequate pipeline settings. Concerning the last, quality thresholds as well as bioinformatic settings (i.e., clustering approach, analysis pipeline, and specific adjustments such as truncation) are responsible for a number of observed differences between studies. Conclusions drawn by comparing one data set to another (e.g., between publications) appear to be problematic and require independent cross-validation using matching V-regions and uniform data processing. Therefore, we highlight the importance of a thought-out study design including sufficiently complex mock standards and appropriate V-region choice for the sample of interest. The use of processing pipelines and parameters must be tested beforehand.

**KEYWORDS** 16S rRNA gene sequencing, amplicon sequencing, variable regions, clustering, bioinformatic settings, microbiome, databases, mock communities

**Citation** Abellan-Schneyder I, Matchado MS, Reitmeier S, Sommer A, Sewald Z, Baumbach J, List M, Neuhaus K. 2021. Primer, pipelines, parameters: issues in 16S rRNA gene sequencing. *mSphere* 6:e01202-20. <https://doi.org/10.1128/mSphere.01202-20>.

**Editor** Susannah Green Tringe, U.S. Department of Energy Joint Genome Institute

**Copyright** © 2021 Abellan-Schneyder et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Klaus Neuhaus, [neuhaus@tum.de](mailto:neuhaus@tum.de).

**Received** 26 November 2020

**Accepted** 4 February 2021

**Published** 24 February 2021

The human gut microbiome is a complex environment hosting a large number of different bacteria. A cost-effective method to determine the bacterial composition of, e.g., human fecal samples is to sequence amplicons targeting the 16S rRNA gene. Microbial compositions of diverse environments, which are influenced by different factors or conditions (e.g., sampling time point, targeted rRNA region, response to health or disease, sequencing strategy, machinery, depth, and read lengths), were also studied with this method (1–7).

The 16S rRNA gene spans about 1,500 bp and is structured in highly conserved regions interspersed with nine variable regions (V-regions), V1 to V9 (8, 9). The conserved regions can be used for primer binding and thus allow for capturing a greater number of different bacterial taxa, sometimes including or not including archaea, while the variable regions permit the discrimination of these taxa within different microbial environments (10). However, differences between the conserved regions and, therefore, differences in primer annealing result in an unequal amplification of bacteria present in a sample (11). Depending on the particular V-region that was targeted, differences in the sequencing results and taxonomic outcome occurred, which led to misinterpretation (12, 13). Further, not every variable region has the same sensitivity, i.e., allowing separation of closely related taxa (14). Concerning archaea, the applicability of certain primer pairs has been covered well in previous studies (12, 15, 16).

Second-generation sequencers, e.g., Illumina's MiSeq, enable sequencing of amplicons up to 600 bp with high accuracy. This length allows targeting about one to three adjacent variable regions of the 16S rRNA gene using "universal" primers for the conserved regions. In a subsequent PCR, sequencing adapters are added to the amplicons (17). After a cleanup step, the amplicon libraries are sequenced. The resulting reads are used to analyze similarities and differences between samples with different microbial compositions (e.g., alpha- and beta-diversity) (18). In contrast, full-length 16S rRNA gene sequencing is possible by using third-generation sequencers, for instance, Oxford Nanopore MinION (19) and the PacBIOs Sequel (20), which were introduced in 2009 and 2008, respectively. The greatest advantage is the long read length (up to 10,000 bp) and sequencing on a single-molecule level in a short time. These long reads enable an improved identification of bacterial taxa, as shown in several recent studies (21–27). Nevertheless, significant drawbacks include the relatively high error rate (up to 15% per sequence) (28, 29), limited applicability in high-throughput studies, higher general costs, and even less standardization of protocols and analysis pipelines. However, despite the widespread use of 16S rRNA gene sequencing, there is a need to better understand the differences between the targeted region and the data analysis pipeline chosen in amplicon sequencing of the 16S rRNA genes.

For short-amplicon sequencing, a literature survey showed that the regions V1-V2/V3 (30, 31), V3-V4/V5 (32–34), and V4 (35, 36) are most commonly used. However, the taxonomic classification differs considerably when targeting different variable regions (37), affecting attempts to perform cross-study comparison and leading to further biases in compositional analysis, where short-amplicon primers are not as universal as desired (11, 38). Since the taxonomic resolution seems to differ for some phyla for different variable regions (39), closely related bacterial species and genera might be indistinguishable (40). Moreover, the choice of bioinformatic processing pipelines and analysis tools is known to influence the results (41–44). Different 16S rRNA gene-specific taxonomic classification methods, such as Mothur (45), Qiime (46), Qiime2 (47), DADA2 (48), and others, were developed. During data processing, sequences are clustered into operational taxonomic units (OTUs) at a threshold of 97% sequence similarity. Sequence representatives, i.e., sequences with the least mismatches to other sequences in a cluster, are used for taxonomic assignment. Amplicon sequence variants (ASVs) or zero-radius OTUs (zOTUs) have been suggested as alternatives to OTUs (48, 49), as they correct for sequencing errors by different denoising approaches. In contrast to OTUs, these clusters are supposed to contain reads originating only from the same bacterial species, enabling a cross-study comparison (49, 50). In any case, after clustering, sequences are classified for

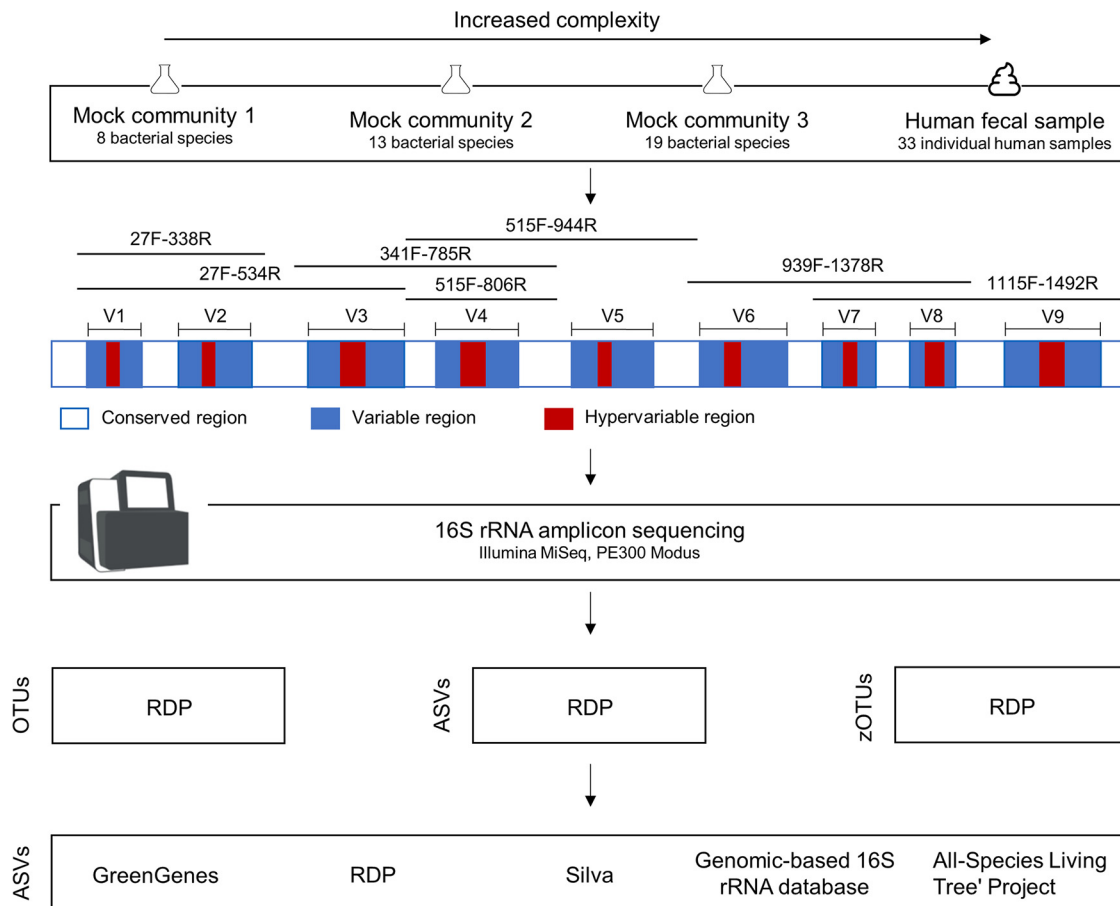
taxonomic assignment using databases of known 16S rRNA gene sequences, e.g., GreenGenes (GG) (51), the Ribosomal Database Project (RDP) (52), Silva (53), the genomic-based 16S rRNA Database (GRD) (54), or The All-Species Living Tree (LTP) (55). Not only different pipelines and reference databases but also settings of a given pipeline influence the results and are an often-overlooked bias in microbiome studies (42, 56–58). Nevertheless, some biases occurring in 16S rRNA gene amplicon sequencing have already been addressed in the past. Well-studied biasing factors, for instance, include sampling and storage procedures (59–63), DNA extraction methods (64–68), choice of variable region and primers (12, 36, 69–72), library preparation and sequencing strategies (73–76), and sequence data processing, including denoising, taxonomic classification, and the use of distinct bioinformatic tools (42, 56–58). Further, the use of negative controls and mock communities as internal standards to detect contamination or aberrancies in the sequencing results was proposed (77–79).

In this study, we joined several of these separate issues to raise awareness that the combination of primer sequence choice, clustering methods, reference database, and analysis parameters must be considered thoroughly to avoid increased bias. Thus, we created a large benchmark data set of 16S rRNA gene amplicon sequences, targeting different V-regions of the 16S rRNA gene, and systematically tested different software tools with different sets of parameters for the analysis. We sequenced three mock communities of increasing complexity with known composition, along with complex human fecal samples for comparison.

## RESULTS

We systematically assessed the global influence of multiple parameters in mock communities of known composition and in human samples (Fig. 1). First, the choice of primers targeting different variable regions of the 16S rRNA gene was evaluated. We show that primer choice influences the taxonomic composition, visible in a multidimensional scaling (MDS) plot of samples originating from the same donor (Fig. 2). Second, we investigated how, and in what magnitude, the use of different clustering approaches and taxonomy assignment methods influences the results for the classification of bacterial taxonomies.

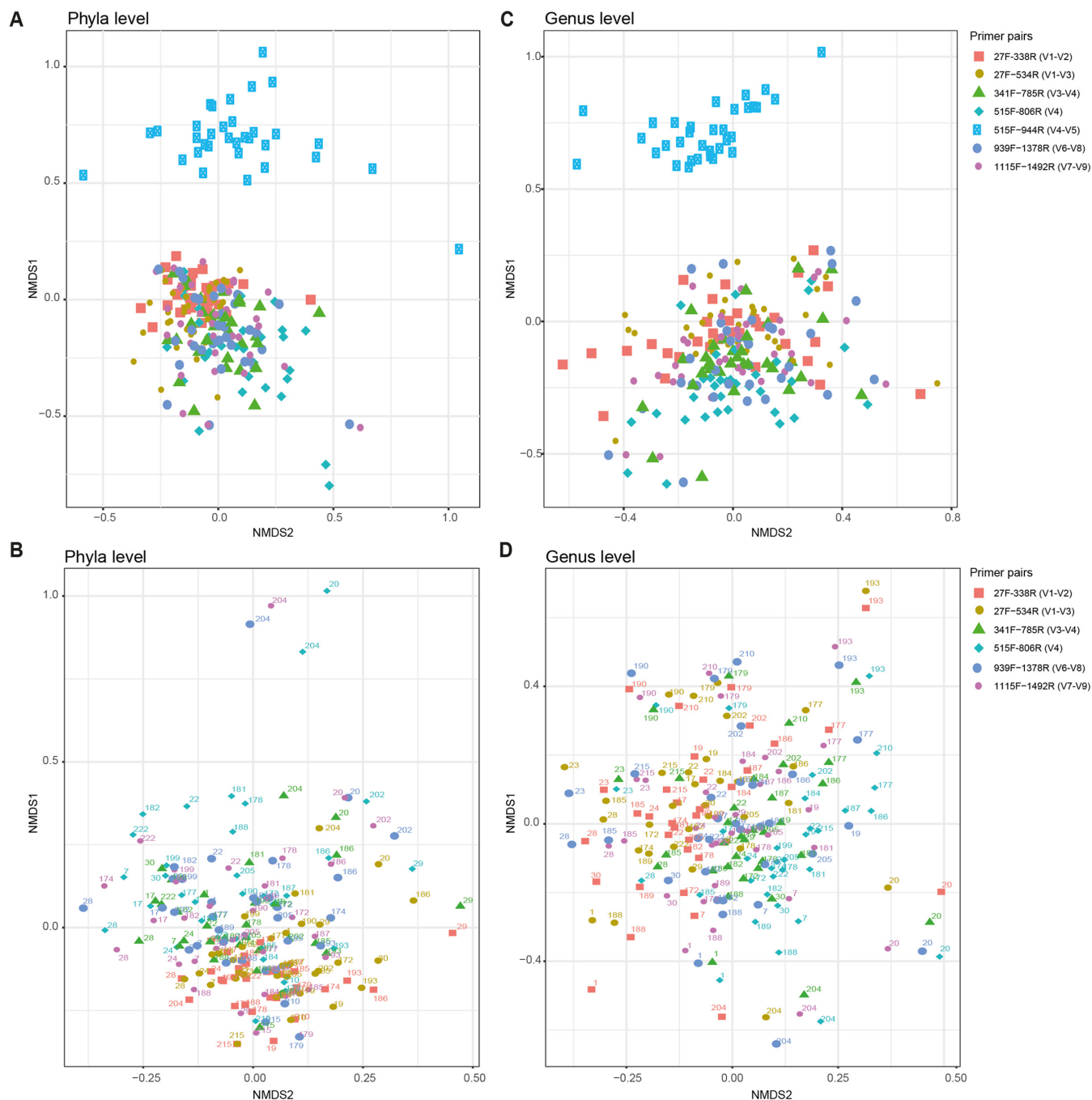
**Primer choice influences the estimated microbial composition.** A set of different 16S rRNA gene sequencing primer pairs covering one, two, or three of the variable regions V1 to V9 is commonly used for the analysis of microbial compositions. Depending on the input material (e.g., human gut samples, water analysis, sludge, food research, etc.), different primer pairs are used. In this study, we investigated seven different primer pairs, 27F-338R (V1-V2), 27F-534R (V1-V3), 341F-785R (V3-V4), 515F-806R (V4), 515F-944R (V4-V5), 939F-1378R (V6-V8), and 1115F-1492R (V7-V9), for the analysis of human gut samples and mock communities (Fig. 1 and Table 1). The use of different primer pairs led to primer-specific and not mainly donor-specific clustering of human stool samples (Fig. 2). These differences varied according to the analyzed taxonomic level. Differences were found to be less pronounced at higher taxonomic levels, e.g., phylum level compared to genus level (Fig. 2A and C). When analyzing samples from the same human donor but sequenced using different primer pairs, some taxa are unique for certain primer pairs. For instance, when analyzing human sample 1 (Fig. 3), *Verrucomicrobia* was detected only when using 341F-785R (V3-V4), 515F-806R (V4), 939F-1378R (V6-V8), and 1115F-1492R (V7-V9) primers and not 27F-338R (V1-V2), 27F-534R (V1-V3), or 515F-944R (V4-V5). Comparisons of samples derived from the same human donor but sequenced using different primer pairs become even more difficult at the genus level (see Fig. S2 in the supplemental material). This was mainly due to differences in the prevalence of genera when using different V-regions. A large number of reads were not classified down to genus level in either one or several V-regions and were thus considered “unknown.” Importantly, the 515F-944R (V4-V5) primer pair seemed to produce results with only a few overlaps with other primer pairs (Fig. 2) and displayed a low abundance of *Bacteroidetes* (Fig. 3; Fig. S2). We analyzed whether this was due to a much lower theoretical coverage of known bacterial species. Therefore, all primers were



