Universiteit
Leiden
The Netherlands

Microbiome-mediated colonization resistance: defense against enteropathogens and multi-drug resistant organisms

Ducarmon, Q.R.

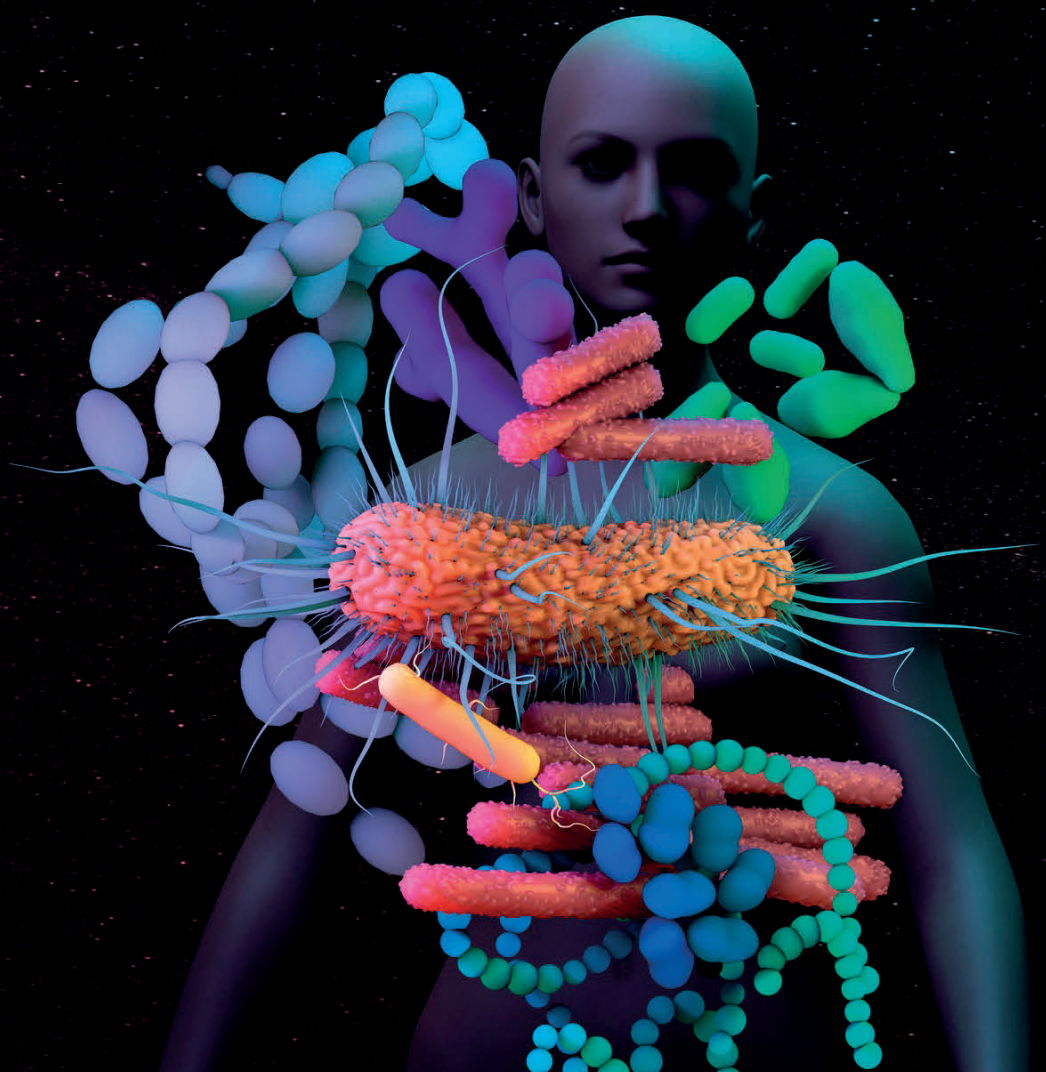
Citation
Ducarmon, Q. R. (2022, March 23). *Microbiome-mediated colonization resistance: defense against enteropathogens and multi-drug resistant organisms*. Retrieved from https://hdl.handle.net/1887/3280022

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3280022

Note: To cite this publication please use the final published version (if applicable).

Part II

Optimization and standardization of laboratory and computational procedures for microbiome research



Chapter 4

Toward standards in clinical microbiota studies: comparison of three DNA extraction methods and two bioinformatic pipelines

DNA isolation and informatics for microbiota studies
mSystems, 2020

Q.R. Ducarmon[a, b], B.V.H. Hornung[a,b], A.R. Geelen[a,b], E.J. Kuijper[a,b], R.D. Zwittink[a,b]

[a]Center for Microbiome Analyses and Therapeutics, Department of Medical Microbiology, Leiden University Medical Center, Leiden, The Netherlands
[b]Experimental Bacteriology, Department of Medical Microbiology, Leiden University Medical Center, Leiden, The Netherlands

mSystems. 2020 Feb 11;5(1):e00547-19.
Supplementary information available online.

Abstract

When studying the microbiome using next generation sequencing, DNA extraction method, sequencing procedures and bioinformatic processing are crucial to obtain reliable data. Method choice has been demonstrated to strongly affect the final biological interpretation. We assessed the performance of three DNA extraction methods and two bioinformatic pipelines for bacterial microbiota profiling through 16S rRNA gene amplicon sequencing, using positive and negative controls for DNA extraction and sequencing, and eight different types of high- or low-biomass samples. Performance was evaluated based on quality control passing, DNA yield, richness, diversity and compositional profiles. All DNA extraction methods retrieved the theoretical relative bacterial abundance with maximum three-fold change, although differences were seen between methods, and library preparation and sequencing induced little variation. Bioinformatic pipelines showed different results for observed richness, but diversity and compositional profiles were comparable. DNA extraction methods were successful for feces and oral swabs and variation induced by DNA extraction methods was lower than inter-subject (biological) variation. For low-biomass samples, a mixture of genera present in negative controls and sample-specific genera, possibly representing biological signal, were observed. We conclude that the tested bioinformatic pipelines perform equally with pipeline-specific advantages and disadvantages. Two out of three extraction methods performed equally well, while one method was less accurate regarding retrieval of compositional profiles. Lastly, we again demonstrate the importance of including negative controls when analyzing low bacterial biomass samples.

Importance

Method choice throughout the workflow of a microbiome study, from sample collection to DNA extraction and sequencing procedures, can greatly affect results. This study evaluated three different DNA extraction methods and two bioinformatic pipelines by including positive and negative controls, and various biological specimens. By identifying an optimal combination of DNA extraction method and bioinformatic pipeline use, we hope to contribute to increased methodological consistency in microbiota studies. Our methods were not only applied to commonly studied samples for microbiota analysis, e.g. feces, but also on more rarely studied, low-biomass samples. Microbiota composition profiles of low-biomass samples (e.g. urine and tumor biopsies) were not always distinguishable from negative controls, or showed partial overlap, confirming the importance of including negative controls in microbiota studies, especially when low bacterial biomass is expected.

Introduction

Humans constantly interact with microbes that are present in the environment and reside on or within the human body. Recently, the attention for microbes has shifted from an exclusive interest in the pathogenicity of specific microbes toward the potential beneficial role of the microbiota in human health [1]. The gastrointestinal tract contains the highest number of microbes and has been the most extensively studied body site of all human microbial communities [2]. However, many other body sites are inhabited by various microbes composing a specific microbiota, such as the oral region, skin and urogenital system. Microbial complexity varies between these niches, e.g. a healthy vaginal microbiota is often mainly composed of a few *Lactobacillus* strains, while gut and skin microbiota are usually more diverse [3].

A limiting factor in current microbiome research is that comparison of various study results is often difficult due to the application of different methodologies and lack of appropriate controls. These differences can affect data outcomes and lead to variation as large as biological differences [4]. Variation can be introduced throughout the entire workflow, from sample collection, storage and processing to data analysis [5-8]. Recently, more attention has been devoted to standardizing the workflow of microbiome research. For instance, it was observed that DNA extraction has a large impact on obtained data [4, 9] and consensus has been achieved regarding the application of bead-beating to increase efficiency of cell wall lysis and thereby improve the yield of Gram-positive bacterial DNA [10]. Nevertheless, various kits and in-house extraction methods are used across different laboratories. Recently, Costea *et al.* evaluated 21 DNA extraction methods

across three continents and suggested one protocol, named protocol Q, as 'golden standard' for human fecal samples [9]. They stated that it was unknown whether this method is optimal for other samples than fecal material, e.g. for low-biomass samples. To evaluate performance of DNA extraction for low-biomass samples, it is crucial to include multiple negative controls to allow for identification of bacterial DNA introduced during the entire workflow, from sample collection to sequencing [11].

