

Team 2: Adam Abdirahman Hassan, Farbod Nematifar, Timothy Bernas, Trushaan Bundhoo, Young Ha Jin

Determining the effect of Hunter Gatherer Diets and Westernized Diets on the Gut Microbiome of Adults

Introduction and Background

The human gut microbiome is known for playing a major role in promoting normal bodily function, affecting metabolism, defending against pathogens, and stimulating the immune system (1). Its role in maintaining homeostasis and regulating immune response make it a pivotal aspect of microbiological research. Previous research has shown diet to be a key factor in determining the composition of the human gut microbiome (2). Our proposed project will analyze microbiome data from two different populations with different lifestyles and diets, with the goal of determining how different diets may contribute to compositional differences in the gut microbiome.

Previous studies noted that the gut microbiome of hunter gatherer communities was much higher in alpha diversity metrics compared to westernized diets. Shannon and Bray-Curtis diversity measures alluded to a very different, and more diverse gut microbiome in an Amazonian hunter-gatherer tribe compared to the US (3). Additional studies have indicated that the reduction in alpha diversity that is seen within Western diets have been associated with various conditions such as obesity and ALS (4,5).

We will be examining data from two separately collected datasets, one containing microbiome data from the Hadza hunter-gatherer tribe in Tanzania, and the other containing microbiome data collected from Colombian individuals. The Colombian data set acts as a good model for a population with a Westernized diet as their biomes have been found to resemble that of American and European populations (6).

The samples in the Tanzania dataset were collected and analyzed with the goal of determining how the population's microbiome composition varied throughout the year, as diet and hunting patterns changed based on the weather (7). Human samples were taken from hand swabs and stool samples, and other samples were taken from honey, animal stomach swabs, or animal feces. The researchers concluded that the gut microbiome of the Hadza were similar to a Western microbiome during the inactive wet season but were far more dissimilar during the hunting dry season (7). The Colombian dataset contains data from fecal samples collected from Colombian adults, whose diet shifted towards a more westernized diet, with the goal of studying if this shift was linked to changes in the gut microbiome composition (8). The researchers found that the microbiome samples of this population reflected a gradient between what is expected from a traditional Colombian diet versus a westernized diet (8).

Given that the Colombia dataset contains only human fecal samples, other sample types in the Tanzania dataset are not of interest to us and will not be included in our analyses. Other common metadata categories in our combined dataset in addition to location will be age and sex. We also aim to assess whether these factors can also be sub-correlated to shifts in microbial composition between the two populations. Age is known to correlate with incidences of metabolic diseases (9), which may be associated with perturbations in gut microbiome composition. Sex has also been linked to differences in the abundance of genera and species present in the gut microbiome (10). Additionally, within the Tanzania data set we removed any fecal samples that came from individuals that were below the age of 18 to ensure that only adults were being compared between the 2 data sets. Moreover, we removed samples in the Columbia dataset that were from smokers (8), as this could be a significant confounding variable due to drops in intestinal pH and increases in oxidative stress found in smokers (11).

Our project is dependent on the fact that there are clear and distinct differences in the microbial species that are present within the gut microbiome of a hunter-gatherer diet vs a Western diet. Specifically, we expect there to be a greater number of unique species present within the gut microbiome of those that have a hunter gatherer diet since it is widely considered to be prevent but microbiome dysbiosis when compared to a Western diet, which tends to promote gut microbiome dysbiosis.

Research Objective

Research Question: What are the primary differences and similarities between the gut microbiome of those that ascribe to a Western diet vs a hunter gatherer diet?

Previous literature has only analyzed each of these diet types individually through a metatextual analysis but never together as a joined data set. Therefore, there is a gap within the literature regarding the direct analysis between 2 separate data sets one referring to a hunter gatherer diet (Tanzania) and the other a Western diet (Colombia) (7,8). Consequently, we simply wanted to compare the data sets and see the differences with respect to gut microbial composition in terms of both alpha diversity and beta diversity. Additionally, we want to get an understanding of the specific microbial taxa that the two diet types have in common and the taxa that are unique to each. Finally, we will determine the exact functional capability of the gut microbiomes for each individual diet type as another point of comparison. Once all these metrics have been determined we will then look to see if the results that we observed can be uniquely understood through sex and age, and if new relationships between the populations arise using these diversity metrics. Ultimately, the purpose of our research is to get a better understanding on how diet shapes our gut microbiome and to see if it matches with findings discovered by other researchers.

