Understanding the role of the human gut microbiome in mediating chronic disease and healthy aging

Johannes P. Johnson-Martínez

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2024

Reading Committee:
Sean M. Gibbons, Chair
Nancy L. Allbritton
Julia Y. Cui

Program Authorized to Offer Degree:
Bioengineering

University of Washington

Abstract

Understanding the role of the human gut microbiome in mediating chronic disease and healthy aging

Johannes P. Johnson-Martínez

Chair of the Supervisory Committee:
Sean Gibbons
Department of Bioengineering

The human gut microbiome is considered by many clinicians and researchers to be one of the final frontiers of discovery in medicine, or a “second brain” for the body. The gut microbiome encodes 100-fold more genes than the human genome and many of these foreign genes influence our phenotypes through the production of small molecules, like neurotransmitters or short-chain fatty acids (SCFAs). The bidirectional relationship between the human gut microbiota and the host, including the gut-brain axis, gut-kidney-heart axis, and gut-liver axis, comprise a complex ecosystem that influences aging and long-term health. Here, I describe my dissertation work, which focused on signatures of chronic disease and healthy aging in the body and how these signatures were related to ecological variation in the gut microbiome. I leveraged deep phenotyping data from the Arivale Scientific Wellness Program, which concluded in 2019. I found significant associations between microbiome composition and function, bowel movement frequency, frailty, and biomarkers of pro-inflammatory diseases and chronic kidney disease (CKD). Uremic toxins and biomarkers of inflammation were elevated in otherwise healthy individuals experiencing constipation or diarrhea, respectively. Microbially-derived uremic toxins enriched in the blood of constipated individuals were also negatively associated with kidney function. Enriched inflammatory proteins and genera, as well as immunological metabolites found in the Arivale cohort blood plasma imply a stressful cellular environment in the frail. Frailty negatively associates with estimated SCFA (propionate) production independent of age. These findings reveal some of the biological networks and mechanisms underlying the previously observed relationships between gut microbiome composition, healthy aging, and chronic disease development, warranting further investigation into the role of the gut microbiome as a mediator of health and disease throughout the lifespan.

DEDICATION

Every keystroke typed and synapse of mine fired during the assembly of this doctoral work and the journey toward completing this dissertation is dedicated to two pivotal figures in my life: to my mother, Victoria Johnson and in memory to my late grandmother, Rosa Rodriguez, who passed away during the completion of this work.

Rosa was born in 1928 and never able to attain a formal fourth grade education in rural Colombia, but she trained as an x-ray assistant during the tuberculosis epidemic in that region, and came home every day to throw away her clothing so as to not infect her family. She worked hard to make sure that her children had food to eat, and it was often not enough. She taught me and my family that education means everything for opportunity, not just for survival, but also for serving a gifted purpose in this world.

My mother, Victoria, suffered a toxic, whole-body chemical poisoning in the early years of her chemical engineering career (as one of the first female chemical engineers in Colombia) that left her disabled in a coma, and she almost did not survive. Thankfully, she recovered to an extent. After she had me, we fled Colombia during the Pablo Escobar revolt in the early 90s and came to the United States of America as first generation immigrants.

Over 20 years, I cared for my grandmother, Rosa, after her diagnosis of vascular dementia, watching her slowly decline over many years with her. Since her poisoning, I have also watched my mother, Victoria, decline from neurological disease to the point that she needs multiple spinal surgeries and is permanently disabled. Throughout their illnesses, they had never relented in supporting my education, my future, and my health and wellbeing above all. They were my pillars of strength underneath me as I increased my level of education despite many obstacles.

These two women always protected me, prioritized me, and raised me to value science, knowledge, integrity, and perseverance. It is because of my mother that I was inspired to become a chemical engineer, just like her. It is because of caring for my grandmother over many years of her life with dementia that I was inspired to pursue graduate school and study chronic diseases related to aging. In the late days of her dementia, she called me “Doctor”, because of how much I took care of her, as if to speak the future into existence.

Everything I am and do in my life is because of these two people. It is my honor to dedicate my doctoral dissertation and research to these heroes in my life.

TABLE OF CONTENTS

- Abstract
- Specific Aims
- Central Hypotheses
- Key Research Approach
- Abbreviations
- Chapter 1: Bowel movement frequency, the gut microbiome, and host molecular phenotypes
- Chapter 2: Frailty, healthy aging, and the gut microbiome
- Conclusion
- Bibliography

DEDICATION 5
TABLE OF CONTENTS 7
ABSTRACT 10
SPECIFIC AIMS 11
Specific Aim #1: 11
  ● Rationale 11
  ● Expected Outcomes 11
Specific Aim #2: 11
  ● Rationale 12
  ● Expected Outcomes 12
CENTRAL HYPOTHESES 12
KEY RESEARCH APPROACH 12
  Data Collection and Analysis 13
  Expected Outcomes 13
ABBREVIATIONS 13
CHAPTER 1: BOWEL MOVEMENT FREQUENCY, THE GUT MICROBIOME, AND HOST MOLECULAR PHENOTYPES 16
1.1 SUMMARY 16
1.2 INTRODUCTION 16
1.3 RESULTS 18

A cohort of generally-healthy individuals.......... 18
Gut microbiome structure and composition across BMF categories.......... 20
Variation in blood metabolites across BMF categories.......... 21
Blood plasma chemistries across BMF categories.......... 22
Blood proteomics across BMF categories.......... 22
Self-reported diet, lifestyle, anxiety and depression histories associated with BMF categories and demographic covariates.......... 22
BMF-associated blood metabolites associated with kidney function in a generally-healthy cohort.......... 24
1.4 DISCUSSION.......... 25
Study limitations.......... 28
1.5 FIGURE TITLES AND LEGENDS.......... 30
Figure 1.......... 30
Figure 2.......... 31
Figure 3.......... 32
Figure 4.......... 33
Figure 5.......... 34
Figure 6.......... 35
Figure 7.......... 36
1.6 SUPPLEMENTAL FIGURES AND TABLES.......... 37
Table S1.......... 37
Figure S1.......... 38
Figure S2.......... 39
Figure S3.......... 40
Figure S4.......... 41
Figure S5.......... 42
Figure S6.......... 43
1.7 METHODS.......... 44
Resource availability.......... 44
Lead contact.......... 44
Materials availability.......... 44
Key Resources Table.......... 44
Data and code availability.......... 45
1.8 EXPERIMENTAL MODEL AND SUBJECT DETAILS.......... 45
Institutional review board approval for the study.......... 45
Generally-healthy cohort.......... 45
1.9 METHODS.......... 47
Gut microbiome data.......... 47
Olink proteomics.......... 49
Metabolon metabolomics.......... 49
Blood plasma chemistries.......... 51

Questionnaire data.. 52
Depression and anxiety health history data...52
1.10 QUANTIFICATION AND STATISTICAL ANALYSIS...53
Statistical analyses.. 53
CHAPTER 2: FRAILTY, HEALTHY AGING, AND THE GUT MICROBIOME.............................55
2.1 SUMMARY.. 55
2.2 INTRODUCTION.. 56
2.3 RESULTS.. 57
FIS and multi-omics analysis... 57
Frailty associations with microbiome uniqueness, α-diversity, and estimated SCFA production fluxes...58
Significant genera variation across frailty.. 58
Significant blood metabolomics variation across frailty.. 59
Significant blood proteomics variation across frailty... 59
2.4 DISCUSSION... 59
Frailty associations with uniqueness, alpha-diversity, and estimated SCFA fluxes.............. 59
Differentially abundant genera linked to varying SCFA production and pathobiont enrichment 60
Blood metabolites associated with amino acids, inflammation, and lipid metabolism........... 61
Blood proteomics associated with inflammation.. 62
Study limitations..64
2.5 FIGURE TITLES AND LEGENDS...65
Figure 8..65
Figure 9..66
Figure 10..67
Figure 11..68
Figure 12..69
2.7 METHODS...70
Data and code availability..70
Institutional review board approval for the study... 70
Frailty cohort selection...70
2.8 QUANTIFICATION AND STATISTICAL ANALYSIS...71
Statistical analyses.. 71
FIS.. 72
CONCLUSION...73
BIBLIOGRAPHY.. 75

ABSTRACT

The human gut microbiome is considered by many clinicians and researchers to be one of the final frontiers of discovery in medicine, or a “second brain” for the body [1] . The gut microbiome encodes 100-fold more genes than the human genome and many of these foreign genes influence our phenotypes through the production of small molecules, like neurotransmitters or short-chain fatty acids (SCFAs) [2,3] . The bidirectional relationship between the human gut microbiota and the host, including the gut-brain axis, gut-kidney-heart axis, and gut-liver axis, comprise a complex ecosystem that influences aging and long-term health [4,5] . Here, I describe my dissertation work, which focused on signatures of chronic disease and healthy aging in the body and how these signatures were related to ecological variation in the gut microbiome. I leveraged deep phenotyping data from the Arivale Scientific Wellness Program, which concluded in 2019. I found significant associations between microbiome composition and function, bowel movement frequency, frailty, and biomarkers of pro-inflammatory diseases and chronic kidney disease (CKD). Uremic toxins and biomarkers of inflammation were elevated in otherwise healthy individuals experiencing constipation or diarrhea, respectively. Microbially-derived uremic toxins enriched in the blood of constipated individuals were also negatively associated with kidney function. Enriched inflammatory proteins and genera, as well as immunological metabolites found in the Arivale cohort blood plasma imply a stressful cellular environment in the frail. Frailty negatively associates with estimated SCFA (propionate) production independent of age. These findings reveal some of the biological networks and mechanisms underlying the previously observed relationships between gut microbiome composition, healthy aging, and chronic disease development, warranting further investigation into the role of the gut microbiome as a mediator of health and disease throughout the lifespan.

SPECIFIC AIMS

Specific Aim #1:

Quantify and interpret the multi-omic signatures (using Count Regression for Correlated Observations with the Beta-binomial [CORNCOB] on 16s rRNA gene amplicon sequence variants counts, Linear Models for Microarray Data [LIMMA] on untargeted blood plasma metabolomics, LIMMA on blood plasma proteomics, LIMMA on clinical blood laboratory chemistries, and POLR on diet/lifestyle/health history survey data) of an otherwise healthy cohort based upon categorical bowel movement frequency (BMF) variation, accounting for gender, age, BMI, eGFR, the first three principal components of genetic ancestry, batch effects from vendors, C-reactive protein (CRP), glycohemoglobin A1c (A1C), and LDL cholesterol (LDL) levels. Analyze any relationships between BMF and kidney function (eGFR).

- **Rationale:** This aim elucidates any strong signal that appears when regressing each ‘omic against categorical BMF and identifies significant metabolites to test in a mediation analysis with eGFR.
- **Expected Outcomes:** Respective lists of summary statistical features and names of metabolites, chemistries, proteins, and genera that contribute to driving BMF variation; a contribution to the understanding of biochemical associations with BMF in the body.

Specific Aim #2:

Quantify and interpret the multi-omic signatures (using CORNCOB on amplicon sequence variants counts, LIMMA on untargeted blood plasma metabolomics, LIMMA on blood plasma proteomics, and LIMMA on clinical blood laboratory chemistries) of a 500+ cohort based on frailty variation. Analyze any relationships between frailty and the blood and gut microbiota accounting for age, gender, BMI, and the first four principal components of genetic ancestry.

- **Rationale:** This aim elucidates any strong signal that appears when regressing each ‘omic against continuous frailty (70-pt metric) and identifies significant metabolites, proteins, and genera that are associated with frailty.
- **Expected Outcomes:** Respective lists of summary statistical features and names of metabolites, chemistries, proteins, and genera that contribute to driving frailty variation; a contribution to the understanding of biochemical associations with frailty and aging in the body.

CENTRAL HYPOTHESES

1. **BMF aberrations, an early prodromal symptom of neurodegenerative diseases (ND) and a highly prevalent comorbidity in chronic diseases like CKD, will exhibit biochemical signatures even in the relatively healthy before later disease onset.**
2. **Analyzing Frailty Index Score (FIS), a 70-pt measure of vulnerability in succumbing to debilitating disease and mortality, will unveil biochemical patterns indicative of inflammation, oxidative stress, and other factors that deviate between healthy aging and frail aging.**

KEY RESEARCH APPROACH

Generalized linear regression, ordinal logistic regression, and beta-binomial regression will be employed to determine statistical results and data visualizations using established bioinformatics pipelines. These include: Jupyter Notebooks/R/Python, custom software packages for each, LIMMA, CORNCOB, POLR, generalized linear regression modeling (GLMs) to generate outputs.

Data Collection and Analysis

The data have been collected in a repository during the Arivale Scientific Wellness study (which ended in 2019). They are available as Jupyter/Python data snapshots. 16s and ‘omic data must be prepared for regression analysis and data visualization.

Expected Outcomes

This research approach, given the sample size of 500+ participants and the resolution of the BMF data, should answer the research questions regarding characterizing BMF and frailty multi-omic signatures. This will allow future researchers to justify designing more complex studies to obtain longitudinal and causal studies on BMF and frailty as they relate to the human gut microbiota.

ABBREVIATIONS

1. Bowel Movement Frequency (BMF) - number of bowel movements per week.
2. Estimated glomerular filtration rate (eGFR) - measure of kidney function.
3. Inflammatory Bowel Disease (IBD) - chronic disease of inflammation in all or part of the gastrointestinal tract.
4. Gastroesophageal reflux disease (GERD) - a long-term condition that occurs when stomach acid returns into the esophagus. Left untreated, it can cause mutagenesis from acid damage to the esophagus.
5. Irritable Bowel Syndrome (IBS) - an intestinal disorder caused by exclusive or mixed symptoms of diarrhea, constipation, bloating, gas, or other symptoms.
6. Frailty Index Score (FIS) - defined differently per study via scores. Measures the vulnerability of individuals succumbing to debilitating illness and death while accounting for age.

7. Short-chain fatty acids (SCFA) - anti-inflammatory/signaling molecules (e.g. butyrate, propionate, indole acetate) produced by gut microbiota to immunomodulate and regulate gut motility and inflammation.
8. Linear Models for Microarray Data (LIMMA) - data analysis, linear models and differential expression for microarray and RNA-seq data. Used in linear regression modeling.
9. Low-density lipoprotein cholesterol (LDL) - a clinical laboratory measurement of unhealthy cholesterol levels in the blood plasma.
10. C-reactive protein (CRP) - a clinical laboratory measurement of a blood plasma protein indicative of inflammation in the body.
11. Complete blood count (CBC) - a clinical laboratory measurement of number of cells in blood.
12. Complete metabolic panel (CMP) - a 14-point clinical laboratory panel of measurements of important health biomarkers to diagnose conditions such as infection and anemia.
13. Glycohemoglobin A1c (A1C) - a clinical laboratory measurement of the percentage of red blood cells coated in sugar; a measure of blood sugar levels in the previous months.
14. Count Regression for Correlated Observations with the Beta-binomial (CORNCOB) - R package for modeling count-based gut microbiome data using a beta-binomial mixture distribution.
15. Principal components of genetic ancestry (PC, e.g. PC1-PC4) - the principal components of genetic ancestry of the Arivale cohort.
16. Cardiovascular disease (CVD) - pathology of the heart and vasculature.
17. Autism Spectrum Disorder (ASD) - neurological and developmental disorder that affects how people interact, learn and behave.
18. Chronic kidney disease (CKD) - long-term kidney dysfunction or impairment that is progressive.

19. Colonic transit time (CTT) - quantitative measurement of time elapsed for bolus to pass through colon.
20. Proportional odds logistic regression (POLR) - method of regression used on data with ordinal variables.
21. Neurodegenerative disease (ND) - progressive neurological disease that causes physical, emotional, and/or mental/cognitive impairment over time.
22. Parkinson’s disease (PD) - a classic ND characterized by dopaminergic dysfunction and a build-up of α-synuclein protein, now believed to interact with the vagus nerve.
23. Tumor necrosis factor (TNF) - protein involved in immunomodulation.
24. Immunoglobulin A (IgA) - one of the five primary immunoglobulins involved in mucosal homeostasis and immunity.
25. 3-indoxyl sulfate (3-IS) - a uremic toxin; biomarker for kidney decline.
26. P-cresyl sulfate (PCS) - a uremic toxin; biomarker for kidney decline.
27. P-cresyl glucuronide (PCG) - a uremic toxin; biomarker for kidney decline.
28. Phenylacetylglutamine (PAG) - uremic toxin; induces inflammatory stress.
29. Lipopolysaccharides (LPS) - large, amphipathic glycoconjugates; they stimulate Toll-like receptor 4 (TLR4) immune response and are a major component of the outer membrane of gram-negative bacteria.
30. Toll-like receptor 4 (TLR4) - an immune cell transmembrane protein responsible for modulating innate immune response.
31. Wingless-related integration site (WNT) signaling pathway - plays crucial roles in various cellular processes, including inflammation, cancer, and arteriosclerosis.
32. dickkopf WNT signaling pathway inhibitory 1 (DKK1) - a potent inhibitor of the Wingless-related integration site (WNT) signaling pathway.
33. insulin-like growth factor binding protein 2, 36kDa (ILGFBP2) - regulates IGF activity, immune response, and insulin sensitivity.

CHAPTER 1: BOWEL MOVEMENT FREQUENCY, THE GUT MICROBIOME, AND HOST MOLECULAR PHENOTYPES

The following chapter is adapted from Johnson-Martínez et al., 2024, *Cell Reports Medicine* [6] .

1.1 SUMMARY

Bowel movement frequency (BMF) directly impacts the gut microbiota and is linked to diseases like CKD or dementia. In particular, prior work has shown that constipation is associated with an ecosystem-wide switch from fiber fermentation and short chain fatty acid production to more detrimental protein fermentation and toxin production. Here, we analyze multi-omic data from generally-healthy adults to see how BMF affects their molecular phenotypes, in a pre-disease context. Results show differential abundances of gut microbial genera, blood metabolites, and variation in lifestyle factors across BMF categories. These differences relate to inflammation, heart health, liver function, and kidney function. Causal mediation analysis indicates that the association between lower BMF and reduced kidney function is partially mediated by the microbially-derived toxin 3-indoxyl sulfate (3-IS). This result, in a generally-healthy context, suggests that toxins associated with abnormal BMF precede organ damage and may be drivers of chronic, aging-related diseases.

1.2 INTRODUCTION

The gut microbiome influences human health in a number of ways, from mediating early life immune system development [7,8], to determining personalized responses to nutritional interventions [9,10] and influencing the central nervous system [11,12]. Bowel movement frequency (BMF), defined as the frequency of defecation, is a major determinant of the composition of the human gut microbiota, and can be affected by diet, hydration, physical activity, host mucus production, microbe- and host-derived small molecules (e.g., short chain fatty acids, bile acids, or neurotransmitters), and peristaltic smooth muscle contractions in the gastrointestinal tract

[13-14,15]. Aberrant BMFs have been implicated in the development of various chronic diseases [16–18].

