

# A unique and reliable fecal DNA extraction method for 16S rRNA gene and shotgun metagenomic sequencing in the analysis of the human gut microbiome

**Céline Elie**

Bioaster Technology Research Institute

**Magali Perret**

Bioaster Technology Research Institute

**Karen Louis**

Bioaster Technology Research Institute

**Asmaà Fritah-Lafont**

Bioaster Technology Research Institute

**Philippe Leissner**

Bioaster Technology Research Institute

**Carole Vachon**

bioMerieux SA

**Hervé Rostaing**

bioMerieux SA

**Frédéric Reynier**

Bioaster Technology Research Institute

**Gaspard Gervasi**

bioMerieux SA

**Adrien Saliou** (✉ [adrien.saliou@bioaster.org](mailto:adrien.saliou@bioaster.org))

Bioaster Technology Research Institute <https://orcid.org/0000-0002-7681-4989>

---

## Research

**Keywords:** human microbiome, gut microbiota, stool samples, DNA extraction, standardization, metagenomics, 16S rRNA gene amplicon sequencing, shotgun metagenomic sequencing

**Posted Date:** August 6th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-52279/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** The gut microbiome is widely analyzed using high-throughput sequencing, such as 16S rRNA gene amplicon sequencing and shotgun metagenomic sequencing (SMS). DNA extraction is known to have a large impact on the metagenomic analyses. The aim of this study was to select a unique and best performing DNA extraction protocol for both metagenomic sequencing methods. In that context, four commonly used DNA extraction methods were compared for the analysis of the gut microbiota. Commercial versions were evaluated against modified protocols using a stool preprocessing device (SPD, bioMérieux) in order to facilitate DNA extraction. Stool samples from nine healthy volunteers and nine patients with a *Clostridium difficile* infection were extracted with all protocols and sequenced with both metagenomic methods. Protocols were ranked using wet- and dry-lab criteria, including quality controls of the extracted genomic DNA, alpha-diversity, accuracy using a mock community of known composition and repeatability across technical replicates.

**Results:** Independently of the sequencing methods used, SPD significantly improved efficiency of the four tested protocols compared with their commercial version, in terms of extracted DNA quality, accuracy of the predicted composition of the microbiota (notably for Gram-positive bacteria), sample *alpha*-diversity, and experimental repeatability. The best overall performance was obtained for the S-DQ protocol, SPD combined to the DNeasy PowerLyser PowerSoil protocol from QIAGEN.

**Conclusion:** Based on this evaluation, we recommend to use the S-DQ protocol, to obtain standardized and high quality extracted DNA in the human gut microbiome studies.

## Background

In recent years, advances in next-generation sequencing (NGS) have revolutionized the analysis of complex microbial ecosystems including the gut microbiota, leading to advanced understanding of its role in health and disease [1–6]. Alterations in the composition and diversity of the gut microbiota communities have been correlated with a large number of diseases, such as inflammatory bowel disease [7–11], irritable bowel syndrome [12, 13], metabolic disorders (e.g, type 2 diabetes (T2D), obesity and nonalcoholic fatty liver disease (NALFD)) [14–19], and more recently, cancer [20–30].

Nevertheless, metagenomic methods are known to be prone to errors at different steps of the workflow, from sample collection [31–34], DNA extraction [35, 36], library preparation and sequencing [37, 38] to data analysis [39, 40]. In order to facilitate the implementation of these methods into clinical routine practice, standardized methods are urgently needed [41–45].

The choice of the DNA extraction method has been demonstrated to strongly affect the detection of bacterial communities [35, 36, 46–48]. DNA extraction is a sophisticated process, including sample weighing, sample homogenization, bacterial cell lysis, and DNA purification, for which each step still requires improvements and guidelines. For instance, the standard weighing procedure can be tedious and time-consuming to collect the same volume of fecal material for all samples. Sample homogenization

could also have an impact on the bacteria that can be detected [49, 50]. Surprisingly, few studies have reported the use of commercial devices to standardize the handling of fecal samples prior to DNA extraction. [31, 51, 52]. Also, as the cell wall of Gram-positive bacteria is composed of a thick layer of peptidoglycan, bead-beating is now recommended to improve the lysis [53, 54]. Nevertheless, beads can vary in size and material (e.g, ceramic, glass, zirconia or silica), which may also play a role in the lysis efficiency. Even if commercial solutions provide standardized methods for bacterial lysis and DNA purification, some laboratories still use in-house protocols, making difficult the selection of one gold-standard protocol.

Recently, twenty-one DNA extraction protocols were compared in a multicentric study across three continents, using shotgun metagenomic sequencing (SMS) [35]. From the analysis of two healthy individuals, the authors proposed a QIAGEN protocol, named Q, as the standardized reference protocol for DNA extraction in human gut microbiome studies. However, it has been shown that different fecal samples may vary in terms of bacterial composition (gram-positive vs gram-negative cells) [35, 55], microbial load (high vs low bacterial cells per fecal material) [56, 57], disease-related clinical status (healthy vs sick individuals) [58–60] and stool consistency (separate hard lumps vs watery) [61–63]. A comparison study with a higher number of individuals including both healthy and sick donors would be of both clinical and technological interests to address the variability and heterogeneity of fecal samples.

Although SMS has the potential to deeply investigate microbial communities [64, 65], amplicon sequencing targeting the 16S rRNA gene is often the preferred and the most cost-effective metagenomic method in the analysis of clinical cohorts [66, 67]. Obviously, these sequencing methods have their own limitations and biases, which are important to consider for the selection of one DNA extraction protocol in human gut microbiome studies.

To address these considerations, our study evaluated four commercially available DNA extraction methods, using both SMS and 16S rRNA amplicon sequencing. These protocols were tested as recommended by the manufacturers, but also with an upstream stool preprocessing device (SPD), designed to facilitate DNA extraction [51]. The protocols were evaluated according to wet-lab as well as dry-lab criteria, using nine healthy individuals and nine *Clostridium difficile* infected patients.

## Methods

### Stool samples.

Fecal samples from nine healthy volunteers and nine patients with *Clostridium difficile* infection (CDI) were provided by a certified testing laboratory in France and tested for *Clostridium difficile* toxins. Upon reception, each fecal sample was freshly aliquoted into 24 tubes (8 protocols x 3 replicates) and frozen at -80 °C until extraction, the - 80 °C storage being known to maintain a stable microbial community for long-term period [68].

## Microbial mock community.

The microbial mock community was prepared by mixing nine bacteria (**Table 2**), including four easy-to-lyse Gram-negative bacteria (*Pseudomonas aeruginosa*, *Escherichia coli*, *Salmonella enterica* and *Rhizobium radiobacter*) and five more difficult to lyse Gram-positive bacteria (*Lactobacillus fermentum*, *Enterococcus faecalis*, *Staphylococcus aureus*, *Listeria innocua* and *Bacillus subtilis*). Bacterial cells were obtained from ATCC and cultivated according to ATCC's recommendations. The number of viable cells was estimated by plate counting. The mock community was prepared by mixing between  $2.7 \times 10^7$  and  $3.6 \times 10^8$  cells of nine bacteria and stored at  $-80^\circ\text{C}$  until extraction.

## DNA extraction.

Four commercial protocols were compared in this study, according to the manufacturers' recommendations: the NucleoSpin Soil kit (#740780.50, protocol May 2016/Rev. 06, Macherey-Nagel), the DNeasy PowerLyzer PowerSoil Kit (#12855-100, protocol 07272016, QIAGEN), the QIAamp Fast DNA Stool kit (#51604, QIAGEN, protocol modified from [36]) and the ZymoBIOMICS DNA Mini kit (#D4300, protocol 1.1.0, ZymoResearch). These protocols were also tested in combination with a stool preprocessing device (SPD, #421061, bioMérieux, [51]). This device was designed to facilitate and standardize fecal sample preparation before nucleic acid extraction. It includes a spoon for a 200 mg calibrated sample and a vial containing a buffer for sample resuspension, glass beads for homogenization and two filters for retaining fecal debris. After 5 minutes hands-on-time, the filtrate is ready-to-use for downstream DNA extraction. Protocols of extraction methods as well as SPD are detailed in **Supplementary Methods**. DNA was extracted in triplicates from fecal samples and from the microbial community. A260/A280 ratio was assessed using the DropSense 96 system (Trinean). Genomic DNA size was assessed using the Genomic DNA ScreenTape (#5067–5364, Agilent) on the 2200 TapeStation system (Agilent). DNA concentrations were estimated using the QuantiFluor One dsDNA kit (#E4870, Promega) with the GloMax system (Promega).

