Undergraduate Journal of Experimental Microbiology and Immunology (UJEMI) Vol. 29:1-15

Colombian Prediabetic Patients Classified Using Glycosylated Hemoglobin or Fasting Plasma Glucose Present Distinct Microbiome Compositions

Loujain Bilal, Ayesha Lalani, Parvin Malhi, Maya Ruehlen, Aiden Simard

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

SUMMARY Prediabetes is the precursor condition to Type II diabetes (T2D), a chronic metabolic disorder characterized by insulin resistance and insufficient uptake of glucose from the bloodstream. Current diagnostic methods glycosylated hemoglobin (HbA1c) and fasting plasma glucose (FPG) are the most widely used clinical tools yet can be inconsistent in determining prediabetic or T2D status. While current research highlights the association between gut microbiome composition and T2D, and to a lesser extent prediabetes, the difference between the microbiome composition of patients classified using HbA1c and FPG is currently unknown. To address this knowledge gap, we used 16S rRNA sequencing data collected from stool samples of a Colombian cohort, investigating the differences in the gut microbiome composition in patients classified using HbA1c compared to FPG. Although the alpha and beta diversity analysis revealed no significant differences on a community level, the core microbiome, differential abundance, and indicator taxa analyses displayed differences in gut microbial composition when comparing HbA1c and FPG prediabetic definitions with minimal overlap in identified taxa. Further, functional analysis suggested distinct metabolic profiles among HbA1c compared to FPG-classified prediabetes patients. Taken together, our study demonstrates that diagnosing prediabetes using the two diagnostic tools reveals distinct compositions of the gut microbiome, impacting both taxonomic and functional levels, underscoring the importance of employing multiple diagnostic tools in clinical practice to optimize prediabetes management.

INTRODUCTION

Type II diabetes (T2D) presents a growing public health challenge, impacting over 500 million individuals globally (1) incurring up to $10,000 in annual out-of-pocket expenses. (2). Projections indicate that by 2050, worldwide cases will surpass one billion, positioning T2D among the top 10 leading causes of global mortality and disability (1). This chronic metabolic disorder is marked by insulin resistance, leading to persistently elevated blood sugar levels (2). Prediabetes is an insulin-related condition preceding T2D, where blood sugar levels are higher than normal but not high enough to be considered diabetes (3). It is estimated by the Centers for Disease Control that 80% of individuals with prediabetes are unaware of their health condition (3). Those with diabetes or prediabetes face heightened risks of severe medical complications such as cardiovascular and renal diseases (4).

While the boundary between prediabetes and diabetes is somewhat subjective, clinicians generally define prediabetes based on either glycosylated hemoglobin levels (HbA1c) or fasting plasma glucose concentrations (FPG) (5). HbA1c refers to the percentage of hemoglobin in the bloodstream that is carrying a glucose molecule. It is a long-term diagnostic tool since it reports an individual's average blood glucose over a 3-month period (5). FPG refers to the actual concentration of glucose in the bloodstream following an 8 to 12-hour fast and is considered more of an acute diagnostic tool (5).

Previous research has shown that a patient's HbA1c and FPG often do not align, and that while each diagnostic method is viable for assessing high blood sugar to an extent, there are valid concerns surrounding their usage as diagnostic criteria for prediabetes (5). It is generally accepted that HbA1c at or above 5.7% is considered prediabetic trending towards diabetes (4). Likewise, per the World Health Organization, FPG at or above 6.1 mmol/L is considered

Published Online: September 2024

Citation: Bilal, Lalani, Malhi, Ruehlen, Simard. 2024. Colombian prediabetic patients classified using glycosylated hemoglobin or fasting plasma glucose present distinct microbiome compositions. UJEMI 29:1-15

Editor: Shruti Sandilya and Ronja Kothe, University of British Columbia

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

prediabetic towards diabetes (5). However, it has been reported by Ho-Pham et al., that HbA1c and FPG may be categorizing distinct groups of diabetic individuals given that notable differences exist in the classification of diabetes when comparing HbA1c and FPG measurements (6). Suggesting that these tools may identify distinct diabetic populations with varying microbial and physiological processes (6). With these discrepancies combined with the fact that both of these diagnostic tools are utilized in a clinical setting, we wanted to compare their predictive effectiveness by examining taxonomic and functional differences in the gut microbiome of prediabetic and healthy patients.

Microbes of the gut and their metabolites can have a strong influence on the host metabolism and consequently T2D (7, 8). As such, numerous studies have highlighted the importance of a healthy gut microbiota in managing and protecting against insulin resistance in the context of prediabetes and T2D (9, 10). Previous studies have established a significant association between the gut microbiome composition and the development and progression of T2D. A comprehensive review by Gurung et al. found the genera *Ruminococcus*, *Fusobacterium,* and *Blautia* were consistently increased in T2D populations while the genera, *Roseburia*, *Faecalibacterium, Bifidobacterium, Bacteroides,* and *Akkermansia* were decreased (11). Other studies describe a decrease in taxa involved in the production of butyrate, an important short-chain fatty acid (SCFA), which is involved in reducing inflammatory responses, and maintaining the integrity of the intestinal barrier (12, 13). Furthermore, T2D-specific taxa was observed in certain tissue in T2D individuals with severe obesity (14). Depleted Gram-positive bacteria and an increased abundance of opportunistic Gram-negative Enterobacteriaceae was found in adipose tissue (14). These studies highlight the potential pathogenic effects of certain taxa or microbiome alterations on the metabolic health of the host.

In contrast, certain bacterial taxa have been described to be protective against T2D including *Lactobacillus fermentum, plantarum* and *casei*, *Roseburia intestinalis, Akkermansia muciniphila* and *Bacteroides fragilis* which are all associated with improved glucose metabolism, insulin sensitivity and anti-inflammatory activity (15). *Bifidobacterium* and *Bacteroides* are the most consistently reported in the literature as potentially protective against T2D (11). Genera such as *Lactobacillus* can be challenging to analyze as the association with T2D is taxa-specific and studies have reported inconsistent findings (11). This can be partly explained by the use of medications as some diabetes drugs such as metformin have been described to alter the gut microbial composition (16). Overall, these studies highlight the complexity of the gut microbiome and its association with T2D while also presenting the potential for microbial-based T2D therapeutics.

Considering this background and knowledge regarding diagnostic tools and the gut microbiome in T2D, we sought to investigate how patients classified with prediabetes using HbA1c or FPG testing might differ in the context of their gut microbiome composition. No previous studies have been conducted to research this link, highlighting a prominent knowledge gap in the field of T2D and microbiome research. In this study, we investigated the differences in the gut microbiome composition of patients categorized using HbA1c compared to FPG. Given the inconsistencies between HbA1c and FPG tests to characterize T2D and prediabetes, we expect to observe a difference in the microbial composition as described by our diversity and taxa analyses, of prediabetic and T2D patients classified using either diagnostic tool.

METHODS AND MATERIALS

Dataset acquisition. The dataset used in this investigation was taken from a publication in *Scientific Report* by researchers de la Cuesta-Zuluaga *et al.,* titled, “Gut microbiota is associated with obesity and cardiometabolic disease in a population in the midst of Westernization” (17). Their study sought to investigate the gut microbiome composition of Colombian adults whose diets are undergoing westernization (17). Individuals that were sampled were located in a variety of Colombian cities, notably, Bogotá, Medellin, Cali, Barranquilla and Bucaramanga (17). The dataset consists of 441 Colombian adult stool samples (17). Within their accessible dataset, 16S rRNA sequencing data from the gut and T2D-related metadata were available (17). Microbial sequences were obtained using the 515F and 806R primers to isolate the V4 region of the 16S rRNA gene.

