Volume 9:1-12

UJEMI+

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

Exploring non-WEIRD populations gut microbiota: the relative importance of obesity metrics on gut microbiome diversity and composition in Colombian adults

Ghazaleh Shoaib, Sana Samadi[1], Jong Yoon Joo[2] and Jen Wang[2]

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

[1] These authors contributed equally to this work as first co-authors
[2] These authors contributed equally to this work as second co-authors

SUMMARY Previously, gut microbiome research has predominantly focused on Western, educated, industrialized, rich and democratic (WEIRD) societies, which limits the generalizability and identification of patterns across ethnicities. This study explores the link between obesity and microbiome variation in an ethnically representative group, which is compared with other global datasets. In this study, we investigated the effect of obesity on the gut microbiome using a dataset with 442 Colombian adults to provide a valuable example of diversity in research. We first analyzed the alpha and beta diversity of obesity-related and general factors. This allowed us to see which ones significantly affected gut microbiota. This revealed some individual obesity-linked predictors affected variation in composition and diversity. However, when the dataset was filtered into obese and non-obese individuals, no specific microbial community compositional differences were found. Despite this, core microbiome analysis revealed certain gut bacterial species were consistently found in obese or non-obese groups. Finally, using model selection to contextualize obesity-related metrics among other predictors, we found that some obesity metrics significantly explained diversity but not composition. This study suggests that although there may be a significant link between obesity and gut microbial variation in WEIRD populations, the patterns may be potentially different in non-WEIRD populations such as Colombian adults.

INTRODUCTION

The diversity and composition of the gut microbiome is associated with various lifestyle factors, beyond just obesity, including age, diet, smoking, and exercise (1). However, the specific determinants and the extent of their impact on gut microbial variation remains unclear. When assessing the effects of obesity on the diversity/composition of the gut microbiome, it is imperative to consider the impact of other potential predictors. To evaluate the relative impact of various combinations of predictors within a dataset, researchers may employ model selection as an appropriate tool. Model selection compares various candidate "models" or combinations of predictors to determine which model best explains variation in the data. This can be applied to gut microbiome research to determine the combination of metadata categories or lifestyle factors that best explain microbial variation.

Historically, research on links between the gut microbiome and obesity has been done on Western, educated, industrialized, rich and democratic (WEIRD) societies (2-4). However, it

Published Online: September 2023

Citation: Shoaib, Samadi, Joo, and Wang. 2023. Exploring non-WEIRD populations gut microbiota: the relative importance of obesity metrics on gut microbiome diversity and composition in Colombian adults. UJEMI+ 9:1-12

Editor: Lauren Pugsley, University of British Columbia

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

remains unclear whether the findings apply universally across different ethnic groups. Notably, WEIRD populations have been found to have reduced gut microbial diversity compared to non-WEIRD populations (2). This may be due to cultural differences such as diet, pollution, exercise, sanitation disparities and stressors, all of which impact the gut microbiome (2). To address this gap in knowledge, we explore the link between gut microbial variation and obesity in anon-WEIRD population using data from Colombian adults. This adds to the growing body of literature on diverse populations in science (5). As a disclaimer, the term “non-WEIRD” (Western, educated, industrialized, rich and democratic) is used to highlight populations beyond the typical North American and European societies that have been the historical focus of research. However, we recognize the potential misinterpretation of this term to refer to Colombia, which is industrialized and non-rural. While this is true, we refer to Columbia as non-WEIRD as it represents a distinct socio-cultural context with notable differences in lifestyle and diet. We aim to contribute valuable insights into obesity and gut microbial variation in an ethnically representative group such as Columbia, which is what the term “non-WEIRD” attempts to encapsulate. We pursued three main aims in our study. The first aim was to examine alpha and beta diversity across all metadata categories. This was done to better understand which predictors individually affect microbial diversity and composition. We hypothesized that obesity-related metrics would impact microbiome diversity and composition. Our analysis revealed that several obesity-related metrics significantly impact alpha and beta diversity. To further explore this in our second aim, we conducted a comparison of gut microbial composition between "obese" and "non-obese" individuals, as defined by global health statistics. It is important to note that the Colombian government defines Obesity using only BMI in statistics. However, to achieve a more robust analysis, we used body fat percentage and waist circumference for a better representation of obesity, aligning with globally-accepted definitions. Considering the widely-established links in other populations (mostly WEIRD), we predicted obesity would have a significant impact on microbiome variation. For our third aim, it was necessary to understand the relative importance of obesity metrics compared to other predictors. To achieve this, we employed model selection, allowing us to determine the combination of predictors that best explains microbial variation.

Our results contribute to understanding the distinctive connections between obesity and the gut microbiome within a non-WEIRD population. Through a comprehensive analysis of this dataset and a special focus on the relative importance of obesity, our study offers valuable insights into the factors influencing the gut microbiome.

METHODS AND MATERIALS

Dataset and metadata. The dataset used in this study was derived from the research article “Gut microbiota is associated with obesity and cardiometabolic disease in a population amid Westernization” by Cuesta-Zuluaga *et al.* (6). It comprised of 16S rRNA gene sequencing data from fecal samples collected from 441 individuals (227 men and 214 women) between the ages of 18 and 65, residing in Colombia. The metadata accompanying the dataset was used as input for Quantitative Insights Into Microbial Ecology 2 (QIIME2) analysis (7). The original raw DNA reads from the study by Cuesta-Zuluaga *et al.* can be found at the SRA-NCBI under BioProject PRJNA417579. The R code to reproduce statistical analyses for this study is available in the supplemental.

