



Article

Deciphering the Impact of Defecation Frequency on Gut Microbiome Composition and Diversity

Gwoncheol Park ^{1,2,3,†} , Seongok Kim ^{1,2,†} , WonJune Lee ^{1,2}, Gyungcheon Kim ^{1,2} and Hakdong Shin ^{1,2,*}

¹ Department of Food Science & Biotechnology, College of Life Science, Sejong University, Seoul 05006, Republic of Korea; gp21p@fsu.edu (G.P.); skim01@sejong.ac.kr (S.K.); leewonjune@sejong.ac.kr (W.L.); paulkc12@gmail.com (G.K.)

² Carbohydrate Bioproduct Research Center, Sejong University, Seoul 05006, Republic of Korea

³ Department of Health, Nutrition & Food Sciences, College of Education, Health & Human Sciences, Florida State University, Tallahassee, FL 32306, USA

* Correspondence: hshin@sejong.ac.kr; Tel.: +82-2-6935-2525

† These authors contributed equally to this work.

Abstract: This study explores the impact of defecation frequency on the gut microbiome structure by analyzing fecal samples from individuals categorized by defecation frequency: infrequent (1–3 times/week, $n = 4$), mid-frequent (4–6 times/week, $n = 7$), and frequent (daily, $n = 9$). Utilizing 16S *rRNA* gene-based sequencing and LC-MS/MS metabolome profiling, significant differences in microbial diversity and community structures among the groups were observed. The infrequent group showed higher microbial diversity, with community structures significantly varying with defecation frequency, a pattern consistent across all sampling time points. The *Ruminococcus* genus was predominant in the infrequent group, but decreased with more frequent defecation, while the *Bacteroides* genus was more common in the frequent group, decreasing as defecation frequency lessened. The infrequent group demonstrated enriched biosynthesis genes for aromatic amino acids and branched-chain amino acids (BCAAs), in contrast to the frequent group, which had a higher prevalence of genes for BCAA catabolism. Metabolome analysis revealed higher levels of metabolites derived from aromatic amino acids and BCAA metabolism in the infrequent group, and lower levels of BCAA-derived metabolites in the frequent group, consistent with their predicted metagenomic functions. These findings underscore the importance of considering stool consistency/frequency in understanding the factors influencing the gut microbiome.

Keywords: stool defecation; defecation frequency; gut microbiome; gut metabolome



Citation: Park, G.; Kim, S.; Lee, W.; Kim, G.; Shin, H. Deciphering the Impact of Defecation Frequency on Gut Microbiome Composition and Diversity. *Int. J. Mol. Sci.* **2024**, *25*, 4657. <https://doi.org/10.3390/ijms25094657>

Academic Editor: Walter Wahli

Received: 20 February 2024

Revised: 2 April 2024

Accepted: 18 April 2024

Published: 25 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The gut microbiome, which comprises 10–100 trillion microorganisms, plays a crucial role in contributing to human health through individual or collective functions of these microbiota [1–3]. Human gut microbiota varies significantly among individuals, based on host properties, such as age, nationality, gender, or body mass index (BMI) [4]. These properties have become obstacles to studying the relationships between the human microbiome and health. To overcome this challenge, recent studies have introduced enterotyping, which stratifies multi-dimensional human gut microbial communities into three distinct enterotypes by the predominance of specific genera: *Bacteroides* (enterotype 1), *Prevotella* (enterotype 2), and *Ruminococcus* (enterotype 3) [4–7]. These enterotypes represent densely populated areas of community composition in a multi-dimensional space, aiming to simplify the complexity of the gut microbiota by reducing individual variation [4]. This approach has enabled the identification of the relationships and/or connections between the human microbiome and various diseases or diets, including the development of diagnostics for cancer and implications for weight loss, as well as the introduction of personalized

nutrition in obesity management [8–11]. Previous reports have demonstrated that Koreans are clustered into two enterotypes, dominated by either *Bacteroides* or *Prevotella* [12,13].

A variety of factors associated with bowel movements, including the colonic transit time [14], and stool consistency [15], have been reported to affect the gut microbiome. Previous studies have shown that defecation frequency is considered as one of the factors that can affect the gut microbial community and the more frequent the defecation, the lower the diversity of the gut microbiota [16,17]. Given that previous studies have only considered single timepoint fecal samples and lacked a multi-omics approach, we aimed to comprehensively investigate the correlation between defecation frequency and the structure of the gut microbiome. We achieved this by incorporating bioinformatics data from a single participant, analyzed through longitudinal observations, along with metabolomic profiles.

In this study, we recruited 27 Korean volunteers, analyzing six longitudinal fecal samples from each subject to investigate the relationship between defecation frequency and gut microbial profiles. Consistent with previous reports, we observed that a higher defecation frequency was associated with a reduction in microbial diversity, and the predicted functional genes and the gut microbial structure varied depending on the defecation frequency.

2. Results

2.1. General Characteristics of the Subjects

Twenty adults, aged 22–31, with a body mass index (BMI) ranging from 18.42 to 27.47 kg/m², were enrolled in this study. All subjects had not been prescribed antibiotics for one month prior to the experiment and were stratified into three groups based on their natural defecation frequency, as determined through a questionnaire: Group 1 (1–3 times/week, *n* = 4), Group 2 (4–6 times/week, *n* = 7), and Group 3 (everyday, *n* = 9). The defecation frequency remained consistent among all the subjects throughout the study period, which lasted up to 23 days. The sampling schedules are detailed in Supplementary Figure S1. It should be noted that the stool collection interval varied among the subjects. However, our primary aim was to capture stable characteristics of the gut microbiome associated with defecation frequency over one month, rather than to track short-term fluctuations. Consequently, no statistical difference in the gut microbial structure was observed at any sampling timepoints within each defecation frequency group, as reflected by the intergroup distance (Supplementary Figure S2). No significant differences were found in regard to age (one-way ANOVA test; *p*-value = 0.540) and BMI (*p*-value = 0.836), according to the defecation frequency groups. The general characteristics of all the participants are listed in Table 1.

Table 1. General characteristics of subjects included in this study (age, sex, BMI).

	1–3 Times a Week	4–6 Times a Week	Everyday	<i>p</i> Value	
Total no. of subjects	<i>n</i> = 4	<i>n</i> = 14	<i>n</i> = 14	-	
No. of enrolled subjects	<i>n</i> = 4	<i>n</i> = 7	<i>n</i> = 9	-	
Sex	Male, <i>n</i> (%)	2 (50.0)	1 (14.3)	5 (55.6)	-
	Female, <i>n</i> (%)	2 (50.0)	6 (85.7)	4 (44.4)	-
Age, average ± SD	24.0 ± 1.2	24.7 ± 1.7	23.9 ± 1.2	0.540	
BMI, average ± SD	22.1 ± 1.5	21.3 ± 3.0	21.4 ± 1.1	0.836	

2.2. Identification of Enterotypes among All Subjects

A total of 2,596,610 sequence reads were obtained from the fecal samples, with an average of over 20,000 sequence reads per sample. These reads were binned into 870 amplicon sequence variants (ASVs) (Table 2).

Table 2. Reads and ASV counts for each group.

