Title

Stability of the Fecal and Oral Microbiome over 2 Years at -80°C for Multiple Collection Methods.

Journal

Cancer Epidemiology, Biomarkers and Prevention, 32(3)

Authors

Abnet, Christian
Vogtmann, Emily
Sinha, Rashmi
et al.

Publication Date

2023-03-06

DOI

10.1158/1055-9965.EPI-22-0883

Peer reviewed

HHS Public Access

Author manuscript

Cancer Epidemiol Biomarkers Prev. Author manuscript; available in PMC 2023 September 13.

Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2023 March 06; 32(3): 444–451.

doi:10.1158/1055-9965.EPI-22-0883.

Stability of the fecal and oral microbiome over two years at −80°C for multiple collection methods

Semi Zouiouich[1], **Doratha A Byrd**[2], **Xing Hua**[3], **Smriti Karwa**[1], **Yunhu Wan**[1], **Jianxin Shi**[1], **Gregory C Humphrey**[4], **Gail L Ackermann**[4], **Rob Knight**[4], **Christian C Abnet**[1], **Emily Vogtmann**[1], **Rashmi Sinha**[1]

[1]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland

[2]Department of Cancer Epidemiology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida

[3]Fred Hutchinson Cancer Research Center, Seattle, Washington

[4]Department of Pediatrics, University of California, San Diego, California.

Abstract

Background: In prospective cohorts, biological samples are generally stored over long periods before an adequate number of cases have accrued. We investigated the impact of sample storage at −80°C for two years on the stability of the V4 region of the 16S rRNA gene across seven different collection methods (i.e., no additive, 95% ethanol, RNAlater stabilization solution, fecal occult blood test cards, and fecal immunochemical test tubes for feces; OMNIgene ORAL tubes and Scope mouthwash for saliva) among 51 healthy volunteers.

Materials and Methods: Intraclass correlation coefficients (ICCs) were calculated for the relative abundance of the top three phyla, the 20 most abundant genera, three alpha diversity metrics, and the first principal coordinates of three beta diversity matrices.

Results: The subject variability was much higher than the variability introduced by the sample collection type, and storage time. For fecal samples, microbial stability over two years was high across collection methods (range, ICCs=0.70–0.99), except for the samples collected with no additive (range, ICCs=0.23–0.83). For oral samples, most microbiome diversity measures were stable over time with ICCs above 0.74; however, ICCs for the samples collected with Scope mouthwash were lower for two alpha-diversity measures, Faith's phylogenetic diversity (0.23) and the observed number of operational taxonomic units (0.23).

Conclusions: Fecal and oral samples in most used collection methods are stable for microbiome analyses after two years at 80°C, except for fecal samples with no additive.

Corresponding author, Semi Zouiouich, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 9609 Medical Center Dr. Rockville, MD 20850, semi.zouiouich@nih.gov.

Authors' contributions: Conceptualization, D.B., J.S., R.K., C.A., E.V. and R.S.; laboratory methodology G.H., G.A. and R.K.; bioinformatics, Y.W.; formal analysis, S.Z., S.K., X.H., and J.S.; writing—original draft preparation, S.Z.; writing—review and editing D.B., J.S., Y.W., E.V. and R.S.; All authors have read and agreed to the published version of the manuscript.

Conflict of interest: The authors declare no potential conflicts of interest.

Impact: This study provides evidence that samples stored for an extended period from prospective studies are useful for microbiome analyses.

Keywords

Oral microbiome; Fecal microbiome; Stability; Collection methods

Introduction

Most of the current knowledge on the human microbiome and its role in the development of chronic diseases is based on cross-sectional studies with small sample sizes and a lack of standardized protocols (1–3). Consequently, findings are generally too inconsistent to draw firm conclusions regarding the microbiome's contribution to disease etiology (4,5) or to rule out reverse causation. Repeated, prospectively collected samples, using the best, standard methods, from population-based cohort studies are necessary for better understanding of the temporal nature of microbiome-disease associations. To this end, previous studies have investigated various collection methods and their impact on multiple estimates of microbial composition in fecal (6–12) and oral samples (13–16) to support large-scale microbiome specimen collection in population-based studies.

While immediately freezing fecal samples at −80°C without additive is considered as the current putative "gold standard", this is likely infeasible in larger population-based studies. Recent prospective cohorts started to collect fecal samples using alternative methods such as OMNIgene GUT kit, 95% ethanol, fecal occult blood test (FOBT) cards, or fecal immunochemical test (FIT) tubes. The latter two collection methods support the possibility to create prospective cohorts by partnering with colorectal screening programs. Millions of FOBT and FIT samples are being collected each year in colorectal cancer screening registries, providing rich opportunities to establish large epidemiologic studies with serially-collected fecal samples. FOBT and FIT have been found to be reproducible, relatively concordant with samples stored with no additive, and stable after short periods at room temperature (8,9,11,17) supporting their collection for microbiome research after use for colorectal cancer screening.

Contrary to fecal sample collection, there is currently no gold standard collection method for studying the oral microbiome. Previous studies have compared various sample collection and storage methods to characterize oral microbiome using saliva and oral wash samples (13–15). Two commonly used collection methods, OMNIgene ORAL kit and Scope mouthwash, were recently compared and had differences in relatives abundances of specific taxa (13,15) but similarities in alpha and beta diversity metrics. Overall, the use of Scope mouthwash for large epidemiologic studies could represent a cheaper and more practical alternative to OMNIgene and saliva samples containing no additive and frozen immediately after collection.

In prospective cohorts, biological samples are generally stored over long periods before an adequate number of cases have accrued. At present, only a few studies have investigated the stability of fecal and oral microbiome over long periods in a freezer (18–20). Therefore, we investigated fecal and oral microbiome stability in samples stored for two years in a −80°C

freezer, collected via seven different collection methods (i.e., no additive, 95% ethanol, RNA*later* stabilization solution, FOBT cards, and FIT tubes for fecal samples; OMNIgene ORAL tubes and Scope mouthwash for oral samples), using 16S rRNA gene amplicon sequencing data, among 51 volunteers.

Materials and Methods

Study population.

A detailed report of this population was described previously (9). Briefly, aliquots of oral and fecal samples from 51 healthy volunteers from the Mayo II study were used in the current project (Table 1). All participants provided written informed consent and the study was approved by the Mayo Clinic Studies Institutional Review Board and the NCI Office of Human Subjects Research (approval number 12189).

