Fecal Sample Collection Methods and Time of Day Impact Microbiome Composition and Short Chain Fatty Acid Concentrations

Jacquelyn Jones (jacquelyn.jones@postgrad.curtin.edu.au)
Curtin University
Stacey Reinke
Edith Cowan University
Alishum Ali
Curtin University
Debra Palmer
Telethon Kids Institute
Claus T. Christophersen
Curtin University

Research Article

Keywords: human gut microbiome, OMNIgeneGUT, amplicons

DOI: https://doi.org/10.21203/rs.3.rs-199162/v1

Abstract

Associations between the human gut microbiome and health outcomes continues to be of great interest, although fecal sample collection methods which impact microbiome studies are sometimes neglected. Here, we expand on previous work in sample optimization, to promote high quality microbiome data. To compare fecal sample collection methods, amplicons from the bacterial 16S rRNA gene (V4) and fungal (ITS2) region, as well as short chain fatty acid (SCFA) concentrations were determined in fecal material over three timepoints. We demonstrated that spot sampling of stool results in variable detection of some microbial members, and inconsistent levels of SCFA; therefore, sample homogenization prior to subsequent analysis or subsampling is recommended. We also identify a trend in microbial and metabolite composition that shifts over two consecutive stool collections less than 25h apart. Lastly, we show significant differences in bacterial composition that result from collecting stool samples in OMNIgeneGUT tube (DNA Genotec) or Stool Nucleic Acid Collection and Preservation Tube (NORGEN) compared to immediate freezing. To assist microbiome investigators plan their fecal sample collection and storage procedures for multiple analyses, we recommend participants to collect the first full bowel movement of the day and freeze the sample immediately after collection.

Introduction

Our understanding of the relationship between the human gut microbiome and host continues to expand from explorations which describe inhabitants, to studies which demonstrate the involvement of the microbiome in human health and disease and disorders. Some examples include neurological disorders such as depression [1], Alzheimer's disease [2] and Autism Spectrum Disorder [3], as well as inflammatory diseases such as food allergies [4], and inflammatory bowel diseases [5]. Advancements in microbiome studies have been accelerated by increased sequencing capabilities [6], along with sensitive analytical techniques tailored for the quantification of metabolites in fecal material [7,8]. Short chain fatty acids (SCFA) are metabolites produced exclusively by resident bacteria, and are important for proper gut barrier functioning. Therefore, SCFA are also associated with dysbiosis, and other inflammatory disorders [5]; and investigating the gut microbiome by combining microbial sequencing data and metabolomic approaches has been an important step in unraveling links between resident bacteria, SCFA, and health outcomes (Arrieta et al., 2018; De Filippis et al., 2016; Lagkouvardos et al., 2015).

However, stool, which is used as proxy for the distal colon microbiome, is a complex matrix of endo- and exogenous material containing a heterogeneous distribution of microorganisms [13], which is susceptible to changes during and after collection. Microbiome profiles may be misrepresented due to subsampling of non-homogenized stool as seen in Gorzelak (2015) where large variations in bacterial abundance detected via qPCR in non-homogenized stool samples were significantly reduced after stool homogenization. In addition, the effects of sub-sampling stool may be further amplified when performing metabolomic analyses, as highly sensitive techniques are used [8,14].

Stool collection by participants may be an undesirable yet necessary aspect of partaking in a microbiome study. Providing participants with a clean and simple collection method should increase compliance, but also maintain sample integrity. Some commercial stool collection tubes allow for easy collection and short term (~14 days) ambient temperature storage; however, some of these have been associated with changes in proportions of bacterial phyla [15]. A final consideration is the level of inter-individual differences that occur in the fecal microbiome over a week [16], and even from day to day [17], meaning that collection periods may need to span a number of days, or be collected at a particular time in the day to accurately capture the inherent variability. As far as the authors are aware spatial and short-term temporal variability of bacterial and fungal communities has never been evaluated together with SCFA composition. To address this gap, this study will assess the effects of five fecal sample collection methods, as well as consecutively collected whole stool samples (less than 25 hours apart), on the variability of the fecal microbiome. The comparisons will be drawn from bacterial and fungal community composition as well as SCFA profiling.

Methods

Study Design

Six healthy female volunteers, aged 25 – 40 years, who had not taken antibiotics in the last 3 months or probiotics in the last month prior to recruitment into this study provided fecal samples with written informed consent. The study protocol was approved by the Human Research Ethics Committee (HRE2018-0791) from Curtin University, Western Australia, and methods were performed in accordance with the relevant guidelines and regulations. Each participant collected 3 fecal samples over 2 or 3 consecutive days using the provided fecal sample collection kit. All stools were collected at the participants home and frozen immediately (-20°C) in a portable freezer. Collection of the first stool required collecting one complete bowel movement, and from this stool collecting 3 small aliquots in the provided collection tubes. For the second and third collection one complete bowel movement was collected individually (Figure 7). Once the collection was complete, the freezer was transported to Curtin University and the stool was transferred to a -80 °C freezer upon arrival.

