# R Final Project

Erica Faucher

**A comparison of bulk stool and FTA card stool sampling methods in terms of bacterial diversity**

Abstract

This project examines the human gut microbiome, which is influenced by a person's diet, delivery mode, geography, and food supply, to name a few. FTA cards and bulk stool samples are two methods used to analyze the bacterial make up of the microbiome; however, it has not been tested whether FTA cards have an effect on bacterial diversity in samples. The stool samples of seven participants were taken using both methods, and then sequenced using 16S rRNA gene sequencing. The data was analyzed in R in terms of bacterial diversity by phyla. Multiple graphs were created to show the different sampling methods' bacterial makeup, as well as statistical results. It was found that there is a statistical difference between the FTA cards and bulk stool sample methods in terms of bacterial diversity.

Introduction

The human gut microbiome contains a multitude of different microbes, especially bacteria. The diversity of gut bacteria varies from person to person, and is affected by many factors, such as delivery mode (whether a person was born via cesarian section or vaginal birth), infant feeding method (breast milk or formula), diet, geography, food supply, stress, and antibiotics (Cresci and Bawden, 2016). Stool samples are typically used for analyzing the gut microbiome with gene sequencing. Stool samples are typically collected with the bulk stool sample method, which requires immediate freezing. However, this method is not feasible for hard-to-reach populations, such as people who live in more rural areas. Bulk stool also cannot be mailed, since the samples must be processed within 24 hours, or as soon as possible. Fast technology for analysis of nucleic acids (FTA) cards are a newer and alternative method. A chemical on the cards fixes the bacteria and allows them to be stored at room temperature. FTA cards have the potential to make stool sample collection easier for researchers to process, as well as allow for more gut microbiome research to be done in harder to reach populations.

The efficacy of FTA cards against bulk stool sample collection, in terms of bacterial diversity and richness, has not been conducted. Bacterial diversity in the gut is important because of its relationship with one's overall health. Lower bacterial diversity in the gut microbiome, as well

as an increase of pro-inflammatory bacteria, is associated with Alzheimer's disease. The gut microbiome may have a role in the regulating neuroinflammatory disorders, such as multiple sclerosis and Parkinson's disease. Changes to the gut microbiome can affect the progression of illnesses in the body, as well as neurotransmitters like serotonin (Menezes and Shah, 2024).

This project focuses on analyzing seven stool samples that were taken with both the bulk stool and FTA card sampling methods. It was predicted that there will be no difference in bacterial diversity between the two stool sampling methods.

## Methods

The gold standard bulk stool sample collection method was compared to the FTA card collection method, and the microbial differences between the two methods was analyzed. In the summer of 2023, seven volunteers donated their stool samples, which were collected with both the bulk stool and FTA card sample methods. The bulk stool samples were homogenized, made into three aliquots, and stored at -80$^{o}$C within 24 hours of collection. The person who produced the stool sample smeared a small amount of the sample into the four circles on the FTA card, which immediately fixed the sample and allowed it to be stable at room temperature for a long period of time. 16S rRNA gene sequencing was then conducted on all the samples.

The data was transferred into an Excel spreadsheet, which was uploaded into R. The data was then relabeled, and the taxonomy column was split so that there was one column for each taxonomic level. Upon creating graphs of the data, it was determined that the data set was quite large, and that a smaller subset of the data would be analyzed in the context of this paper. Thus, only the phyla of bacteria in the samples was compared.

One bar graph was created for the average bacteria phyla in all of the bulk samples, and another bar graph for the average number of bacteria in all the FTA card samples. A third bar graph was created for the average bacteria in all samples, FTA and bulk. A Shannon Diversity Index was run to examine the difference in diversity for the two sample methods for each phylum. A second Shannon test was done to compare the diversity between the sampling methods, and a box plot was produced to visualize the results. A Wilcox test was then run to test the statistical significance of the diversity differences between the two methods. The code for all of these graphs and analyses can be located on Git Hub at https://github.com/emf2003/R_Final_Project.git.

## Results

Figure 1 (below) depicts a bar graph of the average number of bacteria in each phyla for all seven participants' bulk stool samples. The phyla with the highest averages across all samples include Firmicutes and Bacteroidota. Sample 1 has the highest average of bacteria in the Actinobacteriota and Verrucomicrobiota phyla. Sample 1 also has the lowest average of Bacteroidota bacteria. Sample 6 has the lowest average number of bacteria in the Firmicutes phylum.