Defecation Frequency	1–3 Times	4–6 Times	Everyday	Total
No. of samples	24	38	54	116
No. of sequences	507,975	865,445	1,223,190	2,596,610
Average sequence \pm SD	21,166 \pm 4922	22,775 \pm 5060	22,652 \pm 3901	22,385 \pm 4569
No. of features (ASVs)	459	486	532	870
Average features (ASVs) \pm SD	124 \pm 23	102 \pm 17	94 \pm 21	103 \pm 23

To stratify the population for a better understanding of the complexity of the gut microbiota by minimizing the impact of individual variability, the enterotype of the subjects was determined using baseline samples. The samples were clustered based on the relative abundance at the genus level using the Jensen–Shannon divergence (JSD) distance and the partitioning around medoids (PAM) clustering algorithm. The optimal number of clusters was estimated using the Calinski–Harabasz (CH) index (Figure 1A). The structures of the microbiota from all the samples were separated into two distinct clusters (Figure 1B; 27 subjects with a *Bacteroides*-dominant type and five subjects with a *Prevotella*-dominant type). To further corroborate the clustering across all the subjects based on their enterotypes, we also employed the Bray–Curtis distance metric. All the subjects were separated by their Bray–Curtis distance and marked with their respective enterotypes for the principal coordinates analysis (PCoA). Consistent with the JSD index shown in Figure 1B, a similar separation in the PCoA plot was observed (shown in PC1, 15.4% variance) using the Bray–Curtis index (Figure 1C). Considering that *Prevotella*-dominant and *Bacteroides*-dominant types uniquely feature a predominance of either *Prevotella* or *Bacteroides*, respectively, we expected a distinct *Bacteroides* and *Prevotella* ratio (PB ratio, $Prevotella / (Prevotella + Bacteroides)$) based on their enterotype classification. Accordingly, we calculated the PB ratio for each participant. While most subjects generally clustered into two distinct groups, several in the *Bacteroides*-dominant cluster (e.g., Sub. 083, 094, 107, and 109) showed ambiguous PB ratio values that set them apart from others in same cluster (Figure 1D). Considering that human gut microbiota varies greatly across individuals depending on host properties, which could skew data analysis, we decided to filter the subjects based on the following criteria: (1) those with distinct *Bacteroides* dominance (PB ratio < 0.02); (2) those whose microbiota dominance is maintained across all sampling timepoints. Accordingly, based on the criteria, only samples from a total of 20 subjects with the *Bacteroides*-dominant enterotype were used for further analysis and the consistency of the *Bacteroides*-dominant enterotype across all six sampling points for these subjects was confirmed (Supplementary Figure S3).

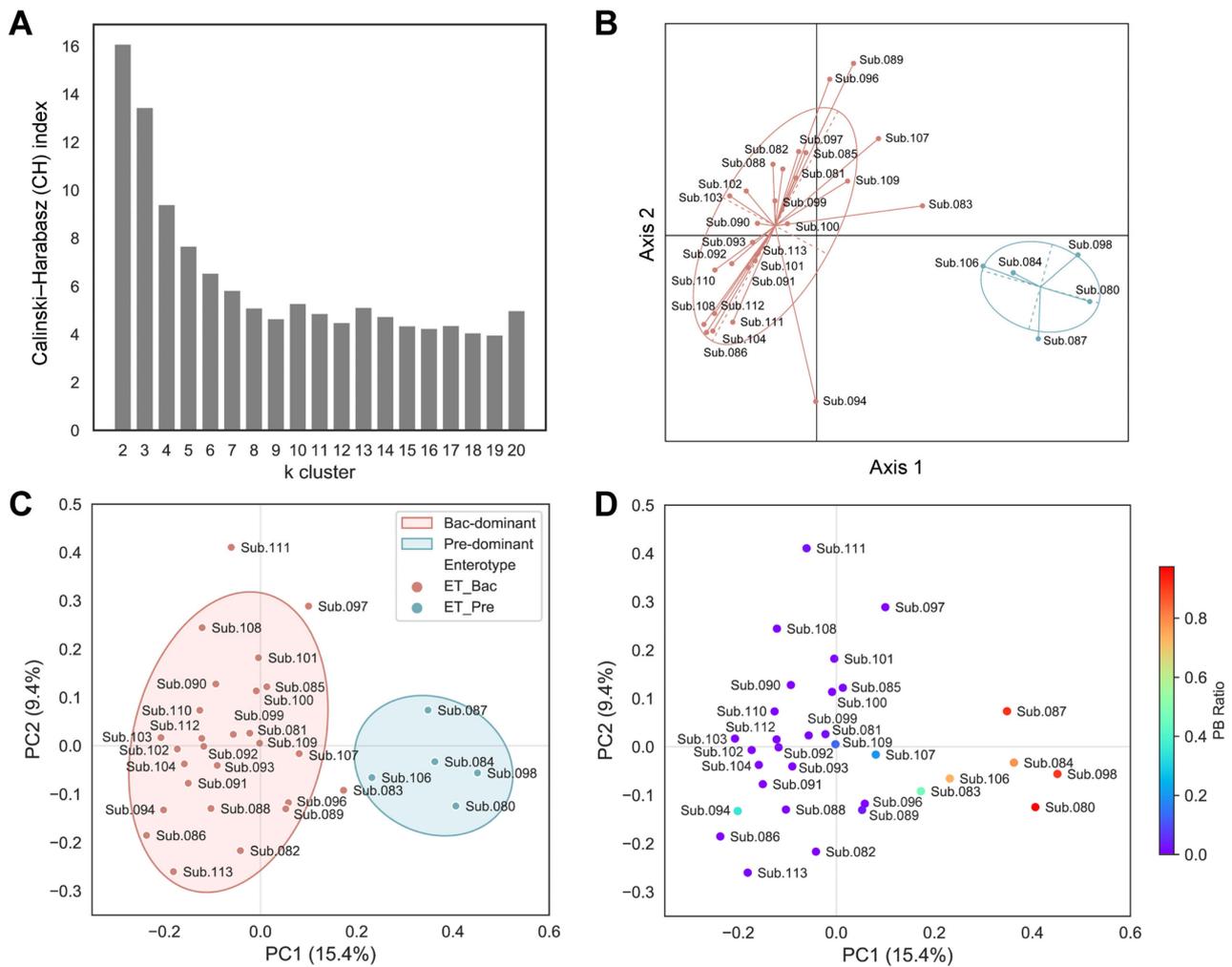


Figure 1. Gut enterotype classification of subjects. **(A)** Evaluated Calinski–Harabasz (CH) index for every number of k clusters to determine the optimal number of clusters. **(B)** Principal coordinates analysis (PCoA) plot of enterotyping data based on relative genus abundance using the Jensen–Shannon divergence (JSD) distance matrix and the partitioning around medoids (PAM) clustering algorithm. **(C,D)** PCoA plots based on the Bray–Curtis distance metric. **(C)** Samples are colored according to the enterotyping results. Covariance ellipses were projected for each cluster and the bounds of the cluster were marked by two standard deviations (2σ) in each direction from the mean of the cluster. **(D)** Samples are colored according to the PB ratio (*Prevotella*/*Prevotella* + *Bacteroides*). Coloring toward red or purple indicates *Prevotella*-dominant or *Bacteroides*-dominant subjects, respectively.

2.3. Microbial Diversity and Structure Differences According to Defecation Frequency

To determine the relationship between microbial diversity and defecation frequency, the alpha diversity was measured using Faith’s PD and observed features indices. The microbial diversity of the ‘1–3 times a week’ group was significantly higher than in the other two groups for both metrics, while no difference was observed between the ‘4–6 times a week’ and ‘everyday’ groups. These results demonstrate that higher defecation frequency is associated with lower microbial diversity in the gut (Figure 2). In line with this trend, we also observed a negative relationship between microbial diversity and defecation frequency at each sampling timepoint, albeit with slight variability (Supplementary Figure S4).