Sample preparation

All stool samples were thawed at 4°C, and transferred on ice to a EuroClone Biological safety cabinet to limit potential contamination. To assess variability between aliquots collected at home, each of the small aliquots were individually homogenized for 30 seconds with a sterile plastic scoop, and stool (0.25 g) was collected into separate tubes for each of two downstream analyses (metabarcoding, and SCFA quantification). The remaining stool from the initial three aliquots was combined and manually homogenized together for 30 seconds with a sterile plastic scoop, and collected again for two separate analyses. All samples were immediately frozen to -80°C. Prior to preparation, whole stool samples were ranked on the Bristol Stool form chart. To assess collection methods, from each unhomogenized stool,

feces were collected into one OMNIgeneGUT tube (DNA genotec) (collection method O) and one Stool Nucleic Acid Collection and preservation Tube (NORGEN) (collection method N), and were stored at room temperature for 12 days. The remaining stool from each sample was individually homogenized while within the plastic collection bag for 1 minute, and then subsequent aliquots of stool (0.25 g) were collected for each of the three analyses and immediately frozen to -80°C (collection method F).

Short Chain Fatty Acid Quantification

Homogenized fecal material (0.25 g) was transported on ice to the Science Analytical Facility at Edith Cowan University, Western Australia. Here SCFA were extracted, and 2-ethyl butyric acid was used as an internal standard. Extracts were then stored at 4°C for no more than 48h prior to analysis on Thermo Scientific GC-MS (ISQ) using a Thermo Scientific TG-Wax column (30mx0.25mmx0.25□m), and a seven-point calibration.

Fecal DNA Extraction

Immediately prior to DNA extraction, frozen stool samples were thawed on ice, and stool samples stored in preservation tubes were shaken by hand for 10 sec. DNA was extracted by using QIAamp Power Fecal DNA kit (QIAGEN, Hilden, Germany) using QIAamp Power Fecal DNA IRT protocol for QIAcube (QIAGEN), as well as OMNIgeneGUT microbial DNA purification protocol using QIAGEN QIAamp PowerFecal DNA kit, both according to the manufacturer's instructions with one modification at step 3 of the IRT protocol, tubes were vortexed for 20 sec to incorporate beads and stool prior to heating. Extraction controls were also processed following the same protocol as frozen stool samples.

Bacterial and Fungal Library preparation and Sequencing

The V4 region of bacterial DNA and mock communities (ATCC MSA-1002) were targeted and amplified using 16S primers 515F [37] and 806R [38], each with a 6-8 bp unique barcode. The PCR reactions contained of 1x PCR buffer (Applied Biosystems), 2mM $MgCl_2$ (Applied Biosystems), 0.25nM dNTP (Bioline), 0.4mg/mL BSA (Thermo Fisher Scientific) 0.4 mM primer (Integrated DNA Technologies), 0.12x SYBER, and 2U AmpliTaq Gold™ DNA polymerase (Thermo Fisher Scientific). Reactions contained 2 mL of template DNA which was previously screened and optimized for efficiency by qPCR [39], and had a final volume of 25 mL. Fungal DNA and mock communities [40] were amplified using ITS2 primers FSeq and RSeq [23]. PCR reactions were the same as for bacterial amplification except 3mL of template DNA was added to each reaction.

The reactions for both bacterial and fungal amplicons were performed on StepOnePlus Real-Time PCR system (Applied Biosystems), and under the following conditions for bacterial amplicons: denaturing at 94ºC for 3 min, followed by 30 cycles of 94ºC for 40 sec, annealing at 53ºC for 40 sec, and extension at

72ºC for 60 sec. The cycling program for fungi was as follows: denaturing at 94ºC for 3 min, followed by 35 cycles of 94ºC for 40 sec, 55 ºC for 40 sec, 72ºC for 80 sec. Both amplicons underwent a final extension at 72ºC for 10 minutes. Individual Bacterial and Fungal libraries were prepared by blending together in equimolar concentrations. Illumina compatible adaptors were ligated to the DNA fragments (Lucigen, Middelton, WI, USA), and the resulting amplicons were size selected using Pippin Prep (Sage Science). The QIAquick PCR purification column clean up kit (Qiagen, Germantown, MD), was used to purify the DNA library before sequencing, which was performed at Curtin University, Western Australia, using the Illumina MiSeq platform and V2 500 cycle kit (Illumina, San Diego, CA, USA) with 2 × 250 bp paired-end read lengths.

