Volume 10:1-11

UJEMI+

The Undergraduate Journal of Experimental Microbiology & Immunology (+Peer Reviewed)

The Diet of the Hadza Tribe is Higher in Gut Microbial Diversity, but Lower in Functional Diversity when Compared to a Westernized Diet

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

Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada

SUMMARY The Western diet has recently been a point of scrutiny in the scientific community due to its association with various metabolic diseases such as obesity, type 2 diabetes, cardiovascular disease, and certain cancers. However, the unique hunter-gatherer diet of the Hadza tribe has been noted to be protective against such diseases, which has been partially attributed to their unique gut microbiome. While previous studies have compared the gut microbiomes of the Hadza to those of Western countries, very few have looked into its relationship to the gut microbiome's metabolic pathways and functional diversity. In our study, we confirm the literature's finding that the Hadza microbiome, when compared to a westernized gut microbiome, is higher in alpha diversity and demonstrates high beta diversity. Additionally, we found multiple beneficial taxa to be present in higher amounts in the Hadza gut microbiome with more dysbiotic species being present in the westernized one. Further, our study indicates that the Hadza surprisingly demonstrated lower functional diversity, while the westernized dataset contained more upregulated metabolic pathways.

INTRODUCTION

The Western diet is one defined by its abundance of pre-packaged foods, refined sugar, candy, fried foods, processed meat, and sugary beverages (1). The diet has been a point of scrutiny due to its association with chronic metabolic diseases such as obesity, type 2 diabetes, cardiovascular disease, and certain cancers (1). The gut microbiota has been known to play a protective role in the development of these diseases, primarily by the production of short-chain fatty acids (SCFAs) that have demonstrated protective effects against chronic diseases such as obesity and hypertension (2). Contrasting the Western diet is the hunter-gatherer lifestyle of the Hadza. This Tanzanian tribe is one of the few remaining hunter-gatherer cultures, with their diet consisting mainly of meat, berries, honey, and tubers (3). Although they have little to no access to modern medicine, the Hadza's rate of metabolic disease, infectious disease, and nutritional deficiencies remain low, having partially been attributed to their unique gut microbiome (3). Therefore, great interest lies in researching the unique diet and corresponding gut microbiome of the Hadza, in hopes of researching what fundamental differences are key to reducing metabolic diseases.

We decided to compare different datasets, one containing gut microbiome data from the Hadza tribe in Tanzania and the other containing data from Colombian individuals. The Colombian dataset came from a study carried out by *de la Cuesta-Zuluaga* where they stated that the gut microbiome for Colombian individuals is in the midst of Westernization (4). Specifically, they concluded that the gut microbiome sample from Colombian individuals across different urban populations contained a microbial composition that resembled a gradient between what is expected for a more traditional diet and a westernised one (4). The

Published Online: September 2024

Citation: Hassan, Nematifar, Bernas, Bundhoo, Jin. 2024. The diet of the Hadza tribe is higher in gut microbial diversity, but lower in functional diversity when compared to a westernized diet. UJEMI+ 10:1-11

Editor: Shruti Sandilya, University of British Columbia

Address correspondence to: https://jemi.microbiology.ubc.ca/

Hadza dataset was from a study by *Smits et al*, and investigated the seasonal cycling of the tribe's gut microbiome between the wet season and the dry season. The researchers concluded that the shift in diet from consuming more meat in the dry season to more fruit in the wet season contributed to seasonal cycling of various enzymatic pathways (5).

While previous studies have compared the unique microbiome of the Hadza to other communities, gaps are still present in this newly developing area of research. One study by *Schnorr et al.* compared Hadza microbiome samples with Italian microbiome samples in the context of SCFA production. However, this study was limited by sample size, only including 27 Hadza fecal samples and 16 Italian fecal samples. Another study by *Carter et al.* compared a sizable 167 Hadza microbiome samples to those from California and Nepal (6). However, this study did not investigate whether a relationship was present between diversity metrics and functional diversity.

Our study aimed to bridge this gap by investigating the relationship between diversity metrics previously described in the literature and how this impacts functional diversity between Hadza and westernized microbiomes. Using Shannon's alpha diversity and weighted UniFrac beta diversity metrics, we confirmed that the Hadza had overall greater gut microbial diversity and significantly differs when compared to Western gut microbiomes. Additionally, we performed a core microbiome analysis to further show that there was a high number of species present in the Hadza microbiome, but not in the Colombian microbiome. Despite signs of increased diversity in other areas, functional analysis revealed that many metabolic pathways in the Hadza microbiome were significantly downregulated compared to the Colombian microbiome. Our findings illustrate that between the two microbiomes, while Hadza gut microbiomes are higher in alpha diversity, a westernized gut microbiome is higher in functional diversity.

METHODS AND MATERIALS

Dataset and metadata information. The raw FASTQ files for both the Hadza (Tanzania) and Colombian data sets were obtained from studies carried out by Smits et al, and de la Cuesta-Zuluaga et al respectively (3, 4). The Tanzanian or Hadza data set contained 350 fecal samples from 188 individuals (both men and women) over a span of 12 months over the wet and dry season, including those over the age of 3 years. All of the samples were collected from two separate camps within the central Rift Valley of Tanzania referred to as Sengeli and Hukamako camps. There were also samples collected via hand swabs or from non-human sources such as from honey, animal stomach swabs, or animal feces from cows. The Colombian data set contained fecal samples that all came from 441 individuals (both men and women) aged 18-62 from the largest urban centers of Colombia from July to November, encompassing a wet and dry season. The participants included in the trial were of ranging BMIs (no underweight individuals), heights, were not taking antibiotics/parasitics, lacked neurodegenerative diseases, and were not pregnant. DNA was extracted from fecal samples and through next-gen miSEQ sequencing, the V4 region of the 16S rRNA gene was amplified to generate amplicon sequence variants (ASVs) for samples in each dataset.