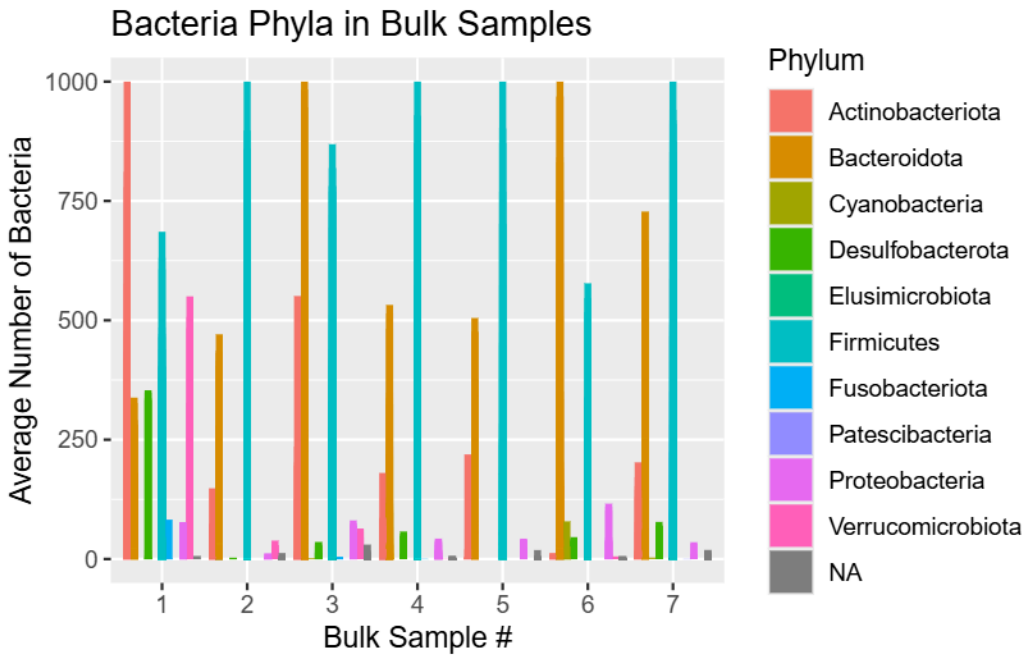


Figure 1: The average number of bacteria in each phyla in all bulk stool samples.

Figure 2 (below) shows the average number of bacteria in each phyla for the FTA card samples across all participants. The phyla with the highest averages across all samples were Firmicutes and Bacteroidota. Sample 1 had the highest average number of Actinobacteriota , Verrucomicrobiota, and Desulfobacterota, but the lowest average number of Bacteroidota. Sample 6 had the lowest average number of bacteria in the Firmicutes phylum.