## 16S rRNA gene library preparation and sequencing.

16S rRNA gene libraries was prepared according to Illumina's protocol (# 15044223 RevB, [69]). In order to minimize the risk of cross-contamination and pipetting errors, the workflow was automated using a high-throughput liquid handler; the Freedom EVO NGS workstation (TECAN) [70]. Briefly, V3-V4 hypervariable regions were first amplified from 12.5 ng of genomic DNA, using the following primers: (i) Forward Primer: TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGAGGCAGC-AG and (ii) Reverse Primer: GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVG-GGTWCTAAT and 2x KAPA HiFi HotStart ReadyMix (Kapa Biosystems). PCR cycle conditions were  $95^\circ\text{C}$  for 3 min, 25 cycles of ( $95^\circ\text{C}$  for 30 seconds,  $55^\circ\text{C}$  for 30 seconds  $72^\circ\text{C}$  for 30 seconds), then a final extension of  $72^\circ\text{C}$  for 5 minutes. The libraries were purified using AMPure XP beads (Beckman Coulter). Dual indexes and sequencing adapters

from the Illumina Nextera XT index kits (Illumina) were added in a second PCR using 2x KAPA HiFi HotStart ReadyMix (Kapa Biosystems). Cycle conditions were 95 °C for 3 minutes, 8 cycles of (95 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 30 seconds), then a final extension of 72 °C for 5 min. Ready-to-sequence libraries were purified using AMPure XP beads (Beckman) and quantified by fluorescence using the QuantiFluor One dsDNA kit (# E4870, Promega) with the GloMax system (Promega). Quality control was performed using a 2200 TapeStation system with the DNA 1000 screenTape (# 5067–5582, Agilent). The library pool was quantified by qPCR with the KAPA Library Quantification Kit for Illumina platforms (Kapa Biosystems). Sequencing was performed on a MiSeq system (Illumina) with the MiSeq Reagent v3 kit (600 cycles) in a 2 × 300 bp mode.

## Shotgun metagenomic library preparation and sequencing.

SMS libraries were prepared using the Nextera XT DNA Library Preparation Kit (# FC-131-1096, Illumina), following Illumina's instructions (protocol # 15031942 v03 February 2018). Briefly, 1 ng of genomic DNA was used for the tagmentation reaction for a total volume of 20 µl. After 5 min at 55 °C, the reaction was stopped by adding 5 µl of the Neutralize Tagment (NT) Buffer. A limited-cycle PCR amplification was then performed to amplify the tagmented DNA (addition of 15 µl of Nextera PCR Master Mix (NPM)) and to add Illumina sequencing adapters (addition of 5 µl of both Index 1 primer and Index 2 primer from the Nextera XT index kit, Illumina) for a total volume of 50 µl. The following PCR cycle program was used: 72 °C for 3 minutes, 95 °C for 30 seconds, 12 cycles of (95 °C for 10 seconds, 55 °C for 30 seconds, 72 °C for 30 seconds), 72 °C for 5 minutes. SMS libraries were quantified using the QuantiFluor One dsDNA kit (# E4870, Promega) with the GloMax system (Promega). The quality of libraries was assessed using the High Sensitivity DNA kit on the Agilent 2100 Bioanalyzer. Sequencing was performed on a NextSeq500 system (Illumina) with the NextSeq 500/550 High Output v2 kit (300 cycles) in 2 × 150 bp.

## 16S rRNA gene profiling.

After quality control with FastQC (v0.11.3), overlapping paired-end reads were merged with PEAR (v0.9.10). Quality trimming and filtering of amplicons were performed with trimmomatic (v0.36) and SGA (v0.9.9). The following parameters were used: maximum of 20 low-quality base calls in the whole sequence, no ambiguous bases (N), a minimum Phred quality score of 15 over a 4 bp sliding window, a minimum average quality score of 25, and a minimum length of 100 bp after trimming. During the PCR amplification process, artefactual sequences can be generated from multiple parent sequences, and are called chimeric sequences. These sequences were removed using the UCHIME *de novo* algorithm, which is integrated in the usearch6.1 QIIME pipeline. Sequences were grouped into operational taxonomic units (OTUs) based on a 97% identity threshold using usearch6.1 through the "pick\_open\_reference\_otus.py" QIIME script. OTUs recruiting less than 0.005% of the total number of sequences were filtered out, as recommended by [71]. Taxonomic annotation was performed in QIIME using the RDP classifier trained on SILVA rRNA reference database (v123).

# Shotgun metagenomic profiling.

Reads were trimmed and filtered based on the sequence quality and length, exactly as for the 16S rRNA gene data analysis. Reads were annotated using the Centrifuge software [72] on the NCBI Refseq genome database (v.2018, complete genomes and scaffolds [73]).

## Statistical analysis.

Annotated tables were normalized by a “total count” method (at the OTU level for 16S rRNA gene sequencing, at the species level for SMS). All subsequent analyses were performed in R (version 3.3.1). The repeatability was assessed by calculating a coefficient of variation for each bacterium present in all replicates of a condition for every patient. *Alpha*-diversity (Shannon indices) was calculated for each sample using the vegan package. The accuracy of the protocols was evaluated on the mock community sample by calculating the Euclidean distance between expected and predicted abundances (log<sub>2</sub>), using the “philtropy” R package. Differentially abundant bacteria between protocols with or without the SPD were identified using the DESeq2 package. For each criterion (except for *alpha*-diversity), the statistical significance of the differences between protocols was computed with a pairwise Wilcoxon rank test. For multiple comparisons, *P-values* were corrected by Benjamini Yakuteli correction and adjusted *P-values* below 0.05 were considered statistically significant. The *alpha*-diversity values varied greatly from one patient to another, so the patient effect was controlled in a linear model using the “limma” package, and statistics were computed with the empirical Bayes method.

## Results

### Study design

In our study, four commercial DNA extraction protocols were evaluated based on the supplier’s recommendations: the NucleoSpin Soil kit (Macherey-Nagel, named MN), the DNeasy PowerLyzer PowerSoil kit (Qiagen, named DQ), the QIAamp Fast DNA Stool kit (Qiagen, named QQ), and the ZymoBIOMICS DNA Mini kit (ZymoResearch, named Z). In order to facilitate the first steps of DNA extraction, they were also tested with an upstream stool processing device, named SPD (See **Supplementary Methods** for detailed protocols). The resulting protocols were named as follows: S-MN stands for SPD + MN, S-DQ for SPD + DQ, S-QQ for SPD + QQ and S-Z for SPD + Z.

We analyzed fecal samples from nine healthy volunteers and nine patients suffering from CDI. A defined mixture of bacterial species (mock community) was also prepared and sequenced to assess the efficiency and accuracy of DNA extraction, by comparing the observed bacterial abundances to the theoretical ones. DNA extraction protocols were first compared using 16S rRNA gene amplicon sequencing for a total of 456 samples (18 fecal samples and 1 mock community, in triplicates). DNA extraction protocols were also evaluated using SMS as a read-out for a reduced number of samples (n =

56), including fecal samples from six individuals (three healthy individuals and three CDI + patients) and the mock community (Fig. 1).

## Quality and quantity of extracted DNA

When selecting a DNA extraction protocol, sufficient genomic DNA of high quality is desirable for preparing metagenomics libraries. In the present study, we evaluated the DNA yield, DNA fragment size and DNA quality. A protocol that performs poorly on these criteria would likely skew measured bacterial compositions, as only a small portion of bacterial communities present in the original sample would be analyzed.

Considerable variability was found in the extraction yield for the tested protocols (Fig. 2.a), which is in line with previous studies [36]. Except for MN, DNA extraction protocols in combination with SPD seemed to recover as much or more DNA yields compared to their commercial versions. Notably, increases were observed for S-QQ (p-value < 0.1) and S-Z (p-value < 0.05), compared to QQ and Z, respectively. A same DNA yield was recovered for the protocol DQ with and without the use of SPD (p-value > 1). SPD seemed to negatively affect the extraction yield when coupled with the protocol MN (p-value < 0.01). Out of the eight extraction protocols tested, protocols S-MN and Z significantly recovered the lowest DNA concentrations.

