Dietary Protein Sources Shape Gut Microbiome: Insights from Meta-analysis and Machine Learning

Samson Adejumo[1], Pritha Das[1], Sangita Ghimire[1], Casey Kin Yun Lim[2], Judy Malas[1], Jacob Allen[2], Jarrad Hampton-Marcell[1]

[1]Department of Biological Sciences, University of Illinois Chicago, Chicago, IL 60607

[2]Department of Health and Kinesiology, University of Illinois at Urbana-Champaign, Urbana, IL

Corresponding Author

Jarrad Hampton-Marcell

950 S. Halsted St.

Science & Engineering Laboratory East, Room 4283

Chicago, IL 60607

marcell2@uic.edu

Abstract

Background: Dietary proteins, specifically protein fermentation, are commonly associated with gut microbiota composition; however, identifying microbial features that influence microbial diversity, composition, and structure and their implications in shaping host metabolism remains understudied and poorly understood. Plant- and animal-based proteins differ in amino acid composition and digestibility, yet their impact on microbial community structure and taxonomic shifts lacks systematic analysis. This meta-analysis applies machine learning to identify microbial taxa and their importance in predicting plant- versus animal-based protein diets in murine models, providing insights into bacterial contributions to protein metabolism and gut ecosystem stability.

Methods: Amplicon sequencing data and metadata from 10 peer-reviewed murine microbiota studies (n = 187 mice) were analyzed to investigate the effects of plant- and animal-based protein diets. Study selection and data extraction followed PRISMA guidelines. 16S rRNA gene sequencing data were used to assess microbial community structure, diversity, and composition. Alpha diversity (Shannon, Simpson, Chao1) was analyzed using the Wilcoxon Rank-Sum test, while beta diversity was assessed using Aitchison distance with PCA for ordination. ANOSIM and PERMANOVA tested for statistical differences. Random Forest classification and LEfSe analysis identified microbial taxa significantly associated with dietary protein sources ($p < 0.05$, LDA > 3.0)

Results: Plant-protein diets significantly increased microbial diversity (Shannon index, ANOVA, $p < 0.001$) and enriched 33 genera, including *Bacteroides, Muribaculaceae,* and *Allobaculum.* In contrast, animal-protein diets enriched seven genera, such as *Colidextribacter, Clostridium_sensu_stricto_1,* and *Rikenellaceae_g* ($p < 0.01$). Random Forest classification achieved 96.4% accuracy, identifying key microbial signatures distinguishing dietary groups. LEfSe analysis confirmed these differences, reinforcing distinct microbial adaptations to dietary protein sources.

Conclusion: This study identifies diet-specific microbial biomarkers, demonstrating how dietary protein sources shape gut microbiota composition. These findings highlight the role of diet-microbiome interactions and suggest potential applications for microbiome-targeted nutritional strategies.

Keywords: Dietary protein, Gut microbiome, Microbial diversity, Meta-analysis, Machine learning, Random Forest classification, LEfSe analysis.

1.0 Introduction

The gut microbiome is a dynamic and diverse microbial ecosystem that plays a central role in dietary nutrient metabolism. While dietary carbohydrates are well known to drive saccharolytic fermentation and short-chain fatty acid (SCFA) production, the microbial metabolism of dietary proteins remains less understood (Rowland *et al.*, 2018). Dietary proteins serve as essential substrates for both the host and microbiota, influencing microbial composition, metabolic pathways, and functional interactions within the gut environment (Scott *et al.*, 2013; Wu *et al.*, 2011). Unlike carbohydrates, which undergo rapid fermentation in the proximal colon, dietary proteins that escape digestion in the upper gastrointestinal tract enter the colon, where they undergo microbial metabolism. The metabolic consequences of protein fermentation depend on gut microbial composition, protein source, and interactions with other dietary components (Beaumont *et al.*, 2017).

Microbial protein metabolism occurs through two primary pathways: saccharolytic-associated fermentation and proteolytic-associated degradation. In saccharolytic-associated pathways, protein-derived amino acids are fermented into SCFAs and other bioactive metabolites, contributing to microbial cross-feeding, gut barrier integrity, and immune regulation (Canfora *et al.*, 2019). In contrast, proteolytic-associated pathways degrade proteins into nitrogenous byproducts such as ammonia, hydrogen sulfide, and p-cresol, which have been linked to gut dysbiosis, epithelial barrier dysfunction, and inflammatory processes (Windey *et al.*, 2012; Rooks & Garrett, 2016). While SCFA production is widely considered beneficial, the effects of proteolytic fermentation vary depending on the amino acids metabolized and the microbial taxa involved in their catabolism.

Not all dietary protein sources influence the gut microbiota equally. Plant proteins, which are enriched with dietary fiber and phytochemicals, promote SCFA production, contributing to anti-inflammatory effects, improved gut barrier function, and reduced systemic inflammation (den Besten *et al.*, 2013; Vinelli *et al.*, 2022). However, their lower digestibility may limit the availability of essential amino acids, potentially affecting their ability to meet protein requirements (Scott *et al.*, 2013; Wang *et al.*, 2020; Rutherfurd *et al.*, 2015). By contrast, animal proteins provide essential amino acids critical for muscle synthesis, particularly for athletes and elderly populations (Phillips *et al.*, 2016). However, their higher sulfur-containing amino acid content is associated with increased proteolytic fermentation, leading to greater production of nitrogenous waste products such as hydrogen sulfide and ammonia, which may negatively impact gut homeostasis (Kitada *et al.*, 2019; Ling *et al.*, 2023). These compositional shifts suggest that dietary protein sources play a critical role in shaping gut microbial metabolism, yet the specific bacterial taxa driving these processes remain incompletely characterized.

Despite increasing interest in protein-microbiota interactions, most studies have either focused on individual protein types or inferred functional pathways without systematically assessing

microbial community shifts across different protein sources. While high-protein diets are known to modulate gut microbial diversity, relatively few studies have characterized the taxonomic composition of bacteria responsible for amino acid fermentation versus proteolytic metabolism. This knowledge gap limits our understanding of how microbial taxa contribute to protein metabolism and gut homeostasis.

To address this gap, this study conducts a meta-analysis of murine microbiome datasets to identify microbial taxa associated with plant- and animal-based protein diets. By integrating microbial sequencing data across multiple studies, we aim to systematically characterize how dietary proteins shape microbial community composition and to identify key taxa involved in protein metabolism. We hypothesize that plant-based proteins enrich bacterial taxa linked to amino acid fermentation and SCFA production, whereas animal-based proteins favor taxa involved in proteolytic metabolism and nitrogenous waste production. While this study does not directly analyze microbial metabolites or functional pathways, it provides essential insights into taxonomic shifts in response to dietary protein sources, establishing a foundation for future mechanistic investigations.

2.0 Materials and Methods

2.1 Search Strategy

This meta-analysis followed the Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) guidelines (Page *et al.*, 2021) to ensure a transparent, systematic, and reproducible study selection process. PRISMA provides a structured framework for identifying, screening, and selecting studies in a manner that minimizes bias and enhances the reliability of findings. A comprehensive literature search was conducted to identify murine studies published in different databases from inception to March 31, 2024. Peer-reviewed, English-written articles were retrieved from (1) Web of Science, (2) PubMed, (3) MEDLINE Complete, (4) Scopus, (5) Google Scholar, and (6) manual searching. The Boolean search string used was: *(“protein fermentation” OR “protein digestion” OR “protein intake” OR “dietary protein” OR “protein metabolism” OR “protein consumption” OR “protein diet”) AND (“microbiome” OR “microflora” OR “microbial” OR “commensal” OR “microbiota” OR “bacteria”) AND (“gut” OR “intestinal” OR “gastrointestinal” OR “intestine” OR “digestive tract” OR “GIT” OR “enteric” OR “ileum” OR “colon” OR “colonic” OR “duodenum” OR “jejunum” OR “fecal” OR “faecal”) AND (“mice” OR “rat” OR “rats” OR “mouse” OR “rodent” OR “rodents” OR “murine”).*

2.2 Selection Criteria

The study selection process followed PRISMA guidelines and consisted of four phases: (1) Identification of relevant records, (2) Screening of abstracts for eligibility, (3) Full-text eligibility assessment, and (4) Final inclusion in the meta-analysis. The PRISMA flow diagram (Figure 1) summarizes the number of records identified, screened, excluded, and retained for analysis. Studies were included if they met the following criteria: (1) involved murine models (mice or rats), (2) examined dietary protein interventions, (3) utilized 16S rRNA gene sequencing, (4) had clearly distinguishable treatment and control groups, and (5) provided gut microbiota composition data. Studies were excluded if they were editorials, letters, reviews, comments, or other meta-analyses, lacked a protein intervention, involved non-murine or in vitro studies, or had unpublished data or unclear dietary group differentiation. Search results were exported to

Mendeley for initial deduplication, followed by a second deduplication in Zotero. The remaining records were manually screened by title and abstract to exclude studies irrelevant to the research objectives.

2.3 Bioinformatics and Data Processing

Raw sequencing data (FASTQ files) were retrieved from NCBI and ENA databases using SRA accession numbers reported in selected studies. Bioinformatics processing was performed using QIIME2-amplicon-2024.5 (Bolyen *et al.*, 2019; Ziemski *et al.*, 2022). Sequence processing included demultiplexing and quality filtering using the q2-demux plugin, followed by denoising with DADA2 (v1.28.0) to generate amplicon sequence variants (ASVs) (Callahan *et al.*, 2016). Taxonomic classification was performed using the SILVA-138 database, with a Naïve Bayes classifier (q2-feature-classifier) implementing the classify-sklearn algorithm (Bokulich *et al.*, 2018). ASVs with confidence scores below 70% were excluded to improve data reliability. Processed data—including feature tables, phylogenetic trees, and taxonomy files—were imported into R Studio 4.4.1 for downstream analysis. The phyloseq (v1.48.0) package was used for microbiome data integration, while ggplot2 (v3.5.1), patchwork (v1.3.0), ggpubr (v0.6.0), pheatmap (v1.0.12), and dplyr (v1.1.4) were used for data visualization and statistical analyses (McMurdie & Holmes, 2012; Wen *et al.*, 2023).

2.4 Statistical Analysis

2.4.1 Alpha and Beta Diversity

Alpha diversity indices (Shannon, Simpson, Chao1) were used to assess microbial richness and diversity. Comparisons between plant- and animal-protein diets were performed using the Wilcoxon Rank-Sum test. Beta diversity was assessed using Aitchison distance, a Euclidean-based metric applied to centered log-ratio (CLR)-transformed compositional microbiome data. Principal Component Analysis (PCA) was used for ordination to visualize microbial community differences. Analysis of Similarities (ANOSIM) tested whether microbial composition varied more between dietary groups than within groups (Samuthpongtorn *et al.*, 2022). Permutational analysis of variance (PERMANOVA) was conducted to quantify the variance in microbial composition explained by dietary protein sources (Tang *et al.*, 2016).

2.4.2 Generalized Linear Mixed Model (GLMM)

To account for variability across studies, a Generalized Linear Mixed Model (GLMM) was implemented using the lme4 (v1.1-34) package in R. In this model, dietary protein source was treated as a fixed effect, while isolation source (fecal, ileal, colonic, cecal, or small intestinal content) was included as a random effect to control for site-specific variations in microbial composition. This approach allowed for the identification of microbial diversity differences that are primarily attributed to dietary protein source, while reducing confounding effects from differences in sampling location (Madi *et al.*, 2023; Sweeny *et al.*, 2023).

2.4.3 Receiver Operating Characteristic (ROC) Analysis

Receiver Operating Characteristic (ROC) analysis was conducted to evaluate the predictive power of microbial alpha diversity metrics in distinguishing dietary protein sources. ROC curves were generated using the pROC (v1.18.0) package in R, with the Area Under the Curve (AUC) quantifying classification accuracy. This analysis assessed whether alpha diversity indices (Observed ASVs, Shannon, Simpson, ACE) could reliably differentiate plant- and animal-based protein diets (Li *et al.*, 2022; Zaheer *et al.*, 2018; Deng *et al.*, 2022).

2.4.4 Machine Learning and Microbial Biomarker Identification

A Random Forest (RF) supervised machine learning model was applied to classify dietary protein sources based on microbial composition (Manandhar *et al.*, 2021). This ensemble learning algorithm constructs multiple decision trees and aggregates their predictions to enhance classification accuracy. The dataset was randomly split into 70% training and 30% testing sets. Model performance was evaluated using a confusion matrix, with accuracy, sensitivity, and specificity as key metrics. Feature importance was ranked using the Mean Decrease Accuracy (MDA) metric, which quantifies each taxon's contribution to classification performance.

Additionally, LEfSe (Linear Discriminant Analysis Effect Size) analysis was performed to identify microbial biomarkers distinguishing plant- and animal-based protein diets, applying a significance threshold of $\alpha = 0.05$ and absolute LDA score > 3.0 (Segata *et al.*, 2011).

3.0 Results

3.1 Characteristics of Selected Studies

The literature search identified 5,380 articles across PubMed (n = 339), EMBASE (n = 540), Scopus (n = 2,741), Web of Science (n = 1,710), and manual searches (n = 50). After screening, 2,696 abstracts were excluded due to duplicates (n = 982), reviews (n = 1,188), conference abstracts (n = 309), meta-analyses (n = 8), and letters or editorials (n = 209).