**FIG 1** Overview of the analysis strategies used in this study. DNAs from different sample types with increasing complexity (i.e., 3 mock communities and 33 human stool samples) were extracted. Amplicons were generated using different primer pairs targeting different V-regions and sequenced on an Illumina MiSeq. Afterwards, the impacts of different clustering approaches and reference databases on the microbial profiles were investigated.

evaluated *in silico* for their theoretical coverage on all bacterial genera using the Silva database. While the theoretical coverage for 515F-944R (V4-V5) primers was lower than for the primer pairs 27F-338R (V1-V2), 27F-534R (V1-V3), 341F-785R (V3-V4), and 515F-806R (V4), we found the theoretical coverage for primer pairs 939F-1378R (V6-V8) and 1115F-1492R (V7-V9) to be even lower (Table S2). Thus, we believe that the low coverage of *Bacteroidetes* is the main reason for primer pair 515F-944R (V4-V5) to form an outlier.

**Clustering approaches have minor influence on taxonomic profiles.** In addition to the 97%-identity OTU approach, ASV clustering gained a lot of attention in the latest studies (43). Due to its improved resolution and thus better comparability of results between different studies, it is nowadays a popular and often favored method. In this study, we tested whether different clustering approaches have an influence on the assigned taxonomic profiles for the ZIEL-I mock community. Thus, we compared ASVs, zOTUs, and OTUs. Overall, the clustering methodology seemed to have only a minor effect on the assigned taxonomic composition compared to the effect of primer choice (Fig. 4A). Again, the 515F-944R (V4-V5) primer pair showed profiles distinct from those found for all other primer pairs used, no matter which clustering was used. Differences observed for each clustering approach were mainly due to identification problems at the genus level. When using the ASV approach for clustering the data, *Bacillus* could not be classified down to genus level. In contrast, this was possible when using zOTU and OTU approaches. Similarly, *Enterococcus* was not assigned correctly by the 27F-534R (V1-V3) primer pair using the ASV approach. Overall, we found that ASVs



**FIG 2** NMDS plots for the microbiome composition of human samples. Sample similarity is shown at phylum level (A and B) and at genus level (C and D). Different primer pairs are indicated to the right for all panels. Top panels (A and C) include processing the V4-V5 region, while for the bottom panels (B and D) this region has been omitted since results using 515F-944R primers (blue squares in panels A and C) fall separately from all other clusters. Labeling of the samples in the bottom panels (B and D) is based on donor number.

performed best for most of the other genera, as differences between theoretical values and expected amounts of the distinct taxa were the smallest here (Table S3). The additional analysis of a human sample subset resulted in results comparable to those for the ZIEL-I mock community (example of one representative sample is shown in Fig. 4B). Differences in taxonomic profiles are more dependent on primer pairs used than on clustering approach. Smaller variations occurred mostly due to problems assigning genera; e.g., identification of members of the *Lachnospiraceae* family on the genus level is not possible for zOTUs when using primer pairs 515F-944R (V4-V5) and



**TABLE 1** V-region-specific forward and reverse primers and annealing temperature for 1st step PCR

V-region	Forward primer	Reverse primer	Forward sequence (5'–3')	Reverse sequence (5'–3')	Specificity	Annealing temp (°C)	Reference
V1-V2	27F	338R	AGA GTT TGA TYM TGG CTC AG	GCT GCC TCC CGT AGG AGT	Universal <sup>a</sup>	57	Salter et al. (115)
V1-V3	27F	534R	AGA GTT TGA TYM TGG CTC AG	ATT ACC GCG GCT GCT GG	Universal	57	Walker et al. (84)
V3-V4	341F	785R	CCT ACG GGN GGC WGC AG	GAC TAC HVG GGT ATC TAA TCC	Universal	55	Klindworth et al. (70)
V4	515F	806R	GTG CCA GCM GCC GCG GTA A	GGA CTA CHV GGG TWT CTA AT	Universal	53	Caporaso et al. (116)
V4-V5	515F	944R	GTG CCA GCM GCC GCG GTA A	GAA TTA AAC CAC ATG CTC	Bacterial	53	Fuks et al. (117)
V6-V8	939F	1378R	GAA TTG ACG GGG GCC CGC ACA AG	CGG TGT GTA CAA GGC CCG GGA ACG	Bacterial	58	Lebuhn et al. (118)
V7-V9	1115F	1492R	CAA CGA GCG CAA CCC T	TAC GGY TAC CTT GTT ACG ACT T	Bacterial	51	Turner et al. (119)

<sup>a</sup>Universal, binds to archaea and bacteria.

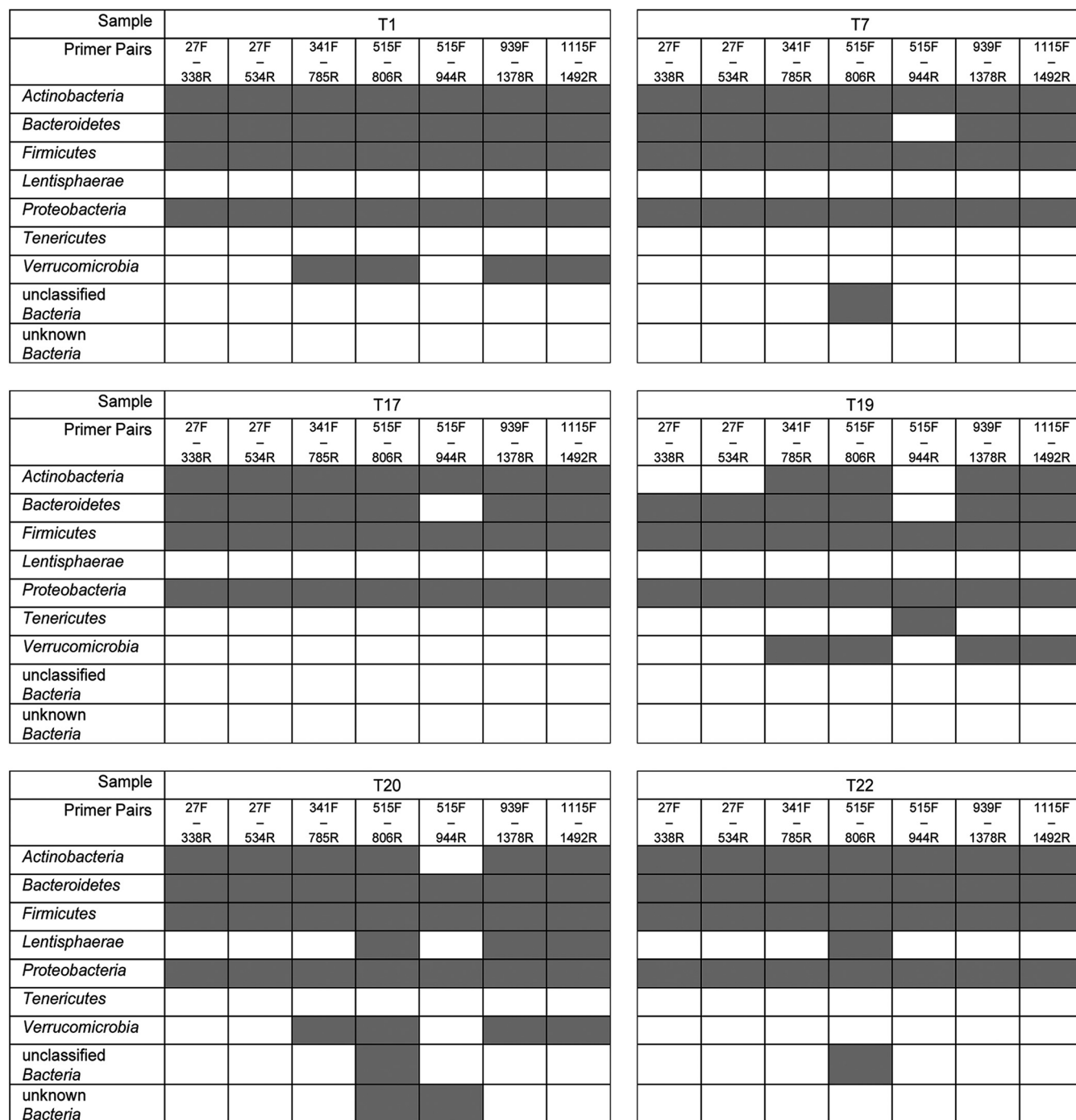
1115F-1492R (V7-V9). Still, neither OTU nor zOTU clustering caused a larger bias, and thus, the influence of clustering is limited.

**Sample taxonomies are influenced by reference databases.** Ideally, the 16S rRNA gene sequences should reflect the organism the sequence came from. However, this depends not only on the primer pairs used or how sequence data were extracted from the raw data but also on the quality of the reference database and thus the taxonomic classification. We systematically tested five different databases commonly used: GG, RDP, Silva, GRD, and LTP.

When analyzing the Zymo mock community, which includes only eight different bacteria, we observed just a few minor differences in the assigned taxonomy for different V-regions used. Further, differences were relatively minor using different reference databases in the analysis (Fig. 5A). Using RDP for primer pair 515F-806R (V4), *Bacillus* could not be classified at the genus level but was at least assigned to *Bacillales* at the family level. The classification of *Escherichia/Shigella* was most accurate when using Silva or RDP as a reference database; thus, it displayed the lowest deviation from the ideal composition of the mock community. GG could not identify *Escherichia/Shigella* and *Listeria* at the genus level and showed poor results. When using the Zymo mock community, GG might be dismissed as an inferior database, but all other parameters seemed to have no major impact. However, as a mock community of only eight bacterial species provides only limited insights, we used two further, more complex mock communities.

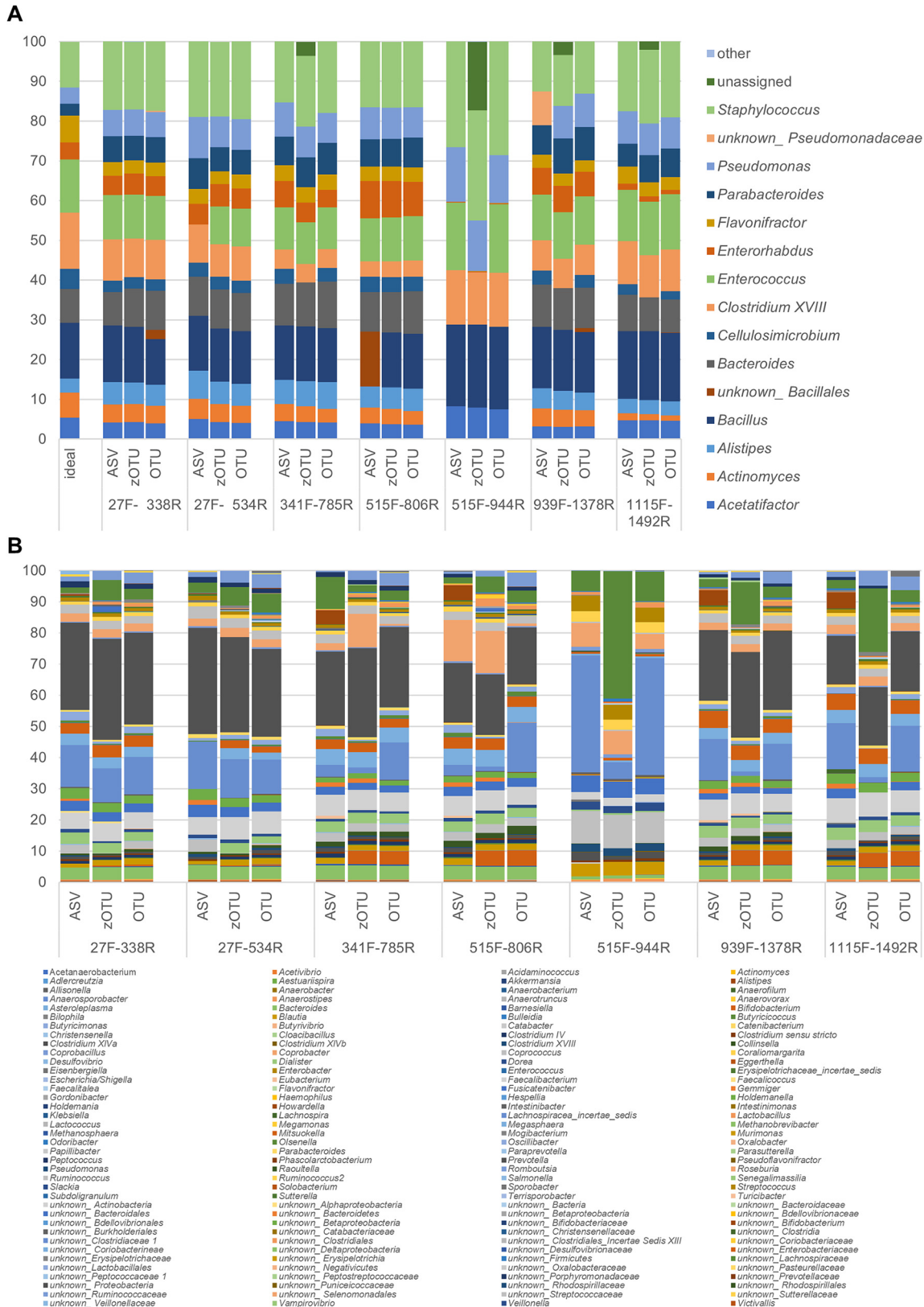
The ZIEL-I mock community consists of 13 species in 13 genera (Fig. 5B) and uses bacteria, which would be expected in the gut. Analyzing this, GG performed worst again. No genus-level classification for *Acetatifactor*, *Bacillus*, *Clostridium*, and *Pseudomonas* was possible using GG as a reference. GRD classified neither *Bacillus* nor *Pseudomonas* down to genus level. The other databases worked reasonably well but with some differences between V-regions. As before, 515F-944R (V4-V5) data performed worst. Only 4 to 8 taxa were classified at genus level, whereas between 9 and 13 taxa (Table 2) were identified when analyzing the data generated by using the primer pair 341F-785R (V3-V4). *Actinomyces*, *Alistipes*, *Bacteroides*, *Cellulosimicrobium*, *Parabacteroides*, and *Flavonifractor* were not detected with the primer pair 515F-944R (V4-V5) at the genus level irrespective of the reference database used.