As part of optimizing the procedures for 16S rRNA gene amplicon sequencing-based microbiota studies in our facility, we evaluated three DNA extraction methods and two bioinformatic pipelines using various positive controls and negative controls. In addition, we applied these DNA extraction methods to various biological specimens.

Results and discussion

Mock communities pass quality control

We evaluated three different DNA extraction methods and two bioinformatic pipelines for microbiota profiling through 16S rRNA gene amplicon sequencing (Fig 1) using several positive and negative controls. Included positive controls were two bacterial mock communities and one DNA standard. Included negative controls were DNA extraction controls and sequencing controls. Quality control (QC) passing (DNA concentration and intact genomic fragment) were evaluated to determine extraction method performance. It was expected that positive controls would pass QC, while negative controls would not. Regarding mock communities, all extractions using Zymo and Q passed QC, while for Magna one extraction did not pass QC for both the ATCC mock community and Zymo mock community (Table S3). This was not unexpected, as mock communities were diluted for extraction using Magna and, therefore, DNA concentrations were lower. Negative extraction controls did not pass QC for Q and Magna, but they did for Zymo. This likely represents a higher contamination load during the extraction process for Zymo, which was also reflected by higher DNA concentrations (Table S3). A full overview of all samples included in this study, their QC passing and DNA concentrations can be found in Table S4.

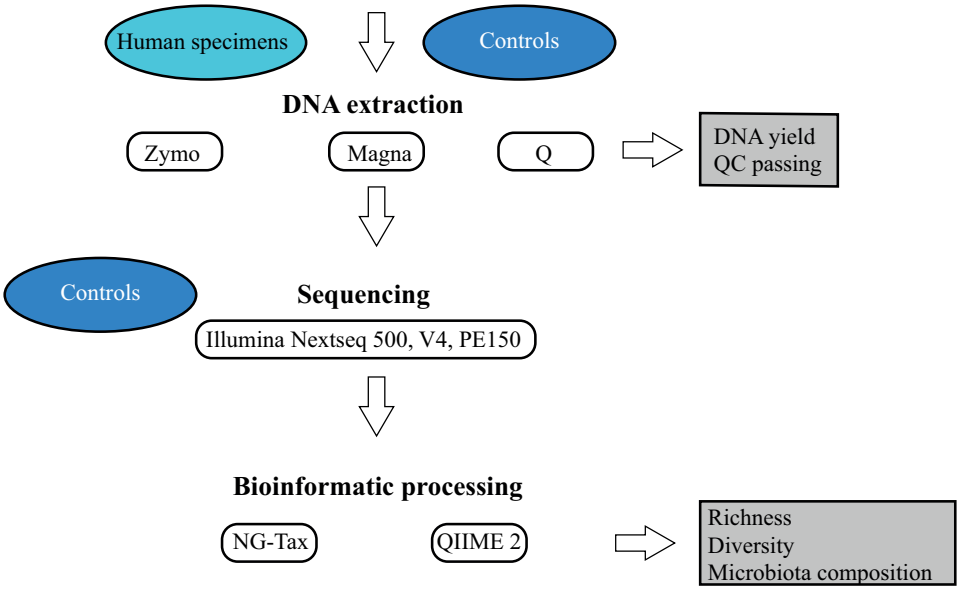


Figure 1: Study design workflow. DNA was extracted from human specimens and positive and negative controls using three different DNA extraction methods. DNA extraction performance was assessed on DNA yield and QC passing. Extracted DNA, and positive and negative sequencing controls were sequenced. Raw sequencing data was processed using two bioinformatic pipelines. Performance was assessed on microbiota composition, richness and diversity.

Positive controls: Classification, richness, diversity and relative species abundance

Primer choice in combination with bioinformatic pipeline choice may limit correct classification of all bacterial species in mock communities

Performance of the three extraction methods in combination with two bioinformatic pipelines, NG-Tax and QIIME 2, was evaluated on correctly identifying richness, diversity and relative abundances from bacterial mock communities and a DNA standard. Richness and diversity were computed at the OTU level and at genus level. Analysis of compositional profiles was performed at genus level. Both pipelines failed to classify one organism from either mock community; NG-Tax did not detect *Cutibacterium* from the ATCC mock, while QIIME 2 did not detect *Salmonella* from the Zymo mock. The inability to detect *Cutibacterium* is most likely a combination of different internal settings and filtering steps in the computational pipelines and a primer choice issue, since the universal 515F and 806R primers are known to poorly amplify *Cutibacterium acnes* [12]. Poor amplification of *C. acnes* results in limited read numbers, which may be filtered out during bioinformatic processing. These issues could likely be solved by choosing primers targeting different 16S rRNA gene regions, or by using adapted V4 region primers which do allow for accurate amplification of *Cutibacterium* [12, 13]. Regarding

QIIME 2 and the inability to detect *Salmonella*, there was an *Enterobacteriaceae* family with approximately expected relative abundance for *Salmonella*, and we were therefore confident this represented *Salmonella*. This *Enterobacteriaceae* family was subsequently included as *Salmonella*, and designated *Enterobacteriaceae (Salmonella)*. This classification error likely resulted from the fact that *Enterobacteriaceae* members cannot always be discriminated based on the 16S rRNA V4 region [14].

DNA standard and Zymo mock community data can be recovered independent of extraction protocol or pipeline

The Zymo mock and DNA standard consist of respectively cell material or DNA of eight bacterial species and two fungal species. As the 16S rRNA gene was targeted, fungi should not be detected. Therefore, theoretical richness is eight and theoretical Shannon diversity was calculated to be 2.01.

Regarding the DNA standard, NG-Tax overestimated OTU-based richness for both duplicates, DNA 1 and DNA 2 (Fig 2A, table S3). Richness was however accurately retrieved at genus level (Fig 2C). The same was observed regarding diversity, which was overestimated at the OTU level (Fig 2B), but accurate at genus level (Fig 2D). QIIME 2 approached theoretical richness and diversity values at the OTU level (Fig 2A+B, table S3). Richness slightly improved at genus level (Fig 2C), while diversity did not differ from OTU-based diversity (Fig 2D). Thus, QIIME 2 better estimated richness and diversity at the OTU level, while NG-Tax performed better at genus level (Table S3). This likely stems from NG-Tax finding an inflated richness due to assignment of multiple OTUs from a single organism (e.g. multiple *Enterococcus* OTUs). When OTUs are collapsed at genus level, this is no longer a problem, probably explaining why NG-Tax can perform better at genus level, while performing worse at the OTU level.

Compositional profiles of DNA 1 and DNA 2 are highly similar to theoretical abundance (Fig 3A+B). To quantify differences in compositional profiles, Bray-Curtis dissimilarity and Kullback-Leibler divergence (Fig 3C-F) [15] and fold errors for each taxon (Fig 4) were determined. For the dissimilarity and divergence values, a value of zero represents an identical microbiota composition to the theoretical expectation. NG-Tax obtained values closer to zero than QIIME 2 for both DNA 1 and DNA 2, although the difference is minimal (Fig 3 and Table S2) and the performance of both pipelines can therefore be regarded as equal. A similar conclusion can be drawn from the fold errors (Fig 4), since both pipelines accurately retrieved expected relative abundance, with all genera having a fold error between -1.5 and 1.5 (Table S3).

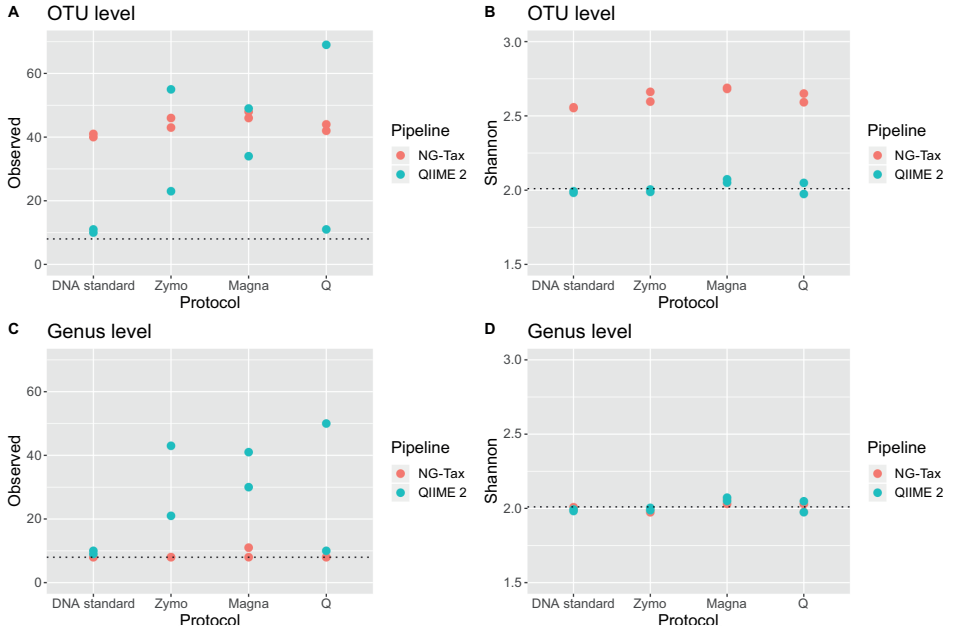


Figure 2: Richness (observed OTUs) and diversity (Shannon) computed for Zymo DNA and Zymo mock at OTU level (A+B) and at genus level (C+D) for each combination of bioinformatic pipeline and DNA extraction method. Dashed lines indicate theoretical values.