In practice, a best performing protocol would be a protocol for which the highest number of samples could be prepared for sequencing. Here, for a given protocol, we measured the percentage of samples whose DNA concentration was superior to 5 ng/μl, threshold corresponding to the minimal DNA concentration recommended to prepare 16S rRNA gene sequencing libraries (**Table 1**). In our hands, none of the tested protocols was able to retrieve DNA for all the samples, with a concentration superior to this threshold. Except for S-MN, the best performances were observed when the protocols were combined with SPD. S-Z recovered enough DNA material for 88% of samples, followed by MN (86%), S-QQ (82%) and S-DQ (81%).

Regarding the fragment size of DNA, variations were also observed between the extraction protocols. QQ and MN protocols yielded the shortest DNA fragments with a median size around 12,000 bp, which was shorter than S-QQ (p-value > 0.1) and significantly shorter than the other ones (p-value < 0.01, Fig. 2.b). The longest DNA fragment sizes were observed for S-MN with an average size of 21,000 bp, followed by DQ, S-DQ and Z with DNA fragments around 18,000 bp (p-value > 0.1).

We also assessed DNA purity using the A260/280 ratio. A ratio of 1.8, which is generally accepted as “pure” for DNA, was observed for S-DQ (Fig. 2.c). A ratio below 1.8 was observed for the protocols MN, S-MN, Z, S-Z and DQ, which may indicate the presence of protein, phenol or other contaminants. A ratio close to 2 was assessed for QQ and S-QQ, suggesting the possible presence of RNA in samples (p-value < 0.01 in comparison with the other protocols). Except for MN, the protocols combined with SPD generated DNA of purity equal or superior to their standard versions.

# Observed microbial diversity and performance in extracting Gram-positive bacteria

In addition to the wet-lab criteria, the extraction quality was also evaluated using 16S rRNA gene amplicon and SMS data, by investigating the observed microbial diversity of samples (Fig. 3). This *alpha*-diversity has been recently described as a good indicator of DNA extraction performance, being positively correlated with the Gram-positive bacteria extraction [35].

We observed a lower microbial diversity for CDI + patients compared to healthy volunteers (**Data not shown**), confirming the results of previous studies [74]. As a considerable variability was found within each group of individuals, we corrected the individual effect in the statistical model to emphasize differences between extraction protocols. Interestingly, the *alpha*-diversity was equal or the highest when the samples were extracted with a SPD-associated protocol, independently of the sequencing method used. Regarding the 16S rRNA gene data, the median *alpha*-diversity values were above 4.5 for S-QQ, S-MN, S-Z and S-DQ, and below 4.4 for QQ, MN, Z and DQ (at least p-value < 0.05, Fig. 3.a). For SMS data, SPD seemed to improve the *alpha*-diversity values for MN, Z and DQ but with limited effect on the QQ protocol (Fig. 3.b).

We then evaluated if the observed diversity was associated with an effective Gram-positive bacteria recovery. For this purpose, we assessed the ratio *Firmicutes/Bacteroidetes*, two main phyla, commonly found in the gut microbiota. *Firmicutes* and *Bacteroidetes* are phyla of bacteria, which are, for the most part Gram-positive and Gram-negative, respectively. In theory, the ratio *Firmicutes/Bacteroidetes* should be improved by a protocol performing well for the extraction of Gram-positive bacteria [75]. Remarkably, this ratio was increased for the four protocols combined with SPD in comparison to their standard versions, in both 16S and SMS data (Fig. 4). To quantify more precisely the SPD effect on microbial community composition, DESeq2 was used to test the differential abundance of taxa between standard vs SPD-combined protocols. For each patient, the relative abundance of the *Firmicutes* phylum increased significantly, whereas the *Bacteroidetes* phylum decreased significantly with the use of SPD. This analysis was also performed at the family level, where SPD led to a significant decrease of Gram-negative families and a significant increase of Gram-positive families (**Supplementary Table 1**). Altogether, our results were consistent with a positive effect of SPD on the observed *alpha*-diversity, by improving the recovery of Gram-positive bacteria.

## Extraction protocol accuracy

In order to estimate the accuracy of the extraction protocols, a mock community consisting of nine bacterial species of known respective abundances was prepared and sequenced. The protocol accuracy was estimated by calculating the Euclidean distance (the lower the distance, the better the prediction) between observed and expected abundances at the genus level (Fig. 5). Interestingly, the bacterial abundances were better predicted using SMS than using 16S rRNA gene sequencing. Independently of the metagenomics methods, these predictions were even better when SPD was used upstream for the

protocols QQ, MN and Z. The same effect was observed with DQ combined with SPD but only from SMS data. Based on 16S rRNA gene data, DQ was the most accurate protocol, followed by S-MN, S-QQ and S-DQ. Regarding SMS data, S-DQ was the best performing protocol, followed by S-QQ, S-MN and S-Z. In our hands, QQ was the less accurate protocol in all the conditions tested. Detailed bacterial abundances at the genus level are plotted in **Supplementary Fig. 1**.

## Protocol repeatability

Using 16S rRNA gene sequencing data, the eight protocols were next evaluated for repeatability across the variations of bacterial abundances between triplicates of a same stool sample (Fig. 6). The calculated coefficient of variation was the highest using the four standard protocols (QQ, MN, Z and DQ). We observed a significant increase ( $p$  value  $< 0.01$ ) of the repeatability when the protocols were coupled with SPD compared to their standard versions. The median of the coefficient of variation was divided by 1.57 between QQ (13.2) and S-QQ (8.4), 1.35 between Z (11.7) and S-Z (8.7), 1.21 between MN (10.9) and S-MN (9) and 1.24 between DQ (12.6) and S-DQ (10.2). S-QQ was the most repeatable protocol, followed by S-Z, S-MN and S-DQ.

## Protocols overall performance

In our study, eight DNA extraction protocols were evaluated using both wet- and dry-lab criteria, with two sequencing read-outs. Taking altogether, no single protocol performed the best for all tested criteria. To help in data interpretation, we ranked the protocols according to a score which was assigned to each criterion based on the observed results (Fig. 7). For a given criterion, the scores ranged from 0 (the worst result obtained in our dataset) to 10 (the best result obtained in our dataset). These scores were then plotted using a spider chart: a score of 0 represents the center, whereas a score of 10 is the vertex. The generated areas were then used to help in selecting the best-overall performing DNA extraction protocol.

Overall, the protocols combined with SPD performed better compared to their standard version (Fig. 7). In our hands, S-DQ showed the best overall performance (Fig. 7.a). This modified protocol performed well for the quality of the extracted DNA, the observed diversity and the accuracy. However, the S-DQ performance was slightly inferior to other protocols regarding DNA yield but this difference was only significant compared to S-Z ( $p$ -value  $< 0.1$ , Fig. 2.a), which performed poorly on the other criteria. Even if S-DQ was not the best protocol for this criterion, enough DNA material was produced to prepare and sequence the metagenomics libraries. S-DQ was also found to be less repeatable than S-Z, S-MN and S-QQ but the slight difference was only significant compared to S-QQ (Fig. 6), which performed poorly on other important criteria, including accuracy and diversity.

Considering the standard versions of the protocols, DQ had the best overall performance (Fig. 7.b). This protocol performed well in terms of accuracy from both metagenomics sequencing methods. QQ generated the highest diversity in SMS results, but the difference with DQ was not significant ( $p$ -value  $> 0.1$ , Fig. 3.b). DQ generated lower diversity than MN in 16S rRNA gene sequencing results, but this

difference was also not significant (p-value > 0.1, Fig. 3.a). Finally, MN and Z were slightly more repeatable than DQ, but not significantly (p-value > 0.1, Fig. 6).

## Discussion

DNA extraction is a crucial step of the metagenomics workflow, known to be influenced by many parameters, which are difficult to evaluate exhaustively. In addition to in-house protocols, new commercial solutions are now emerging, making difficult the choice of a good protocol for the gut microbiota. Benchmarking protocols is thus crucial to understand the potential biases and to avoid errors during data interpretation. Recent gut microbiome studies compared various DNA extraction protocols but were limited to a low number of fecal samples, mainly from healthy individuals [27–30]. As a consequence, the performance of such protocols may not be guaranteed for a clinical cohort.