The ZIEL-II mock community increased the complexity of the comparison by including 19 bacteria in 18 genera. Furthermore, we purposely included species which showed difficulties in past experiments (data not shown). Again, the 515F-944R (V4-V5) primer pair showed inadequate performance irrespective of the database. Using the Silva database, 14 to 18 taxa were classified at genus level for primer pair 341F-785R (V3-V4), whereas only 7 to 9 taxa were found for data corresponding to primer pair 515F-944R (V4-V5) at genus level (Table 2). *Akkermansia* could not be identified using the 27F-338R (V1-V2) primers (Fig. 5C). *Microbacterium* was underrepresented when using the 341F-785R (V3-V4) primers. *Enterobacter* and *Ruminococcus* were best classified by Silva. Generally, most accurate taxonomic classifications were possible when using Silva or RDP as the reference database. Silva even had the smallest amount of unknown genus-level identifications, followed by RDP, LTP, GRD, and GG.

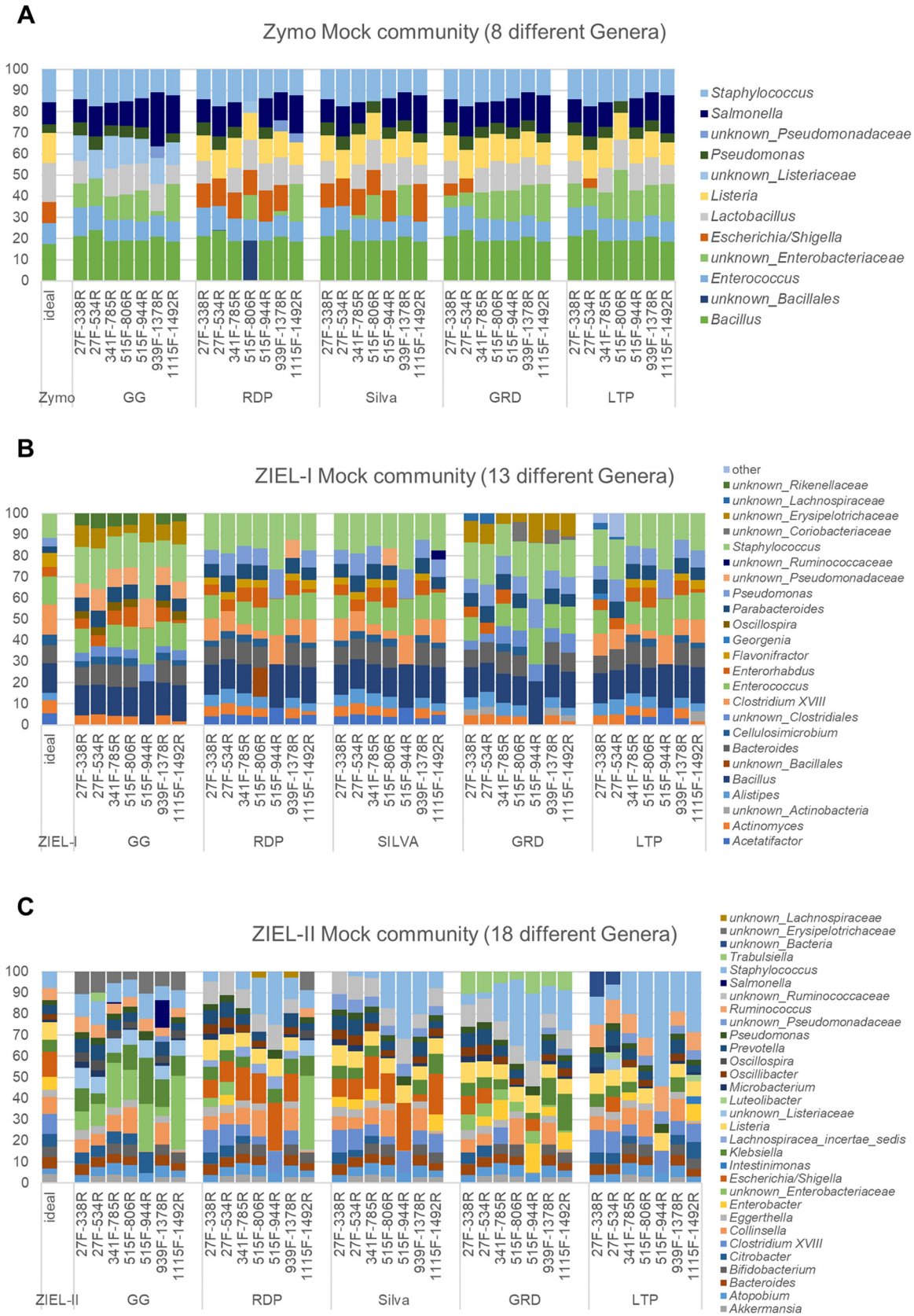


**FIG 3** Presence-and-absence map of human samples on phylum level for different V-regions. Gray represents present taxa, and white represents absent taxa. Primers and their V-region spanning are given in Table 1.

**Specific pipeline settings have minor influences on taxonomic classification.** As clustering methodologies showed a minor influence and the use of different reference databases a more severe impact on taxonomical profiles, we also assessed the potential influence of specific pipeline parameters. As ASVs performed slightly better than zOTUs and OTUs, we focused our comparison on ASVs. Processing steps include removal of primers and adapters, trimming of low-quality reads, chimera removal, and merging of paired-end reads. The removal of all primer and adapter sequences is required for ASV production. Incorrect removal or insufficient trimming leads to loss of sequences in the merging and







**FIG 5** Comparison of mock communities sequenced over different V-regions, processed using different databases as references (GG, GreenGenes; RDP, Ribosomal Database Project; GRD, the genomic-based 16S rRNA database; LTP, The All-Species Living Tree Project) at genus level. Primers and their V-region spanning are given in Table 1.

**TABLE 2** Number of ASVs and number of assigned taxa<sup>a</sup>

Primer pairs	Database	Zymo (8 species)		ZIEL-I (13 species)		ZIEL-II (19 species)	
		ASVs	Assigned taxa	ASVs	Assigned taxa	ASVs	Assigned taxa
27F-338R (V1-V2)	GG	19	6 [2]	19	9 [4]	34	13 [3]
	RDP	19	8	19	13	34	16 [2]
	Silva	19	8	19	13	34	17
	GRD	19	8 [1]	33	9 [4]	34	14 [3]
	LTP	19	7 [1]	33	11 [2]	34	13 [2]
27F-534R (V1-V3)	GG	23	5 [2]	36	8 [4]	45	15 [3]
	RDP	23	7 [1]	36	12	45	17 [1]
	Silva	23	7	36	12	45	18 [1]
	GRD	24	7 [1]	36	8 [4]	45	16 [3]
	LTP	23	7 [1]	36	9 [2]	45	15 [3]
341F-785R (V3-V4)	GG	12	6 [2]	22	9 [4]	33	15 [3]
	RDP	12	8	22	13	33	17 [1]
	Silva	12	8 [1]	22	13	33	18 [1]
	GRD	12	7 [1]	22	10 [2]	33	14 [3]
	LTP	12	7 [1]	22	13	33	17 [1]
515F-806R (V4)	GG	9	6 [2]	18	9 [4]	25	14 [3]
	RDP	9	5 [3]	18	12 [1]	25	15 [2]
	Silva	9	7 [1]	18	12 [1]	25	15 [1]
	GRD	9	7 [1]	18	9 [3]	25	12 [3]
	LTP	9	6 [1]	18	13	25	14 [1]
515F-944R (V4-V5)	GG	9	6 [2]	9	4 [3]	18	7 [3]
	RDP	9	8	9	7	18	9 [1]
	Silva	9	8	9	7	18	8 [1]
	GRD	9	7 [1]	9	4 [3]	18	8 [3]
	LTP	9	7 [1]	9	7	18	7 [1]
939F-1378R (V6-V8)	GG	17	5 [3]	21	9 [4]	33	13 [4]
	RDP	17	7 [2]	21	12 [1]	33	15 [2]
	Silva	17	7 [1]	21	13	33	16 [1]
	GRD	17	7 [1]	21	8 [4]	33	15 [3]
	LTP	17	7 [1]	21	13	33	16 [1]
1115F-1492R (V7-V9)	GG	15	6 [2]	18	9 [4]	36	13 [3]
	RDP	15	6 [2]	18	13	36	16 [3]
	Silva	15	8	18	12 [1]	36	17 [1]
	GRD	15	7 [1]	18	8 [4]	36	15 [3]
	LTP	15	7 [1]	18	12 [1]	36	16 [1]

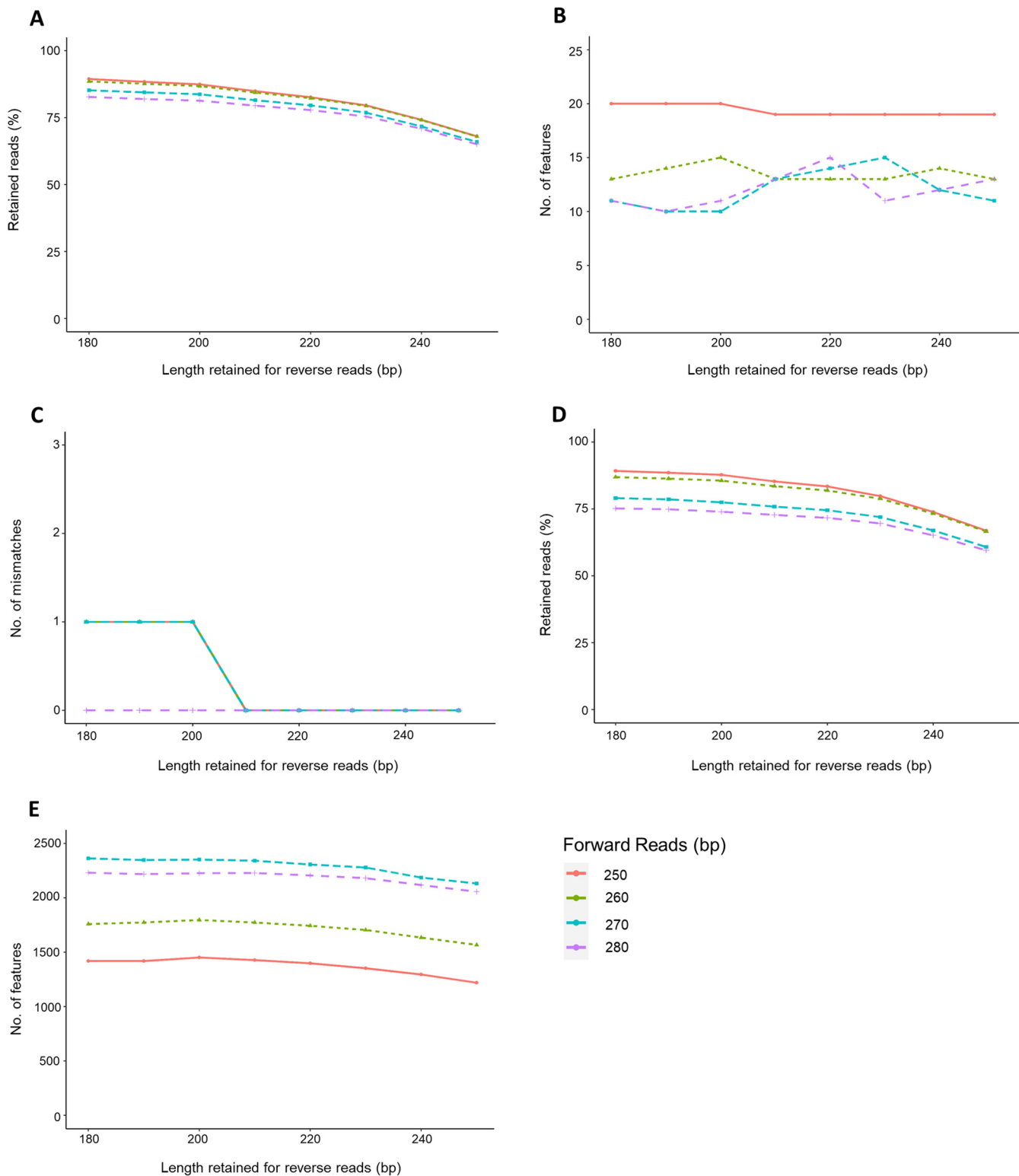
<sup>a</sup>Assigned taxa are at the genus level; brackets indicate that taxa are unknown at the genus level. The Zymo, ZIEL-I, and ZIEL-II mock communities contain 8, 13, and 19 bacterial species, respectively (for ZIEL-II, 18 at genus level, when *Escherichia/Shigella* fall into one cluster). Shading in green indicates good identification (the darker the better), while yellow and darker shading indicates inferior outcomes.

chimera removal steps. Ambiguous nucleotides would, for example, cause a problem, as default merging settings require a minimum overlap length of 20 bp and identical sequences in forward and reverse reads. Still, we expected the truncation step to have the largest impact on the results. In general, truncation is important to reduce the influence of low-quality bases at the end of the sequence reads. The truncated length for forward and reverse reads can be decided based on two factors: quality scores and amplicon length. However, there is a trade-off between read quality and read length for efficient merging. In this study, we performed the truncation step with different combinations of truncated length for forward and reverse reads for the ZIEL-I mock community for the V4 region (primer pair 515F-806R). Different ranges of forward (250 to 280 bp) and reverse (180 to 250 bp) read lengths were selected based on the quality ( $q$ ) score ( $\geq 20$ ) and amplicon length. We found that changes in the forward and reverse truncated lengths directly influence the percentages of sequence counts retained after that step (Fig. 6A). For instance, when the forward read length is set to 250 bp and the reverse read length to be 180 bp, 90% of the input reads were retained. The percentage of retained reads gradually decreased from 90% to 68% when increasing the reverse read length. The same trend was observed for forward 260-bp and reverse truncated length combinations (180 to 250 bp). However, using a forward read length of 270 bp or 280 bp combined with a reverse read length between 180 and 250 bp resulted in a lower percentage of retained reads, ranging from 85% to 65%. The lower number and, thus, reduced percentage of retained reads are mostly due to a decreased number of reads passing the filter. Subsequently, only this decreased number of reads was processed during denoising and merging steps (see Table S3).