Experimental Aims

Aim 1: Determine if gut microbial composition differs between adults in Tanzania (hunter gatherer diet) and adults in Colombia (Westernized diet) using alpha- and beta-diversity measures.

Our first aim is to combine the Tanzania dataset with the Colombia dataset to answer whether a hunter-gatherer lifestyle microbiome is significantly different from a Westernized diet microbiome. This will be quantified by analyzing both Shannon's and Weighted-Unifrac diversity metrics in R, along with a permanova statistical analysis. We know that a hunter-gatherer lifestyle is associated with higher alpha diversity, unlike Western diets. We also know that the different lifestyles of the two groups would most likely lead to vastly different gut microbial communities. Therefore, we predict that alpha diversity scores will be higher in the Tanzania dataset, and we also predict a high beta diversity score between the Tanzania and Colombia groups.

Aim 2: Determine the taxa that are common between the 2 groups (Tanzania and Colombia) as well as the taxa that are unique to the 2 individual groups.

Our second aim is to show how similar or dissimilar the gut microbiome taxonomic composition is between the two datasets, further building upon the goals outlined in aim 1. Specifically, we expect that the amount of overlap between the two diet types, in terms of taxa, would be very small since the 2 populations radically differ in the foods that they consume, detailed in the papers that analyzed the Columbia and Tanzania data sets respectively (7,8). Moreover, there should be a greater number of unique taxa within the Tanzania data set since it has been reported within the literature that hunter gatherers have a greater degree of microbial diversity than Westernized diet types (3). Thus, the primary form of analysis that will be carried out is a core microbiome analysis comparing the Tanzania and Colombia samples in terms of taxa/species. Consequently, to visualize the results of the core microbiome analysis a Venn Diagram will be generated to show unique and overlapping species between the datasets. Afterwards, to determine if taxa of interest (unique or are found in common between the 2 data sets) are significant, an indicator taxa analysis will be performed.

Aim 3: - Determine the differences and or similarities in the functional capabilities of the gut microbial communities between adults from Tanzania versus adults from Colombia.

Expanding upon our second aim, we want to look more closely at exactly what these differences and similarities entail with regards to the functional diversity present in these gut microbiomes. With vast differences in diet, we can expect to see vast differences in the metabolic pathways that are present in the microbiome. On the other hand, we expect to see only a little, if any, overlap in metabolic pathways present between the two diets due to the differences in their overall compositions. Western diets have a much higher proportion of highly processed food and carbohydrates compared to hunter gatherer diets, which primarily consist of animal products supplemented by foraging. While both diets contain a significant proportion of red meat, the quality and quantity present in both diets are likely different enough to see little to no overlap in metabolic pathways present regarding red meat consumption (12, 13). There are multiple studies highlighting the metabolic pathway composition of Western diets and hunter gatherer diets separately but few if any that directly compare them to each other. The primary form of analysis will utilize PICRUST2, an R package developed to predict the presence and abundance of gene families associated with higher level functions, such as metabolic pathways, based on amplicon sequence variants found in the provided samples (14). We will use the package to predict the presence of higher-level pathways that are unique to either the Tanzanian adults or Colombian adults as well as any higher-level pathways that may be present in both of their gut microbiomes.

Aim 4: Determine if age and sex have a significant effect on the composition of the gut microbiome of the 2 data sets (Tanzania and Colombia).

In addition to exploring general compositional differences between the microbiomes of individuals with westernized diets or hunter-gatherer diets, further segmented analysis will occur across age and sex. Within the Hunter-Gatherer data set, Weighted-Unifrac analysis showed that there was significant difference between sexes that could be attributed to sex-based divisions in labor (15). For example, women mostly forage for plants and spend most of their time in camp. Thus, the increased *Treponema* within their biomes could be attributed to diets rich in plant fiber. The data set demonstrated no compositional differences amongst age groups.