Deconvolution

Unique molecular barcodes were used to demultiplex reads with no mismatches allowed. Cutadapt [41] was used to remove primers, and the remaining reads were quality filtered, trimmed, and merged using DADA2 [42]. Reads with ambiguous bases, or with more than two expected errors were discarded. Amplicon sequence variants (ASVs) were inferred from the reads using the pseudo-pooled method, and merged with a minimum overlap of 60bp allowing for one mismatch (16S V4), and 30bp with no mismatches (ITS2). Amplicons were retained at a minimum length of 150, and 251, base pairs for ITS2, and 16S V4, datasets respectively. Chimera errors were also removed with DADA2 using the default method. Classification for 16S sequence variants was performed using the Genome Taxonomy reference database (release 95) formatted for use with DADA2 (https://zenodo.org/record/3951383#.X7Hs49sRVTY), while the UNITE general FASTA release for fungi Version 18.11.2018 [43] was used for ITS2 sequence classification, each with a minimum of 50% bootstrapping. Contamination was removed from all sequences with one run of the function remove.count in microDecon [44]. Any ASVs with unassigned phylum, or with a prevalence less than 1 in 5% of samples were filtered out, as were fungal samples with less than 1000 reads.

Statistical Analysis

Sequence counts were used to determine richness and alpha diversity indices (Chao1, and Shannon) for bacterial microbiomes as applied in the Phyloseq package in R [45]. Correlation between library size and diversity estimate were tested for, and alpha diversity measures with significant Pearson correlation ($p < 0.01$) to reads per sample were rarefied to lowest sample depth prior to calculation for those alpha diversity measures (Chao1). Beta diversity between collection methods was estimated using Euclidian distances of center log ratio (CLR) transformed data as well as Bray-Curtis similarity, and visualized using PCO. SCFA concentration data were $\log_{10}$ transformed, and total SCFA concentration was checked for normality with the Shapiro-Wilk test prior to paired t-test.

To evaluate differentially abundant taxa between homogenization method (aliquots and whole stool), the effect size estimate as a log2 fold change was calculated in DESeq2 statistical package [46] with a Benjamini-Hochburg adjustment for multiple testing, and a design controlling for subject. Statistical differences between taxa abundance and community diversity due to homogenization methods and collection methods were further tested using ANOVA (MicrobiomeAnalyst) and PERMANOVA (PRIMER7) respectively. Projection to latent structures discriminant analysis was used to test for differences in microbiome composition due to collection method; and regularize canonical correlation analysis (integrated to maximize correlation between latent variables) was used to integrate SCFA and bacteriome data, both in MixOmics [47].

Results

Overview of microbiome taxonomy and SCFA concentrations

Stool samples yielded bacterial communities (bacteriome) from all individuals and sampling methods, while fungal communities (mycobiome) were successfully sequenced in 53 of 78 samples, but with uneven library size (~100x). Overall the fecal bacteriome had a higher number of ASVs than the mycobiome (Supplementary Table 1). Across all individuals and collection methods, the most abundant bacterial families were Bacteroidaceae and Lachnospiraceae, which made up 38% and 10% of the bacteriome, while the most abundant fungal families were Saccharomycetaceae 90%, and Phaffomycetaceae 7%. To account for technical bias in library preparation, a single sample from one individual was also processed in duplicate for each gene region. Bacterial alpha diversity estimates for this replicate sample were more similar than the fungal replicate, while fungal replicates also had low richness, indicating that both the rarity of this community, and the library preparation may impact the interpretation of relative abundance of fungal communities (Supplementary Table 1). Bacterial and Fungal mock communities were also sequenced as positive controls, which allowed reads of the mock community samples to be compared to the known composition of the mock community (Supplementary Figure 1). Of the 20 bacterial species known to be in the mock community, 18 were correctly detected to family level and 15 to genus level, leaving 2 species of the mock community unidentified. Of the 19 fungal species in the mock community, 13 were sequenced and correctly resolved to family and genus (however, R. irregularis only had 11 reads), leaving 6 species of the mock community unidentified.

The mycobiome signature of each subject was not as distinct as the bacteriome (Figure 1). While fungal communities tended to cluster by subject, an analysis by PCO, shows the mycobiome of individuals 1 and 6 overlaps, which seems to be driven by both individuals having a composition of ≥99% Saccharomyces. An analysis of Beta diversity (Euclidian, and Bray-Curtis) between individuals over the collection period – with all sample types and sample points – showed significant differences in both bacterial and fungal communities (PERMANOVA $p < 0.02$).