The association between the percentage of reads retained and the number of ASVs obtained after those processing steps was also evaluated. The slight differences in the retained percentage of reads for different truncated length combinations did not drastically affect the number of features obtained. The total number of ASVs varied from 10 to 20 for different combinations of truncated lengths for the ZIEL-I mock community. Using truncated lengths of 250 bp and 180 bp for forward and reverse reads, respectively, resulted in 20 ASVs, while other length combinations obtained only 10 to 15 ASVs (Fig. 6B).

To check whether the observed differences in detected ASVs (e.g., 10 versus 20) arose from contaminated reads not corresponding to bacteria included in the ZIEL-I mock community, we performed a local BLAST search. We checked the reads produced by different forward and reverse read combinations against the reference sequence and used a cutoff of  $\geq 97\%$  identity,  $\geq 90\%$  coverage, and E value of  $\leq 0.00001$ . BLAST results of each forward and reverse read combination showed that 91 to 100% of the ASVs were mapped against the reference sequence of the mock community. The highest number of mismatches was found to be 1 (Fig. 6C). Only a very few nonhits, which did not reach the above-mentioned BLAST cutoffs, were obtained. Nevertheless, truncation for each amplicon length should be tested since low-quality bases impair read clustering.

Mock communities will, irrespective of the number of species added, never fully reflect complex microbial communities. Thus, we analyzed whether truncation showed an impact on a complex microbial community similar to that for the mock community used before. To this end, we used the previously analyzed 33 human stool samples as the test set. We found that the percentage of reads retained after truncation showed lower variations than for the mock community. The largest number of reads retained was identified for setting 250 bp and 180 bp for forward and reverse reads, respectively (Fig. 6D). Interestingly, when using 250 bp for the forward read, the percentage of retained reads decreased from 89% to 67% when increasing the reverse read length from 180 to 250 bp. Thus, insufficient removal of low-quality read sections (i.e., wrong bases) inhibits merging. The number of ASVs varied from 1,219 (250 bp forward/250 bp reverse) to 2,363 (for 270 bp/180 bp) for different combinations of truncated lengths (Fig. 6E), which led us to investigate whether different numbers of ASVs affect taxonomic assignments at genus level. Toward this end, we analyzed the number of generated ASVs for 280-bp forward reads in combination with different reverse read lengths. The number of ASVs varied from 2,057 (for 280 bp/250 bp) to 2,231 (for 280 bp/180 bp).



**FIG 6** (A and B) The effects of different lengths of forward and reverse reads after truncation on the percentage of sequences retained after denoising (A) and number of features obtained (B) for the ZIEL-I mock community. The numbers of mismatches obtained after local BLAST search against reference sets are shown; these were used in order to test the accuracy of the ASV predictions (C). (D and E) Analysis of human data set on retained reads after denoising and truncation (D) and number of features obtained (E) for each read-length combination.

**TABLE 3** Bacterial taxa at genus level influenced by primer choice and selected reference database<sup>a</sup>

Genera	27F-338R (V1-V2)		27F-534R (V1-V3)		341F-785R (V3-V4)		515F-806R (V4)		515F-944R (V4-V5)		939F-1378R (V6-V8)		1115F-1492R (V7-V9)	
	RDP	Silva	RDP	Silva	RDP	Silva	RDP	Silva	RDP	Silva	RDP	Silva	RDP	Silva
<i>Acetatifactor</i>	○	○	○	○	○	○	-	-	-	-	-	-	○	○
<i>Actinomyces</i>	-	-	○	○	-	-	-	-	x	x	-	-	-	-
<i>Akkermansia</i>	x	x	○	○	○	○	-	-	x	x	-	-	-	-
<i>Alistipes</i>	-	-	-	-	-	-	-	-	x	x	-	-	○	○
<i>Atopobium</i>	-	-	-	-	-	-	-	-	-	-	-	-	○	○
<b><i>Bacillus</i></b>	○	○	○	○	○	+	x	+	-	-	○	○	○	○
<b><i>Bacteroides</i></b>	○	○	+	+	+	+	+	+	x	x	○	○	○	○
<i>Bifidobacterium</i>	-	-	-	-	+	+	○	○	x	x	+	+	+	+
<i>Cellulosimicrobium</i>	-	-	-	-	-	-	○	○	x	x	-	-	-	-
<i>Citrobacter</i>	○	○	○	○	-	-	x	x	x	x	x	x	x	x
<b><i>Clostridium XVIII</i></b>	○	○	○	○	-	-	-	-	+	+	-	-	○	○
<i>Collinsella</i>	○	○	-	-	+	+	-	-	-	-	-	-	-	-
<i>Eggerthella</i>	-	-	○	○	○	+	+	+	x	x	+	+	-	-
<i>Enterobacter</i>	x	○	-	-	x	○	x	x	x	x	x	x	x	-
<b><i>Enterococcus</i></b>	○	○	○	○	○	○	○	○	○	○	○	○	+	+
<i>Enterorhabdus</i>	○	○	○	○	-	-	-	-	-	-	-	-	-	-
<b><i>Escherichia/Shigella</i></b>	○	○	○	○	-	-	○	○	-	-	○	○	x	-
<i>Flavonifractor</i>	-	-	-	-	-	-	-	-	x	x	-	-	-	x
<i>Klebsiella</i>	○	○	○	○	○	○	x	x	○	x	○	○	-	-
<i>Lactobacillus</i>	-	-	x	x	-	-	○	○	-	○	-	-	-	-
<b><i>Listeria</i></b>	+	+	○	○	○	○	○	○	○	○	○	○	○	○
<i>Microbacterium</i>	-	-	-	-	-	-	○	○	x	x	-	-	-	+
<i>Oscillibacter</i>	-	-	-	-	-	-	-	-	x	x	○	○	x	-
<i>Parabacteroides</i>	-	-	-	-	-	-	-	-	x	x	-	-	-	-
<i>Prevotella</i>	-	-	-	-	○	○	+	+	x	x	-	-	○	○
<b><i>Pseudomonas</i></b>	-	-	-	-	-	-	-	-	-	-	x	x	-	-
<i>Ruminococcus</i>	x	x	x	x	x	x	x	x	x	x	x	x	-	x
<i>Salmonella</i>	+	+	-	-	○	+	x	x	-	-	-	-	○	-
<b><i>Staphylococcus</i></b>	+	○	-	+	○	○	-	-	-	-	+	-	○	○
Database	RDP	Silva	RDP	Silva	RDP	Silva	RDP	Silva	RDP	Silva	RDP	Silva	RDP	Silva

<sup>a</sup>RDP (left column for each V-region) and Silva (right column for each V-region) were used as reference databases. +, <5% difference from reference (shaded green); ○, 5 to 25% difference from the reference (shaded white); -, >25% difference from the reference (shaded light brown); x, not detected at genus level (shaded dark brown). In bold are bacterial genera present in more than one mock community; therefore, mean values were calculated for these species to estimate their performance.

The number of different genera (including unknown and unclassified entries) varied from 131 (for 280 bp/250 bp) to 143 (for 280 bp/190 and 200 bp).

**Selection of primer, pipeline, parameters, and complexity of the ecosystem influences taxonomic classification.** Using three different mock communities, we were able to show differences in taxonomic compositions that were due to differences in used primer pairs, reference databases, clustering methods, or specific settings. We determined a set of bacterial taxa which are biased due to primer choice as well as the reference database (Table 3). Of note, we observed that there is a strong association between the correct assignment of taxa and the complexity of the mock community. For example, *Staphylococcus* was included in all three mock communities. This species was well characterized when using the Zymo mock community but poorly represented when using the more complex mock communities ZIEL-I and ZIEL-II (Table S3). Moreover, we evaluated the influence of specific primers and their comparability in a large population-based cohort ( $n=1,976$  subjects). Amplicon sequencing was performed targeting the V1-V2 and the V3-V4 regions of the 16S rRNA gene (1). For the V3-V4 region, the same primer set as in this work was used. However, for V1-V2, the same primer region was used but the forward primer (27F) did not include the degenerated bases Y and M (80). This led, for example, to a complete loss of identification of *Bifidobacterium* but to an identification of *Akkermansia*. These findings strengthen our hypothesis that methodological settings influence the outcome and, thus, the results that are generated out of 16S rRNA gene sequencing data. We would like to highlight the need for transparency to increase reproducibility and comparability.

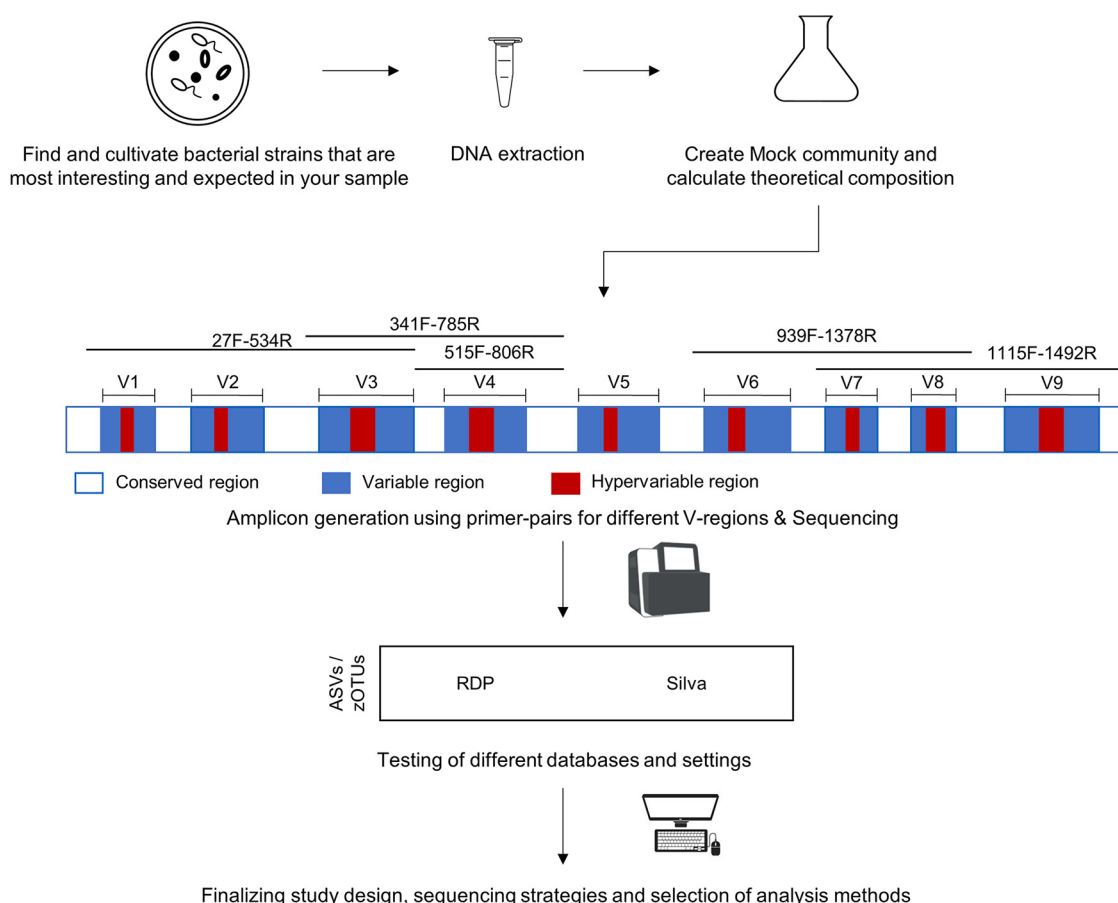


## DISCUSSION

For short-amplicon 16S rRNA gene sequencing, primers spanning more than one V-region are commonly used, which enhances precision in identifying bacteria compared to a single region. Some of the most frequently used primer pairs enclose V1-V3, V3-V4, and V3-V5, which were used in large population-based cohorts, e.g., the Human Microbiome Project and others (1, 33, 34). Nevertheless, each different primer pair or V-region used will cause bias in the data. In addition, sampling and sample storage, sample processing (including DNA extraction and amplicon generation), sequencing analysis, and data processing introduce further bias. In the last 10 years, many of these factors were studied for a variety of ecosystems, e.g., the human gut (31, 40, 59, 68, 74, 81–84), oral and skin microbiomes (64, 85, 86), food-related ecosystems (87, 88), and environmental microbiomes such as water, marine environments, and sludge (16, 69, 72, 89–91). Nevertheless, the combination of different bias-causing factors was rarely studied. In this study, we analyzed the effects of choice of primer, reference databases, clustering method, and specific pipeline settings in combination on human stool samples and mock communities with increasing complexity using recent approaches. We wanted to highlight the contribution of each these factors to the precision of taxonomic assignment, providing the scientific community with up-to-date guidelines for experimental design and data analysis. Anticipating conclusions, each experimental setting (e.g., cohort and environment) needs to be tested up front for best performance using different experimental settings and strategies.