Metadata filtering and grouping. To investigate the differences between HbA1c and FPG diagnostic tools in relation to gut microbiome composition among a Colombian cohort, a metadata file was formatted and filtered in R (version 2023.12.1+402) (18) using the tidyverse package (19). Beginning with filtering, primary metadata categories HbA1c and FPG were retained, along with secondary variables including Body-Mass Index (BMI), High-Density Lipoprotein (HDL), Cholesterol, C-Reactive Protein (CRP), and Triglycerides as these have all been shown to be markers associated with T2D (20). Following filtering, two new metadata categories were generated to represent a prediabetes diagnosis using HbA1c (HbA1c_Prediabetic) or FPG (FPG_Prediabetic). The prediabetes thresholds for diagnosis were based on the industry standards: an HbA1c value of 5.7% or greater (4) and an FPG value of 6.1 mmol/L or greater was deemed prediabetic (5). Each metadata category was represented as a new column in the metadata file, with each sample receiving either "yes" if they met or surpassed the prediabetes threshold for either diagnostic marker or a "no" if they fell short of the set threshold. Metadata filtering and formatting had no impact on total study sample size as all 441 samples were retained. As noted in Table 1, 135 individuals were classified prediabetic using HbA1C, and 21 individuals were classified using FPG.

TABLE 1. Classifying prediabetes using HbA1C and FPG to define experimental cohorts. Individuals were classified as prediabetic using either HbA1c or FPG thresholds. An HbA1c value of 5.7% or greater and an FPG value of 6.1 mmol/L or greater was deemed prediabetic. A total of 441 participants were divided into either a Prediabetic or Healthy cohort in both the HbA1c and FPG condition.

Disease State	HbA1c Sample Sizes	FPG Sample Sizes
Prediabetic	135	21
Healthy	306	420

Data processing using the QIIME 2 pipeline. The sequence data was processed using the Quantitative Insights into Microbial Ecology version 2 (QIIME 2) (21) and details can be found in the supplemental QIIME 2 script. The raw single-ended sequence data was imported using a manifest file and demultiplexed using QIIME 2 (version 2023.7). Sequence quality control was performed using the Divisive Amplicon Denoising Algorithm 2 (DADA2) (22). Amplicon sequence variants (ASVs) were determined using a truncation length of 230 base pairs, which maintained a medium Phred quality score of 37. Subsequently, a feature table was generated, and mitochondrial and chloroplast sequences were filtered out to produce a filtered table. The ASVs were taxonomically classified using a Naive Bayes classifier (trained on truncated full-length sequences of 230 base pairs) (23) derived from the SILVA 138 99% database, specifically for the V4 region of the 16S rRNA gene (24). Amplification was performed using the 515F primer (GTGCCAGCMGCCGCGGTAA) and the 806R primer (GGACTACHVGGGTWTCTAAT). A rooted phylogenetic tree was constructed by aligning ASVs using Multiple Alignment Fast Fourier Transform (MAFFT) and FastTree (25, 26). The filtered feature table, taxonomy table, and rooted phylogenetic tree were exported from QIIME 2 to R for further analysis.

Alpha and beta diversity analysis using QIIME 2. To determine the community diversity between the prediabetic and non-prediabetic populations in both the HbA1c and FPG-classified cohorts, core alpha and beta metrics were generated using QIIME 2 (21). The samples were alpha-rarefied at a sampling depth of 26,493 to retain 8,689,704 (53.68%) features in 328 (74.38%) samples. The core metric files were generated using the QIIME 2 core phylogenetic diversity metrics tool (27). Files were exported to a local computer and visualized using QIIME 2 View and further analyses were performed on RStudio for plots of interest (28).

Alpha and beta diversity metrics and statistical analysis in RStudio. The QIIME 2 output files including the feature table, metadata, taxonomy, and phylogenetic tree were imported into RStudio (18). A phyloseq object was created using these files and the phyloseq package (29). Samples were filtered to remove non-bacterial sequences and exclude ASVs with less than 5 reads and samples with fewer than 100 reads. Alpha diversity was analyzed using Observed, Chao1 and Shannon indexes and pairwise Kruskal-Wallis tests were used to determine statistical significance using the phyloseq package (29). Bray-Curtis and weighted UniFrac indexes were used to assess beta diversity and the PERMANOVA (permutational multivariate ANOVA) test was used to determine statistical significance using the vegan package (30). The ggplot2 package was used to visualize the community diversity through box plots and Principal Coordinate Analysis (PCoA) plots (31).

Core microbiome analysis. Core microbiome analysis was performed in RStudio using the microbiome package (version 1.23.1) (32). Details can be found in the supplemental R script. Phyloseq object was created using the phyloseq package (version 1.46.0) (33). The detection thresholds were set to 0.0 and prevalence thresholds were set to 0.3 (34). A Venn diagram was created using ggVennDiagram (version 1.5.2) (35) to visualize shared and unique taxa between healthy and prediabetic cohorts using either the HbA1c or FPG diagnostic tools.

Differential abundance analysis. To identify the differences in the abundance of each taxa between both HbA1c and FPG prediabetic and healthy samples, differential abundance analysis (DAA) using the DESeq2 (36) package was conducted. First, the existing phyloseq object was transformed to remove the zeros and then converted to DESeq objects specific to HbA1c or FPG Prediabetic. The DESeq function was used to run the analysis using the healthy samples as references. The analysis output was graphed on a volcano plot showing the taxa that increased or decreased in abundance in the prediabetic samples compared to the healthy samples, highlighting those that were statistically significantly abundant ($P \leq 0.05$).

Indicator taxa analysis. To determine if there were species that were more prevalent and/or abundant in HbA1c and FPG prediabetic and healthy samples, indicator taxa analysis was run using the existing phyloseq object and the phyloseq (33) and indicspecies (37) packages in R. First, the phyloseq object was filtered as described in the alpha and beta analysis. The filtered phyloseq table was then converted to relative abundance. The multipatt function was used to cluster samples into HbA1c prediabetic and FPG prediabetic groups. A full list of indicator species classified down to the genus level for each prediabetic diagnosis definition group was generated using the summary command. Those identified as significant ($P \leq 0.05$) indicator taxa were formatted into a table.

PICRUSt2 analysis. The PICRUSt2 (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States 2) plugin tool in QIIME 2 was employed to predict the gut-microbiome-associated functional pathways (38). First, the SEPP tool in PICRUSt2 placed ASVs from 16S rRNA sequences onto a phylogenetic tree using reference phylogeny (38). The PIC statistical tool then predicted the community's functional potential by estimating the abundance of orthologous gene groups from the SEPP phylogenetic alignment (39). The resulting pathway abundance data was then imported into RStudio for further functional analysis using the ggpicrust2 package (40). Using a metadata file which subsets samples based on a prediabetes diagnosis (yes/no), a DAA of the pathway abundance data was conducted using the DESeq2 method to assess the statistical significance across prediabetic vs. healthy samples (41). Finally, MetaCyc pathways were annotated, and only statistically significant functional pathways were plotted on a Log2FoldChange bar graph.

RESULTS

Prediabetic status was not correlated with changes to the gut microbiome for patients categorized using HbA1c or FPG. To first establish a baseline for the microbial differences in patients with and without prediabetes, we conducted alpha and beta diversity analysis. As seen in Figures 1 and S1, the alpha diversity analyses conducted resulted in no significant differences in richness (Observed and Chao1) and abundance (Shannon) between

prediabetic and non-prediabetic patients in both HbA1c and FPG groups. Similarly, in Figure S2, beta diversity, Bray-Curtis and Weighted UniFrac indexes indicated no significant differences in abundance, richness, or phylogenetic distance between the two diagnostic measures. These findings suggest no differences in gut microbiota diversity on a community level in prediabetic and non-prediabetic patients in HbA1c and FPG-classified cohorts, leading us to investigate further by looking at individual taxonomic differences.