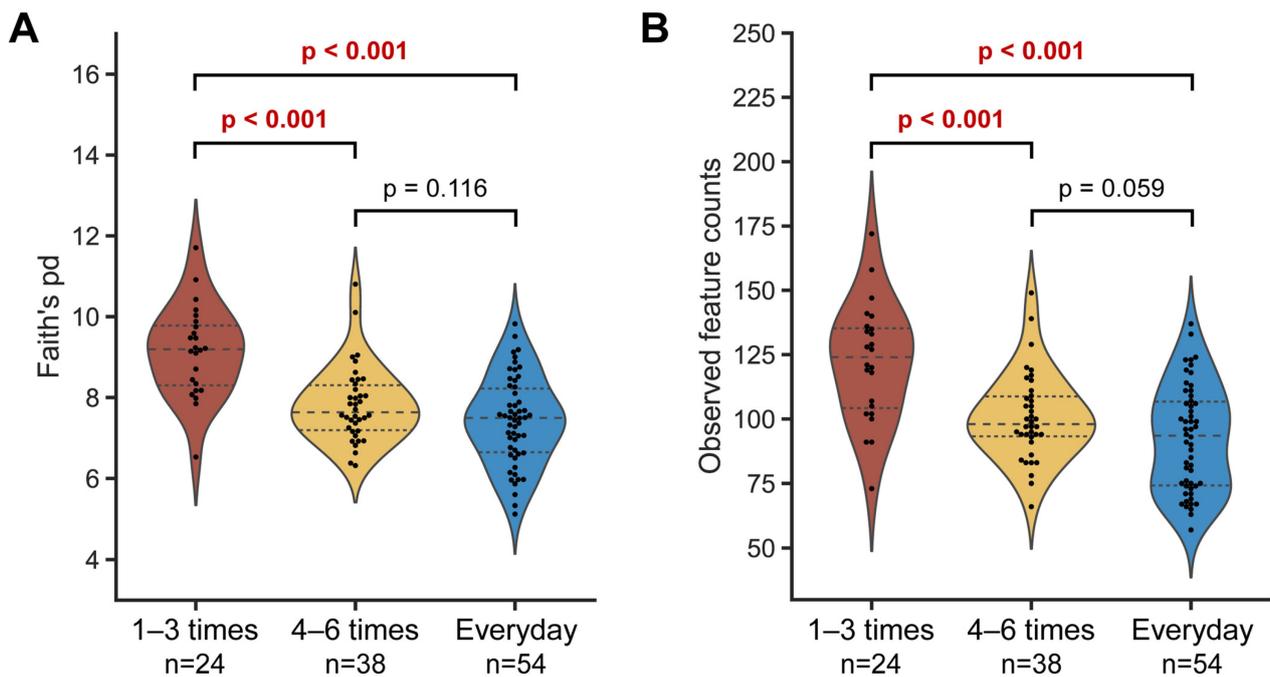


Figure 2. Alpha diversity of gut microbiota according to defecation frequency. Faith's phylogenetic distance (A) and observed features metrics (B) were used to plot the graphs. Each single dot represents an individual. The dashed line represents the median and the dotted lines denote interquartile ranges (IQRs). Statistical analysis was conducted using the pairwise Kruskal–Wallis test.

To evaluate the differences among the samples in regard to the bacterial communities, PCoA analysis based on the weighted UniFrac distance was performed, revealing that each group was well separated, and their community structures significantly differed depending on the defecation frequency (Figure 3A, PERMANOVA $p < 0.001$). The separation also remained consistent at each sampling timepoint (Figure S4A). We then examined the pairwise distances between the groups, according to the defecation frequency: intragroup distances for the everyday samples were significantly greater than the other two groups (Figure 3B); the intergroup distance of the everyday group compared to the 1–3 times a week group was larger than that of 4–6 times a week; and the difference between the pairwise distance was shown to be statistically significant (Figure 3C). Consistent with the weighted UniFrac distance, a similar separation according to the defecation frequency was also observed using the unweighted UniFrac distance in the PCoA plot and a similar trend remained at each sampling timepoint (Supplementary Figures S5 and S6).

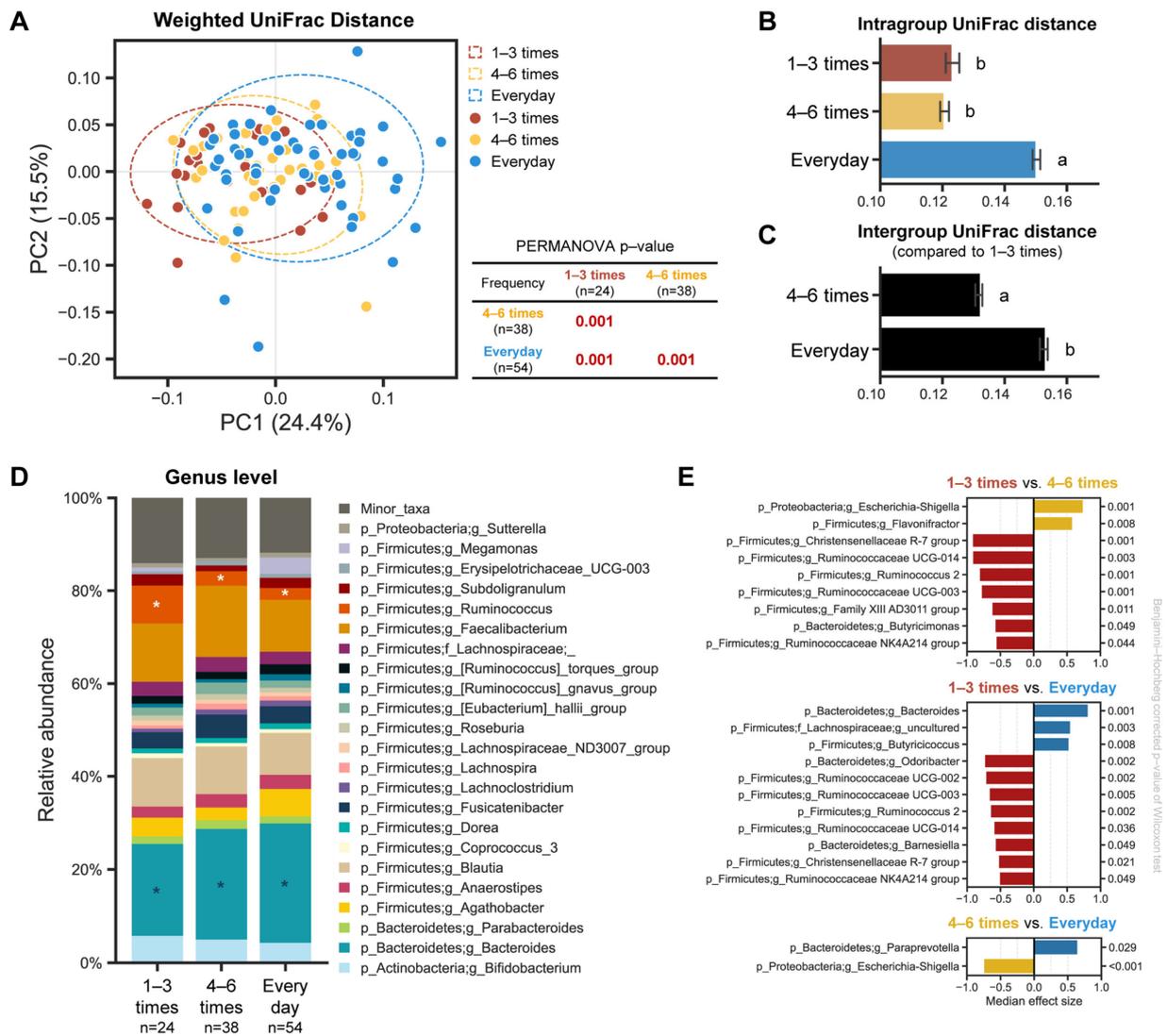


Figure 3. Microbial structure and composition of each group, depending on the defecation frequency. (A) PCoA plot based on weighted UniFrac distance. PERMANOVA was used to test the statistical differences among the groups. Intragroup (B) and intergroup (C) distances, using weighted UniFrac distance, across the three different groups. Data shown and error bars are mean \pm SEM (non-parametric *t*-test, $p < 0.001$). Different alphabets indicate significant differences between the groups ($p < 0.001$). (D) Relative abundance plot of bacterial taxa at the genus level across all three groups. Asterisks (white/blue) denote differentially abundant taxa, according to the defecation frequency. (E) Bar chart showing differentially abundant taxa between the groups using ALDEx2 (median effect size > 0.5 and Benjamini–Hochberg corrected p -value of Wilcoxon test < 0.05).