A total of 2,684 full-text articles were assessed, with 2,668 excluded due to lack of an appropriate control group (n = 237), use of non-murine models (n = 1,761), inability to retrieve data (n = 534), or absence of a protein intervention (n = 136).

From the remaining studies, 10 (Masuoka *et al.*, 2020; Xie *et al.*, 2020; Ijaz *et al.*, 2020; Ruocco *et al.*, 2020; Choi *et al.*, 2021; Tan *et al.*, 2022; Yin *et al.*, 2022; Zhu *et al.*, 2022; Zhong *et al.*, 2022; Zhang *et al.*, 2022) were included in the meta-analysis based on availability of sequencing data, well-defined dietary protein interventions, and complete metadata. These studies, as summarized in Table 1, examined the effects of nine dietary protein types, including BCAAs, beef, casein, chicken, EAAs, low-protein diets, pork, sausage, and soy protein. The study selection process is summarized in Figure 1.

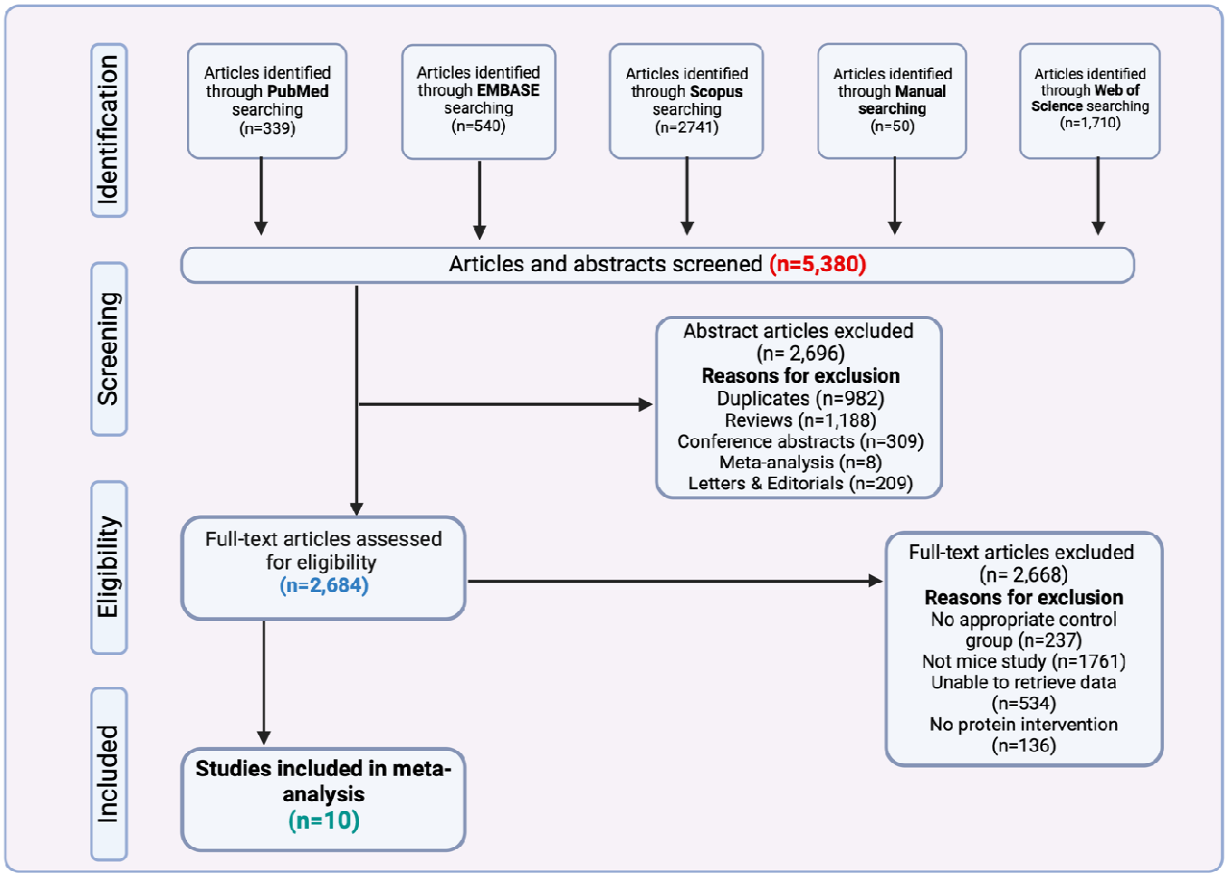


Figure 1: PRISMA Flow Diagram for Study Selection

Table 1: Characteristics of studies included in the meta-analysis

Country, Year, First Author	Title	Bioproject accession	Protein source	Sample number	Age of mice (weeks)
Australia, 2022 Jian Tan	Dietary protein increases T-cell-independent sIgA production through changes in gut microbiota-derived extracellular vesicles	PRJEB39583	Casein	8	6
Japan, 2020 Hiroaki Masuoka	The influences of low protein diet on the intestinal microbiota of mice	PRJDB8898	Casein	30	9
China, 2022 Yantao Yin	Dietary oxidized beef protein alters gut microbiota and induces colonic inflammatory damage in C57BL/6 mice	PRJNA872483	Beef	24	5
China, 2022 Yifan Zhu	Reno-Protective Effect of Low Protein Diet Supplemented With a-Ketoacid Through Gut Microbiota and Fecal Metabolism in 5/6 Nephrectomized Mice	PRJNA810785	Low Protein diet	10	8

China, 2020 Yunting Xie	Processing Method Altered Mouse Intestinal Morphology and Microbial Composition by Affecting Digestion of Meat Proteins	PRJNA607338	Casein, Pork, Sausage, Soy protein	30	4
Canada, 2021 Béatrice S.-Y. Choi	Feeding diversified protein sources exacerbates hepatic insulin resistance via increased gut microbial branched-chain fatty acids and mTORC1 signaling in obese mice	PRJEB37442	Casein	12	10
China, 2020 Muhammad Umair Ijaz	Meat Protein in High-Fat Diet Induces Adipogenesis and Dyslipidemia by Altering Gut Microbiota and Endocannabinoid Dysregulation in the Adipose Tissue of Mice	PRJNA548036	Beef, Casein, Chicken, Pork	27	7
China, 2022 Wuling Zhong	High-protein diet prevents fat mass increase after dieting by counteracting Lactobacillus-enhanced lipid absorption	PRJNA868153	Casein	14	8
China, 2022 Ranran Zhang	Oral administration of branched-chain amino acids ameliorates high-fat diet-induced metabolic-associated fatty liver disease via gut microbiota-associated mechanisms	PRJNA852193	Branched-chain amino acids	12	6
Italy, 2020 Chiara Ruocco	Manipulation of Dietary Amino Acids Prevents and Reverses Obesity in Mice Through Multiple Mechanisms That Modulate Energy Homeostasis	PRJEB25686	Essential amino acids, Casein	20	8

3.2 Composition of Core Gut Bacteria in Protein-fed mice

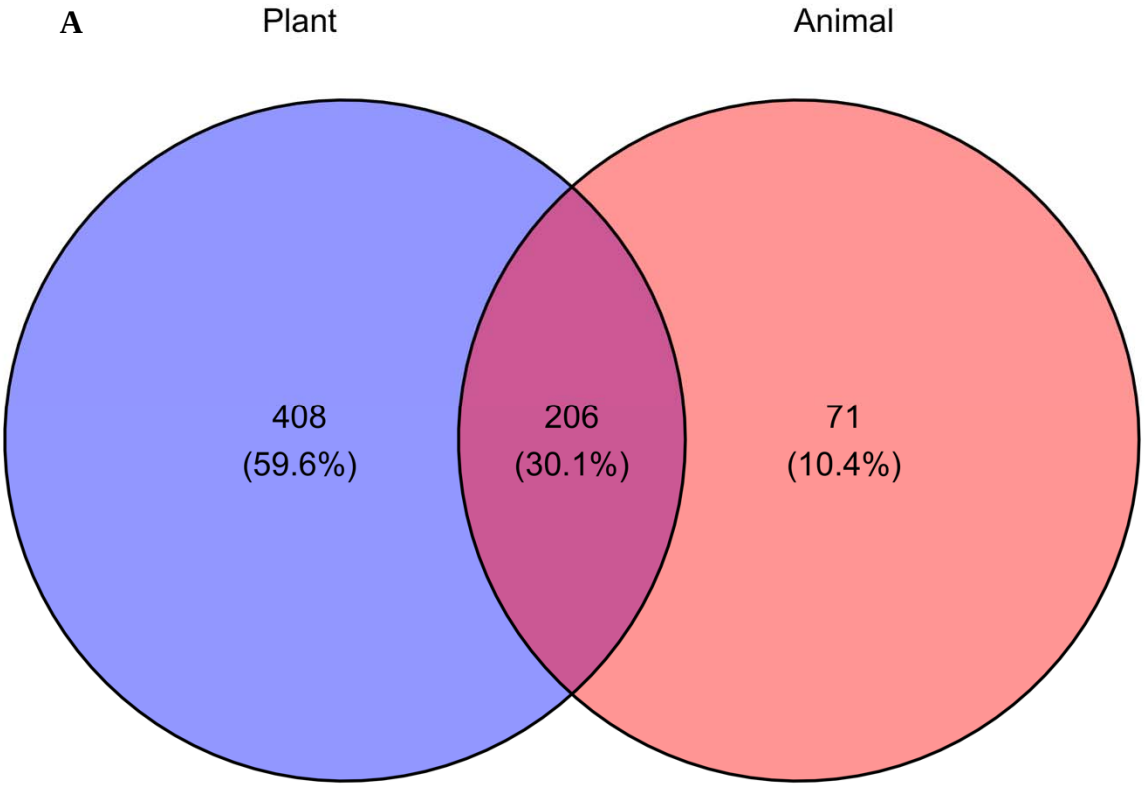
A total of 9,575,175 sequencing reads were obtained from 187 protein-fed mouse samples, generating 16,630 amplicon sequence variants (ASVs), with an average of 51,204 reads per sample. After denoising and chimera removal, 10,123 ASVs remained.

Among the 685 bacterial genera identified, plant protein-fed mice supported a significantly greater number of unique taxa (408 genera, 59.6%) compared to animal protein-fed mice (71 genera, 10.4%), indicating greater taxonomic diversity (Figure 2A). Additionally, plant-protein diets supported a significantly higher number of low-abundance taxa than animal protein diets, reflecting enhanced ecological richness and metabolic versatility (Figure 2B vs. Figure 2C).

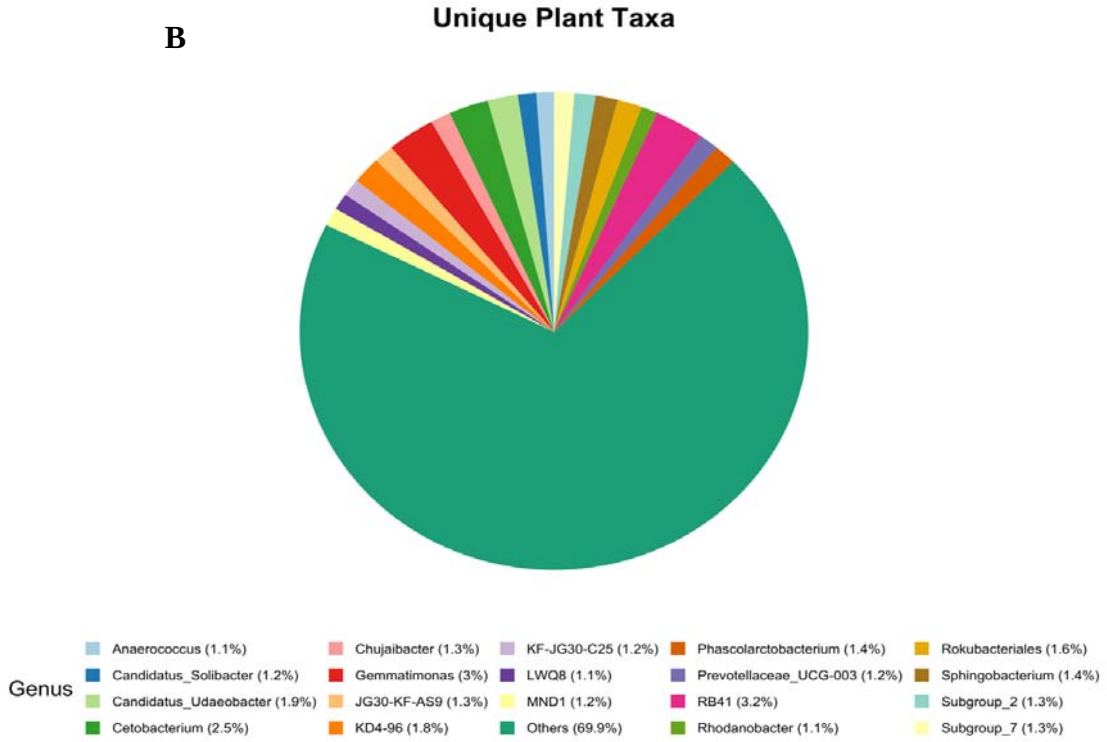
Despite these differences, both dietary groups shared a core set of dominant genera, including *Muribaculaceae*, *Faecalibaculum*, and *Akkermansia*, though their relative abundances varied by diet. *Muribaculaceae*, a prominent saccharolytic fermenter, was more abundant in plant-protein-fed mice (35.64%) than in animal-protein-fed mice (21.93%), suggesting a preference for plant-derived polysaccharides. Conversely, animal protein-fed mice showed enrichment of *Lactococcus* (7.34%), *Ligilactobacillus* (4.24%), and *Staphylococcus* (3.25%), which are associated with proteolytic metabolism and potential pro-inflammatory activity. In contrast,

Dubosiella (4.79%), *Allobaculum* (4.79%), *Bacteroides* (3.02%), and *Alloprevotella* (2.15%) were more abundant in plant protein-fed mice, supporting their role in fermenting complex carbohydrates and maintaining gut integrity (Figure 2D; Supplementary Table 1).