Metadata filtering and grouping. The original metadata belonging to the Colombia datasets contained many different categories ranging from BMI class (overweight/obese/healthy), body fat percentage, caloric intake, cardiometabolic status (4). The original metadata for the Hadza dataset had categories that specified the time in which the samples were collected from, the source it derived from (human vs non-human), the sex/age of the human samples, and the geographic location it was collected within. The metadata files for each dataset were joined together on the basis of shared categories such as age, sex, and fecal samples (3). Every other metadata category that didn't fall within those parameters was not included within our analysis. We then removed any fecal samples from within the Tanzania dataset that came from individuals that were below the age of 18 to ensure that only adults were being compared. Lastly, we removed samples in the Columbia dataset that were from smokers, to eliminate possible confounds. The newly generated combined metadata table was then used for the rest of our analysis.

Data processing using QIIME2. A manifest file was created with the samples of interest before being imported into Quantitative Insights into Microbial Ecology 2 (QIIME2) version 2023.7 (7). The manifest file was used to import and demultiplex sequences of interest for processing. 632 samples with a sequencing depth of 117,562 were present in the data before denoising. ASVs were denoised using the Divisive Amplicon Denoising Algorithm 2 (DADA2) and trimmed to 210 bp based on to maintain a Phred score of at least 35 (8). This resulted in the retention of 628 samples at a sampling depth of 95,000. The DADA2 representative sequences output was passed through the SILVA 138-99 classifier to complete taxonomic analysis of the ASVs (9). The generated feature table was then filtered to exclude mitochondrial and chloroplast DNA sequences. Sequences were then alpha-rarefied at a depth of 22,176, retaining 52.42% of the features in 77.23% of the samples for further analysis.

Alpha diversity analysis. The data processed in QIIME2 was imported into RStudio (version 2023.06) for further analysis. The metadata file, taxonomy file, phylogenetic tree file, and the feature table file were combined to form a phyloseq object using the phyloseq R package (10). Alpha diversity was then measured using the Shannon's Diversity Index and represented with a boxplot. Statistical analysis was performed using a pairwise Kruskal-Wallis test using the ggpubr RStudio package (11).

Beta diversity analysis. Beta diversity analysis was performed via weighted UniFrac to generate a PCoA plot. Statistical significance was measured through PERMANOVA using the *adonis2* function from the *Vegan* R package (12).

Proportionality Testing. To determine that the populations were fit to be compared in terms of age and sex, proportionality testing in R (version 2023.06) was done using the tidyverse (13) and dplyr (14) packages. A chi-square test was done in R to compare the proportion of males and females in the groups. A Wilcoxon test was used to examine the age distribution in the groups.

Linear Regression. To determine if age and sex were confounding variables for beta and alpha diversity, linear regression models were performed in R (version 2023.06) using the carData, car, and MASS libraries.

Core microbiome. Core microbiome analysis was carried out within R using the *microbiome* package (version 1.22.0) (15). We used an abundance threshold of 0.001 and a prevalence of 0.10 and only kept taxa that were deemed significant (P value < 0.05). Results were visualised with a Venn diagram using the *ggVenndiagram* package (version 1.2.3) (16).

Indicator taxa analysis. Indicator taxa analysis was carried out using the *indicspecies* R package (version 1.7.14) (17). Data was filtered based on the two different locations in Tanzania and Colombia. Afterwards, ASVs were then filtered and grouped on the basis of genus and only significant outputs (P value < 0.05) were displayed in table 1 with their corresponding indicator taxa value.

Metagenome prediction generation and visualisation. Metagenome predictions were generated using the Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt2) pipeline (18). The inputs for PICRUSt2, a FASTA file compiling all sequences of interest and a table of those sequence abundances, were generated using the DADA2 pipeline with a left trim length of 0 and a truncation length of 210 (8). Visualisations of the PICRUSt2 outputs were generated in R using the ggpicrust2 R package (19). Differential abundance analysis was performed using the LinDA method and annotated using the Kyoto Encyclopedia of Genes and Genomes (KO) database.

RESULTS

Greater diversity within hunter gatherer gut microbiome than westernized microbiome. In order to indicate overall differences in microbiome, we ran a core microbiome analysis. This indicated large compositional differences between the Hadza and

Westernized datasets. They share 114 species, compared to the Hadza and Westernized datasets having 176 and 77 unique to themselves, respectively (Figure 1). Additionally, it was observed that there is a greater degree of diversity within the gut microbiome of those