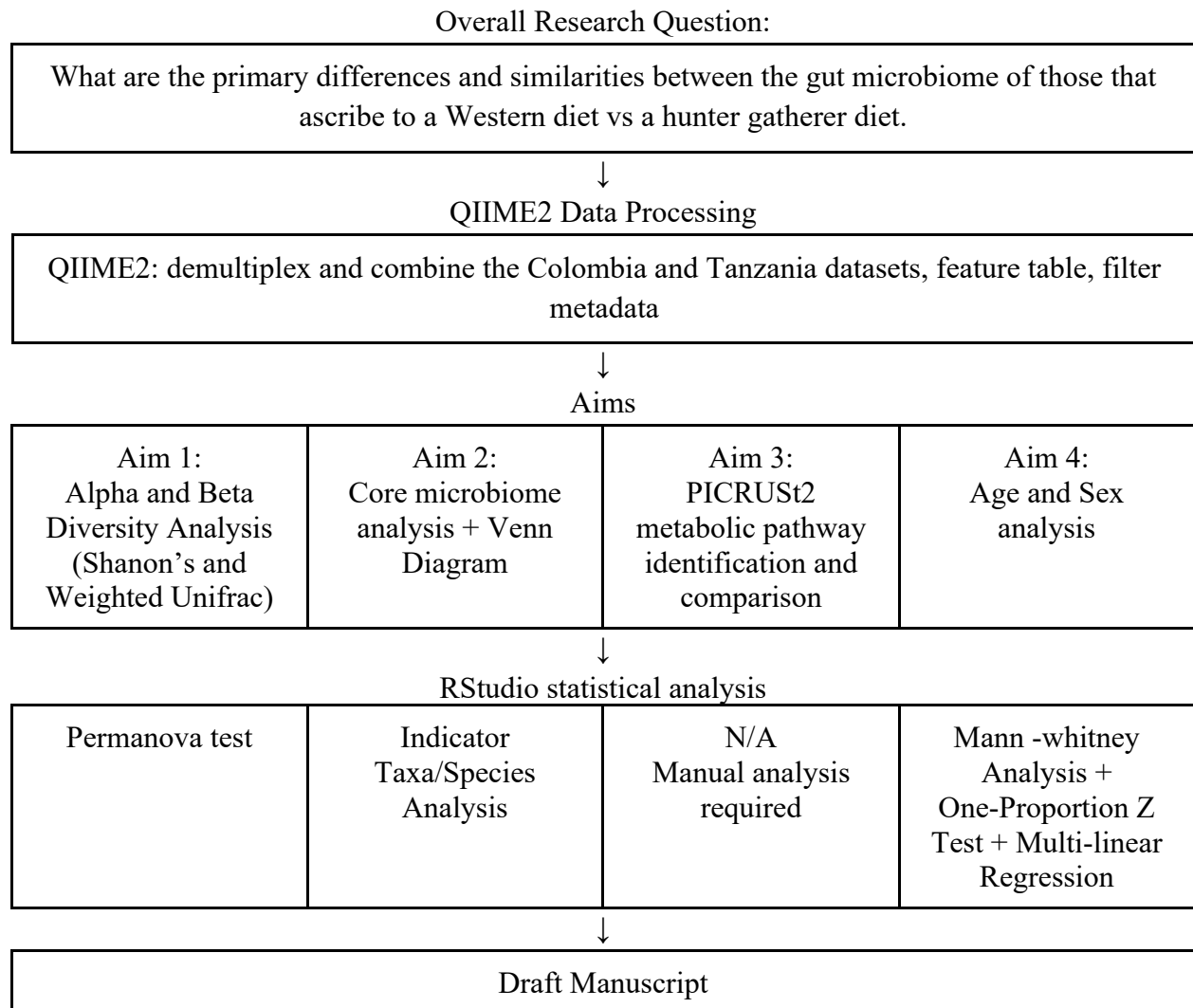
The Colombian data was found to have greater alpha diversity amongst women in comparison to men. There was also a strong positive correlation between age and alpha diversity that plateaus after approximately 40 years (6).

These findings provide good foundational logic for conducting cross-population analysis with hopes of identifying sex and age-based differences between the populations. To compare the two data sets, initial comparisons of the average age of the two data sets must be done, as well as a Mann-Whitney U test to examine the age-tendencies of the two data sets. Furthermore, to ensure a similar proportion of men to women, a one-proportion z-test can be done. Should the

populations not be deemed proportional in terms of age and sex samples can be taken out to make them comparable.