2.4. Microbial Composition Difference and Relationship with Defecation Frequency

We observed differences in the microbial composition of the gut microbiota based on the defecation frequency. To further dissect which bacterial taxa are differentially prominent in each group, we analyzed the relative bacterial abundance and discriminant taxa using ANOVA-like differential expression 2 (ALDEx2) at the genus level. The *Ruminococcus* genus was significantly overrepresented in the ‘1–3 times a week’ group compared to the other two groups, and its abundance decreased as the defecation frequency increased. Although the *Ruminococcus* genera was identified as a significant component of the gut microbiota in our subjects, its presence alone was not sufficient to alter their established enterotypes (Supplementary Figure S3). In contrast, the *Bacteroides* genus was highly prominent in the ‘everyday’ group compared to the other two groups, and its abundance decreased as the

defecation frequency became less frequent (Figure 3D). Likewise, the same trend was also observed at each single timepoint (Supplementary Figure S3). The ‘1–3 times a week’ group showed more overrepresented bacterial taxa when compared to the ‘4–6 times a week’ or ‘everyday’ groups, respectively (Figure 3E).

2.5. Predictive KEGG Functional Profiling

The Phylogenetic Investigation of Communities by Reconstruction of Unobserved States 2 (PICRUSt2) was used to predict the KEGG functional bacterial gene profiles from the 16S *rRNA* dataset in this study. To assess whether distinct bacterial genes are associated with defecation frequency, we performed a pairwise comparison between the infrequent (‘1–3 times a week’) and the frequent (‘4–7 times a week’) groups. Statistically significant pathways were selected if LDA > 2 and the FDR-adjusted *p*-value was <0.05. The infrequent group showed higher differential enrichments of bacterial genes related to the biosynthesis of fatty acids, vancomycin group antibiotics, peptidoglycan, branched-chain amino acids (BCAAs), lysine, aromatic amino acids, pantothenate and CoA, and terpenoid backbone. Additionally, genes associated with metabolism of propanoate, sulfur-containing amino acids, pyruvate, methane, histidine, glycerophospholipid, pyrimidine, nicotinate and nicotinamide, and retinol were, statistically, significantly more prevalent in the infrequent group when compared to the frequent group. In contrast, the frequent group exhibited a higher prevalence of the bacterial genes related to the metabolism of BCAAs, glyoxylate and dicarboxylate, glutathione, amino sugar and nucleotide sugar, vitamin B6, hexose, lipoic acid, biotin, and the biosynthesis of ubiquinone and other terpenoid quinone, folate, lipopolysaccharide, and polyketide sugar units. Notably, the interesting difference in the overlapping pathways between the two groups includes a higher enrichment of BCAA biosynthesis in the infrequent group and a higher prevalence of BCAA degradation in the frequent group (Figure 4A).

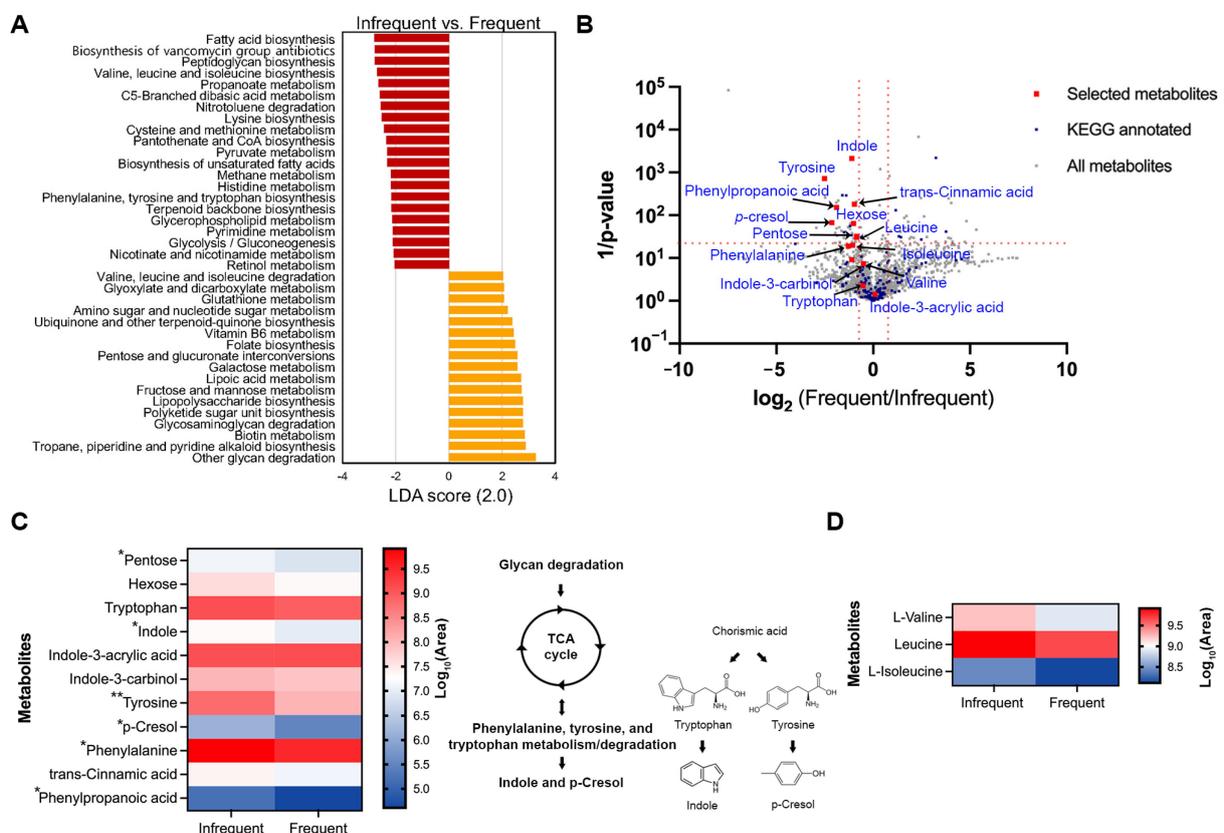


Figure 4. Predictive KEGG functional pathways that are differentially present in pairwise comparisons. (A) Bar chart represents distinctly abundant KEGG pathways that were selected based

on the following criteria: $LDA > 2$ and adjusted p -value < 0.05 . (B) Volcano plots of differential metabolites between the infrequent and the frequent group. All metabolites and those with KEGG annotations are indicated by gray and blue squares, respectively. The metabolites that overlap with predictive functional KEGG pathways generated using 16S *rRNA* data are shown in red squares. Metabolites were identified as candidates based on a $\log_2(\text{Frequent}/\text{Infrequent})$ ratio > 1 and statistical significance ($p < 0.05$) based on unpaired t test after two-stage step-up correction method (indicated by red dotted lines). (C,D) Heatmaps of differential metabolites in phenylalanine/tyrosine/tryptophan biosynthesis and degradation (C), and in branched-chain amino acid metabolism (D). Metabolites with asterisks are statistically significant based on unpaired t test after two-stage step-up correction; * $p < 0.05$; ** $p < 0.01$.

2.6. Differences in the Metabolites Related to Defecation Frequency

To investigate whether specific metabolites are associated with defecation frequency, we analyzed the metabolic profiles, using LC-MS/MS on the fecal samples. Six randomly selected samples from adults, aged 23–34, with a body mass index (BMI) ranging from 17 to 24 kg/m², were included in the analysis. They were divided into two groups based on their defecation frequency: an infrequent group ('1–3 times a week', $n = 3$) and a frequent group ('4–7 times a week', $n = 3$). Consistent with the result that the abundance level of *Ruminococcus* was inversely related with that of *Bacteroides* as the defecation frequency increases (Figure 3D), the same tendency was observed in the participants for the metabolite analysis (Supplementary Figure S7). In our dataset, 814 metabolites in negative mode and 891 metabolites in positive mode were identified using the Compound Discoverer software 3.3 SP2 (Figure 4B, gray squares; Supplementary Dataset, Table S1), with 154 metabolites annotated with a KEGG ID from open databases, such as MZcloud, ChemSpider, KEGG, and the Human Fecal Metabolome database (Figure 4B, blue squares; Supplementary Dataset, Table S1). Consistent with Figure 4A, the pathways for fructose and mannose/galactose metabolism and valine/leucine/isoleucine degradation were significantly more prevalent in the frequent group. In contrast, pathways for phenylalanine/tyrosine/tryptophan and valine/leucine/isoleucine biosynthesis were significantly dominant in the infrequent group (Supplementary Dataset, Table S2). Among the metabolites associated with phenylalanine, tyrosine, and tryptophan metabolism, compounds such as pentose (a precursor for the TCA cycle), indole, tyrosine, p-cresol, phenylalanine, and phenylpropanoic acid showed significant enrichment in the infrequent group compared to the frequent group (Figure 4B,C; red squares). Similarly, among BCAAs, only leucine was statistically more prevalent in the frequent group, although the other two were not statistically significant (Figure 4B–D; red squares).