Similar analyses were performed for the Zymo mock to evaluate performance of DNA extraction methods in combination with the bioinformatic pipelines. All DNA extraction methods, independent of pipeline, resulted in OTU-based richness above 20 for most samples, far higher than theoretical expectance (Fig 2A). This is especially noteworthy for QIIME 2, as it was highly accurate in retrieving correct richness for the DNA standard, in contrast to NG-Tax. Zymo and Q protocols in combination with NG-Tax retrieved accurate genus level-based richness, while a slightly inflated richness was observed for Magna (Fig 2C). No extraction method was consistent in retrieving correct genus level-based richness in combination with QIIME 2. Regarding diversity, all DNA extractions, independent of pipeline, retrieved highly accurate values at genus level (Table S3). At the OTU level, however, the NG-Tax pipeline resulted in overestimation of diversity independent of DNA extraction method, and can therefore be considered a result of bioinformatic processing. Magna extraction resulted in Bray-Curtis and Kullback-Leibler values closer to zero than Zymo and Q, independent of pipeline (Fig 3C-F and Table S3). A similar conclusion can be drawn from the fold errors, which are lowest for Magna and pipeline-independent (Fig 4 and Table S3).

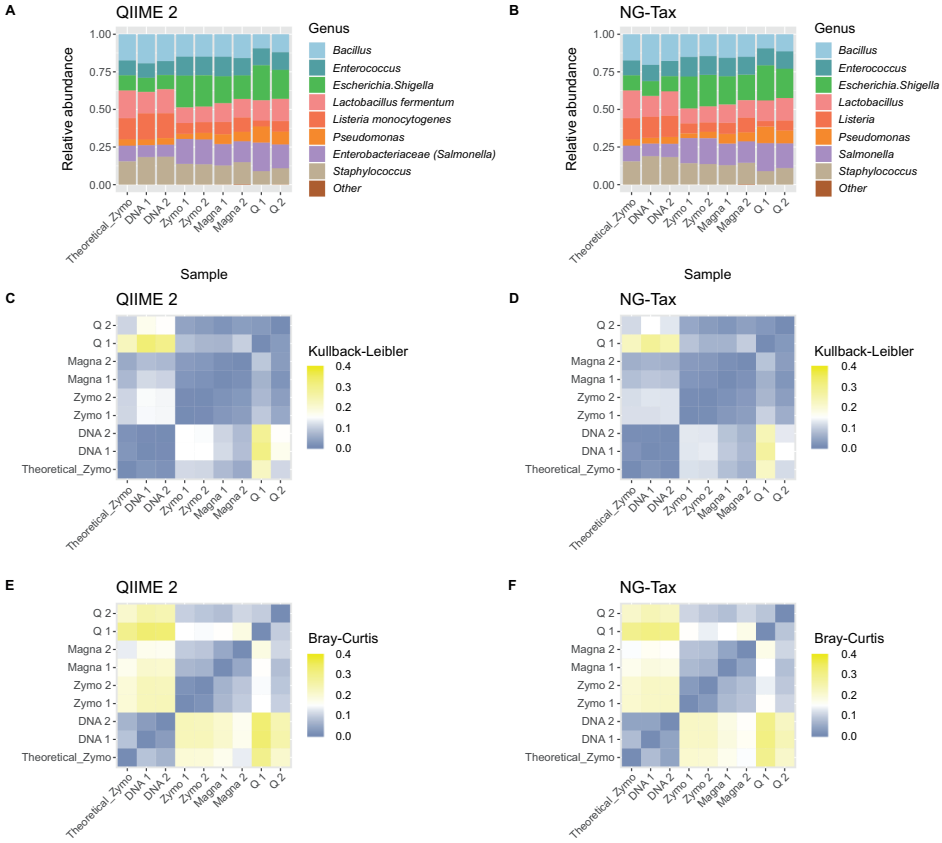


Figure 3: Compositional profiles at the genus level for QIIME 2 (A) and NG-Tax (B) for Zymo mock, theoretical composition is indicated in the first bar graph. Comparison of compositional profiles expressed by Kullback-Leibler divergence (C+D) and Bray-Curtis dissimilarity (E+F) per pipeline. QIIME 2 results are shown in figure C+E, NG-Tax results are shown in figure D+F. For both Kullback-Leibler and Bray-Curtis measures, 0 indicates an identical compositional profile, while higher numbers indicate more dissimilar profiles.

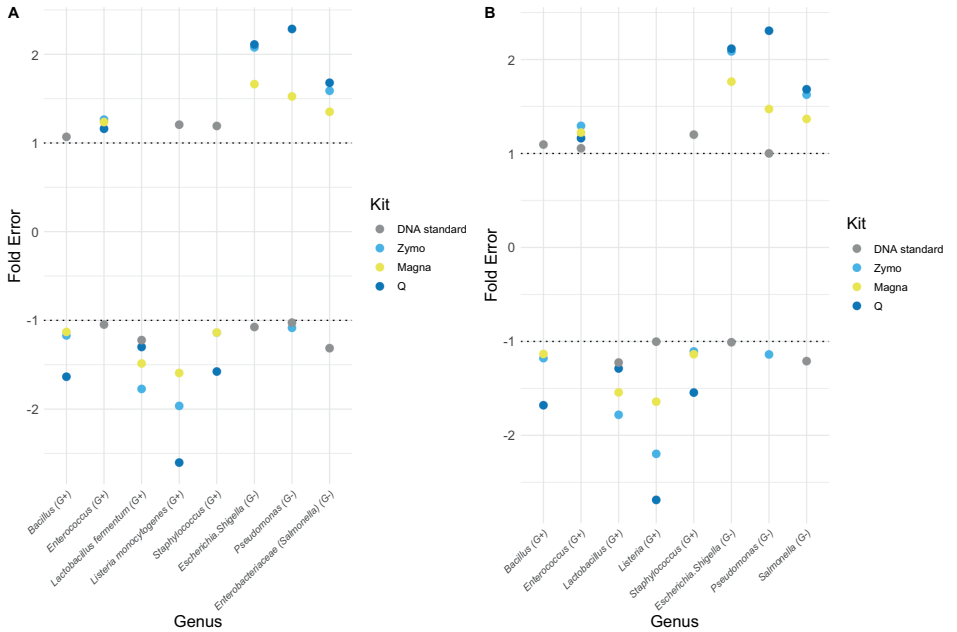


Figure 4: Fold error per bacterium as compared to theoretical values for QIIME 2 (A) and NG-Tax (B). Genera are ordered based on being Gram-positive or Gram-negative. A value above 1 represents overestimation, and a value below -1 represents underestimation.

Taken together, results obtained from the DNA standard indicate that QIIME 2 and NG-Tax perform equally well in general, except for overestimation of the OTU level richness and diversity when using NG-Tax. Results obtained from the Zymo mock, which is a better representation of the full procedure for a microbiota study, indicate that richness is most accurate at the genus level using Zymo or Q in combination with the NG-Tax pipeline. In addition, bacterial microbiota composition profiles are best retrieved using Magna, followed by Zymo, and are pipeline-independent.

In concordance with current literature [9] and independent of extraction method, a general underestimation of Gram-positive bacteria was observed, with *Enterococcus* being the sole exception (Fig 4). This is most likely due to incomplete cell wall lysis of Gram-positive bacteria. Based on the DNA standard and the Zymo mock, we conclude that Zymo and Magna in combination with either pipeline are the best performing combinations (Table S3). However, when high-throughput DNA extraction is required (e.g. for large cohort studies), Magna may be preferred from a practical point of view, although it overestimates richness independent of pipeline.

In general, overestimation of OTUs may stem from the 100% identity setting for clustering, combined with the natural divergence of the 16S rRNA gene [16, 17]. There is no current consensus on OTU identity setting, and cut-offs between 97% and 100% are most commonly used [18]. An advantage of the 100% cut-off is that unique taxa differing a single nucleotide are clustered into different OTUs. A disadvantage is that, as intragenomic diversity in the 16S rRNA gene is common within bacterial genomes, a 100% cut-off can lead to multiple OTUs stemming from a single bacterium and thereby inflate richness [17]. In addition, using a 100% cut-off can theoretically inflate richness due to sequencing errors and requires computational denoising. Apart from biological explanations, the different algorithms and internal filtering steps used in QIIME 2 and NG-Tax can affect the outcome for richness.