Once the sets have been deemed proportional, a linear regression will be conducted using specific age groups or sex against the selected diversity metrics. The initial age split will be 18-40, and 40 onward because of the correlation between age and alpha diversity until 40 years. Based on the results, further sub-grouping can be done to assess differences in diversity.

Flow Chart

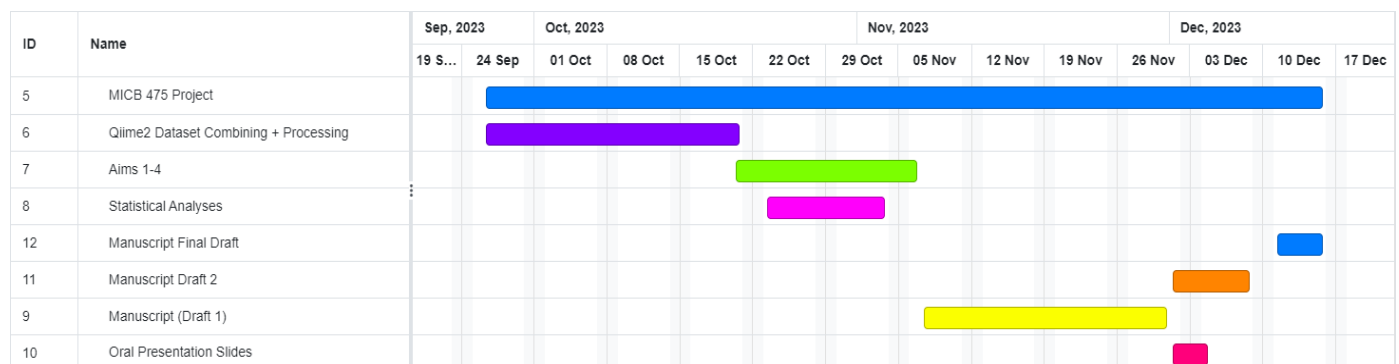


Approach

Aim 1: Does the gut microbial composition differ between adults in Tanzania (hunter gatherer diet) and adults in Colombia (Westernized diet)	Aim 2: Determine the taxa that are common between the 2 groups (Tanzania and Colombia) as well as the taxa that are unique to the 2 individual groups.	Aim 3: What are the differences and or similarities in the functional capabilities of the gut microbial communities between adults from Tanzania vs Colombia.	Aim 4: Does age and sex have a significant effect on the composition of the gut microbiome of the 2 data sets (Tanzania and Colombia)
A1-1: Generation of new meta data and manifest table joining the Colombia and Tanzania datasets based on similar metadata columns in preparation for QIIME analysis. (Joined data on the basis of sex, age, and sample type (fecal samples))	A2-1: Generation of new meta data and manifest table joining the Colombia and Tanzania datasets based on similar metadata columns in preparation for QIIME analysis. (Joined data on the basis of sex, age, and sample type (fecal samples))	A3-1: Generation of new meta data and manifest table joining the Colombia and Tanzania datasets based on similar metadata columns in preparation for QIIME analysis. (Joined data on the basis of sex, age, and sample type (fecal samples))	A4-1: Generation of new meta data and manifest table joining the Colombia and Tanzania datasets based on similar metadata columns in preparation for QIIME analysis. ((Joined data on the basis of sex, age, and sample type (fecal samples))
A1-2: Demultiplexing and creating new feature table (QIIME2)	A2-2: Demultiplexing and creating new feature table (QIIME2)	A3-2: Demultiplexing and creating new feature table (QIIME2)	A4-2: Demultiplexing and creating new feature table (QIIME2)
A1-3: Generate rooted trees, and taxonomy files.	A2-3: Generate rooted tree and taxonomy files.	A3-3: Generate predicted pathways and their abundances using PICRUST2. (R analysis)	A4-4: Mann-Whitney U Test to compare age tendency across data sets, and One-Proportion Z Test to compare sex proportions across groups (R analysis)
A1-4: Generate Phyloseq object and filter data to remove	A2-4: Generate Phyloseq object and filter data to remove	A3-4: Manual analysis of predicted pathways,	A4-4: Linear regression of diversity metrics