Abnormally high BMF (e.g., diarrhea, defined as more than three watery stools per day), has been associated with lower gut microbiome alpha-diversity, inflammation, increased susceptibility to enteric pathogens, and poorer overall health [19–22]. Abnormally low BMF (e.g. constipation, defined as fewer than three hard, dry stools per week), has been associated with higher gut microbiome alpha-diversity, with an enrichment in microbially-derived urinary metabolites known to be hepatotoxic or nephrotoxic, and with an increased risk for several chronic medical conditions, including neurodegenerative disorders and CKD [16,23–25]. Indeed, constipation is a known risk factor for CKD severity and end-stage renal disease (ESRD) progression [26,27]. In one study, up to 71% of dialysis patients suffered from constipation [28], while the prevalence of constipation in the general population was 14.5% in adults under 60 years old and 33.5% in those over 60 [29]. A nationwide study of veterans found an incrementally higher risk for renal disease progression in those who reported increasingly severe constipation [30]. However, while it is clear that morbidity and mortality risk worsen with constipation in those with active CKD, potential connections between BMF and the development and early-stage kidney disease are not yet established.

Both constipation and CKD associate with declines in gut microbiota-mediated short-chain fatty acid (SCFA) production and a rise in the production of amino acid putrefaction byproducts, including several toxic microbe-host co-metabolites, such as 3-indoxyl sulfate (3-IS), p-cresyl sulfate (PCS) and phenylacetylglutamine (PAG), which all have been implicated in CKD progression [31–33]. This is consistent with an established microbiota-wide transition from saccharolytic to proteolytic fermentation in constipated individuals due to the exhaustion of dietary fiber in stool [16,34]. Thus, while the potential relationship between BMF and organ function in healthy populations is not fully understood, the gut metabolic phenotype associated with lower BMF suggests a connection.

In this study, we focus on categories of self-reported BMF in a large population of generally-healthy individuals with a wide range of molecular phenotypic data in order to quantify the phenotypic impact of BMF on blood plasma metabolites, blood proteins, clinical chemistries, and gut microbiome composition in a pre-disease context. By exploring the molecular phenotypic consequences of BMF variation in a generally-healthy cohort, along with BMF-associated demographic, dietary, lifestyle, and psychological factors, we aimed to identify early-stage biomarkers and potential therapeutic targets for the monitoring and prevention of certain chronic, non-communicable diseases, like CKD.

1.3 RESULTS

A cohort of generally-healthy individuals

3,955 Arivale Scientific Wellness program participants with BMF data were initially considered in this analysis. Arivale, Inc. (USA), was a consumer scientific wellness company that operated from 2015 until 2019. Briefly, participants consented to having their health, diet, and lifestyle surveyed through an extensive questionnaire, along with blood and stool sampling for multi-omic and blood plasma chemistries data generation (**Fig. 1**). Any respondents that indicated "true" or affirmatively to any of the following questionnaire features were excluded from the analysis (i.e., they were not considered "generally-healthy"): taking blood pressure, cholesterol, or laxative medication or having self- or family-history of bladder or kidney disease (i.e. kidney cancer, bladder infections, polycystic kidney disease or PKD, kidney stones, kidney failure or kidney disease), inflammatory bowel disease (IBD; both Crohn's Disease and Ulcerative Colitis), irritable bowel syndrome (IBS), celiac disease, diverticulosis, gastroesophageal reflux disease (GERD), or peptic ulcers (i.e., these individuals were not considered 'generally-healthy'–see Supplement, **Table S1**). There were 1,425 participants who met these exclusion criteria and had necessary covariate data. Across all Arivale participants that had available demographic and survey information, 82.8% of those individuals identified as

"White" (N = 2,562), 8.5% identified as "Asian" (N = 262), 3.2% identified as "Black or African-American" (N = 98), 0.2% identified as "American Indian or Alaska Native" (N = 9), 0.65% identified as "Native Hawaiian or other Pacific Islander" (N = 20), and 4.7% identified as "Other" (N = 144). 93.6% of these individuals identified as "Non-Hispanic" (N = 2,897) and 6.4% identified as "Hispanic" (N = 198, 55.6% of which self-identify as "White"). Respondents were in the United States, predominantly from the Pacific West, and their ages ranged from 19 to 89 years old. 65.1% were female with a mean ± s.d. body mass index of 27.15 ± 5.89 (**Table S1**). 1,062 of these individuals had gut microbiome data, 486 had blood metabolomics data, 823 had proteomics data, 1,425 had clinical chemistries data, and 1,420 had survey data (derived from questionnaires). Self-reported BMF values (responses to typical number of bowel movements per week) were grouped into four categories (**Fig. 1**), which we labeled as: "constipation" (≤ 2 bowel movements per week), "low-normal" (3-6 bowel movements per week), "high-normal" (1-3 bowel movements per day), and "diarrhea" (4 or more bowel movements per day). We first looked at potential associations between BMF and relevant covariates: gender, age, BMI, estimated glomerular filtration rate (eGFR), low-density lipoprotein blood plasma levels (LDL), C-reactive protein blood plasma levels (CRP), hemoglobin A1c blood plasma levels (A1C), and the first three principal components of genetic ancestry (PC1, PC2, and PC3) (N = 1,425; **Fig. 2**). When BMF was coded as an ordinal dependent variable and regressed using ordered proportional odds logistic regression [35] (POLR), only BMI (POLR, FDR-corrected p = 1.82E-3), age (POLR, FDR-corrected, p = 2.07E-3), gender (POLR, FDR-corrected p = 3.68E-16), and the first three principal components of genetic ancestry (PC1, PC2, and PC3; POLR, FDR-corrected p < 0.0001) showed significant, independent associations with BMF (**Fig 2**), with females, older individuals, and individuals with lower BMIs tending to report lower BMFs (**Fig. 2**). All covariates listed above were included in downstream regressions, regardless of whether or not they showed an independent association with BMF. The high-normal BMF group was chosen as the reference for all downstream regressions where BMF was encoded as a categorical variable.

eGFR was also regressed against BMF and the other covariates to determine which were significant associated with eGFR, and the covariates with significant p-values included gender, age, BMI, LDL, A1C, PC1, PC2, and PC3 (GLM, $p < 0.05$).

Gut microbiome structure and composition across BMF categories

We looked at a subcohort of individuals that met our health exclusion criteria with 16S amplicon sequencing data from stool (N = 1,062). Amplicon sequence variant (ASV) richness (GLM, p = 2.85E-3, linear β_{BMF} = -65.9E-3) and Shannon diversity (GLM, p = 1.07E-3, linear β_{BMF} = -3.25E-1) were negatively associated with BMF, independent of the covariates listed above, and with BMF encoded as an ordinal variable with a linear coefficient (**Fig. 3**). Pielou's evenness, on the other hand, was positively associated with BMF (GLM, p = 8.5E-3, linear β_{BMF} = 2.6E-3), independent of covariates (**Fig. 3**).

Differential abundance analysis of commensal gut bacterial genera across BMF categories was conducted using beta-binomial regression [36] (CORNCOB) with BMF encoded as a categorical variable with the high-normal group as the reference category. Of the 135 genera that passed our prevalence filter (i.e., detection across ≥ 30% of individuals), 59 were significantly associated with BMF (49 of which had genus-level taxonomic annotations; see **Table S1** for detailed list of β-coefficients and p-values), independent of covariates and following an FDR correction for multiple tests on the likelihood ratio test (LRT) p-values (FDR-corrected $p < 0.05$). We z-score normalized the centered log-ratio (CLR) abundances of the 49 annotated genera across all samples and then plotted the average z-score within each BMF bin for each taxon as a heatmap (**Fig. 4**). We also provide supplemental boxplots, showing CLR abundances across BMF categories, of the top 10 most abundant taxa and 10 taxa with the smallest p-values from the 49 mentioned above (**Fig. S1-S2**). In order of descending abundance, the following taxa were significantly enriched in constipated individuals, compared to the high-normal BMF category (Wald Test, FDR-corrected β_{BMF} $p < 0.05$): *Ruminiclostridium_9,* *Ruminococcacaeae_UCG-005, Ruminococcaceae_NK41214_group, Family_XIII_AD3011_group, Romboutsia, Ruminocaccaeae_UCG-004, UBA1819, Negativibacillus, DTU089, GCA-900066225, Candidatus_Soleaferrea, Anaerotruncus, Defluviitaleaeceae_UCG-011, Eisenbergiella, Pygmalobacter, Peptococcus, Hydrogenoanaerobacterium, Anaerofustis*, and *DNF00809*. *Lachnospiraceae_ND3007_group* and *Lachnospiraceae_UCG-004* were significantly depleted in constipated individuals. Several more were associated with enrichment or depletion in the low-normal BMF category, compared to the reference category (**Fig. 4**; see Supplement). There was no significant difference between the high-normal and diarrhea categories for any of the genera, which could be due to low sample size in the diarrhea category (i.e., we were likely underpowered to detect those associations).

Variation in blood metabolites across BMF categories

Blood metabolite vs. BMF regression analyses were run using a generalized linear modeling (GLM) framework in LIMMA, with BMF as a categorical independent variable, along with the same set of covariates mentioned above. Of the metabolites that passed our abundance and prevalence filters (N = 956, see **Method Details**), 9 separate metabolites were significantly associated with BMF (all 9 showed differential abundance between low-normal and high-normal categories, which is the comparison we were most powered for), independent of covariates and following an FDR correction for multiple tests (GLM, FDR-corrected $p < 0.05$, **Fig. 5**, **S3**). The annotated metabolites tended to show a decreasing trend with increasing BMF, while the unannotated metabolites and 3-IS showed more varied relationships (e.g. monotonic and non-monotonic) with BMF (**Fig. 5, S3**). PCS, PAG, PCG, and 3-IS were significantly enriched in the low-normal BMF category, compared to the reference category (**Fig. 5, S3**). 75 different metabolites were significantly associated with eGFR, independent of covariates and following the same FDR correction for multiple tests (linear regression, FDR-corrected $p < 0.05$, **Fig. 5,**

S3). Only one of these eGFR-associated metabolites overlapped with any of the BMF-associated metabolites: 3-IS.

Blood plasma chemistries across BMF categories

Of the 55 blood plasma chemistries filtered for prevalence (see **Method Details**), 21 were significantly associated with diarrhea (e.g., omega-6 fatty acid, homocysteine, total protein, and bilirubin) and one (omega-6/omega-3 ratio in the blood) was associated with the low-normal BMF category, relative to the reference category, after adjusting for all covariates and for multiple testing (**Fig. 6, S4**; N = 1,425, GLM, FDR-corrected $p < 0.05$).

Blood proteomics across BMF categories

None of the 274 blood proteins that passed our prevalence filter (see **Method Details**) showed significant associations with BMF after adjusting for all covariates and for multiple testing (N = 823, GLM, FDR-corrected $p < 0.05$).

Self-reported diet, lifestyle, anxiety and depression histories associated with BMF categories and demographic covariates

99 survey questions (see **Supplement**; questions with sparse data were filtered out) on health, diet, and lifestyle were examined from 1,420 generally-healthy individuals from the Arivale cohort in order to identify covariate-independent associations with BMF. Tests were run using the “polr” package in R (ordinal regression) [35], including the same set of covariates from the prior analyses, and with BMF coded as a categorical variable with high-normal BMF as the reference group (**Fig. 7A**). Response categories for each question ascended ordinally in value or intensity (i.e., low to high), so that a positive association represented an increase in that variable. Across the 99 questions, the top results with significant odds ratios related to BMF categories were displayed relative to high-normal BMF (**Fig. 7A**), colored by the variable

category (“Diet/Lifestyle” or “Health/Digestion”). BMI, age, gender, and other covariates were also associated with many of these questionnaire-derived features, independent of BMF (**Fig. 7A**). In particular, females tended to eat more vegetables and fruit in a week and had a higher diarrhea frequency. Males, on the other hand, showed higher weekly snack intake and easier bowel movements (**Fig. 7A**). Unsurprisingly, constipation (lowest BMF range) was negatively associated with reported ease of bowel movement and diarrhea was positively associated with self-reported diarrhea frequency (i.e., these were separate questions on the questionnaire) (**Fig. 7A**). Those with higher weekly snack intake were more likely to be in the low-normal BMF category, and those with higher weekly vegetables intake, weekly fruit intake, greater ease of bowel movements, and those with higher self-reported diarrhea frequency were more likely to be in the high-normal BMF category (**Fig. 7A**). Higher diarrhea frequency was significantly associated with having a higher BMI and with being younger relative to the rest of the cohort, while being older made one more likely to report having greater ease of bowel movement (**Fig. 7A**). Finally, those with low LDL values (better cholesterol health) were more likely to report higher fruit intake and those with low CRP (low inflammation) values were more likely to report higher vegetables intake (**Fig. 7A**). These findings showcase a variety of common-sense dietary and lifestyle factors that could be leveraged to manage BMF, cardiometabolic, and immune health.

A subset of participants self-reported their history of depression and anxiety, including: “self-current”, “self-past”, and “family” history of depression and anxiety (N = 2,096, see Supplement; 11 questions related to anxiety and 23 related to depression). After logistic regression, 3 “true or false”-response questions related to history of depression in self and family history appeared marginally significant (logistic regression, FDR-corrected $p < 0.1$), with a self-reported “true” response to a “family history of depression” showing a marginal association with constipation (logistic regression, FDR-corrected < 0.1), a self-reported “true” response to a “sibling history of depression” showing a significant association with diarrhea (logistic regression, FDR-corrected < 0.05), and a self-reported "true" response to "recent ailments; self-history of depression" showing a marginal association with low-normal BMF (logistic regression, FDR-corrected < 0.1). Similarly, the same approach yielded a single marginal association between a "true" response to "self past history of anxiety disorder" and low-normal BMF (logistic regression, FDR-corrected < 0.1). Each of these associations were relative to the high-normal BMF reference category.

BMF-associated blood metabolites associated with kidney function in a generally-healthy cohort

Using the nine BMF-associated metabolites (ordered in ascending p-value: PCS, X - 23997, PAG, X - 11850, PCG, X - 12216, 3-IS, X - 11843, and X - 21310), an analysis was performed on all of the generally-healthy Arivale participants with paired BMF, eGFR, and blood metabolomic data (N = 572). Using OLS, eGFR was regressed against BMF (encoded as a numerical variable between 1, 2, 3, or 4, with 1 being constipation, 2 being low-normal, 3 being high-normal, and 4 being diarrhea) and the nine BMF-related metabolites, which yielded a significant overall model (**Fig. S6B**; OLS, $R^2 = 0.082$, $p = 2.42E-7$). Two of the BMF-associated metabolites showed significant beta-coefficients in the model: X - 12216 and 3-IS (**Fig. S6B**; OLS, $\beta_{X\ -\ 12216} = -1.98$, $p = 1.20E-2$ and $\beta_{3\text{-IS}} = -9.69$, $p = 1.96E-5$, respectively). These negative coefficients indicated that higher baseline levels of these blood metabolites were associated with lower kidney function.

Finally, given that microbially-derived 3-IS was independently associated with both eGFR and BMF, we hypothesized that 3-IS may be mediating, in part, the impact of BMF on eGFR. To test this hypothesis, we ran a causal mediation analysis (using the mediation library in R [37]; see **Methods**) on the generally-healthy Arivale individuals with BMF, eGFR, and the blood metabolomics data (N = 572; **Fig. 7B; S7**). BMF categories were merged into a "Low" (low-normal BMF and constipation) and a "High" categories (high-normal BMF and diarrhea participants) in order to consolidate the BMF categories with very small Ns (i.e., constipation and diarrhea). The total effect of the overall model did not quite pass our significance threshold of alpha < 0.05 (total effect, p = 0.064), but we saw a significant average direct effect of BMF on eGFR (ADE = -4.458, p = 0.012) and a highly significant average causal mediation effect of BMF via 3-IS on eGFR (ACME = 1.343, p < 2E-16; **Fig. 7B**).

1.4 DISCUSSION

In this study, we delve into the multi-omic fingerprint of cross-sectional BMF variation in a large, generally-healthy population (**Fig. 1**). We find that aberrant BMFs were associated with variation in the ecological composition of the gut microbiota, plasma metabolite levels, clinical chemistries, diet, lifestyle, and psychological factors (**Figs. 4-7A**). Overall, we observe an enrichment of microbially-derived uremic toxins in blood resulting from protein fermentation in the guts of individuals with lower BMFs. These toxins have been implicated in disease progression and mortality in CKD [27,38] and many of the same metabolites have been associated with other chronic diseases, like neurodegeneration [39,40].

Of the core set of covariates used in our regression analyses, only age, gender, BMI, and genetic ancestry PCs 1-3 were independently associated with BMF, with females, individuals with lower BMIs, and younger individuals showing lower average BMFs (**Fig. 2**). Consistent with these results, women are known to be at higher risk of constipation and kidney dysfunction [41,42]. In a prior study, individuals with lower BMIs were shown to produce less motilin (i.e., a hormone involved in gut motility) and were more likely to suffer from constipation [43]. Lower BMFs have also been linked to inflammation, oxidative stress, and cardiovascular disease risk [44,45]. The associations between BMF and the first three principal components of genetic ancestry indicate a relationship between host genetics and BMF variation, which is further supported by a prior GWAS study [46].

Independent of these covariates, several gut bacterial genera enriched in individuals with lower BMFs (CORNCOB, $p < 0.001$), such as *Christensenellaceae_R-7_group*, *Anaerotruncus*, *Blautia*, *Family_XIII_AD3011_group* (Anaerovoracaceae family), and *Methanobrevibacter*, were previously found to be enriched in Parkinson's disease (PD) patients, who often suffer from chronic constipation [47]*.* *Desulfovibrio*, which has been shown to be enriched in several disease states [48], was elevated at lower BMF (**Fig. 4**). Another set of genera were depleted in lower BMF categories, such as *Bacteroides*, *Lachnoclostridium*, *Lachnospiraceae_ND3007_group*, *Lachnospiraceae_UCG-004*, and *Veillonella*, which are all important contributors to SCFA production [49–52]. This reduction in SCFA producers is consistent with the switch away from saccharolytic fermentation towards proteolytic fermentation in the case of constipation [16]. Reduced SCFA production is known to weaken smooth muscle contractions that drive peristalsis [53–55], acting as a positive feedback on constipation.