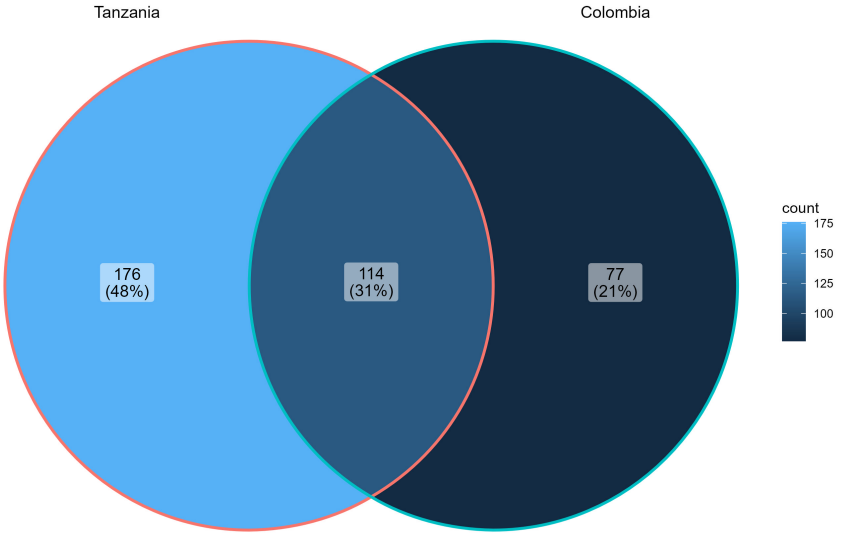
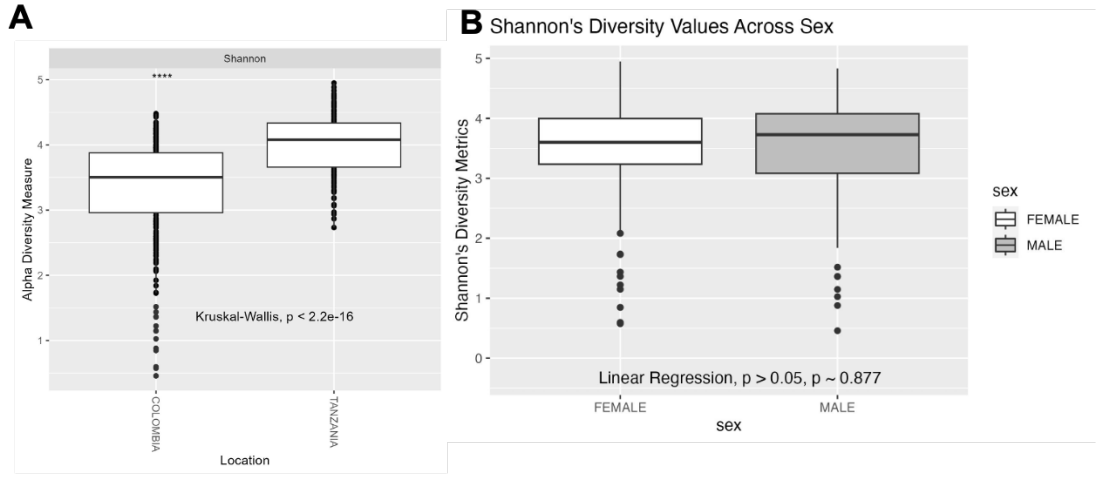


FIG. 1 More unique taxa within Hadza hunter gatherer gut microbiome than westernized Colombian. Core microbiome analysis was carried out on fecal samples from the Tanzania and Colombia datasets. An abundance threshold of 0.001 and a prevalence of 0.10 were used. Top indicator taxa for each individual diet type are detailed within Table 1.

that follow a hunter gatherer diet (Hadza) than those who adhere to a westernized diet, we carried out alpha and beta diversity-based analysis using the Shannon and weighted UniFrac metrics. Shannon's Diversity Index was significantly higher (p value = 2.2 e-16) within the Hadzan gut microbiome when compared to the Colombian microbiome (Figure 2A). These



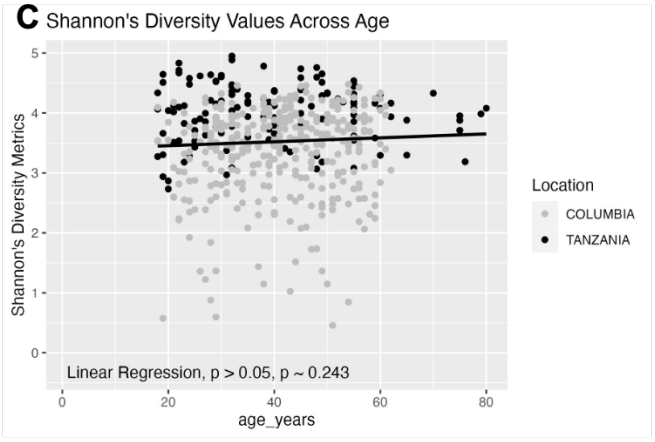


FIG. 2 The Hadza gut microbiome has higher alpha diversity than the Westernized gut microbiome. Alpha diversity was measured through the Shannon diversity index (Panel A). Significance measured through a pairwise Kruskal-Wallis test, $p < 0.05$. Sex (Panel B) and age (Panel C) do not correlate with a significant difference in alpha diversity as measured through linear regression analysis.

results reveal that the Hadzan gut microbiome has a greater degree of species richness and evenness when compared to their westernized diet counterparts. We also observed through our weighted UniFrac analysis and the resulting PCoA plot that samples that belong to the Hadza tend to cluster together and samples that correspond to Colombia tend to cluster together with a minimal amount of overlap between the two locations (Figure 2A). Therefore, the results that we acquired primarily reveal that, in terms of microbial composition, the two diet types are significantly different from each other (p value = 0.001). Our findings show that there is a greater overall diversity within the Hadzan gut microbiome than their Colombian contemporaries.

Age and sex differences do not contribute to differences in gut-microbiome alpha diversity. To determine if there are other confounding variables driving differences in gut microbiome alpha diversity, linear regressions were performed across Shannon's diversity values along with age or sex to determine if these metrics were drivers of alpha diversity differences across the entire dataset. Age (p value = 0.243) and sex (p value = 0. 877) demonstrated there was no significant shift in alpha diversity when segmenting for these variables (Figure 2B, 2C). Location was the only driver of differences in alpha diversity values between the populations.

Age does not demonstrate a significant shift in beta diversity values, but sex does. To determine whether age and sex contribute to significant differences in beta-diversity, weighted UniFrac analyses with PERMANOVA were performed and the resulting PCoA plots were generated. Age did not meaningfully contribute to differences in beta diversity (p = 0.9) (Figure 3B). Sex based analysis did demonstrate that sex does lead to significant differences in beta diversity between men and women (p = 0.01) (Figure 3C). Though, it is not as significant as location. However, qualitative observations in the differences between the diet of Tanzanian men and women may account for this, and this presents an interesting branch for future exploration. In conclusion, location is the primary driver of beta-diversity but sex-based differences in beta diversity exist that should be explored further.