First, the effect of different primer pairs on the corresponding microbial profile was evaluated. Irrespective of the reference database, the primer pair 341F-785R (V3-V4) slightly outperformed the other combinations and is, therefore, a justified choice for human gut samples. This is also in accordance with Thijs et al. (71), who suggested the primer pair 341F-785R to be a good match for soil and plant-associated bacterial microbiome studies, and Rausch et al. (92), who recommended the use of the V3-V4 region over V1-V2. The sequences produced by using the primer pair 515F-944R (V4-V5) performed well when analyzing the microbiota profile of the Zymo mock community but showed poor performance on the more complex ZIEL-I and ZIEL-II mock communities, suggesting that the primer combination may not be suitable for complex microbial ecosystems at all. This highlights also the importance of including mock communities in routinely performed 16S rRNA gene analysis, as a theoretical sequence analysis by Yang et al. (14) suggested the V4-V5 region to be a good match based on its robustness in representing the full-length 16S rRNA sequences and, therefore, theoretically seemed to be a good primer pair. However, it did not perform well when real samples were used.

Obviously, mock communities do not fully reflect the complexity of a microbial community as it is seen in, e.g., human stool samples. Therefore, we included 33 human fecal samples in our analysis as well. Here, phylum-level classification is robust across the use of different primer pairs targeting different V-regions for *Bacteroidetes* (except 515F-944R), *Proteobacteria*, and *Firmicutes*. In contrast, the detection of *Actinobacteria*, *Tenericutes*, *Lentisphaerae*, and *Verrucomicrobia* varied across the use of different primer pairs, highlighting that the choice of primer should be considered carefully. Intraindividual comparison at genus level showed a high degree of variability across the different targeted regions. This was due to many unknown or unclassified taxa at genus level as well as a generally large number of different taxa. This highlights the need for ecosystem-specific reference databases (93, 94) and new bioinformatic tools that can integrate data across V-regions by taking into account region-specific bias. Here, we notice a need for large-scale studies covering multiple V-regions, which would allow for training taxonomic classifiers that can dynamically account for any region-specific bias. This would possibly be obsolete by sequencing the full-length 16S rRNA gene, although sequencing would still be influenced by the primer choice, i.e., 27F and 1492R, for nearly full-length sequencing. Full-length 16S rRNA gene sequencing is possible by using third-generation sequencing strategies (24, 26, 27) or by the generation of short reads that are later *de novo* assembled to a synthetic full-length sequence (95). Those methods seemingly offer taxonomic identification down to species or even strain level (27). Both approaches are not yet well



**FIG 7** Recommended validation strategy before starting new microbiome studies, especially for uncommon environments. Even existing commonly used parameter combinations might be reevaluated. Thus, complex mock communities should be used and sequenced, testing a variety of different primer pairs for best performance within the environment of interest. Despite their being of minor influence, we still recommend using clustering approaches that include denoising steps (e.g., DADA2 generating ASVs) and recommend the seemingly well-curated and up-to-date databases RDP and Silva as references.

established for high-throughput sequencing and are not cost-efficient, reproducible, or easy in handling and thus need further investigation to be competitive. Further, long-read sequencing still suffers from comparably high error rates (29, 96).

It is known that the use of different bioinformatic pipelines can have an impact on the determined microbiota composition (40, 43, 65, 97). However, the influence of reference databases for taxonomic prediction was, to our knowledge, not intensively studied. In this study, we evaluated the performance of five different databases using three different mock communities. We tested the ability of each database to identify the correct taxonomy and assessed how well the known diversity of the mock samples could be captured by each database. Our finding illustrated that the Silva and RDP databases were the most accurate 16S rRNA gene databases, showing similar performances consistently superior to those of GRD, LTP, and GG in terms of true positives at the genus level. GG failed to classify *Escherichia/Shigella*, *Listeria*, *Acetatifactor*, *Bacillus*, *Clostridium*, and *Pseudomonas*, in line with the results of Park and Won (98), who found GG to be subpar compared to Silva. GG was last updated in 2013, and any usage is highly questionable.

In addition to the above, we found that quality assessment for each particular database could be conducted only when using a variety of V-regions and a sufficient complex mock community. Low-complexity mock communities using common bacteria did not reveal database issues. Thus, low-complexity mock communities might be used as positive controls in existing pipelines for general quality monitoring, but they are not recommended for detecting fundamental issues when setting up a new study,

**TABLE 4** Composition of the ZIEL-I mock community<sup>a</sup>

Species	Amt of gDNA used (ng)	Genome size (bp)	16S rRNA gene copy no.	Theoretical abundance (%)
<i>Actinomyces bowdenii</i>	12	3,103,770	3	6.3
<i>Enterorhabdus mucosicola</i>	12	3,009,822	2	4.3
<i>Cellulosimicrobium cellulans</i>	12	3,850,000	3	5.1
<i>Bacteroides sartorii</i>	12	5,377,291	7	8.5
<i>Alistipes</i> sp.	12	3,734,239	2	3.5
<i>Bacillus subtilis</i>	12	4,215,606	9	14.0
<i>Parabacteroides goldsteinii</i>	12	6,751,539	7	6.8
<i>Flavonifractor plautii</i>	12	4,306,691	2	3.0
<i>Clostridium ramosum</i>	12	3,235,195	7	14.2
<i>Enterococcus hirae</i>	12	2,962,227	6	13.3
<i>Acetatifactor muris</i>	12	6,013,646	5	5.4
<i>Staphylococcus warneri</i>	12	2,860,455	5	11.4
<i>Pseudomonas</i> sp.	12	6,342,352	4	4.1

<sup>a</sup>Genome sizes were determined according to entries in EzBioCloud (101), and 16S rRNA gene copy number was determined according to entries in rrnDB (103).

pipeline, or laboratory. Further, concerning other body sites (or environments), specific mock communities of sufficient complexity should be used. Certainly, the addition of ubiquitous bacteria, like the skin commensal *Cutibacterium acnes* in humans and other such bacteria, should be considered.

A third factor influencing taxonomic assignment is constituted by the denoising and OTU clustering steps of data analysis. To investigate this aspect, we compared classical OTUs generated by  $\geq 97\%$  clustering Qiime1, ASVs generated by DADA2 denoising (48), and zOTUs generated by the USEARCH denoising algorithm (49, 99). The numbers of features identified by these clustering approaches were nearly identical across all three approaches for the tested mock community. ASV clustering performed well in the human data sets despite the increased complexity, supporting the results of previous studies (42, 100), which suggests that ASVs are the current best choice, as they showed the highest accordance with the theoretical composition of the tested mock community. However, zOTUs performed very similarly and are more robust and user-friendly concerning the input.

Specific settings, e.g., the truncation length, influence the number of reads retained for further analysis steps, as we have demonstrated. Selecting a suitable truncation

**TABLE 5** Composition of the ZIEL-II mock community<sup>a</sup>

Species	Amt of gDNA used (ng)	Genome size (bp)	16S rRNA gene copy no.	Theoretical abundance (%)
<i>Prevotella copri</i>	12	3,784,859	4	4.2
<i>Collinsella aerofaciens</i>	12	2,463,631	5	8.0
<i>Atopobium parvulum</i>	12	1,543,805	1	2.6
<i>Eggerthella lenta</i>	12	3,500,501	3	3.4
<i>Bifidobacterium longum</i>	12	2,402,802	3	4.9
<i>Clostridium ramosum</i>	12	3,703,302	9	9.6
<i>Staphylococcus epidermidis</i>	12	2,520,741	5	7.8
<i>Klebsiella pneumoniae</i>	12	5,589,189	8	5.6
<i>Escherichia coli</i> LF82	12	4,881,487	7	5.6
<i>Shigella flexneri</i>	12	4,551,801	7	6.1
<i>Oscillibacter valericigenes</i>	12	4,470,622	3	2.6
<i>Akkermansia muciniphila</i>	12	2,760,363	3	4.3
<i>Ruminococcus gnavus</i>	12	3,415,781	5	5.8
<i>Bacteroides vulgatus</i>	12	5,063,322	7	5.4
<i>Pseudomonas aeruginosa</i>	12	6,612,169	4	2.4
<i>Citrobacter freundii</i>	12	5,300,882	8	5.9
<i>Enterobacter cloacae</i>	12	5,030,416	8	6.3
<i>Listeria welshimeri</i>	12	2,819,373	6	8.4
<i>Microbacterium flavum</i>	12	6,818,507	2	1.2

<sup>a</sup>Genome sizes were determined according to entries in EzBioCloud (101), and 16S rRNA gene copy number was determined according to entries in rrnDB (103).

length is of importance, as too-short reads have short or missing overlaps that lead to problems during merging. Conversely, too-long reads can be difficult to merge, as they show lower sequence quality. The varying number of detected ASVs for different truncation lengths is linked to the trade-off between incorporating reads of lower quality and the sensitivity for detecting low-abundance genera. By systematically reducing the reverse read length, the number of rarely observed sequences increased, as sequencing errors decrease. This highlights an important role for this parameter in the reproducibility of analysis results. To assess this potential bias, we suggest using sufficiently complex mock communities of known composition to determine suitable truncation lengths. Further, it is important to report this parameter (as well as all others) with respect to reproducibility of analysis results.

In summary, our results across 3 mock communities and 33 human samples suggest using primers for the V3-V4 region, which show good overall performance for human gut samples. As a reference database, we recommend using either Silva or RDP. Even though only minor differences were observed between clustering methods, we currently recommend using ASVs or zOTUs, with negligible difference between the two. Regarding pipeline settings, we suggest that truncated length combinations should be tested for the primer pairs used in each study. For example, we would suggest for V4 reads truncated to 250 bp and 180 bp for forward and reverse, respectively. However, the last settings depend on the amplicon lengths of the V-regions. To guarantee comparable and reliable results, we recommend creating specific (i.e., reflecting the targeted microbial environment) and sufficiently complex mock community to test whether the study design and the analysis pipelines will be suitable for the bacterial community of interest or type of sample desired (Fig. 7).

## MATERIALS AND METHODS

**Preparation of human gut samples.** Stool samples were obtained from healthy volunteers (33 subjects) and collected in stool sample tubes (Sarstedt AG & Co.). Tubes had been pre-filled under a clean bench with 8 ml of stabilizing buffer (1,400 ml of Milli-Q water supplemented with 60 ml of 0.5 M EDTA, 37.5 ml of 1 M sodium citrate, and 1.05 kg of ammonium sulfate [pH 5.2] and sterile filtered using a 0.2- $\mu$ m filter). A stainless steel mixing bead of 5.5 mm (MP Biomedicals) was added to facilitate homogenization of the crude stool in the stabilizing fluid. The stool was directly resuspended by shaking and vortexing. All samples were aliquoted (in 600- $\mu$ l portions) and stored at  $-80^{\circ}\text{C}$  until DNA extraction.

**Preparation of mock communities.** A mock community is a defined *in vitro*-created mixture of microbial cells. For validation, three different mock communities were used, (i) the ZymoBIOMICS microbial community DNA standard (Zymo Research; catalog no. D6306) with 8 bacterial species, (ii) a more complex in-house mock community (ZIEL-I) including 13 different bacterial species (Table 4), and (iii) another in-house mock community (ZIEL-II) with even more increased complexity including 19 different bacterial species (Table 5). For the in-house mock communities, common gut-related bacterial species were used. The mock community ZIEL-II included such species, which seemed to be influenced by targeted V-region in preliminary results (data not shown). Bacteria were cultured as described in Table S1 and harvested after 2 to 3 days by centrifugation. Pellets were resuspended in stabilizing buffer and stored at  $-80^{\circ}\text{C}$  until further processing. After genomic DNA (gDNA) extraction was performed for each strain separately (see below), strain identities were verified by Sanger sequencing. Afterwards, mock communities were constructed by pooling 12 ng of bacterial gDNA per strain. The theoretical composition was calculated according to the formula described for the Zymo mock community by Zymo Research:  $16\text{S rRNA gene copy number} = \text{total genomic DNA (g)} \times \text{unit conversion constant (bp/g)/genome size (bp)} \times 16\text{S rRNA gene copy number per genome}$ . Genome sizes were determined by the 16S reference database EzBioCloud (101). If the genome size for the species included was not available in the database, the closest relative (based on 16S rRNA gene identity) was used for genome size estimation instead. In cases in which only the genus of the bacterium used in the mock community is known, mean genome sizes including all species listed in the database of the genus were used. The 16S rRNA gene copy number was determined from rrnDB (102, 103) as a reference database, also using the closest relative as a surrogate or using mean values of 16S rRNA gene numbers if specific values were not available. Overall, the three different mock communities were sequenced (see below) in duplicates (ZIEL-I) or triplicates (Zymo and ZIEL-II). For further analyses (see below), we used the mean values of the taxonomic compositions of the replicates (all replicates are shown in Fig. S3 to S5).

**Extraction of gDNA.** gDNA was isolated with a modified protocol of Godon et al. (104) as described previously (105). Briefly, either 600  $\mu$ l of pure bacterial culture or 600  $\mu$ l of frozen stool samples (i.e., bacteria in stabilizer fluid) was thawed on ice and vortexed. Samples were transferred into a 2-ml bead-beating tube (MP Biomedicals), and 250  $\mu$ l of 4 M guanidinium thiocyanate and 500  $\mu$ l of 5% sodium *N*-laurylsarcosine were added. The mixture was incubated at  $70^{\circ}\text{C}$  for 60 min with shaking (700 rpm). Next, cells were disrupted by bead-beating using a FastPrep24 instrument (MP Biomedicals). Bead-beating was conducted three times for 40 s at 6.5 m/s, with cooling with dry ice. Processed samples were stored

on ice. Subsequently, 15 mg of polyvinylpyrrolidone was added to each sample, with brief mixing. Samples were centrifuged for 3 min at  $15,000 \times g$  and  $4^{\circ}\text{C}$ , and the supernatant was transferred into a fresh 2-ml sample tube. To every sample,  $5 \mu\text{l}$  of RNase A (10 mg/ml) was added and samples were incubated for 20 min at  $37^{\circ}\text{C}$  with moderate shaking (700 rpm). DNA was purified using gDNA columns (Macherey-Nagel) following the manufacturer's instructions. Finally, gDNA was eluted in  $100 \mu\text{l}$  of elution buffer provided in the kit. Concentrations and purity were checked using the NanoDrop system (Thermo Scientific), and samples were stored at  $4^{\circ}\text{C}$  (up to 5 days) or at  $-20^{\circ}\text{C}$  thereafter.