Consistent with our microbiome results, we found gut microbiome-derived protein fermentation byproducts, like PCS, PAG, and 3-IS, were enriched in the blood of individuals with lower BMFs (**Fig. 5**) [56–58] . PCS has been associated with deteriorating kidney function and with damage to nephrons as well as cognitive decline and neuroinflammation [59,60]. 3-IS has been associated with vascular disease and mortality in CKD patients [61]. PAG has been associated with CKD progression and mortality [32,33,56,57] . Ultimately, we see an enrichment in microbially-derived uremic toxins in the blood of generally-healthy individuals with lower BMFs.

Most of the clinical chemistry-BMF associations showed relative enrichment in the higher-BMF category, and these features tended to reflect hepatic and nephrotic function. For example, high bilirubin can indicate liver disease from the overactive breakdown of red blood cells, but interestingly, higher bilirubin levels in serum coincide with a lower risk for CKD development and progression, which coincides with our observation that the lowest BMF categories had higher levels of uremic toxins but lower bilirubin levels [62]. Other metrics, like creatinine levels and linoleic acid levels, correlate positively with BMF and negatively with kidney function [63–65]. In fact, most of the laboratory values, such as the mean corpuscular hemoglobin concentration (MCHC), which measures the concentration of blood cells, can indicate kidney or liver disease [66]. It is interesting to note that biomarkers indicating kidney disease risk and progression were enriched at lower BMFs and biomarkers indicating liver disease risk and progression were enriched at higher BMFs in a generally-healthy population, showing how aberrant BMF in either direction may increase chronic disease risk.

In addition to demographic factors associated with BMF, the questionnaire results indicate dietary and lifestyle factors that are known to influence BMF, like fruit and vegetable intake (i.e., sources of dietary fiber and polyphenols) [42,44]. We observed a lower fruit and vegetable intake and an increased likelihood of snacking in the low-normal BMF category compared to the high-normal BMF category [29,42]. We also found that constipation and diarrhea were marginally (and in one case, significantly) associated with self-reported measures of depression and anxiety, which aligns with prior work showing higher prevalence of anxiety and depression (between 22-33%) on the Hospital Anxiety and Depression Scale (HADS) and the Mini International Neuropsychiatric Interview (MINI) in patients with chronic constipation [67].

Blood levels of 3-IS were independently associated with both BMF and eGFR, which led us to the hypothesis that 3-IS may mediate the potential influence of BMF on eGFR. Indeed, we observed a significant average direct effect of BMF on eGFR (ADE, $p = 0.012$) and a highly significant average causal mediation effect for 3-IS (ACME, $p < 2E-16$; **Fig. 7B**). Together, these results indicate that aberrant BMF-associated increases in 3-IS are associated with declining kidney function in a generally-healthy cohort, which is consistent with similar associations that have been observed between 3-IS and poorer outcomes in CKD patients [61].

Bowel movement abnormalities, such as constipation or diarrhea, have been linked to diseases ranging from enteric infections [22] to many chronic diseases like CKD, IBD, and neurodegenerative conditions like Alzheimer's and PD [39,68,69]. Indeed, even in the context of our generally-healthy cohort, we see the build up of microbially-derived uremic toxins in the blood of individuals with lower BMFs. Perhaps most concerning was our observation that aberrant BMF-associated microbial metabolite 3-IS was also associated with lower eGFR values. Our results underscore common-sense dietary and lifestyle changes, like increasing intake of fruits and vegetables, which may help to normalize BMF and perhaps reduce BMF-associated risk. We suggest that chronic constipation or diarrhea may be underappreciated drivers of organ damage and chronic disease, even in healthy populations, although additional work is required to rigorously quantify how BMF impacts disease risk.

Study limitations

There are some important limitations to consider when interpreting the results of this study. The generally-healthy cohort studied here was overwhelmingly "White", predominantly female, and from the West Coast of the U.S.A., which limits the generalizability of our results. In addition, the diet, lifestyle, and mood data were self-reported and subject to biases and errors, and are not indicative of clinical diagnoses, although BMF was binned into four coarse-grained categories in an attempt to mitigate self-reporting bias. In fact, BMF is not quite synonymous with transit time through the gut, which can be measured through means like the "blue dye method" [13] , although BMF still appears to be a useful and interesting measure of self-reported bowel habit differences in this study when binned in such coarse-grained categories. We had limited representation in the constipation and diarrhea categories, which reflects the "generally-healthy" nature of this cohort, but this also limited our statistical power for detecting associations in these groups. Detailed renal assessments beyond eGFR and clinical laboratory chemistries were also not available in this study. Additionally, the dietary variables that were associated with better BMF outcomes (i.e., increased dietary fiber intake, in the form of fruits and vegetables) are not devoid of clinical risk and may not be appropriate for everyone. For example, high-fiber diets can sometimes lead to bloating and inflammation in IBD patients [70]. Furthermore, CKD patients are often coached to limit their intake of fiber-rich foods because they can contain high levels of potassium and phosphorus [71]. However, low-fiber diets may act as a positive feedback on constipation and inflammation. This highlights the importance of intervening at the prodromal stage, before disease manifests, when a healthy, largely plant-based diet is well-tolerated. While we find some evidence for microbially-derived, BMF-associated uremic toxins in blood influencing kidney function in a generally-healthy cohort, more work is needed to establish a link between longer-term BMF management and chronic disease risk. In the mediation analysis, we did not see a strong total model effect, despite seeing significant direct and mediation effects. This kind of result is expected when the total and mediation effects are of a similar magnitude, when there are opposing effect directions between treatments and mediators, or when there are other more complicated effects (e.g., non-linear associations) [72]. Ultimately, future intervention trials should be done to assess the potential for managing BMF throughout the lifespan as a strategy to reduce chronic disease risk.

1.5 FIGURE TITLES AND LEGENDS

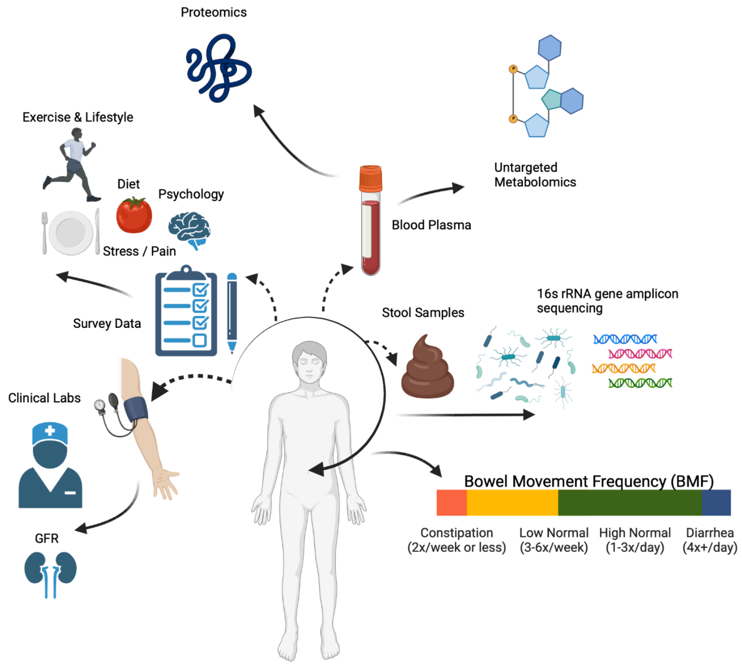


Figure 1. Data collection strategy. Arivale participants were sampled for blood plasma and stool, in addition to filling out extensive diet, health, and lifestyle questionnaires. Clinical chemistries, untargeted metabolomics, and proteomics data were generated from blood plasma samples. Gut microbiome 16S rRNA amplicon sequencing data were generated from stool samples collected using at-home kits. BMF data were extracted from the questionnaire data as self-reported frequencies per week or day.

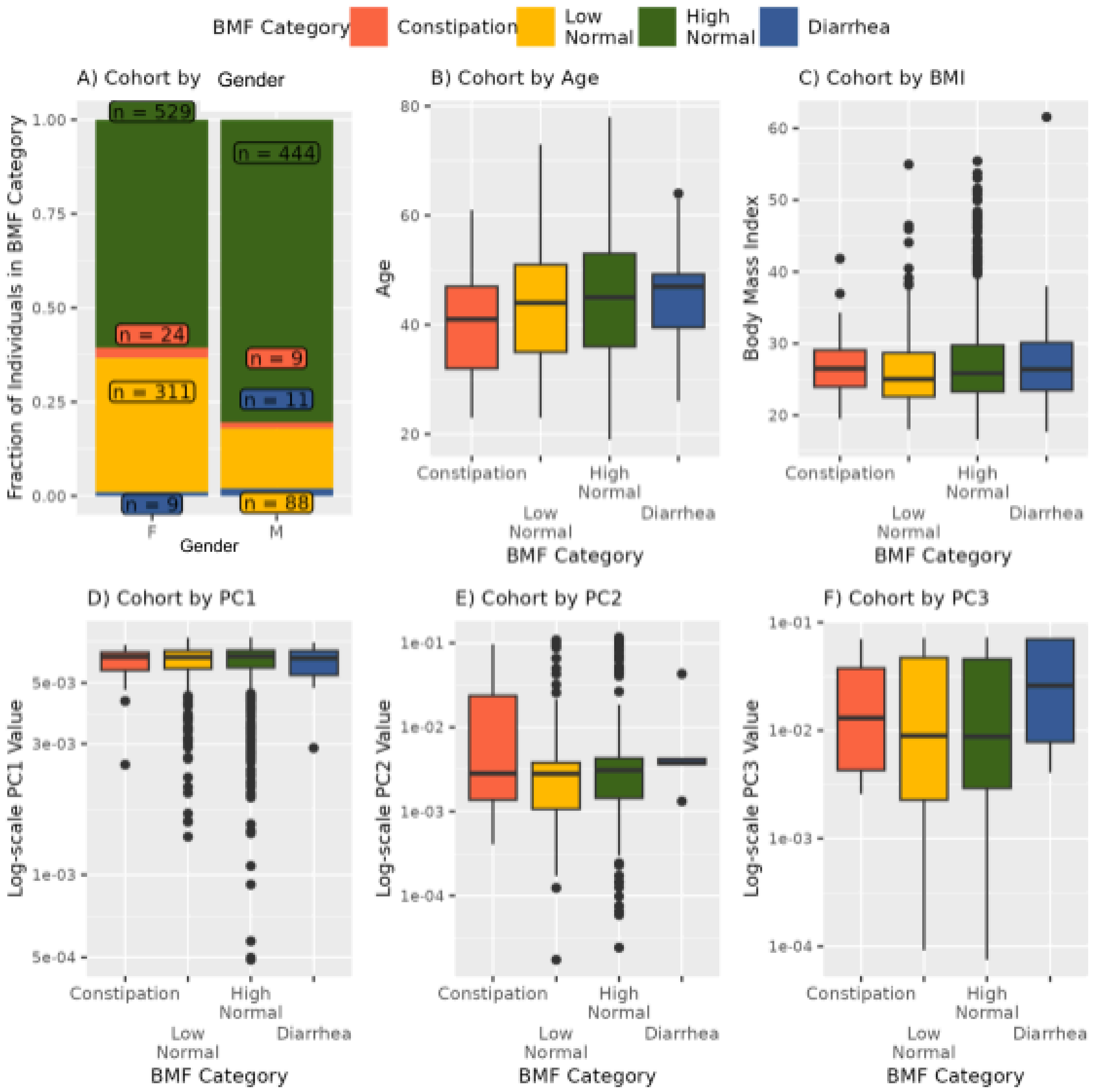


Figure 2. Plotting covariates that showed a significant association with BMF: gender, age, BMI, and the first three principal components of genetic ancestry (PC1-PC3) (A-F). POLR was used to regress BMF against the covariates (gender, age, BMI, eGFR, LDL, CRP, A1C, plus the first three principal components of genetic ancestry in the cohort, PC1, PC2, PC3). The result was that gender ($p = 3.68E-16$), BMI ($p = 1.82E-3$), age ($p = 2.075E-3$), and PCs1-3 ($p < 0.00001$, respectively) were significantly associated with BMF.

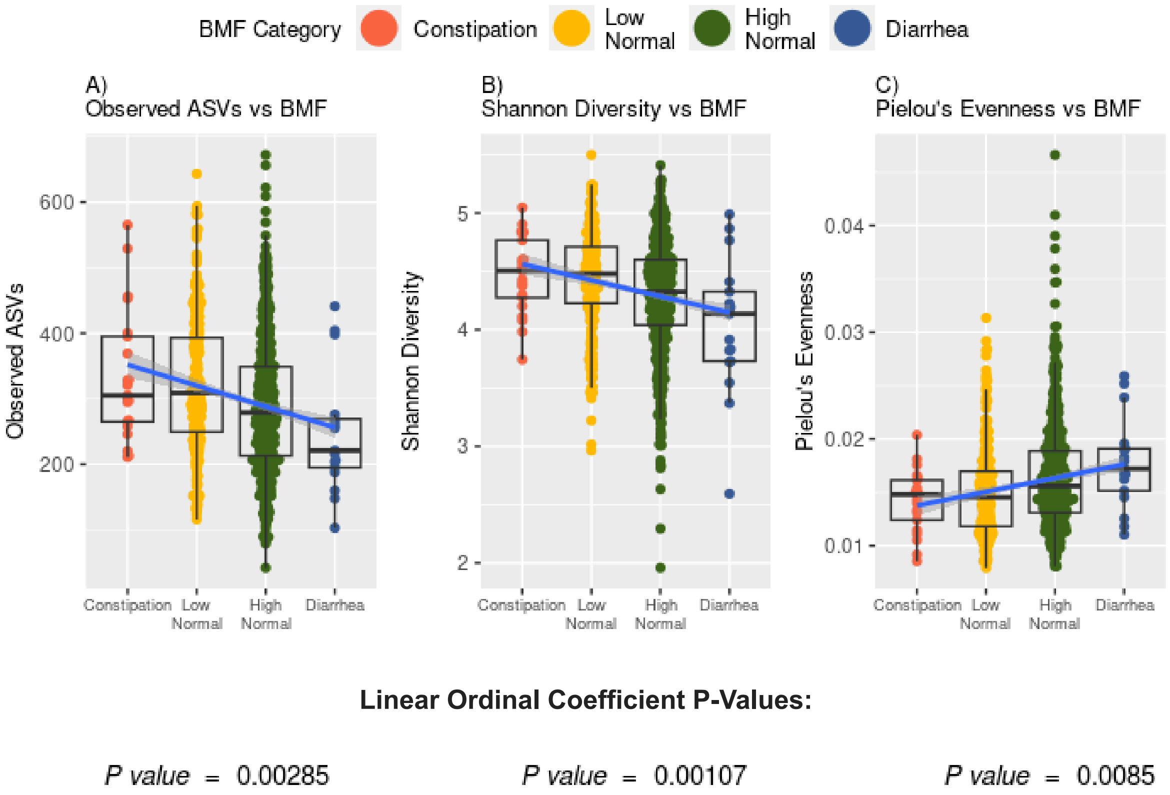


Figure 3. Associations between gut microbiome alpha-diversity and BMF. (A) Richness of amplicon sequence variants (ASVs) across BMF categories (ordinal BMF variable, Linear Regression, p = 2.85E-3). (B) Shannon Diversity across BMF categories (ordinal BMF variable, Linear Regression, p = 1.07E-3). (C) Pielou's Evenness across BMF categories (ordinal BMF variable, Linear Regression, p = 8.5E-2).

Figure 4. Heatmap of average z-scored CLR abundances within each BMF category for all annotated genera significantly associated with BMF. 46 significant taxa, in order of decreasing average relative abundance, with their z-scored, CLR-transformed abundances averaged within each BMF category plotted as a heatmap. Covariates included gender, age, BMI, eGFR, LDL, CRP, A1C, and PCs1-3. Asterisks denote the individual FDR-corrected significance threshold for the Wald Test p-value of the β_{BMF}-coefficient for each BMF category, relative to the high-normal reference category. Rows without asterisks showed a significant overall model (FDR p-value <0.05), despite a lack of significance for the individual coefficients. (***): $p < 0.0001$, (**): $0.0001 < p < 0.01$, (*): $0.01 < p < 0.05$.

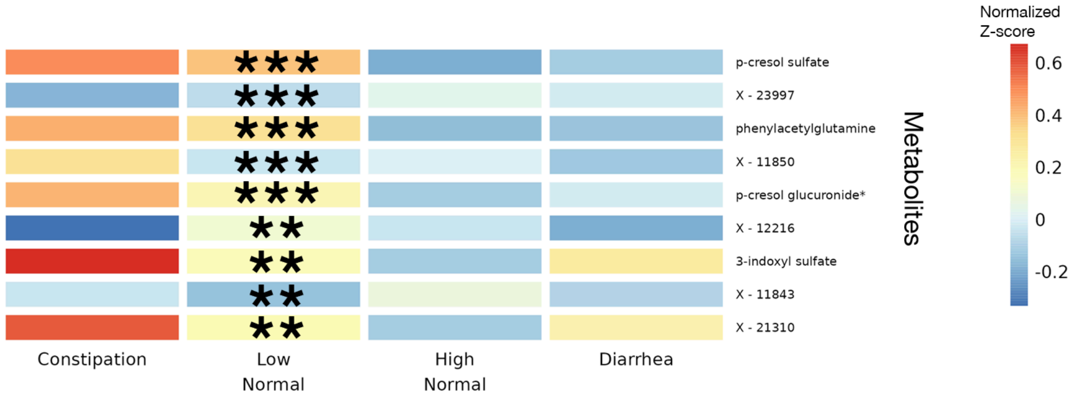


Figure 5. Heatmap of average z-scored blood plasma metabolites levels within each BMF category for all metabolites significantly associated with BMF. 9 significant blood plasma metabolites, with average z-scores within each BMF category plotted as a heatmap. Significant associations were identified using LIMMA, with FDR-corrected p-values of the ratio test between the main model and the null model. Here, the covariates included gender, age, BMI, eGFR, LDL, CRP, A1C, and PCs1-3. Asterisks denote metabolites with significant β_{BMF} coefficient(s) in the linear regression model after FDR correction. (***): $p < 0.0001$, (**): $0.0001 < p < 0.01$, (*): $0.01 < p < 0.05$.