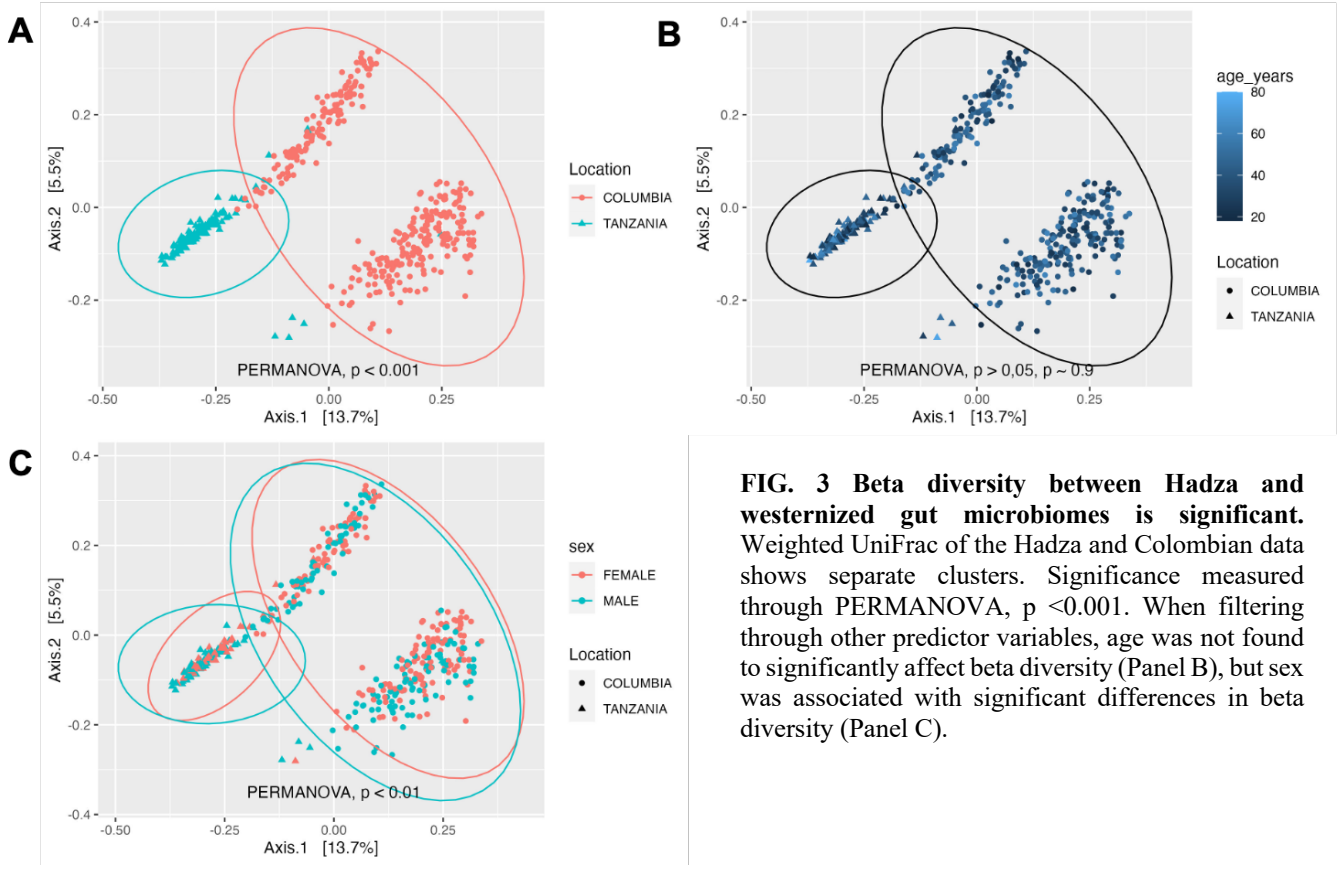


FIG. 3 Beta diversity between Hadza and westernized gut microbiomes is significant. Weighted UniFrac of the Hadza and Colombian data shows separate clusters. Significance measured through PERMANOVA, p <0.001. When filtering through other predictor variables, age was not found to significantly affect beta diversity (Panel B), but sex was associated with significant differences in beta diversity (Panel C).

More unique taxa within gut microbiome of hunter-gatherer diet than westernized diet. Microbiome analysis indicated a greater amount of unique taxa within the hunter gatherer gut microbiome (176 different taxa) compared to the westernized microbiome (77 different taxa). We discovered that the two different diet types had some degree of taxa in common (114 different taxa) but they had a larger degree of dissimilarity than similarity (Figure 3). We also observed that there are more unique taxa within the Hadzan gut microbiome than the Colombian (Figure 3). Thus, the gut microbiome of the two different diet types is compositionally different from each other and corroborates the idea that gut microbial diversity is greater within the hunter gatherer diet type.

Hunter gatherer diet primarily contains beneficial taxa from the *Lachnospiraceae* family while the westernized diet does not. In order to get a better understanding of the specific microbial taxa that are unique to the Hadzan and Colombian gut microbiome, an indicator taxa analysis was performed. We observed that the top 10 taxa within the Tanzanian data set were overrepresented with members that belonged to the family *Lachnospiraceae* (Table 1). We also discovered that the Colombian microbiome primarily had bacteria belonging to notable genera such as *Escherichia-Shigella* and *Collinsella* (Table 1). Thus, the lack of beneficial taxa belonging to the family *Lachnospiraceae* within the Colombian gut microbiome may lead to adverse health outcomes in diseases such as colorectal cancer (20).