SCFA concentrations were determined from whole stool as well as surface collected aliquots, and overall the average molar ratio of acetate, propionate, and butyrate was 78:12:10 respectively. The mean concentrations of individual or total SCFA in µmol per gram of faeces was not significantly different between collection methods; and in all subjects, acetic acid was most variable, ranging from an average of 103 to 697 µmol g^{-1}.

Comparison of surface aliquots and whole stool sampling methods

To assess the impact of sampling method on a-priori grouping by individual, a hierarchical cluster analysis was performed on bacterial and fungal communities (Figure 2). Bacterial communities from the same subject grouped together, with a SIMPROF test identifying significantly different sub clusters for 5 of 6 individuals. Fungal communities also clustered according to individual, but these groups were less similar than their respective bacterial groups. Furthermore, the aliquots from participant 6 clustered more closely to participant 1 than to its own respective whole stool sample. The mycobiome and bacteriome from the combined aliquot clustered according to individual, however did not seem to align consistently with the other aliquots.

To assess the heterogeneity of SCFAs and bacterial diversity within a single stool, the coefficient of variation (CV) for these measures was compared across three aliquots from a single stool and three separate stools (collected over 3 days) per individual (Supplementary Table 2). Acetic acid and valeric acid, were found to be as variable along a single stool as they were across three bowel movements, whereas propionic was more variable across bowel movements. Overall, the SCFA concentrations were more variable across 3 bowel movements than along a single stool except in individuals 1 and 6. Shannon diversity was less variable along a single stool (5 of 6 individuals), Chao1 species richness was more variable along a single stool (4 of 6 individuals), and Phylogenetic diversity was equally variable along a single stool (3 of 3 individuals). This trend was further assessed by integrating SCFA data with bacteriome data through rCCA (distance between features were relatively short, indicating the strong agreement between datasets), and the plotted canonical variates show the variability between the surface collected samples was still evident in subjects 1, and 6 (Figure 3).

Within each individual the composition of microorganisms within the aliquots were not identical to each other, or to the whole stool from which they were sub sampled. DESEq2 was used to compare differentially abundant ASVs between aliquots and whole stool. This method identified 12 bacterial, and 16 fungal features with a $\log_2$ fold change greater than |2.5|. Of these, 5 bacterial and 1 fungal ASVs were significantly enriched in the whole stool compared to aliquots, and 1 fungal ASV was enriched in stool aliquots compared to whole stool (Table 1).

Bacterial community composition is affected by collection methods

Differences in bacterial communities due to collection method where visualized using PCO (Figure 4) showing a clear separation between directly frozen samples (method F) and those collected with either the Norgen (collection method N), or OMNIgene tubes (collection method O). Significant differences in beta diversity were tested using PERMANOVA, and stool collected with method F were significantly different ($p < 0.01$) to both those collected with the N or O methods. The N and O method where also significantly different to each other ($p = 0.04$). In whole stool samples, the overall most abundant families were Bacteroidaceae (F 38%, N 43%, O 39%), Ruminococcaceae (F 7%, N 17%, O 33%), and Lachnospiraceae (F 10%, N 15%, O 10%), with the abundance of Ruminococcacae significantly increased (ANOVA $P < 0.001$, FDR = 0.007) due to collection using the N and O methods compared to the F method. A number of taxa were also recovered differentially between the three collection methods (Supplementary Table 3), including some high-ranking taxa (Figure 5).

Short term changes to Microbiome composition and SCFA concentration

All 6 participants collected two bowel movements consecutively within a 25-hour period, with 5 of the 6 individuals producing two bowel movements within 10 hours. The total concentration of SCFA was significantly higher in the second stool sample compared to the first using a paired t-test ($p = 0.04$). While not significant, bacterial richness ($p=0.45$) and diversity estimates ($p=0.95$) were also lower in the second stool collection for 4 of 6 individuals (Figure 6). Fungal communities from three individuals which were successfully profiled consecutively did not show any trend between richness and diversity measures.

Stool form according to the Bristol Stool Form Scale (BSFS) was also recorded during sample processing, and most individuals (4 of 6) did not have uniform stool types collected over the three time points. SCFA concentrations clustered in a PCA according to BSFS, and when the SCFA data was integrated with 16S ASV data using rCCA to maximise correlation, the resulting correlation coefficients also grouped loosely according to stool type (Supplementary Figure 2).