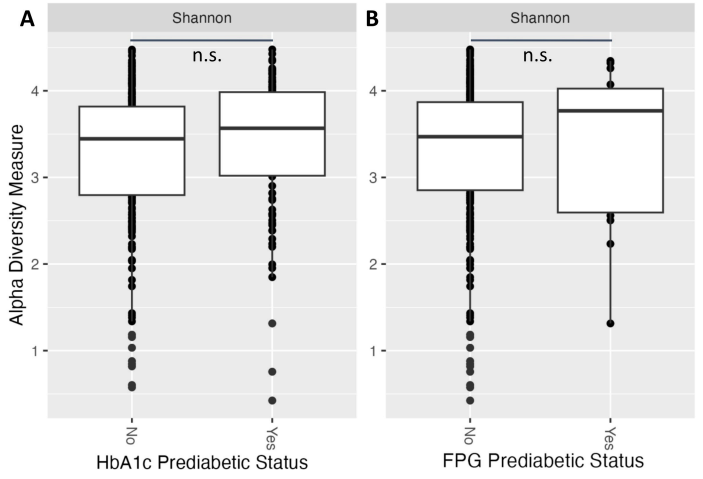
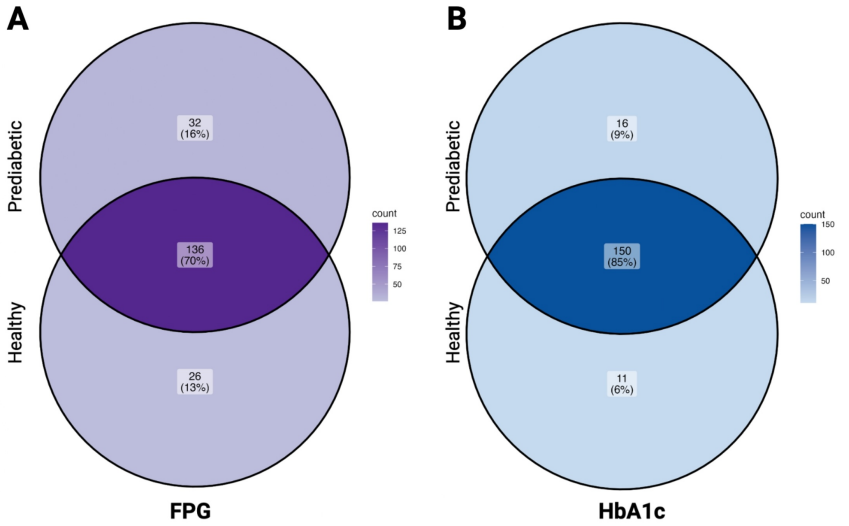


FIG. 1 Prediabetic status was not correlated with changes to the gut microbiome for patients classified using HbA1c or FPG. Used the alpha diversity metric, Shannon, to compare the microbial richness of prediabetic and non-prediabetic patients based on **(A)** HbA1c-classified patients, P = 0.083. **(B)** FPG-classified patients, P = 0.350. Significance was determined with the Kruskal-Wallis test using a threshold of $P \leq 0.05$. n.s. indicates no significant difference.

Core microbiomes of prediabetic and healthy patients show taxonomic differences when categorized using the different diagnostic tools. To determine which taxa were shared between healthy and prediabetic cohorts, as well as between the two diagnostic tools, HbA1c and FPG, a core microbiome analysis was conducted. The analysis determined that a large portion of the microbiome was shared between healthy and prediabetic patients when using either diagnostic tool (Figure 2). When comparing diagnostic tools, fewer ASVs were exclusive to the prediabetic cohort when using HbA1c, compared to using FPG to categorize. When analyzing ASVs common to these two exclusively prediabetic cohorts, there were only seven ASVs that overlapped when comparing diagnostic tools. This shows there are taxonomic differences within the core microbiome in prediabetic patients when FPG is used to classify prediabetes compared to HbA1c.



C

	FPG	HbA1c
Prediabetes Taxa	32	16
% Prediabetes	16	9
Overlap Taxa	7	

FIG. 2 Prediabetic patients exhibit taxonomic differences in core microbiome between when categorized using FPG compared to HbA1c. Core microbiome analysis was conducted on 441 individuals, 21 of which were classified as prediabetic using HbA1c and 135 of which were classified as prediabetic using FPG. Analysis was completed using a detection threshold of 0.0 and a prevalence threshold of 0.3. **(A)** 32 taxa exclusive to prediabetic microbiomes when using FPG to categorize. **(B)** 16 taxa exclusive to prediabetic microbiomes when using HbA1c to categorize. **(C)** Summary of the provided Venn diagrams. Taxonomic overlap of 7 ASVs when comparing the exclusively prediabetic cohorts.

FPG prediabetic diagnostic tool resolves a greater difference in the abundance of taxa between prediabetic and healthy samples than HbA1c. A differential abundance analysis was conducted to compare the differences in the abundance of each taxon between prediabetic and healthy samples for both diagnostic definitions (Figure 3). An alpha significance value of $P \leq 0.05$ was used to identify significantly differentially abundant ASVs. When using the HbA1c prediabetic diagnosis, 4 ASVs had decreased in abundance in prediabetic compared to healthy samples (Figure 3A). However, for the FPG prediabetic diagnosis, 78 ASVs decreased in abundance (Figure 3B). Therefore, the FPG prediabetic definition reveals a greater difference in the abundance of taxa in prediabetic compared to healthy samples as HbA1c. Furthermore, only 1 ASV, classified from the Genus *Clostridia UCG-014,* was found to be common among the significantly decreased taxa from both diagnosis definitions, emphasizing their distinct characteristics. Although both definitions are used to diagnose prediabetic patients, gut microbial abundance differences were seen when looking at patients who are prediabetic based on one definition compared to the other.

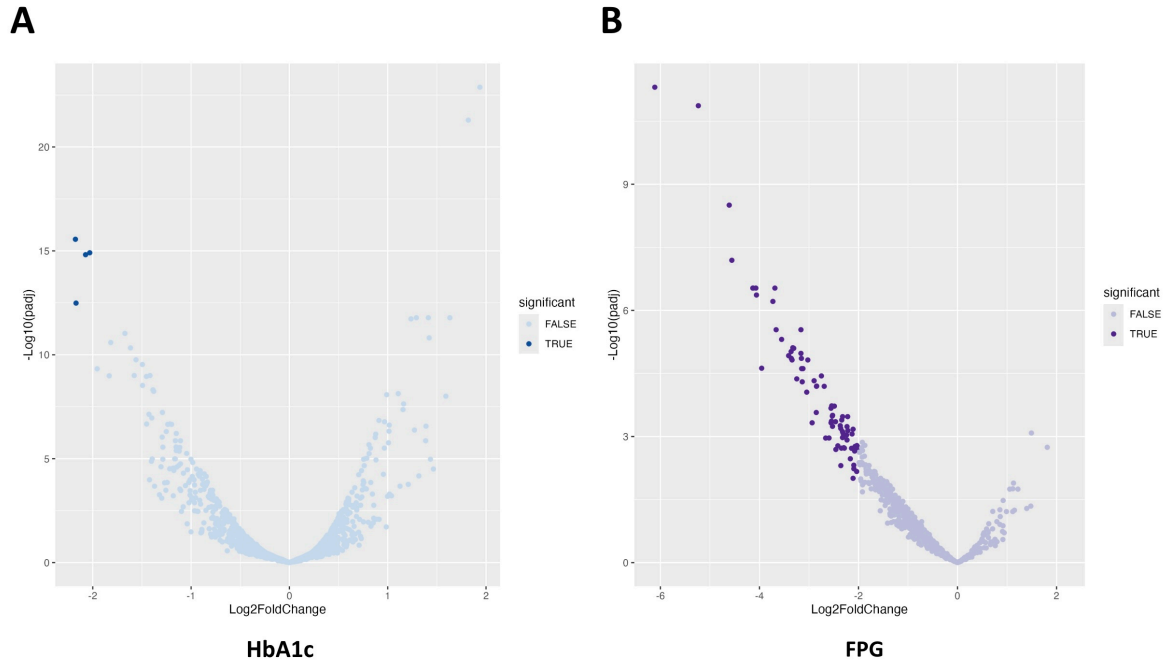


FIG. 3 FPG and HbA1c prediabetic classifications identified differences in the number of decreased abundance taxa in prediabetic vs. healthy samples. Differential abundance analysis compared taxonomic abundances between **(A)** HbA1c prediabetic and **(B)** FPG prediabetic to healthy samples. The x-axis indicates the Log2FoldChange of ASV abundance measures in prediabetic compared to healthy samples with the greater fold-change indicating either an increase in abundance (positive) or a decrease in abundance (negative). The y-axis is the -log10 of the adjusted *P* value, indicating whether the differential abundance was significant ($P \leq 0.05$). Darker blue and purple points represent the HbA1c and FPG significantly differentially abundant taxa (TRUE) respectively. In HbA1c prediabetic samples, 4 taxa significantly decreased compared to healthy, while in FPG prediabetic samples, 78 taxa significantly decreased compared to healthy.