TABLE. 1 Top 10 indicator taxa within Hadza and westernized gut microbiome

Tanzania							
Rank	**IV**	**P**	**Phylum**	**Class**	**Order**	**Family**	**Genus**
#1	0.949	0.005	Bacteroidota	Bacteroidia	Bacteroidales	*Prevotellaceae*	*Prevotellaceae NK3B31 group*
#2	0.948	0.005	Spirochaetota	Spirochaetia	Spirochaetales	*Spirochaetaceae*	*Treponema*
#3	0.934	0.005	Bacteroidota	Bacteroidia	Bacteroidales	*Rikenellaceae*	*Rikenellaceae RC9 gut group*
#4	0.916	0.005	Bacteroidota	Bacteroidia	Bacteroidales	*Prevotellaceae*	*Alloprevotella*
#5	0.880	0.005	Bacteroidota	Bacteroidia	Bacteroidales	*Prevotellaceae*	*Prevotellaceae UCG-003*
#6	0.877	0.005	Firmicutes	Clostridia	Lachnospirales	*Lachnospiraceae*	*Marvinbryantia*
#7	0.876	0.005	Bacteroidota	Bacteroidia	Bacteroidales	*Prevotellaceae*	*Genus uncultured*
#8	0.869	0.005	Firmicutes	Clostridia	Lachnospirales	*Lachnospiraceae*	*Eubacterium hallii group*
#9	0.863	0.005	Proteobacteria	Gammaproteobacteria	Aeromonadales	*Succinivibrionaceae*	*Succinivibrio*
#10	0.834	0.005	Firmicutes	Clostridia	Lachnospirales	*Lachnospiraceae*	*Lachnospiraceae ND3007 group*
Colombia							
#1	0.952	0.005	Actinobacteriota	Coriobacteriia	Coriobacteriales	*Coriobacteriaceae*	*Collinsella*
#2	0.950	0.005	Verrucomicrobiota	Verrucomicrobiae	Verrucomicrobiales	*Akkermansiaceae*	*Akkermansia*
#3	0.940	0.005	Bacteroidota	Bacteroidia	Bacteroidales	*Rikenellaceae*	*Alistipes*
#4	0.939	0.005	Actinobacteriota	Actinobacteria	Bifidobacteriales	*Bifidobacteriaceae*	*Bifidobacterium*
#5	0.936	0.005	Proteobacteria	Gammaproteobacteria	Enterobacterales	*Enterobacteriaceae*	*Escherichia-Shigella*
#6	0.916	0.005	Firmicutes	Bacilli	Lactobacillales	*Streptococcaceae*	*Streptococcus*
#7	0.895	0.005	Firmicutes	Clostridia	Oscillospirales	*Ruminococcaceae*	*UBA1819*
#8	0.893	0.005	Actinobacteriota	Actinobacteria	Actinomycetales	*Actinomycetaceae*	*Actinomyces*
#9	0.875	0.005	Proteobacteria	Gammaproteobacteria	Enterobacterales	*Ruminococcaceae*	*Incertae_Sedis*
#10	0.872	0.005	Actinobacteriota	Coriobacteriia	Coriobacteriales	*Enterobacteriaceae*	*N/A*

Greater functional diversity in gut microbiomes of Western diets consumers (Colombian adults) compared to hunter gatherer diet consumers (Tanzanian adults). Principal component analysis (Figure 4A) indicated that there is a significant overlap in the functional diversity of the Colombia and Tanzania gut microbiome samples, with Colombia clustering almost entirely inside Tanzania. Additionally, Colombia has a much wider range

of functional diversity outside its primary cluster. At least 30 metabolic pathways are significantly more enriched in the Colombia gut microbiome when compared to the Tanzania gut microbiome, indicating more pathways with unique functions are present in the Colombia gut microbiome (Figure 4B).