Fecal and oral specimen collection.

The fecal and oral sample collection methods were explained in detail in two previously published studies (9,15). As described before, the participants collected stool at the clinic and immediately delivered it to the study coordinator for processing. The fecal specimens for each individual were mixed manually using a spatula, and aliquoted to each of the five different collection methods - Sarstedt feces tube (Numbrecht, Germany) containing no additive, 2.5 mL of 95% ethanol (Sigma-Aldrich, St. Louis, Missouri), 2.5 mL of RNA*later* Stabilization Solution, FOBT cards (Beckman Coulter, Brea, California), and FIT tubes (Polymedco, Inc., Cortlandt Manor, New York). Each participant provided saliva using an OMNIgene ORAL kit (DNA Genotek Inc., Ottawa, Ontario, Canada) and a Scope mouthwash (Procter & Gamble Co., Cincinnati, Ohio) in person in the laboratory. An outline of the samples used for this study is presented in Table 1. For fecal samples, two aliquots of samples collected with no additive, 95% ethanol, RNAlater and FIT, and three aliquots of samples collected with FOBT were frozen immediately at −80°C (Day 0). In addition, two aliquots of samples collected with 95% ethanol, RNAlater and FIT, and three aliquots of samples collected with FOBT remained at room temperature for 96 hours (Day 4). At the end of the four days, the remaining aliquots were frozen at −80°C. For oral samples, one aliquot of samples collected with OMNIgene tube and Scope mouthwash was frozen immediately at −80°C (Day 0) and one aliquot of Scope mouthwash remained at room temperature for 96 hours (Day 4) before being frozen at −80°C. No aliquot of OMNIgene tube remained at room temperature for 96 hours as the cost was prohibitive. In total, accounting for all the time points, for fecal samples collected with 95% ethanol, FIT, no additive and RNAlater, 12 aliquots were produced; for fecal samples collected with FOBT cards, 18 aliquots were created.

DNA extraction and sequencing.

DNA extraction and sequencing were performed at the University of California, San Diego (La Jolla, California). The methods for DNA extraction, PCR amplification, and sequencing were performed as described by Caporal et al. (21) and were explained in detail previously (8,9,15). Briefly, DNA was extracted using the MO-BIO PowerMag Soil DNA Isolation Kit. The universal bacterial primer set 515F/806R was used to PCR amplify the V4 region

of the 16S rRNA gene. All barcoded amplicons were pooled with equal concentrations for sequencing on Illumina's HiSeq. As listed in Table 1, one set of each sample was initially extracted upon receipt (baseline), after a year frozen at −80°C (year 1), and after two years frozen at −80°C (year 2). All of the 16S rRNA gene amplicon sequencing was done at the same time, at the end of year 2. The average coverage was approximately 34,000 reads per sample.

Bioinformatic processing.

Bioinformatics processing was performed using QIIME2 2019.1 (22). Reads were demultiplexed and quality filtered using DADA2 1.9.3 with a Phred quality score of 33 (23). After error correction, chimera removal, and removal of phiX sequences through DADA2 and non-bacteria sequences through lack of taxonomic classification, the cleaned read files were joined to make a single sequence variant table, with each sequence variant representing a unique pair sequence. Taxonomic classification was performed with a naive Bayes classifier trained on the SILVA v132 99% (operational taxonomic unit) OTU database that includes only the V4 region (defined by the 515F/806R primer pair) (24).

Based on rarefaction curves, alpha and beta diversity measures were calculated after rarefaction to 27,000 reads per sample for both fecal and oral samples; no samples were excluded after rarefaction. Alpha diversity measures, i.e., observed OTUs to estimate taxa richness, Shannon Diversity to estimate taxa richness and evenness, and Faith's Phylogenetic Diversity to consider phylogenetic differences in samples, were calculated using the R phyloseq package (25). For beta diversity measures, the Bray-Curtis distance was calculated to capture differences in microbial abundances between samples using the R vegan package (26). Unweighted UniFrac, and weighted UniFrac were calculated to capture differences between samples based on sequence distances using the R GUniFrac package (27).

Statistical analysis.

To quantify the percentage of microbiome variability explained by subject, sample collection type, and the number of years frozen, a distance-based coefficient of determination R^2 was estimated using a permutational multivariate analysis of variance ('adonis' function, vegan package, R). This summary of the overall variability between sample characteristics was based on three beta diversity distances (i.e., Bray-Curtis, unweighted UniFrac, and weighted UniFrac).

The stability over time in a freezer of the oral and fecal samples, for each collection method, was calculated using intraclass correlation coefficients (ICCs) using a linear mixed effects model. The ICCs were calculated based on (i) the square root of the relative abundances of the three most dominant phyla (Firmicutes, Bacteroidetes, and Actinobacteria for the fecal samples, and Firmicutes, Bacteroidetes, and Proteobacteria for the oral samples) and the 20 most prevalent genera, which were present in at least 10% of fecal samples with relative abundance of >0.001; (ii) three alpha diversity metrics (observed OTUs, Shannon Diversity, and Faith's Phylogenetic Diversity); and (iii) the first multidimensional scaling axis, also called first principal coordinate (PC1), of three beta diversity metrics (Bray-Curtis, 22.8% and 9.1% of explained variability for fecal and saliva samples respectively; unweighted

UniFrac, 8.9% and 3.2% of explained variability for fecal and saliva samples respectively; weighted UniFrac, 20.8% and 9.2% of explained variability for fecal and saliva samples respectively). The 95% confidence intervals (95% CI) were estimated using 1,000 bootstrap samples.

We conducted a differential abundance analysis using the Wilcoxon signed-rank test to identify the bacterial taxa at the phylum, family and genus level which were differentially abundant among baseline, year 1, and year 2 samples. Taxa read counts were normalized into proportions before analysis and taxa with a prevalence less than 10% or a detection threshold for absence/presence <0.001 were excluded from testing. False discovery rate (FDR) control using the Benjamini-Hochberg procedure was used to correct for multiple testing. We replicated this differential abundance analysis using a methodology called Analysis of Compositions of Microbiomes with Bias Correction (ANCOM-BC) to estimates the unknown sampling fractions among samples and model the absolute abundance of genera. Results from the ANCOM-BC were not substantially different from results from the Wilcoxon signed-rank test.