ATCC mock is recovered incorrectly, independent of extraction protocol or pipeline

The ATCC mock consists of 20 unique bacterial species, with four of them belonging to two genera (*Staphylococcus* and *Streptococcus*). Therefore, theoretical richness at OTU level would be 20, but eighteen at the genus level. In addition, these 20 unique bacterial species come from different environments, including gut, oral and skin microbiota.

No values close to the theoretical profiles for the ATCC mock for any extraction method/bioinformatic pipeline were observed, and one sample from Q consisted almost entirely of non-classifiable reads (Fig 5), indicating sample-related issues. *Bacillus* was highly overrepresented in all other samples, with a relative abundance over 30% in Zymo and Magna extracted samples, while 6.13% is expected. Curiously, after the first mechanical lysis step in Q, we could culture *Bacillus cereus* and *Cutibacterium acnes* (identification scores of 1.90 and 2.00, respectively), and *Bacillus cereus* (identification score 2.05) after mechanical lysis in Zymo. This is clinically important, as it means that infectious materials cannot be considered safe or non-infectious after mechanical lysis. As culturing of *B. cereus* indicates that cell wall lysis was incomplete, it would be expected that its relative abundance was underestimated, contrarily to what was observed. Another research group recently reported a similar overrepresentation of *Bacillus* in the ATCC community [19]. ATCC itself was also unable to retrieve abundances close to theoretical expectation, neither with 16S rRNA gene amplicon sequencing nor with shotgun sequencing [20]. Several reasons could explain this discrepancy between theoretical profiles and obtained profiles. For example, physical cell-to-cell interactions or presence of different metabolites may interfere with DNA extraction [16, 21]. Therefore, based on this synthetic community, no conclusions on the optimal extraction-pipeline combination could be made. This proposed positive control prompts the question whether mock communities are always reliable for assessing performance of DNA extraction methods. As can be observed from the Zymo mock, DNA extraction kits do not necessarily inflict observed deviations, but may rather be a result of mock community-specific properties.

Outcomes may depend on extraction kit / community type combination, indicating the potential necessity to use a positive control that strongly resembles the investigated microbiome.

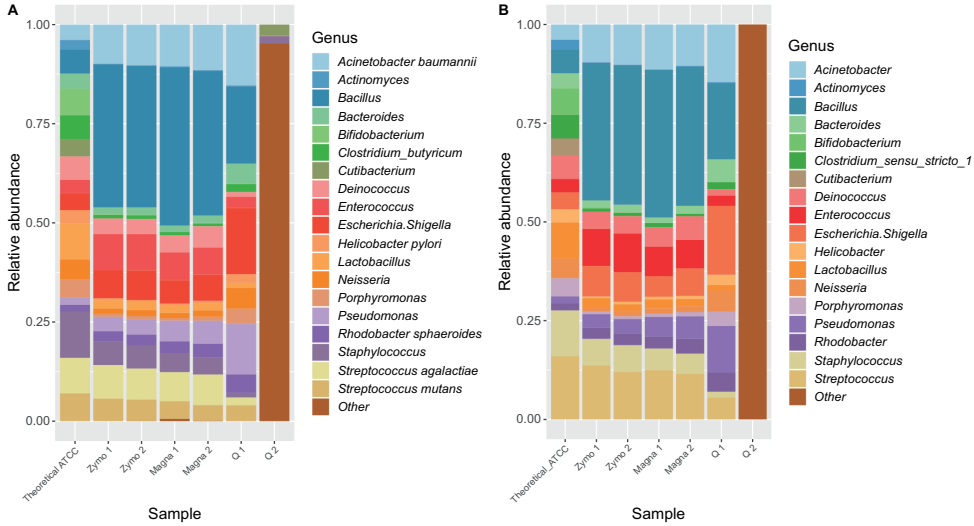


Figure 5: Compositional profiles at the genus level for QIIME 2 (A) and NG-Tax (B) for the ATCC mock. Genus '*Other*' is the sum of the relative abundance of all genera not listed in the legend.

Negative controls: inconsistently contaminated

Negative controls were taken along for each extraction method to check for kit-specific contaminants, which is especially relevant for deciding whether low-biomass samples contain real microbiota. Regarding Zymo, clear kit-contaminants were *Pseudomonas* and *Delftia* (Fig S2A+C), consistent across the different pipelines at the genus level, and with previous findings [11, 22]. For Magna and Q, specific contaminants were less obvious, although *Pseudomonas* was present. Generally, negative controls mostly consisted of genera commonly found in gut and oral microbiota, most of them also previously described as contaminants [11]. In addition, negative sequencing controls were taken along, and here no consistent contaminants could be observed (Fig S2B+D). Potential contamination sources are multifold, such as kit contamination, index hopping, or well-to-well contamination [23, 24]. Index-hopping is however not a likely source of contamination, as the negative control for Magna was sequenced in different lanes, and profiles look highly similar (Fig S2A+C). Additionally, we did not observe index-hopping in our positive controls.

One of the contaminants we identified has not been previously described as a contaminant,

namely *Clostridioides*. This likely represents *C. difficile*, and contamination by this bacterium can be explained by the fact that DNA extractions were performed in our National Reference Laboratory for *C. difficile*, which probably contains minor amounts of *C. difficile* spores during most time points. *C. difficile* contamination on laboratory surfaces has also recently been described in another clinical microbiology laboratory [25].

By incorporating this information with the Zymo positive controls, it can be concluded that Zymo and Magna are most optimal. Magna most accurately captured the expected community profile, while kit-specific contaminants are clear and easy to discriminate from biological signal using Zymo (Table S2).When investigating different biological sample types it would be ideal to use a kit for which kit contaminants do not overlap with the biological signal, e.g. *Pseudomonas* contamination when studying sputum samples from cystic fibrosis patients who are frequently colonized with *Pseudomonas* spp. However, this would require contaminants to be stable across batches, which has been shown to not be the case [22].

Automatic Magna extraction yields lowest DNA for biological samples

Twenty-seven biological samples were available per extraction protocol (Table S1) and Q was most successful in passing QC (22/27), followed by Zymo (20/27) and Magna (17/27) (Table S3), although differences were not statistically significant (Cochran's Q-test, p=0.178). QC passing was based on DNA concentration and intact genomic fragments. DNA concentrations were on average lowest for Magna, while yields were comparable between Q and Zymo (Figure S1). Processing of raw sequencing data from biological samples was performed using the NG-Tax pipeline at the genus level.

Fecal microbiota analysis is only slightly affected by the applied DNA extraction methods

DNA extracted from fecal samples using the three different protocols all passed QC. Magna, Zymo and Q achieved an average concentration of approximately 29 ng/µl, 111 ng/µl and 212 ng/µl, respectively (Fig. S1). While DNA yield varied between extraction methods, all were sufficient for sequencing. Microbiota profiles were comparable between extraction methods for each sample (Figure S3A). In addition, differences in compositional profiles were quantified using Kullback-Leibler divergence (Figure 6A). This heatmap shows that technical variation induced by DNA extraction method is much lower than biological variation between feces samples. Profiles of the feces donors contained many bacterial genera commonly present in fecal microbiomes [26, 27]. Healthy fecal microbiomes largely consist of Bacteroidetes and Firmicutes phyla (~90%), while Actinobacteria and Proteobacteria are present in smaller proportions. At the genus level, *Bacteroides*, *Prevotella* and *Faecalibacterium* are among the most prevalent genera [3], all of which were found in high abundance herein.