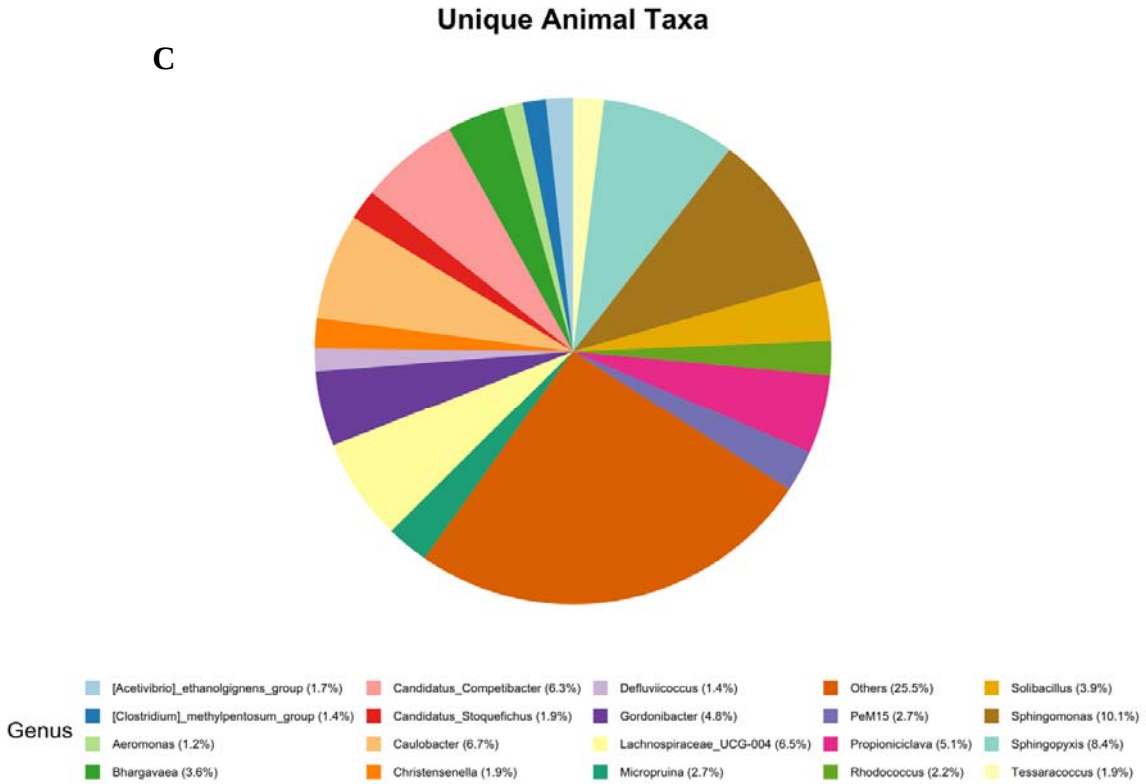
Across all samples, the most dominant phyla were Firmicutes (58.48%) and Bacteroidota (21.88%), followed by Proteobacteria (7.08%), Actinobacteriota (4.13%), Acidobacteriota (1.34%), and Verrucomicrobiota (1.28%). The Firmicutes-to-Bacteroidota (F: B) ratio, a frequently cited marker of metabolic health, was significantly higher in animal protein-fed mice (63.3:19.9) than in plant protein-fed mice (47.3:23.5) (Figure 2E, Supplementary Table 2). While an elevated F:B ratio has historically been linked to obesity, inflammation, and gut barrier dysfunction (Walters *et al.*, 2014; Ley *et al.*, 2006), emerging evidence suggests that this ratio alone is not a definitive marker of microbial health, emphasizing the need to assess overall microbiome composition and function.



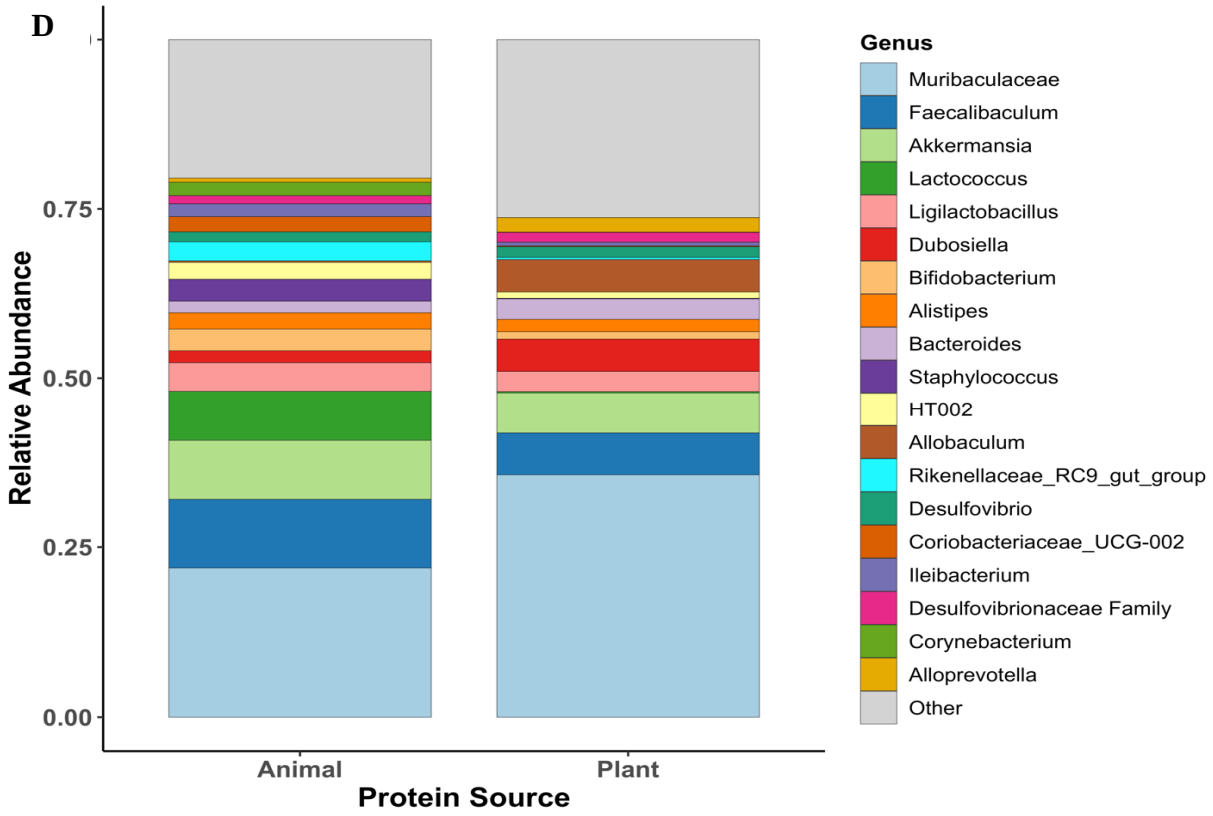
B


Unique Plant Taxa
Genus
Anaerococcus (1.1%)
Candidatus_Solibacter (1.2%)
Candidatus_Udaeobacter (1.9%)
Cetobacterium (2.5%)
Chujaibacter (1.3%)
Gemmatimonas (3%)
JG30-KF-AS9 (1.3%)
KD4-96 (1.8%)
KF-JG30-C25 (1.2%)
LWQ8 (1.1%)
MND1 (1.2%)
Others (69.9%)
Phascolarctobacterium (1.4%)
Prevotellaceae_UCG-003 (1.2%)
RB41 (3.2%)
Rhodanobacter (1.1%)
Rokubacteriales (1.6%)
Sphingobacterium (1.4%)
Subgroup_2 (1.3%)
Subgroup_7 (1.3%)

C


Unique Animal Taxa
Genus
[Acetivibrio]_ethanolgignens_group (1.7%)
[Clostridium]_methylpentosum_group (1.4%)
Aeromonas (1.2%)
Bhargavaea (3.6%)
Candidatus_Competibacter (6.3%)
Candidatus_Stoquefichus (1.9%)
Caulobacter (6.7%)
Christensenella (1.9%)
Defluviicoccus (1.4%)
Gordonibacter (4.8%)
Lachnospiraceae_UCG-004 (6.5%)
Micropruina (2.7%)
Others (25.5%)
PeM15 (2.7%)
Propioniciclava (5.1%)
Rhodococcus (2.2%)
Solibacillus (3.9%)
Sphingomonas (10.1%)
Sphingopyxis (8.4%)
Tessaracoccus (1.9%)

D


Relative Abundance
0.00
0.25
0.50
0.75
Animal
Plant
Protein Source
Genus
Muribaculaceae
Faecalibaculum
Akkermansia
Lactococcus
Ligilactobacillus
Dubosiella
Bifidobacterium
Alistipes
Bacteroides
Staphylococcus
HT002
Allobaculum
Rikenellaceae_RC9_gut_group
Desulfovibrio
Coriobacteriaceae_UCG-002
Ileibacterium
Desulfovibrionaceae Family
Corynebacterium
Alloprevotella
Other

E

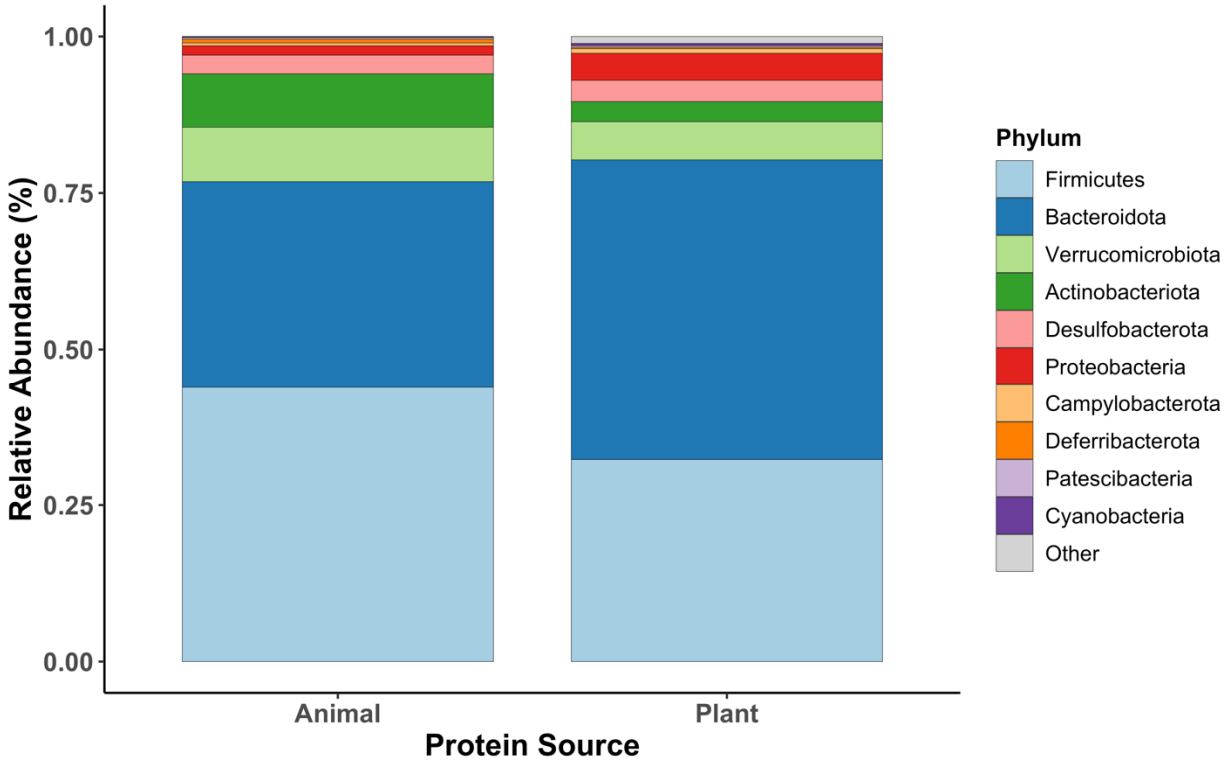

Relative Abundance (%)
0.00
0.25
0.50
0.75
1.00
Animal
Plant
Protein Source
Phylum
Firmicutes
Bacteroidota
Verrucomicrobiota
Actinobacteriota
Desulfobacterota
Proteobacteria
Campylobacterota
Deferribacterota
Patescibacteria
Cyanobacteria
Other

Figure 2. Gut microbial profiles differ between plant- and animal-based protein diets.
(A) Venn diagram showing the number of shared and unique bacterial genera in plant- and animal protein-fed mice. Pie charts showing the relative abundance of the top 20 unique taxa in plant-fed (B) and animal-fed (C) mice. Bar plots showing differentially abundant genera by diet group at the genus level (D) and phylum level (E).