For simplicity, for the above-described analyses, we primarily focused on samples frozen immediately at −80°C; however, we repeated our analysis on samples that were frozen after four days at room temperature (Supplementary Table 1 and 2). As described previously, duplicates or triplicates of fecal samples were made for each collection method (Table 1). To simplify the interpretation of our results, we randomly selected one set of aliquots over the three timepoints (baseline, year 1, year 2) for each of the 51 individuals; we replicated our analysis in the set of aliquots which was discarded from the main analysis (Supplementary Table 3). We also compared the ICC among the duplicates. Overall, duplicates were very comparable for most collection methods, even after one year or two years of storage with ICCs varying between 0.7 and > 0.9. However, the comparison of duplicates of fecal samples collected with no additive varied depending on the time of storage. At baseline, ICCs varied between 0.8 and >0.9 except for Bray-Curtis (0.4) and weighted UniFrac (0.3), which might be driven by changes in proportions of certain taxa. After one year of storage, ICCs varied between 0.6 and >0.9 except for the relative abundance of Bacteroidetes (0.1). After two years of storage, ICCs varied between 0.2 and 0.9 except for Bacteroidetes (<0.1).

Data availability

Data are available at the National Center for Biotechnology Information Sequence Read Archive (accession number PRJNA846182).

Results

Percent variability explained by subject, collection type, and storage time.

Based on Bray-Curtis and weighted and unweighted UniFrac distance matrices, interindividual variability explained 55% to 65% of the overall variability in fecal samples and 51% to 78% in oral samples, which was higher than the variability introduced by the sample collection type (4% to 13% for fecal samples, and 3% to 8% for oral samples) and the storage time (<1% to 2% for both fecal and oral samples) for all measures of beta

diversity (Figure 1). When removing the stool samples stored with no additive from the analysis, interindividual variability explained 65% to 79% of the overall variability in fecal samples, while the variability introduced by the sample collection type (2% to 8%) and the storage time (<1%) was lower.

Stability of fecal samples over two years of storage.

Stability as measured by ICCs over two years in a freezer was generally high for all fecal collection methods, excluding no additive, with ICCs ranging from 0.73 to 0.99 (Figure 2, Supplementary Table 4). However, for the samples collected with no additive, the ICCs for all microbial measures were lower. ICCs for phyla Bacteroidetes and Firmicutes were 0.23 and 0.32 respectively; the ICC for the Shannon Diversity was 0.42; and finally, ICCs for Bray-Curtis, unweighted UniFrac and weighted UniFrac were 0.43, 0.83 and 0.45, respectively. In most collection methods, the ICCs for the relative abundance of the top 20 genera were generally high, except for *Streptococcus*, with ICCs varying from 0.05 to 0.93 depending on the collection method (Supplementary Table 5). However, for the samples collected with no additive, the ICCs were lower for the top 20 genera with ICCs ranging from 0.18 to 0.78.

Stability of oral samples over two years of storage.

For the oral samples, most microbiome diversity measures were stable over time with ICCs above 0.74 (Figure 3, Supplementary Table 6). However, ICCs for the samples collected with Scope mouthwash were lower for the Faith's phylogenetic diversity and the observed OTUs, with ICCs equal to 0.23 for the two alpha-diversity metrics. For the relative abundance of the top 20 genera, the ICCs were generally above 0.71, except for *Bacteroides* (ICC = 0.11 and ICC = 0.19), *Faecalibacterium* (ICC = 0.11 and ICC = 0.08) and *Roseburia* (ICC = 0.15 and ICC = 0.09) in Scope mouthwash and OMNIgene tube, respectively (Supplementary Table 7).

The replication of this analysis in samples frozen after four days at room temperature produced similar results, which are presented in Supplementary Table 1 and 2. The ICCs for duplicated or triplicated fecal samples – discarded from the main analysis – were comparable to the ICCs from the main analysis (Supplementary Table 3).

Differential abundance comparisons for fecal samples.

Comparing the relative abundances at the phylum and genus level, most collection methods seemed to have similar composition when comparing samples at baseline, year 1, and year 2 (Figure 4) except for samples stored with no additive. In fecal samples stored with no additive, 27 genera statistically significantly differed across two years, including *Bacteroides*, *Faecalibacterium* or *Bifidobacterium* (Supplementary Table 8). For example, the average relative abundance of *Bacteroides* decreased from 28.13% at baseline to 2.25% at year 2 while the relative abundance of *Blautia* increased from 5.73% at baseline to 16.84% at year 2. For other fecal collection methods, there were substantially fewer taxa that had statistically significant shifts across the two years. Additionally, these differences mostly included low abundance genera with relative abundances <1%; most notably, for FIT

samples, we observed an average increase in the relative abundance of *Faecalibacterium*, from 8.62% at baseline to 14.96% at year 2.

Differential abundance comparisons for oral samples.

For oral samples, we observed a statistiscally significant increase in the relative abundance of *Actinomyces* in both OMNIgene tubes and Scope mouthwash (1.14% and 1.93% at baseline to 1.91% and 5.14% at year 2). In addition, in Scope mouthwash samples, we detected alterations in the relative abundances of more highly abundant taxa. For example, the relative abundance of the phyla Actinobacteria and Firmicutes increased from 8.76% and 42.07% at baseline to 12.96% and 50.02% at year 2, respectively. At the genus level, the relative abundance of *Streptococcus* increased from 23.31% at baseline to 32.32% at year 2.

Discussion

Among 51 healthy volunteers from the Mayo II study, microbial variability was largely explained by subject differences, followed by the sample collection type and minimally by the storage time in both fecal and oral samples. For many of the collection methods, the stability of fecal and oral microbiome in samples stored for two years at −80°C was generally high for the relative abundance of three phyla, three alpha-diversity metrics, three beta-diversity matrices, and the relative abundances of most of the top 20 genera. However, for the fecal samples collected with no additive, the ICCs were lower for the relative abundance of Bacteroidetes and Firmicutes, Shannon Diversity and, Bray-Curtis and weighted UniFrac. In addition, ICCs for the samples collected with Scope mouthwash were lower for the Faith's phylogenetic diversity and the observed OTUs. Many of the relative abundances of the top 20 genera also had lower ICCs for fecal samples collected with no additive. When comparing the relative abundances at the phylum, family and genus level, all the collection methods for both fecal and oral samples, except for fecal samples collected with no additive, had overall similar composition when comparing samples at baseline, year 1, and year 2. However, differences in low abundant genera were observed in all the collection methods for both fecal and oral samples. These results suggest that studies must consider long-term stability before selecting appropriate collection methods when storing fecal and oral samples.