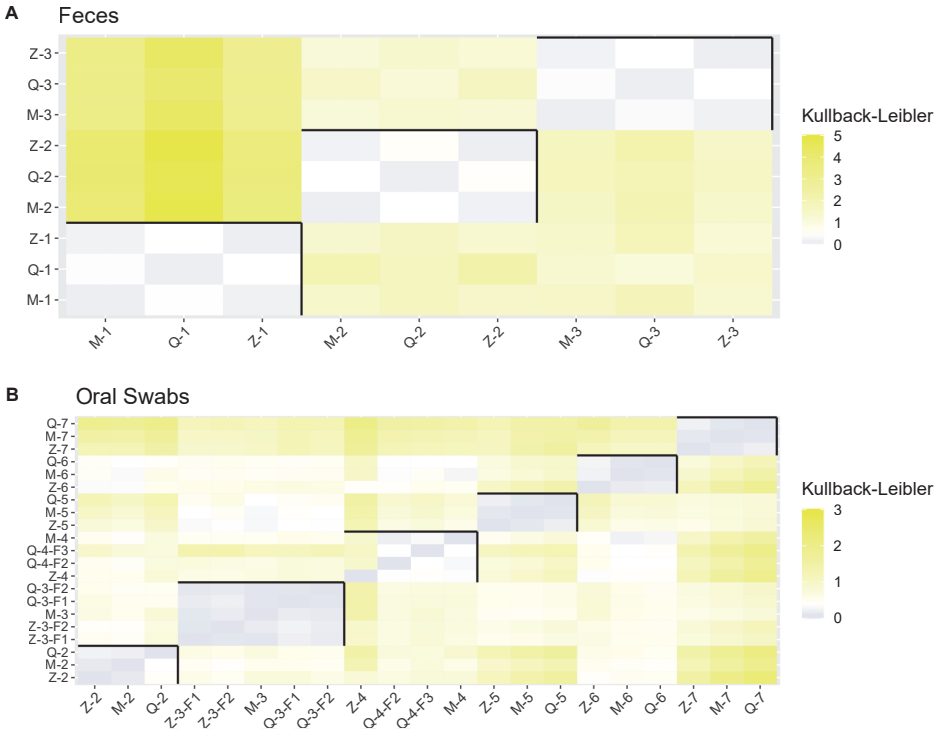


Figure 6: Kullback-Leibler divergence heatmap of feces (A) and oral swabs (B). Black lines group unique biological samples. Gray indicates highly similar composition, while yellow indicates divergence in composition. F1-F2-F3 represent samples which have been sequenced in duplicate, but on different flow cells.

Microbiota profiles of oral swabs are consistent, despite low DNA yields

Out of eighteen DNA extractions, fifteen extractions passed QC for oral swabs. Only for Zymo, all extractions passed QC. DNA yields were highly variable for all extraction methods, ranging from 0.12 to 6.34 ng/µl. Half of the extractions (nine/eighteen) yielded a concentration below one ng/µl. All compositional profiles were dominated by *Streptococcus*, *Prevotella* spp., *Haemophilus* and *Veillonella*, which was individual-independent. In addition, technical variation induced by DNA extraction and subsequent steps was lower than biological variation (Fig 6B). The oral microbiota, like the gut microbiota, is highly diverse. Nevertheless, a certain core of genera (e.g. *Streptococcus* spp. and *Prevotella* spp.*)* is present in most people, all of which were found in our study [3, 28, 29]. Together, the good QC passing rate, DNA concentrations and consistency of compositional profiles between extraction methods lead us to conclude that all three methods work well for oral swabs.

Applied methodology yields inconsistent results for the urine microbiota

During the last decade, microbiota studies showed that urine contains a bacterial microbiota [30, 31]. Despite using 30-40 ml of urine and centrifugation prior to extraction [32], we were not able to convincingly capture a urinary microbiota for all samples (Fig S3C). DNA concentrations were high for an infected sample (between thirteen and 42 ng/µl), but concentrations for the other samples were between 0.11 and 0.99 ng/µl. Six out of nine samples passed QC. For the infected sample with a high bacterial load, we were able to classify the cause of infection to *Enterobacteriaceae*, which is in agreement with the fact that most UTIs are caused by members of *Enterobacteriaceae*. One urine sample showed high similarity to negative controls for respective kits, with non-classifiable reads for Q and Magna, and high relative abundance of *Pseudomonas* for Zymo (Fig S3C). Another urine sample contained a high *Lactobacillus* relative abundance, which has previously been shown to be prevalent in urine samples [31]. *Lactobacillus* spp. could be cultured in 15% of urine samples collected by a transurethral catheter and was thereby the most prevalent genus cultured [31]. Another small-scale study found that in five out of six patients, *Lactobacillus* was detected in midstream urine samples and its relative abundance was between 22 and 80% [30]. In addition, presence of *Atopobium*, *Gardnerella*, *Prevotella* and *Anaerococcus* point towards an existing urinary microbiota [33]. However, *Pseudomonas,* a common Zymo kit contaminant, was still found in this urine sample, and for Magna more than 25% of reads could not be classified (Fig S3C). This could indicate that the biological signal is not much stronger than contamination, and therefore a mixed profile is observed. Further efforts and method optimization should be undertaken, although this can be difficult to implement in routine work [34]. In addition, culturing could be used as a follow-up method to confirm that contaminants are not viable bacteria, but rather bacterial DNA.

Saliva samples with long storage time and multiple freezing-thawing cycles seem unsuitable for microbiota research

DNA yield from included saliva samples was lower as compared to literature [35, 36] (Fig S1). Only a single DNA extraction had a concentration of slightly above one ng/µl (1.18; Table S4), while all other extractions had concentrations between 0.04 and 0.68 ng/µl. This is most likely associated with storage duration (~fifteen years) and the fact that samples were thawed and refrozen several times. This also explains why only three out of nine DNA extractions passed QC. The included saliva samples were chosen as investigators within our facility were interested to see if microbiota studies could be performed using these samples. Compositional profiles consisted of a mixture of genera present in the normal oral microbiota (*Oribacterium*, *Prevotella_7*, *Prevotella_9* and *Streptococcus*) [3], genera present in our negative controls (*Pseudomonas*, *Delftia*) and non-classifiable reads (Fig S3D). In combination with low DNA yields, it is likely that a mixture between biological signal and contamination signal is present. Therefore,

we consider the applied extraction methods unsuitable for saliva samples with a long duration of storage time and multiple freezing-thawing cycles.

The colorectal cancer microbiota present in biopsies was indistinguishable from negative controls or fecal microbiota

As colorectal cancer development has been associated with specific gut bacteria, we were interested to see if colorectal cancer tissue itself also contained bacteria [37, 38]. DNA concentrations were sufficient for all samples to pass QC, but extracted DNA was likely mostly human-derived. Two of three extraction methods were not successful, as samples extracted using Zymo and Magna showed high similarity to their respective negative controls (Fig S3E). Using Q, *Bacteroides*, *Fusobacterium* and *Gemella* were identified, all being previously associated with colorectal cancer development [37, 39]. Several gut commensals, including *Faecalibacterium* and *Escherichia-Shigella* were present in both the negative controls and these colorectal cancer samples. It is therefore difficult to discriminate whether these are contaminant bacteria, or whether they represent biological signal.

We hypothesized that by spinning down the material, the supernatant would contain more bacteria than the cancer tissue. DNA concentrations of supernatant were between 0.16 and 2.32 ng/µl, and seven out of nine DNA extractions passed QC (Table S4). For one sample, it was clear that across all methods many genera were observed which were present in negative controls (e.g. *Pseudomonas*), or reads could not be classified at all (Fig S3F). A second sample seemed to contain a real microbiota. Profiles were consistent across extraction methods, did not contain many contaminants and had specific bacteria previously linked to colorectal cancer (e.g. *Fusobacterium*) [37]. The third sample showed a profile reflecting a mix between biological signal and technical contamination. Profiles were consistent across methods and contained genera representative of a gut microbiota, but also contained non-classifiable reads and contamination. Therefore, profiles are likely a mixture of biological signal and technical contamination, and further optimization is necessary prior to using this sample type for experimental studies. We have the same recommendation for colorectal cancer sample types as for urine, as discussed above.

It remains unclear whether HPV-negative vulvar squamous cell carcinoma biopsies contain a bacterial microbiota

Vulvar squamous cell carcinoma (VSCC) has different etiological pathways, of which one is associated with human papilloma virus (HPV). The counterpart is non-virally related and is frequently associated to lichen sclerosis, a benign chronic inflammatory lesion and *TP53* mutations [40, 41]. We extracted DNA from HPV-negative VSCC tissue as a pilot study to determine if investigating the relationship between bacterial microbiota and HPV-negative VSCC would be potentially feasible. DNA concentrations were high

(Fig S1), only for three extractions below one ng/µl, and eight out of nine extractions passed QC. However, DNA was probably again largely human-derived. This was reflected in the obtained microbiota profiles, as most reads were not classified or the profiles showed high similarity to negative controls (e.g. high abundance of *Pseudomonas*) (Fig S3G). Therefore, it is unlikely that this cancer tissue contains bacteria, or bacteria are so lowly abundant that they are overshadowed by contamination load. In general, the vulvar microbiota has not been extensively studied. A recent study on vulvar microbiota observed that *Lactobacillus*, *Corynebacterium*, *Finegoldia*, *Staphylococcus and Anaerococcus* are most abundant on this body site, but the use of negative controls was not reported [42]. These genera are also part of the vaginal microbiota, and might be sampling contamination or reflect high similarity between vulvar and vaginal microbiota. A large amount of formalin-fixed VSCC materials are stored in a biobank at our facility. To investigate whether this sample collection could be used for microbiota profiling, DNA was extracted from three formalin-fixed VSCC samples. DNA concentrations were all below 0.3 ng/µl, and only two out of nine extractions passed QC (Table S4). One sample extracted with Q was excluded from further analysis, as no reads were present after sequencing. Extraction and sequencing of formalin-fixed material poses additional problems, as DNA molecules could be highly fragmented and too short for amplicon sequencing of the V4 region [43]. For Zymo, samples resembled negative controls, with *Delftia* and *Pseudomonas* being highly abundant (Fig S3H). The same samples had completely different microbiota profiles when using protocol Q or Magna. Both extraction methods showed genera commonly found in the lower urogenital tract, including *Streptococcus*, *Prevotella* and *Gordonia* [3, 27]. However, many of these genera were also detected in negative controls. In combination with low DNA yield and inconsistent profiles across extraction methods, we conclude that no reliable bacterial microbiota profile could be identified in these samples. For both VSCC types, we suggest the same way forward as for urine samples.