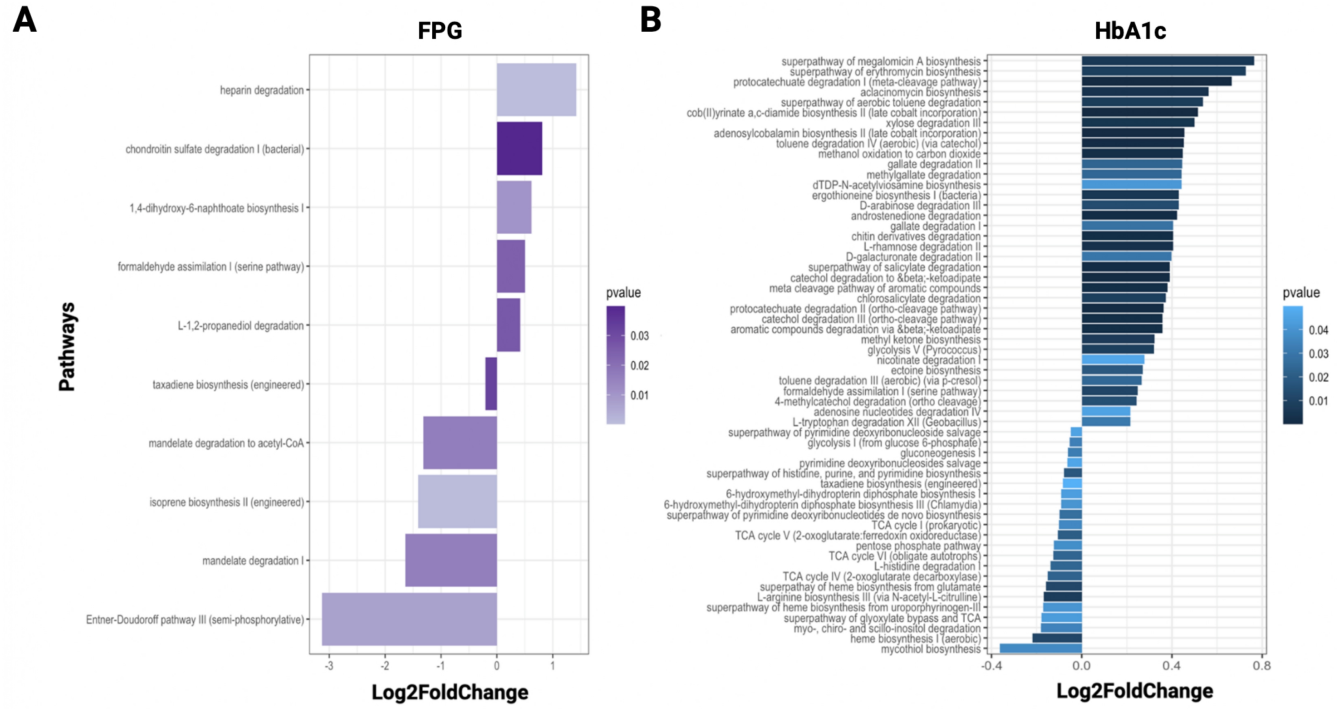
There are distinct indicator genera between HbA1c and FPG-classified prediabetic patients. Table 2 displays the statistically significant ($P \leq 0.05$) indicator ASVs classified by their phylum, family, and genus for HbA1c prediabetic (n=9), FPG prediabetic (n=7) and healthy samples (n=1). Comparing the indicator taxa among the prediabetic groups, no ASVs overlapped with each other. However, looking at the classification of the taxa, the phyla Bacteroidota, Proteobacteria and Firmicutes and the family *Lachnospiraceae* are seen in both definitions indicating potential commonalities in the prediabetic microbial gut composition at broader taxonomic levels. Narrowing down to the genus level, no overlap can be seen in the HbA1c and FPG prediabetic indicator taxa. This suggests that while there may be shared taxonomic phyla (Bacteroidota, Proteobacteria and Firmicutes), the specific genera within these groups differ between the two definitions of prediabetes. This could indicate subtle differences in the microbiome composition associated with two different prediabetic

definitions, FPG and HbA1c. Therefore, depending on the prediabetic definition, the specific microbial variants identified may vary.

TABLE 2. HbA1c and FPG prediabetic classification revealed unique indicator taxa for prediabetic and healthy samples, with no genera overlap. Indicator taxa analysis was conducted for HbA1c prediabetic, FPG prediabetic and healthy samples. Indicator taxa with *P* values below the set threshold of $P \leq 0.05$ are shown alongside their classified phyla, family and genera. Indicator value scores ASVs based on their abundance and prevalence in a particular group to assign association with higher indicator values representing high abundance and prevalence.

Phylum	Family	Genus	Indicator Value	P-Value
HbA1c Prediabetic				
Bacteroidota	*Prevotellaceae*	*Prevotellaceae UCG-003*	0.271	0.03
Bacteroidota	*Barnesiellaceae*	*Coprobacter*	0.295	0.035
Actinobacteriota	*Actinomycetaceae*	*F0332*	0.331	0.005
Elusimicrobiota	*Elusimicrobiaceae*	*Elusimicrobium*	0.292	0.025
Proteobacteria	*Acetobacteraceae*	*Roseomonas*	0.168	0.025
Proteobacteria	*Moraxellaceae*	*Acinetobacter*	0.259	0.02
Firmicutes	*Lachnospiraceae*	*Syntrophococcus*	0.264	0.005
Proteobacteria	*Aeromonadaceae*	*Aeromonas*	0.142	0.045
Proteobacteria	*Yersiniaceae*	*NA*	0.255	0.03
FPG Prediabetic				
Firmicutes	*Caldicoprobacteraceae*	*Caldicoprobacter*	0.217	0.045
Firmicutes	*Erysipelotrichaceae*	*NA*	0.571	0.005
Bacteroidota	*Marinifilaceae*	*Sanguibacteroides*	0.467	0.01
Proteobacteria	*Rhizobiaceae*	*Phyllobacterium*	0.296	0.01
Synergistota	*Synergistaceae*	*Cloacibacillus*	0.592	0.015
Firmicutes	*Peptostreptococcaceae*	*Paeniclostridium*	0.372	0.01
Firmicutes	*Lachnospiraceae*	*Bacteroides pectinophilus*	0.323	0.01
Healthy				
Proteobacteria	*Oxalobacteraceae*	*Oxalobacter*	0.258	0.03

Prediabetes patients characterized using the HbA1c diagnostic tool have more dysregulated microbial metabolic pathways than patients categorized with the FPG tool. Functional pathways of gut microbiomes of individuals categorized as prediabetic by either the HbA1c or FPG diagnostic tool were analyzed using PiCRUSt2 ASV alignment. Using SEPP and PIC tools, 16S rRNA amplicon sequencing data was aligned with reference phylogeny, and analyzed for abundance, respectively. MetaCyc pathway predictions were conducted, identifying five pathways that were upregulated and five that were downregulated when using FPG as the diagnostic tool (Figure 4A and 4C). Conversely, when using HbA1c to classify, there were 36 upregulated pathways and 22 downregulated ones (Figure 4B and 4C). Notably, HbA1c characterized patients display more microbial pathway alterations than the FPG cohort, demonstrating that patients classified using these tools further differ on a functional level in the context of the microbiome. Further, there were only two dysregulated pathways in common between the HbA1c and FPG diagnostic tools. Notably, formaldehyde assimilation I, which is a serine pathway, was upregulated in both the prediabetic cohort classified using FPG as well as HbA1c (Figure 4A and 4B). Another pathway, taxadiene biosynthesis is downregulated in both of these conditions (Figure 4A and 4B). This data reveals distinct pathway alterations in prediabetic patients when using different diagnostic criteria as shown by minimal overlap in pathway alterations and a drastic number of functional changes shown in HbA1c-categorized individuals compared to FPG-categorized individuals.