**Primer selection and *in silico* testing.** Primers for commonly used V-regions were chosen after a literature survey. *In silico* tests of primer specificity were conducted using Silva TestPrime 1.0 (<http://www.arb-silva.de/search/testprime/>) using standard settings with zero mismatches.

**Library preparation of different variable regions of the 16S rRNA gene.** For amplification of the variable regions (Fig. S1) and addition of adapter binding sites for sequencing, a 1st-step PCR was performed in a  $50\text{-}\mu\text{l}$  total volume. Each reaction mixture contained 24 ng of gDNA,  $1 \times$  Phusion HF buffer, 0.2 mM deoxynucleoside triphosphates (dNTPs),  $0.125 \mu\text{M}$  each forward and reverse primer, 7.5% dimethyl sulfoxide (DMSO), and  $0.25 \mu\text{l}$  of Phusion HF II DNA polymerase (Thermo Fisher). PCR was performed as follows:  $98^{\circ}\text{C}$  for 40 s, 15 cycles of  $98^{\circ}\text{C}$  for 20 s, the V-region specific annealing temperature (Table 1) for 40 s, and  $72^{\circ}\text{C}$  for 40 s, and a final extension step at  $72^{\circ}\text{C}$  for 2 min.

Barcodes enabling multiplexing were added in the 2nd-step PCR. For this, a  $100\text{-}\mu\text{l}$  PCR mixture was prepared using  $10 \mu\text{l}$  of the 1st-step PCR product,  $1 \times$  Phusion HF buffer, 0.2 mM dNTPs,  $0.125 \mu\text{M}$  each forward and reverse barcode primer, 0.25% DMSO, and  $0.5 \mu\text{l}$  of Phusion HF II DNA polymerase. PCR conditions were  $98^{\circ}\text{C}$  for 40 s, 10 cycles of  $98^{\circ}\text{C}$  for 20 s,  $55^{\circ}\text{C}$  for 40 s, and  $72^{\circ}\text{C}$  for 40 s, and a final extension step at  $72^{\circ}\text{C}$  for 2 min. Further details and work time estimations are found in the work of Reitmeier et al. (105).

**Library quality check and sequencing.** For validation and quality assurance,  $8 \mu\text{l}$  of the 2nd-step PCR product was loaded onto a 1.5% agarose gel. The remaining  $92 \mu\text{l}$  of the 2nd-step PCR product was purified with AMPure XP beads using a ratio of 1.8 times (i.e., addition of  $180 \mu\text{l}$  of beads to  $100 \mu\text{l}$  of sample). Concentrations of the final PCR products were measured in triplicates using a Qubit (Thermo Fisher). Each sample was adjusted to 0.5 nM, and all samples were pooled and sequenced in paired-end modus for  $2 \times 300$  bp (PE300) using a MiSeq system (Illumina, Inc.) following the manufacturer's instructions. The final DNA concentration of the library was 12 pM, and 15% (vol/vol) PhiX was added.

**Primer-specific feature classifiers.** User-generated feature classifiers accounting for unique characteristics introduced by sample preparation, sequencing primer, and read length perform generally better than the naive classifiers trained on full-length sequences (106). In order to improve the taxonomic classification, five different databases were used to generate primer-specific feature classifiers, namely, GreenGenes (GG) (51), the Ribosomal Database Project (RDP) (52), Silva (53), the genomic-based 16S rRNA Database (GRD) (54), and The All-Species Living Tree (LTP) database (55). Feature classifiers were built for each V-region or primer pair using the *q2-feature-classifier* (107), which is a naive Bayes taxonomic classifier implemented in Qiime2-2019.10 (47).

**OTU clustering using Qiime1.** We consider Qiime-UCLUST (108) a popular example of an OTU-generating method as well as the recently proposed USEARCH-UNOISE3 (49, 99) (described below). Qiime-UCLUST clusters sequence reads at  $\geq 97\%$  sequence identity. UCLUST clustering was performed in Qiime1 as follows. Forward and reverse primer sequences and the low-quality reads ( $q \leq 20$ ) of demultiplexed paired-end reads were removed by cutadapt 2.10 (109). The trimmed reads were joined by *multiple\_join\_paired\_ends.py* to create a single fasta file of all samples using *multiple\_split\_libraries\_fastq.py*. OTU abundance tables were generated using the UCLUST clustering method through the script *pick\_de\_novo\_otus.py* script in Qiime1. OTU mapping files along with representative sequences, alignment of sequences, and taxonomic alignment files were generated during the *de novo* clustering steps. The RDP database was used as a reference database for defining OTUs at  $\geq 97\%$  sequence similarity.

**zOTU generation using UNOISE.** USEARCH-UNOISE3 aims to reconstruct exact biological sequences from the samples into zOTUs. Paired-end raw reads were merged using the *fastq\_mergepairs* script of USEARCH version 11 (108), and the primer sequences were removed using the *fastx\_truncate* script. Merging and primer removal steps were conducted before quality filtering, as primer removal reduces the expected errors and merging before quality filtering improves the base call error estimates captured in the overlapping regions as suggested by the author of USEARCH/UPARSE (110). Processed reads were deduplicated and *de novo* clustered into zOTUs. RDP database (project release 11) was used for taxonomic assignment of the representative zOTU sequences.

**ASV generation using nf-core/ampliseq pipeline.** The three mock communities and human data sets were analyzed using the *nf-core/ampliseq* nextflow pipeline (111, 112). *nf-core/ampliseq* is a Qiime2-based end-to-end solution for processing 16S rRNA gene amplicon sequencing data. The quality of raw sequencing reads was assessed by FastQC (113). Primer sequences and bases with low-quality scores were trimmed using cutadapt (109). The DADA2 (48) package wrapped inside the *nf-core/ampliseq* pipeline was used for denoising and constructing ASVs. Based on the quality profile and amplicon length, truncated lengths for forward (250 to 280 bp) and reverse reads (180 to 260 bp) were used in the DADA2 denoising steps to study the relationships between the truncated lengths and number of ASVs generated.

**Data visualization using Rhea.** Data visualization was performed with the R-based pipeline Rhea (114), a collection of R-scripts for 16S rRNA gene sequencing data analysis. After normalization, alpha-diversity and beta-diversity were determined and visualized. Taxonomic classification was conducted down to genus level.

**Data visualization for human samples.** To determine differences of the microbiota composition by targeting different V-regions, a multivariate analysis was performed using the *vegan* R-package. Therefore, a Bray-Curtis distance between samples was calculated based for relative abundance values



on phylum and genus levels and grouped according to targeted V-region. First, two dimensions of the nonmetric MDS (NMDS) plot were visualized by using *ggplot2*, and data points were labeled according to targeted V-region.

**Data availability.** Raw sequencing data are available at the Sequence Read Archive under the accession number [PRJNA674596](https://www.ncbi.nlm.nih.gov/sra/PRJNA674596).

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**FIG S1**, PDF file, 0.1 MB.

**FIG S2**, PDF file, 0.7 MB.

**FIG S3**, PDF file, 0.1 MB.

**TABLE S1**, PDF file, 0.02 MB.

**TABLE S2**, PDF file, 0.02 MB.

**TABLE S3**, PDF file, 0.1 MB.

**TABLE S4**, PDF file, 0.1 MB.

**TABLE S5**, PDF file, 0.2 MB.

## ACKNOWLEDGMENTS

We thank Thomas Clavel, Theresa Streidl, and David Wylensek (Research Group Functional Microbiome, RWTH Aachen), Annemarie Siebert and Michaela Kreitmeier (Chair of Microbial Ecology, TUM), and Nico Gebhardt (Chair of Nutrition & Immunology, TUM) for providing bacterial strains. Further, we thank Annika Naumann, Andrea Isabel Proaño Vasco, and Caroline Ziegler for excellent technical assistance.

I.A.-S. was funded by the ZIEL—Institute for Food & Health with a grant for a doctorate position and partially funded by a grant of the Research Foundation of Dairy Science at the Technical University of Munich (VFMF), both given to K.N. J.B. was partially funded by VILLUM Young Investigator Grant no. 13154. This work was partially funded by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation)—Projektnummer 395357507—SFB 1371.

## REFERENCES

- Reitmeier S, Kiessling S, Clavel T, List M, Almeida EL, Ghosh TS, Neuhaus K, Grallert H, Linseisen J, Skurk T, Brandl B, Breuninger TA, Troll M, Rathmann W, Linkohr B, Hauner H, Laudes M, Franke A, Le Roy CI, Bell JT, Spector T, Baumbach J, O'Toole PW, Peters A, Haller D. 2020. Arrhythmic gut microbiome signatures predict risk of type 2 diabetes. *Cell Host Microbe* 28:258–272.e6. <https://doi.org/10.1016/j.chom.2020.06.004>.
- Hamady M, Knight R. 2009. Microbial community profiling for human microbiome projects: tools, techniques, and challenges. *Genome Res* 19:1141–1152. <https://doi.org/10.1101/gr.085464.108>.
- Shokralla S, Spall JL, Gibson JF, Hajibabaei M. 2012. Next-generation sequencing technologies for environmental DNA research. *Mol Ecol* 21:1794–1805. <https://doi.org/10.1111/j.1365-294X.2012.05538.x>.
- Duvallet C, Gibbons SM, Gurry T, Irizarry RA, Alm EJ. 2017. Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. *Nat Commun* 8:1784. <https://doi.org/10.1038/s41467-017-01973-8>.
- Goodrich JK, Davenport ER, Beaumont M, Jackson MA, Knight R, Ober C, Spector TD, Bell JT, Clark AG, Ley RE. 2016. Genetic determinants of the gut microbiome in UK twins. *Cell Host Microbe* 19:731–743. <https://doi.org/10.1016/j.chom.2016.04.017>.
- Hiergeist A, Reischl U, Gessner A, Priority Program 1656 Intestinal Microbiota Consortium/ quality assessment participants. 2016. Multi-center quality assessment of 16S ribosomal DNA-sequencing for microbiome analyses reveals high inter-center variability. *Int J Med Microbiol* 306:334–342. <https://doi.org/10.1016/j.ijmm.2016.03.005>.
- Nelson MC, Morrison HG, Benjamins J, Grim SL, Graf J. 2014. Analysis, optimization and verification of Illumina-generated 16S rRNA gene amplicon surveys. *PLoS One* 9:e94249. <https://doi.org/10.1371/journal.pone.0094249>.
- Woese CR, Kandler O, Wheelis ML. 1990. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A* 87:4576–4579. <https://doi.org/10.1073/pnas.87.12.4576>.
- Janda JM, Abbott SL. 2007. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol* 45:2761–2764. <https://doi.org/10.1128/JCM.01228-07>.
- Baker GC, Smith JJ, Cowan DA. 2003. Review and re-analysis of domain-specific 16S primers. *J Microbiol Methods* 55:541–555. <https://doi.org/10.1016/j.mimet.2003.08.009>.
- Martinez-Porchas M, Villalpando-Canchola E, Ortiz Suarez LE, Vargas-Albores F. 2017. How conserved are the conserved 16S-rRNA regions? *PeerJ* 5:e3036. <https://doi.org/10.7717/peerj.3036>.
- Fischer MA, Güllert S, Neulinger SC, Streit WR, Schmitz RA. 2016. Evaluation of 16S rRNA gene primer pairs for monitoring microbial community structures showed high reproducibility within and low comparability between datasets generated with multiple archaeal and bacterial primer pairs. *Front Microbiol* 7:1297. <https://doi.org/10.3389/fmicb.2016.01297>.
- Martinez-Porchas M, Vargas-Albores F. 2017. An efficient strategy using k-mers to analyse 16S rRNA sequences. *Heliyon* 3:e00370. <https://doi.org/10.1016/j.heliyon.2017.e00370>.
- Yang B, Wang Y, Qian P-Y. 2016. Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. *BMC Bioinformatics* 17:135. <https://doi.org/10.1186/s12859-016-0992-y>.
- Pausan MR, Csorba C, Singer G, Till H, Schöpf V, Santigli E, Klug B, Högenauer C, Blohs M, Moissl-Eichinger C. 2019. Exploring the archaeome: detection of archaeal signatures in the human body. *Front Microbiol* 10:2796. <https://doi.org/10.3389/fmicb.2019.02796>.
- Bahram M, Anslan S, Hildebrand F, Bork P, Tedersoo L. 2018. Newly designed 16S rRNA metabarcoding primers amplify diverse and novel archaeal taxa from the environment. *Environ Microbiol Rep* 11:487–494. <https://doi.org/10.1111/1758-2229.12684>.
- Berry D, Ben Mahfoudh K, Wagner M, Loy A. 2011. Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. *Appl Environ Microbiol* 77:7846–7849. <https://doi.org/10.1128/AEM.05220-11>.
- Liu Z, Lozupone C, Hamady M, Bushman FD, Knight R. 2007. Short pyrosequencing reads suffice for accurate microbial community analysis. *Nucleic Acids Res* 35:e120. <https://doi.org/10.1093/nar/gkm541>.
- Branton D, Deamer DW, Marziali A, Bayley H, Benner SA, Butler T, Di Ventra M, Garaj S, Hibbs A, Huang X, Jovanovich SB, Krstic PS, Lindsay S,