Sample groups with and without biological signal cluster apart

Lastly, we performed t-distributed stochastic neighbor embedding (t-SNE) clustering using Bray-Curtis measures on all samples used in the present study (Fig 7) [44]. Based on microbiota composition as measured by Bray-Curtis, t-SNE projects points in a two-dimensional space, while maintaining local structures present in high-dimensional space. Clear clusters could be identified for Zymo positive controls, feces, oral swabs and ATCC mock (all but one sample) (Fig 7). Other biological samples and negative controls were more dispersed throughout the plot, indicating that either more biological or technical variation was present. This is in agreement with our detailed analysis, showing that their microbiota cannot necessarily be distinguished from the negative controls. This highlights the importance of including negative controls in microbiota studies, which has previously been shown in two studies aiming to unravel the placental microbiota [45, 46],

and is increasingly recognized in the field. It is currently unclear whether a placental microbiota exists, but when comparing placental samples of healthy deliveries to included negative controls, microbiota compositions could not be distinguished [45, 46].

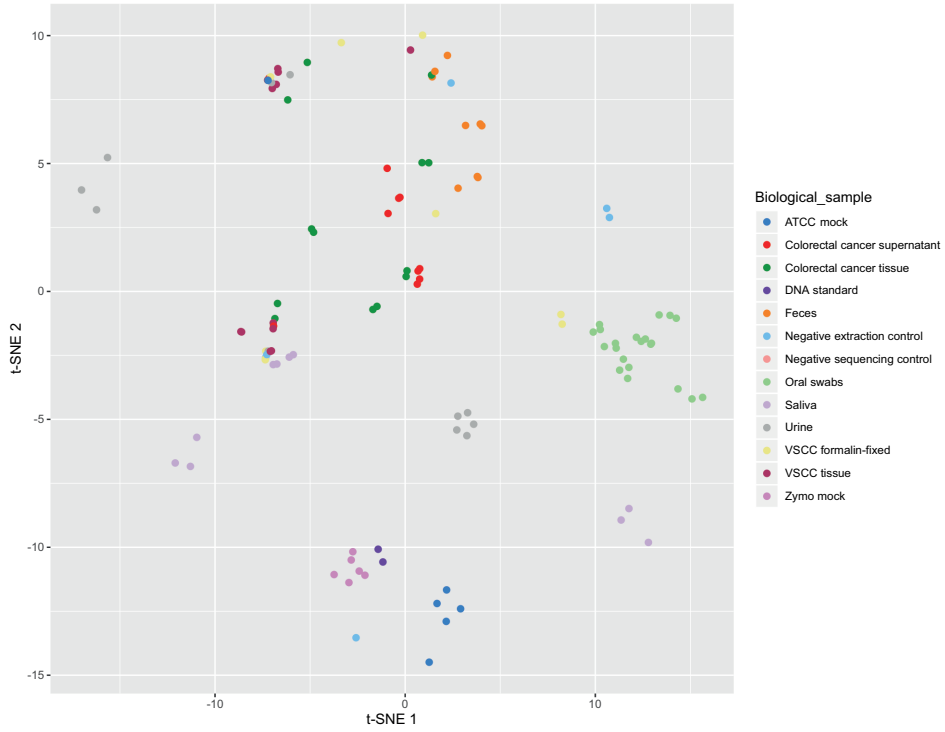


Figure 7: Bray-Curtis distance measures visualized by t-distributed stochastic neighbour embedding **(**t-SNE) for all samples. Each dot in the plot represents a single sample, and short distances between samples indicate high similarity.

Strengths and limitations

The current study had several strengths and limitations. By using a positive control of cell material with a corresponding DNA standard, we differentiated variation induced from sequencing procedures and DNA extraction. We demonstrate the importance of using positive and negative controls in microbiota studies, and show that negative controls are crucial for interpretation of low-biomass samples. Another strength of the study was that for several higher biomass biological samples (feces and oral swabs), we showed that technical variation was much smaller than biological variation. A shortcoming of the study is that we did not perform any other quantification next to 16S

rRNA gene sequencing (e.g. qPCR), which may be particularly useful for quality control of the ATCC mock. Furthermore, the current study used only three unique samples of most biological sample types. Especially for samples for which DNA extraction was challenging (urine samples, colorectal cancer supernatant), a higher number of unique samples would have allowed for a more thorough evaluation.

Conclusion

The current study evaluated three DNA extraction methods and two bioinformatic pipelines for bacterial microbiota profiling using several positive and negative controls, and a range of biological specimens. All three extraction methods quite accurately retrieved theoretical abundance of the Zymo mock, but not of the ATCC mock. For DNA extraction, we recommend using the Zymo and Magna protocol, since they showed good overall performance for all samples. Sequencing procedure only induced minor variation, as shown using a DNA standard. We furthermore showed that the NG-Tax and QIIME 2 pipelines perform equally well overall, each having their specific flaws.
By including negative controls and comparing these with low-biomass samples, we evaluated whether low-biomass samples consisted of technical noise, biological signal or a mixture. In most cases, identification of a unique microbiota was not achieved, highlighting the importance of negative controls and sufficiently sensitive methods. The results from this study can help other microbiome study groups to select an appropriate DNA extraction method and bioinformatic pipeline. Lastly, we hope this study contributes to further awareness of the usage of controls, especially when studying low-biomass samples.

Materials and Methods

Sample collection and pre-processing
Eight different biological specimens were included in this study, namely feces, urine, saliva, oral swabs, colorectal cancer tissue, colorectal cancer supernatant, vulvar squamous cell carcinoma tissue and formalin-fixed vulvar squamous cell carcinoma. Of each biological specimen, three unique samples were included. Only for oral swabs, six unique samples were included (Table S1). These samples were anonymized and treated according to the medical ethical guidelines described in the Code of Conduct for Proper Secondary Use of Human Tissue of the Dutch Federation of Biomedical Scientific Societies. A detailed overview of sample types, sample processing and storage conditions can be found in Table S1.

Mock communities and DNA standard

Two mock communities (ZymoBiomics Microbial Community Standard, Zymo Research, Irvine, California, USA and 20 Strain Even Mix Whole Cell Material ATCC® MSA2002™, ATCC, Wesel, Germany) were included as positive controls for DNA extraction. Exact composition and relative abundances of 16S rRNA gene copies was provided on the product sheet for ZymoBiomics Microbial Community standard (hereafter referred to as Zymo mock), while for ATCC® MSA2002™ (hereafter referred to as ATCC mock) we calculated expected 16S rRNA gene profiles based on genomic information (Table S2). ZymoBiomics Microbial Community DNA Standard (hereafter referred to as DNA standard) was taken along as a positive sequencing control.

DNA extraction

Procedures

Cancer samples were pre-processed for DNA extraction comparably to a recent study on pancreatic cancer microbiota [47], urine samples according to a recent publication on how to study urinary microbiota [32] and other samples according to in-house methods for sample processing (Table S1). For solid cancer samples, the beating steps during pre-processing were performed using a Qiagen TissueLyser LT (Qiagen Benelux, Venlo, the Netherlands) at 50Hz for one minute (Table S1). As single saliva samples did not contain sufficient volume for multiple extractions, several samples of the same individual were pooled to obtain the appropriate volume. DNA was extracted in duplicate from three unique samples for each biological material, only for oral swabs from six unique samples, and from the two mock communities. DNA was extracted using three different extraction protocols (see Protocols section), and for each protocol a negative extraction (no sample) was included in duplicate. The DNA standard was taken along in duplicate. DNA was quantified using a Qubit 3.0 Fluorometer (Invitrogen, Breda, the Netherlands) and the Qubit™ dsDNA HS Assay Kit (Thermo Fisher, Landsmeer, the Netherlands). A schematic overview of the study setup is shown in Figure 1.