C

	FPG	HbA1c
Upregulated Pathway	5	36
Overlapping Upregulated Pathways	1	
Downregulated Pathways	5	22
Overlapping Downregulated Pathways	1	

FIG. 4 HbA1c and FPG diagnostic tools yield distinct microbial pathway dysregulation in prediabetes patients. PiCRUSt2 analysis was conducted to predict functional pathway abundances in patients characterized using the FPG and HbA1c tools. Pathways were deemed significant if they met a threshold of $P \leq 0.05$, and significantly upregulated and downregulated pathways were plotted as Log2FoldChange for **(A)** FGP-prediabetic samples and **(B)** HbA1c-prediabetes samples. **(C)** Summary of the altered functional pathways from FPG and HbA1c.

DISCUSSION

This study aimed to explore the differences in the gut microbiome composition among a Colombian cohort classified as prediabetic using HbA1c as compared to FPG, utilizing data collected by de la Cuesta-Zuluaga et al. (17).

Firstly, it was observed that there were no significant differences in community diversity between prediabetic and healthy individuals, irrespective of whether they were classified using HbA1c or FPG (Figure 1, Figure S1 and Figure S2). Specifically, no significant difference in Shannon diversity between prediabetic and healthy individuals was found (Figure 1A and 1B), which aligns with findings of previous research by Kitten et al., who also found no significant differences in Shannon diversity when comparing Mexican Americans with and without T2D (42). Additionally, the analysis of the Chao1 index and Observed Features showed no significant differences in alpha diversity between prediabetic and healthy individuals, regardless of the diagnostic tool used (Figure S1A and S1B). Similarly, Zhang et al. found no difference in Chao1 diversity between individuals with normal glucose tolerance and those who were prediabetic (43). Furthermore, beta diversity analysis using the Bray-Curtis and weighted UniFrac indexes demonstrated no significant differences between the prediabetic and healthy groups classified with either HbA1c or FPG (Figure S2A and S2B), consistent with previous literature (44, 45). These findings suggest that prediabetes, as classified by either HbA1c or FPG, does not appear to have a significant impact on the overall diversity of the gut microbiota in the Colombian cohort.

Despite observing no difference in the overall community gut microbiome diversity in prediabetic samples compared to healthy, other research has shown that individuals with T2D differ in specific gut microbial composition from non-diabetic ones and that the development of gut microbial dysbiosis could be a clinical manifestation of T2D (46). Therefore, to identify specific bacterial taxa associated with the two prediabetes definitions within our sample cohort, we conducted taxonomic analyses including the core microbiome, differential abundance, and indicator taxa analyses. Our core microbiome analysis revealed differences in the unique microbiome of prediabetic individuals when comparing the diagnostic tools. Seven ASVs were identified to be common among the prediabetic-only groups while 25 and 9 ASVs remained unique to FPG and HbA1c prediabetic samples respectively (Figure 2). In addition, the results of the indicator analysis highlighted that bacteria belonging to the phyla Firmicutes, Bacteroidota and Proteobacteria had associations with prediabetes (Table 2) which was further supported by our differential abundance analysis. Here, we observed that although there were no overlapping ASVs, bacteria of the same phyla, as in the indicator analysis, had significantly decreased in abundance compared to the healthy controls when using both HbA1c and FPG prediabetic definitions. These findings suggest that while overall gut microbiome diversity may not differ in prediabetic individuals compared to healthy controls, there are specific microbial taxa and patterns associated with prediabetes, which may vary depending on the diagnostic criteria used. This contributes to the understanding of the complex relationship between the gut microbiome and prediabetes.

Moreover, the common ASV between the decreased abundance ASVs in both definitions was classified as the class *Clostridia UCG-014* from the Firmicutes phylum. Findings from other researchers support this as Larsen et al. indicated that the proportions of phylum Firmicutes and class *Clostridia* were significantly reduced in the diabetic group compared to the control group (47). However, other studies, like Sedighi et al., contradict this and showcase an increasing abundance of Firmicute bacteria with diabetes (48). Both studies utilized different cohorts from different countries with varying factors that can affect the specific gut microbial differences like diet, lifestyle, genetics, and the parameters used to define prediabetes. Thus, these findings underscore the complexity and variability in gut microbial composition associated with prediabetes.

Studies have observed discrepancies in identifying prediabetic populations of the two diagnostic tools FPG and HbA1c. Ho-Pham et al. discussed a significant discordance in the diagnosis of diabetes between FPG and HbA1c measurements from blood samples of participants in Vietnam suggesting they may identify two separate prediabetic populations with varying microbial and physiological processes (6). Additionally, White et al. described the differences of both measures in the sensitivity to diagnose prediabetes across different cohorts with FPG appearing to underestimate the burden of undiagnosed diabetes. (49). This can be seen in the prediabetic sample sizes from the classification of our cohort. The FPG definition identified 21 prediabetic samples and HbA1c identified 135 samples. The discrepancy of the two definitions in identifying prediabetic samples and furthermore their unique core microbiome, indicator and differentially abundant taxa revealed from our analyses may be explained by the different physiological processes that the two definitions use to define prediabetes (6). FPG measures blood glucose at a point in time after no consumption of sugar, helping to identify a problem with the patient's sugar metabolism, leading to diabetes, and can be impacted by acute illnesses and stress (50). Whereas HbA1c looks at blood glucose levels over a longer period (2-3 months) and senses long-term signs of diabetes like obesity, heart conditions, high cholesterol, etc (50). Therefore, someone who is diagnosed as prediabetic using HbA1c may be unique physiologically and thus microbially to someone diagnosed with FPG. This highlights the importance of considering multiple diagnostic criteria and individualizing diagnostic approaches based on clinical judgment and risk factors. Further investigations into how those processes affect the gut microbiome are needed to fully understand the mechanisms.

Interestingly, we observed that a single ASV was found to be shared between the core microbiome and indicator taxa of FPG-classified prediabetic individuals (Table 2, Figure 2). The organism was classified as being a part of the *Erysipelotrichaceae* family. *Erysipelotrichaceae* is a family of bacteria that has been largely implicated in metabolic disorders particularly obesity (51). However, there is a growing body of research that suggests

that this family may be important in the context of T2D. Lippert et al. explored the relationship between the gut microbiome and various glucose metabolism disorders and found that there was a greater abundance of the *Erysipelotrichaceae* family in individuals with impaired glucose tolerance including individuals with T2D as compared to healthy individuals who had normal glucose tolerance (52). The researchers also found that the *Erysipelotrichaceae* family was shared in the core microbiome of individuals with pre-diabetes/diabetes, obesity, and metabolic syndrome (52). Similarly, in a study conducted by Graval et al. this family was found to have a significantly increased abundance in T2D vs healthy individuals (53). In contrast to these findings, in our study, we found that there was a non-significant decrease in the abundance of this family in pre-diabetic individuals as compared to healthy controls (Figure 3). Nevertheless, these inconsistencies may in part be explained by the relationship between microbes and T2D being species-specific (54). For instance, while the genus *Lactobacillus* as a whole has shown a positive association with T2D in some studies, it is important to note that individual species within the genus may have varying effects on health (54). Some species of *Lactobacillus* have been found to possess anti-inflammatory properties and may confer benefits to metabolic health (54). Therefore, the overall impact of *Lactobacillus* on T2D risk and metabolic health likely depends on the specific species present and their interactions with the host and other members of the gut microbiota (54). By extension, this variability in the effects of specific microbial taxa underscores the complexity of microbial associations with prediabetes and T2D. This may partially explain why we observed a decrease in the abundance of the *Erysipelotrichaceae* family in our study compared to other research, as taxonomic resolution to the genus and species level within this family is necessary to better understand its relationship with prediabetes and T2D.