nonbacterial sequences, samples with too little reads, and ASVs that are too rare. (R analysis)	nonbacterial sequences, samples with too little reads, and ASVs that are too rare. (R analysis)	comparison between pathways present between the Tanzania and Colombia datasets	across age groups, and sex. (R analysis)
A1-5: Generate rarefaction curve and determine proper sampling depth (R analysis)	A2-5: Generate rarefaction curve and determine proper sampling depth. (R analysis)		
A1-6: Carry out alpha diversity analysis (Shannon) and beta diversity analysis (Weighted Unifrac) (R analysis)	A2-6: Undergo core microbiome analysis on samples from Tanzania and Colombia to see the taxa they have in common via Venn diagram analysis. (R analysis)		
	A2-7: Statistical significance of taxa/species of interest using indicator species analysis. (R analysis)		

Timeline for Completing Work



Feasibility and Training

The QIIME2 component of the project shouldn't be that much of a challenge except for combining the Tanzania and Colombia datasets which has already been done. We have already received guidance through the recordings and tutorials on how to do most if not all the techniques required to carry out our aims with regards to data manipulation and analysis.

The R component of the project should also be relatively straightforward, albeit we will have to teach ourselves how to use PICRUST2 to carry out our third aim. Other than that, we have already received guidance on how to do most if not all the techniques required to carry out our aims from the recordings and tutorials.

Should we need any additional training in QIIME2 or R our first approach will be to consult google to see if the issue is something we can easily solve on our own. Otherwise, we will save our questions for Chad during our weekly meetings and if the issue is especially urgent, whoever is free will ask the teaching team during the earliest office hours.

Dataset Overview

As stated in the introduction, our combined dataset contains data collected by researchers investigating the seasonal cycling of the Hadza tribe gut microbiome (7), in addition to data collected by researchers assessing the effects of diet westernization in the Colombian population (8). Sequences in each sample were demultiplexed via QIIME2 before subsequent denoising using the DADA2 denoising tool (16). Results of the denoising can be found in Table 1. The maximum read length represented in the samples is 251 bp; at least 98% of all samples are this length. Based on the read quality plot (Figure 1) the reads will all be truncated to 210 bp, as after this point the quality score begins to drop noticeably. The data was filtered during combination of the datasets pre-processing, so all the samples in the combined dataset will be used in our analysis. Diversity analysis will be carried out at a rarefaction depth of 22172, as this maximizes both retained features and samples. 52.42% of the features in the dataset are retained in 77.23% of the samples (Figures 2 & 3).

Table 1: Denoising results

	Before denoising	After denoising
Total sample size	632	628
Sequencing depth	4305 - 117,562	95,000