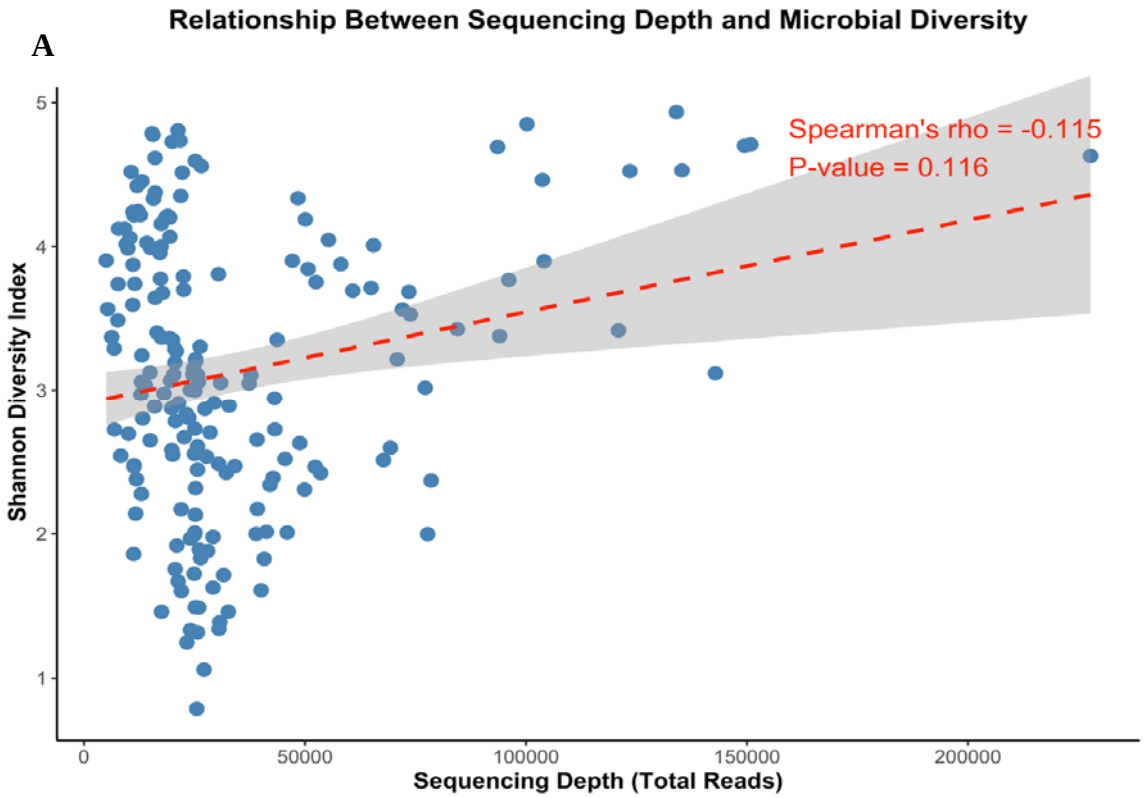
3.3 Dietary Protein Sources Influence Microbial Alpha Diversity

The Generalized Linear Mixed Model (GLMM) assessed the effects of dietary protein source, protein level, host age, and country of origin on Shannon diversity, incorporating isolation source as a random effect. Among the tested models, Model 3 provided the best fit, yielding the lowest AIC (398.08) and BIC (420.70) and the highest log-likelihood (-192.04). The chi-squared test confirmed that adding protein level, host age, and country of origin significantly improved model performance ($p = 1.453e-11$). Among fixed effects, protein source had the strongest positive effect (1.25097), followed by protein level (0.40397) and host age (0.30001), while country of origin (-0.55194) was negatively associated with microbial diversity (Supplementary Table 3).

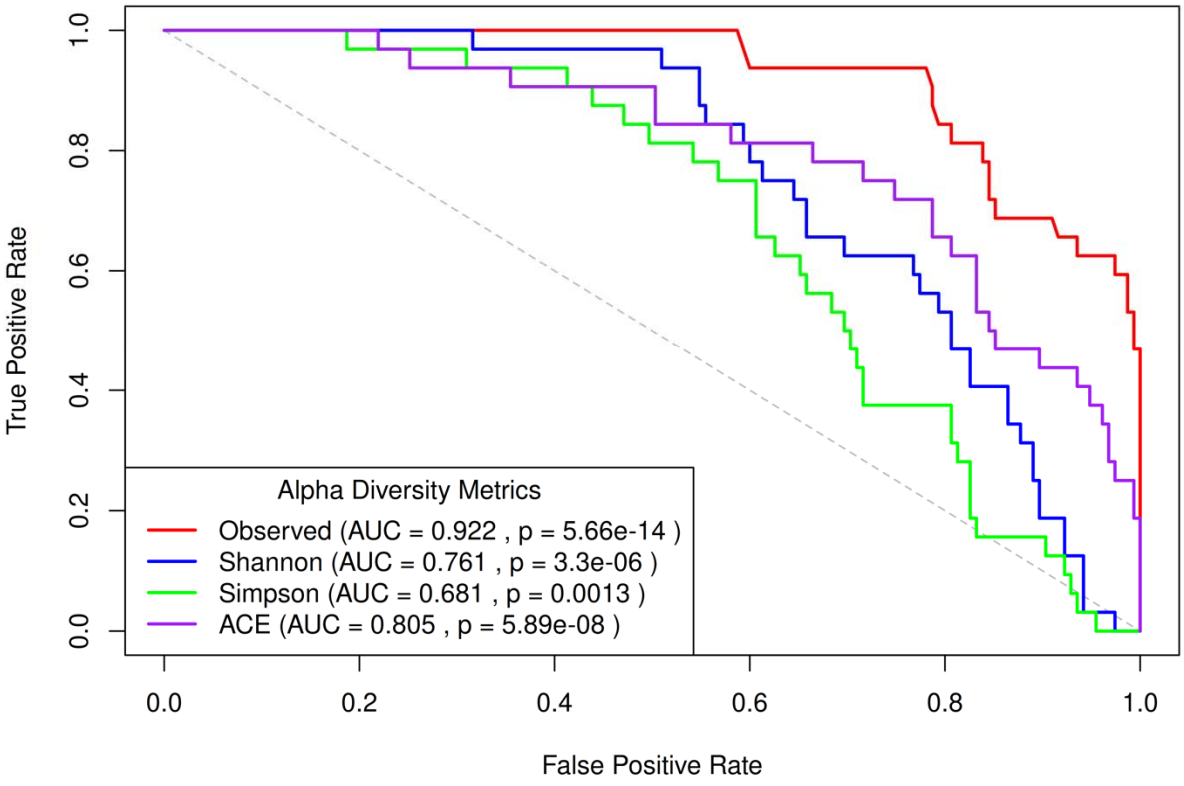
Sequencing depth did not significantly influence diversity estimates, as confirmed by Spearman rank correlation between sequencing depth and the Shannon diversity index ($\rho = 0.107$, $p = 0.145$; Figure 3A). This indicates that microbial richness and evenness differences were driven by dietary protein sources rather than technical biases (Zaheer *et al.*, 2018; Deng *et al.*, 2022).

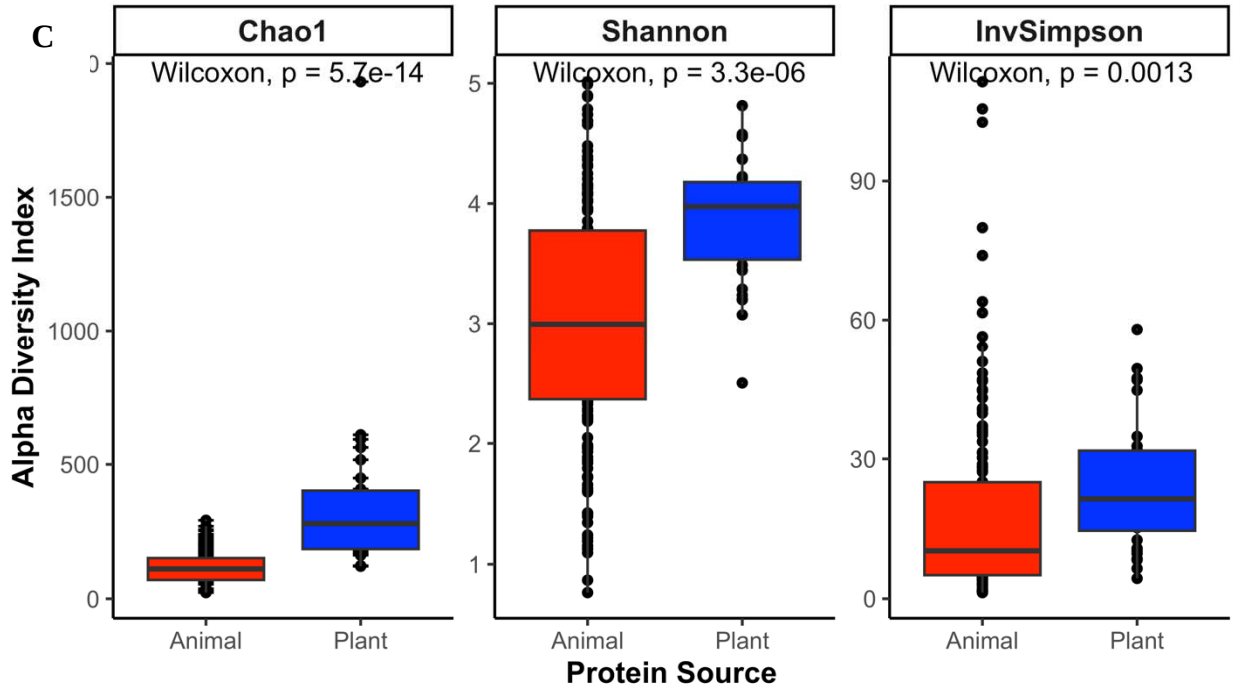
ROC analysis validated the predictive capacity of alpha diversity metrics, with Observed Richness achieving the highest classification accuracy (AUC = 0.922, $p = 5.66e-14$), followed by ACE (AUC = 0.805, $p = 5.89e-08$), Shannon Index (AUC = 0.761, $p = 3.3e-06$), and Simpson Index (AUC = 0.681, $p = 0.0013$) (Figure 3B-C).

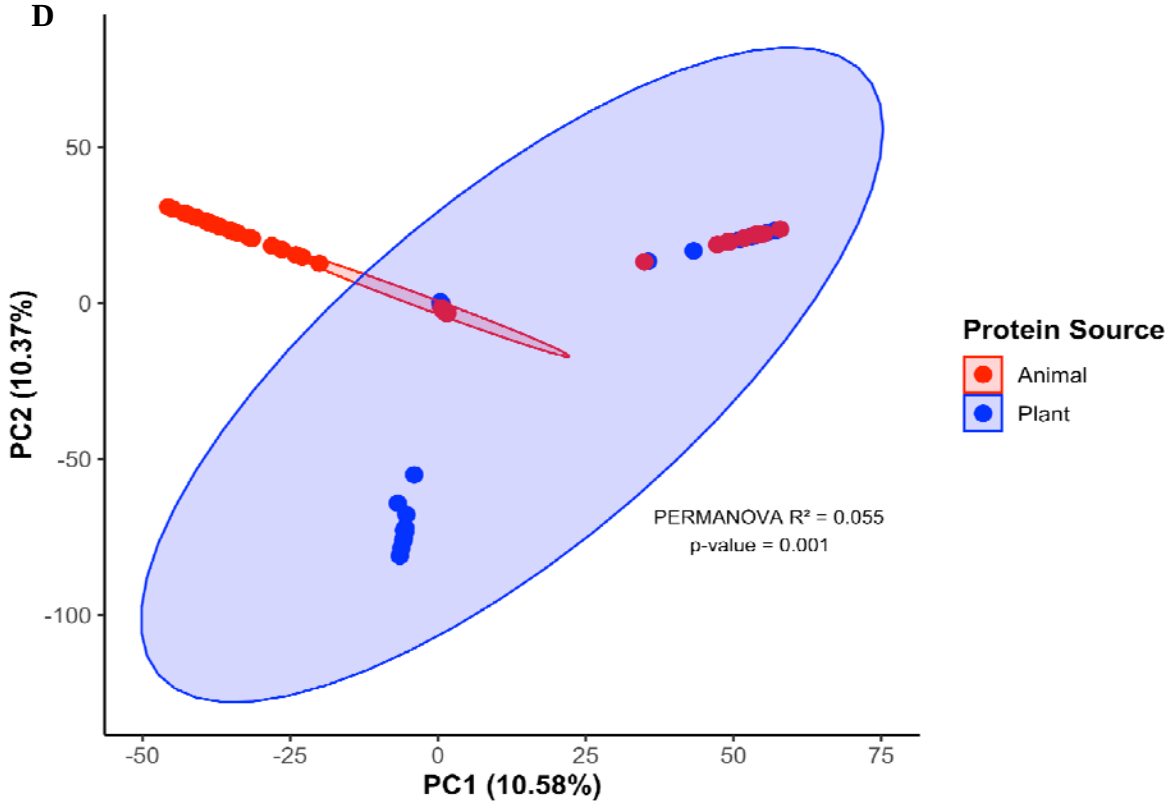
A


Relationship Between Sequencing Depth and Microbial Diversity
Spearman's rho = -0.115
P-value = 0.116
Shannon Diversity Index
Sequencing Depth (Total Reads)
0
50000
100000
150000
200000
1
2
3
4
5