Some previous studies have evaluated the short-term, from a few days to a few months, as well as long-term stability of several years, of fecal samples for microbial analyses using no additive and RNAlater (17–19,27–29). For fecal samples collected in RNAlater from 24 individuals, −80°C storage for up to five years had little effect on the gut microbiota composition (19). For fecal samples collected with no additive, in nine infant fecal samples stored at −80°C for two years, significant changes in the microbial community, with a minor reduction in the observed number of taxa and variations in the abundance of some specific taxa were seen (20). Our results were consistent with previous findings suggesting that the technical variability introduced by storage effects (6 months to two years) is lower than the interindividual microbiota variability (19,28). In more complex microbiota from adults, we also found significant changes in relative abundances of specific genera and estimated moderate to poor stability of several diversity metrics in fecal samples stored

with no additive for two years. Although considered as the “gold standard” method, the stability of samples stored immediately with no solution is questionable for long periods of storage, even when left at −80°C. Indeed, as previously described, freezing fecal samples prior to DNA extraction might impact the relative abundances of Gram-positive and Gram-negative bacteria (29,31). This is an important consideration for future microbiome cohorts as immediately freezing fecal samples with no additive at −80°C is likely to be infeasible in the context of large population-based studies.

Our results suggest that the collection of fecal samples in FIT tubes and FOBT cards is feasible as long-term storage is compatible with fecal microbiome analyses in these commonly used sample collection methods from colorectal cancer screening registries. For fecal samples collected with FOBT and FIT, we had good to excellent stability of the gut microbiome over two years of storage −80°C; although, we observed alterations in the relative abundance of a few low abundant genera in both methods, and an increase in the relative abundance of *Faecalibacterium* from 8.62% at baseline to 14.96% at year 2 in FIT samples. This supports the hypothesis that colorectal cancer screening registries in which millions of FOBT or FIT samples are being collected each year, might provide rich opportunities to establish cohorts of individuals with pre-diagnostic collected fecal samples (8,9,11,12). Indeed, to attain adequate power to detect disease associations with microbiome profiles, future studies will likely have to recruit a large cohort or follow the cohort for a long times, from several years to decades (32). Long-term stability of the gut microbiome over a long period of storage at −80°C is therefore an important paremeter to take into account when establishing prospective population-based cohorts. In addition, FOBT cards stored at room temperature for 10 weeks as previously showed excellent stability when compared with the immediately frozen cards (12); this indicates that future microbiome population-based cohorts would not have to immediately freeze fecal samples after collection, which might be required in large prospective cohorts.

To our knowledge, no study has considered the impact of long-term storage of oral samples on oral microbiome characteristics. Our findings suggest that the duration in which saliva samples remain stored at −80°C has only limited impact on the stability of the oral microbiome for both OMNIgene ORAL kit and Scope mouthwash. However, for samples collected with Scope mouthwash, two alpha-diversity metrics showed low stability; these results might be driven by the variation of the relative abundances of specific bacteria after two years of storage. Specifically, when comparing the Scope mouthwash samples at baseline and at year 2, we observed alterations in the relative abundances of genera with relative abundance >1%, i.e., higher levels of *Streptococcus* and *Actinomyces* after two years of storage. Interestingly, although both OMNIgene ORAL kit and Scope mouthwash contain different antimicrobial agents that inhibit the growth of bacteria (13), we observed low stability of the relative abundances of the genera *Bacteroides*, *Faecalibacterium*, and *Roseburia* in the two collection methods. In the absence of replication and validation of these results, it is unclear whether these changes were related to actual variations of the relative abundance of these three taxa during storage or to other technical variations.

This study has several limitations. First, only healthy donors were included in this study. As such, the results presented may not be representative of less diverse microbiome

compositions of diseased individuals. Second, the duration of storage at −80°C was limited to two years. Although, our results suggest that two-years freezing is, for most collection methods, compatible with fecal and oral microbiome analyses, future microbiome prospective cohorts might store fecal and oral samples for decades before sufficient numbers of diseased cases are reached. Third, the lack of technical duplicates for oral samples prevented the separation of technical variation due to long-term storage from other technical variability. Though, technical reproducibility in OMNIgene ORAL kits and Scope mouthwash was usually high in previous studies (13,15,16), and typical epidemiologic studies are based on single-point sampling with no technical duplicates. Finally, we used 16S rRNA gene sequencing to characterize the microbial metrics, while other profiling methods such as whole-genome shotgun metagenomics are becoming more common. However, 16S rRNA gene sequencing remains the most affordable method to study the bacterial microbiome diversity, and other scientists may find these results helpful in designing large epidemiologic cohorts for microbiome studies.

This analysis also has several important strengths. First, with 51 healthy volunteers, there was greater statistical power to determine long-term stability of diverse microbiome metrics than any previous study, in both fecal and oral samples. Second, to our knowledge, our study was the first to consider the impact of long-term storage of oral samples on oral microbiome characteristics. Third, our analysis was the first to include a wide variety of collection methods, comprising FOBT and FIT, which represent viable methods for the opportunistic collection of fecal samples after colorectal cancer screening.

Currently, immediately freezing fecal samples collected with no additive is considered as the best collection method to evaluate associations between the gut microbiome and diseases; however, our findings suggest that fecal samples should be preserved via RNAlater, 95% ethanol, FIT tubes or FOBT cards when stored over long periods of time. When stored via these methods, our findings support that the fecal microbiome will remain stable for up to two years of storage at −80°C. While we cannot make a direct simple recommendation of a specific method on our analysis, we have highlighted several issues that investigators should be aware of when storing fecal samples for long period. Overall, we recommand that future studies use the same collection method to be able to compare results across cohorts. Our results also suggest that storing saliva samples at −80°C for two years has only limited impact on the stability of the oral microbiome for both OMNIGene ORAL kit and Scope mouthwash. Future studies are needed to replicate our findings across longer follow up periods.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial support/Grants:

This work was supported by the Intramural Research Program of the National Cancer Institute at the National Institutes of Health.