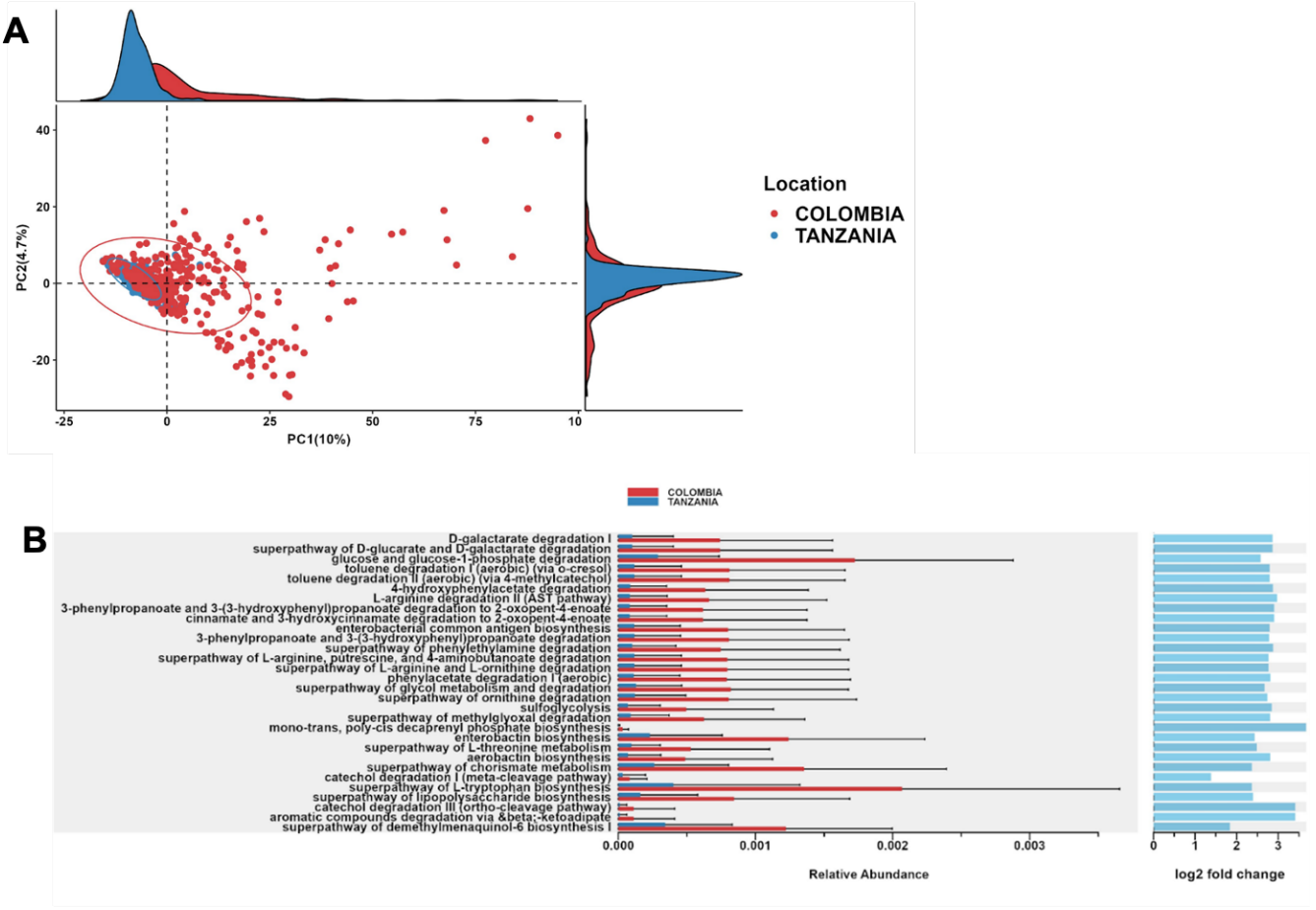


FIG. 4 Greater functional diversity within westernized Colombian gut microbiomes than Hadza hunter gather gut microbiomes. Principal component analysis was carried out on fecal samples from the Hadza and Colombia datasets, filtering for P values less than 0.05 (panel A, top). Errorbar plot of the top 30 enriched pathways with the smallest P adjusted values grouped by location and annotated using the KO database (panel B, bottom).

DISCUSSION

In order to initially prove that there are significant compositional differences between the gut microbiome between hunter-gatherer and westernized diet types, we carried out alpha and beta diversity analyses between the two datasets. Using the Shannon diversity index as a measure of alpha diversity, we found that Tanzanian gut microbiomes had significantly higher alpha diversity in comparison to Colombia (Figure 2A). These results were expected since there have been studies that have echoed this sentiment such as the one by *Schnorr et al* which analysed the gut microbiome between Italians and Hadza where they also saw a greater degree of Shannon diversity within the Hadza tribe (3). We also demonstrated via weighted UniFrac analysis that gut microbial samples that belong to the Hadza clustered together while the Colombian samples generally clustered together too. However, a minor discrepancy that we revealed is that the Colombian samples are a bit more stratified, having two distinct subclusters that are both within different regions on the PCOA plot.

From the results, it was determined that location is the primary driver of significant differences in alpha and beta diversity between the populations (Figure 2). However, sex was also a significant driver of beta-diversity. Other studies done on the original Hadza data set had similar findings, as weighted UniFrac analysis showed that there were significant

differences between men and women's gut microbiome beta diversity values that could be attributed to sex-based divisions in labour (3). For example, women mostly forage for plants and spend most of their time in camp. Thus, sex-based differences within their biomes could be attributed to women eating foods rich in plant fiber, while men consumed more meat products (3). The significant difference between the composition of Hadza male and female microbiomes could drive the differences in clustering found in our study.

We further investigated the differences between the gut microbiomes of a hunter-gatherer and westernized diet by looking at the specific taxa that are key members via indicator taxa analysis. Our findings revealed that the hunter-gatherer microbiome had significantly increased levels of genera such as *Eubacterium hallii*, a group that is responsible for the production of SCFAs, such as propionate (21). The microbiomes also contained significantly more members belonging to the family *Lachnospiraceae*, another family responsible for the production of SCFAs such as butyrate and propionate. In contrast, the westernized Colombia dataset had an increased presence of dysbiotic species such as members from the *Escherichia-Shigella* genera, which are known to contain opportunistic pathogens. *Wang et al* found that a greater abundance of *Escherichia-Shigella* in colorectal patients was correlated with a lower abundance of *Lachnospiraceae*, which could potentially play a role in SCFA production (20). The westernized dataset also contains another unique genus in the form of *Collinsella* which has been associated with poor health outcomes. Specifically, a study carried out by *Karlsson et al* revealed that those suffering from symptomatic atherosclerosis had increased abundance of species that belonged to the genus *Collinsella* (22). In conclusion, our taxonomic analysis reveals that the westernized gut microbiome lacks SCFA producers and instead contains more taxa that are associated with cancer and cardiovascular disorders.

To determine the differences in metabolic pathways between westernized and Hadza microbiomes, functional analysis via PICRUST2 was performed. We observed a greater increase in functional diversity in the Colombian gut microbiome compared to the Hadza. A large portion of these enriched pathways within the Colombian gut microbiome belong to degradation pathways, specifically ones involved in the degradation of various sugars. This upregulation in sugar degradation pathways matches the increased sugar consumption associated with westernized diets that have been correlated with higher rates of obesity and diabetes (23). In addition to this, the degradation of L-arginine was enriched in the Colombia gut microbiome (Figure 4B). L-arginine can be found in relatively high concentrations in a number of foods not available to the Hadza people, such as seafood, nuts, seeds and rice (24). Additionally, the lower degree of functional diversity found within the Hadza gut microbiome may be indicative of a much higher degree of functional redundancy due to a diet that is entirely restricted to the local flora and fauna. The gut microbiota of the Hadza are highly specialised to their unique diets (24). This greater functional redundancy may be a strong indicator of stability and resilience to any perturbations in the gut microbiome (25). With regard to sudden and short-term changes in the diets of both groups, it is likely that their gut microbiomes will quickly return to pre-intervention states after returning to their regular diet (25). On the other hand, the impacts of long term, more permanent changes in diet are harder to predict due to a lack of research exploring this topic. Overall, our results indicate that while the Hadza gut microbiome is higher in overall diversity metrics, their gut microbiomes demonstrate lower functional diversity, mostly attributing to their diet.