2.7. Random Forest Prediction

To test whether the defecation frequency of the samples can be predicted based on their microbial compositional features, a random forest classifier (RFC) model was employed using the ASVs table with 50 decision trees. The model generated an overall accuracy of 60%, surpassing the 45% baseline (random prediction without complex computations). Regarding the stool groups, the 'everyday' group showed the highest accuracy (78%), followed by the '4–6 times a week' group (72%) (Figure 5A). The best performance was observed in the '4–6 times a week' group (area under the curve, AUC: 0.83), followed by the 'everyday' group (AUC: 0.76) (Figure 5B). The 'everyday' group showed the best separation in the random forest class probability histogram (Figure 5C).

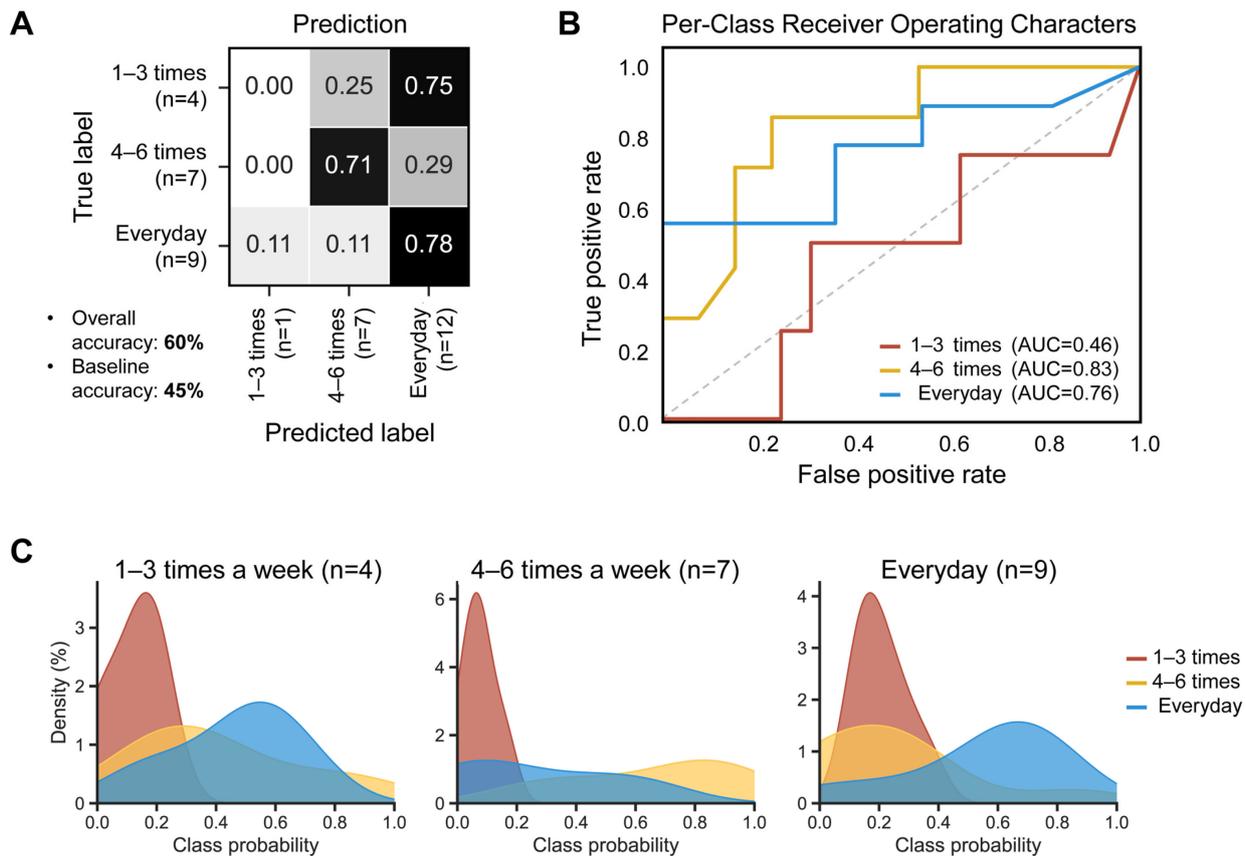


Figure 5. Diagnostic accuracy of gut microbial structure using the baseline fecal samples, according to defecation frequency. The higher the accuracy, the darker the shade. (A) The confusion matrix of the random forest classifier. (B) Receiver operating characteristic (ROC) curves, evaluating ability to predict defecation frequency using random forest classification. Classification accuracy was assessed by the area under the curve (AUC). The dashed line represents the points of no discrimination on a ROC curve. (C) Random forest class probability histograms for ‘1–3 times a week’ (red), ‘4–6 times a week’ (yellow), and ‘everyday’ (blue) according to the defecation frequency.

3. Discussion

This study involved collecting longitudinal fecal samples from each individual over three weeks (six samples), instead of collecting them at a single timepoint. This approach helps control the variations in gut microbiota that could occur within an individual. Our results show clear and lasting differences in the gut microbial profiles according to the defecation frequency. We have demonstrated that less frequent defecation is associated with a richer population of microbes in the gut (Figure 2). In addition, the microbial structure is distinct, depending on the defecation frequency, based on weighted/unweighted UniFrac distances (Figures 3A and S5A), consistent with previous reports [16–18]. These results strongly indicate differences in either the microbial composition or abundance depending on the defecation frequency. It was interesting to observe the gradual movement of the samples from the infrequent to the frequent defecation group towards the PC1 axis, suggesting the possibility that the ‘4–6 times a week’ group has intermediate microbial structure characteristics between the ‘1–3 times a week’ and ‘everyday’ groups (Figure 3A). We found two different taxa that are differentially present depending on the defecation frequency: the *Ruminococcus* and *Bacteroides* genus (Figure 3D,E). Considering that intragroup distances in the ‘1–3 times a week’ and ‘4–6 times a week’ groups are not significantly variable, these specific taxa could potentially be considered as biomarkers for stratifying groups according to defecation frequency.

An overrepresentation in the methane metabolism pathway was observed in the '1–3 times a week' group compared to the '4–6 times a week' group (with no observation in the 'everyday' group) (Figure 4). This is in line with the fact that increased methane production was observed in people with constipation [18,19]. Previous studies have demonstrated the limited bioavailability of polysaccharides, as well as extensive protein catabolism as the retention time increases in in vitro models of the gut [20,21]. In addition to previous findings, recent studies have revealed the relationship between the colonic transit time (CTT), a proxy for defecation frequency, and bacterial metabolism, where a longer CTT is associated with higher colonic protein catabolism, as reflected by the presence of urinary metabolites like *p*-cresol sulfate and *p*-cresol glucuronide. These metabolites are products of tyrosine metabolism, as identified by urinary metabolic phenotyping in the gut [14,18,22]. Furthermore, a longer CTT is positively associated with the higher enrichment of *Ruminococcaceae* and *Ruminococcus* [14]. In this regard, it is noteworthy that the predicted functional profile related to tyrosine metabolism was overrepresented in the '1–3 times a week' group compared to the '4–6 times a week' group (Figure 4). In agreement with previous findings, our data also showed that the *Ruminococcus* genus and *Ruminococcaceae* family were significantly overrepresented in the '1–3 times a week' group, and the functional genes involved in tyrosine metabolism were highly enriched in the '1–3 times a week' group compared to the '4–6 times a week' group (with no observation in the 'everyday' group). To address the limitation of genomic analysis based solely on 16S *rRNA* data, metabolomic analysis was further employed to support the functional gene predictions. Metabolite profiling indicated a higher enrichment of tyrosine metabolism in the infrequent group, evidenced by the increased level of metabolites involved in phenylalanine/tyrosine/tryptophan metabolism. In contrast, the frequent group showed a decreased level of branched-chain amino acids (valine, leucine, and isoleucine), whereas the infrequent group exhibited high levels of these amino acids (Figure 4B–D). These metabolic profiles strongly corroborate the functional KEGG pathway predicted using 16S *rRNA* data (Figure 4B–D). By integrating metabolite profiling with taxa analysis, we observed an overrepresentation of indole and *p*-cresol in the infrequent group, where the abundance of *Bacteroides* was low and that of *Ruminococcus* was high, compared to the frequent group (Figures 4 and S7). These findings are in line with previous findings, demonstrating a negative correlation between the abundance of *Bacteroides* and the levels of *p*-cresol and indole, and a positive correlation with *Ruminococcus* (Figures 3 and 4; [14,23,24]). Considering that *p*-cresol and indole, known as uremic toxins, can potentially cause chronic kidney diseases and cardiovascular diseases, and as such should be detoxified in the body [25,26], frequent stool defecation could be preferable for health. Previous studies have demonstrated that circulating BCAA levels are positively correlated with the development of obesity, and *Bacteroides* spp. are shown to promote BCAA catabolism in brown fat, thus preventing obesity [27,28]. Our findings, which indicate a higher abundance of *Bacteroides* and decreased levels of BCAA in the frequent group compared to the infrequent group, are in strong alignment with these studies.