Our study is the first, to our knowledge, to compare four commercial DNA extraction protocols using both metagenomics sequencing methods on an adequate number of stool samples for statistical analysis and biological conclusion (n = 18). In an effort to streamline fecal preparation prior to DNA extraction, the commercial protocols were also tested in combination with a stool preprocessing device. As recommended by recent studies, we also included a positive control, the mock community, so that we could reliably assess the accuracy of extraction protocols. The mock was made up of nine bacterial species and processed alongside fecal specimens. The eight protocols tested were ranked based on wet- and dry-lab criteria. The global aim was to identify one method that perform well and generate the most accurate and reproducible data.

In addition to healthy donors, patients suffering from a *Clostridium Difficile* Infection (CDI) were also recruited, allowing to test the protocols on samples with various microbial composition, consistency and biomass. CDI is a burning issue, as *Clostridium difficile*, a gram-positive bacterium, is the leading cause for diseases from mild diarrhea to pseudomembranous colitis in hospitalized patients [76]. Fecal microbiota transplant (FMT) is emerging as a new option for recurrent CDI [77]. Identifying which bacteria are already present (recipient) and have been transferred (donor) is essential and requires the use of highly sensitive, robust and fast metagenomics techniques [78, 79].

In our study, a total of 456 and 56 samples were analyzed using 16S rRNA gene sequencing and SMS, respectively, allowing to have unprecedented comparison results. Even if, as expected, SMS is more sensitive and accurate in bacterial detection, our present findings indicate good agreement between the two sequencing methods but also between the samples from the two groups of individuals. Interestingly, our results show that no single DNA extraction protocol performs well on all the criteria tested, which complicates the identification of the best performing protocol. Considering the strategy of selection described above, we consider S-DQ as being the best-performing protocol overall for extracting human fecal samples. This protocol generated an amount of good quality DNA that was compatible with subsequent library preparations for all samples. Regarding the dry-lab criteria, S-DQ combined the best

results in terms of alpha-diversity, extraction of gram-positive bacteria, repeatability and accuracy in bacterial detection.

Remarkably, the bioinformatics analysis also shed light on the added value of the stool preprocessing device for the all the extraction protocols. In our study, the protocols in combination with SPD have in common the first steps of the procedure. This includes the shaking and the mechanical lysis with zirconia and silica beads 0.1 mm. In such combination, we observe an increase of the observed alpha-diversity. Our results are in good agreement with Costea et al. who showed that these parameters of the protocol were positively associated with the observed diversity, which is a good indicator of an efficient lysis [35]. Biased protocols are also known to cause overrepresentation of gram-negative bacteria due to the inefficient lysis of gram-positive bacteria. For the SPD-combined protocols, we observed an increase of the relative abundance of Gram-positive bacteria and a corresponding decrease in the relative abundance of Gram-negative bacteria, which led to an increase of the *Firmicutes/Bacteroidetes* ratio. The SPD can therefore provide more accurate characterization of the microbiota by reducing the ratio bias. In terms of repeatability, SPD also showed promising results. This device and similar approaches would be of particular interest to limit variations when several experimenters, and even different labs in case of multi-centric studies, perform DNA extraction. Lastly, the use of our in-house mock community, composed of both gram-positive and gram-negative bacteria cells, made possible to benchmark the protocols in terms of bacterial abundance predictions. Our results demonstrate that SPD in combination with any of the four protocols is more accurate in assessing the bacterial abundances than the protocols in their standard version.

We are also aware that all the protocols may not have been tested in optimal parameters. The commercial protocols were tested using the beads provided in the kit on a Retsch system for 5 minutes. In our hands, protocol Z was the worst performer according to wet-lab criteria. Today, Zymo Research recommends other bead-beating protocols than the one tested, which could have improved its performance [80].

Such device prior DNA extraction may add additional costs to the DNA extraction reactions but from our perspective, getting unbiased microbiome data is priceless. In our data set, we have also shown that the DQ protocol is the best protocol among the commercial solutions.

## Conclusion

We have shown that both metagenomics methods are in good agreement when comparing DNA extraction protocols.

We recommend the S-DQ protocol to extract microbial DNA from human stool samples. While we have only tested S-DQ on fecal samples, we suppose that it might also work well with other types of microbiota samples, although some modifications may be necessary.

SPD appears to be a new way to improve the overall performance of any DNA extraction protocol. We propose to now include stool preprocessing devices in new microbiome studies to streamline and standardize DNA extraction.

## Abbreviations

- SMS

Shotgun Metagenomic Sequencing

- SPD

Stool preprocessing Device

- T2D

Type 2 Diabetes

- NAFLD

nonalcoholic fatty liver disease

- Q

Protocol Qiagen from [35]

- CDI

Clostridium difficile infection

- PCR

Polymerase Chain Reaction

- NPM

Nextera PCR Master Mix

- OTU

Operational Taxonomic Units

- NT

Neutralize Tagment

- MN

The NucleoSpin Soil kit (Macherey-Nagel)

- DQ

The DNeasy PowerLyzer PowerSoil kit (Qiagen)

- QQ

The QIAamp Fast DNA Stool kit (Qiagen)

- Z

The ZymoBIOMICS DNA Mini kit (ZymoResearch)

- S-MN

SPD in combination with MN

- S-DQ

SPD in combination with DQ

- S-QQ

SPD in combination with QQ

- S-Z

SPD in combination with Z.

## Declarations

### Ethics approval and consent to participate:

Patients were initially recruited for diagnostic purpose and were informed regarding collection, storage and possible use for research activities. Fecal samples were retrieved for our study according to the French legal and medical ethical guidelines. These samples were declared by BIOASTER to the French Ministry of Higher Education, Research and Innovation (N°DC-2018-3240).

### Consent for publication:

Not applicable

### Availability of data and materials:

The datasets generated during the current study are available on the BioProject database (ID PRJNA648321), at the following link:

<http://www.ncbi.nlm.nih.gov/bioproject/648321>

### Competing interests:

The authors declare that they have no competing interests

### Funding:

This research project has received funding from the French Government through the Investissement d'Avenir program (grant n°ANR-10-AIRT-03) and from bioMérieux.

### Authors' contributions:

A.S. and G.G. designed the project. M.P. and A.S. performed the experimental work. C.E., M.P., A.F.-L., P.L., C.V., H.R., F.R., G.G. and A.S. analyzed results and wrote the manuscript. K.L. was in charge of the recruitment of clinical samples.

# Acknowledgements

The authors thank Cécile Chauvel (xDATA Team, BIOASTER) for the statistical analyses and Adrien Villain (OMICS Hub, BIOASTER) for critical reading of the manuscript. We thank Johann Pellet and Pierre Veyre (xDATA Team, BIOASTER) for their involvement in the data and computing management and the IN2P3 Computing Center (CNRS, Lyon-Villeurbanne, France) for the provisioning and excellent performance of computing infrastructure essential to our analyses.

## References

1. Gilbert JA, et al. Current understanding of the human microbiome. *Nat Med.* 2018;24(4):392–400.
2. Chiu CY, Miller SA. Clinical metagenomics. *Nat Rev Genet.* 2019;20(6):341–55.
3. Salk JJ, Schmitt MW, Loeb LA. Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. *Nat Rev Genet.* 2018;19(5):269–85.
4. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet.* 2016;17(6):333–51.
5. Loman NJ, Pallen MJ. Twenty years of bacterial genome sequencing. *Nat Rev Microbiol.* 2015;13(12):787–94.
6. Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet.* 2010;11(1):31–46.
7. Sartor RB. Mechanisms of disease: pathogenesis of Crohn's disease and ulcerative colitis. *Nat Clin Pract Gastroenterol Hepatol.* 2006;3(7):390–407.
8. Halfvarson J, et al. Dynamics of the human gut microbiome in inflammatory bowel disease. *Nat Microbiol.* 2017;2:17004.
9. Vieira-Silva S, et al., *Quantitative microbiome profiling disentangles inflammation- and bile duct obstruction-associated microbiota alterations across PSC/IBD diagnoses.* *Nat Microbiol,* 2019.
10. Manichanh C, et al. The gut microbiota in IBD. *Nat Rev Gastroenterol Hepatol.* 2012;9(10):599–608.
11. Lavelle A, Sokol H. Gut microbiota-derived metabolites as key actors in inflammatory bowel disease. *Nat Rev Gastroenterol Hepatol.* 2020;17(4):223–37.
12. Simren M, et al. Intestinal microbiota in functional bowel disorders: a Rome foundation report. *Gut.* 2013;62(1):159–76.
13. Mayer EA, et al. Towards a systems view of IBS. *Nat Rev Gastroenterol Hepatol.* 2015;12(10):592–605.
14. Musso G, Gambino R, Cassader M. Interactions between gut microbiota and host metabolism predisposing to obesity and diabetes. *Annu Rev Med.* 2011;62:361–80.
15. Larsen N, et al. Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. *PLoS One.* 2010;5(2):e9085.
16. Qin J, et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature.* 2012;490(7418):55–60.