Discussion

Analysis of the fecal microbiome is now commonly complemented by an additional analysis of microbial metabolites such as SCFA. To ensure these data can be represented together without the impact of spatial and temporal variability of the fecal material, collection and storage methods for stool samples must be considered. Our results found sporadic detection of low abundance bacterial and fungal species in unhomogenized stool. Further, SCFA concentrations were also shown to vary considerably across a single stool.

The level of variability (CV) in microbial diversity and SCFA concentration across a single stool, was compared per participant to the variability across three separate bowel movements. It was expected that temporal shifts in community structure over three timepoints would be larger than replicate sampling from a single stool. While Shannon Diversity was more variable for 5 of 6 individuals among whole stool samples, Richness based on Chao1 was more variable along a single stool for 4 of 6 individuals. As well, SCFA concentrations were more variable within a single stool than across three separate bowel movements for two individuals. When the bacterial and SCFA data was integrated using rCCA, the intraindividual variability between the aliquots was also evident. Clustering of subsamples from individuals 2, 4, and 5 were very tight, indicating little difference in community structure due to sampling method. Although, samples from subjects 1, 3, and 6 were less tightly clustered, implying community structure changes along the surface of the stool in these individuals that are sensitive to sampling method. This demonstrates that for some individuals, heterogeneity of microorganisms and microbial metabolites in stool may be as great as that observed over the course of two days, which will become apparent if samples are collected by sub-sampling a small volume of stool. This is consistent with reports of heterogeneity in mucosal bacteriome [18], fecal microbiome [19,20], and metabolite concentrations [8]. As fecal material moves through the colon, the exterior surface is exposed to the mucus layer secreted by epithelial cells. This mucus (which is a niche for commensal microbes) accumulates in fecal material, and has been proposed as a mechanism for the patchy recovery of microbial species along the surface of stool [13].

While the surface of the stool may have more variable richness and diversity, the beta diversity of bacterial communities between individuals was still significantly different, indicating that differences due to subsampling are less pronounced than differences between individuals. This is consistent with similar work, where beta diversity (weighted UniFrac) was compared across 9 stool subsampling locations with no significant differences observed [21]. Fungal communities however did not seem to be structured according to the individual to the same extent as bacterial communities, but was structured in one of two ways: Dominated by *Saccharomyces cerevisiae* (≥99%), or by hosting a more even abundance of genera including S. cerevisiae, and either *Kazachstania servazzii* and *Cyberlindnera jadinii*, or *Hanseniapora uvarum* and *Torulaspora delbrueckii*. In another study targeting fungi in the gut using the ITS1 region, three main mycobiome types were found: either dominated by *Candida albicans*, or *Saccharomyces cerevisiae*, or multi species type [22]. In the present study, *Candida* spp were not found at greater than 1% of the total community in any individual. However, *Candida apicola* was also not identified in the fungal mock community which could indicate an unknown technical bias against this group, although presumably not due to primer bias as low abundance was detected and this primer set has been used successfully for other Candida species [23].

The microbiome is often scrutinized for small community changes in association with host-related biological factors such as diet or disease. These changes in microbial signatures are often detected in less abundant taxa, or only within particular groups of bacteria and can vary among individuals. Most bacterial ASVs with large differential abundance were found to be enriched in whole stool compared to

surface aliquots, and all but one Alphaproteobacteria were classified as Clostridia. The Internal regions of stool have previously been shown to harbor significantly higher abundances of Firmicutes and *Bifiobacteria* spp compared to the external surface [19]. In this study, the external surface of stool was likely targeted by surface aliquot collections, rather than the internal regions of stool, and if the internal regions of stool harbor larger abundances of Firmicutes, this might explain some of the differences seen between the surface aliquots and the homogenized whole stool. On the other hand, half of all fungi with large differential abundance were found to be reduced in whole stool compared to the surface aliquots; and of these all but one Dothideomycetes were classified as Saccharomycetes, indicating Saccharomycetes may be a mucosal associated fungus in the gut.

The long-term view of the healthy human gut microbiome seems to show a dynamic community which retains prolonged stability, but is punctuated by periods of disturbance [24,25] On a shorter timescale, diet has been shown to cause fluctuations in microbial species [26], as well as SCFA concentrations [27], but what those shifts may look like across consecutive stools has not been previously explored. While only a small proportion of women defecate more than once a day, defecation frequency is known to be higher in men [28], and positively associated with vigorous physical activity, as well as plant based or high fiber diets [29]. Therefore, the time of day that samples are collected may need to be indicated in sample collection protocols provided to participants. In this study, all women claimed to regularly defecate more than once a day, and the second stool of the day (collected on average 8 hours after) had significantly higher total SCFA concentrations. The second stool also tended to have lower bacterial richness and Shannon diversity index compared to the first stool, although these differences were not significant. Similarly, a recent study assessing the microbiome and SCFA concentrations at a single timepoint in 441 adults found that lower bacterial diversity was associated with higher SCFA concentrations [30]. The association between bacteria and SCFA concentration seen in this study supports the idea that bacterial metabolites are linked to the circadian clock [31], and demonstrates why time of stool sample collection may be particularly important in individuals who defecate more than once per day.