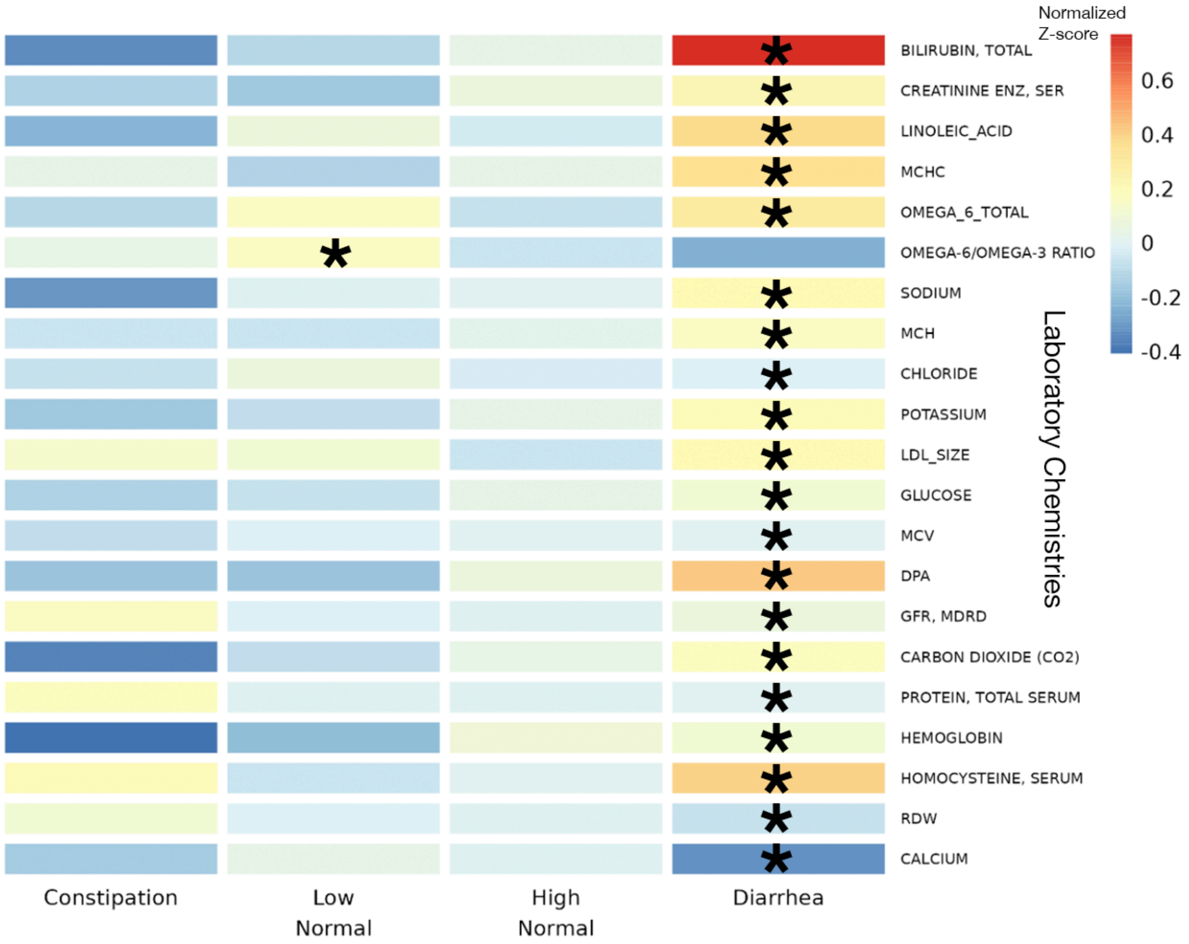


Figure 6. Heatmap of average z-scored clinical chemistries within each BMF category for all chemistries significantly associated with BMF. 22 BMF-associated chemistries, identified using LIMMA models with FDR-corrected p-values of the ratio test between the main model and the null model, with average z-scores within each BMF category plotted as a heatmap. Here, the covariates included gender, age, BMI, eGFR, LDL, CRP, A1C, and PCs1-3. Asterisks denote FDR-corrected p-value thresholds for metabolites with significant β_{BMF} coefficient(s) in the linear regression model. (***): $p < 0.0001$, (**): $0.0001 < p < 0.01$, (*): $0.01 < p < 0.05$.

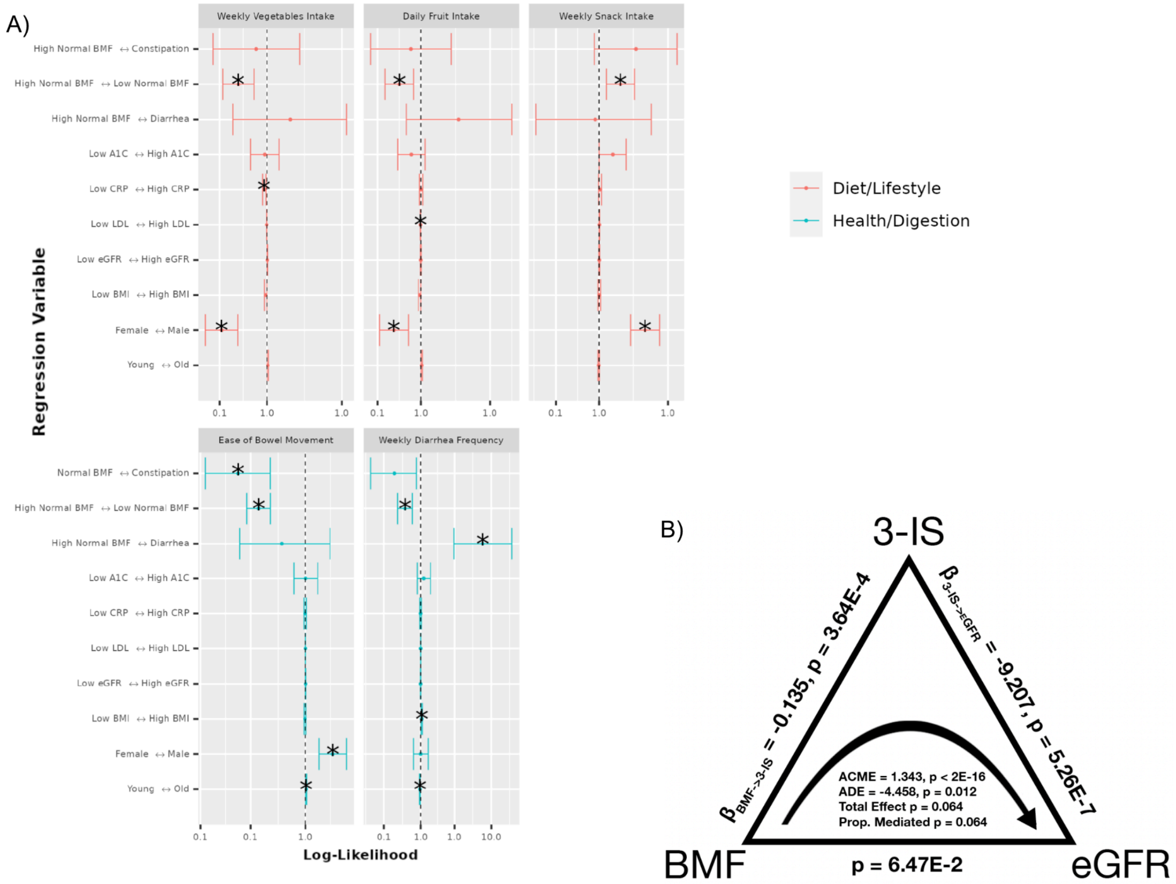


Figure 7. Ordinal regression odds ratio for health, diet, and lifestyle survey data vs BMF and covariates, and causal mediation analysis, with BMF as the treatment variable, 3-IS as the mediator variable, and eGFR as the response variable. A) Bowel Movement Frequency (BMF) categories are shown by question type (diet/lifestyle or health/digestion). "High-normal" BMF (7-21/week) is the reference. Vertical ticks show directional associations in likelihood between variables across the horizontal axis. The center line (x = 1.0) indicates equal likelihood of increased values on either side. Confidence intervals not crossing the line are significant associations (FDR $p < 0.05$). B) BMF affects eGFR directly (average direct effect, ADE) and indirectly (average causal mediated effect, ACME) through 3-IS (a metabolite). Both effects are significant (N = 572); the total effect was not significant (N = 572; ADE -4.458, p = 0.012; ACME 1.343 $p < 2E-16$). The total effect and the proportion mediated terms did not pass our significance threshold of $\alpha = 0.05$.

1.6 SUPPLEMENTAL FIGURES AND TABLES

Covariates:	Mean ± standard deviation, or % across Arivale:
Gender	65.1% Female
BMI	27.2 ± 5.89
Age	46.36 ± 12.96
eGFR	89.07 ± 20.20
CRP	2.40 ± 4.76
LDL	114.17 ± 33.77
A1C	5.49 ± 0.57
Highlighted exclusionary criteria:	
Percent with self-reported kidney disease:	3.00% (119 out of 3,955 participants with BMF data available withheld from cohort)
Percent IBS or IBD:	3.23% (128 out of 3,955 participants with BMF data available withheld from cohort)
Exclusionary features (988 out of 3,955 participants with BMF data, or 25% of the initial BMF cohort, answered affirmatively to any of these and were excluded from the analyses. The final N of remaining participants after merging with covariates was N = 1,425 for the final baseline cohort):	
Self - current history - bladder infection	
Self - current history - kidney disease	
Self - current history - kidney infection	
Self - current history - kidney stones	
Self - current history - bladder/kidney - other	
Self - current history - polycystic kidney disease (PKD)	
Self - current history - urinary incontinence	
Self - current history - kidney cancer	
Self - current history - celiac disease	
Self - current history - colonic Crohn's disease	
Self - current history - diverticulosis	
Self - current history - gastroesophageal reflux disease (GERD)	
Self - current history - ileal Crohn's disease	
Self - current history - irritable bowel syndrome (IBS)	
Self - current history - inflammatory bowel disease (IBD)	
Self - current history - ulcerative colitis	
Self - current history - peptic ulcer	
Self - laxatives usage	
Self - anticoagulation or cholesterol drugs usage	
Self - blood pressure drugs usage	

Table S1. The modeling covariates and exclusionary criteria, related to Figure 1. Out of the 3,955 total Arivale participants that had BMF data, 3.00% self-reported kidney disease (the kidney-related questions in the exclusionary features) and 3.23% self-reported IBS or IBD. An initial baseline cohort of 3,132 participants that had health history survey questionnaire data was available. The participants that answered affirmatively to the exclusionary features were removed from the analysis, resulting in 25% of the initial cohort with BMF data being filtered down to N = 1,561, and subsequently, a final baseline cohort of 1,425 individuals after merging for covariates.

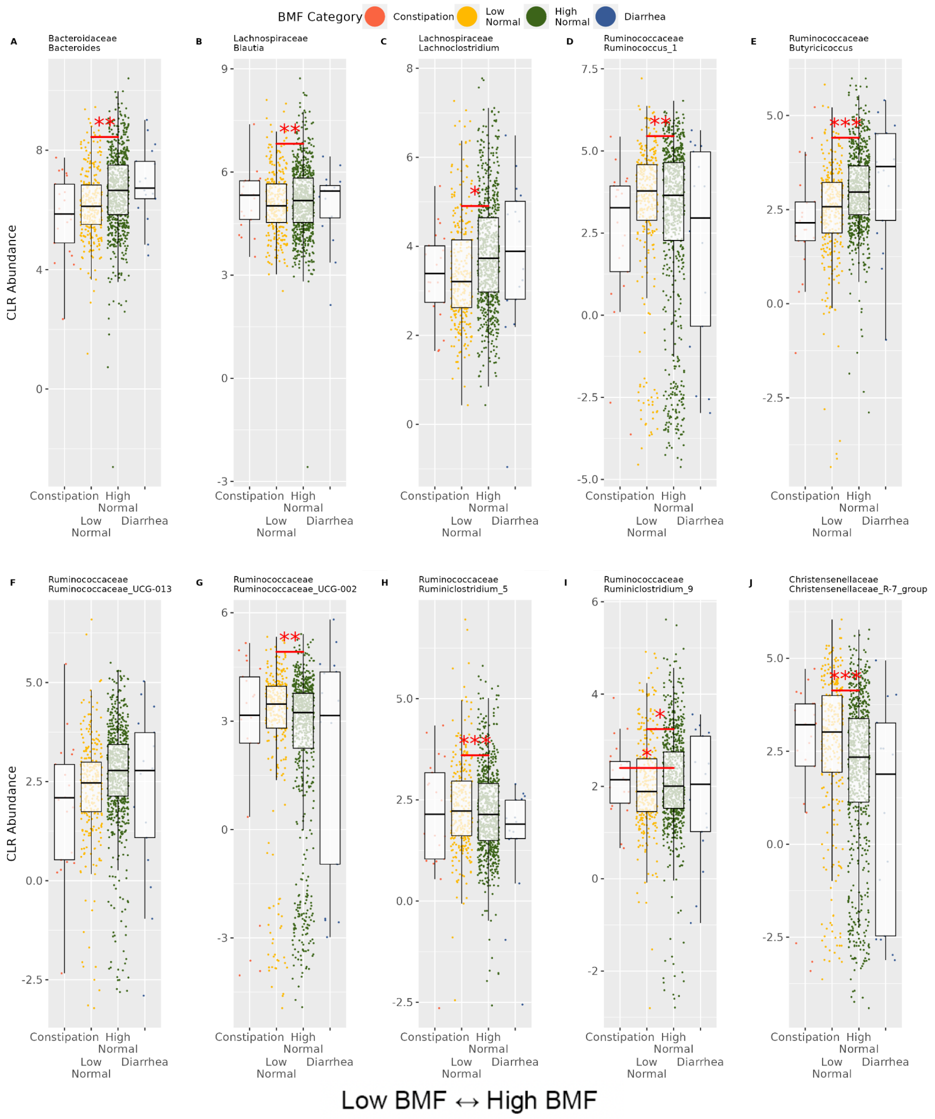


Figure S1. The top 10 most abundant genera significantly associated with BMF (A-J). Related to Figure 4. Significant genera from the CORNCOB analysis in order of decreasing CLR-transformed abundance. The line in each plot denotes significant differences from the reference category (“High Normal” BMF), and asterisks denote FDR-corrected significance threshold. (***): $p < 0.0001$, (**): $0.0001 < p < 0.01$, (*): $0.01 < p < 0.05$. The horizontal axes are annotated as four BMF categories: “Constipation” (BMF = 1-2✕ per week), “Low Normal” (BMF = 3-6✕ per week), “High Normal” (BMF = 1-3✕ per day) which is the reference category in regression, and “Diarrhea” (BMF = 4✕ or more per day).

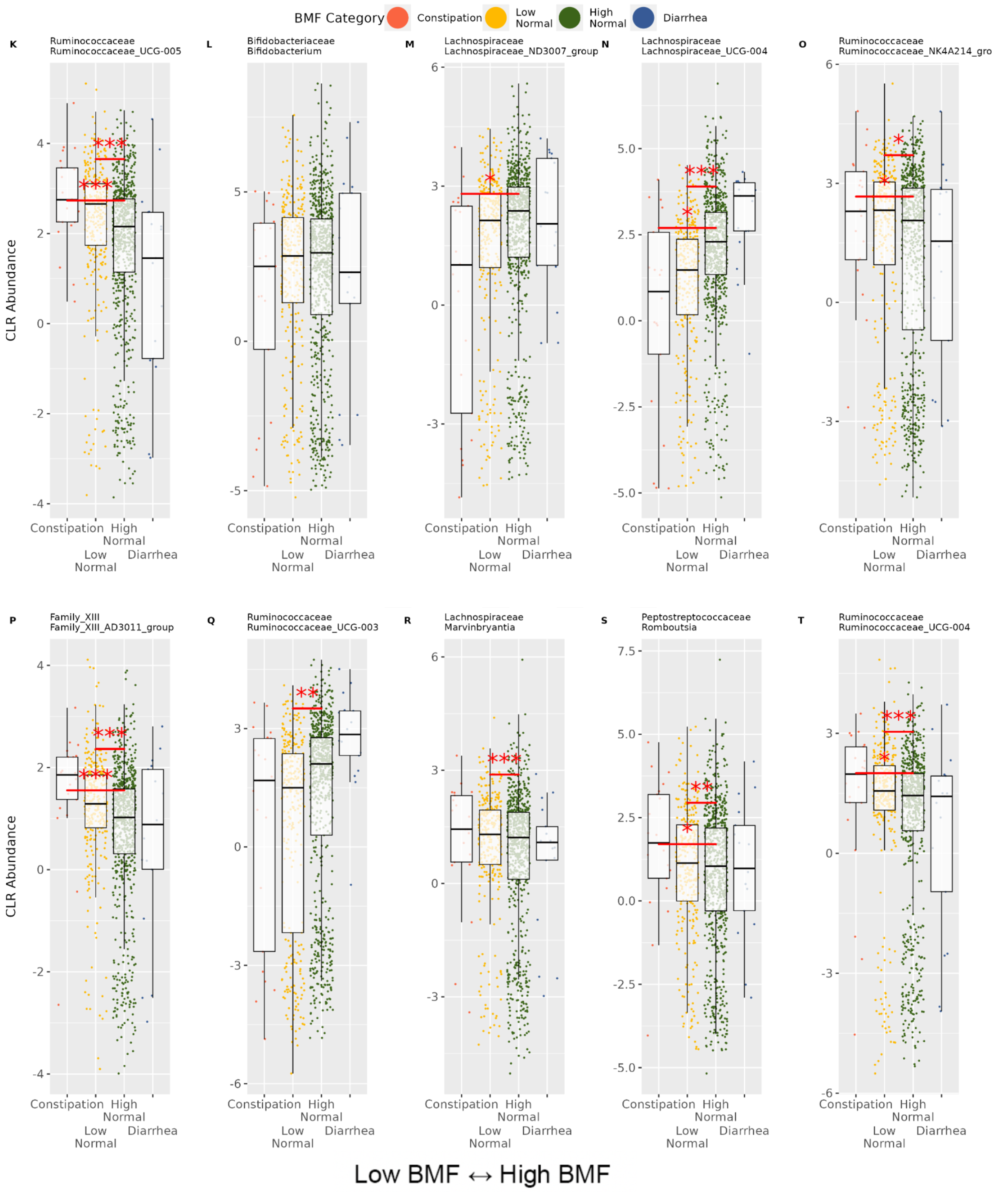


Figure S2. The top 11-20 most abundant genera associated with BMF (K-T), related to Figure 4. Significant genera from the CORNCOB analysis in order of decreasing CLR-transformed abundance. The line in each plot denotes significant differences from the reference category (“High Normal” BMF), and asterisks denote FDR-corrected significance threshold. (***): p < 0.0001, (**): 0.0001 < p < 0.01, (*): 0.01 < p < 0.05. The horizontal axes are annotated as four BMF categories: “Constipation” (BMF = 1-2✕ per week), “Low Normal” (BMF = 3-6✕ per week), “High Normal” (BMF = 1-3✕ per day) which is the reference category in regression, and “Diarrhea” (BMF = 4✕ or more per day).