B


Alpha Diversity Metrics
Observed (AUC = 0.922 , p = 5.66e-14)
Shannon (AUC = 0.761 , p = 3.3e-06)
Simpson (AUC = 0.681 , p = 0.0013)
ACE (AUC = 0.805 , p = 5.89e-08)
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0


C
Chao1
Wilcoxon, p = 5.7e-14
Shannon
Wilcoxon, p = 3.3e-06
InvSimpson
Wilcoxon, p = 0.0013
Alpha Diversity Index
Animal
Plant
Protein Source


D
PC2 (10.37%)
PC1 (10.58%)
PERMANOVA R² = 0.055
p-value = 0.001
Protein Source
Animal
Plant

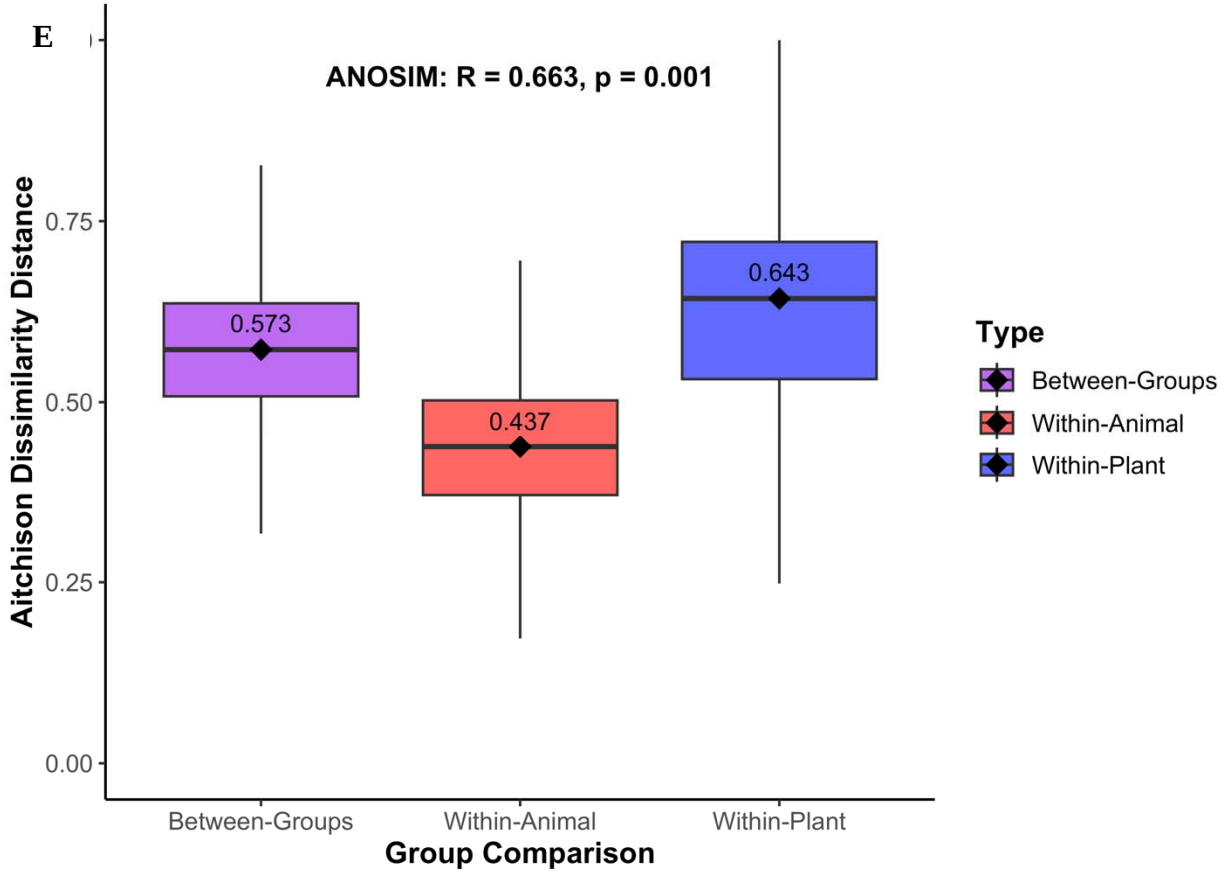


Figure 3: Microbial Diversity and Community Structure in Protein-Based Diets

Spearman correlation between sequencing depth and Shannon diversity (A). ROC curves of alpha diversity indices highlight their predictive capacity for dietary protein sources (B). Boxplots show significantly higher richness and evenness in plant-protein samples (C). PCoA based on Aitchison distance demonstrates distinct clustering of microbial communities between dietary groups, with ellipses indicating 95% confidence intervals (D). ANOSIM analysis confirms significant differences in community structure (E; $R = 0.663$, $P < 0.001$).

3.4 Dietary Protein Sources Shape Beta Diversity and Microbial Community Structure

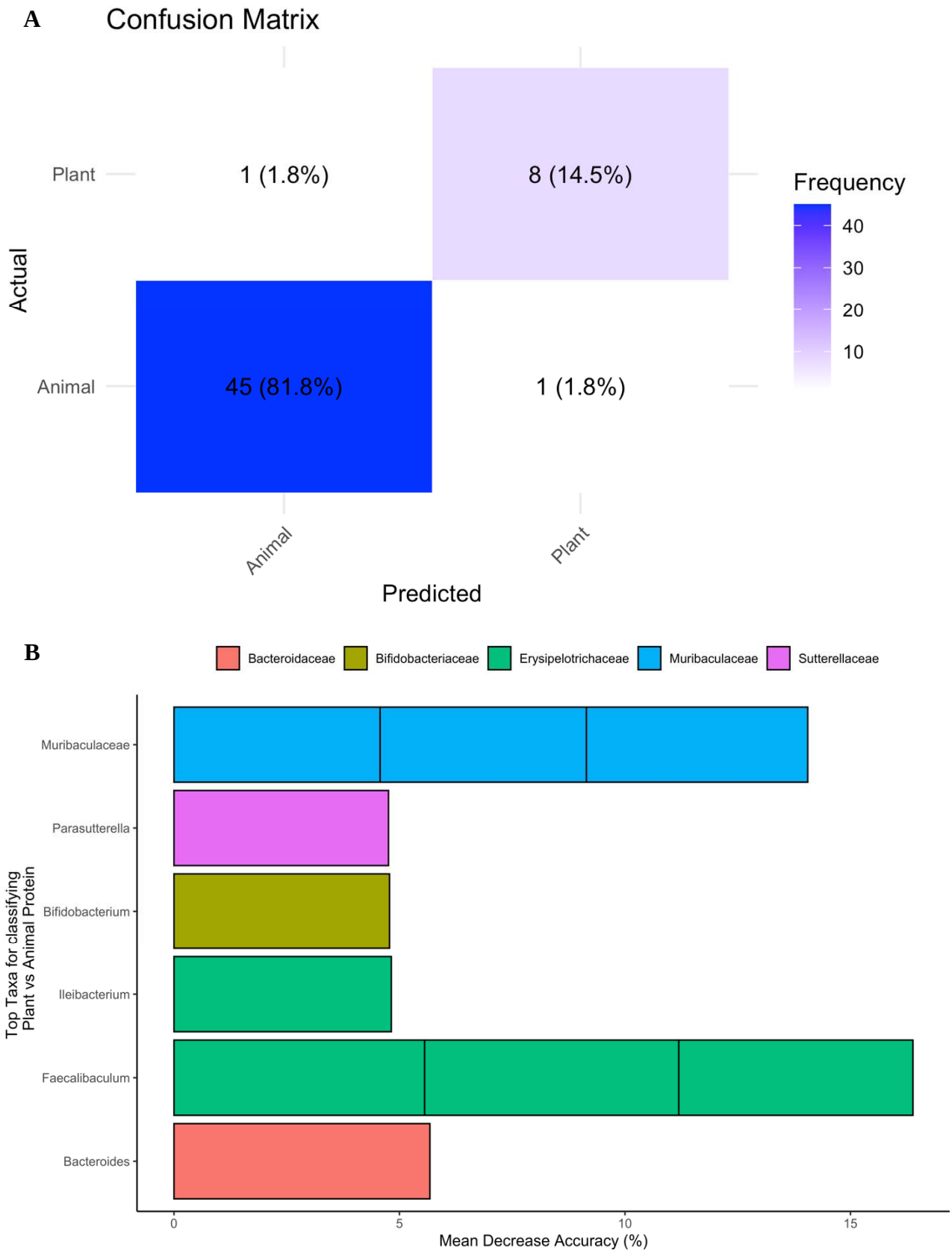
Beta diversity analysis revealed distinct microbial clustering patterns between plant- and animal-protein diets, indicating diet-driven differences in microbial composition. Principal Component Analysis (PCA) using Euclidean distances showed clear separation between groups, with the first two principal components explaining 10.58% and 10.37% of the variance, respectively (Figure 3D). Between-group differences were further validated by PERMANOVA ($R^2 = 0.055$, $p = 0.001$) and ANOSIM ($R = 0.663$, $p = 0.001$), confirming significant compositional divergence between plant- and animal-protein-fed microbiomes (Figure 3E).

3.4 Random Forest Model and LEfSe Analysis Highlight Microbial Signatures of Protein Diets

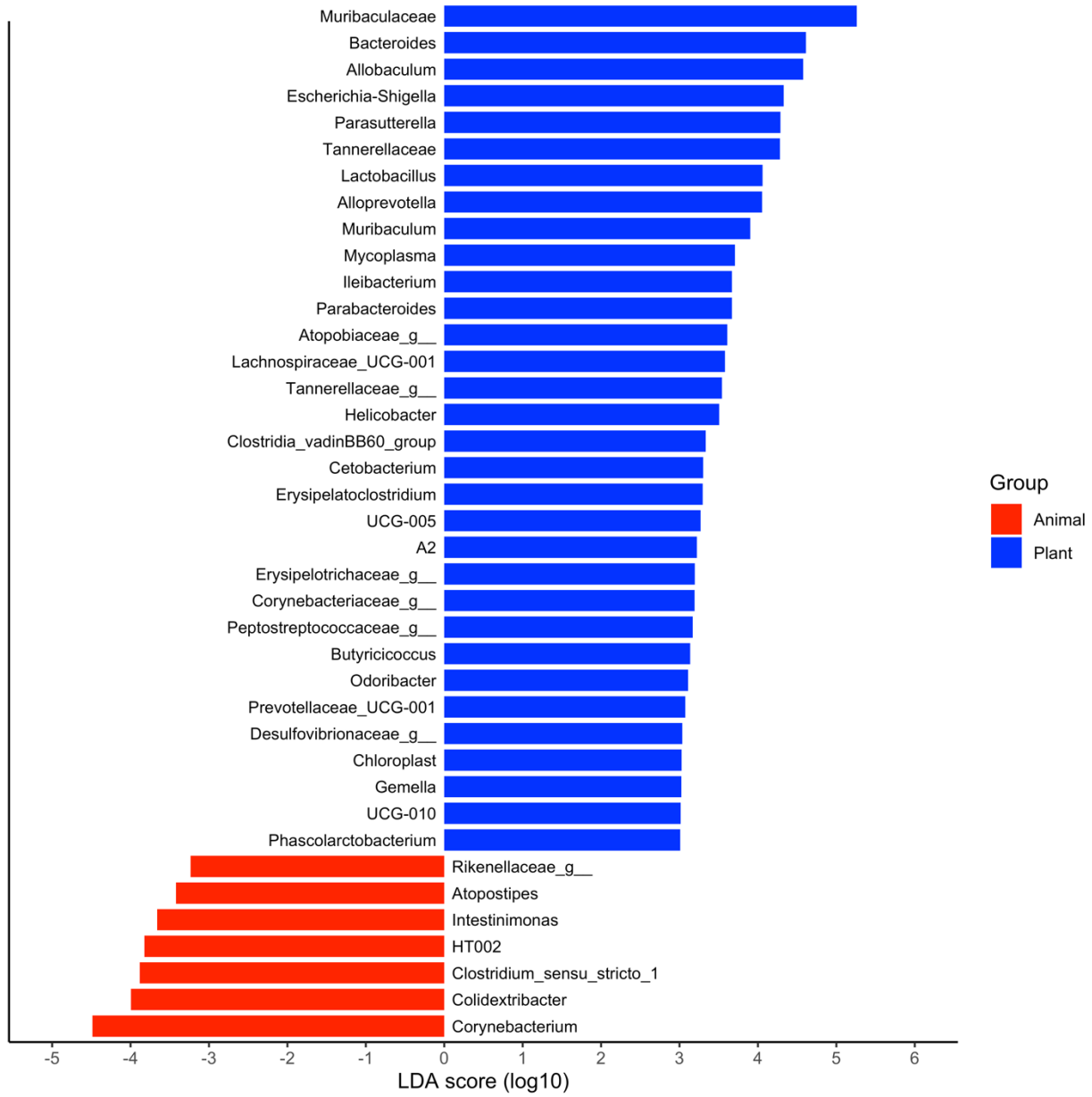
The Random Forest (RF) model achieved a classification accuracy of 96.36% in distinguishing plant- and animal-protein diets, as evidenced by the confusion matrix (Figure 4A). This high predictive capability was driven by microbial taxa strongly associated with each dietary group. The Top 10 amplicon sequence variants (ASVs) identified through Mean Decrease Accuracy (MDA) analysis included *Bacteroides* (5.67%), *Faecalibaculum* (5.63%), *Faecalibaculum* (5.55%), *Faecalibaculum* (5.19%), *Muribaculaceae* (4.91%), *Ileibacterium* (4.82%), *Bifidobacterium* (4.78%), *Parasutterella* (4.76%), *Muribaculaceae* (4.57%), and *Muribaculaceae* (4.57%), (Figure 4B; Supplementary Table 4). To validate the RF model's robustness, a random classification model was generated as a control, achieving an accuracy of only 40.0%. This contrast highlights that the observed RF model performance was driven by biologically meaningful differences rather than random chance (Supplementary Figure 3).