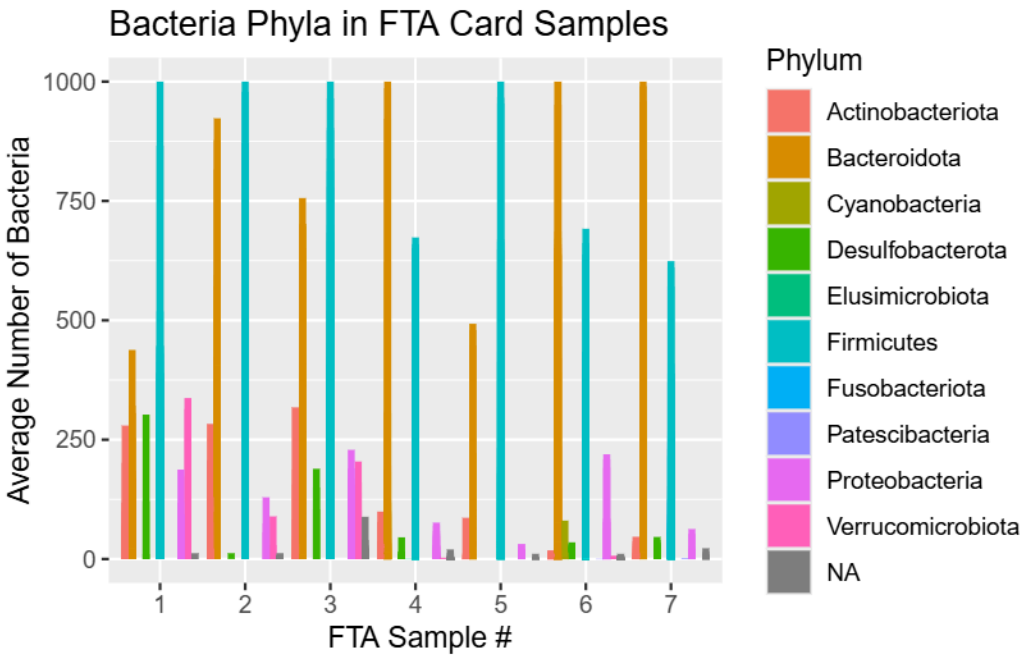


Figure 2: Average number of bacteria, by phyla, in all FTA card stool samples.

Figure 3 (below) shows the average number of bacteria in the bulk and FTA samples for all participants. The average number of bacteria in the bulk samples was higher than the average number of bacteria in the FTA card stool samples, across all samples. The FTA cards for samples 1, 2, and 3 have the lowest average numbers of bacteria. FTA sample 6 has the highest average number of bacteria compared to the other FTA card samples. Bulk sample 2 has the highest average number of bacteria overall.

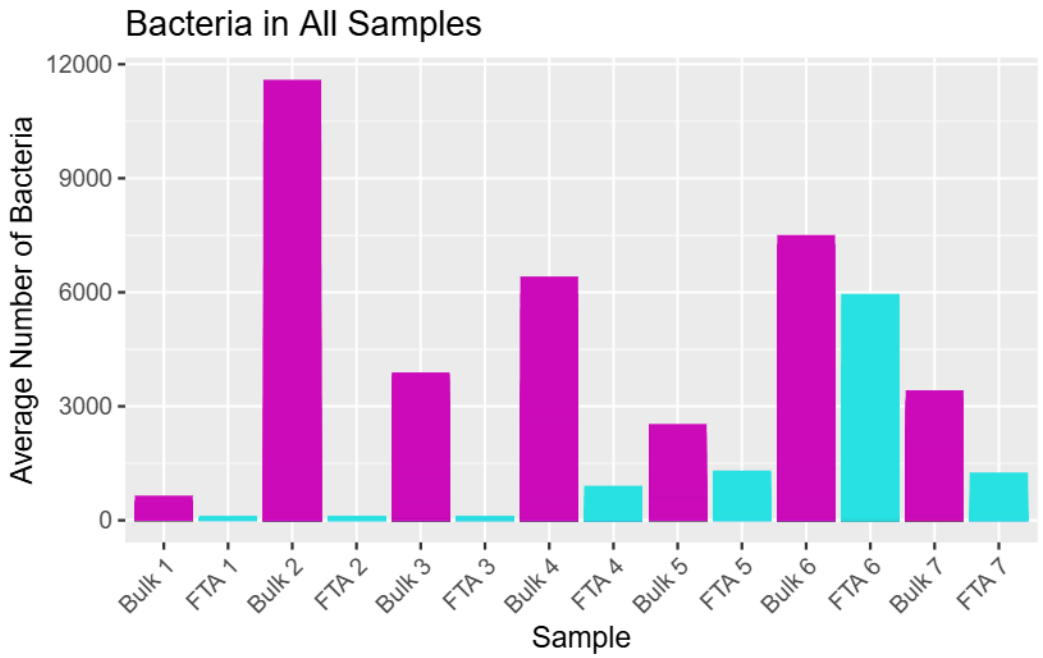


Figure 3: Average number of bacteria in FTA and bulk stool samples for all participants.

A Shannon Diversity Index was run and was organized by phylum. Phyla Patescibacteria and Cyanobacteria had the lowest values compared to the other phyla. The bulk method shows higher values than the FTA values for all phyla except Patescibacteria and Verrucomicrobiota.