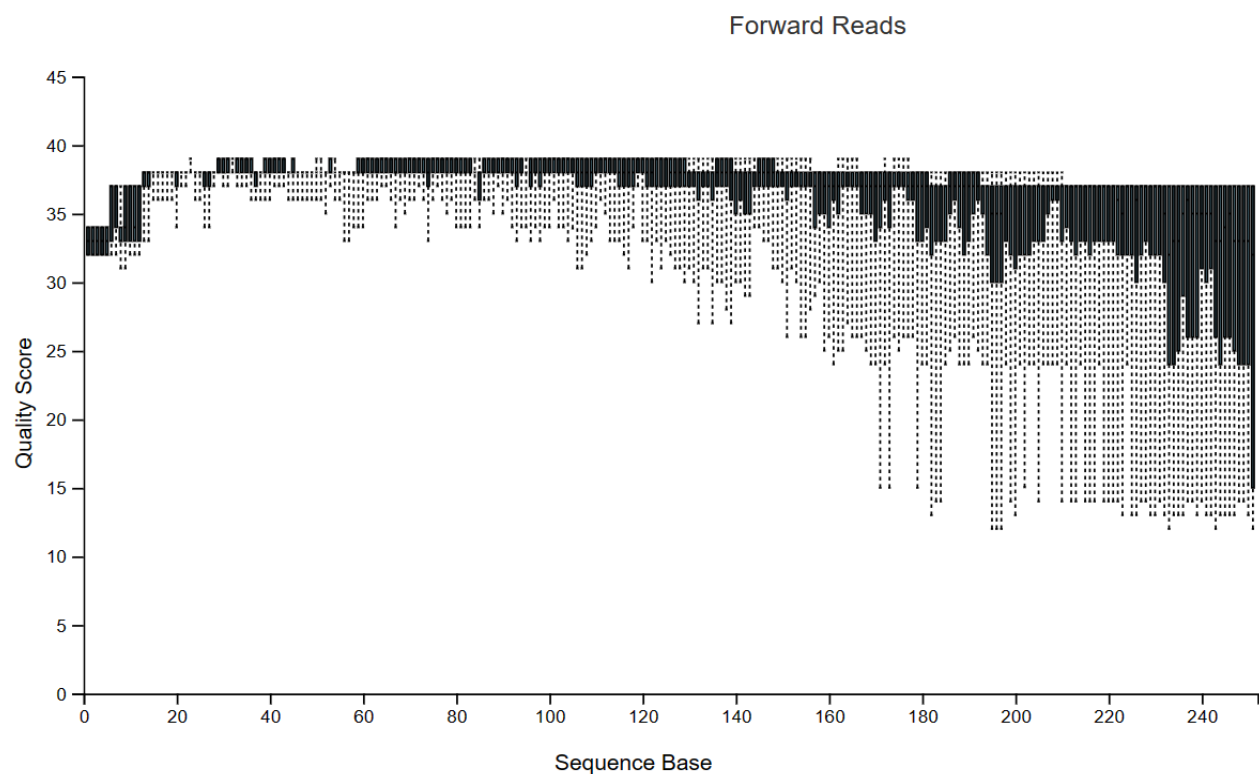


Figure 1: Read Quality Plot

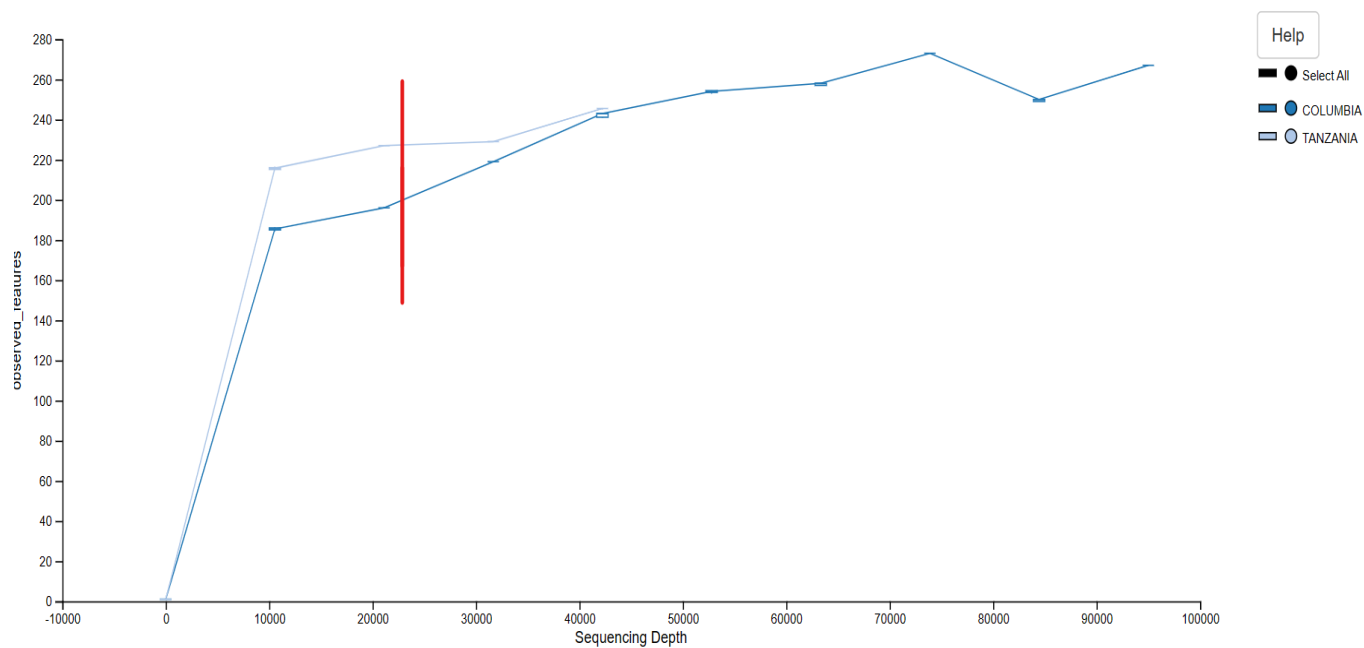


Figure 2: Alpha rarefaction curve showing observed features based on sequencing depth. Red line indicates chosen rarefaction depth of 22172.