In order to further resolve the taxa, we identified to be predictive of prediabetes based on FPG to the genera or species level, we utilized the Basic Local Alignment Search Tool (BLAST) to analyze the nucleotide sequence corresponding to that ASV (55). We found strong alignment with an unknown species within the *Anaerorhabdus* genus with a maximum score to be 416, a query coverage of 100%, an extremely low E value and 100% identity. To our knowledge, there is no current research that has explored the relationship between the *Anaerorhabdus* genus and pre-diabetes and therefore future studies could further characterize this genus and investigate the mechanisms by which this genus may be involved in prediabetes.

Furthermore, while using FPG as a diagnostic tool revealed more taxonomic differences than using HbA1c, utilizing HbA1c highlighted more functional differences in comparisons between prediabetic and healthy individuals. When FPG was used as a diagnostic tool for prediabetes, five pathways were found to be upregulated and five were downregulated (Figure 4A and 4C). Meanwhile, 36 upregulated pathways and 22 downregulated ones were found when HbA1c was used as a diagnostic tool (Figure 4B and 4C). This comparison indicates that when FPG is used, a smaller number of pathways are affected, while when HbA1c is used, a larger number of pathways are affected. This suggests that HbA1c might capture a broader spectrum of alternations in biological processes associated with prediabetes, as reflected in the microbiome. Perhaps this can be attributed to the fact that HbA1c reflects average blood glucose levels over three months as compared to FPG which provides a snapshot of glucose levels at a specific point in time (5). The longer-term perspective captured by HbA1c may allow for the detection of more subtle and sustained changes in metabolic processes, leading to a broader range of affected pathways.

Moreover, there was minimal overlap between the pathways that were altered in prediabetic as compared to healthy individuals when comparing between the two diagnostic tools. Specifically, only two altered pathways were common among the diagnostic tools, namely formaldehyde assimilation I was found to be upregulated in prediabetic patients and taxadiene biosynthesis was downregulated (Figure 4A and 4B). Formaldehyde assimilation I is a metabolic pathway responsible for converting formaldehyde into other compounds. Hipkiss et al. found that individuals with T2D have an increased production of formaldehyde, a reactive and potentially toxic compound that can induce cellular stress and damage (56). The upregulation of formaldehyde assimilation I that has been observed in prediabetic individuals may indicate a response to increased formaldehyde levels in related metabolic

pathways and serve as a way to prevent its accumulation. Conversely, the downregulation of taxadiene biosynthesis may reflect a prioritization of metabolic pathways, reallocating resources away from non-essential processes like secondary metabolite production, such as taxadiene biosynthesis, towards pathways crucial for managing the stress response provoked by prediabetes and T2D (57). The remaining pathways that were upregulated and downregulated when HbA1c was used as a diagnostic tool did not overlap with those that were altered in FPG-classified individuals (Figure 4C). Consequently, this highlights the potential of different diagnostic markers to reveal distinct aspects of prediabetes-related physiological changes in the microbiome.

Although both HbA1c and FPG serve as diagnostic markers for prediabetes in clinical practice, our research indicates that they influence taxonomic and functional changes within the microbiota in distinct ways. The FPG prediabetic diagnostic tool revealed more pronounced differences in taxa abundance between prediabetic and healthy samples compared to HbA1c. Conversely, patients characterized with HbA1c exhibited more dysregulated microbial metabolic pathways than those classified with FPG. These distinct microbial signatures associated with different diagnostic criteria underscore the need for tailored approaches in understanding the gut microbiome's role in metabolic disorders. Relying solely on one diagnostic marker may not offer a comprehensive understanding of the complex interplay between prediabetes and the microbiome. Our findings highlight the importance of considering a combination of diagnostic markers to more accurately assess an individual's prediabetic status and associated health risks. HbA1c reflects long-term glycemic control (5), providing insights into chronic metabolic dysregulation, while FPG offers immediate glucose levels (5) that can impact microbial dynamics in the short term. Combining these markers may allow clinicians to assess both the chronic and acute metabolic influences on the microbiota, providing a more holistic view of prediabetes-related changes. By incorporating multiple diagnostic markers, clinicians can gain a more comprehensive understanding of an individual's prediabetic status, stratify individuals based on their health risks, and develop personalized treatment approaches that address the specific metabolic and microbial imbalances present in each individual with prediabetes.

Limitations Along with our preliminary findings regarding the differences in the diabetic diagnostic measures and the gut microbiome composition in T2D patients, it is important to consider the limitations of this study. Firstly, the dataset provided by de la Cuesta-Zuluaga et al. is not well characterized for a diabetes cohort (17). The sample size of prediabetic patients was limited as many patients did not have diabetes as indicated by their HbA1c and FPG levels. To expand our sample size of prediabetic patients, we pooled in diabetic patients which limits our ability to differentiate between microbial differences in patients experiencing insulin resistance and those with well-established T2D. Moreover, patients were not clinically diagnosed during sample collection and therefore, using clinical parameters, we classified patients as prediabetic and diabetic using their HbA1c and FPG values. This may have resulted in inaccurate grouping of patients as T2D diagnosis requires a clinician and often takes into account medical history, symptoms, more than one test or tests taken at multiple time points (58).

Additionally, there may have been confounding variables in our data set that we did not control for including sex, age, genetics, use of medications (e.g. antibiotics, probiotics), mode of delivery, infant feeding, and other chronic conditions which have all been described to impact the microbial composition (59, 60). Particularly relevant to our study is the use of diabetes drugs as Lee et al. describe their ability to alter the gut microbiome composition (16) highlighting its potential impact in our analysis as a confounder. Although it is unlikely to control for all the variables that can influence the gut microbiome, our study is limited by its inability to fully account for these confounding factors, which may introduce bias and affect the validity and generalizability of our findings.

Lastly, our dataset is collected from a unique population in Colombia that is undergoing Westernization as defined by de la Cuesta-Zuluaga et al. (17). This limits its generalizability to populations that consume an inherently Westernized diet such as Canada and the United States or conversely, those who have not undergone Westernization. Regional differences and

individual diets play significant roles in the microbiome composition that should always be taken into consideration when considering gut microbial diversity (61).

Conclusions Our study investigated whether diagnosing prediabetes using HbA1c versus FPG reveals differences in gut microbiome composition within a Colombian cohort. Interestingly, we observed no significant differences in microbial diversity between prediabetic and healthy individuals, regardless of which diagnostic marker was considered. However, notable disparities in taxonomic composition emerged when comparing both diagnostic tools. While some ASVs were shared among prediabetic individuals categorized by both HbA1c and FPG, several were unique to each diagnostic group. Moreover, indicator taxa associated with prediabetes showed no genera overlap between the two diagnostic criteria which may imply the presence of two distinct microbial profiles associated with each diagnostic tool. Additionally, our investigation revealed distinct patterns in microbial taxa and functional pathways between FPG- and HbA1c-categorized individuals. Specifically, FPG-categorized individuals exhibited a higher number of taxa showing both upregulation and downregulation in prediabetic individuals compared to healthy individuals. In contrast, HbA1c-categorized individuals demonstrated more alterations in functional pathways than those with FPG. Our work demonstrates that diagnosing prediabetes using HbA1c and FPG reveals differences in the composition of the gut microbiome on a taxonomic and functional level, thus supporting our hypothesis. This present work emphasizes the need for using multiple diagnostic tools in a clinical setting to enhance the efficacy of prediabetes management.