References

1. Huybrechts I, Zouiouich S, Loobuyck A, Vandenbulcke Z, Vogtmann E, Pisanu S, et al. The Human Microbiome in Relation to Cancer Risk: A Systematic Review of Epidemiologic Studies. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2020 Oct;29(10):1856–68.
2. Robinson CK, Brotman RM, Ravel J. Intricacies of assessing the human microbiome in epidemiologic studies. Ann Epidemiol. 2016 May;26(5):311–21. [PubMed: 27180112]
3. Vandeputte D, Tito RY, Vanleeuwen R, Falony G, Raes J. Practical considerations for large-scale gut microbiome studies. FEMS Microbiol Rev. 2017 Aug 1;41(Supp_1):S154–67. [PubMed: 28830090]
4. Ahn J, Sinha R, Pei Z, Dominianni C, Wu J, Shi J, et al. Human gut microbiome and risk for colorectal cancer. J Natl Cancer Inst. 2013 Dec 18;105(24):1907–11. [PubMed: 24316595]
5. Mascitti M, Togni L, Troiano G, Caponio VCA, Gissi DB, Montebugnoli L, et al. Beyond Head and Neck Cancer: The Relationship Between Oral Microbiota and Tumour Development in Distant Organs. Front Cell Infect Microbiol. 2019;9:232. [PubMed: 31297343]
6. Sinha R, Chen J, Amir A, Vogtmann E, Shi J, Inman KS, et al. Collecting Fecal Samples for Microbiome Analyses in Epidemiology Studies. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2016 Feb;25(2):407–16.
7. Sinha R, Abnet CC, White O, Knight R, Huttenhower C. The microbiome quality control project: baseline study design and future directions. Genome Biol. 2015;16:276. [PubMed: 26653756]
8. Vogtmann E, Chen J, Kibriya MG, Chen Y, Islam T, Eunes M, et al. Comparison of Fecal Collection Methods for Microbiota Studies in Bangladesh. Appl Environ Microbiol. 2017 May 15;83(10):e00361–17. [PubMed: 28258145]
9. Vogtmann E, Chen J, Amir A, Shi J, Abnet CC, Nelson H, et al. Comparison of Collection Methods for Fecal Samples in Microbiome Studies. Am J Epidemiol. 2017 Jan 15;185(2):115–23. [PubMed: 27986704]
10. Wang Z, Zolnik CP, Qiu Y, Usyk M, Wang T, Strickler HD, et al. Comparison of Fecal Collection Methods for Microbiome and Metabolomics Studies. Front Cell Infect Microbiol. 2018;8:301. [PubMed: 30234027]
11. Byrd DA, Chen J, Vogtmann E, Hullings A, Song SJ, Amir A, et al. Reproducibility, stability, and accuracy of microbial profiles by fecal sample collection method in three distinct populations. PloS One. 2019;14(11):e0224757.
12. Zouiouich S, Mariadassou M, Rué O, Vogtmann E, Huybrechts I, Severi G, et al. Comparison of Fecal Sample Collection Methods for Microbial Analysis Embedded within Colorectal Cancer Screening Programs. Cancer Epidemiol Biomarkers Prev. 2022 Feb 7;31(2):305–14. [PubMed: 34782392]
13. Yano Y, Hua X, Wan Y, Suman S, Zhu B, Dagnall CL, et al. Comparison of Oral Microbiota Collected Using Multiple Methods and Recommendations for New Epidemiologic Studies. mSystems. 2020 Jul 7;5(4):e00156–20. [PubMed: 32636335]
14. Luo T, Srinivasan U, Ramadugu K, Shedden KA, Neiswanger K, Trumble E, et al. Effects of Specimen Collection Methodologies and Storage Conditions on the Short-Term Stability of Oral Microbiome Taxonomy. Appl Environ Microbiol. 2016 Sep 15;82(18):5519–29. [PubMed: 27371581]
15. Vogtmann E, Chen J, Kibriya MG, Amir A, Shi J, Chen Y, et al. Comparison of Oral Collection Methods for Studies of Microbiota. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2019 Jan;28(1):137–43.
16. Vogtmann E, Hua X, Zhou L, Wan Y, Suman S, Zhu B, et al. Temporal Variability of Oral Microbiota over 10 Months and the Implications for Future Epidemiologic Studies. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2018 May;27(5):594–600.
17. Gudra D, Shoaie S, Fridmanis D, Klovins J, Wefer H, Silamikelis I, et al. A widely used sampling device in colorectal cancer screening programmes allows for large-scale microbiome studies. Gut. 2019 Sep 1;68(9):1723–5. [PubMed: 30242040]

18. Kia E, Wagner Mackenzie B, Middleton D, Lau A, Waite DW, Lewis G, et al. Integrity of the Human Faecal Microbiota following Long-Term Sample Storage. PloS One. 2016;11(10):e0163666.
19. Tap J, Cools-Portier S, Pavan S, Druesne A, Öhman L, Törnblom H, et al. Effects of the long-term storage of human fecal microbiota samples collected in RNAlater. Sci Rep. 2019 Jan 24;9(1):601. [PubMed: 30679604]
20. Shaw AG, Sim K, Powell E, Cornwell E, Cramer T, McClure ZE, et al. Latitude in sample handling and storage for infant faecal microbiota studies: the elephant in the room? Microbiome. 2016 Jul 30;4(1):40. [PubMed: 27473284]
21. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 2012 Aug;6(8):1621–4. [PubMed: 22402401]
22. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019 Aug;37(8):852–7. [PubMed: 31341288]
23. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods. 2016 Jul;13(7):581–3. [PubMed: 27214047]
24. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013 Jan;41(Database issue):D590–6. [PubMed: 23193283]
25. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PloS One. 2013;8(4):e61217.
26. Dixon P. VEGAN, a package of R functions for community ecology. J Veg Sci. 2003;14(6):927–30.
27. Chen J, Bittinger K, Charlson ES, Hoffmann C, Lewis J, Wu GD, et al. Associating microbiome composition with environmental covariates using generalized UniFrac distances. Bioinforma Oxf Engl. 2012 Aug 15;28(16):2106–13.
28. Carroll IM, Ringel-Kulka T, Siddle JP, Klaenhammer TR, Ringel Y. Characterization of the fecal microbiota using high-throughput sequencing reveals a stable microbial community during storage. PloS One. 2012;7(10):e46953.
29. Fouhy F, Deane J, Rea MC, O'Sullivan Ó, Ross RP, O'Callaghan G, et al. The Effects of Freezing on Faecal Microbiota as Determined Using MiSeq Sequencing and Culture-Based Investigations. PLOS ONE. 2015 Mar 6;10(3):e0119355.
30. Lauber CL, Zhou N, Gordon JI, Knight R, Fierer N. Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples. FEMS Microbiol Lett. 2010 Jun;307(1):80–6. [PubMed: 20412303]
31. Bahl MI, Bergström A, Licht TR. Freezing fecal samples prior to DNA extraction affects the Firmicutes to Bacteroidetes ratio determined by downstream quantitative PCR analysis. FEMS Microbiol Lett. 2012 Apr;329(2):193–7. [PubMed: 22325006]
32. Sinha R, Goedert JJ, Vogtmann E, Hua X, Porras C, Hayes R, et al. Quantification of Human Microbiome Stability Over 6 Months: Implications for Epidemiologic Studies. Am J Epidemiol. 2018 Jun;187(6):1282–90. [PubMed: 29608646]