- Ling XS, Mastrangelo CH, Meller A, Oliver JS, Pershin YV, Ramsey JM, Riehn R, Soni GV, Tabard-Cossa V, Wanunu M, Wiggan M, Schloss JA. 2008. The potential and challenges of nanopore sequencing. *Nat Biotechnol* 26:1146–1153. <https://doi.org/10.1038/nbt.1495>.
20. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, Dewinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulsson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J, Turner S. 2009. Real-time DNA sequencing from single polymerase molecules. *Science* 323:133–138. <https://doi.org/10.1126/science.1162986>.
  21. Kai S, Matsuo Y, Nakagawa S, Kryukov K, Matsukawa S, Tanaka H, Iwai T, Imanishi T, Hirota K. 2019. Rapid bacterial identification by direct PCR amplification of 16S rRNA genes using the MinION nanopore sequencer. *FEBS Open Bio* 9:548–557. <https://doi.org/10.1002/2211-5463.12590>.
  22. Curren E, Yoshida T, Kuwahara VS, Leong SCY. 2019. Rapid profiling of tropical marine cyanobacterial communities. *Reg Stud Mar Sci* 25:100485. <https://doi.org/10.1016/j.risma.2018.100485>.
  23. Benítez-Páez A, Portune KJ, Sanz Y. 2016. Species-level resolution of 16S rRNA gene amplicons sequenced through the MinION™ portable nanopore sequencer. *Gigascience* 5:4. <https://doi.org/10.1186/s13742-016-0111-z>.
  24. Cuscó A, Catozzi C, Viñes J, Sanchez A, Francino O. 2019. Microbiota profiling with long amplicons using Nanopore sequencing: full-length 16S rRNA gene and the 16S-ITS-23S of the *rrm* operon [version 2; peer review: 2 approved, 3 approved with reservations]. *F1000Res* 7:1755. <https://doi.org/10.12688/f1000research.16817.2>.
  25. Martijn J, Lind AE, Schön ME, Spiertz I, Juzokaite L, Buniks I, Pettersson OV, Ettema TJG. 2019. Confident phylogenetic identification of uncultured prokaryotes through long read amplicon sequencing of the 16S-ITS-23S rRNA operon. *Environ Microbiol* 21:2485–2498. <https://doi.org/10.1111/1462-2920.14636>.
  26. Callahan BJ, Wong J, Heiner C, Oh S, Theriot CM, Gulati AS, McGill SK, Dougherty MK. 2019. High-throughput amplicon sequencing of the full-length 16S rRNA gene with single-nucleotide resolution. *Nucleic Acids Res* 47:e103. <https://doi.org/10.1093/nar/gkz569>.
  27. Johnson JS, Spakowicz DJ, Hong BY, Petersen LM, Demkowicz P, Chen L, Leopold SR, Hanson BM, Agresta HO, Gerstein M, Sodergren E, Weinstock GM. 2019. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun* 10:5029. <https://doi.org/10.1038/s41467-019-13036-1>.
  28. Rhoads A, Au KF. 2015. PacBio sequencing and its applications. *Genomics Proteomics Bioinformatics* 13:278–289. <https://doi.org/10.1016/j.gpb.2015.08.002>.
  29. Bowden R, Davies RW, Heger A, Pagnamenta AT, de Cesare M, Oikkonen LE, Parkes D, Freeman C, Dhalla F, Patel SY, Popitsch N, Ip CLC, Roberts HE, Salatino S, Lockstone H, Lunter G, Taylor JC, Buck D, Simpson MA, Donnelly P. 2019. Sequencing of human genomes with nanopore technology. *Nat Commun* 10:1869. <https://doi.org/10.1038/s41467-019-09637-5>.
  30. Hoffmann C, Hill DA, Minkah N, Kirn T, Troy A, Artis D, Bushman F. 2009. Community-wide response of the gut microbiota to enteropathogenic *Citrobacter rodentium* infection revealed by deep sequencing. *Infect Immun* 77:4668–4678. <https://doi.org/10.1128/IAI.00493-09>.
  31. Alcon-Giner C, Caim S, Mitra S, Ketskemety J, Wegmann U, Wain J, Belteki G, Clarke P, Hall LJ. 2017. Optimisation of 16S rRNA gut microbiota profiling of extremely low birth weight infants. *BMC Genomics* 18:841. <https://doi.org/10.1186/s12864-017-4229-x>.
  32. Nossa CW, Oberdorf WE, Yang L, Aas JA, Paster BJ, Desantis TZ, Brodie EL, Malamud D, Poles MA, Pei Z. 2010. Design of 16S rRNA gene primers for 454 pyrosequencing of the human foregut microbiome. *World J Gastroenterol* 16:4135–4144. <https://doi.org/10.3748/wjg.v16.i33.4135>.
  33. Jumpstart Consortium Human Microbiome Project Data Generation Working Group. 2012. Evaluation of 16S rDNA-Based Community Profiling for Human Microbiome Research. *PLoS One* 7:e39315. <https://doi.org/10.1371/journal.pone.0039315>.
  34. Methé BA, Nelson KE, Pop M, Creasy HH, Giglio MG, Huttenhower C, Gevers D, Petrosino JF, Abubucker S, Badger JH, Chinwalla AT, Earl AM, FitzGerald MG, Fulton RS, Hallsworth-Pepin K, Lobos EA, Madupu R, Magrini V, Martin JC, Mitreva M, Muzny DM, Sodergren EJ, Versalovic J, Wollam AM, Worley KC, Wortman JR, Young SK, Zeng Q, Aagaard KM, Abolude OO, Allen-Vercoe E, Alm EJ, Alvarado L, Andersen GL, Anderson S, Appelbaum E, Arachchi HM, Armitage G, Arze CA, Ayvaz T, Baker CC, Begg L, Belachew T, Bhanagiri V, Bihan M, Blaser MJ, Bloom T, Bonazzi VR, Brooks P, Buck GA, Buhay CJ, The Human Microbiome Project Consortium, et al. 2012. A framework for human microbiome research. *Nature* 486:215–221.
  35. Claesson MJ, O'Sullivan O, Wang Q, Nikkila J, Marchesi JR, Smidt H, de Vos WM, Ross RP, O'Toole PW. 2009. Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS One* 4:e6669. <https://doi.org/10.1371/journal.pone.0006669>.
  36. Ghyselinck J, Pfeiffer S, Heylen K, Sessitsch A, De Vos P. 2013. The effect of primer choice and short read sequences on the outcome of 16S rRNA gene based diversity studies. *PLoS One* 8:e71360. <https://doi.org/10.1371/journal.pone.0071360>.
  37. Bukin YS, Galachyants YP, Morozov IV, Bukin SV, Zakharenko AS, Zemskaya TI. 2019. The effect of 16S rRNA region choice on bacterial community metabarcoding results. *Sci Data* 6:190007. <https://doi.org/10.1038/sdata.2019.7>.
  38. Barb JJ, Oler AJ, Kim H-S, Chalmers N, Wallen GR, Cashion A, Munson PJ, Ames NJ. 2016. Development of an analysis pipeline characterizing multiple hypervariable regions of 16S rRNA using mock samples. *PLoS One* 11:e0148047. <https://doi.org/10.1371/journal.pone.0148047>.
  39. Pinna NK, Dutta A, Monzoorul Haque M, Mande SS. 2019. Can targeting non-contiguous V-regions with paired-end sequencing improve 16S rRNA-based taxonomic resolution of microbiomes?: an *in silico* evaluation. *Front Genet* 10:653. <https://doi.org/10.3389/fgene.2019.00653>.
  40. Plummer E, Twin J, Bulach DM, Garland SM, Tabrizi SN. 2015. A comparison of three bioinformatics pipelines for the analysis of preterm gut microbiota using 16S rRNA gene sequencing data. *J Proteomics Bioinform* 8:12. <https://doi.org/10.4172/jpb.1000381>.
  41. Marizzoni M, Gurry T, Provasi S, Greub G, Lopizzo N, Ribaldi F, Festari C, Mazzelli M, Mombelli E, Salvatore M, Mirabelli P, Franzese M, Soricelli A, Frisoni GB, Cattaneo A. 2020. Comparison of bioinformatics pipelines and operating systems for the analyses of 16S rRNA gene amplicon sequences in human fecal samples. *Front Microbiol* 11:1262. <https://doi.org/10.3389/fmicb.2020.01262>.
  42. Prodan A, Tremaroli V, Brolin H, Zwinderman AH, Nieuwdorp M, Levin E. 2020. Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. *PLoS One* 15:e0227434. <https://doi.org/10.1371/journal.pone.0227434>.
  43. Sierra MA, Li Q, Pushalkar S, Paul B, Sandoval TA, Kamer AR, Corby P, Guo Y, Ruff RR, Alekseyenko AV, Li X, Saxena D. 2020. The influences of bioinformatics tools and reference databases in analyzing the human oral microbial community. *Genes* 11:878. <https://doi.org/10.3390/genes11080878>.
  44. Balvocüte M, Huson DH. 2017. SILVA, RDP, Greengenes, NCBI and OTT—how do these taxonomies compare? *BMC Genomics* 18:114. <https://doi.org/10.1186/s12864-017-3501-4>.
  45. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75:7537–7541. <https://doi.org/10.1128/AEM.01541-09>.
  46. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenkov T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–336. <https://doi.org/10.1038/nmeth.f.303>.
  47. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, et al. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 37:852–857. <https://doi.org/10.1038/s41587-019-0209-9>.
  48. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* 13:581–583. <https://doi.org/10.1038/nmeth.3869>.

49. Edgar RC. 2018. Updating the 97% identity threshold for 16S ribosomal RNA OTUs. *Bioinformatics* 34:2371–2375. <https://doi.org/10.1093/bioinformatics/bty113>.
50. Callahan BJ, McMurdie PJ, Holmes SP. 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J* 11:2639–2643. <https://doi.org/10.1038/ismej.2017.119>.
51. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72:5069–5072. <https://doi.org/10.1128/AEM.03006-05>.
52. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR, Tiedje JM. 2014. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* 42:D633–D642. <https://doi.org/10.1093/nar/gkt1244>.
53. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:D590–D596. <https://doi.org/10.1093/nar/gks1219>.
54. Laboratory for Integrated Bioinformatics, Center for Integrative Medical Sciences. 2015. GRD—Genomic-based 16S ribosomal RNA database, Riken (Japan). <https://metasystems.riken.jp/grd/>.
55. Yarza P, Richter M, Peplies J, Euzéby J, Amann R, Schleifer K-H, Ludwig W, Glöckner FO, Rosselló-Móra R. 2008. The All-Species Living Tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. *Syst Appl Microbiol* 31:241–250. <https://doi.org/10.1016/j.syapm.2008.07.001>.
56. Bailén M, Bressa C, Larrosa M, González-Soltero R. 2020. Bioinformatic strategies to address limitations of 16S rRNA short-read amplicons from different sequencing platforms. *J Microbiol Methods* 169:105811. <https://doi.org/10.1016/j.mimet.2019.105811>.
57. Escobar-Zepeda A, Godoy-Lozano EE, Raggi L, Segovia L, Merino E, Gutiérrez-Rios RM, Juárez K, Licea-Navarro AF, Pardo-Lopez L, Sanchez-Flores A. 2018. Analysis of sequencing strategies and tools for taxonomic annotation: defining standards for progressive metagenomics. *Sci Rep* 8:12034. <https://doi.org/10.1038/s41598-018-30515-5>.
58. Bokulich NA, Rideout JR, Mercurio WG, Shiffer A, Wolfe B, Maurice CF, Dutton RJ, Turnbaugh PJ, Knight R, Caporaso JG. 2016. mockrobiota: a public resource for microbiome bioinformatics benchmarking. *mSystems* 11:e00062-16. <https://doi.org/10.1128/mSystems.00062-16>.
59. Gorzelak MA, Gill SK, Tasnim N, Ahmadi-Vand Z, Jay M, Gibson DL. 2015. Methods for improving human gut microbiome data by reducing variability through sample processing and storage of stool. *PLoS One* 10:e0134802. <https://doi.org/10.1371/journal.pone.0134802>.
60. Sinha R, Chen J, Amir A, Vogtmann E, Shi J, Inman KS, Flores R, Sampson J, Knight R, Chia N. 2016. Collecting fecal samples for microbiome analyses in epidemiology studies. *Cancer Epidemiol Biomarkers Prev* 25:407–416. <https://doi.org/10.1158/1055-9965.EPI-15-0951>.
61. Burz SD, Abraham AL, Fonseca F, David O, Chapron A, Béguet-Crespel F, Cénard S, Le Roux K, Patrascu O, Levenez F, Schwintner C, Blottière HM, Béra-Maillet C, Lepage P, Doré J, Juste C. 2019. A guide for *ex vivo* handling and storage of stool samples intended for fecal microbiota transplantation. *Sci Rep* 9:8897. <https://doi.org/10.1038/s41598-019-45173-4>.
62. Choo JM, Leong LE, Rogers GB. 2015. Sample storage conditions significantly influence faecal microbiome profiles. *Sci Rep* 5:16350. <https://doi.org/10.1038/srep16350>.
63. Fouhy F, Deane J, Rea MC, O'Sullivan O, Ross RP, O'Callaghan G, Plant BJ, Stanton C. 2015. The effects of freezing on faecal microbiota as determined using MiSeq sequencing and culture-based investigations. *PLoS One* 10:e0119355. <https://doi.org/10.1371/journal.pone.0119355>.
64. Teng F, Darveekaran Nair SS, Zhu P, Li S, Huang S, Li X, Xu J, Yang F. 2018. Impact of DNA extraction method and targeted 16S-rRNA hyper-variable region on oral microbiota profiling. *Sci Rep* 8:16321. <https://doi.org/10.1038/s41598-018-34294-x>.
65. Ducarmon QR, Hornung BVH, Geelen AR, Kuijper EJ, Zwartink RD. 2020. Toward standards in clinical microbiota studies: comparison of three DNA extraction methods and two bioinformatic pipelines. *mSystems* 5:e00547-19. <https://doi.org/10.1128/mSystems.00547-19>.
66. Gryp T, Glorieux G, Joossens M, Vaneechoutte M. 2020. Comparison of five assays for DNA extraction from bacterial cells in human faecal samples. *J Appl Microbiol* 129:378–388. <https://doi.org/10.1111/jam.14608>.
67. Sinha R, Abu-Ali G, Vogtmann E, Fodor AA, Ren B, Amir A, Schwager E, Crabtree J, Ma S, Abnet CC, The Microbiome Quality Control Project Consortium, Knight R, White O, Huttenhower C. 2017. Assessment of variation in microbial community amplicon sequencing by the Microbiome Quality Control (MBQC) project consortium. *Nat Biotechnol* 35:1077–1086. <https://doi.org/10.1038/nbt.3981>.
68. Costea PI, Zeller G, Sunagawa S, Pelletier E, Alberti A, Levenez F, Tramontano M, Driessen M, Herczeg R, Jung F-E, Kultima JR, Hayward MR, Coelho LP, Allen-Vercocoe E, Bertrand L, Blaut M, Brown JRM, Carton T, Cools-Portier S, Daigneault M, Derrien M, Druésne A, de Vos WM, Finlay BB, Flint HJ, Guamer F, Hattori M, Heilig H, Luna RA, van Hylckama Vlieg J, Junick J, Klymiuk I, Langella P, Le Chatelier E, Mai V, Manichanh C, Martin JC, Mery C, Morita H, O'Toole PW, Orvain C, Patil KR, Penders J, Persson S, Pons N, Popova M, Salonen A, Saulnier D, Scott KP, Singh B, Slezak K, Veiga P, Versalovic J, Zhao L, Zoetendal EG, Ehrlich SD, Dore J, Bork P. 2017. Towards standards for human fecal sample processing in metagenomic studies. *Nat Biotechnol* 35:1069–1076. <https://doi.org/10.1038/nbt.3960>.
69. Parada AE, Needham DM, Fuhrman JA. 2016. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environ Microbiol* 18:1403–1414. <https://doi.org/10.1111/1462-2920.13023>.
70. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO. 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 41:e1. <https://doi.org/10.1093/nar/gks088>.
71. Thijs S, Op De Beeck M, Beckers B, Truyens S, Stevens V, Van Hamme JD, Weyens N, Vangronsveld J. 2017. Comparative evaluation of four bacteria-specific primer pairs for 16S rRNA gene surveys. *Front Microbiol* 8:494. <https://doi.org/10.3389/fmicb.2017.00494>.
72. Tremblay J, Singh K, Fern A, Kirton ES, He S, Woyke T, Lee J, Chen F, Dangel JL, Tringe SG. 2015. Primer and platform effects on 16S rRNA tag sequencing. *Front Microbiol* 6:771. <https://doi.org/10.3389/fmicb.2015.00771>.
73. Gohl DM, Vangay P, Garbe J, MacLean A, Hauge A, Becker A, Gould TJ, Clayton JB, Johnson TJ, Hunter R, Knights D, Beckman KB. 2016. Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. *Nat Biotechnol* 34:942–949. <https://doi.org/10.1038/nbt.3601>.
74. Jones MB, Highlander SK, Anderson EL, Li W, Dayrit M, Klitgord N, Fabani MM, Seguritan V, Green J, Pride DT, Yooshep S, Biggs W, Nelson KE, Venter JC. 2015. Library preparation methodology can influence genomic and functional predictions in human microbiome research. *Proc Natl Acad Sci U S A* 112:14024–14029. <https://doi.org/10.1073/pnas.1519288112>.
75. D'Amore R, Ijaz UZ, Schirmer M, Kenny JG, Gregory R, Darby AC, Shakya M, Podar M, Quince C, Hall N. 2016. A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling. *BMC Genomics* 17:55. <https://doi.org/10.1186/s12864-015-2194-9>.
76. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 79:5112–5120. <https://doi.org/10.1128/AEM.01043-13>.
77. Yeh Y-C, Needham DM, Sieradzki ET, Fuhrman JA. 2018. Taxon disappearance from microbiome analysis reinforces the value of mock communities as a standard in every sequencing run. *mSystems* 3:e00023-18. <https://doi.org/10.1128/mSystems.00023-18>.
78. Karstens L, Asquith M, Davin S, Fair D, Gregory WT, Wolfe AJ, Braun J, McWeeney S. 2019. Controlling for contaminants in low-biomass 16S rRNA gene sequencing experiments. *mSystems* 4:e00290-19. <https://doi.org/10.1128/mSystems.00290-19>.
79. Sinha R, Abnet CC, White O, Knight R, Huttenhower C. 2015. The microbiome quality control project: baseline study design and future directions. *Genome Biol* 16:276. <https://doi.org/10.1186/s13059-015-0841-8>.
80. Thaiss Christoph A, Zeevi D, Levy M, Zilberman-Schapira G, Suez J, Tengeler Anouk C, Abramson L, Katz Meirav N, Korem T, Zmora N, Kuperman Y, Biton I, Gilad S, Harmelin A, Shapiro H, Halpern Z, Segal E, Elinav E. 2014. Transkingdom control of microbiota diurnal oscillations promotes metabolic homeostasis. *Cell* 159:514–529. <https://doi.org/10.1016/j.cell.2014.09.048>.
81. Bellali S, Lagier JC, Raoult D, Bou Khalil J. 2019. Among live and dead bacteria, the optimization of sample collection and processing remains essential in recovering gut microbiota components. *Front Microbiol* 10:1606. <https://doi.org/10.3389/fmicb.2019.01606>.
82. Ma J, Sheng L, Hong Y, Xi C, Gu Y, Zheng N, Li M, Chen L, Wu G, Li Y, Yan J, Han R, Li B, Qiu H, Zhong J, Jia W, Li H. 2020. Variations of gut microbiome profile under different storage conditions and preservation periods: a multi-dimensional evaluation. *Front Microbiol* 11:972. <https://doi.org/10.3389/fmicb.2020.00972>.