Preliminary data processing and ASV determination. The raw sequence data was processed using the QIIME2 pipeline (version 2021.11). Single-end fastq files in Phred33V2 format were imported and demultiplexed single-end fastq files using the. Subsequently, th e demultiplexed samples were then denoised with the Divisive Amplicon Denoising Algorithm 2 (DADA2) method to determine amplicon sequence variants (ASVs) (8). A truncation length of 250 base pairs was selected.

Taxonomic classification and filtering. ASVs were classified using a Naïve Bayes classifier was trained using the SILVA (SSU and LSU rRNA sequences of Bacteria, Archaea and Eukarya) reference database (version 138-99) (9, 10). Prior to training, the reference

sequences were first trimmed according to the 16S rRNA gene V4 region primers (515f: 5'-GTGCCAGCMGCCGCGGTAA-3', 806R: 5'GGACTACHVGGGTWTCTAAT-3') using the 'qiime feature-classifier extract-reads' command. The trained classifier was obtained using the 'qiime feature-classifier fit-classifier-naive-bayes' command (11). The trained classifier was used to classify ASVs with the 'qiime feature-classifier classify-sklearn' command (12). Taxonomic bar plots were generated using the 'qiime taxa barplot' command, and Eukaryota, mitochondria, and chloroplasts were removed using the 'qiime taxa filter-table' command (7).

Phylogenetic tree construction and rarefaction. To perform phylogenetic diversity analysis, a phylogenetic tree was constructed using the 'qiime phylogeny align-to-tree-mafft-fasttree' command. This process involved aligning the representative sequences using Multiple Alignment using Fast Fourier Transform (MAFFT), and constructingan unrooted tree using FastTree (13). The tree was then rooted using the midpoint rooting method. Alpha rarefaction curves were generated using the 'qiime diversity alpha-rarefaction' command, which allows for assessing the sequencing depth and sample coverage (14). Based on the alpha rarefaction curve, a rarefied table was generated with a sampling depth of 20,655 (14). The ASV table, rooted phylogenetic tree, taxonomy, and representative sequences were converted and exported for further analysis in R.

Alpha and beta-diversity analysis on all metadata categories. The QIIME2 output files of feature table, metadata, taxonomy, and phylogenetic tree were imported into R. A phyloseq object was created using these four files (15). A custom R loop was implemented to conduct a linear regression model for all 35 metadata categories as predictors. P-value corrections were done using the Benjamini-Hochberg method, and a richness plot was created and saved for each metadata category. A similar loop was created to run PERMANOVA tests for all 35 metadata categories as predictors. Unweighted Unifrac was used as the distance metric. P-value corrections were done using the Benjamini-Hochberg method. A Principal Coordinate Plot (PCoA) plot was created for each predictor. Tables with predictors and their adjusted P-values were also created for alpha and beta diversity analysis.

Analyzing compositional differences between obese and non-obese individuals. Samples were filtered to only include individuals who either satisfied both waist circumference and body fat percentage obesity threshold (obese); or neither of them (non-obese). For females, samples with Body Fat Percentage $\geq$ 25% and Waist Circumference $\geq$ 80cm were considered obese, while for males, samples with Body Fat Percentage $\geq$ 30% and Waist Circumference $\geq$ 90cm were considered obese. Subsequently, phyloseq object was created by combining the filtered metadata with the feature table, taxonomy table, and phylogenetic tree. PCoA analysis was performed using the Unweighted Unifrac algorithm as the input.

Core microbiome analysis on obese and non-obese individuals. The core bacterial taxa in obese and non-obese individuals were found by calculating the core microbiome using the "microbiome" package in R (16). The parameters were set a frequency of 0.001 and the prevalence at 0.10. The results for obese and non-obese individuals were compared using a Venn diagram generated with the "ggVennDiagram" package (17).

Performing model selection on the dataset. All subsequent analyses were conducted using the phyloseq object. The full alpha-diversity model was created by conducting a linear model on all 35 predictors from the metadata using the Shannon index. The full beta-diversity model was created by conducting a PERMANOVA test on all 35 metadata predictors using the Weighted Unifrac metric (18). To select the final modelsmodels we implemented a stepwise regression by Akaike's Information Criteria (AIC) scoring. The alpha diversity model used the 'stepAIC()' function from the MASS package, while the beta diversity model used a custom loop to drop predictors from the model based on AIC score (https://github.com/kdyson/R_Scripts/blob/0130c64cfc1437c340a1237889456f4da31da871/AICc_PERMANOVA.R).

RESULTS

Metadata categories related to obesity, general health, or location significantly influenced microbial diversity and composition. Alpha diversity analysis was performed on all 35 metadata categories using Shannon as the diversity metric. The results revealed that three metadata categories had a significant impact on diversity: "Adiponectin Levels" (Shannon, p=0.019), "Animal Protein Intake" (Shannon, p=0.033), and "City" (Shannon, p=0.000) (Table 1). "Adiponectin levels" have been found to have underlying links to obesity, "Animal Protein Intake" is connected to overall health status, and "City" is location-related. Notably, only the cities of Bogota and Medellin had significant effects on diversity within the city category.

TABLE. 1 Metadata categories related to obesity, general health, or location showed significance in alpha diversity. Linear model with Shannon Index was performed individually on all 35 predictors within the dataset. Estimates indicate the increase or decrease in Shannon diversity. Significant results are indicated by an asterisk (* = P<0.05, ** = P<0.01, *** = P<0.001). n = 442.