LEfSe analysis provided complementary insights by identifying 33 genera enriched in plant-protein diets, including *Muribaculaceae, Bacteroides, Parasutterella, Parabacteroides,* and *Allobaculum.* In contrast, 7 genera were enriched in animal-protein diets, such as *Corynebacterium, Colidextribacter, Rikenellaceae_g, Atopostipes, Intestinimonas, HT002,* and *Clostridium_sensu_stricto_1* (Figure 4C, Supplementary Table 5). A cladogram generated through LEfSe highlighted distinct phylogenetic patterns, with plant-protein diets linked to reduced *Firmicutes* abundance and increased representation of *Proteobacteria*, *Campylobacterota*, and *Bacteroidota* (Figure 4D). Together, Random Forest and LEfSe analyses provided ecologically relevant insights into the microbial biomarkers associated with dietary protein sources.

These analyses revealed significant differences in the relative abundance of key taxa between dietary groups (Wilcoxon test, $p < 0.05$). For instance, *Bacteroides*, *Muribaculaceae*, and *Parasutterella* were enriched in plant-protein diets, whereas *Bifidobacterium* and *Ileibacterium* were more abundant in animal-protein diets. Notably, *Faecalibacterium* showed no significant differences between the two dietary groups.


A
Confusion Matrix
Plant
Animal
Actual
1 (1.8%)
8 (14.5%)
45 (81.8%)
1 (1.8%)
Frequency
40
30
20
10
Animal
Plant
Predicted
B
Bacteroidaceae
Bifidobacteriaceae
Erysipelotrichaceae
Muribaculaceae
Sutterellaceae
Muribaculaceae
Parasutterella
Bifidobacterium
Ileibacterium
Faecalibaculum
Bacteroides
Top Taxa for classifying
Plant vs Animal Protein
0
5
10
15
Mean Decrease Accuracy (%)

C


Muribaculaceae
Bacteroides
Allobaculum
Escherichia-Shigella
Parasutterella
Tannerellaceae
Lactobacillus
Alloprevotella
Muribaculum
Mycoplasma
Ileibacterium
Parabacteroides
Atopobiaceae_g__
Lachnospiraceae_UCG-001
Tannerellaceae_g__
Helicobacter
Clostridia_vadinBB60_group
Cetobacterium
Erysipelatoclostridium
UCG-005
A2
Erysipelotrichaceae_g__
Corynebacteriaceae_g__
Peptostreptococcaceae_g__
Butyricicoccus
Odoribacter
Prevotellaceae_UCG-001
Desulfovibrionaceae_g__
Chloroplast
Gemella
UCG-010
Phascolarctobacterium
Rikenellaceae_g__
Atopostipes
Intestinimonas
HT002
Clostridium_sensu_stricto_1
Colidextribacter
Corynebacterium
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
LDA score (log10)
Group
Animal
Plant

D

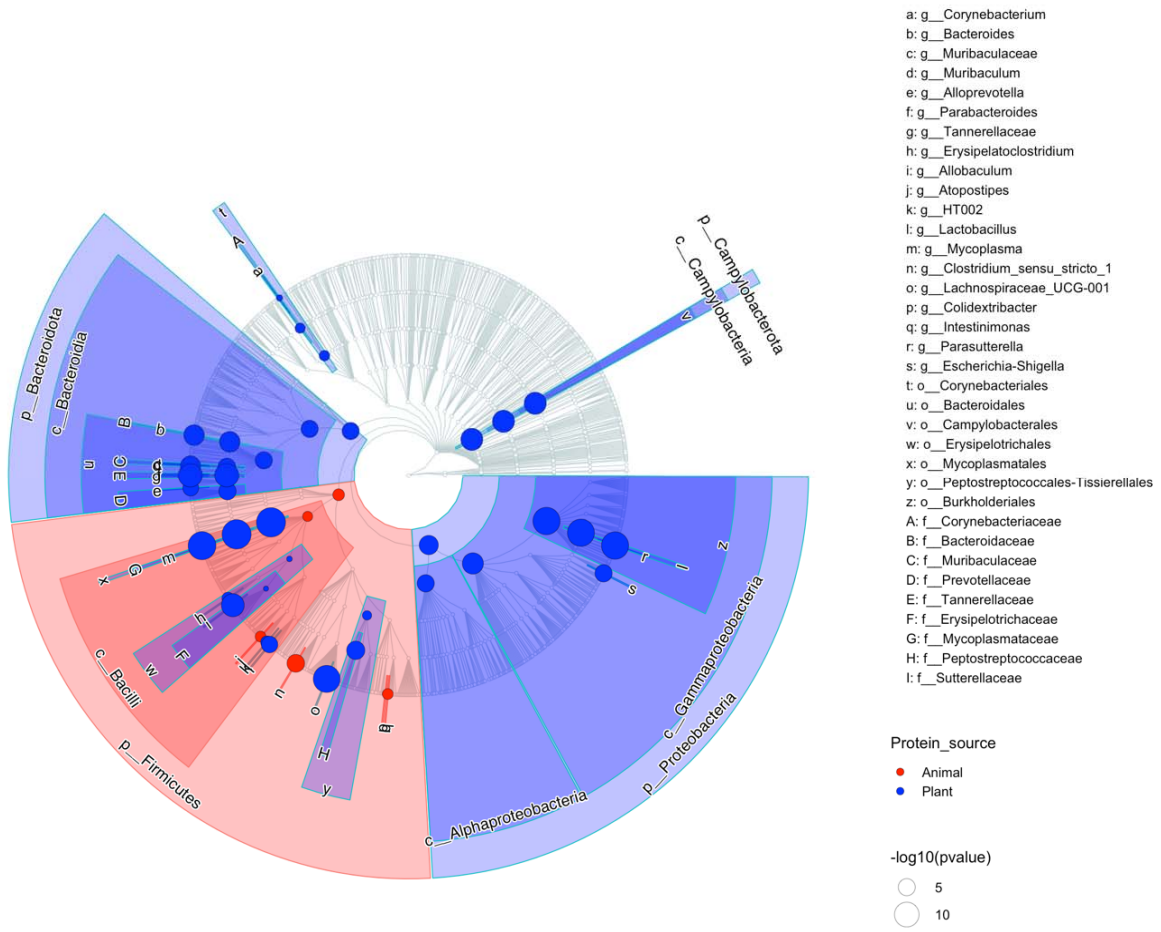


Figure 4. Key Microbial Taxa Identified by Machine Learning and LEfSe

Confusion matrix showing Random Forest (RF) model performance in classifying dietary protein sources (A). The top 10 microbial taxa driving RF accuracy are presented with taxa genera categorized by color (B). LDA using LEfSe highlights taxa enriched in each diet (LDA score ≥ 3.0, $p < 0.05$) (C). Cladogram shows microbial taxa enriched in plant-based (blue) and animal-based (red) diets, arranged hierarchically from phylum to genus with node sizes reflecting statistical significance (-log10[p-value]) (D).

4.0 Discussion

This meta-analysis provides insights into how dietary protein sources shape gut microbiota composition and diversity. While previous studies have largely examined individual protein types, this analysis integrates multiple murine studies to identify consistent microbial patterns associated with plant- and animal-based protein diets. Using differential abundance analysis and statistical modeling, we highlight key microbial taxa linked to distinct protein sources, offering a broader ecological perspective on gut microbial responses to dietary proteins (De Filippis *et al.*, 2016; Canfora *et al.*, 2019). These findings reinforce dietary protein as a major environmental factor influencing gut microbiota composition and functional metabolic interactions.

4.1 Impact of Dietary Protein Sources on Gut Microbial Diversity

Alpha diversity analysis revealed that plant-protein diets were associated with greater microbial richness and evenness, reinforcing prior findings that fiber- and polyphenol-rich diets support a more diverse bacterial community (Singh *et al.*, 2017; Gilbert *et al.*, 2018). In contrast, animal-protein diets exhibited lower microbial diversity, suggesting a more specialized microbiota adapted to nitrogen-rich environments (Conlon, 2014; Tomova *et al.*, 2019).

Further supporting these observations, GLMM analysis identified dietary protein source as the strongest predictor of microbial diversity, with protein level (expressed as % kcal of total energy intake) and host age also playing significant roles. The negative association between geographic origin and diversity suggests potential regional or environmental influences on microbiome responses to dietary protein intake (Madi *et al.*, 2023; Sweeny *et al.*, 2023). Additionally, sequencing depth did not significantly affect diversity estimates, confirming that observed differences were biologically driven rather than technical artifacts (Zaheer *et al.*, 2018; Deng *et al.*, 2022). The high classification accuracy of alpha diversity metrics in ROC analysis further supports microbial diversity as a biomarker for dietary influences on gut microbiota (Li *et al.*, 2022).

Beyond richness and evenness, dietary protein significantly shaped microbial composition. Beta diversity analysis revealed clear compositional differences between plant- and animal-protein diets, suggesting that dietary protein intake influences not only microbial diversity but also microbiota structure and organization. PCA analysis showed distinct clustering, while PERMANOVA and ANOSIM confirmed significant between-group differences. These findings align with previous research showing that plant-based diets promote greater microbial heterogeneity, whereas animal-protein diets support a more uniform microbiome dominated by nitrogen-metabolizing microbes (Wu *et al.*, 2011; Johnson *et al.*, 2019). This suggests that dietary protein not only affects microbial diversity but also drives distinct microbiome structures and ecological niches.

4.2 Microbial Biomarkers and Protein Metabolism

Machine learning (Random Forest) and statistical (LEfSe) approaches enabled the identification of microbial taxa distinguishing plant- and animal-protein diets. Random Forest provided a data-driven ranking of microbial features based on their predictive importance, while LEfSe identified taxa consistently enriched in each dietary group, accounting for both biological significance and effect size (Segata *et al.*, 2011). The high classification accuracy of the Random Forest model (96.36%) suggests that microbial biomarkers can serve as predictive indicators of dietary protein intake, reinforcing their role in gut microbial nitrogen metabolism (Topçuoğlu *et al.*, 2020).

4.2.1 Microbial Adaptations to Plant-Protein Diets

Plant-protein diets were associated with microbes that favor saccharolytic fermentation and nitrogen recycling, promoting microbial diversity and metabolic stability (De Filippis *et al.*, 2016; Canfora *et al.*, 2019). ***Muribaculaceae***, a predominant fiber-fermenting taxon enriched in plant-protein diets, plays a crucial role in gut and systemic health by producing short-chain fatty acids (SCFAs), particularly propionate and butyrate. These metabolites reduce systemic inflammation, enhance metabolic flexibility, and maintain colonic health by supporting mucus layer integrity (Lagkouvardos *et al.*, 2019). Beyond saccharolytic activity, *Muribaculaceae* contributes to bile acid metabolism and aromatic amino acid degradation, underscoring its role in host-microbiota interactions and metabolic regulation (Zhu *et al.*, 2024). In addition, Muribaculaceae facilitates nitrogen recycling by incorporating ammonia and nitrogenous

compounds into biosynthetic pathways, preventing toxic accumulation. Through glutamine synthetase activity, it assimilates ammonia into glutamine, redirecting nitrogen into microbial biomass rather than harmful byproducts like ammonia and hydrogen sulfide (Smith *et al.,* 2022). This function helps stabilize the gut environment and prevent the overgrowth of proteolytic bacteria. Moreover, *Muribaculum intestinale,* a key member of this family, produces cardiolipins that induce antigen-specific cytokine responses, suggesting a role in immune modulation and host defense (Huda *et al.,* 2022).