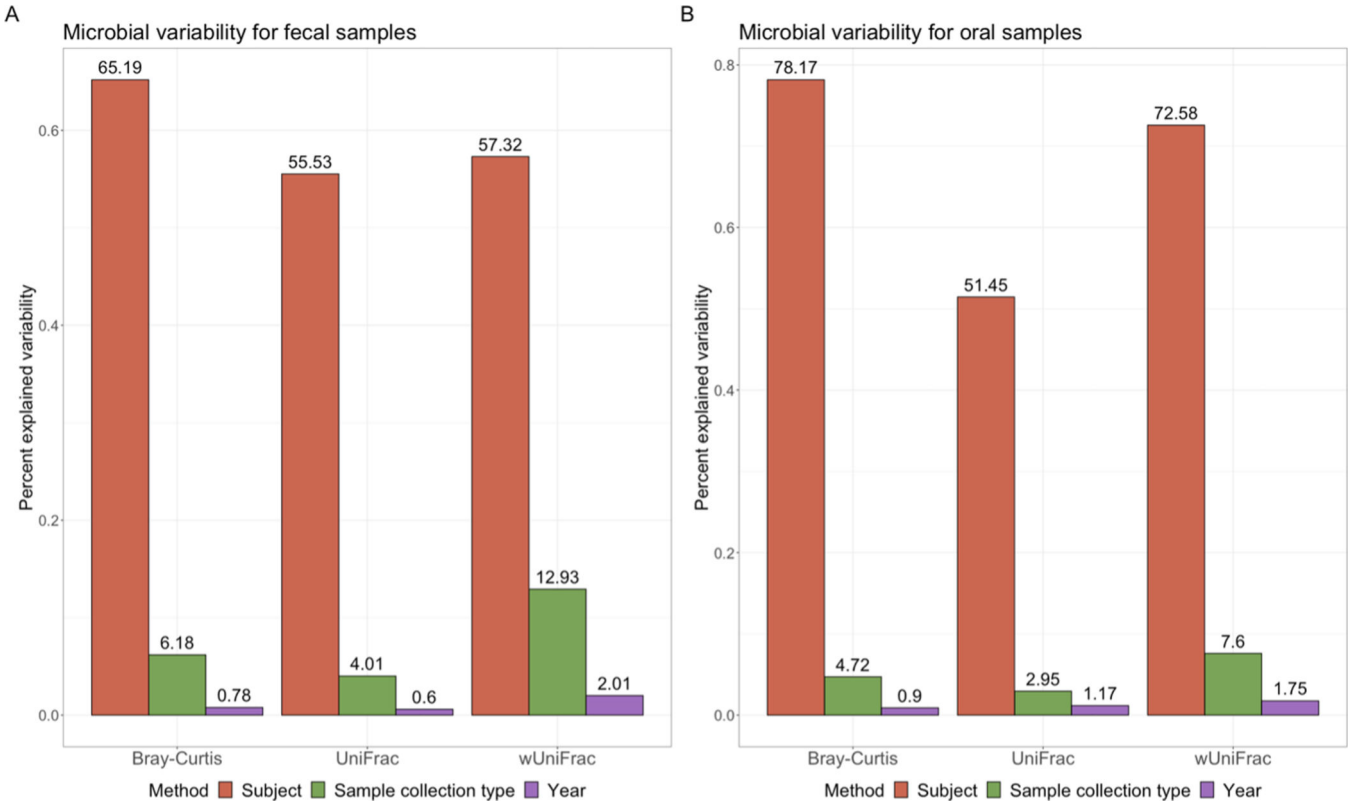

Figure 1.

Percent variability explained by subject, collection type, and storage time using a distance-based coefficient of determination (R^2) for beta diversity estimates from Bray-Curtis, weighted and unweighted UniFrac distance matrices for (A) fecal and (B) oral samples.