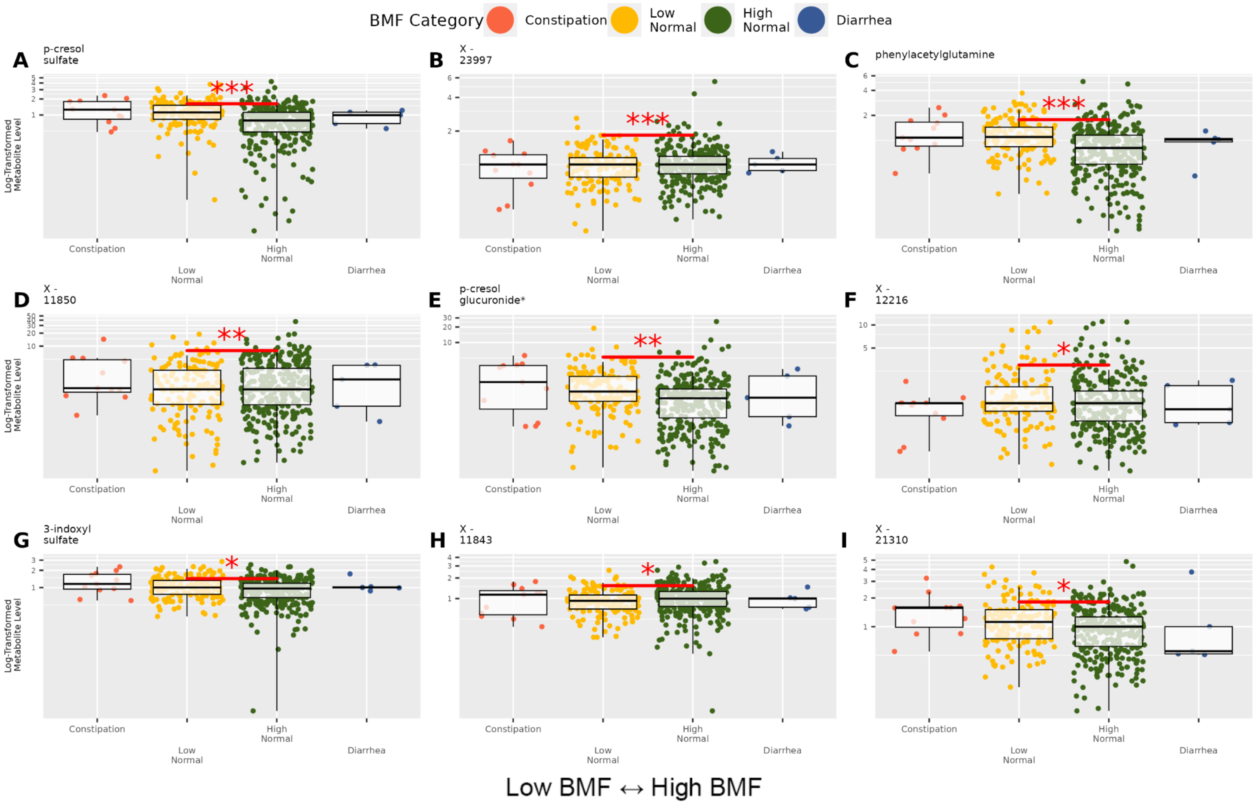


Figure S3. Significant BMF-associated plasma metabolites boxplots (A-I), related to Figure 5. Significant plasma metabolites from the LIMMA analysis. The horizontal axes are annotated as four BMF categories: “Constipation” (BMF = 1-2✕ per week), “Low Normal” (BMF = 3-6✕ per week), “High Normal” (BMF = 1-3✕ per day) which is the reference category in regression, and “Diarrhea” (BMF = 4✕ or more per day). Red significant comparison lines across each plot denote significant differences from the reference category (“High Normal” BMF), and asterisks denote FDR-corrected significance threshold. (***): $p < 0.0001$, (**): $0.0001 < p < 0.01$, (*): $0.01 < p < 0.05$.

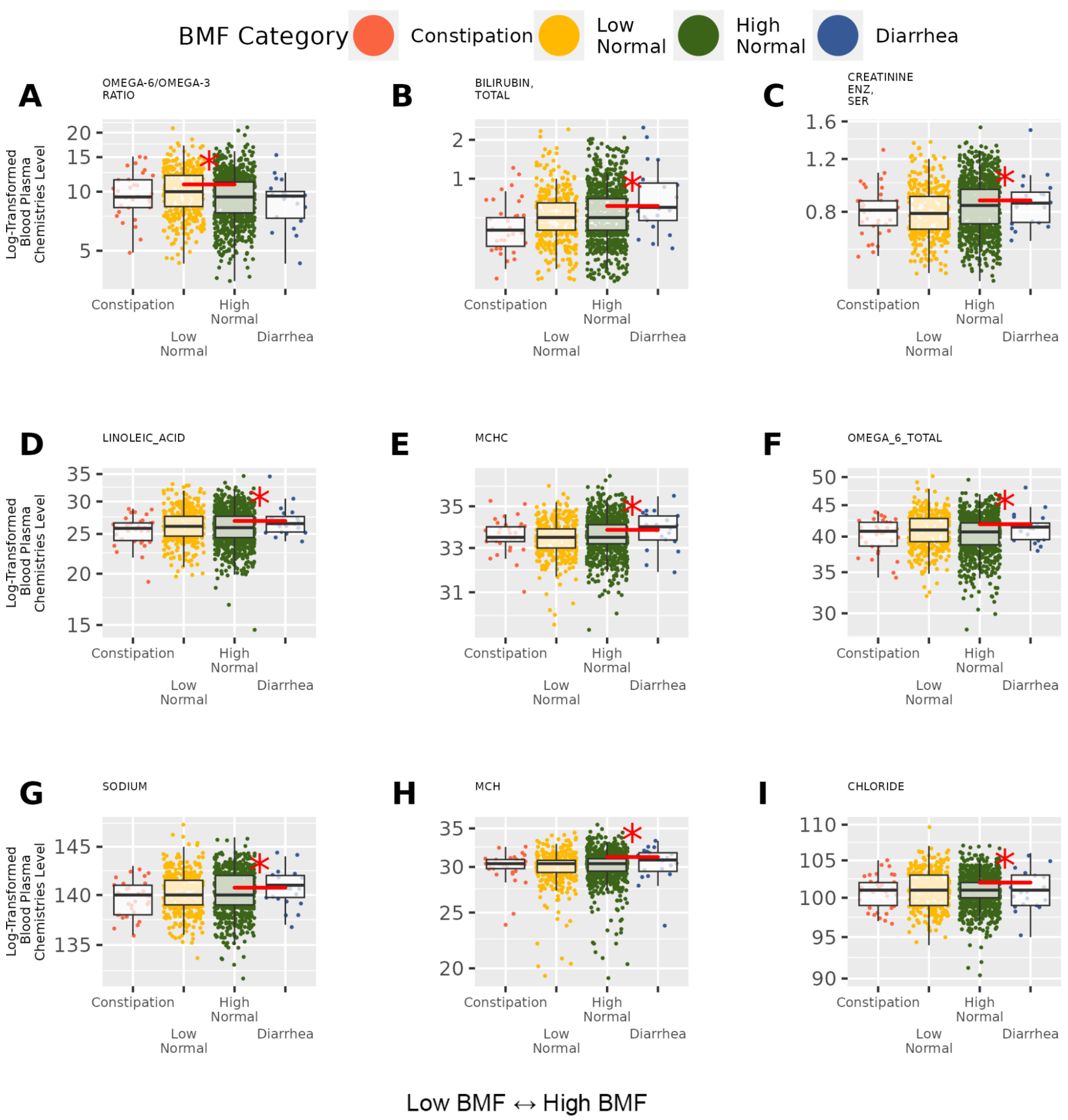


Figure S4. Significant BMF-associated clinical chemistries boxplots (A-I), related to Figure 6. Significant clinical chemistries from the LIMMA analysis. The horizontal axes are annotated as four BMF categories: “Constipation” (BMF = 1-2✕ per week), “Low Normal” (BMF = 3-6✕ per week), “High Normal” (BMF = 1-3✕ per day) which is the reference category in regression, and “Diarrhea” (BMF = 4✕ or more per day). Red significant comparison lines across each plot denote significant differences from the reference category (“High Normal” BMF), and asterisks denote FDR-corrected significance threshold. (***): $p < 0.0001$, (**): $0.0001 < p < 0.01$, (*): $0.01 < p < 0.05$.