Decreasing bacterial richness has also been found to correlate with decreasing stool firmness, or a higher Bristol Stool form value, based on fecal samples from 53 women [32]. As well, the BSFS has also been shown to be a good predictor of whole-gut transit time, with high stool form values correlating to longer transit times [28,33,34]. A more recent study also found when stool form had a Bristol Stool value of less than three it was correlated with greater transit times, indicating that stool form can help predict whole-gut or colonic transit times [35]. While this study had a small sample size, it was interesting to note that both SCFA and bacterial phylogenetic diversity grouped according to stool form, and when these data were integrated through rCCA this trend was also observed. If microbial diversity and SCFA concentration are also linked to stool form and potentially transit time, assessing stool form at the point of sample processing could be a simple way to add valuable information to downstream multivariate analysis, and help explain sample clustering. Further, to reduce within-day variability that could potentially distort a long-term study, participants could be instructed to collect at a similar time, such as the first bowel movement of the day.

Directly freezing stool samples with no additional solution is considered the gold standard method for storing stool, while Norgen and in OMNIgeneGUT tubes offer a convenient method of collecting fecal material from remote participants. Regardless of collection method, all whole stool samples were dominated by Bacteroidaceae, but the second most abundant family Ruminococcaceae were significantly expanded in samples collected with both the O and N methods compared the F method, indicating that the two preservation methods may impact fecal microbiomes in a similar way. Each of the three collection methods was significantly different, although the differences between the directly frozen samples compared to either of the other two other room temperature preservation methods was most obvious.

Two previous studies have also compared fecal bacterial communities collected using OMNIgeneGUT kits which were frozen prior to processing with samples which were immediately frozen. One study found storage methods, contributed to the significant differences between samples based on Bray-Curtis dissimilarity measure, and that those collected in OMNIgeneGUT kits had a significant increase in Proteobacteria [36]; while another study found that samples stored in OMNIgeneGUT tubes resulted in microbiome profiles with decreased Actinobacteria and increased Lenthisphaerae compared to those that were frozen without stabilization [15]. Within our study, the preservation tubes were kept at room temperature – in accordance with manufacturer's instructions – and at the phyla level Actinobacteria were also reduced in fecal samples collected with both the O and N methods. It is more likely then, that the reduction in Actinobacteria is a result of storing in a preservation liquid, rather than the storage temperature.

Stool sample collection methods must not sacrifice sample "viability" for convenience, therefore, where possible we recommend collecting stool in bulk and freezing immediately. As well, during sample processing technicians can record the stool form according to BSFS, and homogenize the entire sample prior to subsampling for analysis. This method eliminates any subsampling bias due to heterogenous distribution of microbes in stool, and provides enough material for multiple assays. Additionally, because this method is less hands-on for participants, it may increase compliance if multiple collections are required. For studies where it is not possible to store a large quantity of bulk stool or where frozen transportation of stool is not viable, commercial preservation tubes may be an attractive alternative. In this circumstance it is recommended to only use a single tube type and insure a standard protocol. Furthermore, if OMNIgeneGUT or Norgen collection kits are used, researchers should be cautious in interpreting the reduced abundance of Firmicutes and Actinobacteria. Lastly, collection protocols should consider that some individuals can regularly have more than one bowel movement per day, and those participants should be instructed, where practical, to collect stool at a similar time.

Declarations

Data Availability

ASV tables, metadata and sequences reads are available in the figshare repository and can be accessed currently with the private link https://figshare.com/s/0372d215da9fb7c5588d. ASV tables and metadata will be made publicly available in the figshare repository at the time of publication, while sequence reads will be made publicly available on the NCBI Short Read Archive.

Acknowledgments

We are grateful to all participants of this study. We would also like to extend our thanks to Johney Lo for his assistance with statistical analysis, Sinéad Jones for her artwork used in Figure 7, and Rose Lines for her technical assistance with laboratory work.

Author Contributions

C.T.C, S.R, J.J and D.P contributed to study concept. J.J recruited participants and performed all laboratory work. A.A and J.J participated in data processing, and J.J performed all analyses and drafted the manuscript. All authors contributed to reviewing the manuscript and approved the final manuscript.

Additional Information

Competing Interests: The authors declare no competing interests.