Metadata Category	Estimate	P-value	Significance
Adiponectin (ng/mL)	0.026	0.019	*
Age (years)	0.006	0.096	ns
Ages Groups (years)[a]	0.136	0.102	ns
BMI (kg/m^2)	-0.004	0.677	ns
BMI Class Obese	-0.058	0.584	ns
BMI Class Overweight	-0.154	0.186	ns
Body Fat Percentage (%)	0.005	0.555	ns
Calorie Intake (kcal/day)	0.000	0.694	ns
Cardiometabolic Status Healthy	0.041	0.630	ns
City (Bogota)	0.479	0.000	***
City (Bucaramanga)	0.047	0.715	ns
City (Cali)	-0.227	0.118	ns
City (Medellin)	0.311	0.016	*
Diastolic Blood Pressure (mmHg)	0.000	0.987	ns
Fiber (g/day)	0.012	0.132	ns
Glucose (mg/dL)	0.001	0.771	ns
Hemoglobin a1c (%)	0.071	0.258	ns
C-reactive protein (mg/L)	-0.011	0.279	ns
Insulin (IU/mL)	0.000	0.928	ns
Latitude (°)	-0.017	0.303	ns
Total Cholesterol (mg/dL)	0.001	0.357	ns
High-density lipoprotein (mg/dL)	0.004	0.215	ns
Low-density lipoprotein (mg/dL)	0.001	0.303	ns
Very low density lipoprotein (mg/dL)	-0.001	0.637	ns
Triglycerides (mg/dL)	0.000	0.526	ns
Medication (Yes)[b]	0.014	0.871	ns
Carbohydrates (% Daily Value)[c]	0.003	0.802	ns
Total Protein (% Daily Value)[c]	0.008	0.779	ns
Total Fat (% Daily Value)[c]	-0.014	0.419	ns
Animal Protein (% Daily Value)[c]	-0.016	0.033	*
Monounsaturated Fat (% Daily Value)[c]	-0.031	0.474	ns
Polyunsaturated Fat (% Daily Value)[c]	-0.066	0.230	ns
Saturated Fat (% Daily Value)[c]	-0.005	0.855	ns
Sex (Male)[d]	-0.092	0.271	ns
Smoker (Yes)[e]	-0.074	0.562	ns
Stool Consistency Hard	0.009	0.968	ns
Stool Consistency Normal	0.374	0.127	ns
Stool Consistency Soft	0.137	0.715	ns
Systolic Blood Pressure (mmHg)	-0.001	0.641	ns
Metabolic Equivalent of Time (min/week)	0.000	0.774	ns
Waist Circumference (cm)	-0.001	0.706	ns

[a] done in reference to the age group 41-62
[b] done in reference to "No" to medications
[c] % of daily calorie consumption
[d] done in reference to "Female"
[e] done in reference to "No" to smoking

Beta diversity analysis was also performed on all 35 metadata categories using Unweighted Unifrac as the distance metric. The results showed that 15 unique predictors significantly influenced microbial composition (Table 2). Among those 15, some predictors were found to be closely linked to obesity, namely, "body mass index (BMI)" (PERMANOVA, 0.012), "Calorie Intake" (PERMANOVA, 0.047), and "Waist Circumference" (PERMANOVA, 0.003). Another subset of significant predictors, such as "Blood Pressure" (PERMANOVA, 0.002), "Fiber Intake" (PERMANOVA, 0.002), and "Medications" (PERMANOVA, 0.001), was related to general health status. Furthermore, two of the 15 significant predictors, "City" (PERMANOVA, 0.000) and "Latitude" (PERMANOVA, 0.000), were location-related. Because of the observed association between alpha and beta diversity and obesity metrics, we next asked whether obesity itself has any effect on gut microbial composition. To test this, we constructed an obesity definition using globally accepted metrics and examined compositional differences. This contrasts with the above analysis, where metrics were looked at individually.

TABLE. 2 Metadata categories related to obesity, general health, or location showed significance in beta diversity. PERMANOVA (Unweighted Unifrac distance metric) was performed individually on all 35 predictors within the dataset. R2 values indicate the proportion of microbial composition explained by each predictor. Significant results are indicated by an asterisk (* = P<0.05, ** = P<0.01, *** = P<0.001, **** = P<0.0001). n = 442.

Metadata Category	R2	P-Value	Significance
Adiponectin (ng/mL)	0.003	0.100	ns
Age (years)	0.003	0.207	ns
Age range	0.003	0.076	ns
BMI (kg/m^2)	0.004	0.012	*
BMI Class	0.007	0.022	*
Body Fat Percentage (%)	0.003	0.313	ns
Calorie Intake (kcal/day)	0.003	0.047	*
Cardiometabolic Status	0.003	0.122	ns
City	0.023	0.000	****
Diastolic Blood Pressure (mmHg)	0.004	0.003	**
Fiber (g/day)	0.004	0.002	**
Glucose (mg/dL)	0.003	0.555	ns
Hemoglobin a1c (%)	0.003	0.208	ns
C-reactive protein (mg/L)	0.003	0.157	ns
Insulin (IU/mL)	0.002	0.905	ns
Latitude (degrees)	0.005	0.000	***
Total Cholesterol (mg/dL)	0.003	0.323	ns
High-density lipoprotein (mg/dL)	0.003	0.040	*
Low-density lipoprotein (mg/dL)	0.003	0.479	ns
Very low density lipoprotein (mg/dL)	0.003	0.020	*
Triglycerides (mg/dL)	0.004	0.015	*
Medication	0.005	0.001	***
Carbohydrates (% Daily Value)	0.003	0.158	ns
Total Protein (% Daily Value)	0.002	0.952	ns
Total Fat (% Daily Value)	0.003	0.059	ns
Animal Protein (% Daily Value)	0.004	0.002	**
Mono-unsaturated Fat (% Daily Value)	0.003	0.246	ns
Poly-unsaturated Fat (% Daily Value)	0.003	0.195	ns
Saturated Fat (% Daily Value)	0.003	0.146	ns
Sex	0.005	0.000	****
Smoker	0.003	0.085	ns
Stool Consistency	0.013	0.000	****
Systolic Blood Pressure (mmHg)	0.004	0.002	**
Metabolic equivalent of time (min/week)	0.003	0.485	ns
Waist Circumference (cm)	0.004	0.003	**