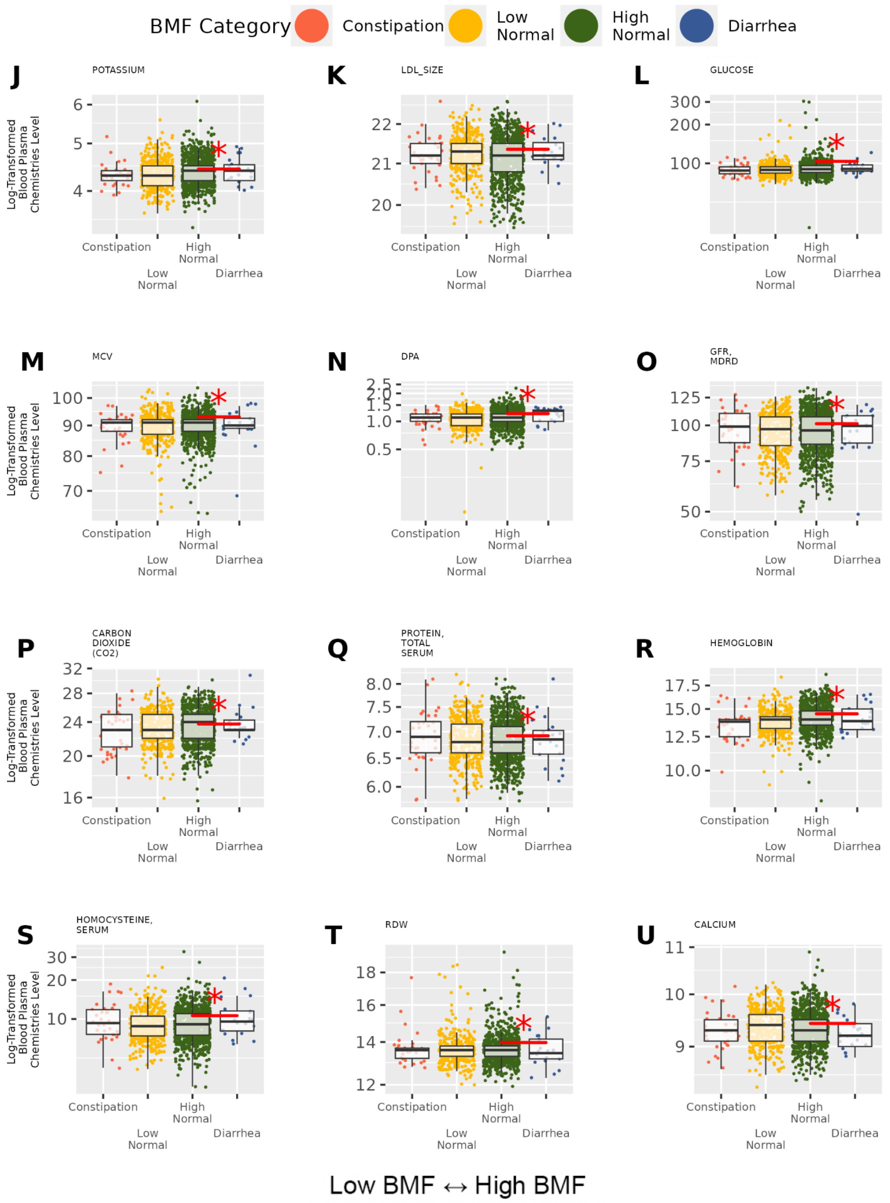


Figure S5. The remaining significant BMF-associated clinical chemistries boxplots (J-U), related to Figure 6. The remaining significant clinical chemistries from the LIMMA analysis. The horizontal axes are annotated as four BMF categories: “Constipation” (BMF = 1-2✕ per week), “Low Normal” (BMF = 3-6✕ per week), “High Normal” (BMF = 1-3✕ per day) which is the reference category in regression, and “Diarrhea” (BMF = 4✕ or more per day). Red significant comparison lines across each plot denote significant differences from the reference category (“High Normal” BMF), and asterisks denote FDR-corrected significance threshold. (***): $p < 0.0001$, (**): $0.0001 < p < 0.01$, (*): $0.01 < p < 0.05$.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   eGFR   R-squared:                       0.082
Model:                            OLS   Adj. R-squared:                  0.067
Method:                 Least Squares   F-statistic:                     5.547
Date:                Sun, 18 Feb 2024   Prob (F-statistic):           2.42e-07
Time:                        07:29:22   Log-Likelihood:                -2465.4
No. Observations:                 572   AIC:                             4951.
Df Residuals:                     562   BIC:                             4994.
Df Model:                           9
Covariance Type:            nonrobust
=========================================================================================
                            coef    std err          t      P>|t|      [0.025      0.975]
-----------------------------------------------------------------------------------------
const                   115.0755      4.841     23.770      0.000     105.566     124.585
bowel                    -3.9902      1.496     -2.667      0.008      -6.929      -1.051
p-cresol sulfate         -2.6898      2.473     -1.088      0.277      -7.548       2.168
X - 23997                 1.7076      1.423      1.200      0.231      -1.087       4.502
phenylacetylglutamine     2.2073      2.247      0.982      0.326      -2.207       6.622
X - 11850                -0.4421      0.359     -1.232      0.218      -1.147       0.263
p-cresol glucuronide      0.3677      0.490      0.750      0.454      -0.595       1.330
X - 12216                -1.9827      0.787     -2.520      0.012      -3.528      -0.437
3-indoxyl sulfate        -9.6859      2.249     -4.307      0.000     -14.104      -5.268
X - 11843                 0.0527      0.554      0.095      0.924      -1.036       1.141
==============================================================================
Omnibus:                       39.704   Durbin-Watson:                   1.841
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               22.585
Skew:                          -0.333   Prob(JB):                     1.25e-05
Kurtosis:                       2.290   Cond. No.                         45.4
==============================================================================
```

Figure S6. A) Flow chart for cohort selection of baseline Population. B) OLS regression resulting from eGFR ~ BMF-associated metabolites + BMF, related to Method Details. A) Individuals with the full complement of covariate data (gender, age, BMI, and CRP, LDL, A1C, and PCs 1-3) were further filtered for having available baseline data for each of the following: surveys, microbiome profiles, proteomics, clinical chemistries (e.g. complete blood count, or CBC; and comprehensive metabolic panel, or CMP) and metabolomics. The "generally-healthy" exclusion criteria were then imposed (38.5% excluded; see Method Details), along with sparsity or non-missingness minimums for the features in the 'omics data (≥ 30% prevalence for gut microbiome data, metabolomics and clinical chemistries; ≥ 50% prevalence for proteomics; and ≥ 90% prevalence and ≥ 10% affirmative for binary responses in the survey questions). These filters resulted in the final sub-cohort numbers shown on the right side of the figure in blue outlines. Additionally, the eGFR and BMF data frames were merged with the metabolomics data frame and filtered by the "generally-healthy" exclusionary criteria for the 9 BMF-associated metabolites eGFR regression and mediation analysis. B) The p-value for the overall generalized-linear model (eGFR ~ BMF-related metabolites) was significant (N = 572, p = 2.42E-7, R^2 = 0.082) and so were the p-values of the individual β-coefficients for 3-IS ($\beta_{3\text{-IS}}$ = -9.69, p = 1.96E-5), BMF (denoted "bowel"; β_{BMF} = -3.99, p = 7.88E-3), and X - 12216 ($\beta_{X\text{ - }12216}$ = -1.98, p = 1.20E-2).

1.7 METHODS

Resource availability

Lead contact

Additional requests and information regarding resources, experimental materials, reagents, and assay vendors should be directed to and will be fulfilled by the lead contact, Sean Gibbons (sgibbons@isbscience.org).

Materials availability

This study did not generate reagents.

Key Resources Table

Reagent or resource	Source	Identifier
Software and Algorithms		
Custom Python and Jupyter Notebooks and R markdown notebooks	This study	https://github.com/Gibbons-Lab/Aberrant-BMF-Cell-Reports
R and RStudio	Posit.co [73]	https://posit.co/download/rstudio-desktop/
CORNCOB R Package	Martin et al. [36]	https://github.com/statdivlab/corncob
Bioconductor R Package	Bioconductor [36,74]	https://www.bioconductor.org/install/
tidyverse R Packages	tidyverse [75]	https://github.com/tidyverse/tidyverse/
LIMMA R Package	Smyth et al. [76,77]	https://kasperdanielhansen.github.io/genbioconductor/html/limma.html
POLR / MASS R Package	Venables and Ripley et al. [35]	https://rdocumentation.org/packages/MASS/versions/7.3-60.0.1

Data and code availability

- Code used to analyze 16S rRNA gene amplicon sequencing data can be found at https://github.com/gibbons-lab/mbtools.
- Code used to run the statistical analyses described in this paper is available at https://github.com/Gibbons-Lab/Aberrant-BMF-Cell-Reports .
- Qualified researchers can access the full Arivale deidentified dataset, including all raw data, supporting the findings in this study for research purposes through signing a Data Use Agreement (DUA). Inquiries to access the data can be made at data-access@isbscience.org and will be responded to within 7 business days.

1.8 EXPERIMENTAL MODEL AND SUBJECT DETAILS

Institutional review board approval for the study

The procedures for this study were reviewed and approved by the Western Institutional Review Board, under the institutional review board study number 20170658 for the Institute for Systems Biology and 1178906 for Arivale, Inc.

Generally-healthy cohort

All study participants were subscribers in the Arivale Scientific Wellness program (2015-2019) and provided informed consent for the use of their anonymized, de-identified data for research purposes. Participants were community-dwelling, residents of Washington State and California (which are slightly leaner and healthier than other parts of the USA), over the age of 19, non-pregnant, but were not screened for the presence or absence of any particular disease. Participants provided detailed questionnaire data that included self-reported information about medical conditions and medications, along with blood and stool samples that were used to generate blood plasma metabolomics, proteomics, clinical laboratory chemistries, and gut microbiome data (**Fig 1** and **Table S1**). 3,955 individuals with at least one BMF data point available from the Arivale program were initially considered for this analysis. Their data were merged in Python (Jupyter Notebooks) with those with eGFR data (N = 6,133) and those with available survey (N = 5,764), ASVs (N = 3,694), proteomics (N = 2,859), clinical laboratory chemistries (N = 4,881), or metabolomics (N = 2,043) data samples to yield dataframes representative of the cohorts at the intersections of these merges. The final dataframes for the regressions for survey (N = 1,420), ASVs (N = 1,062), proteomics (N = 523), clinical laboratory chemistries (N = 1,425), and metabolomics (N = 486) cohorts were used for downstream regression analyses (**Fig S6**).

Only baseline time point samples were used for each participant for the baseline ‘omics analyses. A 30% prevalence filter was implemented across the gut microbiome, blood plasma metabolomics, proteomics, chemistries, and ordinal questionnaire data analyses. This meant that each final feature in the data could contain no more than 70% missing data from the final cohort of samples in order to be retained for downstream analysis. For microbiome analyses, a filtered subcohort of 1,062 individuals with ASV-level taxa counts, BMF, gender, age, eGFR, BMI, LDL, CRP, A1C, and genetic ancestry data were selected. This filtering resulted in a total of 135 genera. For the metabolomics analysis, a cohort of 486 participants with BMF, gender, age, eGFR, BMI, CRP, LDL, A1C, PC1, PC2, and PC3, and blood metabolomics data were selected. 956 metabolites were retained for downstream analyses. 274 blood proteins that met the prevalence (≥ 50%) filter in the cohort of 823 individuals were retained for downstream analyses. A ≥ 30% prevalence filter was applied to yield 1,425 samples with blood plasma clinical laboratory chemistries data, resulting in 55 features retained for downstream analyses. Similarly, for ordinal regression of the questionnaire data (e.g. diet, lifestyle, and stress/pain/health factors,) using the respective R package, polr [35], we collated all the responses and filtered out questions that contained more than 10% “NAs” (≥ 90% prevalence; and for binary variables in downstream depression/anxiety analyses: ≥ 10% affirmative or “True” responses). We also excluded binary response variables for the general survey questionnaire analysis (separate from the anxiety/depression analysis, which leveraged binary response features), which are incompatible with ordinal regression, resulting in 138 variables across 1,420 participants, in addition to having paired data on age, gender, eGFR, BMI, BMF, CRP, LDL, A1C, PC1, PC2, and PC3. The final features considered needed to retain at least 2 non-missing factors (or categories) and contain at least 10 responses per category, which resulted in 99 features. BMF data was captured from responses to a survey question on how many bowel movements an individual has per week, on average. The available responses to this question were: (1) Twice per week or less; (2) 3-6 times per week; (3) 1-3 times daily; or (4) 4 or more times daily. While the normal range of BMF encompasses both the second and third responses to this question (i.e., between three times a week and three times a day) [78], we chose to define 1-3 times per day (high-normal) as the reference group for the purposes of regression.

Finally, we imposed disease-related exclusion criteria, in order to generate a “generally-healthy” sub-cohort. These include any participants who reported affirmative or “true” to a history of taking cholesterol, laxative, or blood pressure medication, as well as those who reported a self- or family- history presence of the following diseases: bladder or kidney disease, inflammatory bowel disease (IBD), celiac disease, diverticulosis, gastroesophageal reflux disease (GERD), irritable bowel syndrome (IBS), or peptic ulcers (See **Table S1** in Supplement). 988 (25%) out of the initial 3,955 Arivale individuals with BMF data were excluded by these filters.

1.9 METHODS

Gut microbiome data

Fecal samples from Arivale participants were collected (described in Diener et al [19] and detailed here) from proprietary at-home kits developed by two microbiome vendors (DNA Genotek and Second Genome). Using the KingFisher Flex instrument, the MoBio PowerMag Soil DNA isolation kit [79] (QIAGEN) enabled the isolation of stool DNA from 250 ml of homogenized human feces, after performing an additional glass bead-beating step. Qubit measurement and spectrophotometry were also performed using an A260/A280 absorbance ratio.

16S amplicon sequencing was run on a MiSeq [80] (Illumina, USA) with either paired-end 300-bp protocol (DNA Genotek) or paired-end 250-bp protocol (SecondGenome). The FASTQ files were provided by the Illumina Basespace platform after the phiX reads were removed with basecalling. Length cutoffs of 250-bp for the forward reads and 230-bp for the reverse reads were employed. Any reads with more than 2 expected errors or ambiguous base calls under the Illumina error model were eliminated. Over 97% of the reads passed these filters, resulting in approximately 200,000 reads per sample.

Final truncated and filtered reads were then used to infer amplicon sequence variants (ASVs) with DADA2 [81–83]. Each sequencing run separately resulted in its own error profiles. The final ASVs and counts were then joined, with chimeras removed using DADA2's "consensus" strategy. After this step, ~16% of reads were removed. Taxonomic assignment of ASVs was then achieved using the naive Bayes classifier in DADA2 with the SILVA database (version 128) [84].

Nearly 90% of the ASVs were classified down to the genus level, which was the taxonomic level chosen for this analysis. 3,694 samples across 609 taxa were available from these methods, which were then filtered down to 135 taxa after using a 30% prevalence filter (no more than 70% of data was permitted to be missing per filtered taxa). Samples were rarefied to an even depth of 13,703 reads prior to calculating alpha-diversity metrics (using the "rarefy_even_depth( )" function in the phyloseq R package [85]; rng seed = 111). ASV richness (Observed ASVs), Shannon Diversity, and Pielou's evenness were calculated. Merging with covariate data resulted in 1,062 samples with 135 taxa for downstream analyses.

Olink proteomics

Blood plasma proteomic data were generated by Olink Biosciences using the ProSeek Cardiovascular II, Cardiovascular III, and Inflammation arrays. The proteins were filtered down to 274 proteins and 823 samples, retaining proteins with ≥ 50% prevalence across samples and samples with the full set of covariate data. Post-filtering, NAN values were assumed to be below detection and imputed to be the median across samples for that particular protein. The values used for the proteomics analysis were from protein readings previously batch-corrected and normalized based on the overlapping reference samples within the batch plates (i.e., a set of Arivale plasma samples that are run with each batch). The corrected values were also scale-shifted to the reference sample and the original delivered data (using the seventh run as a baseline). Olink's Proximity Extension Assay (PEA), a 2-antibody-barcode technology [81,82,86], is used to tag protein biomarkers with a proximity probe (which binds specifically to the target protein biomarker) and an extension probe (which carries a DNA barcode sequence) as described by Illumina in conjunction with Olink [87–89]. Once both probes bind to each other due to a protein-protein interaction or by proximity, they trigger the activation of the extension probe, beginning the hybridization of the probe with a detection bead's complementary DNA sequence. Each bead contains an individual identifier, which allows target proteins to be decoded according to a barcode. These methods are also described further in Zubair et al [81].

Metabolon metabolomics

Metabolon obtained metabolomics data on the previously mentioned plasma samples using preparation, quality control, and collection methods described in previous studies [82]. During sample processing, the plasma samples were thawed and proteins were removed using methanol extraction. Samples were then divided into 5 fractions including a backup fraction. Organic solvents were removed using TurboVap and measurements were then performed using high-performance liquid chromatography (HPLC) and high-resolution mass-spectrometry (MS) [90]

. Four separate measurements were performed using different fractions combinations: positive-ion and negative-ion modes optimized for both hydrophobic and hydrophilic compounds. Batch correction was performed using quality control samples (i.e., a set of Arivale plasma samples that were run with each batch) and abundance data were normalized to these quality control samples. Metabolites were annotated according to 3 standards: Tier 1, matching to an internal standard; Tier 2, matching to a published MS spectrum; or Tier 3, matching to a known chemical formula. Unknown metabolites were unannotated and labeled with an "X - " label followed by an identifier [91]. 956 total metabolites showed at least 70% prevalence across 486 samples. In this analysis, missing values were imputed to be the median of the non-missing samples for each metabolite, and final downstream metabolites were log-transformed and merged with the full set of covariates.

For the multi-linear regression and causal mediation analysis, those with paired eGFR, BMF-associated metabolomics results, and BMF were filtered using the "generally-healthy" exclusionary criteria and the previously mentioned prevalence filtering for metabolomics. The remaining individuals (**Fig. 7B, S7**; N = 572) were processed in a multi-linear regression (OLS) with eGFR ~ BMF (encoded as a value between 1 and 4 with 4 being diarrhea or the highest BMF) + the obtained metabolomics values for the 9 BMF-associated metabolites (**Fig. S6).** The other multi-omics covariates (gender, age, BMI, CRP, LDL, A1C, and PC1-PC3) were not considered for the subsequent mediation analysis (**Fig. 7B**; N = 562), which was performed using a mediation model with the mediate( ) function from the mediation package in R [31]. Using this modeling function, the outcome model was specified as eGFR ~ 3-IS + BMF (where BMF was encoded as a binary categorical variable, with "Low" including those with low-normal BMF and constipation, and "High" containing those with high-normal BMF and diarrhea. "Low" was the control value for BMF and "High" was the treatment value) and the mediation model was assumed to be 3-IS ~ BMF. ACME and ADE values were obtained from the model and reported using the diagram in **Fig. 7B**. A GLM was also performed between eGFR ~ BMF, 3-IS ~ BMF, and eGFR ~ 3-IS to obtain the β-coefficients and p-values for the relationships between the mediated variables (**Fig. 7B**). Ultimately, we also performed a similar mediation analysis as before, but with the outcome model including eGFR ~ 3-IS + BMF + vegetables intake and a mediation model containing 3-IS regressed against BMF + vegetables intake. This modeling strategy was applied to those with questionnaire survey data (N = 571) on vegetable eating habits (respondents claiming to eat 1 or less vegetables per day were in the "Low" treatment group, while those eating more vegetables than that daily were in the "High" control group) for the participants that self-responded to the inquiry of daily vegetable eating habits, implying a relationship between dieting factors and BMF on eGFR values through the proxy of 3-IS.

Blood plasma chemistries

LabCorp and Quest [92] phlebotomists collected blood from Arivale participants using methods described previously by Wilmanski et al and others [19]. Individuals were asked to abstain from alcohol, vigorous exercise, monosodium glutamate and aspartame at least 24 hours prior to drawing of the blood, as well as fasting at least 12 hours beforehand. Blood samples were collected for clinical chemistries, metabolomics and proteomics at the same time, and within 21 days of stool sampling. BMI was calculated from weight and height using the following formula $BMI = \frac{weight\ (kg)}{(height\ (m))^2}$. 4,881 samples and 68 laboratory values were filtered down using the same prevalence filtering as the metabolomics data. 1,425 samples and 55 chemistries were retained. The final 55 features were log-transformed, with missing samples imputed to be the median value of the non-missing samples. These features were merged with the full set of covariates. eGFR was calculated based on the CKD Epidemiology Collaboration (CKD-EPI) creatinine Equation, as recommended by the current guidelines of the National Kidney Foundation [93]: $eGFR_{cr} = 142 \times \min(Scr/\kappa, 1)^{\alpha} \times \max(Scr/\kappa, 1)^{-1.200} \times 0.9938^{Age} \times 1.012$ [if female], where Scr

= standardized serum creatinine in mg/dL, κ = 0.7 (female) or 0.9 (male), and α = -0.241 (female) or -0.302 (male).

Questionnaire data

3,482 self-reported questionnaire features were retrieved across 5,764 Arivale participants. After health and prevalence filtration, 138 downstream features remained, which were subsequently filtered down again to 99 final features by removing factored features with fewer than 10 responses per level and keeping features with at least 2 non-missing levels to the factor. Category responses were organized and numbered to be ordinally ascending in magnitude or intensity, with relatively even-spaced differences in magnitude between categories wherever possible (i.e. for a factored feature with levels from 1,…,n, the level labeled “1” represented responses such as “Strongly Disagree”, “Never”, “None”, or the lowest frequency/intensity, and the level labeled “n” represented responses such as “Strongly Agree”, “Always”, or the greatest frequency or intensity). These features were merged with the full set of covariate data.

Depression and anxiety health history data

We used logistic regression to scrutinize associations between 23 (anxiety) and 35 (depression) independent binary (“true” or “false”) self-reported questions based on “self-current”, “self-past”, and “family” histories of depression or anxiety, with depression or anxiety encoded as a binary dependent variable, and BMF encoded as a categorical independent variable, and with the standard set of covariates.

1.10 QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses

Using Python (Jupyter Notebooks) [94] and R (RStudio) [73] as well as bioinformatics packages like Phyloseq and Bioconductor [74,85] and tidyverse [75], multiple linear regressions were performed to characterize these variables. The response variables were either: centered log ratio-transformed bacterial genus data, log-transformed plasma metabolomics data, batch-corrected plasma proteomics data, log-transformed plasma chemistries data, or ordinal response variables from questionnaire data, depending on the analysis. For the blood proteomics, plasma chemistries, and metabolite associations, generalized linear regression models were run using the LIMMA package in R [76]. BMF was encoded as a categorical variable (or in the case of analyzing alpha-diversity, it was also computed as an ordinal variable with a linear model coefficient) with categories: 1 = constipation (1-2 bowel movements per week), 2 = low-normal (3-6 bowel movements per week), 3 = high-normal (1-3 bowel movements per day), and 4 = diarrhea (4 or more bowel movements per day). To begin characterizing the main variables in the cohorts: BMF and eGFR, a POLR regression [35] (N = 1,425) was performed on BMF (encoded as an ordinal variable with categories "Constipation", "Low Normal", "High Normal", and "Diarrhea" BMF in ascending order of magnitude) ~ eGFR + other covariates (gender, age, BMI, CRP, LDL, A1C, PC1, PC2, and PC3). Similarly, a GLM (N = 1,425) was computed for eGFR ~ BMF (also encoded ordinally) + other covariates (gender, age, BMI, CRP, LDL, A1C, PC1, PC2, and PC3). These were used to determine the significant covariates affecting each subsequent analysis (**Fig. 2**). Next, in each baseline regression, the following covariates were all included: gender, age, BMI, eGFR, CRP, LDL, A1C, PC1, PC2, and PC3. Gut bacterial genus-level counts were modeled with a beta-binomial distribution using the CORNCOB package in R [36]. For the questionnaire data (ordinal response categories across diet, exercise, stress, pain, and other lifestyle factors), polr in R was used for the ordinal regression analysis (POLR) [35]. For the anxiety and depression data, which were binary in response (“True” or “False”; Non-responders to each feature were not considered and features were filtered to have at least 5 non-missing responses for each binary outcome), logistic regression was performed using the “glm(family = “binomial”)” function in R. All questionnaire and anxiety/depression response modeling results were FDR-corrected for significance. Finally, for the Arivale cohort, the initial time point or baseline value for eGFR was obtained alongside the initial or earliest time point sample for the BMF-related metabolites. eGFR was regressed against the BMF-associated metabolites in an OLS-based linear regression to determine visible effects of these metabolites on our available samples. Finally, a mediation analysis was run using the mediate( ) function in the mediation library available for R [37] on the individuals who met the “generally-healthy” exclusion criteria with paired eGFR, BMF, and 3-IS data. BMF was the treatment variable, 3-IS was the mediator, and eGFR was the response variable. ACME, ADE, total effect and proportion mediated were determined with nonparametric bootstrap confidence intervals.

CHAPTER 2: FRAILTY, HEALTHY AGING, AND THE GUT MICROBIOME

2.1 SUMMARY

In the past century, human life expectancy has more than doubled, resulting in an ever-growing and aging population [95] . This demographic shift presents unique healthcare challenges. Aging has been linked to an unraveling of the intricate systems biology underlying organismal homeostasis, leading to lower stress resiliency, increased inflammation, organ system dysfunctions, the accumulation of senescent cells, and more, resulting in a steady rise in frailty and risk of death over the lifespan [96–101]. Despite significant advances in data collection and integration across various multi-omics and spatial biology platforms, substantial gaps remain in our understanding of the host-microbe interplay in aging and frailty. Here, we investigate the relationship between markers of frailty and other multi-omic features, including the gut microbiome, across >1,000 generally-healthy individuals from the Arivale cohort. Gut bacterial genera, gut microbiome uniqueness, estimates of gut microbiome SCFA production, untargeted metabolomic features, and proteomic features were independently regressed against a 70-point frailty index score (FIS), controlling for age, gender, BMI, and the first four principle components of genetic ancestry. Our results revealed a pronounced inflammatory signature across all 'omics data types, with enrichment of pro-inflammatory gut bacterial genera, metabolites, and proteins in frail individuals, concurrent with a worsening in cardiometabolic health markers. We also find strong associations between cardiometabolic health, inflammation, and FIS. These findings underscore the importance of integrating gut microbiome research into human longevity studies and emphasize the need for future frailty research designs that incorporate longitudinal multi-omic data collection to elucidate the temporal dynamics of host-microbe interactions in the context of aging and frailty.

2.2 INTRODUCTION

Aging is a complex process characterized by hallmark traits that progressively manifest throughout adulthood, including genomic instability, telomere attrition, epigenetic alterations, loss of proteostasis, dysregulated nutrient sensing, stem cell exhaustion, cellular senescence, altered intercellular communication, mitochondrial dysfunction, and gut microbiome disruption. These traits are mechanistically interconnected through their effects on oxidative stress and serve to further exacerbate inflammation and tissue/organ/cellular damage over time [96–101] .

Inflammation, sometimes described as “inflammaging”, is a key driver of frailty [102] . Low-grade, chronic inflammation elevates levels of pro-inflammatory cytokines like interleukin-6 (IL-6), interleukin-18 (IL-18), tumor necrosis factor-α (TNF-α), and CRP. These cytokines disturb tissue homeostasis, contributing to various aging-related processes and diseases. IL-6 enrichment, for example, has been strongly linked to frailty, disability, and mortality in the elderly [103–105103–107] . In “inflammaging,” oxidative stress and reactive oxygen species (ROS) exacerbate inflammation, creating a positive feedback cycle accelerating frailty and aging. Fibrosis from chronic inflammation can also occur, which is when excessive deposition of extracellular matrix proteins build-up and stiffen the tissues, impairing organ function [108,109] . All these cascading effects intensify severity and likelihood of frailty during the aging process.

The microbiota of the gastrointestinal tract play a major role in chronic inflammation and aging [110–113] . By fermenting indigestible fibers from our diet, the microbiota subsequently release SCFAs into the gut environment, which control multiple processes, including functioning as a fuel source for other bacteria, a signal for manipulating microbial functions, combating pathogens, protecting and maintaining intestinal epithelial integrity, and regulating immune cells [114] . For example, SCFA production regulates the differentiation of CD4+ (helper) T cells and activation of CD8+ (cytotoxic lymphocytes) T cells. The absence of commensal microbes that produce SCFA leads to impaired nutrient absorption and a deficiency in maturation of intestinal immune cells, tilting the immune response of the intestine toward pro-inflammatory cytokines and away from producing antimicrobial peptides (AMP) and immunoglobulin A (IgA) [115] . All of these signal a systemic inflammatory environment that promotes disease.

Recent advances in the aging field have revealed that the gut microbiota appears to diverge into two distinct phenotypic trajectories late in life. One trajectory is characterized by increased beta-diversity uniqueness and depletion of core gram-negative taxa present earlier in life, which has been associated with lower frailty and healthier aging. The other trajectory is marked by decreased uniqueness, maintenance of core gram-negative taxa, and higher frailty, indicating less-than-healthy aging [116] . While the possibility of microbiome-mediated healthy aging interventions, like fecal microbiota transplantation (FMT) from the young to the old, is intriguing, the observed divergence in trajectories suggests a more nuanced approach is needed. A deeper understanding of the mechanisms driving the observed healthy aging signature in the gut is key to unlocking the therapeutic potential of the microbiome to extend the human healthspan.

2.3 RESULTS

FIS and multi-omics analysis

We explored associations between a 70-point FIS, calculated for the Arivale cohort, and molecular phenotypic features (blood proteomes, blood metabolomes, clinical chemistries, and stool microbiomes), while accounting for gender, age, BMI, and the first four principal components of genetic ancestry. The FIS took into account measures from clinical chemistries, family/health history questionnaire responses, and vital signs. Features that raised the FIS included higher LDL cholesterol, higher insulin resistance, higher blood pressure. The Arivale participants were filtered and divided into separate subcohort dataframes (Chapter 1.7-1.8) for each multi-omic analysis: gut microbiome, metabolomics, and proteomics. FIS regressions were performed using CORNCOB or LIMMA in R (see Chapter 1.7-1.8), with significant FIS associations presented in correlation heatmaps to depict the inter-relationships between these FIS-relevant ‘omic features.

Frailty associations with microbiome uniqueness, α-diversity, and estimated SCFA production fluxes

Propionate and butyrate flux were negatively correlated with FIS (independent of age in the case of propionate), but negatively associated with PC4 of genetic ancestry. SCFA fluxes [mmol/gDW/h] simulated from MICOM [117] (butyrate and propionate) for the Arivale cohort (assuming a standard European diet) increased with rising uniqueness values (Weighted-UniFrac; **Fig. 9**). Butyrate and propionate have anti-inflammatory effects in the host that could counteract the trend of rising inflammation that is linked with the aging process [91] .

Significant genera variation across frailty

Data on 609 gut genera were obtained and then merged and filtered down to 582 for analysis (See Chapter 1 methods). 582 gut bacterial genera were then analyzed across samples (CORNCOB, N = 1,204), and 80 showed a significant association with FIS, independent of gender, age, BMI, and PC1-PC4. The significant genera and their correlations with FIS, its covariates, related features, and several cardiometabolic, kidney, and liver markers from clinical chemistries are depicted in a correlation heatmap (**Fig. 10**). Overall, bacterial genera associated with lower inflammation (e.g., butyrate-producing clostridia) were negatively associated with FIS, while pro-inflammatory genera like bacilli were positively associated with FIS.

Significant blood metabolomics variation across frailty

1,296 metabolites were considered, and filtered and merged (see Chapter 1 methods) to intersect all necessary covariates, resulting in 1,292 final metabolites for analysis. 271 out of 1,292 blood plasma metabolites (LIMMA, N = 539) were significantly associated with FIS, independent of covariates (gender, age, BMI, and PC1-PC4). These metabolites were depicted on a correlation heatmap, along with cardiometabolic health markers, FIS, covariates, α-diversity (Shannon), SCFA flux (butyrate and propionate) and two measures of β-diversity uniqueness (Weighted-UniFrac and Bray-Curtis). The results showed significant correlations with inflammatory and anti-inflammatory metabolites (**Fig. 11**).

Significant blood proteomics variation across frailty

251 of the total 274 blood proteins considered and analyzed were found to be significant drivers of variation with FIS (LIMMA, N = 761), independent of covariates (gender, age, BMI, and PC1-PC4). These significant proteins were plotted on a correlation heatmap, as described above (**Fig. 12**). Similar to the other ‘omics results, proteins known to be associated with inflammation are also associated with FIS.

2.4 DISCUSSION

Frailty associations with uniqueness, alpha-diversity, and estimated SCFA fluxes

Butyrate showed a significant association with FIS (labeled “merge_fi” in **Fig. 8**) until the BMI covariate was added to the model, which ablated the signal. This result supports prior observations that estimated butyrate flux is negatively associated with BMI [118] . Butyrate was significantly associated with age and PC4 and marginally associated with BMI and PC1. Propionate was also associated with PC4 (but not with age) as well as marginally associated with gender (labeled “sex” in **Fig. 8**). Uniqueness showed significant β_{age} and β_{BMI}-coefficients alongside marginal significance for β_{PC1}. Meanwhile, the regression for alpha-diversity detected significance for the β_{age}, β_{BMI}, and β_{PC1}–β_{PC3}-coefficients, as well as β_{FIS} (highly significant) and a marginal association with β_{gender}. FIS sharply decreased with rising Shannon diversity values. Uniqueness is shown plotted against butyrate flux, with a clear positive trend indicated in a scatterplot (**Fig. 9**).

Differentially abundant genera linked to varying SCFA production and pathobiont enrichment

Bacterial genera significantly associated with FIS shared some common traits (**Fig. 10**). There was a cluster of genera with strong negative associations with SCFA production and strong positive correlations with harmful cardiometabolic health markers and FIS: *Lachnochlostridium, Flavonifractor, Parasutterella, Bilophila, Eggerthella, Klebsiella, Fusobacterium,* and *Megasphaera*. While some of these genera, such as *Flavonifractor*, which is known to metabolize flavonoids (beneficial plant compounds), may play beneficial roles in human health, others like *Fusobacterium*, and *Klebsiella*, contain pathogenic species that may cause disease, and are linked to IBD and colorectal cancer [119–121] . High levels of *Bilophila* are also linked to inflammatory states that *Bilophila* thrive in, worsening the effects of the high-fat diet in a recent mice study [122] . Others with weakly positive correlations with FIS play uncertain roles, like *Hungatella*, a recently identified genus that may be involved in drug resistance in colorectal cancer treatment, and *Eggerthella* has shown strain-level effects on dopamine metabolism and drug inactivation [123–125] . However, those with strong negative correlations with FIS included butyrate producers like *Ruminiclostridium* and *Christensenellaceae_R*, which were positively associated with estimated SCFA fluxes and negatively associated with FIS [126,127] . Finally, in a recent frailty analysis, 18 species were identified as having significant effects with frailty, and many were part of a genus that exhibited the same directional correlations as in this analysis (e.g. *Hungatella, Roseburia, Adlercreutzia,* and *Faecalibacterium* [128]). *Faecalibacterium* (i.e. *prausnitzii*) is also associated with longevity, and in alignment with this study, *Faecalibacterium* anticorrelated with frailty [129] . Overall, these results reveal a cohesive health signature: in general, many of the microbes that suppress inflammation are reduced and the microbes that increase inflammation are enriched with higher levels of frailty and poorer cardiometabolic health.

Blood metabolites associated with amino acids, inflammation, and lipid metabolism

Metabolic features that were positively associated with FIS included certain sugars, lipid metabolism, and amino acid catabolism byproducts and derivatives (**Fig. 11**). 1-carboxyethylvaline, 1-carboxyethylisoleucine, modified amino acids from microbial metabolism, indicate oxidative stress or protein damage and correspond to CKD [130] . Elevated levels are linked to CKD and CVD. Glutamate, an excitatory neurotransmitter, was also elevated, and may suggest neurotoxicity or impaired glutamate metabolism, potentially associated with ND [131] . Meanwhile, γ-glutamylvaline, γ-glutamylleucine, and γ-glutamylisoleucine are dipeptides involved in glutathione metabolism, an important antioxidant system. Elevations in these or proline (involved in wound healing, fibrosis, and collagen) may reflect oxidative stress and are associated with cancer, liver disease, and ND [132–135] . Several elevated ceramides, which are sphingolipids involved in cell signaling, apoptosis, and inflammation, were also associated with insulin resistance, CVD, neuroinflammation, and ND [136–140] . N-stearoyl-sphinganine, another enriched sphingolipid, could potentially link to ND via disruption of sphingolipid metabolism. Glycerophospholipids (GPC, GPE, GPI), which are often components of cell membranes and signaling pathways, can implicate various diseases like ND, cancer, and CVD when their levels deviate. Acylcarnitines (myristoylcarnitine, propionylcarnitine) involved in fatty acid oxidation might indicate metabolic stress, mitochondrial dysfunction, diabetes, and CVD as well [141] . Interestingly, a fatty acid derived from the adrenal glands (adrenate 22:4n6), while not fully understood, also appeared elevated in the cohort, and it may be involved in stress responses [142] . Lactate (product of anaerobic glycolysis) also appeared to be enriched, which can indicate metabolic stress and mitochondrial dysfunction as well [143,144] . Additionally, certain cholesterol metabolism derivatives like deoxycholic acid glucuronide, which is involved in bile acid metabolism and cholesterol homeostasis, were also high in the frail. Increased levels could reflect changes in liver function, gut microbiome activity, and CVD [145,146] . There were also metabolites whose presence corresponded to a decrease in FIS, however; docosahexaenoylcholine (DHA, an omega-3 fatty acid) levels, well-known for its anti-inflammatory effects, coincide with this [147–149] . Ergothionine, a potent antioxidant, has also shown similar effects across numerous studies [150,151] . And certain amino acids, like taurine, which can modulate immune responses and reduce oxidative stress, appeared to be anticorrelated with FIS as well [152,153] . Metabolites that correlated negatively or with the lowest FIS values also included ones such as asparagine (diminished levels of this may impact inflammation and intestinal barrier integrity [154]). Despite early evidences, there remain significant knowledge gaps on the causal relationships between many of the negatively correlating amino acid derivative metabolites and inflammation and other disease states related to frailty, prompting a need for future researchers to explore biochemical mechanisms with longitudinal immunological and stool and blood data.

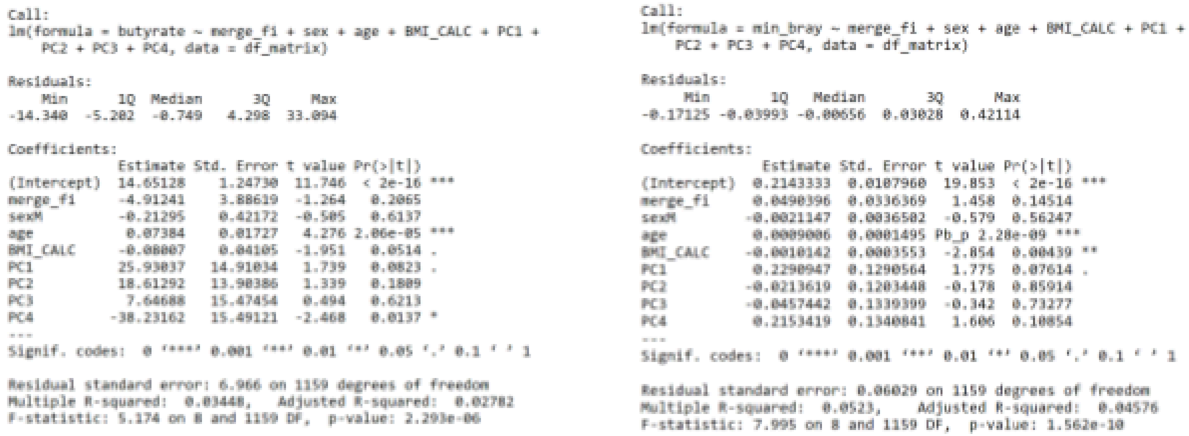
Blood proteomics associated with inflammation

Similar to the other ‘omics results, proteins that were positively associated with FIS indicated higher levels of inflammation. Many of the strongest positive correlations observed with FIS were for growth/signaling factors, immune response/inflammation molecules, and proteases. Myoglobin was one of the strongest positive correlations with increased frailty, and its elevated presence can indicate tissue repair and injury [155–157] . Another was dickkopf WNT signaling pathway inhibitory 1 (DKK1), a potent inhibitor of the WNT signaling pathway, which plays crucial roles in various cellular processes, including inflammation, cancer, and arteriosclerosis [158–160] . Additionally, chemokines such as chemokine (C-C motif) ligand 3 (CCL3), one that attracts and activates immune cells, play a central role in inflammation and were also positively correlated with frailty [161–163] . There were also significantly depleted levels of certain proteins in those with low FIS, particularly pro-fibrotic, pro-inflammatory chemokine (C-C motif) ligand 24 (CCL24), a key regulator of hepatic injury from inflammation [164,165] . More enriched proteins in the frail included immune receptors and signaling molecules like Fc fragment of immunoglobulin G (IgG), low affinity 11b, receptor (CD32, involved in immune complex clearance [166,167]); CD4 (T-cell activator and immunomodulator); tumor necrosis factor receptor (binds with pro-inflammatory cytokine TNF-α [168]); and polymeric immunoglobulin receptor, which is critical for mucosal immunity and IgA transport [169] . Other positively-associated proteins like chymotrypsin C (caldecrin, a digestive protease that breaks down nutrients in the small intestine [170]); fatty acid-binding protein 6, ileal (gut-derived and critical in lipid metabolism and transport in the ileum [171]); and leptin (the hormone critical to regulating appetite and energy [171,172]), were also significantly enriched at higher frailties. Proteins that anticorrelated with frailty included insulin-like growth factor binding protein 2, 36kDa (ILGFBP2), which regulates IGF activity, immune response, and insulin sensitivity; and serpin peptidase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1, also known as SERPINE1, is a gene that expresses plasminogen activator inhibitor 1 (PAI-1) [173]. Deficiencies in this protein or defects in the gene can cause bleeding issues, as this protein inhibits fibrinolysis (the dissolution of blood clotting factors). The clusters of highly frail adults and elevated inflammatory protein levels also corresponded with increases in concerning metrics like higher BMI, bilirubin, HDL, LDL, CRP, and A1C levels, but conversely also higher omega-3 fatty acid index amounts (a protective cardiovascular biomarker [174] ; **Fig. 12**). Meanwhile, sections of anti-correlation with FIS appeared on the heatmap as well, highlighting subtle but significant immune-related related proteins such as glyoxalase 1 (critical in stress response and cancer/inflammation [175]), CD40 ligand (a costimulatory molecule expressed on activated T cells and binds to CD40 on B cells and APCs and can suppress immune activity [176–178]), as well as serine/threonine kinase 4 (used in cell signaling and regulation [179]). CD40 and tumor necrosis factor receptor superfamily, member 10b and 1b (involved in cancer and immune regulation [176,180,181]) were among the proteins with the highest anticorrelations with frailty, while the latter also showed a strong positive correlation with linoleic acid. Linoleic acid is beneficial to cardiovascular health in moderate amounts by lowering blood cholesterol–however, overconsumption can also contribute to inflammation in the body [182] . Overall, several proteins that were positively associated with FIS were indicative of a stressful inflammatory environment in the host (**Fig. 12**).

Study limitations

Many of the same considerations as mentioned in Chapter 1 also apply here. We leveraged the same Arivale cohort that was used in Chapter 1. However, different participants were included/excluded, depending on available FIS and ‘omic data. As before, the cohort studied here was overwhelmingly “White”, predominantly female, and from the West Coast of the U.S.A., which limits the generalizability of our results. Additionally, due to the limitations of this study's FIS (a single score calculated from 70 metrics, many of which lacked repeated measures per participant), it wasn't possible to track individual changes in frailty over time. Similar to the approach in Chapter 1, only the baseline time point was utilized for each individual. Moreover, to maximize the sample size participants were not excluded based on their health status. Future research should prioritize longitudinal sampling of immunological markers alongside gut microbiota profiling in order to improve our mechanistic understanding of host-microbe interplay across the lifespan.

2.5 FIGURE TITLES AND LEGENDS



Butyrate Flux ~ FIS