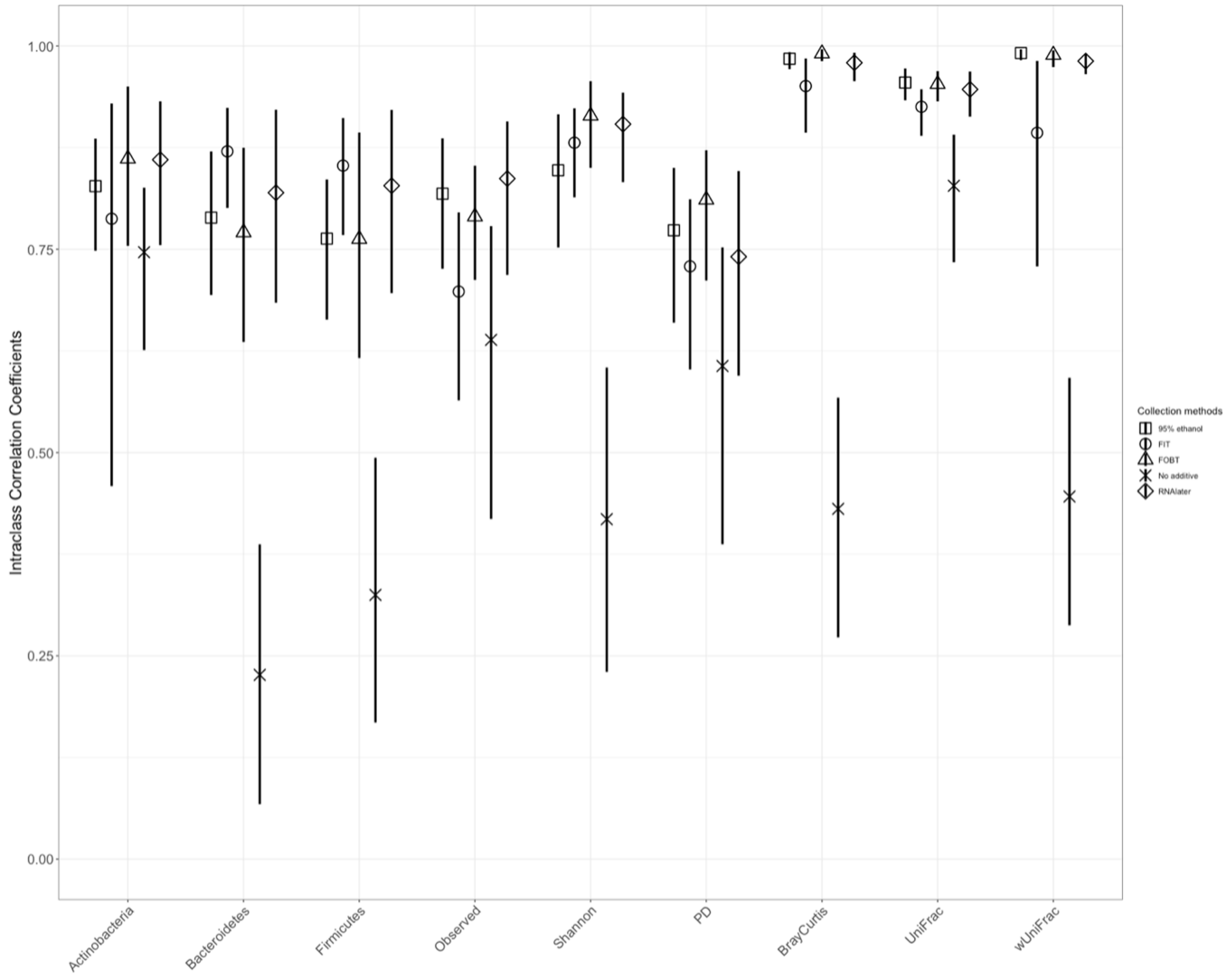

Figure 2.

ICCs based on random effects models for microbiome stability comparing fecal samples extracted upon receipt (baseline), after a year stored at −80°C (year 1), and after two years stored at −80°C (year 2) for five fecal sample collection methods.

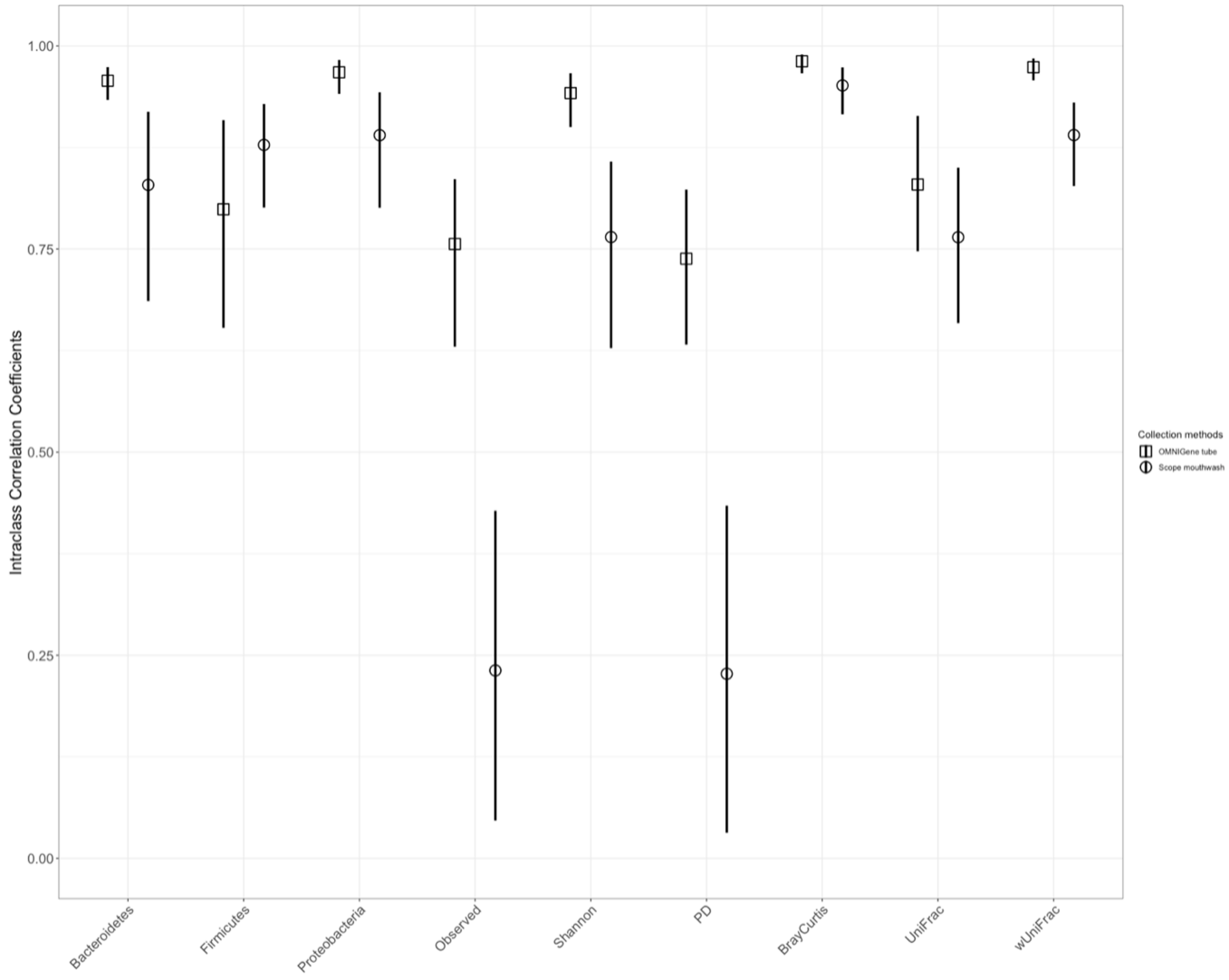

Figure 3.