A strength of this study is the use of longitudinal sample collection from each individual to validate the connection between defecation frequency and gut microbiota, unlike previous studies that considered a single sample at one timepoint. We also employed in silico analysis, using a machine learning model (random forest) trained on the composition of the gut microbiota from individuals to build a prediction model. However, the limitations of this study include an insufficient number of subjects for analyzing *Prevotella*-dominant enterotypes, as well as metabolite analysis. Since each enterotype establishes distinct ecological niches, prolonged defecation frequency may induce markedly varied changes in microbial communities and their metabolic products. Future studies with *Prevotella*-dominant enterotype subjects will help establish a correlation between different microbial compositions/metabolites and defecation frequency. To conduct a more comprehensive analysis of the relationship between defecation frequency and the microbiome profile, future studies should aim to assess defecation frequency using numerical values rather than categorical ones. This approach would facilitate correlational analysis between

defecation frequency and microbial features, including microbial diversity and composition. To deepen our understanding of the comprehensive relationship between defecation frequency and the gut microbiome, further studies employing a comprehensive multi-omics approach with a larger participant cohort are necessary.

4. Materials and Methods

4.1. Sample Collection

Participants were instructed to self-collect fecal samples, twice weekly, over three weeks (23 days at most), at their convenience. Despite this schedule, delays in the collection of these samples, ranging from 1 to 4 days, were recorded for six participants. Sterile cotton-tipped swabs were used for the collection of fresh fecal samples, which were then immediately frozen at -80°C until sequencing. For metabolite analysis, the fecal samples were self-collected by the participants using sterile swabs and subjected to immediate freezing at -80°C until analysis.

4.2. The 16S rRNA Amplicon Sequencing

Total genomic DNA was extracted from the stool samples using the DNeasy PowerSoil HTP 96 kit (Qiagen, Hilden, Germany), according to the manufacturer's instructions, and stored at -80°C for subsequent experiments. The V4 hypervariable region of the 16S rRNA gene was amplified using 515F/806R primers [29]. Amplicons were purified using the NucleoSpin PCR clean-up kit (Macherey-Nagel, Düren, Germany) and subjected to sequencing using the Illumina MiSeq platform (2×300 cycles, paired end). Sample preparations and sequencing steps were performed according to the Earth Microbiome Project (www.earthmicrobiome.org) accessed on 16 February 2023.

4.3. Data Analysis

Analysis of the 16S rRNA sequences was performed using the Quantitative Insights Into Microbial Ecology (QIIME) software package 2-2020.6 [30]. Raw sequencing reads were demultiplexed and filtered on quality using the q2-demux plugin, followed by quality trimming and denoising with DADA2 [31]. All the produced amplicon sequence variants (ASVs) were aligned using MAFFT [32] and used to generate a rooted phylogenetic tree with FastTree 2 [33] for phylogenetic diversity analysis. The q2-feature-classifier plugin was used to trim the ASVs that were not from bacteria using the SILVA 132 database [34,35]. Each ASV was classified taxonomically using a naïve Bayes taxonomy classifier, implemented using the q2-feature-classifier plugin [36]. Faith's phylogenetic diversity [37] and observed features were calculated to measure the alpha diversity, and the unweighted [38]/weighted [39] UniFrac distance were used for the beta diversity. Multiple statistical analyses were performed: (1) a non-parametric Kruskal–Wallis test [40] was used to determine significant differences in the microbial diversity; (2) permutational multivariate analysis of variance (PERMANOVA) was used to evaluate the difference in the community structure (999 random permutations) [41]; (3) a non-parametric *t*-test with 128 Monte Carlo instances was used to run ALDEx2, a tool that detects differentially abundant taxa between samples [42].

4.4. Enterotype Analysis

Baseline samples were clustered based on genus-level collapsed relative abundance using the Jensen–Shannon divergence (JSD) and the partitioning around medoids (PAM) clustering algorithm. The Calinski–Harabasz (CH) index [43] was assessed to determine the optimal number of clusters. The statistical significance of the optimal clustering was evaluated using the silhouette validation technique [44]. Enterotype analysis was performed in the R environment [4], and more detailed information is available at <https://enterotype.embl.de/enterotypes.html> accessed on 1 December 2023.

4.5. Predictive KEGG Function Profiling

The functional prediction of metagenomics profiles from the 16S rRNA gene sequence was computed using the Phylogenetic Investigation of Communities by Reconstruction of Unobserved States 2 (PICRUSt2) [45]. The ASVs table was used as input into the q2-picrust2 plugin of QIIME2. For the inference from the PICRUSt2 results, the Kyoto Encyclopedia of Genes and Genomes (KEGG) ortholog database [46] was used. Significant differences in the predicted functional metagenomic profiles were identified using linear discriminant analysis effect size (LEfSe) analysis (LDA score) [47].

4.6. Sample Preparation for Global Metabolomics

The fecal sample (500 µL) was extracted by mixing an equal volume of 10% of acetonitrile, followed by vigorous shaking for 1 min, then incubated at $-20\text{ }^{\circ}\text{C}$ for 2 h. The resulting mixture was centrifuged at $4\text{ }^{\circ}\text{C}$ and $13,000\times g$ for 15 min, followed by homogenization using a Tissue Lyser (Qiagen) at 30 Hz for 10 min. The supernatant was subjected to LC-MS/MS to profile the metabolites, with quality control (QC) samples used in parallel.

4.7. U-HPLC-MS/MS Conditions

The U-HPLC-MS/MS analyses were performed using a U-HPLC system (Vanquish™ U-HPLC, Thermo Fisher Scientific, Waltham, MA, USA), coupled to a quadrupole mass spectrometer (Thermo Scientific Orbitrap Exploris™ 120 high-resolution/accurate mass spectrometer, interfaced with a heated electrospray ionization (H-ESI) source), at the Biopolymer Research Center for Advanced Materials (BRCAM, Sejong University, Seoul, Republic of Korea). For the chromatographic separation, the prepared samples (10 µL) were injected into a C18 column (Waters, ACQUITY UPLC BEH C18, $2.1\times 100\text{ mm}$, $1.7\text{ }\mu\text{m}$, Waters Corp., Milford, MA, USA), with the column temperature maintained at $45\text{ }^{\circ}\text{C}$ throughout the acquisition period. The mobile phase was eluted by a gradient of water (A) and acetonitrile (B), containing 0.1% acetic acid, with a gradient dilution profile of 95% A (0–2 min), 95–5% A (2–15 min), 5% A (15–17 min), 5–95% A (17–18 min), and 95% A (18–20 min), at a flow rate of 0.3 mL/min. The mass spectrum conditions included a heated capillary of $320\text{ }^{\circ}\text{C}$, a vaporizer temperature of $350\text{ }^{\circ}\text{C}$, a spray voltage of 3.5 KV in positive mode, and 2.5 KV in negative mode; the sheath gas, aux gas, and sweep gas were set at 50, 25, and 1 psi, respectively. The HCD collision energy was set at 15, 30, and 60%, and the scan range covered 55–700 m/z in full scan, with a resolution of 120,000 for MS1 and 15,000 for MS2.