Similarly, ***Bacteroides***, a key genus in the gut microbiota, thrives on dietary fibers and resistant starches from plant-protein diets, driving the production of short-chain fatty acids (SCFAs) such as butyrate, acetate, and propionate (Facchin *et al.,* 2024). These SCFAs play essential roles in reducing gut permeability, maintaining gut barrier integrity, and modulating immune responses. Additionally, Bacteroides metabolizes aromatic amino acids (ArAAs) into anti-inflammatory compounds, such as 4-hydroxyphenylacetate, which support metabolic health and may offer protection against type 2 diabetes and inflammatory bowel disease (Canfora *et al.,* 2019; De Filippis *et al.,* 2016). Beyond saccharolytic fermentation, *Bacteroides* exhibits metabolic flexibility, adapting its energy metabolism to dietary composition. Under fiber-rich conditions, it prioritizes polysaccharide fermentation, producing SCFAs that lower gut pH, which helps inhibit proteolytic bacteria and reduce nitrogenous waste accumulation. Moreover, Bacteroides engages in cross-feeding interactions, where its metabolic byproducts serve as substrates for secondary fermenters, directing excess nitrogen into microbial biomass rather than putrefactive byproducts (Ríos-Covian *et al.,* 2017). This activity stabilizes the carbon-to-nitrogen ratio, preventing nitrogen overload and limiting the formation of harmful metabolites. Conversely, diets low in fiber or high in protein and fat can shift *Bacteroides* metabolism toward proteolytic pathways, leading to increased ammonia and hydrogen sulfide production, which may contribute to gut dysbiosis and inflammation (Zhao *et al.,* 2021).

Furthermore, ***Allobaculum***, a plant-protein-associated taxon, is notable for its and dietary responsiveness and ability to promote balanced gut microbiome. Its abundance positively correlates with lipid metabolism, particularly through angiopoietin-like protein 4 (ANGPTL4) expression in response to high-fat diets (Zheng *et al.,* 2021). However, its role in protein metabolism remains poorly understood, with limited evidence on its involvement in nitrogen assimilation or SCFA production. ***Parasutterella***, enriched in plant-protein diets, influences bile acid metabolism, indirectly regulating microbial nitrogen turnover (Ju *et al.,* 2019). Specifically, *Parasutterella* modulates bile acid conversion, affecting nitrogenous substrate availability for microbial growth and reducing ammonia accumulation. By modifying bile acid pools, *Parasutterella* also limits sulfur-rich substrate availability, thereby reducing hydrogen sulfide production by proteolytic bacteria (Ju *et al.,* 2019).

4.2.2 Microbial Adaptations to Animal-Protein Diets

In contrast, animal-protein diets were associated with microbial taxa that specialize in proteolytic fermentation and nitrogen metabolism, contributing to potentially harmful metabolite production. ***Clostridium_sensu_stricto_1***, a dominant taxon in animal-protein-fed microbiomes, utilizes Stickland fermentation, a pathway that oxidizes and reduces branched-chain and sulfur-containing amino acids, producing ammonia, hydrogen sulfide (H_2S), and branched-chain fatty acids (BCFAs). These metabolic byproducts inhibit butyrate oxidation in colonocytes, disrupt

epithelial tight junctions, and increase oxidative stress, ultimately compromising gut barrier integrity and promoting inflammation (Ridlon *et al.,* 2016; Zhao *et al.,* 2021). Similarly, ***Colidextribacter,*** identified through Random Forest and LEfSe analysis, is increasingly associated with proteolytic fermentation-driven dysbiosis. Studies indicate that it thrives in high-protein dietary environments, where it contributes to gut barrier disruption and microbial imbalance. Notably, *Colidextribacter* metabolizes aromatic amino acids into p-cresol and indoxyl sulfate, compounds linked to microbial dysbiosis and systemic inflammation (Singh *et al.,* 2023). These metabolites, generated via the tryptophan and tyrosine degradation pathways, accumulate in the gut and may be reabsorbed into the bloodstream, acting as uremic toxins. Li et al. (2023) further reported that its enrichment correlates with elevated inflammatory markers, including triglycerides (TG) and glutamic pyruvic transaminase (GPT), indicators of fat accumulation and metabolic stress. In this study, animal-protein diets significantly enriched Colidextribacter, consistent with Zhang et al. (2023), who identified it as a key driver of microbial ecosystem imbalance under proteolytic conditions. Additionally, ***Rikenellaceae_g***, another biomarker of animal-protein diets, participates in nitrogen metabolism and bile acid transformation. This taxon deaminates amino acids into ammonia and amines, increasing colonic pH and creating an environment favorable for alkaline-tolerant proteolytic bacteria. Furthermore, it contributes to secondary bile acid synthesis, intensifying microbial competition and shifting the gut microbiome toward proteolytic rather than saccharolytic metabolism (Yang *et al.,* 2019).

4.3 Implications of Dietary Protein on Gut Microbial Ecology

These findings suggest that plant-protein diets optimize nitrogen utilization, reducing the accumulation of toxic nitrogenous byproducts and enhancing microbial diversity. Conversely, animal-protein diets promote proteolytic fermentation, leading to increased production of ammonia, hydrogen sulfide, and p-cresol, metabolites linked to gut barrier dysfunction and inflammation. The integration of machine learning-based microbial classification provides robust microbial signatures that emphasize the ecological impact of dietary protein sources.

4.4 Limitations and Future Directions

Despite these insights, this study has several limitations. The reliance on murine models limits direct applicability to human microbiota, as gut microbial composition and metabolic interactions differ between species (Nguyen *et al.,* 2015). Additionally, the lack of direct metabolomic measurements restricts the functional validation of microbial fermentation pathways. Future studies should integrate longitudinal human trials, metatranscriptomics, and metabolomics to further elucidate functional microbial responses to dietary proteins (Zhao *et al.,* 2019).

5.0 Conclusion

This meta-analysis highlights dietary protein as a key determinant of gut microbiota composition, with plant-protein diets fostering microbial diversity and animal-protein diets favoring proteolytic fermentation pathways. Future research integrating functional analyses and metabolomic profiling will be essential to fully understand the implications of these microbial adaptations for host metabolic health.

Author Contributions

JHM and SA designed the study. SA, SG and PD were involved in data processing. SA analyzed the data, created figures, and wrote the manuscript. CKYH and JA reviewed the manuscript. All authors participated in interpreting data, revised the manuscript, and agreed to published version of the manuscript.

Funding

This research received no external funding.

Data Availability Statement

Data and bioinformatic workflows are publicly available here:

https://github.com/deejayprof/META_ANALYSIS

Acknowledgements

We acknowledge the support provided by various R libraries that were instrumental in the analysis conducted for this study. Comprehensive details about the data, code, and resources utilized are provided in the data availability section.

Conflicts of Interest

The authors declare no conflicts of interest and affirm that the study's results are presented transparently, accurately, and free from any fabrication, falsification, or inappropriate data manipulation.

Abbreviations

PRISMA: Preferred Reporting Items for Systematic Review and Meta-Analysis

RF: Random Forest

LDA: Linear Discriminant Analysis

SCFAs: Short-Chain Fatty Acids

BCFAs: Branched-Chain Fatty Acids

ANOVA: Analysis of Variance

ROC: Receiver Operating Characteristic

AUC: Area Under the Curve

ACE: Abundance-based Coverage Estimator

PERMANOVA: Permutational Multivariate Analysis of Variance

PCoA: Principal Coordinates Analysis

GLMM: Generalized Linear Mixed Model

LEfSe: Linear Discriminant Analysis Effect Size

qPCR: Quantitative Polymerase Chain Reaction

BCAAs: Branched-Chain Amino Acids

ArAAs: Aromatic Amino Acids

ANOSIM - Analysis of Similarities

ASVs - Amplicon Sequence Variants

rRNA: Ribosomal Ribonucleic Acid

NCBI: National Center for Biotechnology Information

SRA: Sequence Read Archive

ENA: European Nucleotide Archive

QIIME2 - Quantitative Insights Into Microbial Ecology 2

DADA2 - Divisive Amplicon Denoising Algorithm 2

GEMs: Genome-scale metabolic models

MDA: Mean Decrease Accuracy

EAAs: Essential amino acids

References

1. Rowland, I., Gibson, G., Heinken, A., Scott, K., Swann, J., Thiele, I., & Tuohy, K. (2018). Gut microbiota functions: metabolism of nutrients and other food components. European journal of nutrition, 57(1), 1–24
2. Gilbert, M. S., Ijssennagger, N., Kies, A. K., & van Mil, S. W. C. (2018). Protein fermentation in the gut; implications for intestinal dysfunction in humans, pigs, and poultry. American journal of physiology. Gastrointestinal and liver physiology, 315(2), G159–G170.
3. Zhao, J., Zhang, X., Liu, H., Brown, M. A., & Qiao, S. (2019). Dietary Protein and Gut Microbiota Composition and Function. Current protein & peptide science, 20(2), 145–154. https://doi.org/10.2174/1389203719666180514145437
4. Ling, ZN., Jiang, YF., Ru, JN. *et al.*, . Amino acid metabolism in health and disease. Sig Transduct Target Ther 8, 345 (2023). https://doi.org/10.1038/s41392-023-01569-3
5. Kitada, M., Ogura, Y., Monno, I., & Koya, D. (2019). The impact of dietary protein intake on longevity and metabolic health. EBioMedicine, 43, 632–640.
6. Vinelli, V., Biscotti, P., Martini, D., Del Bo', C., Marino, M., Meroño, T., Nikoloudaki, O., Calabrese, F. M., Turroni, S., Taverniti, V., Unión Caballero, A., Andrés-Lacueva, C., Porrini, M., Gobbetti, M., De Angelis, M., Brigidi, P., Pinart, M., Nimptsch, K., Guglielmetti, S., & Riso, P. (2022). Effects of Dietary Fibers on Short-Chain Fatty Acids and

Gut Microbiota Composition in Healthy Adults: A Systematic Review. Nutrients, 14(13), 2559.

7. den Besten, G., van Eunen, K., Groen, A. K., Venema, K., Reijngoud, D. J., & Bakker, B. M. (2013). The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. Journal of lipid research, 54(9), 2325–2340. https://doi.org/10.1194/jlr.R036012
8. Ríos-Covián, D., Ruas-Madiedo, P., Margolles, A., Gueimonde, M., de Los Reyes-Gavilán, C. G., & Salazar, N. (2016). Intestinal Short Chain Fatty Acids and their Link with Diet and Human Health. Frontiers in microbiology, 7, 185.
9. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., … Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ (Clinical research ed.), 372, n71.
10. 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., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., Cope, E. K., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature biotechnology, 37(8), 852–857.
11. Ziemski, M., Adamov, A., Kim, L., Flörl, L., & Bokulich, N. A. (2022). Reproducible acquisition, management and meta-analysis of nucleotide sequence (meta)data using q2-fondue. Bioinformatics (Oxford, England), 38(22), 5081–5091.
12. Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. Nature methods, 13(7), 581–583.
13. Bokulich, N.A., Kaehler, B.D., Rideout, J.R. *et al.*, . Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. Microbiome 6, 90 (2018).
14. Wen, T., Niu, G., Chen, T., Shen, Q., Yuan, J., & Liu, Y. X. (2023). The best practice for microbiome analysis using R. Protein & cell, 14(10), 713–725.
15. McMurdie PJ, Holmes S. Phyloseq: a bioconductor package for handling and analysis of high-throughput phylogenetic sequence data. Pac Symp Biocomput. 2012:235-46. PMID: 22174279; PMCID: PMC3357092.
16. Samuthpongtorn, C., Nopsopon, T., & Pongpirul, K. (2022). Gut microbiome diversity measures for metabolic conditions: a systematic scoping review. Journal of Physiological and Biomedical Sciences , 33 (2). Retrieved from
17. Manandhar, I., Alimadadi, A., Aryal, S., Munroe, P. B., Joe, B., & Cheng, X. (2021). Gut microbiome-based supervised machine learning for clinical diagnosis of inflammatory bowel diseases. American journal of physiology. Gastrointestinal and liver physiology, 320(3), G328–G337.
18. Ley, R. E., Turnbaugh, P. J., Klein, S., & Gordon, J. I. (2006). Microbial ecology: human gut microbes associated with obesity. Nature, 444(7122), 1022–1023.

19. Walters, W. A., Xu, Z., & Knight, R. (2014). Meta-analyses of human gut microbes associated with obesity and IBD. FEBS letters, 588(22), 4223–4233.

20. De Filippis, F., Pellegrini, N., Vannini, L., Jeffery, I. B., La Storia, A., Laghi, L., Serrazanetti, D. I., Di Cagno, R., Ferrocino, I., Lazzi, C., Turroni, S., Cocolin, L., Brigidi, P., Neviani, E., Gobbetti, M., O'Toole, P. W., & Ercolini, D. (2016). High-level adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated metabolome. Gut, 65(11), 1812–1821.