References

1. Valles-Colomer, M. *et al.* The neuroactive potential of the human gut microbiota in quality of life and depression. *Nat. Microbiol.* **4**, 623–632 (2019).
2. Vogt, N. M. *et al.* Gut microbiome alterations in Alzheimer's disease. *Sci. Rep.* **7**, 1–11 (2017).
3. Strati, F. *et al.* New evidences on the altered gut microbiota in autism spectrum disorders. *Microbiome* **5**, 24 (2017).
4. Fieten, K. B. *et al.* Fecal microbiome and food allergy in pediatric atopic dermatitis: A cross-sectional pilot study. *Int. Arch. Allergy Immunol.* **175**, 77–84 (2018).
5. Venegas, D. P. *et al.* Short chain fatty acids (SCFAs)mediated gut epithelial and immune regulation and its relevance for inflammatory bowel diseases. *Frontiers in Immunology* **10**, 277 (2019).
6. D'Amore, R. *et al.* A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling. *BMC Genomics* **17**, 55 (2016).
7. Karu, N. *et al.* A review on human fecal metabolomics: Methods, applications and the human fecal metabolome database. *Analytica Chimica Acta* **1030**, 1–24 (2018).
8. Gratton, J. *et al.* Optimized Sample Handling Strategy for Metabolic Profiling of Human Feces. *Anal. Chem.* **88**, 4661–4668 (2016).

9. De Filippis, F. *et al.* High-level adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated metabolome. *Gut* **65**, 1812–1821 (2016).
10. Arrieta, M. C. *et al.* Associations between infant fungal and bacterial dysbiosis and childhood atopic wheeze in a nonindustrialized setting. *J. Allergy Clin. Immunol.* **142**, 424-434.e10 (2018).
11. Lagkouvardos, I. *et al.* Gut metabolites and bacterial community networks during a pilot intervention study with flaxseeds in healthy adult men. *Mol. Nutr. Food Res.* **59**, 1614–1628 (2015).
12. Tang, Z. Z. *et al.* Multi-omic analysis of the microbiome and metabolome in healthy subjects reveals microbiome-dependent relationships between diet and metabolites. *Front. Genet.* **10**, 454 (2019).
13. Donaldson, G. P., Lee, S. M. & Mazmanian, S. K. Gut biogeography of the bacterial microbiota. *Nat. Rev. Microbiol.* **14**, 20–32 (2016).
14. Couch, R. D. *et al.* The Approach to Sample Acquisition and Its Impact on the Derived Human Fecal Microbiome and VOC Metabolome. *PLoS One* **8**, e81163 (2013).
15. Penington, J. S. *et al.* Influence of fecal collection conditions and 16S rRNA gene sequencing at two centers on human gut microbiota analysis. *Sci. Rep.* **8**, (2018).
16. Flores, G. E. *et al.* Temporal variability is a personalized feature of the human microbiome. *Genome Biol.* **15**, 531 (2014).
17. Caporaso, J. G. *et al.* Moving pictures of the human microbiome. *Genome Biol.* **12**, R50 (2011).
18. Hong, P.-Y., Croix, J. A., Greenberg, E., Gaskins, H. R. & Mackie, R. I. Pyrosequencing-Based Analysis of the Mucosal Microbiota in Healthy Individuals Reveals Ubiquitous Bacterial Groups and Micro-Heterogeneity. *PLoS One* **6**, e25042 (2011).
19. Gorzelak, M. A. *et al.* Methods for improving human gut microbiome data by reducing variability through sample processing and storage of stool. *PLoS One* **10**, (2015).
20. Wesolowska-Andersen, A. *et al.* Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis. *Microbiome* **2**, (2014).
21. Liang, Y. *et al.* Systematic Analysis of Impact of Sampling Regions and Storage Methods on Fecal Gut Microbiome and Metabolome Profiles. *mSphere* **5**, (2020).
22. Motooka, D. *et al.* Fungal ITS1 deep-sequencing strategies to reconstruct the composition of a 26-species community and evaluation of the gut mycobiota of healthy Japanese individuals. *Front. Microbiol.* **8**, 238 (2017).
23. Heisel, T. *et al.* Complementary amplicon-based genomic approaches for the study of fungal communities in humans. *PLoS One* **10**, e0116705 (2015).
24. Fu, B. C. *et al.* Temporal variability and stability of the fecal microbiome: The multiethnic cohort study. *Cancer Epidemiol. Biomarkers Prev.* **28**, 154–162 (2019).
25. Voigt, A. Y. *et al.* Temporal and technical variability of human gut metagenomes. *Genome Biol.* **16**, 73 (2015).
26. David, L. A. *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559–563 (2014).