Microbial gut composition is not significantly affected by obesity status. Since obesity-related metrics were identified as potentially important predictors of both alpha and beta diversity in our previous screen, we next investigated whether beta diversity differed between obese and non-obese individuals. We filtered the dataset to include those who satisfied two obesity metrics (waist circumference, body fat percentage) or neither and performed a PERMANOVA. After filtering there were 267 individuals remaining from the initial 442 sample size. Of these 267 individuals, only 23 were non-obese). Results revealed no significant difference in microbial composition between obese and non-obese individuals (PERMANOVA, p=0.092) (Fig. 1A). Upon observing these results, we then examined whether core microbiome members differed between these two groups. Among the ASV's that were considered core post-filtering, 35% of them were associated only with obese or non-obese individuals (n=83) (Fig. 1B). The remaining 65% were associated with both groups (n=149).

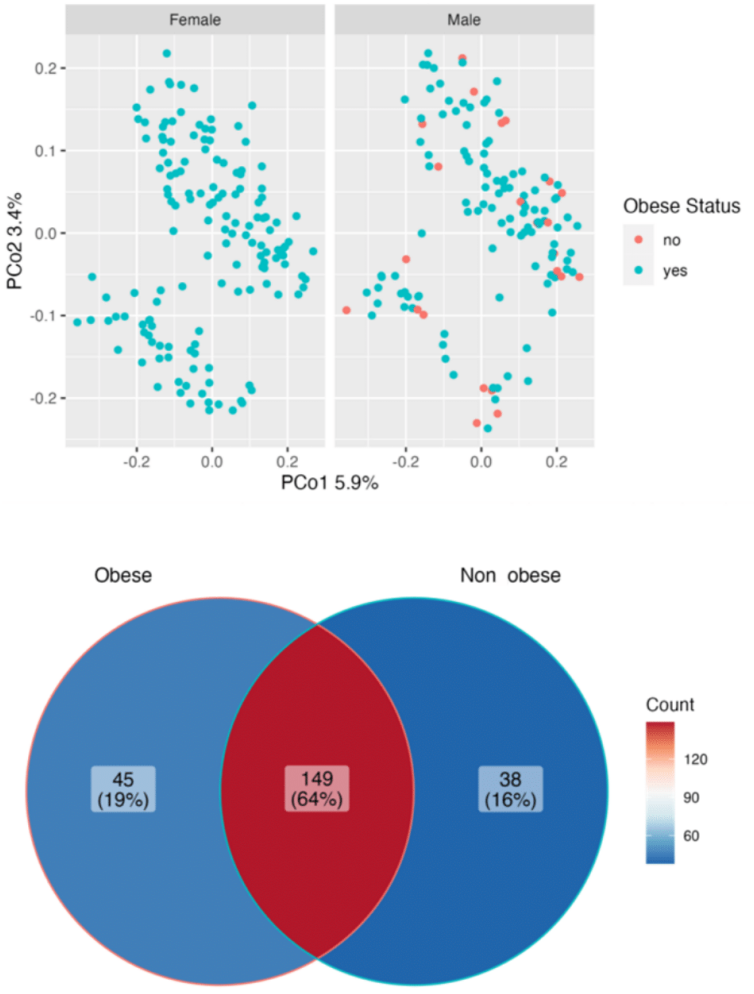

FIG. 1 Obesity does not significantly affect microbial gut composition. A) PCoA plot of obese vs. non-obese individuals (PERMANOVA, p=0.092) using Unweighted Unifrac distance metric (n=267). B) Core microbiome analysis on obese vs. non-obese individuals (prevalence = 0.10, abundance = 0.001) The count represents microbial species. n = 278.

Age, city, sex, and diet have significant effects on Shannon diversity in the final model. Model selection on a full linear model with 35 predictors (AICc=860.9) resulted in a final model with only nine predictors (AICc= -204.4). Out of the nine remaining predictors, four had significant effects on Shannon diversity ($p<0.05$). Another four were found to have no significant effect on Shannon diversity ($p>0.05$). Bogota city (done in reference to Barranquilla) had the largest increase in Shannon diversity (p=0.000), while Cali had the largest decrease (p=0.070) (Table 3). Age increased in Shannon diversity (p=0.013), while other significant predictors, like diet and sex (p= 0.011 for Carbohydrates (% Daily Value), p= 0.009 for Total Fat (% Daily Value), p= 0.001 for Animal Protein (% Daily Value) and p=0.016 for sex), decreased in Shannon diversity.

TABLE. 3 Model selection returned a combination of nine predictors that best explain variance in alpha-diversity. A linear model (response variable = Shannon Index) was performed on the full model of 35 predictors. Adjusted AIC (AICc) score for full model = 860.9. AICc score for the final model = -204.4. Estimates indicate the increase or decrease in Shannon diversity. For the following categorical variables: Cities are in reference to Barranquilla city, "Medication (Yes)" is in reference to "Medication (No)", "Sex (Male)" is in reference to "Sex (Female)". Significant results are indicated by an asterisk (* = P<0.05, ** = P<0.01, *** = P<0.001, **** = P<0.0001). n = 442.