Future Directions This study offers a potential discrepancy between taxonomic diversity when individuals are diagnosed as prediabetic using different diagnostic tools. To further validate HbA1c and FPG as diagnostic tools for T2D in relation to the gut microbiome, a similar study could be conducted with a larger and well-defined dataset. Larger sample sizes would increase confidence in observed results and using a well-defined dataset would prevent researchers from having to define the threshold of prediabetes based on previous literature. This could further elucidate whether patients diagnosed as prediabetic using different diagnostic tools display differences in taxonomic gut microbiome diversity.

Additionally, since this investigation was carried out in a Colombian cohort, future studies could expand the cultural and geographical diversity of the individuals from which samples are collected. This would permit research findings to be more applicable for extrapolation to the general public.

Furthermore, in our study, we observed the relationship between diagnostic tools and the gut microbiome at a single point in time, but future research could expand this investigation into a longitudinal study. Determining how the gut microbiome and taxonomic diversity change over time in cohorts classified using different diagnostic tools could lead to interesting and novel connections between the gut microbiome and T2D.

Finally, several notable taxa were identified during the indicator species investigation. It may be worthwhile for future studies to explore the connection between the *Erysipelotrichaceae* family and hyperglycemic conditions, given that there is existing literature highlighting potential connections to T2D.

DATA AVAILABILITY

Preliminary data processing scripts can be found in the Supplemental QIIME 2 Script and the R diversity analyses, statistical analyses, and PiCRUST analysis scripts can be found in the Supplemental R Script - all of which can be found at https://github.com/loujainbilal/MICB475_Team4

ACKNOWLEDGEMENTS

We would like to thank Dr. Evelyn Sun and Simranpreet Mann for their guidance and assistance during this investigation. We would also like to thank the Department of Microbiology and Immunology at the University of British Columbia for providing the funding and resources for this project. Finally, we would like to thank researchers de la

Cuesta-Zulaga et al. (17) for providing the sequencing data and metadata that allowed this study to be conducted.

CONTRIBUTIONS

Co-authorship of this manuscript should be considered equal between Loujain Bilal (LB), Ayesha Lalani (AL), Parvin Malhi (PM), Maya Ruehlen (MR) and Aiden Simard (AS). All authors contributed equally to designing the study. LB carried out the alpha and beta analysis, writing the corresponding methods and results, generating Figure 1, Supplemental Figures 1 and 2, and writing the introduction and limitations sections. AL executed the differential abundance and indicator species analysis and wrote the corresponding results and discussion sections, along with generating Figure 3 and Table 2. PM contributed to the data processing through the QIIME 2 pipeline and all of its respective troubleshooting, writing the discussion sections for the diversity analyses, the overlapping taxa between the core microbiome and indicator taxa analysis and PICRUSt2 analysis, along with writing the conclusion. MR contributed to writing the introduction, carrying out the core-microbiome analysis and writing the corresponding results and methods section along with generating Figure 2 and writing the future directions. AS contributed to data wrangling and generating the metadata file, writing the methods dataset description, completing the PICRUSt2 analysis, writing the corresponding methods and results and generating Figure 4 and Table 1. All authors contributed to writing the abstract and completing a general review and edits of the draft manuscript.

REFERENCES

1. Global diabetes cases to soar from 529 million to 1.3 billion by 2050 | Institute for Health Metrics and Evaluation. https://www.healthdata.org/news-events/newsroom/news-releases/global-diabetes-cases-soar-529-million-13-billion-2050. Retrieved 24 February 2024.
2. Diabetes and Diabetes-Related Out-of-Pocket Costs - Diabetes Canada. DiabetesCanadaWebsite. https://www.diabetes.ca/advocacy---policies/advocacy-reports/diabetes-and-diabetes-related-out-of-pocket-costs?gad_source=1&gclid=Cj0KCQiAz8GuBhCxARIsAOpzk8ykgfkELtlRO1DTAvKQrGAu0Q5Qx4m8t-iIUEu0knl7l9bW1jJU8b8aAlhaEALw_wcB. Retrieved 24 February 2024.
3. CDC. 2021. Prediabetes - Your Chance to Prevent Type 2 Diabetes. Cent Dis Control Prev. http://bit.ly/2hMpYrt. Retrieved 15 April 2024.
4. Personalized Postprandial Glucose Response–Targeting Diet Versus Mediterranean Diet for Glycemic Control in Prediabetes | Diabetes Care | American Diabetes Association. https://diabetesjournals.org/care/article/44/9/1980/138839/Personalized-Postprandial-Glucose-Response. Retrieved 24 February 2024.
5. **Bansal N.** 2015. Prediabetes diagnosis and treatment: A review. *World J Diabetes* **6**:296–303.
6. **Ho-Pham LT, Nguyen UDT, Tran TX, Nguyen TV.** 2017. Discordance in the diagnosis of diabetes: Comparison between HbA1c and fasting plasma glucose. *PLoS ONE* **12**:e0182192.
7. **Du L, Li Q, Yi H, Kuang T, Tang Y, Fan G.** 2022. Gut microbiota-derived metabolites as key actors in type 2 diabetes mellitus. *Biomed Pharmacother* **149**:112839.
8. **Scheithauer TPM, Rampanelli E, Nieuwdorp M, Vallance BA, Verchere CB, van Raalte DH, Herrema H.** 2020. Gut Microbiota as a Trigger for Metabolic Inflammation in Obesity and Type 2 Diabetes. *Front Immunol* **11**:571731.
9. **Caricilli AM, Saad MJA**. 2013. The Role of Gut Microbiota on Insulin Resistance. *Nutrients* **5**:829–851.
10. **Rathi KK, Kumari N, Javaid MD, Saleem U, Mortensen E, Zhou Y, Maheshwari N.** 2023. Gut microbiome and prediabetes - a review. *Front Bacteriol* **2**.
11. **Gurung M, Li Z, You H, Rodrigues R, Jump DB, Morgun A, Shulzhenko N.** 2020. Role of gut microbiota in type 2 diabetes pathophysiology. *EBioMedicine* **51**:102590.
12. **Karlsson FH, Tremaroli V, Nookaew I, Bergström G, Behre CJ, Fagerberg B, Nielsen J, Bäckhed F**. 2013. Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature* **498**:99–103.
13. **Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, Liang S, Zhang W, Guan Y, Shen D, Peng Y, Zhang D, Jie Z, Wu W, Qin Y, Xue W, Li J, Han L, Lu D, Wu P, Dai Y, Sun X, Li Z, Tang A, Zhong S, Li X, Chen W, Xu R, Wang M, Feng Q, Gong M, Yu J, Zhang Y, Zhang M, Hansen T, Sanchez G, Raes J, Falony G, Okuda S, Almeida M, LeChatelier E, Renault P, Pons N, Batto J-M, Zhang Z, Chen H, Yang R, Zheng W, Li S, Yang H, Wang J, Ehrlich SD, Nielsen R, Pedersen O, Kristiansen K, Wang J.** 2012. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**:55–60.
14. **Anhê FF, Jensen BAH, Varin TV, Servant F, Van Blerk S, Richard D, Marceau S, Surette M, Biertho L, Lelouvier B, Schertzer JD, Tchernof A, Marette A**. 2020. Type 2 diabetes influences bacterial tissue compartmentalisation in human obesity. *Nat Metab* **2**:233–242.
15. **Iatcu CO, Steen A, Covasa M.** 2021. Gut Microbiota and Complications of Type-2 Diabetes. *Nutrients* **14**:166.
16. **Lee CB, Chae SU, Jo SJ, Jerng UM, Bae SK**. 2021. The Relationship between the Gut Microbiome and Metformin as a Key for Treating Type 2 Diabetes Mellitus. *Int J Mol Sci* **22**:3566.