27. McOrist, A. L. *et al.* Fecal Butyrate Levels Vary Widely among Individuals but Are Usually Increased by a Diet High in Resistant Starch. *J. Nutr.* **141**, 883–889 (2011).
28. Heaton, K. W. *et al.* Defecation frequency and timing, and stool form in the general population: a prospective study. *Gut* **33**, 818–824 (1992).
29. Sanjoaquin, M. A., Appleby, P. N., Spencer, E. A. & Key, T. J. Nutrition and lifestyle in relation to bowel movement frequency: a cross-sectional study of 20 630 men and women in EPIC–Oxford. *Public Health Nutr.* **7**, 77–83 (2004).
30. de la Cuesta-Zuluaga, J. *et al.* Higher fecal short-chain fatty acid levels are associated with gut microbiome dysbiosis, obesity, hypertension and cardiometabolic disease risk factors. *Nutrients* **11**, 51 (2019).
31. Murakami, M. & Tognini, P. The Circadian Clock as an Essential Molecular Link Between Host Physiology and Microorganisms. *Frontiers in Cellular and Infection Microbiology* **9**, 469 (2020).
32. Vandeputte, D. *et al.* Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. *Gut* **65**, 57–62 (2016).
33. Degen, L. P. & Phillips, S. F. How well does stool form reflect colonic transit? *Gut* **39**, 109–113 (1996).
34. Lewis, S. J. & Heaton, K. W. Stool form scale as a useful guide to intestinal transit time. *Scand. J. Gastroenterol.* **32**, 920–924 (1997).
35. Saad, R. J. *et al.* Do Stool Form and Frequency Correlate With Whole-Gut and Colonic Transit? Results From a Multicenter Study in Constipated Individuals and Healthy Controls. *Am. J. Gastroenterol.* **105**, 403–411 (2010).
36. Choo, J. M., Leong, L. E. & Rogers, G. B. Sample storage conditions significantly influence faecal microbiome profiles. *Sci. Rep.* **5**, 16350 (2015).
37. Turner, S., Pryer, K. M., Miao, V. P. W. & Palmer, J. D. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. in *Journal of Eukaryotic Microbiology* **46**, 327–338 (John Wiley & Sons, Ltd (10.1111), 1999).
38. Caporaso, J. G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 4516–4522 (2011).
39. Murray, D. C., Coghlan, M. L. & Bunce, M. From Benchtop to Desktop: Important Considerations when Designing Amplicon Sequencing Workflows. *PLoS One* **10**, e0124671 (2015).
40. Bakker, M. G. A fungal mock community control for amplicon sequencing experiments. *Mol. Ecol. Resour.* **18**, 541–556 (2018).
41. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10 (2011).
42. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).
43. Kõljalg, U. *et al.* UNITE: A database providing web-based methods for the molecular identification of ectomycorrhizal fungi. *New Phytol.* **166**, 1063–1068 (2005).

44. McKnight, D. T. *et al.* microDecon: A highly accurate read-subtraction tool for the post-sequencing removal of contamination in metabarcoding studies. *Environ. DNA* **1**, 14–25 (2019).
45. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One* **8**, e61217 (2013).
46. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
47. Rohart, F., Gautier, B., Singh, A. & Lê Cao, K.-A. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLOS Comput. Biol.* **13**, e1005752 (2017).

Table

Table 1. ASVs identified with log2 fold change in gematric mean abundance greater than |2.5| between homogenized whole stool and stool aliquots. * Indicates significant enrichment

ASV	Taxa	log2 Fold Change
Bacteria		
591	Anaerotruncus massiliensis	6.62
790	Anaerovoracaceae	17.92*
802	Anaerovoracaceae	17.92*
70	Eubacterium sp	2.89
4	Fecalibacterium prausnitzii	-3.41
27	Fecalibacterium prausnitzii	2.71
405	Fecalibacterium prausnitzii	17.47*
717	Oscillibacter ruminantium	19.52*
461	Rhizobiaceae	18.36*
Fungi		
19	Alternaria alternata	-2.9
219	Aspergillus niger	3.9
119	Aureobasidium pullulans	17.2*
8	Cyberlindnera jadinii	3.9
4	Eremothecium sinecaudum	-3.5
41	Hanseniaspora uvarum	-24.1
11	Hanseniaspora uvarum	-8.6
5	Kazachstania barnettii	-22.5
2	Kazachstania servazzii	-21.4*
132	Rhodotorula mucilaginosa	9.1
29	Saccharomyces cerevisiae	-23.5
6	Saccharomyces cerevisiae	4.9
17	Sporopachydermia lactativora	5.0
40	Wickerhamomyces ciferrii	7.3