Metadata Category	Estimate	P-Value	Significance
Age (Years)	0.011	0.013	*
Body Fat Percentage (%)	-0.017	0.102	ns
City (Bogota)	0.494	0.000	****
City (Bucaramanaga)	0.029	0.824	ns
City (Cali)	-0.245	0.070	ns
City (Medellin)	0.411	0.001	***
Insulin (μIU/mL)	0.008	0.130	ns
Medication (Yes)	-0.124	0.147	ns
Carbohydrates (% Daily Value)	-0.075	0.011	*
Total Fat (% Daily Value)	-0.089	0.009	**
Animal Protein (% Daily Value)	-0.029	0.001	***
Sex (Male)	-0.221	0.016	*

City, circulating/consumed fat, and stool consistency have significant effects on microbial community composition in the final model. The full PERMANOVA model with all 35 predictors (AICc = -337.1) was reduced to a model with only 6 predictors (AICc = 689.9) (Table 4). Stool consistency (R^2 = 0.037) represented the largest proportion of variation, while animal protein (R^2 = 0.005) represented the lowest. City (R = 0.029) was found to represent the second highest proportion of variation, while triglycerides (R = 0.008) and daily monounsaturated fat intake (R = 0.011) represented a lower percentage in variation. Of the six unique predictors, two were found to have no significant effect on microbial composition (R^2= 0.005 for Animal Protein (% Daily Value) and R^2= 0.007 for sex).

TABLE. 4 Model selection returned a combination of six predictors that best explain variation in beta-diversity. PERMANOVA (Weighted Unifrac distance metric) was performed on the full model of 35 predictors. AICc score for the full model = 689.9. AICc score for the reduced model = -337.1. R^2 values indicate the proportion of variation explained by each predictor. Significant results are indicated by an asterisk (* = P<0.05, ** = P<0.01, *** = P<0.001). n = 442.

Metadata Category	R^2	P-Value	Significance
City	0.029	0.002	**
Triglycerides (mg/dL)	0.008	0.039	*
Animal Protein (% Daily Value)	0.005	0.137	ns
Monounsaturated fat (% Daily Value)	0.011	0.012	*
Sex	0.007	0.054	ns
Stool Consistency	0.037	0.001	***

DISCUSSION

Obesity-linked predictors significantly affected diversity and composition. Alpha and beta diversity analysis of all 35 metadata categories revealed a group of significant predictors with underlying links to obesity (Table 1, 2). The predictor "Adiponectin Levels" significantly affected diversity, while the predictors "Calorie Intake", "BMI", and "Waist Circumference" significantly affected composition. Low adiponectin levels have previously been found to be associated with being overweight/obese and the emergence of non-alcoholic fatty liver disease (NAFLD) (19), both conditions related to excess adipose tissue. We

observed that individuals with lower adiponectin levels had more variable diversity (Fig. S1 Panel D), while individuals with high adiponectin levels had relatively high diversity with less variation. It is speculated that the group with low adiponectin levels may encompass both individuals and individuals who have the conditions of obesity/NAFLD, which may increase variance in microbiome metrics. Furthermore, higher values for BMI, caloric intake, and waist circumference have widely established links to obesity (19, 20). The PCoA plots for these three metrics (Fig. S2 Panels A, C, and N) indicated slight clustering of the data points with higher values. This suggests that individuals with high values in these categories share similar gut microbiome compositions, and also that these points may represent individuals who are obese. The condition of obesity may result in changes to the gut microbiomes of these individuals, making them more compositionally similar.

Not all obesity-related metrics significantly affected gut microbial community diversity and composition. For example, body fat percentage did not significantly influence diversity and composition (Fig. S1 Panel C). However, other obesity-linked metrics did have significant effects (BMI, caloric intake, and waist circumference). Therefore the results may be explained by obesity itself not contributing to microbial composition significantly, but rather other underlying obesity-linked factors might be responsible for shifts in the microbiome.

Previous literature has linked obesity and obesity-related metrics with gut microbial variation, but these studies have mostly been predominantly focused on WEIRD populations (21). Past literature has revealed that non-WEIRD populations have increased diversity and altered composition of the gut microbiome in comparison to WEIRD populations (22, 23). Specifically, these differences have been shown between Latin American populations (which includes Colombians) and North American populations from metropolis areas (22, 23). These differences may be attributed to dietary disparities among the two populations (23), as well as different levels of exposure to bacterial species in water supplies due to the reduced level of safety management in Latin America (24). Together, these factors may explain the differences observed in the above results. Although previous literature has shown WEIRD populations having links between obesity and both alpha and beta diversity, the aforementioned differences between WEIRD and non-WEIRD populations suggests these exact same links should not be expected in both populations. This could explain why body fat percentage was found to be an insignificant predictor of beta diversity in our dataset, while other studies on WEIRD populations have shown strong correlations between body fat and gut microbial composition (25).

There are no significant differences in microbial gut composition between obese and non-obese individuals. Beta diversity analysis on individuals filtered into obese and non-obese categories revealed no significant compositional differences between the two groups (Fig. 1A). These findings diverge from most current literature (26). Various studies have found significant correlations between obesity status and microbial gut composition, but most studies exploring gut microbiome and obesity have been done on WEIRD societies (27, 28) Differences in microbiome variation between ethnic groups are known to exist, and Colombian gut microbiota, in particular, differ from North Americans, Europeans, and Asians (27, 28). Therefore, patterns found linking microbial gut composition and obesity within studies on WEIRD populations should be cautiously applied to other populations, such as in Colombian populations.

Our dataset also included statistical limitations that may have obscured differences in gut microbiome composition between obese and non-obese individuals. Specifically, there are disproportionately more obese than non-obese individuals in the data post-filtering, and the absence of non-obese females (Fig. 1A). These factors may be heavily skewing the results resulting in a non-representative sample. This bias could contribute to the data that contradicts the current literature.