DNA extraction protocols

Detailed protocols, including all minor adaptations, are present in Supplementary Methods. DNA extraction was performed using three methods: 1) the Quick-DNA Fecal/Soil Microbe kit (hereafter referred to as Zymo) (Zymo Research) according to manufacturer instructions with minor adaptations, 2) protocol Q (hereafter referred to as Q) [9] and 3) automated DNA extraction with MagNA Pure 96 ™ (hereafter referred to as Magna) (Roche Diagnostics, Almere, the Netherlands) using the MagNA Pure 96 DNA and viral NA small volume kit (Roche Diagnostics), according to standard operating procedures with minor adaptations. Mock communities were diluted to 10^4-10^5 cells per sample for extraction using Magna. For Q, several buffers and other materials were not provided in the kit and therefore purchased elsewhere, namely BeadBug™ prefilled

tubes with 2.0 mL capacity and 0.1 mm Zirconium beads (Sigma-Aldrich, Zwijndrecht, the Netherlands), RNase A, DNase and protease-free water (10 mg/mL) (Thermo Fisher, the Netherlands) and TE buffer (Thermo Fisher).

MALDI-TOF Mass Spectrometry (Biotyper)

To verify whether all bacteria of the ATCC mock were lysed after the first mechanical lysis step of both Zymo and Q, the lysate was plated on a tryptic soy agar plate containing 5% sheep (VWR International, Amsterdam, the Netherlands), and aerobically and anaerobically incubated at 37°C for five days. The MALDI Biotyper system was used (Bruker Daltonics, Germany) to identify the bacterial species. Samples were prepared in the following way: A bacterial colony was taken from the culturing plate and spread in duplicate on single spots on a Bruker polished steel targetplate. Subsequently, one µl of 70% formic acid was added on each single spot and when dried, one µl prepared Bruker Matrix HCCA according to clinical laboratory protocols was added per spot. The Bruker polished steel targetplate was then used for MALDI-TOF MS Biotyper analysis.

Library preparation and 16S rRNA gene amplicon sequencing

Of each duplicate DNA extraction from biological specimens, the duplicate with highest genomic DNA concentration was used for sequencing. Duplicate samples from controls were both sequenced. Quality control, library preparation and sequencing were performed by GenomeScan B.V. (Leiden, The Netherlands) using the NEXTflex™ 16S V4 Amplicon-Seq Kit (BiooScientific, TX, USA) and Illumina NextSeq 500 (paired-end, 150bp) according to their standard operating procedures. QC passing was based on intact genomic DNA and DNA concentrations measured by GenomeScan B.V. Therefore, those DNA concentrations were used for downstream analysis. Several samples were sequenced on multiple lanes, which is indicated in all relevant figures and tables.

Sequencing data analysis

Read filtering, operational taxonomic unit (OTU)-picking and taxonomic assignment were performed using two different bioinformatic pipelines, QIIME 2 and NG-Tax 0.4 [48, 49], both using the Silva_132_SSU Ref database for taxonomic classification [50]. For both pipelines, a read length of 120 bp was chosen based on quality of reads. The following settings were applied for QIIME 2: forward and reverse read length of 120 bp, quality control using Deblur, identity level of 100% (default). The following settings were applied for NG-Tax: forward and reverse read length of 120 bp, ratio OTU abundance of 2.0 (default), classify ratio of 0.9 (default), minimum threshold of 0.1% (default), identity level of 100% (default), error correction of 98.5 (default). Prior to the NG-Tax run, potential left over primers were removed with cutadapt v. 1.9.1 [51], in paired-end mode, with additional setting -e 0.2 (increased error tolerance, 20%). This setting was required since NG-Tax first creates a smaller custom database, based on the

used primers. During further processing, data has to be primer sequence free, as the primer sequence is removed from the smaller database. Furthermore, all sequences with any deviating barcode in the fastq header were changed to the original barcode to allow inclusion into the NG-Tax pipeline.

The obtained OTU-tables were filtered for OTUs with a number of sequences less than 0.005% of the total number of sequences [52]. Downstream analysis was performed in R (v3.6.1), mainly using the phyloseq (v.1.28.0), microbiome (v.1.6.0) and ggplot2 (v.3.2.0) packages [53-55]. Alpha diversity was computed at both the OTU and genus levels, while analysis of compositional profiles was performed at the genus level. Kullback-Leibler divergence and Bray-Curtis dissimilarity measure heatmaps were computed by first deleting genera that had a relative abundance of zero in all investigated samples (positive controls, feces and oral swabs) and subsequent calculation of the respective measure. All R code is available upon request from the corresponding author.

Notes

Data accessibility

All raw sequencing data used in the current study are deposited in the European Nucleotide Archive with accession number PRJEB34118.

Acknowledgements

We thank all collaborating partners who provided us with clinical biospecimens, namely Liz Terveer, Eric Berssenbrugge, Erik Giltay, Noel de Miranda, Jitske van den Bulk, Natalja ter Haar and Kim Kortekaas. We also thank Eric Claas for support with the use of the MagNA Pure 96 ™ and the clinical diagnostic laboratory for identification of bacteria with MALDI-TOF.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. BH and EK are supported by an unrestricted grant from Vedanta Biosciences Inc. EK has performed research for Cubist, Novartis and Qiagen, and has participated in advisory forums of Astellas, Optimer, Actelion, Pfizer, Sanofi Pasteur and Seres Therapeutics. The companies had no role in the study and writing of this manuscript.

Author contributions

QD, BH, AG, EK and RZ conceptualized and designed the study. QD and AG performed practical laboratory work. BH and RZ processed raw sequencing data. QD analyzed

data, prepared figures and wrote the manuscript under supervision of BH and RZ. All authors interpreted data, read and revised drafts of the manuscript, and approved the final version.

References

1. Guarner F, Malagelada JR. 2003. Gut flora in health and disease. Lancet 361:512-9.
2. Rajilic-Stojanovic M, Smidt H, de Vos WM. 2007. Diversity of the human gastrointestinal tract microbiota revisited. Environ Microbiol 9:2125-2136.
3. Human Microbiome Project C. 2012. Structure, function and diversity of the healthy human microbiome. Nature 486:207-14.
4. Sinha R, Abu-Ali G, Vogtmann E, Fodor AA, Ren B, Amir A, Schwager E, Crabtree J, Ma S, Microbiome Quality Control Project C, Abnet CC, 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.
5. 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.
6. Kennedy NA, Walker AW, Berry SH, Duncan SH, Farquarson FM, Louis P, Thomson JM, Consortium UIG, Satsangi J, Flint HJ, Parkhill J, Lees CW, Hold GL. 2014. The impact of different DNA extraction kits and laboratories upon the assessment of human gut microbiota composition by 16S rRNA gene sequencing. PLoS One 9:e88982.
7. 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.
8. Schirmer M, Ijaz UZ, D'Amore R, Hall N, Sloan WT, Quince C. 2015. Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. Nucleic Acids Res 43:e37.
9. Costea PI, Zeller G, Sunagawa S, Pelletier E, Alberti A, Levenez F, Tramontano M, Driessen M, Hercog R, Jung FE, Kultima JR, Hayward MR, Coelho LP, Allen-Vercoe E, Bertrand L, Blaut M, Brown JRM, Carton T, Cools-Portier S, Daigneault M, Derrien M, Druesne A, de Vos WM, Finlay BB, Flint HJ, Guarner 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.
10. Santiago A, Panda S, Mengels G, Martinez X, Azpiroz F, Dore J, Guarner F, Manichanh C. 2014. Processing faecal samples: a step forward for standards in microbial community analysis. BMC Microbiol 14:112.
11. Eisenhofer R, Minich JJ, Marotz C, Cooper A, Knight R, Weyrich LS. 2019. Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. Trends Microbiol 27:105-117.
12. Kuczynski J, Lauber CL, Walters WA, Parfrey LW, Clemente JC, Gevers D, Knight R. 2011. Experimental and analytical tools for studying the human microbiome. Nat Rev Genet 13:47-58.
13. Zeeuwen P, Boekhorst J, Ederveen THA, Kleerebezem M, Schalkwijk J, van Hijum S, Timmerman HM. 2017. Reply to Meisel et al. J Invest Dermatol 137:961-962.
14. Jovel J, Patterson J, Wang W, Hotte N, O'Keefe S, Mitchel T, Perry T, Kao D, Mason AL, Madsen KL, Wong GK. 2016. Characterization of the Gut Microbiome Using 16S or Shotgun Metagenomics. Front Microbiol 7:459.
15. Kullback S, Leibler RA. 1951. On Information and Sufficiency. The Annals of Mathematical Statistics 22:79-86.
16. Hornung BVH, Zwittink RD, Kuijper EJ. 2019. Issues and current standards of controls in microbiome research. FEMS Microbiol Ecol 95.
17. Sun DL, Jiang X, Wu QL, Zhou NY. 2013. Intragenomic heterogeneity of 16S rRNA genes causes overestimation of prokaryotic diversity. Appl Environ Microbiol 79:5962-9.
18. Patin NV, Kunin V, Lidstrom U, Ashby MN. 2013. Effects of OTU clustering and PCR artifacts on microbial diversity estimates. Microb Ecol 65:709-19.
19. Scherz V, Aeby S, Bertelli C, Greub G. 2019. Microbiota profiling of saliva: unexpected results of a preliminary study, abstr ECCMID, Amsterdam, 2019.
20. ATCC. 2018. Microbiome standard & research solutions. https://www.lgcstandards-atcc.org/~/media/A36E4E940E3F40B1A5E14D6851E9348F.ashx. Accessed April 10th 2019.