Limitations When considering which data to include in our analysis, we ultimately made the decision to include all 350 Tanzanian samples. This introduces our first limitation, which is that we did not consider how different time periods could impact the gut microbiome. Composition in the Hadza can slightly alter in response to different foods consumed during the wet and dry seasons (5). The Tanzanian dataset across different time periods could not be investigated with any reasonable comparison against the Colombia dataset because the Colombia study lacked dates. The authors of the Colombia paper took their samples from July to November (4), but unfortunately did not include the date collected as a metadata category. This makes it impossible to compare Columbia's "wet season samples" and "dry season samples" to those of Tanzania. Fortunately for us, the time periods in which both sets were collected are seasonally similar. The Colombia study took their samples from July to November, which is equivalent to one Colombian wet season and one Colombian dry season.

Coincidentally, 12 months in Tanzania also encompasses one wet season and one dry season (4), which allows for a fair comparison between the two combined datasets. This unfortunately brings in a second limitation which is potential sampling bias. The authors mention that Hadzan microbiome samples from the same individual could be represented twice in the wet and dry season. However, samples from both seasons were necessary for a fair analysis of both populations.

Additionally, indicator taxa analysis resulted in over 200 taxa for each respective diet type. As this is reasonably outside the scope of our project, we only looked at the top 10 species. Therefore, our results may not fully encapsulate the differences in microbial composition between the two populations and how it contributes to their differences in health outcomes.

Another potential limitation was within the analysis of our Colombia dataset. Although we removed metadata categories such as individuals who smoked, we included individuals who were obese and had poor cardiovascular health. While the inclusion of obese individuals and those with poor cardiovascular health are representative of the West, the removal of smokers may not have been representative of the Colombian population or a Westernised population as a whole and could have skewed our results.

Conclusions Our study confirmed the literature's findings that comparing the Hadza gut microbiome with a westernized gut microbiome shows high beta diversity, with the Hadza being higher in alpha diversity. Core microbiome analysis revealed that beneficial families such as *Lachnospiraceae* were far more abundant in the Hadza gut microbiome, while dysbiotic taxa belonging to the genus *Escherichia-Shigella* were far more abundant in the western microbiome. Finally, our functional analysis revealed that while lower in both raw taxa and Shannon's diversity, the Westernised gut microbiome possessed a greater degree of upregulated pathways when compared to the Hadza gut microbiome. This indicates that functional diversity does not necessarily correlate with alpha diversity, and suggests that functional redundancy is a relevant feature of the Hadza gut microbiome. This reveals the relevant factors in the Hadza microbiome, which can contribute to protection from adverse health outcomes and metabolic diseases.

Future Directions For the original analysis, smokers were removed from the data set as prior research indicated that there likely would be statistically significant differences between smokers and non-smokers (26), acting as a possible confound. Reintroduction of that data followed by comparing the compositional (alpha and beta) and functional (PCA plot and pathway) diversity between the gut microbiome of smokers and the Hadza people could potentially provide new insights. While many of the non-smoking Colombians could have a level of gut microbiome dysbiosis due to poor diet, comparing smokers' gut microbiome to the robust and healthy Hadza biome could elucidate specific additional damage from smoking on the gut microbiome as opposed to just poor diet (27).

Furthermore, to further understand how diversity in the gut microbiome could be restored, another study could be done with a non-smoking Colombian cohort where they adopt a "Hadza-esque" diet for a period of one year. Samples would be taken monthly to track changes in each group's respective alpha diversity, overall beta diversity, core-microbiome composition, taxa indicator species, and functional analysis to see if following a hunter-gatherer diet can meaningfully improve microbiome quality (high alpha diversity, increased number of beneficial SCFA producers, etc) and observing its effects on metabolic disease across a year. However, this experiment is rather unfeasible as it would have to consider the various bacterial species present in the drastically environments of Tanzania vs Colombia.

Within the weighted UniFrac analysis, we saw two distinct subclusters of samples within the Colombian dataset (Figure 2A). Specifically, there was one sub-cluster of Colombian samples that clustered near the Hadza gut microbiome and the other subcluster of Colombian samples were some distance away. This discrepancy was most likely due to differences in location, as the original Colombia dataset collected by de la Cuesta-Zuluaga sampled from different urban locations (4). Therefore, further analysis into the Colombian dataset could reveal relevant factors that allow it to cluster closer to the Hadza gut microbiome samples,

such as rural vs urban locations, access to animals, differences in diet, socioeconomic factors, etc.

Additionally, we found a confounding variable in sex across both groups when comparing beta diversity (Figure 2C). Further studies can compare the gut microbiomes of Hadza men with Colombian men, and the gut microbiomes of Hadza women with Colombian women. This would eliminate any sex-based differences in the two groups and would further elucidate the relevant markers of microbial diversity between these two groups. Additionally, to further understand the root of why male and female biomes cluster differently, beta diversity analysis clustered by sex on each individual population could also reveal significant confounds within each respective population.

Lastly, our main finding was that greater functional diversity was correlated with lower alpha diversity when comparing the Hadza and westernized gut microbiomes. Future studies can look into how increased functional diversity could potentially come at the cost of microbial diversity, leaving the host more vulnerable to dysbiosis as microbes could potentially compete for niches within the gut. Other studies can investigate how functional diversity is related to redundant pathways, and whether or not this affects this seemingly backwards relationship between functional and microbial diversity (28).

ACKNOWLEDGEMENTS

We would like to thank the entire MICB475 teaching team for all their guidance and support while we worked on this project, specifically, Dr. Evelyn Sun and Chad Poloni. We would also like to thank Smits et al and de la Cuesta-Zuluaga et al for providing the original datasets and the corresponding metadata. Lastly, we would like to thank the UBC Department of Microbiology and Immunology for providing the facilities, funding and resources we used to carry out our study. We would also like to thank two anonymous reviewers for constructive feedback on this manuscript.