A portion of microbial gut species differs between obese and non-obese individuals. Core microbiome analysis revealed that 35% of the pooled collection of the core microbiome species were associated with either “obese” or “non-obese” individuals (Fig. 1B). This implies that gut microbial composition may be changing based on obesity status. Indicator species analysis revealed significantly higher association of certain bacterial groups with non–obese individuals. For instance, one of the bacterial familiess associated with non-obese

individuals is Rikenellaceae. Rikenellaceae has been previously found in literature to be associated with reduced adipose tissue in non-obese individuals (29). Notably, in contrast to the previous analysis where not all results were found to be consistent with WEIRD populations, in this portion of analysis, all trends observed were consistent with literature on WEIRD populations.

Model selection showed that obesity metrics do not play a large role in the variation of microbial diversity and composition. The final Shannon-diversity model included "Body Fat Percentage", which did not significantly affect microbial diversity. However, it may have been included in the final model due to improving the explanatory power of another predictor. Further analysis revealed that filtering out "Insulin" and "Medication" causes "Body Fat Percentage" to be removed from the final models, suggesting that these predictors share a dependent relationship regarding variation in Shannon diversity (Tables S1-S4). Many medications, such as insulin, can cause weight gain by influencing the redistribution of body fat, leading to an increase in visceral fat accumulation and a decrease in subcutaneous fat (30). Insulin also has an impact on appetite, and can cause diabetic patients to increase their total caloric intake to compensate for low blood glucose levels (31). There is also positive feedback between obesity and insulin insensitivity: visceral fat deposits help build up insulin resistance, meaning obesity may lead to requiring higher doses of insulin (32). Higher insulin intake and obesity have both independently been linked to lower microbial diversity in the gut. However, the relative contribution and interaction between these two health metrics on gut microbiome variation have not been well studied (33, 34). Here, we show that these two health metrics may interact with each other to alter gut microbiome diversity.

Although obesity metrics did not have a significant impact on microbial variation, the impact of diet-related metrics was more profound. The consumption of more "Animal Protein" was an important predictor for both diversity and composition (Tables 3, 4, S1-S4). Populations in WEIRD countries that consume a diet rich in animal protein, carbohydrates, and fat were found to have lower microbial richness and biodiversity. In contrast, microbial richness and biodiversity in non-WEIRD populations, whose diets consist of low amounts of fat and animal protein and high levels of plant protein and fiber, was higher (35, 36). Microbial composition was also found to differ significantly between non-WEIRD populations themselves based on differences in diet (35, 37). Thus, the presence of diet-related metrics, such as "Animal protein", appearing in the final alpha and beta models outlines the significant impact diet has on microbial variation. Despite this, "Body Fat Percentage" appears to be independent of diet, suggesting that obesity does not have a direct association with diet in terms of its influence on the microbiome (Table S1).

Apart from health-related metrics, the non-health related metric "City" is also a strong predictor of both alpha and beta-diversity (Tables 3, 4, S1-S4). The alpha diversity model showed the cities of Bogota and Medellin to have a significant effect on Shannon diversity. Geographic location was previously found to strongly influence microbial diversity , and here we show that geographic location may also explain variation in community composition. This may be explained by the fact that Bogota and Medellin are the two most highly populated cities in Colombia, with populations of 7.4 million and 4 million people, respectively (38, 39). The higher populations of these cities may give residents more economic influence, which could translate to increased access to diverse foods and diets (40). This may result in higher gut microbial diversity in individuals living in these locations (41). Additionally, higher populated cities are often more polluted (42). This affects the exposure of individuals in these cities to the bacterial species found within food supplies (41). Considering how Colombians from different cities are exposed to different microbial species, it is possible that the composition of their gut microbiomes reflects this difference.

Limitations Two major factors pertaining to the sample set contributed to the limitations of this study. Firstly, the dataset does not represent the entire Colombian population, which limits the generalizability of our study. Secondly, the absence of non-obese females and the greater number of obese samples compared to non-obese may lead to uneven results (Fig. 1A). Additionally, this was a cross-sectional study, meaning we are unable to infer causality between obesity and gut microbial variation. This provides only a snapshot of the dynamics between the gut microbiome and the human body. Studies done over a longer time period

would be needed to better understand causal links with obesity. Lastly, using only 16S rRNA gene sequencing identifies bacterial species and provides an incomplete view of the complex human gut microbial community which also includes prokaryotes, eukaryotes, archaea, and viruses (5).

Conclusions This study investigated the associations between obesity-related factors with gut microbial diversity and composition in a Colombian population. Our results revealed that predictors with underlying links to obesity significantly affected microbial diversity and composition. While no significant differences in microbial gut composition were observed between obese and non-obese individuals, core microbiome analysis revealed that some microbial species changed between obese and non-obese individuals. Looking beyond single metrics and analyzing combinations of factors with model selection, we found obesity metrics were not included in the optimal combination of predictors that explain the variation in microbial diversity and composition. Although body fat percentage was included in the final alpha diversity model, it was not a significant predictor. However, it was shown that "city" may be an important factor in determining alpha and beta diversity of the gut microbiome. This may be due to differences in diet and exposure to different microbes in different cities.

Future Directions To better understand the relationship between obesity and gut composition in Colombian populations, it is imperative to obtain an unbiased dataset, not only with a larger number of non-obese individuals, but also with at least some non-obese females. This ensures that observed patterns are a true representation of the population. Additionally, obese individuals had 19% of all ASVs that were found to be part of the core microbiome for any group (pooled) were uniquely associated with them. For future studies, identifying the metabolic pathways of these specific species within the gut environment may provide a deeper understanding of the relationship between obesity and gut microbial composition. Finally, in addition to our current study, conducting more research on non-WEIRD populations, particularly South American populations, is necessary to obtain globally generalizable observations on the interplay between obesity and gut variation.