21. Angelakis E, Bachar D, Henrissat B, Armougom F, Audoly G, Lagier JC, Robert C, Raoult D. 2016. Glycans affect DNA extraction and induce substantial differences in gut metagenomic studies. Sci Rep 6:26276.
22. 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.
23. Costello M, Fleharty M, Abreu J, Farjoun Y, Ferriera S, Holmes L, Granger B, Green L, Howd T, Mason T, Vicente G, Dasilva M, Brodeur W, DeSmet T, Dodge S, Lennon NJ, Gabriel S. 2018. Characterization and remediation of sample index swaps by non-redundant dual indexing on massively parallel sequencing platforms. BMC Genomics 19:332.
24. Minich JJ, Sanders JG, Amir A, Humphrey G, Gilbert JA, Knight R. 2019. Quantifying and Understanding Well-to-Well Contamination in Microbiome Research. mSystems 4.
25. Reigadas E, Vazquez-Cuesta S, Onori R, Villar-Gomara L, Alcala L, Marin M, Martin A, Munoz P, Bouza E. 2019. Clostridioides difficile contamination in the environment of a clinical microbiology laboratory and laboratory workers. Clin Microbiol Infect doi:10.1016/j.cmi.2019.06.027.
26. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto JM, Hansen T, Le Paslier D, Linneberg A, Nielsen HB, Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Li S, Jian M, Zhou Y, Li Y, Zhang X, Li S, Qin N, Yang H, Wang J, Brunak S, Dore J, Guarner F, Kristiansen K, Pedersen O, Parkhill J, Weissenbach J, Bork P, Ehrlich SD, Wang J. 2010. A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464:59-65.
27. Lloyd-Price J, Abu-Ali G, Huttenhower C. 2016. The healthy human microbiome. Genome Med 8:51.
28. Lamont RJ, Koo H, Hajishengallis G. 2018. The oral microbiota: dynamic communities and host interactions. Nat Rev Microbiol 16:745-759.
29. Aas JA, Paster BJ, Stokes LN, Olsen I, Dewhirst FE. 2005. Defining the normal bacterial flora of the oral cavity. J Clin Microbiol 43:5721-32.
30. Wolfe AJ, Toh E, Shibata N, Rong R, Kenton K, Fitzgerald M, Mueller ER, Schreckenberger P, Dong Q, Nelson DE, Brubaker L. 2012. Evidence of uncultivated bacteria in the adult female bladder. J Clin Microbiol 50:1376-83.
31. Hilt EE, McKinley K, Pearce MM, Rosenfeld AB, Zilliox MJ, Mueller ER, Brubaker L, Gai X, Wolfe AJ, Schreckenberger PC. 2014. Urine is not sterile: use of enhanced urine culture techniques to detect resident bacterial flora in the adult female bladder. J Clin Microbiol 52:871-6.
32. Karstens L, Asquith M, Caruso V, Rosenbaum JT, Fair DA, Braun J, Gregory WT, Nardos R, McWeeney SK. 2018. Community profiling of the urinary microbiota: considerations for low-biomass samples. Nat Rev Urol 15:735-749.
33. Bucevic Popovic V, Situm M, Chow CT, Chan LS, Roje B, Terzic J. 2018. The urinary microbiome associated with bladder cancer. Sci Rep 8:12157.
34. Boers SA, Hays JP, Jansen R. 2017. Novel micelle PCR-based method for accurate, sensitive and quantitative microbiota profiling. Sci Rep 7:45536.
35. Lim Y, Totsika M, Morrison M, Punyadeera C. 2017. The saliva microbiome profiles are minimally affected by collection method or DNA extraction protocols. Sci Rep 7:8523.
36. Vesty A, Biswas K, Taylor MW, Gear K, Douglas RG. 2017. Evaluating the Impact of DNA Extraction Method on the Representation of Human Oral Bacterial and Fungal Communities. PLoS One 12:e0169877.
37. Kostic AD, Chun E, Robertson L, Glickman JN, Gallini CA, Michaud M, Clancy TE, Chung DC, Lochhead P, Hold GL, El-Omar EM, Brenner D, Fuchs CS, Meyerson M, Garrett WS. 2013. Fusobacterium nucleatum potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. Cell Host Microbe 14:207-15.
38. He Z, Gharaibeh RZ, Newsome RC, Pope JL, Dougherty MW, Tomkovich S, Pons B, Mirey G, Vignard J, Hendrixson DR, Jobin C. 2019. Campylobacter jejuni promotes colorectal tumorigenesis through the action of cytolethal distending toxin. Gut 68:289-300.
39. Kwong TNY, Wang X, Nakatsu G, Chow TC, Tipoe T, Dai RZW, Tsoi KKK, Wong MCS, Tse G, Chan MTV, Chan FKL, Ng SC, Wu JCY, Wu WKK, Yu J, Sung JJY, Wong SH. 2018. Association Between Bacteremia From Specific Microbes and Subsequent Diagnosis of Colorectal Cancer. Gastroenterology 155:383-390.e8.
40. Abdulrahman Z, Kortekaas KE, De Vos Van Steenwijk PJ, Van Der Burg SH, Van Poelgeest MI. 2018. The immune microenvironment in vulvar (pre)cancer: review of literature and implications for immunotherapy. Expert Opin Biol Ther 18:1223-1233.
41. van de Nieuwenhof HP, van der Avoort IA, de Hullu JA. 2008. Review of squamous premalignant vulvar

lesions. Crit Rev Oncol Hematol 68:131-56.
42. Vongsa R, Hoffman D, Shepard K, Koenig D. 2019. Comparative study of vulva and abdominal skin microbiota of healthy females with high and average BMI. BMC Microbiol 19:16.
43. 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.
44. van der Maaten L, Hinton G. 2008. Visualizing Data using t-SNE. J Mach Learn Res 9:2579-2605.
45. Lauder AP, Roche AM, Sherrill-Mix S, Bailey A, Laughlin AL, Bittinger K, Leite R, Elovitz MA, Parry S, Bushman FD. 2016. Comparison of placenta samples with contamination controls does not provide evidence for a distinct placenta microbiota. Microbiome 4:29.
46. de Goffau MC, Lager S, Sovio U, Gaccioli F, Cook E, Peacock SJ, Parkhill J, Charnock-Jones DS, Smith GCS. 2019. Human placenta has no microbiome but can contain potential pathogens. Nature doi:10.1038/s41586-019-1451-5.
47. Pushalkar S, Hundeyin M, Daley D, Zambirinis CP, Kurz E, Mishra A, Mohan N, Aykut B, Usyk M, Torres LE, Werba G, Zhang K, Guo Y, Li Q, Akkad N, Lall S, Wadowski B, Gutierrez J, Kochen Rossi JA, Herzog JW, Diskin B, Torres-Hernandez A, Leinwand J, Wang W, Taunk PS, Savadkar S, Janal M, Saxena A, Li X, Cohen D, Sartor RB, Saxena D, Miller G. 2018. The Pancreatic Cancer Microbiome Promotes Oncogenesis by Induction of Innate and Adaptive Immune Suppression. Cancer Discov 8:403-416.
48. Ramiro-Garcia J, Hermes GDA, Giatsis C, Sipkema D, Zoetendal EG, Schaap PJ, Smidt H. 2016. NG-Tax, a highly accurate and validated pipeline for analysis of 16S rRNA amplicons from complex biomes[version 1; referees: 2 approved with reservations, 1 not approved]. F1000Research 5.
49. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet C, 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 E, Da Silva R, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley G, Janssen S, Jarmusch AK, Jiang L, Kaehler B, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu Y-X, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton J, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson IIMS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG. 2018. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. PeerJ Preprints 6:e27295v2.
50. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590-6.
51. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17:3.
52. Bokulich NA, Subramanian S, Faith JJ, Gevers D, Gordon JI, Knight R, Mills DA, Caporaso JG. 2013. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. Nat Methods 10:57-9.
53. McMurdie PJ, Holmes S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 8:e61217.
54. Lahti L, Shetty S. 2017. Tools for microbiome analysis in R. Microbiome package version 1.2.1,
55. Wickham H. 2009. ggplot2: Elegant Graphics for Data Analysis. Springer Science & Business Media.