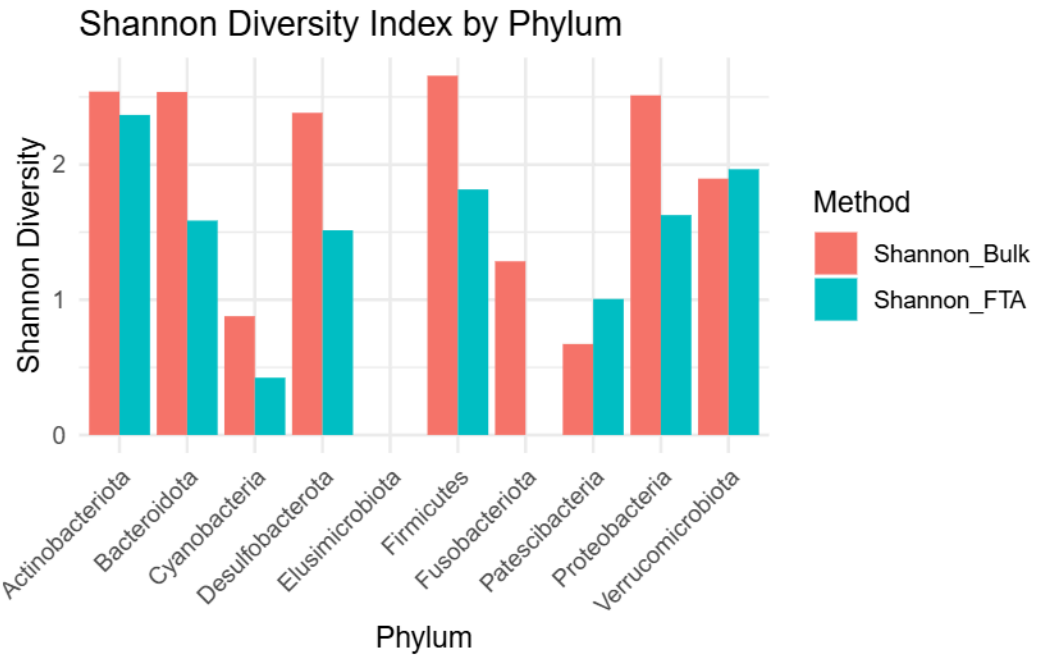


Figure 4: Shannon Diversity Index by Phylum for both sampling methods.

Figure 5 (below) shows a box plot of the results of the second Shannon Diversity Index calculated on the diversity between the two sampling methods. The bulk sampling data has one outlier, shown in the graph as a black dot. The bulk method has a higher median value than the FTA card sampling method. The FTA card method has a larger interquartile range than the bulk sampling method.

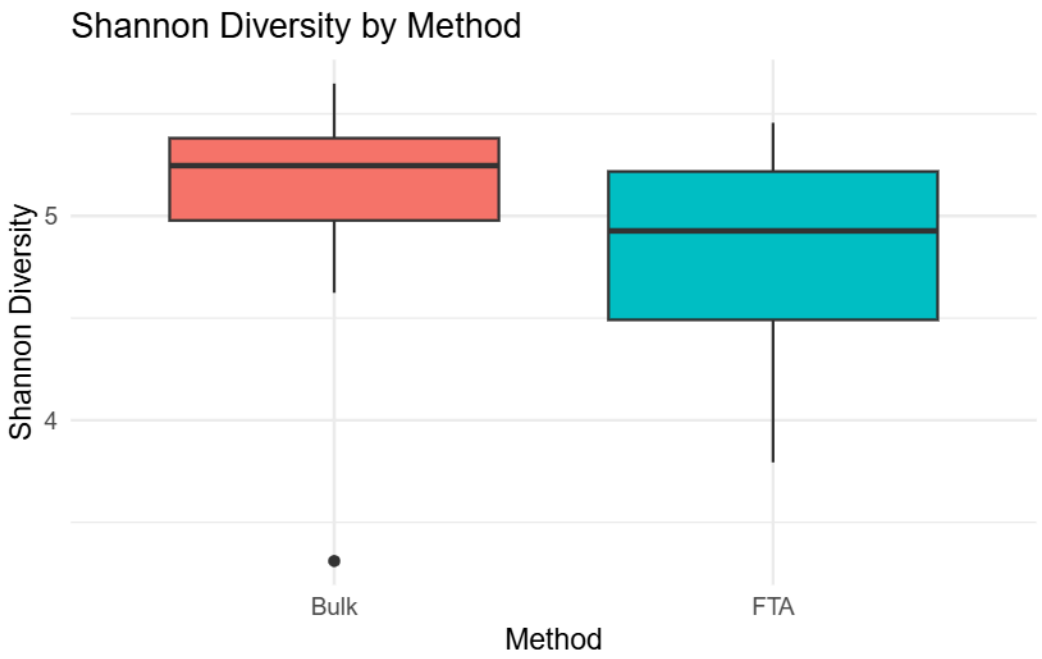


Figure 5: Box plot depicting the Shannon Diversity Index to compare the bacterial diversity for both methods.