ACKNOWLEDGEMENTS

We acknowledge Dr. Melissa Chen, whose support and guidance were instrumental in successfully completing this manuscript and project, and Dr Mihai Cirstea for his work in curating all the metadata files for the MICB 475. We would also like to sincerely thank Dr. Evelyn Sun for her invaluable suggestions for revisions and instruction on R coding. We would also like to acknowledge the Department of Microbiology and Immunology at the University of British Columbia for providing the opportunity of this course-based undergraduate research experience. Lastly, we would like to acknowledge BioProject and Cuesta-Zuluaga J, Corrales-Agudelo V, Velásquez-Mejía EP, Carmona JA, Abad JM, and Escobar JS where we obtained the dataset and got inspiration for our study. We would also like to thank two anonymous reviewers for constructive feedback on this manuscript.

CONTRIBUTIONS

All members equally contributed to the devising of the research topic and corresponding questions. G.S. contributed to the majority of coding for aims 2 and 3, the majority of the writing of the abstract and discussion, and also contributed to writing the introduction, methods, results sections, and supplemental. S.S also contributed to the majority of the coding for aims 2 and 3, the majority of the writing for the introduction, references, discussion and also contributed to writing the abstract, methods, results sections, and supplemental. J.Y.J. performed all initial data analysis within QIIME2, and also contributed to writing the methods, and the majority of the limitations, and conclusion sections. J.W. contributed to the majority of coding for aim 1, and also contributed in coding aim 3, writing the methods, results, and discussion.

REFERENCES

1. Durack J, Lynch SV. 2019. The gut microbiome: Relationships with disease and opportunities for therapy. J Exp Med 216:20–40.
2. Vangay P, Johnson AJ, Ward TL, Al-Ghalith GA, Shields-Cutler RR, Hillmann BM, Lucas SK, Beura LK, Thompson EA, Till LM, Batres R, Paw B, Pergament SL, Saenyakul P, Xiong M,

Kim AD, Kim G, Masopust D, Martens EC, Angkurawaranon C, McGready R, Kashyap PC, Culhane-Pera KA, Knights D. 2018. US Immigration Westernizes the Human Gut Microbiome. Cell 175:962-972.e10.

3. Jumpertz R, Le DS, Turnbaugh PJ, Trinidad C, Bogardus C, Gordon JI, Krakoff J. 2011. Energy-balance studies reveal associations between gut microbes, caloric load, and nutrient absorption in humans. Am J Clin Nutr 94:58–65.
4. Clarke SF, Murphy EF, O'Sullivan O, Lucey AJ, Humphreys M, Hogan A, Hayes P, O'Reilly M, Jeffery IB, Wood-Martin R, Kerins DM, Quigley E, Ross RP, O'Toole PW, Molloy MG, Falvey E, Shanahan F, Cotter PD. 2023. Exercise and associated dietary extremes impact on gut microbial diversity | Gut. https://gut.bmj.com/content/63/12/1913.long..
5. Symonds MRE, Moussalli A. 2011. A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. Behav Ecol Sociobiol 65:13–21.
6. 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.
7. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet C, 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 E, Silva RD, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley G, Janssen S, Jarmusch AK, Jiang L, Kaehler B, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MG, 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 J, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Michael S Robeson II, 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, Hooft JJ van der, Vargas F, Vázquez-Baeza Y, Vogtmann E, Hippel M von, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CH, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG. 2018. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. e27295v2. PeerJ Inc.
8. Callahan BJ, McMurdie PJ, Rosen MJ, et al. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods 13:581–583. doi: 10.1038/nmeth.3869.
9. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glöckner FO. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 35:7188–7196.
10. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596.
11. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA, Caporaso JG. 2018. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. Microbiome 6(90). https://doi.org/10.1186/s40168-018-0470-z
12. q2d2. 2023. Qiime2 q2-feature-classifier https://github.com/qiime2/q2-feature-classifier
13. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30(4):772-80. Doi: https://pubmed.ncbi.nlm.nih.gov/23329690/.
14. Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, Lozupone C, Zaneveld JR, Vázquez-Baeza Y, Birmingham A, Hyde ER, Knight R. 2023. Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome 5(27) https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-017-0237-y.
15. McMurdie PJ, Holmes S. 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLOS ONE 8:e61217.
16. Lahti L, Shetty S, Blake T, Salojarvi J. 2017. Microbiome R package.
17. Gao C-H, Yu G, Dusa A. 2022. ggVennDiagram: A "ggplot2" Implement of Venn Diagram (1.2.2).
18. Modern Applied Statistics with S, 4th ed. https://www.stats.ox.ac.uk/pub/MASS4/. Retrieved 24 April 2023.
19. Ladabaum U, Mannalithara A, Myer PA, Singh G. 2014. Obesity, Abdominal Obesity, Physical Activity, and Caloric Intake in US Adults: 1988 to 2010. Am J Med 127:717-727.e12.
20. Mozaffarian D, Hao T, Rimm EB, Willett WC, Hu FB. 2011. Changes in diet and lifestyle and long-term weight gain in women and men. N Engl J Med 364:2392–2404.
21. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP, Egholm M, Henrissat B, Heath AC, Knight R, Gordon JI. 2009. A core gut microbiome in obese and lean twins. 7228. Nature 457:480–484.
22. Obregon-Tito AJ, Tito RY, Metcalf J, Sankaranarayanan K, Clemente JC, Ursell LK, Zech Xu Z, Van Treuren W, Knight R, Gaffney PM, Spicer P, Lawson P, Marin-Reyes L, Trujillo-Villarroel O, Foster M, Guija-Poma E, Troncoso-Corzo L, Warinner C, Ozga AT, Lewis CM.