17. Aron-Wisnewsky J, et al., *Gut microbiota and human NAFLD: disentangling microbial signatures from metabolic disorders*. Nat Rev Gastroenterol Hepatol, 2020.
18. Canfora EE, et al. Gut microbial metabolites in obesity, NAFLD and T2DM. Nat Rev Endocrinol. 2019;15(5):261–73.
19. Caussy C, et al. A gut microbiome signature for cirrhosis due to nonalcoholic fatty liver disease. Nat Commun. 2019;10(1):1406.
20. Helmkink BA, et al. The microbiome, cancer, and cancer therapy. Nat Med. 2019;25(3):377–88.
21. Yachida S, et al. Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. Nat Med. 2019;25(6):968–76.
22. Routy B, et al. Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. Science. 2018;359(6371):91–7.
23. Zitvogel L, et al. Anticancer effects of the microbiome and its products. Nat Rev Microbiol. 2017;15(8):465–78.
24. Routy B, et al. The gut microbiota influences anticancer immunosurveillance and general health. Nat Rev Clin Oncol. 2018;15(6):382–96.
25. Fulbright LE, Ellermann M, Arthur JC. The microbiome and the hallmarks of cancer. PLoS Pathog. 2017;13(9):e1006480.
26. Gopalakrishnan V, et al. Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. Science. 2018;359(6371):97–103.
27. Thomas RM, Jobin C. Microbiota in pancreatic health and disease: the next frontier in microbiome research. Nat Rev Gastroenterol Hepatol. 2020;17(1):53–64.
28. Hofseth LJ, et al., *Early-onset colorectal cancer: initial clues and current views*. Nat Rev Gastroenterol Hepatol, 2020.
29. Wirbel J, et al. Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. Nat Med. 2019;25(4):679–89.
30. Thomas AM, et al. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. Nat Med. 2019;25(4):667–78.
31. Lim MY, et al. Changes in microbiome and metabolomic profiles of fecal samples stored with stabilizing solution at room temperature: a pilot study. Sci Rep. 2020;10(1):1789.
32. Tap J, et al. Effects of the long-term storage of human fecal microbiota samples collected in RNAlater. Sci Rep. 2019;9(1):601.
33. Moossavi S, et al. Assessment of the impact of different fecal storage protocols on the microbiota diversity and composition: a pilot study. BMC Microbiol. 2019;19(1):145.
34. Martinez N, et al. Filling the gap between collection, transport and storage of the human gut microbiota. Sci Rep. 2019;9(1):8327.
35. Costea PI, et al. Towards standards for human fecal sample processing in metagenomic studies. Nat Biotechnol. 2017;35(11):1069–76.

36. Knudsen BE, et al., *Impact of Sample Type and DNA Isolation Procedure on Genomic Inference of Microbiome Composition*. mSystems, 2016. 1(5).
37. Sze MA, Schloss PD. *The Impact of DNA Polymerase and Number of Rounds of Amplification in PCR on 16S rRNA Gene Sequence Data*. mSphere, 2019. 4(3).
38. Whon TW, et al. The effects of sequencing platforms on phylogenetic resolution in 16 S rRNA gene profiling of human feces. *Sci Data*. 2018;5:180068.
39. Breitwieser FP, Lu J, Salzberg SL. A review of methods and databases for metagenomic classification and assembly. *Brief Bioinform*. 2019;20(4):1125–36.
40. Sczyrba A, et al. Critical Assessment of Metagenome Interpretation-a benchmark of metagenomics software. *Nat Methods*. 2017;14(11):1063–71.
41. Kuczynski J, et al. Experimental and analytical tools for studying the human microbiome. *Nat Rev Genet*. 2011;13(1):47–58.
42. Knight R, et al. Best practices for analysing microbiomes. *Nat Rev Microbiol*. 2018;16(7):410–22.
43. Sinha R, et al. Assessment of variation in microbial community amplicon sequencing by the Microbiome Quality Control (MBQC) project consortium. *Nat Biotechnol*. 2017;35(11):1077–86.
44. Gohl DM. The ecological landscape of microbiome science. *Nat Biotechnol*. 2017;35(11):1047–9.
45. Vandeputte D, et al. Practical considerations for large-scale gut microbiome studies. *FEMS Microbiol Rev*. 2017;41(Supp\_1):S154–67.
46. Fouhy F, et al. 16S rRNA gene sequencing of mock microbial populations- impact of DNA extraction method, primer choice and sequencing platform. *BMC Microbiol*. 2016;16(1):123.
47. Albertsen M, et al. Back to Basics– The Influence of DNA Extraction and Primer Choice on Phylogenetic Analysis of Activated Sludge Communities. *PLoS One*. 2015;10(7):e0132783.
48. Wesolowska-Andersen A, et al. Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis. *Microbiome*. 2014;2:19.
49. Hsieh YH, et al. Impact of Different Fecal Processing Methods on Assessments of Bacterial Diversity in the Human Intestine. *Front Microbiol*. 2016;7:1643.
50. Gorzelak MA, et al. Methods for Improving Human Gut Microbiome Data by Reducing Variability through Sample Processing and Storage of Stool. *PLoS One*. 2015;10(8):e0134802.
51. Feghoul L, et al. Evaluation of a New Device for Simplifying and Standardizing Stool Sample Preparation for Viral Molecular Testing with Limited Hands-On Time. *J Clin Microbiol*. 2016;54(4):928–33.
52. Panek M, et al. Methodology challenges in studying human gut microbiota - effects of collection, storage, DNA extraction and next generation sequencing technologies. *Sci Rep*. 2018;8(1):5143.
53. Kennedy NA, et al. The impact of different DNA extraction kits and laboratories upon the assessment of human gut microbiota composition by 16S rRNA gene sequencing. *PLoS One*. 2014;9(2):e88982.
54. Maukonen J, Simoes C, Saarela M. The currently used commercial DNA-extraction methods give different results of clostridial and actinobacterial populations derived from human fecal samples.

- FEMS Microbiol Ecol. 2012;79(3):697–708.
55. Truong DT, et al. Microbial strain-level population structure and genetic diversity from metagenomes. *Genome Res.* 2017;27(4):626–38.
  56. Vandeputte D, et al. Quantitative microbiome profiling links gut community variation to microbial load. *Nature.* 2017;551(7681):507–11.
  57. Stammler F, et al. Adjusting microbiome profiles for differences in microbial load by spike-in bacteria. *Microbiome.* 2016;4(1):28.
  58. Shreiner AB, Kao JY, Young VB. The gut microbiome in health and in disease. *Curr Opin Gastroenterol.* 2015;31(1):69–75.
  59. Falony G, et al. The human microbiome in health and disease: hype or hope. *Acta Clin Belg.* 2019;74(2):53–64.
  60. Cani PD. Human gut microbiome: hopes, threats and promises. *Gut.* 2018;67(9):1716–25.
  61. Lewis SJ, Heaton KW. Stool form scale as a useful guide to intestinal transit time. *Scand J Gastroenterol.* 1997;32(9):920–4.
  62. Vandeputte D, et al. Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. *Gut.* 2016;65(1):57–62.
  63. Falony G, et al. Population-level analysis of gut microbiome variation. *Science.* 2016;352(6285):560–4.
  64. Quince C, et al. Shotgun metagenomics, from sampling to analysis. *Nat Biotechnol.* 2017;35(9):833–44.
  65. Almeida A, et al. A new genomic blueprint of the human gut microbiota. *Nature.* 2019;568(7753):499–504.
  66. Gohl DM, et al. Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. *Nat Biotechnol.* 2016;34(9):942–9.
  67. Fraher MH, O'Toole PW, Quigley EM. Techniques used to characterize the gut microbiota: a guide for the clinician. *Nat Rev Gastroenterol Hepatol.* 2012;9(6):312–22.
  68. Shaw AG, et al. Latitude in sample handling and storage for infant faecal microbiota studies: the elephant in the room? *Microbiome.* 2016;4(1):40.
  69. Illumina. *16S Metagenomic Sequencing Library Preparation.*  
[https://support.illumina.com/documents/documentation/chemistry\\_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf](https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf).
  70. Tecan. *Automated library preparation for Illumina® 16S metagenomic sequencing.*  
[https://lifesciences.tecan.com/applications\\_and\\_solutions/genomics/ngs\\_sample\\_preparation?p=Literature](https://lifesciences.tecan.com/applications_and_solutions/genomics/ngs_sample_preparation?p=Literature).
  71. Bokulich NA, et al. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat Methods.* 2013;10(1):57–9.