```
Call:
lm(formula = butyrate ~ merge_fi + sex + age + BMI_CALC + PC1 +
    PC2 + PC3 + PC4, data = df_matrix)

Residuals:
    Min      1Q  Median      3Q     Max 
-14.340  -5.202  -0.749   4.298  33.094 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  14.65128    1.24730  11.746  < 2e-16 ***
merge_fi     -4.91241    3.88619  -1.264   0.2065    
sexM         -0.21295    0.42172  -0.505   0.6137    
age           0.07384    0.01727   4.276 2.06e-05 ***
BMI_CALC     -0.08007    0.04105  -1.951   0.0514 .  
PC1          25.93037   14.91034   1.739   0.0823 .  
PC2          18.61292   13.90386   1.339   0.1809    
PC3           7.64688   15.47454   0.494   0.6213    
PC4         -38.23162   15.49121  -2.468   0.0137 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.966 on 1159 degrees of freedom
Multiple R-squared:  0.03448,	Adjusted R-squared:  0.02782 
F-statistic: 5.174 on 8 and 1159 DF,  p-value: 2.293e-06
```

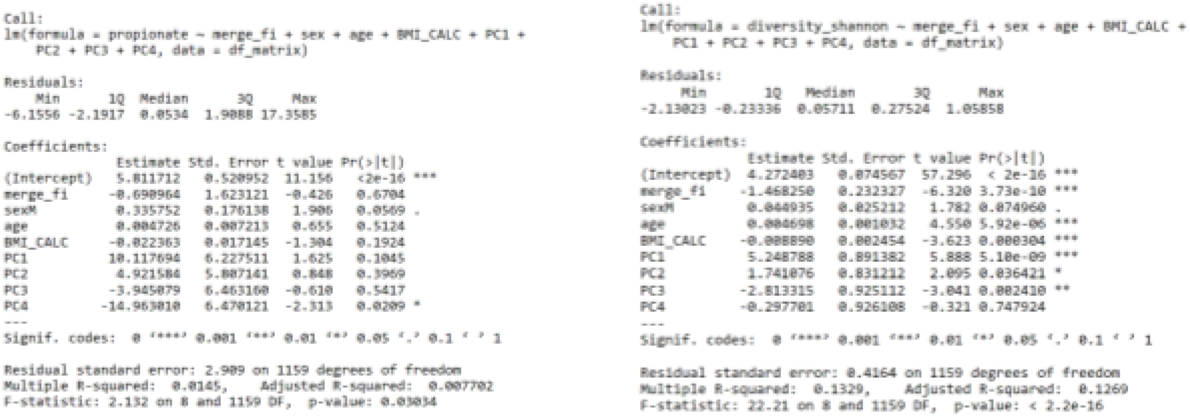
Bray-Curtis Uniqueness ~ FIS