The p-value for the Wilcox test was 0.03359.

Discussion

This project focused on examining the diversity of phyla between the two methods of stool sampling, FTA card and bulk stool. Figures 1 and 2 show that the most common phyla that make up the gut microbiome are Firmicutes and Bacteroidota. However, the FTA samples had higher averages of bacteria belonging to the phyla Verrucomicrobiota and Proteobacteria than in the bulk stool samples. The bulk samples tended to have higher averages of Actinobacteriota than the FTA card samples. As shown in Figure 3, The average number of bacteria in the bulk samples was higher across all participants in the study, indicating that stool sample collection yields a greater number of bacteria than in FTA card sample collection.

The first Shannon Diversity Index (Figure 4) results show that there was overall more diversity in the bulk stool samples than the FTA card samples across all phyla (except for Patescibacteria and Verrucomicrobiota). The second Shannon Diversity Index was run to compare the bacterial diversity between the two sampling methods. The box plot in Figure 5 (the results of the second Shannon, comparing the methods) shows that the bulk sample method has a higher median than the FTA card method, indicating that the bulk sample method has a higher bacterial diversity. The p-value for the Wilcox test was 0.03359, which is less than 0.05, meaning that there is a statistically significant difference between the two sampling methods. There is a

difference in the diversity of phyla between the FTA and bulk stool sampling methods. This goes against the hypothesis, that there would be no difference in diversity between the sampling methods. This means that researchers must take into account this difference when choosing whether to gather stool samples using the FTA card or bulk stool sampling method.

It is important to note that this study has a small sample size (7), which may have an impact on the results. However, smaller sample sizes for studies is not unheard of. A study on sampling methods measuring HIV viral load had 51 participants (Jaumdally *et al,* 2017), and a study on pharmacokinetic properies of insulin icodec had a sample size of 46 (Pieber *et al,* 2023). One study on the effect of *Bifidobacterium breve* M-16V and mood in humans had a sample size of just 30 participants (Mutoh *et al,* 2024).

Another factor that can impact the data is that the participants are aliquoting their samples themselves on FTA cards. It is up to the person who produces the stool sample to accurately swab the sample into the circles on the card. If done incorrectly (ie not swabbing enough sample, or swabbing too much/leaving large pieces of stool), this can affect the number of bacteria in the sample. This can also affect the 16S results. In a bulk stool sample, the participant is still responsible for producing a clean sample (ie a sample not contaminated with urine), but is not responsible for aliquoting the sample.

References

Cresci, G. A., & Bawden, E. (2015). The Gut Microbiome: What we do and don't know. Nutrition in Clinical Practice : Official Publication of the American Society for Parenteral and Enteral Nutrition, 30(6), 734. https://doi.org/10.1177/0884533615609899

Jaumdally, S. Z., Jones, H. E., Hoover, D. R., Gamieldien, H., Kriek, J.-M., Langwenya, N., Myer, L., Passmore, J.-A. S., & Todd, C. S. (2017). Comparison of sampling methods to measure HIV RNA viral load in female genital tract secretions. American Journal of Reproductive Immunology (New York, N.Y.: 1989), 77(3). https://doi.org/10.1111/aji.12619

Menezes, A. A., & Shah, Z. A. (2024). A Review of the Consequences of Gut Microbiota in Neurodegenerative Disorders and Aging. Brain Sciences, 14(12), Article 12. https://doi.org/10.3390/brainsci14121224

Mutoh, N., Moriya, M., Xu, C., Kato, K., Arai, S., Iwabuchi, N., Tanaka, M., Jinghua, Y., Itamura, R., Sakatani, K., & Warisawa, S. (2024). Bifidobacterium breve M-16V regulates the autonomic nervous system via the intestinal environment: A double-blind, placebo-controlled study. Behavioural Brain Research, 460, 114820. https://doi.org/10.1016/j.bbr.2023.114820

Pieber, T. R., Asong, M., Fluhr, G., Höller, V., Kristensen, N. R., Larsen, J. H., Ribel-Madsen, R., Svehlikova, E., Vinther, S., Voortman, M., & Haahr, H. (2023). Pharmacokinetic and pharmacodynamic properties of once-weekly insulin icodec in individuals with type 2 diabetes. Diabetes, Obesity & Metabolism, 25(12), 3716–3723. https://doi.org/10.1111/dom.15266

R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.