72. Kim D, et al. Centrifuge: rapid and sensitive classification of metagenomic sequences. *Genome Res.* 2016;26(12):1721–9.
73. Haft DH, et al. RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res.* 2018;46(D1):D851–60.
74. Han SH, et al. Composition of gut microbiota in patients with toxigenic *Clostridioides (Clostridium) difficile*: Comparison between subgroups according to clinical criteria and toxin gene load. *PLoS One.* 2019;14(2):e0212626.
75. Santiago A, et al. Processing faecal samples: a step forward for standards in microbial community analysis. *BMC Microbiol.* 2014;14:112.
76. Abt MC, McKenney PT, Pamer EG. *Clostridium difficile* colitis: pathogenesis and host defence. *Nat Rev Microbiol.* 2016;14(10):609–20.
77. Kociolek LK, Gerding DN. Breakthroughs in the treatment and prevention of *Clostridium difficile* infection. *Nat Rev Gastroenterol Hepatol.* 2016;13(3):150–60.
78. Weingarden A, et al. Dynamic changes in short- and long-term bacterial composition following fecal microbiota transplantation for recurrent *Clostridium difficile* infection. *Microbiome.* 2015;3:10.
79. Staley C, et al., *Complete Microbiota Engraftment Is Not Essential for Recovery from Recurrent Clostridium difficile Infection following Fecal Microbiota Transplantation.* *mBio,* 2016. 7(6).
80. Research Z. *The Lysis Bias Crisis.* <https://www.zymoresearch.com/blogs/blog/the-lysis-bias-crisis>.

## Tables

Due to technical limitations, the tables are provided in the Supplementary Files section.

## Figures

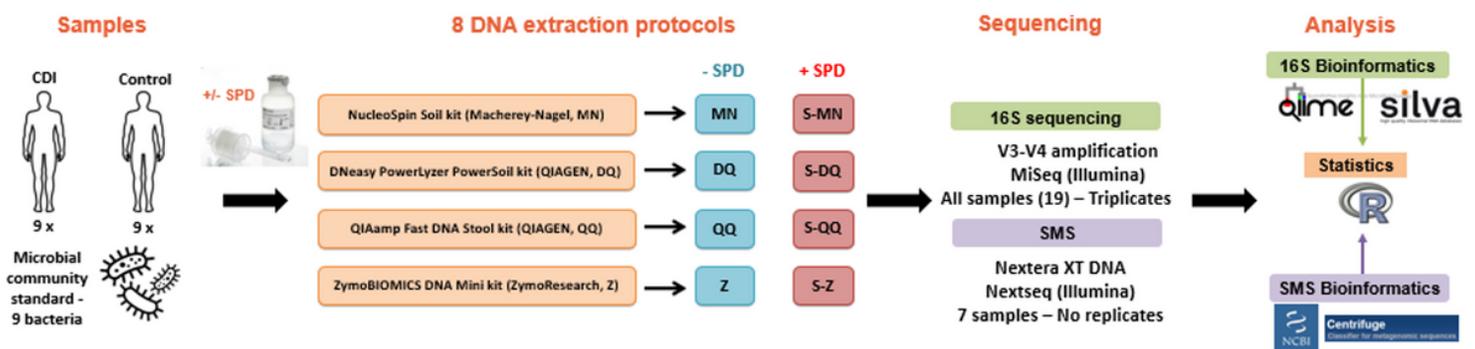
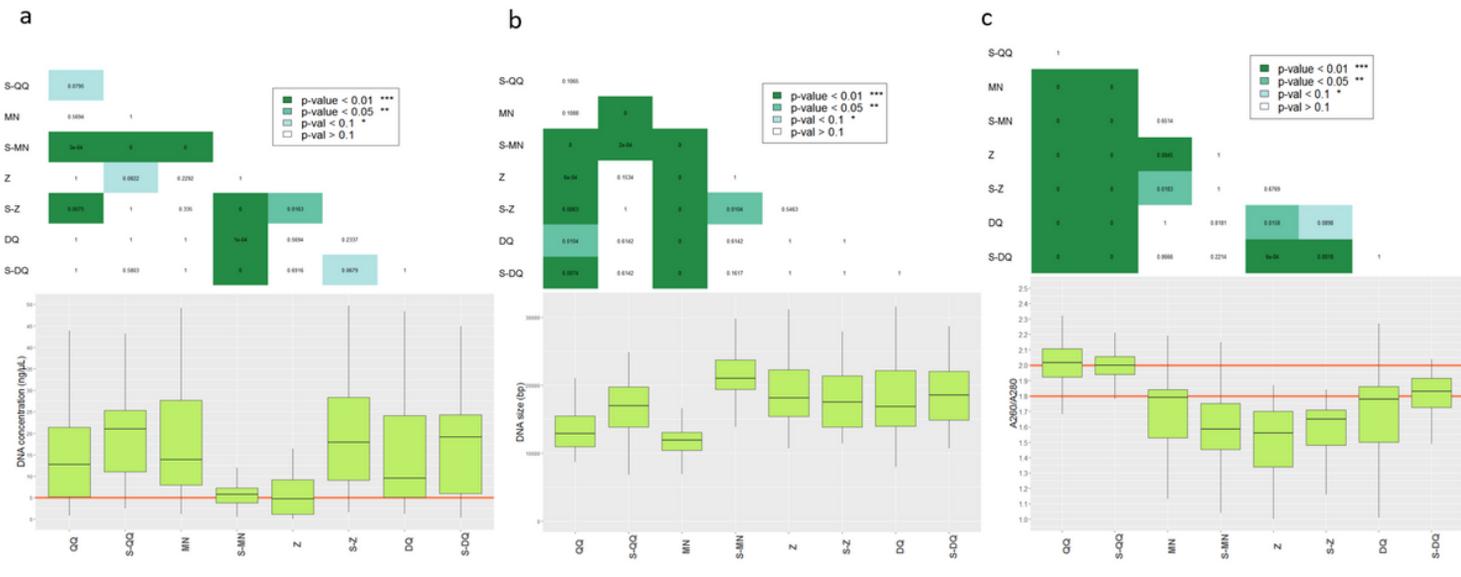