21. Canfora, E. E., Meex, R. C. R., Venema, K., & Blaak, E. E. (2019). Gut microbial metabolites in obesity, NAFLD and T2DM. Nature reviews. Endocrinology, 15(5), 261–273.

22. Sweeny, A. R., Lemon, H., Ibrahim, A., Watt, K. A., Wilson, K., Childs, D. Z., Nussey, D. H., Free, A., & McNally, L. (2023). A mixed-model approach for estimating drivers of microbiota community composition and differential taxonomic abundance. mSystems, 8(4), e0004023.

23. Madi, N., Chen, D., Wolff, R., Shapiro, B. J., & Garud, N. R. (2023). Community diversity is associated with intra-species genetic diversity and gene loss in the human gut microbiome. eLife, 12, e78530.

24. Zaheer, R., Noyes, N., Ortega Polo, R., Cook, S. R., Marinier, E., Van Domselaar, G., Belk, K. E., Morley, P. S., & McAllister, T. A. (2018). Impact of sequencing depth on the characterization of the microbiome and resistome. Scientific reports, 8(1), 5890.

25. Deng, X. L., Frandsen, P. B., Dikow, R. B., Favre, A., Shah, D. N., Shah, R. D. T., Schneider, J. V., Heckenhauer, J., & Pauls, S. U. (2022). The impact of sequencing depth and relatedness of the reference genome in population genomic studies: A case study with two caddisfly species (Trichoptera, Rhyacophilidae, Himalopsyche). Ecology and evolution, 12(12), e9583.

26. Li, Z., Zhou, J., Liang, H., Ye, L., Lan, L., Lu, F., Wang, Q., Lei, T., Yang, X., Cui, P., & Huang, J. (2022). Differences in Alpha Diversity of Gut Microbiota in Neurological Diseases. Frontiers in neuroscience, 16, 879318.

27. Singh, R. K., Chang, H. W., Yan, D., Lee, K. M., Ucmak, D., Wong, K., Abrouk, M., Farahnik, B., Nakamura, M., Zhu, T. H., Bhutani, T., & Liao, W. (2017). Influence of diet on the gut microbiome and implications for human health. Journal of translational medicine, 15(1), 73.

28. Gilbert, J. A., Blaser, M. J., Caporaso, J. G., Jansson, J. K., Lynch, S. V., & Knight, R. (2018). Current understanding of the human microbiome. Nature medicine, 24(4), 392–400.

29. Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., & Huttenhower, C. (2011). Metagenomic biomarker discovery and explanation. Genome biology, 12(6), R60.

30. Tomova, A., Bukovsky, I., Rembert, E., Yonas, W., Alwarith, J., Barnard, N. D., & Kahleova, H. (2019). The Effects of Vegetarian and Vegan Diets on Gut Microbiota. Frontiers in nutrition, 6, 47.

31. Rios-Covian, D., González, S., Nogacka, A. M., Arboleya, S., Salazar, N., Gueimonde, M., & de Los Reyes-Gavilán, C. G. (2020). An Overview on Fecal Branched Short-Chain Fatty

Acids Along Human Life and as Related With Body Mass Index: Associated Dietary and Anthropometric Factors. Frontiers in microbiology, 11, 973.

32. Rios-Covian, D., Salazar, N., Gueimonde, M., & de Los Reyes-Gavilan, C. G. (2017). Shaping the Metabolism of Intestinal *Bacteroides* Population through Diet to Improve Human Health. Frontiers in microbiology, 8, 376. https://doi.org/10.3389/fmicb.2017.00376

33. Ju, T., Kong, J. Y., Stothard, P., & Willing, B. P. (2019). Defining the role of *Parasutterella*, a previously uncharacterized member of the core gut microbiota. The ISME journal, 13(6), 1520–1534. https://doi.org/10.1038/s41396-019-0364-5

34. Johnson, A. J., Vangay, P., Al-Ghalith, G. A., *et al.*, . (2019). Daily sampling reveals personalized diet-microbiome associations in humans. Cell Host & Microbe, 25(6), 789-802.e5. https://doi.org/10.1016/j.chom.2019.05.005

35. Yang, W. Y., Lee, Y., Lu, H., Chou, C. H., & Wang, C. (2019). Analysis of gut microbiota and the effect of lauric acid against necrotic enteritis in Clostridium perfringens and Eimeria side-by-side challenge model. PloS one, 14(5), e0205784.

36. Zhao, L., Zhang, F., Ding, X., *et al.*, . (2019). Gut bacteria selectively promoted by dietary fibers alleviate type 2 diabetes. Science, 359(6380), 1151-1156.

37. Koeth, R. A., Wang, Z., Levison, B. S., Buffa, J. A., Org, E., Sheehy, B. T., Britt, E. B., Fu, X., Wu, Y., Li, L., Smith, J. D., DiDonato, J. A., Chen, J., Li, H., Wu, G. D., Lewis, J. D., Warrier, M., Brown, J. M., Krauss, R. M., Tang, W. H., … Hazen, S. L. (2013). Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nature medicine*, *19*(5), 576–585. https://doi.org/10.1038/nm.3145

38. Tang, Z. Z., Chen, G., & Alekseyenko, A. V. (2016). PERMANOVA-S: association test for microbial community composition that accommodates confounders and multiple distances. Bioinformatics (Oxford, England), 32(17), 2618–2625. https://doi.org/10.1093/bioinformatics/btw311

39. Zheng, Z., Lyu, W., Ren, Y., Li, X., Zhao, S., Yang, H., & Xiao, Y. (2021). Allobaculum Involves in the Modulation of Intestinal ANGPTLT4 Expression in Mice Treated by High-Fat Diet. Frontiers in nutrition, 8, 690138. https://doi.org/10.3389/fnut.2021.690138

40. Scott, K. P., Gratz, S. W., Sheridan, P. O., Flint, H. J., & Duncan, S. H. (2013). The influence of diet on the gut microbiota. *Pharmacological research*, *69*(1), 52–60. https://doi.org/10.1016/j.phrs.2012.10.020

41. Wu, G. D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y. Y., Keilbaugh, S. A., Bewtra, M., Knights, D., Walters, W. A., Knight, R., Sinha, R., Gilroy, E., Gupta, K., Baldassano, R., Nessel, L., Li, H., Bushman, F. D., & Lewis, J. D. (2011). Linking long-term dietary patterns with gut microbial enterotypes. *Science (New York, N.Y.)*, *334*(6052), 105–108. https://doi.org/10.1126/science.1208344

42. Beaumont, M., Portune, K. J., Steuer, N., Lan, A., Cerrudo, V., Audebert, M., Dumont, F., Mancano, G., Khodorova, N., Andriamihaja, M., Airinei, G., Tomé, D., Benamouzig, R., Davila, A. M., Claus, S. P., Sanz, Y., & Blachier, F. (2017). Quantity and source of dietary protein influence metabolite production by gut microbiota and rectal mucosa gene expression: a randomized, parallel, double-blind trial in overweight humans. *The American journal of clinical nutrition*, *106*(4), 1005–1019. https://doi.org/10.3945/ajcn.117.158816

43. Rooks, M. G., & Garrett, W. S. (2016). Gut microbiota, metabolites and host immunity. *Nature reviews. Immunology*, *16*(6), 341–352. https://doi.org/10.1038/nri.2016.42

44. Singh, R. K., Chang, H. W., Yan, D., Lee, K. M., Ucmak, D., Wong, K., Abrouk, M., Farahnik, B., Nakamura, M., Zhu, T. H., Bhutani, T., & Liao, W. (2017). Influence of diet on the gut microbiome and implications for human health. *Journal of translational medicine*, *15*(1), 73. https://doi.org/10.1186/s12967-017-1175-y

45. Gilbert, J. A., Blaser, M. J., Caporaso, J. G., Jansson, J. K., Lynch, S. V., & Knight, R. (2018). Current understanding of the human microbiome. *Nature medicine*, *24*(4), 392–400. https://doi.org/10.1038/nm.4517

46. Conlon, M. A., & Bird, A. R. (2014). The impact of diet and lifestyle on gut microbiota and human health. *Nutrients*, *7*(1), 17–44. https://doi.org/10.3390/nu7010017

47. Topçuoğlu BD, Lesniak NA, Ruffin MT, Wiens J, Schloss PD. 2020. A Framework for Effective Application of Machine Learning to Microbiome-Based Classification Problems. mBio11:10.1128/mbio.00434-20. https://doi.org/10.1128/mbio.00434-20

48. Ridlon, J. M., Harris, S. C., Bhowmik, S., Kang, D. J., & Hylemon, P. B. (2016). Consequences of bile salt biotransformations by intestinal bacteria. *Gut microbes*, *7*(1), 22–39. https://doi.org/10.1080/19490976.2015.1127483

49. Tan, J., Ni, D., Taitz, J., Pinget, G. V., Read, M., Senior, A., Wali, J. A., Elnour, R., Shanahan, E., Wu, H., Chadban, S. J., Nanan, R., King, N. J. C., Grau, G. E., Simpson, S. J., & Macia, L. (2022). Dietary protein increases T-cell-independent sIgA production through changes in gut microbiota-derived extracellular vesicles. *Nature communications*, *13*(1), 4336. https://doi.org/10.1038/s41467-022-31761-y

50. Masuoka, H., Suda, W., Tomitsuka, E., Shindo, C., Takayasu, L., Horwood, P., Greenhill, A. R., Hattori, M., Umezaki, M., & Hirayama, K. (2020). The influences of low protein diet on the intestinal microbiota of mice. *Scientific reports*, *10*(1), 17077. https://doi.org/10.1038/s41598-020-74122-9

51. Yin, Y., Cai, J., Zhou, L., Xing, L., & Zhang, W. (2022). Dietary oxidized beef protein alters gut microbiota and induces colonic inflammatory damage in C57BL/6 mice. *Frontiers in nutrition*, *9*, 980204. https://doi.org/10.3389/fnut.2022.980204

52. Zhu, Y., He, H., Tang, Y., Peng, Y., Hu, P., Sun, W., Liu, P., Jin, M., & Xu, X. (2022). Reno-Protective Effect of Low Protein Diet Supplemented With α-Ketoacid Through Gut Microbiota and Fecal Metabolism in 5/6 Nephrectomized Mice. *Frontiers in nutrition*, *9*, 889131. https://doi.org/10.3389/fnut.2022.889131

53. Xie, Y., Wang, C., Zhao, D., Zhou, G., & Li, C. (2020). Processing Method Altered Mouse Intestinal Morphology and Microbial Composition by Affecting Digestion of Meat Proteins. *Frontiers in microbiology*, *11*, 511. https://doi.org/10.3389/fmicb.2020.00511

54. Choi, B. S., Daniel, N., Houde, V. P., Ouellette, A., Marcotte, B., Varin, T. V., Vors, C., Feutry, P., Ilkayeva, O., Ståhlman, M., St-Pierre, P., Bäckhed, F., Tremblay, A., White, P. J., & Marette, A. (2021). Feeding diversified protein sources exacerbates hepatic insulin resistance via increased gut microbial branched-chain fatty acids and mTORC1 signaling in obese mice. *Nature communications*, *12*(1), 3377. https://doi.org/10.1038/s41467-021-23782-w

55. Ijaz, M. U., Ahmad, M. I., Hussain, M., Khan, I. A., Zhao, D., & Li, C. (2020). Meat Protein in High-Fat Diet Induces Adipogensis and Dyslipidemia by Altering Gut Microbiota and Endocannabinoid Dysregulation in the Adipose Tissue of Mice. *Journal of agricultural and food chemistry*, *68*(13), 3933–3946. https://doi.org/10.1021/acs.jafc.0c00017

56. Zhong, W., Wang, H., Yang, Y., Zhang, Y., Lai, H., Cheng, Y., Yu, H., Feng, N., Huang, R., Liu, S., Yang, S., Hao, T., Zhang, B., Ying, H., Zhang, F., Guo, F., & Zhai, Q. (2022). High-protein diet prevents fat mass increase after dieting by counteracting Lactobacillus-enhanced lipid absorption. *Nature metabolism*, *4*(12), 1713–1731. https://doi.org/10.1038/s42255-022-00687-6

57. Zhang, R., Mu, H., Li, Z., Zeng, J., Zhou, Q., Li, H., Wang, S., Li, X., Zhao, X., Sun, L., Chen, W., Dong, J., & Yang, R. (2022). Oral administration of branched-chain amino acids ameliorates high-fat diet-induced metabolic-associated fatty liver disease *via* gut microbiota-associated mechanisms. *Frontiers in microbiology*, *13*, 920277. https://doi.org/10.3389/fmicb.2022.920277

58. Ruocco, C., Ragni, M., Rossi, F., Carullo, P., Ghini, V., Piscitelli, F., Cutignano, A., Manzo, E., Ioris, R. M., Bontems, F., Tedesco, L., Greco, C. M., Pino, A., Severi, I., Liu, D., Ceddia, R. P., Ponzoni, L., Tenori, L., Rizzetto, L., Scholz, M., … Nisoli, E. (2020). Manipulation of Dietary Amino Acids Prevents and Reverses Obesity in Mice Through Multiple Mechanisms That Modulate Energy Homeostasis. *Diabetes*, *69*(11), 2324–2339. https://doi.org/10.2337/db20-0489