83. Penington JS, Penno MAS, Ngui KM, Ajami NJ, Roth-Schulze AJ, Wilcox SA, Bandala-Sanchez E, Wentworth JM, Barry SC, Brown CY, Couper JJ, Petrosino JF, Papenfuss AT, Harrison LC, ENDIA Study Group. 2018. Influence of fecal collection conditions and 16S rRNA gene sequencing at two centers on human gut microbiota analysis. *Sci Rep* 8:4386. <https://doi.org/10.1038/s41598-018-22491-7>.
84. Walker AW, Martin JC, Scott P, Parkhill J, Flint HJ, Scott KP. 2015. 16S rRNA gene-based profiling of the human infant gut microbiota is strongly influenced by sample processing and PCR primer choice. *Microbiome* 3:26. <https://doi.org/10.1186/s40168-015-0087-4>.
85. Bjerre RD, Hugerth LW, Boulund F, Seifert M, Johansen JD, Engstrand L. 2019. Effects of sampling strategy and DNA extraction on human skin microbiome investigations. *Sci Rep* 9:17287. <https://doi.org/10.1038/s41598-019-53599-z>.
86. Meisel JS, Hannigan GD, Tyldsley AS, SanMiguel AJ, Hodkinson BP, Zheng Q, Grice EA. 2016. Skin microbiome surveys are strongly influenced by experimental design. *J Invest Dermatol* 136:947–956. <https://doi.org/10.1016/j.jid.2016.01.016>.
87. De Filippis F, Parente E, Zotta T, Ercolini D. 2018. A comparison of bioinformatic approaches for 16S rRNA gene profiling of food bacterial microbiota. *Int J Food Microbiol* 265:9–17. <https://doi.org/10.1016/j.ijfoodmicro.2017.10.028>.
88. Xue Z, Kable ME, Marco ML. 2018. Impact of DNA sequencing and analysis methods on 16S rRNA gene bacterial community analysis of dairy products. *mSphere* 3:e00410-18. <https://doi.org/10.1128/mSphere.00410-18>.
89. Fredriksson NJ, Hermansson M, Wilen BM. 2013. The choice of PCR primers has great impact on assessments of bacterial community diversity and dynamics in a wastewater treatment plant. *PLoS One* 8:e76431. <https://doi.org/10.1371/journal.pone.0076431>.
90. Shah M. 2014. An application of sequencing batch reactors in the identification of microbial community structure from an activated sludge. *J Applied Environ Microbiol* 2:176–184.
91. Brandt J, Albertsen M. 2018. Investigation of detection limits and the influence of DNA extraction and primer choice on the observed microbial communities in drinking water samples using 16S rRNA gene amplicon sequencing. *Front Microbiol* 9:2140. <https://doi.org/10.3389/fmicb.2018.02140>.
92. Rausch P, Rühlemann M, Hermes BM, Doms S, Dagan T, Dierking K, Domin H, Fraune S, von Frieling J, Hentschel U, Heinsen F-A, Höppner M, Jahn MT, Jaspers C, Kissoyan KAB, Langfeldt D, Rehman A, Reusch TBH, Roeder T, Schmitz RA, Schulenburg H, Soluch R, Sommer F, Stukenbrock E, Weiland-Bräuer N, Rosenstiel P, Franke A, Bosch T, Baines JF. 2019. Comparative analysis of amplicon and metagenomic sequencing methods reveals key features in the evolution of animal metaorganisms. *Microbiome* 7:133. <https://doi.org/10.1186/s40168-019-0743-1>.
93. Dueholm MS, Andersen KS, McLroy SJ, Kristensen JM, Yashiro E, Karst SM, Albertsen M, Nielsen PH. 2020. Generation of comprehensive ecosystem-specific reference databases with species-level resolution by high-throughput full-length 16S rRNA gene sequencing and automated taxonomy assignment (AutoTax). *mBio* 11:e01557-20. <https://doi.org/10.1128/mBio.01557-20>.
94. F Escapa I, Huang Y, Chen T, Lin M, Kokaras A, Dewhirst FE, Lemon KP. 2020. Construction of habitat-specific training sets to achieve species-level assignment in 16S rRNA gene datasets. *Microbiome* 8:65. <https://doi.org/10.1186/s40168-020-00841-w>.
95. Karst SM, Dueholm MS, McLroy SJ, Kirkegaard RH, Nielsen PH, Albertsen M. 2018. Retrieval of a million high-quality, full-length microbial 16S and 18S rRNA gene sequences without primer bias. *Nat Biotechnol* 36:190–195. <https://doi.org/10.1038/nbt.4045>.
96. Loit K, Adamson K, Bahram M, Puusepp R, Anslan S, Kiiker R, Drenkhan R, Tedersoo L. 2019. Relative performance of MinION (Oxford Nanopore Technologies) versus Sequel (Pacific Biosciences) third-generation sequencing instruments in identification of agricultural and forest fungal pathogens. *Appl Environ Microbiol* 85:e01368-19. <https://doi.org/10.1128/AEM.01368-19>.
97. Almeida A, Mitchell AL, Tarkowska A, Finn RD. 2018. Benchmarking taxonomic assignments based on 16S rRNA gene profiling of the microbiota from commonly sampled environments. *Gigascience* 7:giy054. <https://doi.org/10.1093/gigascience/giy054>.
98. Park S-C, Won S. 2018. Evaluation of 16S rRNA databases for taxonomic assignments using mock community. *Genomics Inform* 16:e24. <https://doi.org/10.5808/GI.2018.16.4.e24>.
99. Edgar RC. 2016. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *bioRxiv* <https://doi.org/10.1101/081257>.
100. Caruso V, Song X, Asquith M, Karstens L. 2019. Performance of microbiome sequence inference methods in environments with varying biomass. *mSystems* 4:e00163-18. <https://doi.org/10.1128/mSystems.00163-18>.
101. Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y, Seo H, Chun J. 2017. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 67:1613–1617. <https://doi.org/10.1099/ijsem.0.001755>.
102. Klappenbach JA, Saxman PR, Cole JR, Schmidt TM. 2001. rrdb: the Ribosomal RNA Operon Copy Number Database. *Nucleic Acids Res* 29:181–184. <https://doi.org/10.1093/nar/29.1.181>.
103. Roller BRK, Stoddard SF, Schmidt TM. 2016. Exploiting rRNA operon copy number to investigate bacterial reproductive strategies. *Nat Microbiol* 1:16160. <https://doi.org/10.1038/nmicrobiol.2016.160>.
104. Godon JJ, Zumstein E, Dabert P, Habouzit F, Moletta R. 1997. Molecular microbial diversity of an anaerobic digester as determined by small-subunit rDNA sequence analysis. *Appl Environ Microbiol* 63:2802–2813. <https://doi.org/10.1128/AEM.63.7.2802-2813.1997>.
105. Reitmeier S, Kiessling S, Neuhaus K, Haller D. 2020. Comparing circadian rhythmicity in the human gut microbiome. *STAR Protoc* 1:100148. <https://doi.org/10.1016/j.xpro.2020.100148>.
106. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA, Gregory Caporaso J. 2018. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* 6:90. <https://doi.org/10.1186/s40168-018-0470-z>.
107. Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73:5261–5267. <https://doi.org/10.1128/AEM.00062-07>.
108. Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460–2461. <https://doi.org/10.1093/bioinformatics/btq461>.
109. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 17:10–12.
110. Edgar RC. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* 10:996–998. <https://doi.org/10.1038/nmeth.2604>.
111. Peltzer A, Straub D, Patel H. 2019. nf-core/ampliseq: Ampliseq version 1.1.2. <https://doi.org/10.5281/zenodo.3585924>.
112. Ewels PA, Peltzer A, Fillinger S, Patel H, Alneberg J, Wilm A, Garcia MU, Di Tommaso P, Nahnsen S. 2020. The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol* 38:276–278. <https://doi.org/10.1038/s41587-020-0439-x>.
113. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
114. Lagkouravos I, Fischer S, Kumar N, Clavel T. 2017. Rhea: a transparent and modular R pipeline for microbial profiling based on 16S rRNA gene amplicons. *PeerJ* 5:e2836. <https://doi.org/10.7717/peerj.2836>.
115. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, Turner P, Parkhill J, Loman NJ, Walker AW. 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol* 12:87. <https://doi.org/10.1186/s12915-014-0087-z>.
116. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci U S A* 108:4516–4522. <https://doi.org/10.1073/pnas.1000080107>.
117. Fuks G, Elgart M, Amir A, Zeisel A, Turnbaugh PJ, Soen Y, Shental N. 2018. Combining 16S rRNA gene variable regions enables high-resolution microbial community profiling. *Microbiome* 6:17. <https://doi.org/10.1186/s40168-017-0396-x>.
118. Leubhn M, Hanreich A, Klocke M, Schlüter A, Bauer C, Pérez CM. 2014. Towards molecular biomarkers for biogas production from lignocellulose-rich substrates. *Anaerobe* 29:10–21. <https://doi.org/10.1016/j.anaerobe.2014.04.006>.
119. Turner S, Pryer KM, Miao VP, Palmer JD. 1999. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *J Eukaryot Microbiol* 46:327–338. <https://doi.org/10.1111/j.1550-7408.1999.tb04612.x>.