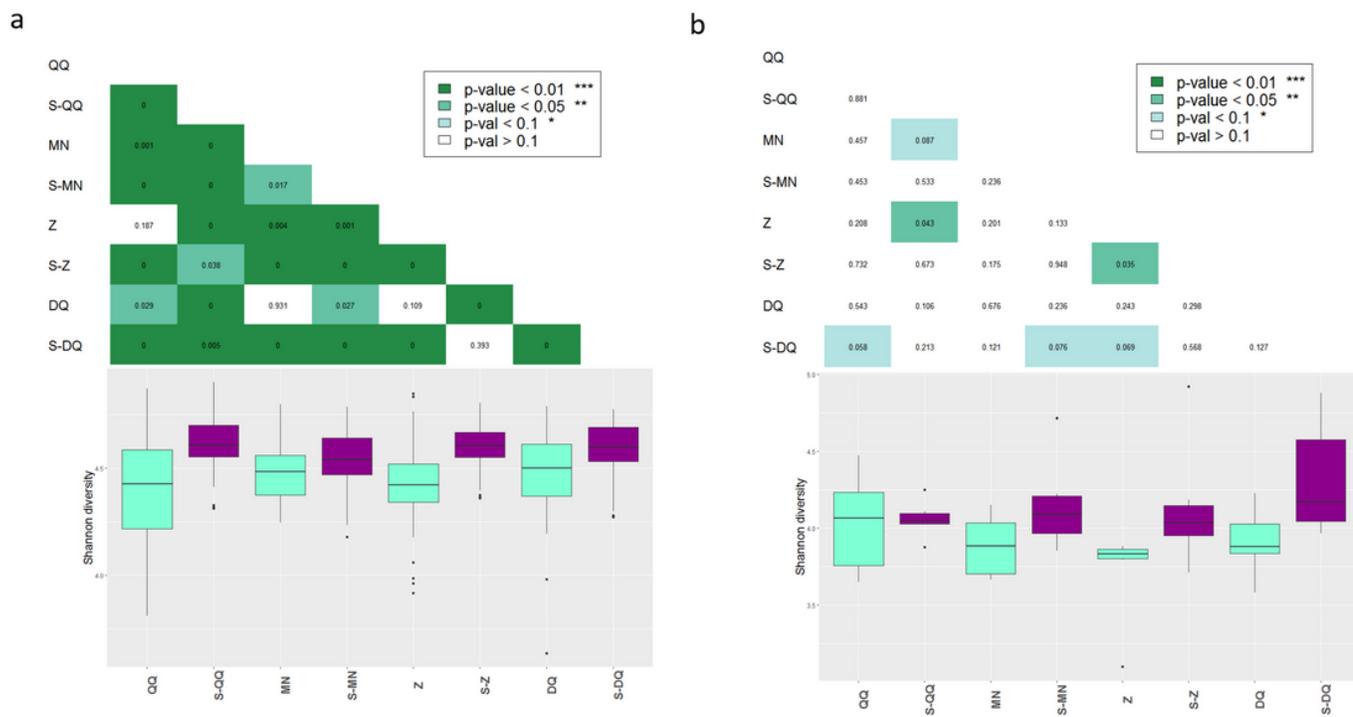
Figure 1

Workflow of sample processing



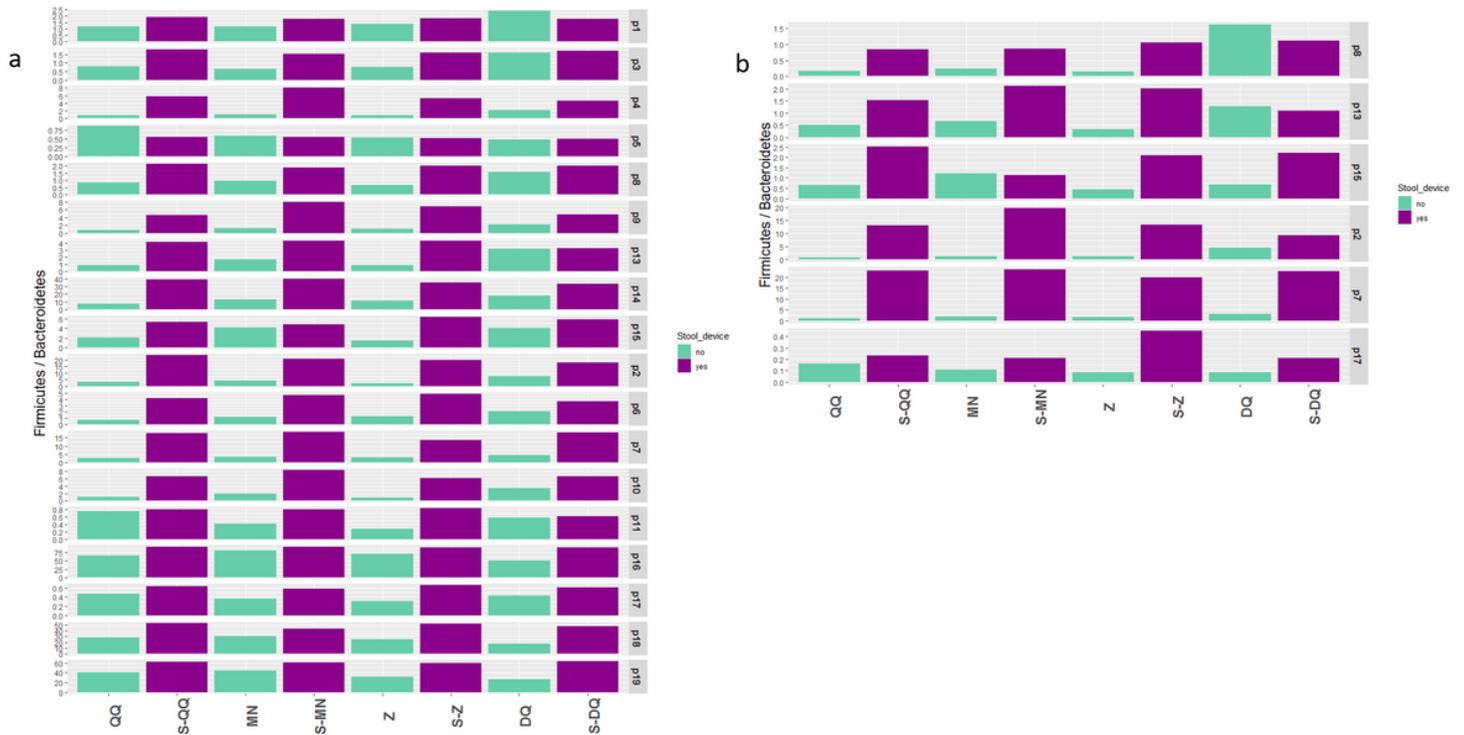
**Figure 2**

Quantity and quality of extracted DNA from 8 different extraction protocols. (a) DNA concentration (ng/μL). The orange horizontal line represents the threshold at 5 ng/μL. (b) DNA size (bp). (c) Absorbance ratios at 260/280. DNA of good quality should have a ratio between the two orange horizontal lines (between 1.8 and 2, preferably close to 1.8). All boxplots are topped by a heatmap showing the pairwise Wilcoxon test p-values. Significant differences between protocols are highlighted in green.



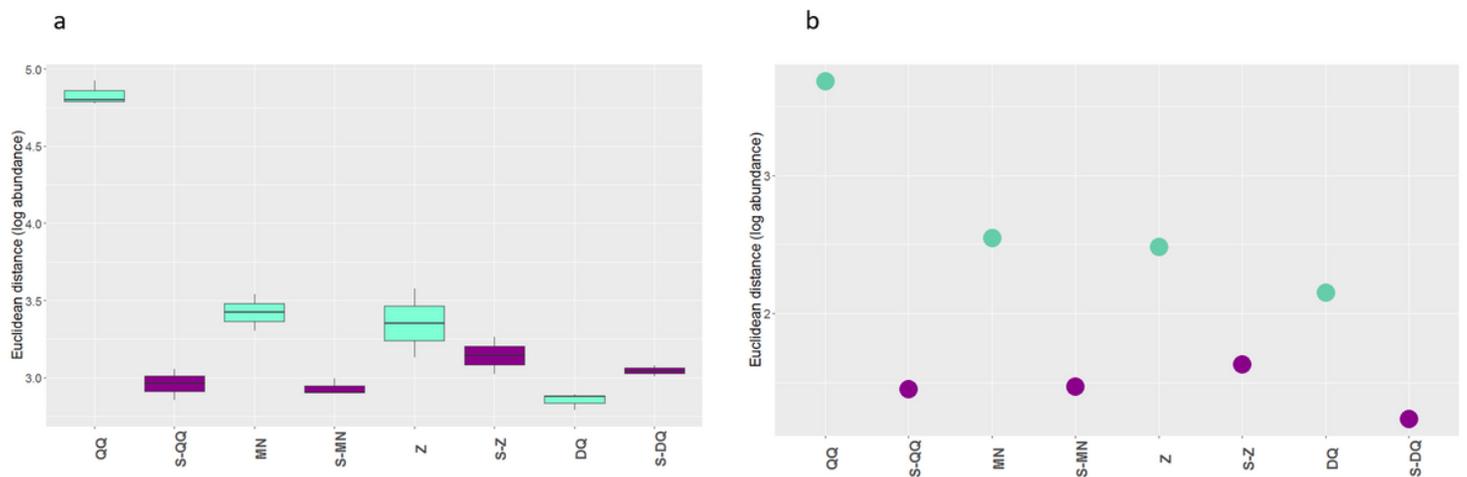
**Figure 3**

Shannon  $\alpha$ -diversity index of sample composition. (a) Shannon diversity calculated from the abundance table of 16S rRNA gene profiling. (b) Shannon diversity calculated from the abundance table of SMS data at species level. Standard protocols are colored in sky blue, whereas protocols associated with SPD are colored in purple. Boxplots are topped by a heatmap showing the pairwise Wilcoxon test p-values. Significant differences between protocols are highlighted in green.