ICCs based on random effects models for microbiome stability comparing oral samples extracted upon receipt (baseline), after a year stored at −80°C (year 1), and after two years stored at −80°C (year 2) for two oral sample collection methods.

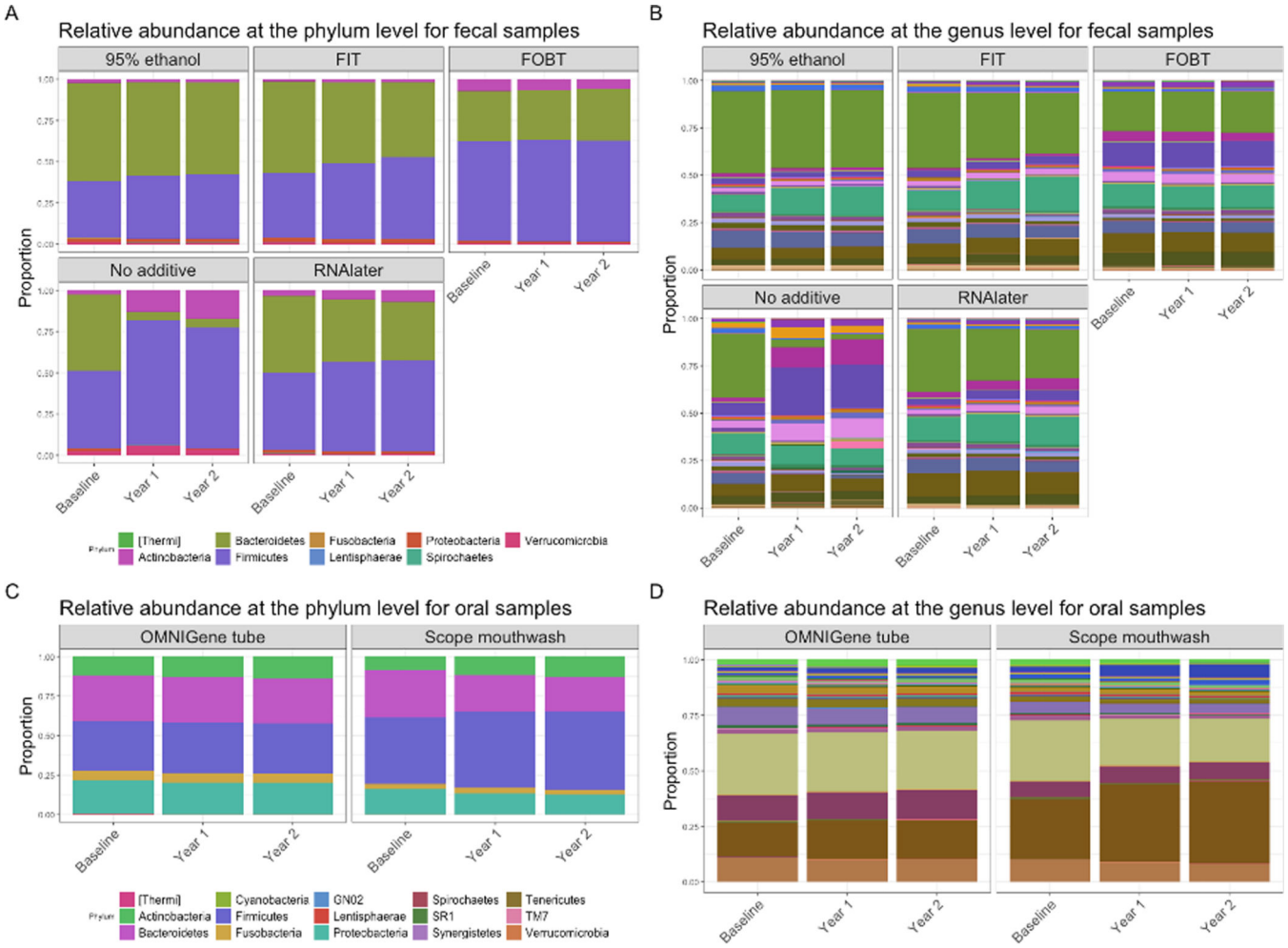

Figure 4.

Relative abundance of taxa at the phylum and genus levels, which were present in at least 10% of samples with relative abundance of >0%, at baseline, and after one and two years of storage at −80°C, for (A, B) fecal (n=51) and (C, D) oral samples (n=51).

Table 1.

Collection methods for fecal and oral samples, and number of aliquots for microbiome analyses in 51 healthy volunteers. For fecal samples, two aliquots of samples collected with no additive, 95% ethanol, RNAlater and FIT, and three aliquots of samples collected with FOBT were frozen immediately at −80°C (Day 0). In addition, two aliquots of samples collected with 95% ethanol, RNAlater and FIT, and three aliquots of samples collected with FOBT remained at room temperature for 96 hours (Day 4). At the end of the four days, the remaining aliquots were frozen at −80°C. For oral samples, one aliquot of samples collected with OMNIgene tube and Scope mouthwash was frozen immediately at −80°C (Day 0) and one aliquot of Scope mouthwash remained at room temperature for 96 hours (Day 4) before being frozen at −80°C.

	Baseline						Year 1						Year 2				
Collection method	Day 0			Day 4			Day 0			Day 4			Day 0				
	N	Missing	Total	N	Missing	Total	N	Missing	Total	N	Missing	Total	N	Missing	Total	N	M
95% ethanol	2×51	0	102	2×51	0	102	2×51	15	87	2×51	16	86	2×51	15	87	2×51	
FIT	2×51	0	102	2×51	2	100	2×51	15	87	2×51	16	86	2×51	15	87	2×51	
FOBT	3×51	0	153	3×51	0	153	3×51	16	137	3×51	29	124	3×51	16	137	3×51	
No additive	2×51	0	102	0	0	0	2×51	16	86	0	0	0	2×51	16	86	0	
RNAlater	2×51	1	101	2×51	1	101	2×51	15	87	2×51	14	88	2×51	15	87	2×51	
Scope mouthwash	51	0	51	51	0	51	51	0	51	51	0	51	51	0	51	51	
OMNIGene tube	51	0	51	0	0	0	51	0	51	0	0	0	51	0	51	0	