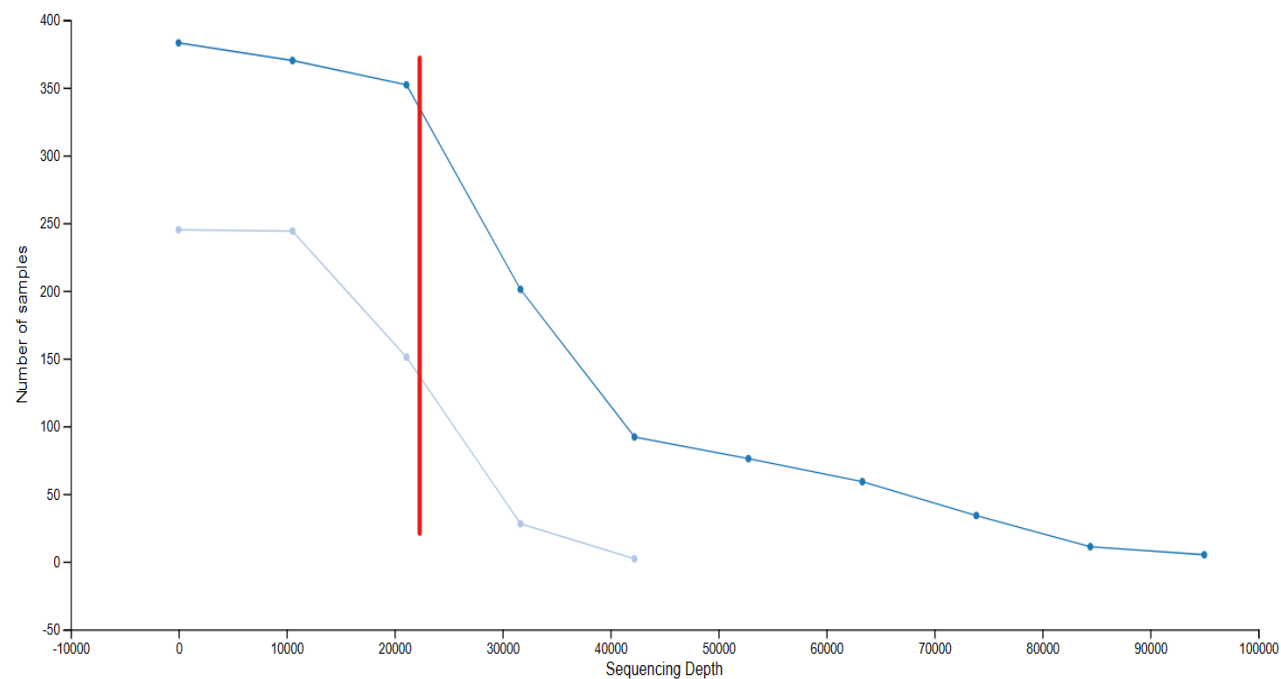


Figure 3: Alpha rarefaction curve showing retained samples based on sequencing depth. Red line indicates chosen rarefaction depth of 22172. Legend found in Figure 1.