4.8. Global Metabolomic Analysis

The raw MS data were processed using Xcalibur version 4.6 and Compound Discoverer 3.3 (Thermo Fisher Scientific, Waltham, MA, USA). Normalization of unknown compound intensities using the QC sample was performed. Enrichment analysis to identify the potential KEGG pathway was performed using MetaboAnalyst (<https://www.microbiomeanalyst.ca/>) accessed on 29 December 2023. Statistical significance was assessed using Student's *t*-test, with *p*-values < 0.05 considered statistically significant.

4.9. Random Forest

For the random forest classifier prediction, supervised classification of the ASVs table generated from the baseline samples of each individual was performed using the q2-sample-classifier [48] plugin, via the nested stratified 4-fold cross-validation [49] classifier grown with 50 trees.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms25094657/s1>.

Author Contributions: Conceptualization, H.S.; validation, G.P., S.K., W.L., G.K. and H.S.; formal analysis, G.P., S.K., W.L. and G.K.; investigation, G.P., S.K., W.L. and G.K.; resources, H.S.; data curation, G.P., S.K., W.L. and G.K.; writing-original draft preparation, G.P.; writing-review and editing,

S.K. and H.S.; visualization, G.P., S.K., W.L. and G.K.; supervision, H.S.; project administration, H.S.; funding acquisition, H.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by a Korea Basic Science Institute grant and the Basic Science Research Program funded by the Ministry of Education (grant No. 2021R1A2C1095215 and No. 2023R1A6C101A045). This study is also supported by the OTTOGI Corporation through Research and Publication Projects.

Institutional Review Board Statement: This study was approved by the Sejong University Institutional Review Board (SJU-HR-E-2017-009). Participation was voluntary and included written informed consent. Participants were Koreans who had not been prescribed antibiotics for one month before the experiment and who were free from chronic diseases, such as diabetes or hypertensive disorders. Information on the subjects, including their sex, age, and BMI, were recorded to check the distribution between the defecation frequency groups.

Informed Consent Statement: Not applicable.

Data Availability Statement: All amplicon sequence data and metadata have been made public through the EMP data portal (Qiita, <https://qiita.ucsd.edu>; Study ID: 15542) accessed on 23 April 2024.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Gill, S.R.; Pop, M.; Deboy, R.T.; Eckburg, P.B.; Turnbaugh, P.J.; Samuel, B.S.; Gordon, J.I.; Relman, D.A.; Fraser-Liggett, C.M.; Nelson, K.E. Metagenomic analysis of the human distal gut microbiome. *Science* **2006**, *312*, 1355–1359. [[CrossRef](#)] [[PubMed](#)]
2. Tremaroli, V.; Backhed, F. Functional interactions between the gut microbiota and host metabolism. *Nature* **2012**, *489*, 242–249. [[CrossRef](#)] [[PubMed](#)]
3. Shreiner, A.B.; Kao, J.Y.; Young, V.B. The gut microbiome in health and in disease. *Curr. Opin. Gastroenterol.* **2015**, *31*, 69–75. [[CrossRef](#)] [[PubMed](#)]
4. Arumugam, M.; Raes, J.; Pelletier, E.; Le Paslier, D.; Yamada, T.; Mende, D.R.; Fernandes, G.R.; Tap, J.; Bruls, T.; Batto, J.M.; et al. Enterotypes of the human gut microbiome. *Nature* **2011**, *473*, 174–180. [[CrossRef](#)] [[PubMed](#)]
5. Wu, G.D.; Chen, J.; Hoffmann, C.; Bittinger, K.; Chen, Y.Y.; Keilbaugh, S.A.; Bewtra, M.; Knights, D.; Walters, W.A.; Knight, R.; et al. Linking long-term dietary patterns with gut microbial enterotypes. *Science* **2011**, *334*, 105–108. [[CrossRef](#)] [[PubMed](#)]
6. Ou, J.; Carbonero, F.; Zoetendal, E.G.; DeLany, J.P.; Wang, M.; Newton, K.; Gaskins, H.R.; O’Keefe, S.J. Diet, microbiota, and microbial metabolites in colon cancer risk in rural Africans and African Americans. *Am. J. Clin. Nutr.* **2013**, *98*, 111–120. [[CrossRef](#)]
7. Koren, O.; Knights, D.; Gonzalez, A.; Waldron, L.; Segata, N.; Knight, R.; Huttenhower, C.; Ley, R.E. A guide to enterotypes across the human body: Meta-analysis of microbial community structures in human microbiome datasets. *PLoS Comput. Biol.* **2013**, *9*, e1002863. [[CrossRef](#)]
8. Zeller, G.; Tap, J.; Voigt, A.Y.; Sunagawa, S.; Kultima, J.R.; Costea, P.I.; Amiot, A.; Bohm, J.; Brunetti, F.; Habermann, N.; et al. Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol. Syst. Biol.* **2014**, *10*, 766. [[CrossRef](#)]
9. Christensen, L.; Roager, H.M.; Astrup, A.; Hjorth, M.F. Microbial enterotypes in personalized nutrition and obesity management. *Am. J. Clin. Nutr.* **2018**, *108*, 645–651. [[CrossRef](#)]
10. Song, E.J.; Han, K.; Lim, T.J.; Lim, S.; Chung, M.J.; Nam, M.H.; Kim, H.; Nam, Y.D. Effect of probiotics on obesity-related markers per enterotype: A double-blind, placebo-controlled, randomized clinical trial. *EPMA J.* **2020**, *11*, 31–51. [[CrossRef](#)]
11. Zou, H.; Wang, D.; Ren, H.; Cai, K.; Chen, P.; Fang, C.; Shi, Z.; Zhang, P.; Wang, J.; Yang, H.; et al. Effect of Caloric Restriction on BMI, Gut Microbiota, and Blood Amino Acid Levels in Non-Obese Adults. *Nutrients* **2020**, *12*, 631. [[CrossRef](#)]
12. Lim, M.Y.; Rho, M.; Song, Y.M.; Lee, K.; Sung, J.; Ko, G. Stability of gut enterotypes in Korean monozygotic twins and their association with biomarkers and diet. *Sci. Rep.* **2014**, *4*, 7348. [[CrossRef](#)]
13. Lee, S.H.; Yoon, S.H.; Jung, Y.; Kim, N.; Min, U.; Chun, J.; Choi, I. Emotional well-being and gut microbiome profiles by enterotype. *Sci. Rep.* **2020**, *10*, 20736. [[CrossRef](#)] [[PubMed](#)]
14. Roager, H.M.; Hansen, L.B.; Bahl, M.I.; Frandsen, H.L.; Carvalho, V.; Gøbel, R.J.; Dalgaard, M.D.; Plichta, D.R.; Sparholt, M.H.; Vestergaard, H. Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. *Nat. Microbiol.* **2016**, *1*, 16093. [[CrossRef](#)]
15. Vandeputte, D.; Falony, G.; Vieira-Silva, S.; Tito, R.Y.; Joossens, M.; Raes, J. Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. *Gut* **2016**, *65*, 57–62. [[CrossRef](#)] [[PubMed](#)]
16. Kwon, H.J.; Lim, J.H.; Kang, D.; Lim, S.; Park, S.J.; Kim, J.H. Is stool frequency associated with the richness and community composition of gut microbiota? *Intest. Res.* **2019**, *17*, 419–426. [[CrossRef](#)]
17. Hadizadeh, F.; Walter, S.; Belheouane, M.; Bonfiglio, F.; Heinsen, F.A.; Andreasson, A.; Agreus, L.; Engstrand, L.; Baines, J.F.; Rafter, J.; et al. Stool frequency is associated with gut microbiota composition. *Gut* **2017**, *66*, 559–560. [[CrossRef](#)]
18. Asnicar, F.; Leeming, E.R.; Dimidi, E.; Mazidi, M.; Franks, P.W.; Al Khatib, H.; Valdes, A.M.; Davies, R.; Bakker, E.; Francis, L.; et al. Blue poo: Impact of gut transit time on the gut microbiome using a novel marker. *Gut* **2021**, *70*, 1665–1674. [[CrossRef](#)] [[PubMed](#)]