CONTRIBUTIONS

Each team member contributed an equivalent amount of effort and time towards the data processing and analysis of this project. Tr.B. provided aim 1A and 2A and wrote the methods for data processing and aim 1A and 2A. A.A.H. provided aim 2, assisted in aim 1, and co-wrote the introduction, discussion, and study limitations. Ti.B. provided aim 3, and wrote the methods for aim 3, and co-wrote the discussion. F.N. provided aims 1B, 2B, 2C, wrote the methods for aims 1B, 2B, and 2C, and co-wrote the discussion and future directions. Y.J. conceptualized the study, wrote the abstract, conclusion, and co-wrote the discussion, future directions, and study limitations. All team members played a part in reviewing and editing the manuscript.

REFERENCES

1. **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.
2. **Ahmadmehrabi S, Tang WHW**. 2017. Gut Microbiome and its Role in Cardiovascular Diseases. Curr Opin Cardiol **32**:761–766.
3. **Schnorr SL, Candela M, Rampelli S, Centanni M, Consolandi C, Basaglia G, Turroni 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.
4. **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.
5. **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.
6. **Carter MM, Olm MR, Merrill BD, Dahan D, Tripathi S, Spencer SP, Yu FB, Jain S, Neff N, Jha AR, Sonnenburg ED, Sonnenburg JL**. 2023. Ultra-deep sequencing of Hadza hunter-gatherers recovers vanishing gut microbes. Cell **186**:3111-3124.e13.
7. **Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J,**

Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu Y-X, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, vander Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol **37**:852–857.

8. **Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP**. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods **13**:581–583.
9. **Pruesse E, Peplies J, Glöckner FO**. 2012. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics **28**:1823–1829.
10. **McMurdie PJ, Holmes S**. 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLOS ONE **8**:e61217.
11. **Kassambara A**. ggplot2 Based Publication Ready Plots. https://rpkgs.datanovia.com/ggpubr/. Retrieved 10 December 2023.
12. **Oksanen J, Simpson GL, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Solymos P, Stevens MHH, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Carvalho G, Chirico M, Caceres MD, Durand S, Evangelista HBA, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill MO, Lahti L, McGlinn D, Ouellette M-H, Cunha ER, Smith T, Stier A, Braak CJFT, Weedon J**. 2022. vegan: Community Ecology Package (2.6-4).
13. **Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen T, Miller E, Bache S, Müller K, Ooms J, Robinson D, Seidel D, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H**. 2019. Welcome to the Tidyverse. Journal of Open Source Software **4**:1686.
14. **Wickham H, François R, Henry L, Müller K, Vaughan D**. 2023. dplyr: A Grammar of Data Manipulation. R package version 1.1.0.
15. **Shetty S, Lahti L**. Introduction to the microbiome R package. https://microbiome.github.io/tutorials/. Retrieved 10 December 2023.
16. **Gao C-H, Yu G, Cai P**. 2021. Ggvenndiagram: An intuitive, easy-to-use, and highly customizable R package to generate Venn diagram. Frontiers in Genetics **12**.
17. **Cáceres MD, Legendre P**. 2009. Associations between species and groups of sites: Indices and statistical inference. Ecology **90**:3566–3574.
18. **Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MGI**. 2020. PICRUSt2 for prediction of metagenome functions. Nat Biotechnol **38**:685–688.
19. **Yang C, Mai J, Cao X, Burberry A, Cominelli F, Zhang L**. 2023. ggpicrust2: an R package for PICRUSt2 predicted functional profile analysis and visualization. Bioinformatics **39**:btad470.
20. **Wang T, Cai G, Qiu Y, Fei N, Zhang M, Pang X, Jia W, Cai S, Zhao L**. 2012. Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. ISME J **6**:320–329.
21. **Engels C, Ruscheweyh H-J, Beerenwinkel N, Lacroix C, Schwab C**. 2016. The Common Gut Microbe Eubacterium hallii also Contributes to Intestinal Propionate Formation. Front Microbiol **7**:713.
22. **Karlsson FH, Fåk F, Nookaew I, Tremaroli V, Fagerberg B, Petranovic D, Bäckhed F, Nielsen J**. 2012. Symptomatic atherosclerosis is associated with an altered gut metagenome. Nat Commun **3**:1245.
23. **Rakhra V, Galappaththy SL, Bulchandani S, Cabandugama PK.** 2020. Obesity and the Western Diet: How We Got Here. Mo Med **117**:536–5381.
24. **Wu G, Bazer FW, Davis TA, Kim SW, Li P, Marc Rhoads J, Carey Satterfield M, Smith SB, Spencer TE, Yin Y**. 2009. Arginine metabolism and nutrition in growth, health and disease. Amino Acids **37**:153–168.
25. **Tian L, Wang X-W, Wu A-K, Fan Y, Friedman J, Dahlin A, Waldor MK, Weinstock GM, Weiss ST, Liu Y-Y**. 2020. Deciphering functional redundancy in the human microbiome. Nat Commun **11**:6217.
26. **Gui X, Yang Z, Li MD**. 2021. Effect of Cigarette Smoke on Gut Microbiota: State of Knowledge. Front Physiol **12**:673341.
27. **Shapiro H, Goldenberg K, Ratiner K, Elinav E**. 2022. Smoking-induced microbial dysbiosis in health and disease. Clin Sci (Lond) **136**:1371–1387.
28. **Wang C, Yu Q-Y, Ji N-N, Zheng Y, Taylor JW, Guo L-D, Gao C**. 2023. Bacterial genome size and gene functional diversity negatively correlate with taxonomic diversity along a pH gradient. Nat Commun **14**:7437.