References

1. Jandhyala SM, Talukdar R, Subramanyam C, Vuyyuru H, Sasikala M, Reddy DN. 2015. Role of the normal gut microbiota. *World J Gastroenterol* 21:8787–8803.
2. Leeming ER, Johnson AJ, Spector TD, Le Roy CI. 2019. Effect of Diet on the Gut Microbiota: Rethinking Intervention Duration. *Nutrients* 11:2862.
3. Contevelle LC, Oliveira-Ferreira J, Vicente ACP. 2019. Gut Microbiome Biomarkers and Functional Diversity Within an Amazonian Semi-Nomadic Hunter–Gatherer Group. *Front Microbiol* 10:1743.
4. Martinez KB, Leone V, Chang EB. 2017. Western diets, gut dysbiosis, and metabolic diseases: Are they linked? *Gut Microbes* 8:130–142.
5. Boddy SL, Giovannelli I, Sassani M, Cooper-Knock J, Snyder MP, Segal E, Elinav E, Barker LA, Shaw PJ, McDermott CJ. 2021. The gut microbiome: a key player in the complexity of amyotrophic lateral sclerosis (ALS). *BMC Med* 19:13.
6. de la Cuesta-Zuluaga J, Kelley ST, Chen Y, Escobar JS, Mueller NT, Ley RE, McDonald D, Huang S, Swafford AD, Knight R, Thackray VG. 2019. Age- and Sex-Dependent Patterns of Gut Microbial Diversity in Human Adults. *mSystems* 4:e00261-19.
7. Smits SA, Leach J, Sonnenburg ED, Gonzalez CG, Lichtman JS, Reid G, Knight R, Manjurano A, Changalucha J, Elias JE, Dominguez-Bello MG, Sonnenburg JL. 2017. Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. *Science* 357:802–806.
8. de la Cuesta-Zuluaga J, Corrales-Agudelo V, Velásquez-Mejía EP, Carmona JA, Abad JM, Escobar JS. 2018. Gut microbiota is associated with obesity and cardiometabolic disease in a population in the midst of Westernization. *Sci Rep* 8:11356.
9. Nagpal R, Mainali R, Ahmadi S, Wang S, Singh R, Kavanagh K, Kitzman DW, Kushugulova A, Marotta F, Yadav H. 2018. Gut microbiome and aging: Physiological and mechanistic insights. *Nutrition and Healthy Aging* 4:267–285.
10. Haro C, Rangel-Zúñiga OA, Alcalá-Díaz JF, Gómez-Delgado F, Pérez-Martínez P, Delgado-Lista J, Quintana-Navarro GM, Landa BB, Navas-Cortés JA, Tena-Sempere M, Clemente JC, López-Miranda J, Pérez-Jiménez F, Camargo A. 2016. Intestinal Microbiota Is Influenced by Gender and Body Mass Index. *PLoS One* 11:e0154090.
11. Fan J, Zhou Y, Meng R, Tang J, Zhu J, Aldrich MC, Cox NJ, Zhu Y, Li Y, Zhou D. 2023. Cross-talks between gut microbiota and tobacco smoking: a two-sample Mendelian randomization study. *BMC Med* 21:163.
12. Clemente-Suárez VJ, Beltrán-Velasco AI, Redondo-Flórez L, Martín-Rodríguez A, Tornero-Aguilera JF. 2023. Global Impacts of Western Diet and Its Effects on Metabolism and Health: A Narrative Review. *Nutrients* 15:2749.
13. Cordain L, Miller JB, Eaton SB, Mann N, Holt SH, Speth JD. 2000. Plant-animal subsistence ratios and macronutrient energy estimations in worldwide hunter-gatherer diets. *Am J Clin Nutr* 71:682–692.
14. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MG. 2020. PICRUST2 for prediction of metagenome functions. *Nature Biotechnology* 38:685–688.
15. Schnorr SL, Candela M, Rampelli S, Centanni M, Consolandi C, Basaglia G, Turrone S, Biagi E, Peano C, Severgnini M, Fiori J, Gotti R, De Bellis G, Luiselli D, Brigidi P, Mabulla A, Marlowe F, Henry AG, Crittenden AN. 2014. Gut microbiome of the Hadza hunter-gatherers. *Nat Commun* 5:3654.
16. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. 2016. Dada2: High-resolution sample inference from Illumina Amplicon Data. *Nature Methods* 13:581–583.

Participation Report

Trushaan: Introduction + Dataset overview

Joshua: Aim1, Timeline

Adam: Aim 2

Timothy: Aim 3

Farbod: Aim 4

Everyone: Editing, flow chart, clarity and grammar