19. Dimidi, E.; Mark Scott, S.; Whelan, K. Probiotics and constipation: Mechanisms of action, evidence for effectiveness and utilisation by patients and healthcare professionals. *Proc. Nutr. Soc.* **2020**, *79*, 147–157. [[CrossRef](#)]
20. Macfarlane, S.; Quigley, M.E.; Hopkins, M.J.; Newton, D.F.; Macfarlane, G.T. Polysaccharide degradation by human intestinal bacteria during growth under multi-substrate limiting conditions in a three-stage continuous culture system. *Fems Microbiol. Ecol.* **1998**, *26*, 231–243. [[CrossRef](#)]
21. Macfarlane, G.T.; Cummings, J.H.; Macfarlane, S.; Gibson, G.R. Influence of retention time on degradation of pancreatic enzymes by human colonic bacteria grown in a 3-stage continuous culture system. *J. Appl. Bacteriol.* **1989**, *67*, 520–527. [[CrossRef](#)] [[PubMed](#)]
22. Davila, A.M.; Blachier, F.; Gotteland, M.; Andriamihaja, M.; Benetti, P.H.; Sanz, Y.; Tome, D. Intestinal luminal nitrogen metabolism: Role of the gut microbiota and consequences for the host. *Pharmacol. Res.* **2013**, *68*, 95–107. [[CrossRef](#)] [[PubMed](#)]
23. Candeliere, F.; Simone, M.; Leonardi, Rossi, M.; Amaretti, A.; Raimondi, S. Indole and p-cresol in feces of healthy subjects: Concentration, kinetics, and correlation with microbiome. *Front. Mol. Med.* **2022**, *2*, 959189. [[CrossRef](#)]
24. Amaretti, A.; Gozzoli, C.; Simone, M.; Raimondi, S.; Righini, L.; Perez-Brocad, V.; Garcia-Lopez, R.; Moya, A.; Rossi, M. Profiling of Protein Degradation in Cultures of Human Gut Microbiota. *Front. Microbiol.* **2019**, *10*, 2614. [[CrossRef](#)] [[PubMed](#)]
25. Falconi, C.A.; Junho, C.; Fogaca-Ruiz, F.; Vernier, I.C.S.; da Cunha, R.S.; Stinghen, A.E.M.; Carneiro-Ramos, M.S. Uremic Toxins: An Alarming Danger Concerning the Cardiovascular System. *Front. Physiol.* **2021**, *12*, 686249. [[CrossRef](#)]
26. Harlacher, E.; Wollenhaupt, J.; Baaten, C.; Noels, H. Impact of Uremic Toxins on Endothelial Dysfunction in Chronic Kidney Disease: A Systematic Review. *Int. J. Mol. Sci.* **2022**, *23*, 531. [[CrossRef](#)] [[PubMed](#)]
27. Newgard, C.B.; An, J.; Bain, J.R.; Muehlbauer, M.J.; Stevens, R.D.; Lien, L.F.; Haqq, A.M.; Shah, S.H.; Arlotto, M.; Slentz, C.A.; et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab.* **2009**, *9*, 311–326. [[CrossRef](#)]
28. Yoshida, N.; Yamashita, T.; Osone, T.; Hosooka, T.; Shinohara, M.; Kitahama, S.; Sasaki, K.; Sasaki, D.; Yoneshiro, T.; Suzuki, T.; et al. *Bacteroides* spp. promotes branched-chain amino acid catabolism in brown fat and inhibits obesity. *iScience* **2021**, *24*, 103342. [[CrossRef](#)]
29. Caporaso, J.G.; Lauber, C.L.; Walters, W.A.; Berg-Lyons, D.; Huntley, J.; Fierer, N.; Owens, S.M.; Betley, J.; Fraser, L.; Bauer, M.; et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **2012**, *6*, 1621–1624. [[CrossRef](#)]
30. Bolyen, E.; Rideout, J.R.; Dillon, M.R.; Bokulich, N.A.; Abnet, C.C.; Al-Ghalith, G.A.; Alexander, H.; Alm, E.J.; Arumugam, M.; Asnicar, F.; et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **2019**, *37*, 852–857. [[CrossRef](#)]
31. Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.; Holmes, S.P. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **2016**, *13*, 581–583. [[CrossRef](#)] [[PubMed](#)]
32. Katoh, K.; Misawa, K.; Kuma, K.; Miyata, T. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066. [[CrossRef](#)] [[PubMed](#)]
33. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLoS ONE* **2010**, *5*, e9490. [[CrossRef](#)] [[PubMed](#)]
34. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glockner, F.O. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2013**, *41*, D590–D596. [[CrossRef](#)] [[PubMed](#)]
35. Yilmaz, P.; Parfrey, L.W.; Yarza, P.; Gerken, J.; Pruesse, E.; Quast, C.; Schweer, T.; Peplies, J.; Ludwig, W.; Glockner, F.O. The SILVA and “All-species Living Tree Project (LTP)” taxonomic frameworks. *Nucleic Acids Res.* **2014**, *42*, D643–D648. [[CrossRef](#)] [[PubMed](#)]
36. Bokulich, N.A.; Kaehler, B.D.; Rideout, J.R.; Dillon, M.; Bolyen, E.; Knight, R.; Huttley, G.A.; Gregory Caporaso, J. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2’s q2-feature-classifier plugin. *Microbiome* **2018**, *6*, 90. [[CrossRef](#)]
37. Faith, D.P. Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* **1992**, *61*, 1–10. [[CrossRef](#)]
38. Lozupone, C.; Knight, R. UniFrac: A new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* **2005**, *71*, 8228–8235. [[CrossRef](#)]
39. Lozupone, C.A.; Hamady, M.; Kelley, S.T.; Knight, R. Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. *Appl. Environ. Microbiol.* **2007**, *73*, 1576–1585. [[CrossRef](#)]
40. Kruskal, W.H.; Wallis, W.A. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* **1952**, *47*, 583–621. [[CrossRef](#)]
41. Anderson, M.J. A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* **2001**, *26*, 32–46.
42. Fernandes, A.D.; Macklaim, J.M.; Linn, T.G.; Reid, G.; Gloor, G.B. ANOVA-like differential expression (ALDEx) analysis for mixed population RNA-Seq. *PLoS ONE* **2013**, *8*, e67019. [[CrossRef](#)] [[PubMed](#)]
43. Caliński, T.; Harabasz, J. A dendrite method for cluster analysis. *Commun. Stat. Theory Methods* **1974**, *3*, 1–27. [[CrossRef](#)]
44. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. [[CrossRef](#)]
45. Douglas, G.M.; Maffei, V.J.; Zaneveld, J.; Yurgel, S.N.; Brown, J.R.; Taylor, C.M.; Huttenhower, C.; Langille, M.G. PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* **2020**, *38*, 685–688. [[CrossRef](#)] [[PubMed](#)]
46. Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **2016**, *44*, D457–D462. [[CrossRef](#)] [[PubMed](#)]

47. Segata, N.; Izard, J.; Waldron, L.; Gevers, D.; Miropolsky, L.; Garrett, W.S.; Huttenhower, C. Metagenomic biomarker discovery and explanation. *Genome Biol.* **2011**, *12*, R60. [[CrossRef](#)] [[PubMed](#)]
48. Bokulich, N.A.; Dillon, M.R.; Bolyen, E.; Kaehler, B.D.; Huttley, G.A.; Caporaso, J.G. q2-sample-classifier: Machine-learning tools for microbiome classification and regression. *J. Open Res. Softw.* **2018**, *3*, 934. [[CrossRef](#)]
49. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.