2015. Subsistence strategies in traditional societies distinguish gut microbiomes. 1. Nat Commun 6:6505.

23. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP, Heath AC, Warner B, Reeder J, Kuczynski J, Caporaso JG, Lozupone CA, Lauber C, Clemente JC, Knights D, Knight R, Gordon JI. 2012. Human gut microbiome viewed across age and geography. 7402. Nature 486:222–227.
24. Hassfurter K. 2017. Progress on Drinking Water, Sanitation and Hygiene: 2017 update and SDG baselines. UNICEF DATA. https://data.unicef.org/resources/progress-drinking-water-sanitation-hygiene-2017-update-sdg-baselines/. Retrieved 24 April 2023.
25. Deurenberg P, Weststrate JA, Seidell JC. 1991. Body mass index as a measure of body fatness: age- and sex-specific prediction formulas. Br J Nutr 65:105–114.
26. Shen J, Obin MS, Zhao L. 2013. The gut microbiota, obesity and insulin resistance, Molecular Aspects of Medicine 34(1):39-58. https://doi.org/10.1016/j.mam.2012.11.001.
27. Borrello K, Lim U, Park S-Y, Monroe KR, Maskarinec G, Boushey CJ, Wilkens LR, Randolph TW, Le Marchand L, Hullar MA, Lampe JW. 2022. Dietary Intake Mediates Ethnic Differences in Gut Microbial Composition. Nutrients 14:660.
28. Bracewell SA, Graeme FC, Johnston EL. 2018. Habitat Complexity Effects on Diversity and Abundance Differ with Latitude: An Experimental Study over 20 Degrees. Ecology 99(9):1964–74. https://www.jstor.org/stable/26626198.
29. Tavella T, Rampelli S, Guidarelli G, Bazzocchi A, Gasperini C, Pujos-Guillot E, Comte B, Barone M, Biagi E, Candela M, Nicoletti C, Kadi F, Battista G, Salvioli S, O'Toole PW, Franceschi C, Brigidi P, Turroni S, Santoro A. Elevated gut microbiome abundance of Christensenellaceae, Porphyromonadaceae and Rikenellaceae is associated with reduced visceral adipose tissue and healthier metabolic profile in Italian elderly. Gut Microbes 13:1880221.
30. Verhaegen AA, Van Gaal LF. 2000. Drugs That Affect Body Weight, Body Fat Distribution, and Metabolism, p. . In Feingold, KR, Anawalt, B, Blackman, MR, Boyce, A, Chrousos, G, Corpas, E, de Herder, WW, Dhatariya, K, Dungan, K, Hofland, J, Kalra, S, Kaltsas, G, Kapoor, N, Koch, C, Kopp, P, Korbonits, M, Kovacs, CS, Kuohung, W, Laferrère, B, Levy, M, McGee, EA, McLachlan, R, New, M, Purnell, J, Sahay, R, Singer, F, Sperling, MA, Stratakis, CA, Trence, DL, Wilson, DP (eds.), Endotext. MDText.com, Inc., South Dartmouth (MA).
31. Jones R. 2007. Insulin-associated weight gain in diabetes – causes, effects and coping strategies.Diabetes, Obesity and Metabolism - Wiley Online Library. https://dom-pubs.onlinelibrary.wiley.com/doi/full/10.1111/j.1463-1326.2006.00686.x?saml_referrer.
32. Wondmkun YT. 2020. Obesity, Insulin Resistance, and Type 2 Diabetes: Associations and Therapeutic Implications. Diabetes Metab Syndr Obes Targets Ther 13:3611–3616.
33. Pinart M, Dötsch A, Schlicht K, Laudes M, Bouwman J, Forslund SK, Pischon T, Nimptsch K. 2021. Gut Microbiome Composition in Obese and Non-Obese Persons: A Systematic Review and Meta-Analysis. Nutrients 14:12.
34. Chen Z, Radjabzadeh D, Chen L, Kurilshikov A, Kavousi M, Ahmadizar F, Ikram MA, Uitterlinden AG, Zhernakova A, Fu J, Kraaij R, Voortman T. 2021. Association of Insulin Resistance and Type 2 Diabetes With Gut Microbial Diversity: A Microbiome-Wide Analysis From Population Studies. JAMA Netw Open 4:e2118811.
35. Lin A, Bik EM, Costello EK, Dethlefsen L, Haque R, Relman DA, Singh U. 2013. Distinct Distal Gut Microbiome Diversity and Composition in Healthy Children from Bangladesh and the United States. PLoS ONE 8:e53838.
36. De Filippo C, Cavalieri D, Di Paola M, Ramazzotti M, Poullet JB, Massart S, Collini S, Pieraccini G, Lionetti P. 2010. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proc Natl Acad Sci U S A 107:14691–14696.
37. Jain A, Li XH, Chen WN. 2018. Similarities and differences in gut microbiome composition correlate with dietary patterns of Indian and Chinese adults. AMB Express 8:104.
38. Guzman L, Bocarejo JP. 2017. Urban form and spatial urban equity in Bogota, Colombia. Transportation Research Procedia 25:4495-4510. doi: 10.1016/j.trpro.2017.05.345.
39. Macrotrends LLC. 2023. Medellin, Colombia Metro Area Population 1950-2023. https://www.macrotrends.net/cities/20827/medellin/population.
40. Hilmers A, Hilmers DC, Dave J. 2012. Neighborhood Disparities in Access to Healthy Foods and Their Effects on Environmental Justice. Am J Public Health 102:1644–1654.
41. Singh RK, Chang H-W, Yan D, Lee KM, Ucmak D, Wong K, Abrouk M, Farahnik B, Nakamura M, Zhu TH, Bhutani T, Liao W. 2017. Influence of diet on the gut microbiome and implications for human health. J Transl Med 15:73.
42. Health Effects Institute. 2022. Air Quality and Health in Cities. https://www.stateofglobalair.org/sites/default/files/documents/2022-08/2022-soga-cities-report.pdf