17. **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.
18. R: The R Project for Statistical Computing. https://www.r-project.org/. Retrieved 14 April 2024.
19. **Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H.** 2019. Welcome to the Tidyverse. J Open Source Softw **4**:1686.
20. **Guo et al.** 2014. Use of HbA1c for Diagnoses of Diabetes and Prediabetes: Comparison with Diagnoses Based on Fasting and 2-Hr Glucose Values and Effects of Gender, Race, and Age.
21. **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.
22. **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.
23. **Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA, Gregory Caporaso J.** 2018. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**:90.
24. **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.
25. **Price MN, Dehal PS, Arkin A.** 2020. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments | *PLOS ONE 5.*
26. **Katoh K, Misawa K, Kuma K, Miyata T.** 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**:3059–3066.
27. Computing diversity metrics — QIIME 2 Cancer Microbiome Intervention Tutorial. https://docs.qiime2.org/jupyterbooks/cancer-microbiome-intervention-tutorial/030-tutorial-downstream/050-core-metrics.html. Retrieved 14 April 2024.
28. QIIME 2 View. https://view.qiime2.org/. Retrieved 15 April 2024.
29. **McMurdie PJ, Holmes S.** 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* **8**:e61217.
30. **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).
31. **Wickham H.** 2016. ggplot2: Elegant Graphics for Data Analysis, 2nd ed. Springer Charm. https://link.springer.com/book/10.1007/978-3-319-24277-4. Retrieved 14 April 2024.
32. 2024. microbiome/microbiome. R. microbiome.
33. **McMurdie PJ, Holmes S.** 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* **8**:e61217.
34. **Neu AT, Allen EE, Roy K.** 2021. Defining and quantifying the core microbiome: Challenges and prospects. *Proc Natl Acad Sci* **118**:e2104429118.
35. **Gao C-H.** 2024. gaospecial/ggVennDiagram. R.
36. **Love MI, Huber W, Anders S.** 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**:550.
37. **Cáceres MD, Legendre P.** 2009. Associations between species and groups of sites: indices and statistical inference. *Ecology* **90**:3566–3574.
38. **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.
39. q2 picrust2 Tutorial. GitHub. https://github.com/picrust/picrust2/wiki/q2-picrust2-Tutorial. Retrieved 14 April 2024.
40. ggpicrust2: an R package for PICRUSt2 predicted functional profile analysis and visualization | Bioinformatics | *Oxford Academic.* https://academic.oup.com/bioinformatics/article/39/8/btad470/7234609. Retrieved 14 April 2024.

41. **Wilkinson TJ, Huws SA, Edwards JE, Kingston-Smith AH, Siu-Ting K, Hughes M, Rubino F, Friedersdorff M, Creevey CJ.** 2018. CowPI: A Rumen Microbiome Focussed Version of the PICRUSt Functional Inference Software. *Front Microbiol* **9**:1095.
42. **Kitten AK, Ryan L, Lee GC, Flores BE, Reveles KR.** 2021. Gut microbiome differences among Mexican Americans with and without type 2 diabetes mellitus. *PLoS ONE* **16**:e0251245.
43. (PDF) Human Gut Microbiota Changes Reveal the Progression of Glucose Intolerance. https://www.researchgate.net/publication/256469051_Human_Gut_Microbiota_Changes_Reveal_the_Progression_of_Glucose_Intolerance. Retrieved 15 April 2024.
44. **Allin KH, Tremaroli V, Caesar R, Jensen BAH, Damgaard MTF, Bahl MI, Licht TR, Hansen TH, Nielsen T, Dantoft TM, Linneberg A, Jørgensen T, Vestergaard H, Kristiansen K, Franks PW, IMI-DIRECT consortium, Hansen T, Bäckhed F, Pedersen O.** 2018. Aberrant intestinal microbiota in individuals with prediabetes. *Diabetologia* **61**:810–820.
45. **Gaike AH, Paul D, Bhute S, Dhotre DP, Pande P, Upadhyaya S, Reddy Y, Sampath R, Ghosh D, Chandraprabha D, Acharya J, Banerjee G, Sinkar VP, Ghaskadbi SS, Shouche YS.** 2020. The Gut Microbial Diversity of Newly Diagnosed Diabetics but Not of Prediabetics Is Significantly Different from That of Healthy Nondiabetics. *mSystems* **5**:e00578-19.
46. **Letchumanan G, Abdullah N, Marlini M, Baharom N, Lawley B, Omar MR, Mohideen FBS, Addnan FH, Nur Fariha MM, Ismail Z, Pathmanathan SG.** 2022. Gut Microbiota Composition in Prediabetes and Newly Diagnosed Type 2 Diabetes: A Systematic Review of Observational Studies. *Front Cell Infect Microbiol* **12**:943427.
47. **Larsen N, Vogensen FK, van den Berg FWJ, Nielsen DS, Andreasen AS, Pedersen BK, Al-Soud WA, Sørensen SJ, Hansen LH, Jakobsen M.** 2010. Gut Microbiota in Human Adults with Type 2 Diabetes Differs from Non-Diabetic Adults. *PLoS ONE* **5**:e9085.
48. **Sedighi M, Razavi S, Navab-Moghadam F, Khamseh ME, Alaei-Shahmiri F, Mehrtash A, Amirmozafari N.** 2017. Comparison of gut microbiota in adult patients with type 2 diabetes and healthy individuals. *Microb Pathog* **111**:362–369.
49. **White KAM, Daneshvari S, Lilyquist J, Luo L, Steffen LE, Bivin A, Gurule N, Ducasa GM, Torres SM, Lindeman R, Sankarappan S, Berwick M.** 2015. Prediabetes: The Variation between HbA1c and Fasting Plasma Glucose. *Int J Diabetol Vasc Dis Res Suppl* **2**:001.
50. HbA1c vs. fasting glucose: understanding the similarities and differences - Blog | Everlywell: Home Health Testing Made Easy. https://www.everlywell.com/blog/hba1c/hba1c-vs-fasting-glucose/. Retrieved 15 April 2024.
51. **Kaakoush NO.** 2015. Insights into the Role of Erysipelotrichaceae in the Human Host. *Front Cell Infect Microbiol* **5**:84.
52. **Lippert K, Kedenko L, Antonielli L, Kedenko I, Gemeier C, Leitner M, Kautzky-Willer A, Paulweber B, Hackl E.** 2017. Gut microbiota dysbiosis associated with glucose metabolism disorders and the metabolic syndrome in older adults. *Benef Microbes* **8**:545–556.
53. **Gravdal K, Kirste KH, Grzelak K, Kirubakaran GT, Leissner P, Saliou A, Casèn C.** 2023. Exploring the gut microbiota in patients with pre-diabetes and treatment naïve diabetes type 2 - a pilot study. *BMC Endocr Disord* **23**:179.
54. **Crudele L, Gadaleta RM, Cariello M, Moschetta A.** 2023. Gut microbiota in the pathogenesis and therapeutic approaches of diabetes. *EBioMedicine* **97**:104821.
55. BLAST: Basic Local Alignment Search Tool. https://blast.ncbi.nlm.nih.gov/Blast.cgi. Retrieved 15 April 2024.
56. **Hipkiss AR**. 2017. Depression, Diabetes and Dementia: Formaldehyde May Be a Common Causal Agent; Could Carnosine, a Pluripotent Peptide, Be Protective? *Aging Dis* **8**:128–130.
57. **Tangvarasittichai S.** 2015. Oxidative stress, insulin resistance, dyslipidemia and type 2 diabetes mellitus. *World J Diabetes* **6**:456–480.
58. CDC. 2023. Diabetes Testing. Cent Dis Control Prev. https://www.cdc.gov/diabetes/basics/getting-tested.html. Retrieved 14 April 2024.
59. **Hasan N, Yang H.** 2019. Factors affecting the composition of the gut microbiota, and its modulation. *PeerJ* **7**:e7502.
60. **Wen L, Duffy A**. 2017. Factors Influencing the Gut Microbiota, Inflammation, and Type 2 Diabetes. *J Nutr* **147**:1468S-1475S.
61. **Senghor B, Sokhna C, Ruimy R, Lagier J-C.** 2018. Gut microbiota diversity according to dietary habits and geographical provenance. *Hum Microbiome J* **7–8**:1–9.