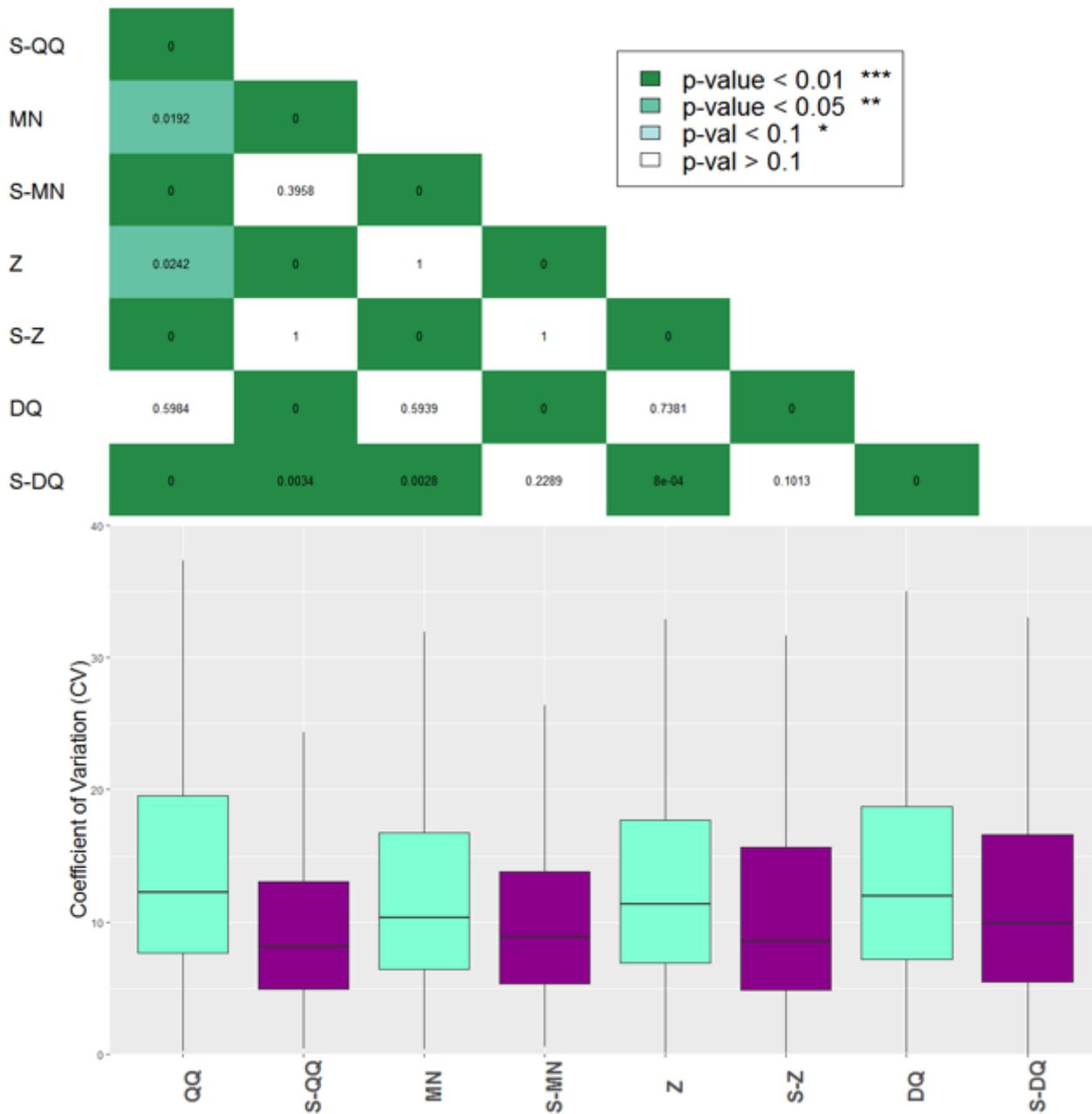
**Figure 4**

Estimation of the Firmicutes/Bacteroidetes ratio. This ratio was calculated for every individual extracted by the 8 different extraction protocols. (a) 16S rRNA gene sequencing data (mean of triplicates). (b) SMS data.



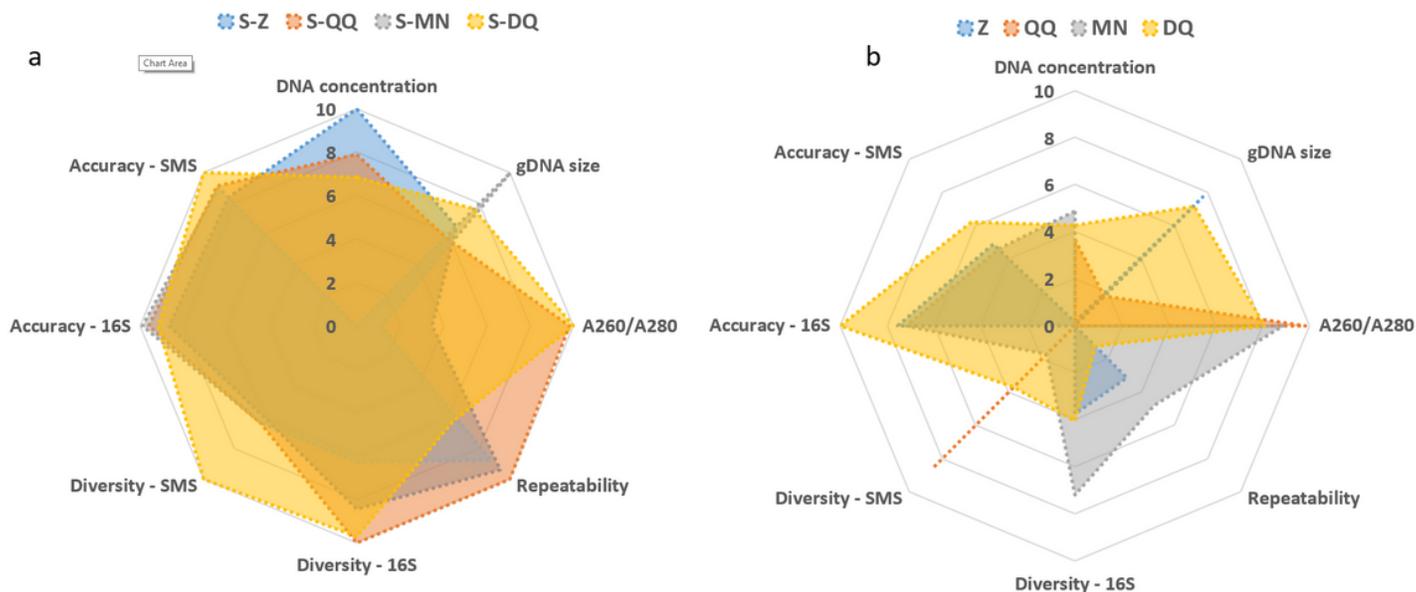
**Figure 5**

Accuracy of the observed bacterial abundance. Euclidean distances were calculated between observed and expected bacterial abundances (log values). (a) Euclidean distances calculated for each of three replicates of each extraction protocol, for 16S rRNA gene sequencing data. (b) Euclidean distances calculated for one replicate of each extraction protocol, for SMS data.



**Figure 6**

Protocol repeatability. A coefficient of variation (SD/mean) was calculated between replicates for each genus within each sample. Standard protocols are colored in sky blue, whereas protocols with SPD are colored in purple. The boxplot is topped by a heatmap showing the pairwise Wilcoxon test p-values. Significant differences between protocols are highlighted in green.



**Figure 7**

Overall performance of the 8 extraction protocols based on all studied criteria. (a) Protocols in combination with SPD. (b) Standard protocols (no SPD). A scale from 0 to 10 was used to compare protocols. Scores of 0 correspond to the weakest values, whereas scores of 10 are the best values of the considered criteria.

## Supplementary Files

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

- [Table1.xlsx](#)
- [Table1.xlsx](#)
- [Table2.xlsx](#)
- [Table2.xlsx](#)
- [SupplementaryTable1.xlsx](#)
- [SupplementaryTable1.xlsx](#)
- [Supplementarymethods.docx](#)
- [Supplementarymethods.docx](#)
- [Supplementaryfigures.docx](#)
- [Supplementaryfigures.docx](#)