```
Call:
lm(formula = min_bray ~ merge_fi + sex + age + BMI_CALC + PC1 +
    PC2 + PC3 + PC4, data = df_matrix)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.17125 -0.03993 -0.00656  0.03028  0.42114 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.2143333  0.0107960  19.853  < 2e-16 ***
merge_fi     0.0490396  0.0336369   1.458  0.14514    
sexM        -0.0021147  0.0036502  -0.579  0.56247    
age          0.0009006  0.0001495 Pb_p 2.28e-09 ***
BMI_CALC    -0.0010142  0.0003553  -2.854  0.00439 ** 
PC1          0.2290947  0.1290564   1.775  0.07614 .  
PC2         -0.0213619  0.1203448  -0.178  0.85914    
PC3         -0.0457442  0.1339399  -0.342  0.73277    
PC4          0.2153419  0.1340841   1.606  0.10854    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.06029 on 1159 degrees of freedom
Multiple R-squared:  0.0523,	Adjusted R-squared:  0.04576 
F-statistic: 7.995 on 8 and 1159 DF,  p-value: 1.562e-10
```



Propionate Flux ~ FIS

```
Call:
lm(formula = propionate ~ merge_fi + sex + age + BMI_CALC + PC1 +
    PC2 + PC3 + PC4, data = df_matrix)

Residuals:
    Min      1Q  Median      3Q     Max 
-6.1556 -2.1917  0.0534  1.9088 17.3585 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)   5.811712   0.520952  11.156   <2e-16 ***
merge_fi     -0.690964   1.623121  -0.426   0.6704    
sexM          0.335752   0.176138   1.906   0.0569 .  
age           0.004726   0.007213   0.655   0.5124    
BMI_CALC     -0.022363   0.017145  -1.304   0.1924    
PC1          10.117694   6.227511   1.625   0.1045    
PC2           4.921584   5.807141   0.848   0.3969    
PC3          -3.945079   6.463160  -0.610   0.5417    
PC4         -14.963010   6.470121  -2.313   0.0209 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.909 on 1159 degrees of freedom
Multiple R-squared:  0.0145,	Adjusted R-squared:  0.007702 
F-statistic: 2.132 on 8 and 1159 DF,  p-value: 0.03034
```

Shannon Diversity ~ FIS

```
Call:
lm(formula = diversity_shannon ~ merge_fi + sex + age + BMI_CALC +
    PC1 + PC2 + PC3 + PC4, data = df_matrix)

Residuals:
     Min       1Q   Median       3Q      Max 
-2.13023 -0.23336  0.05711  0.27524  1.05858 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  4.272403   0.074567  57.296  < 2e-16 ***
merge_fi    -1.468250   0.232327  -6.320 3.73e-10 ***
sexM         0.044935   0.025212   1.782 0.074960 .  
age          0.004698   0.001032   4.550 5.92e-06 ***
BMI_CALC    -0.008890   0.002454  -3.623 0.000304 ***
PC1          5.248788   0.891382   5.888 5.10e-09 ***
PC2          1.741076   0.831212   2.095 0.036421 *  
PC3         -2.813315   0.925112  -3.041 0.002410 ** 
PC4         -0.297701   0.926108  -0.321 0.747924    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4164 on 1159 degrees of freedom
Multiple R-squared:  0.1329,	Adjusted R-squared:  0.1269 
F-statistic: 22.21 on 8 and 1159 DF,  p-value: < 2.2e-16
```

Figure 8. Summary of GLM models of diversity, uniqueness, and SCFA flux regressed against FIS and its covariates.

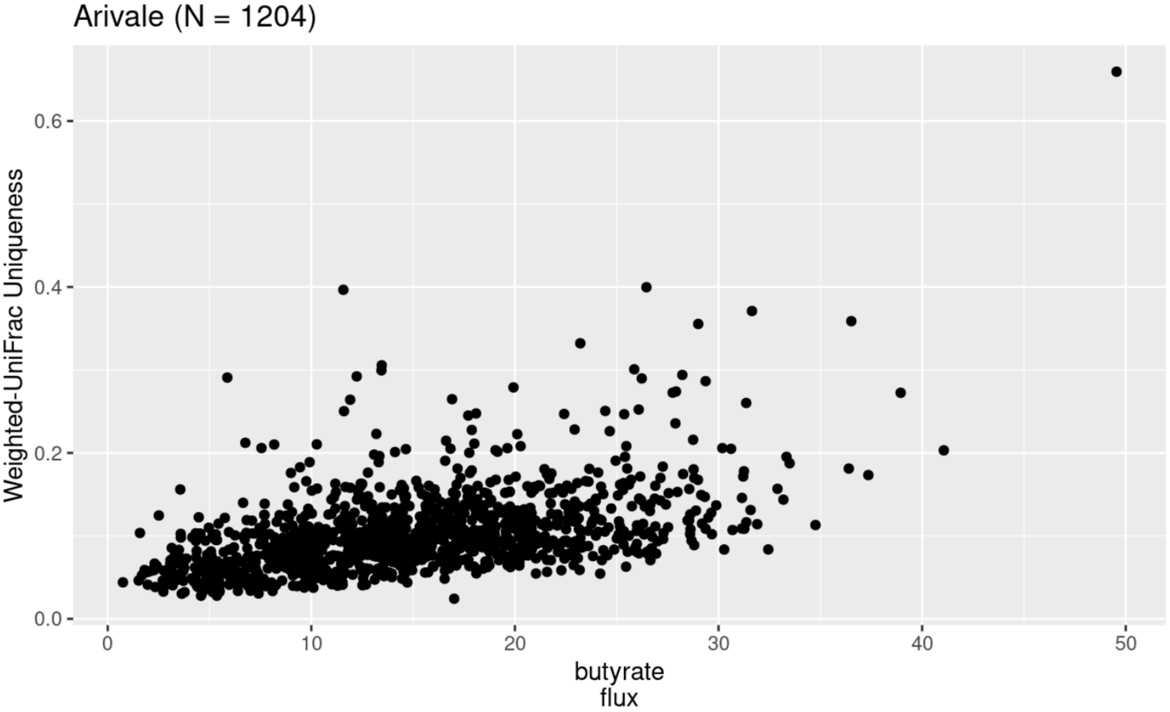


Figure 9. Scatterplot of microbiota uniqueness (Weighted-UniFrac) versus simulated SCFA flux. Flux is MICOM-simulated in units of [mmol/gDW/h].

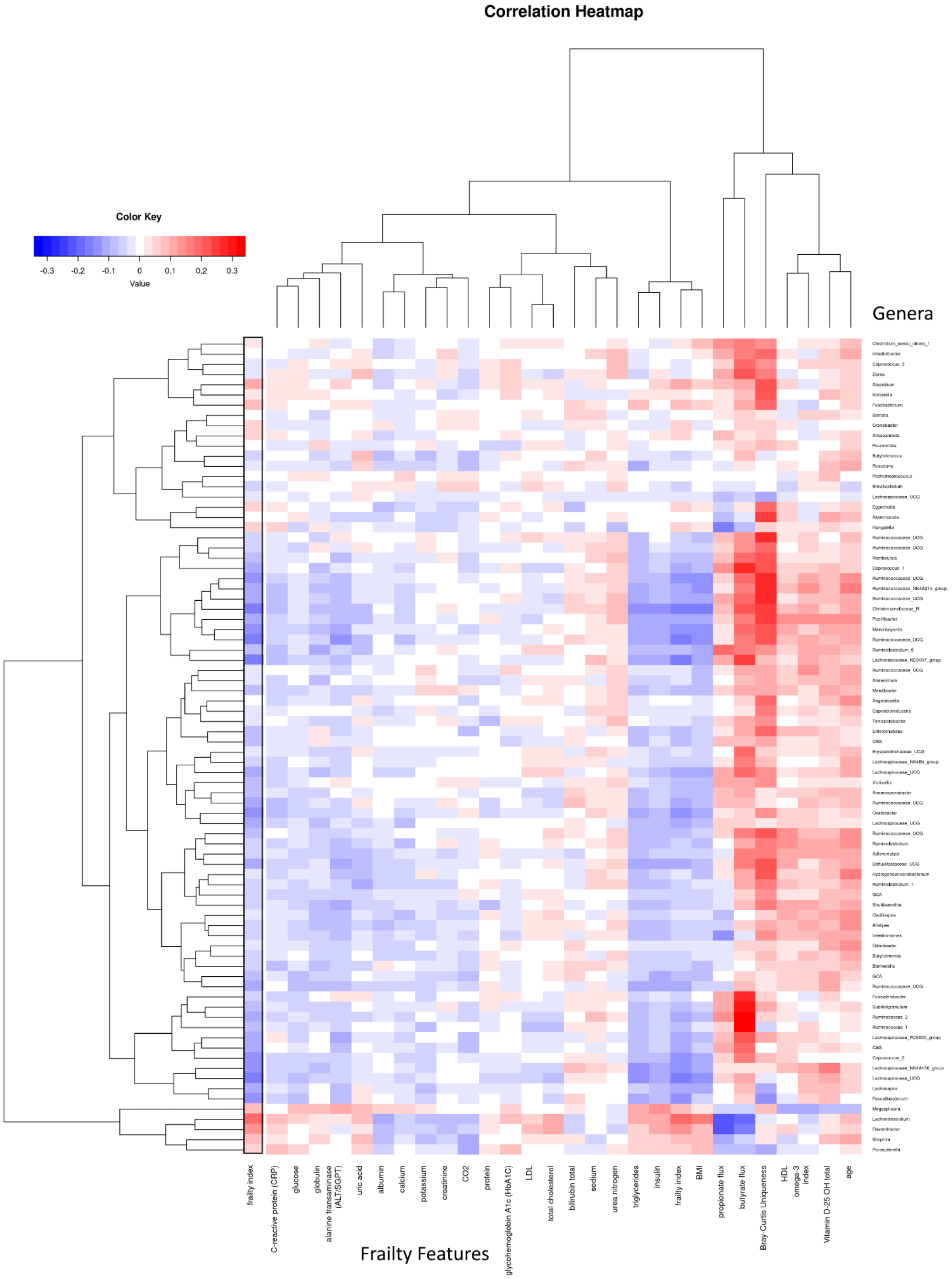


Figure 10. Correlation heatmap of significant annotated genera versus frailty measures, Bray-Curtis uniqueness, MICOM-simulated butyrate and propionate fluxes, and modeling covariates. Significant (FDR-controlled) differences in abundance across taxa were computed using CORNCOB, and simulated butyrate and propionate flux values were determined using MICOM. Elements are grouped dendritically by similarity.

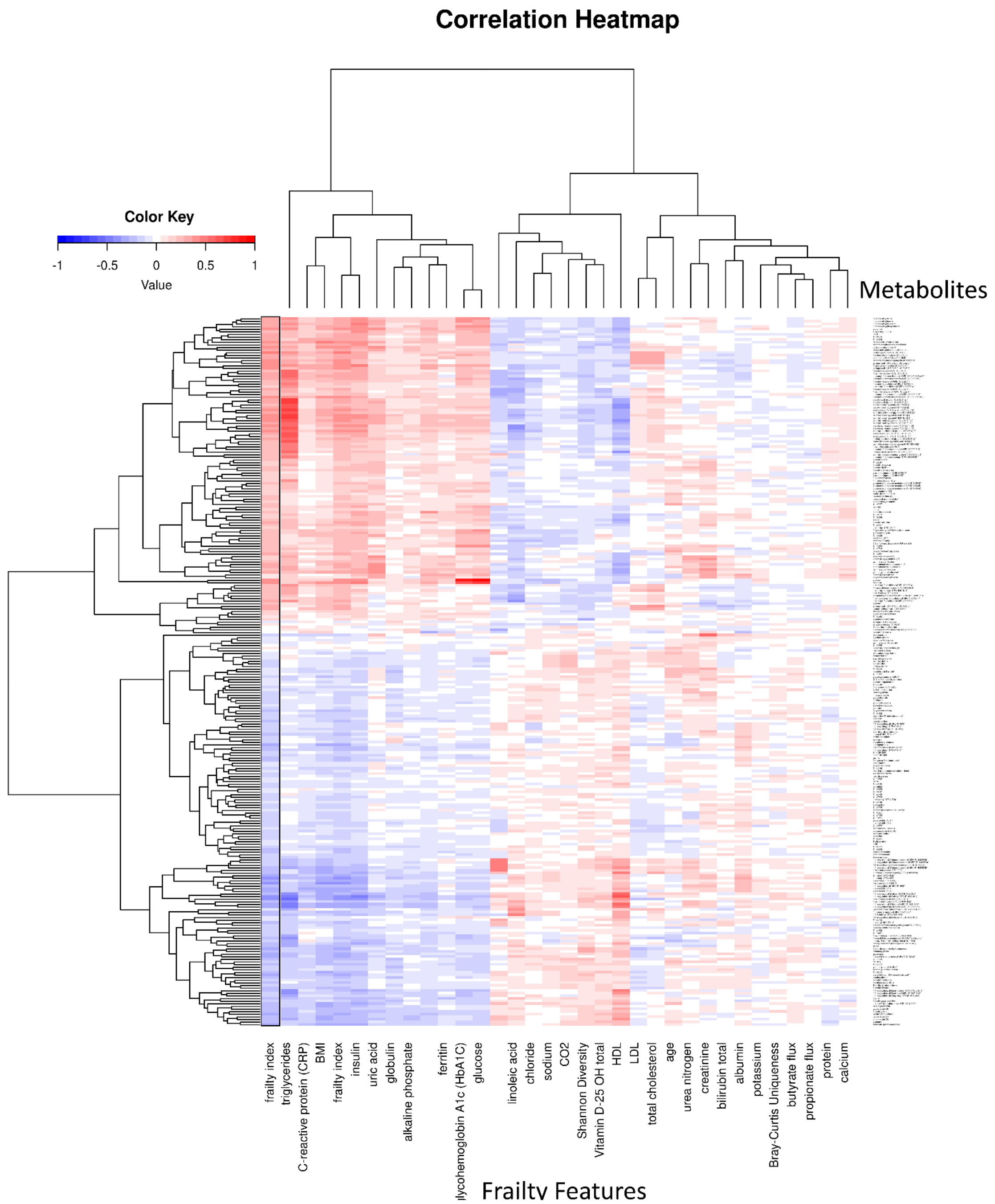


Figure 11. Correlation heatmap of significant blood plasma metabolites versus frailty measures, Bray-Curtis uniqueness, MICOM-simulated butyrate and propionate fluxes, and modeling covariates. Significant (FDR-controlled) differences in metabolites across samples were computed using LIMMA in R, and simulated butyrate and propionate flux values were determined using MICOM. Elements are grouped dendritically by similarity.

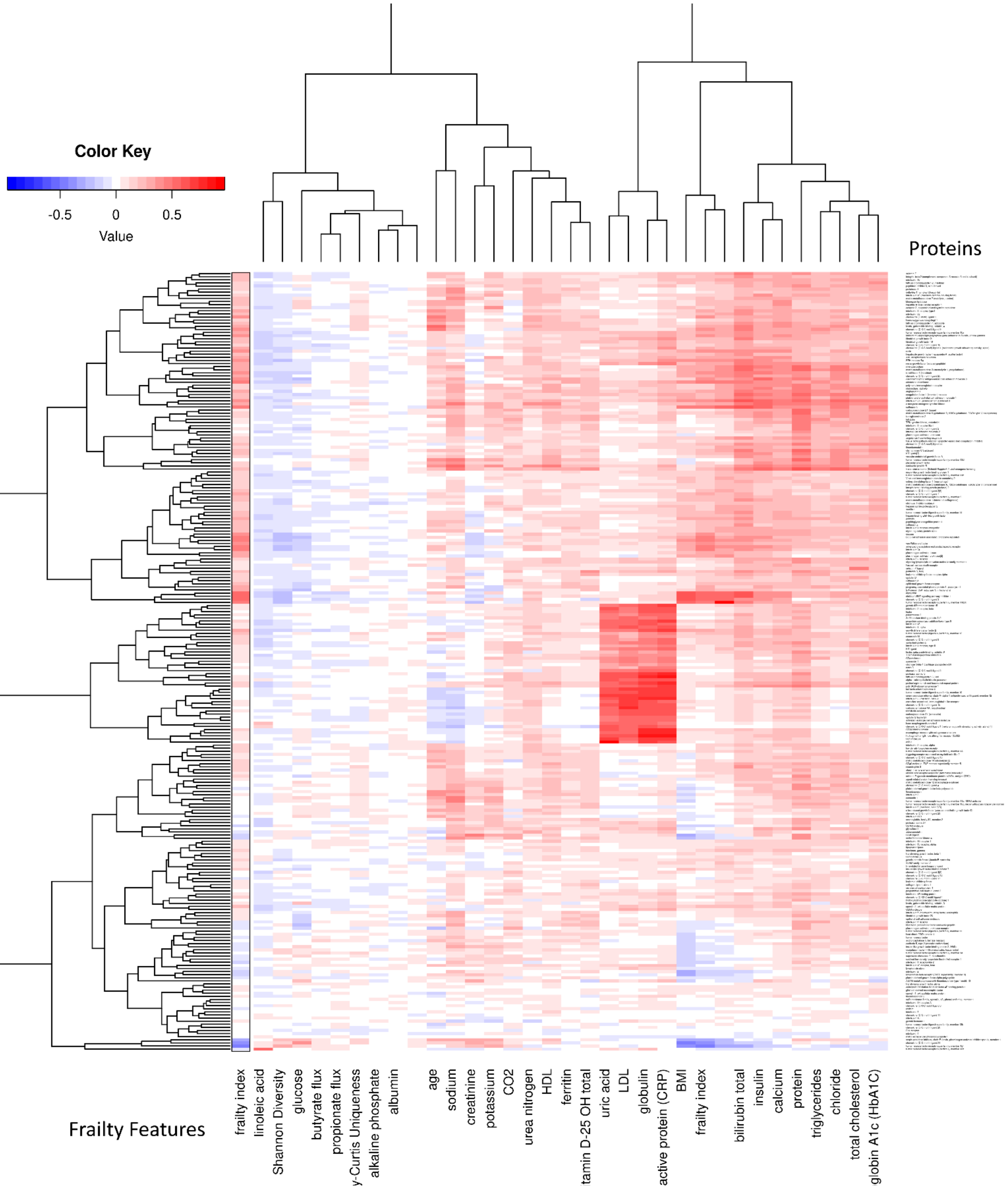


Figure 12. Correlation heatmap of significant blood plasma proteins versus frailty measures, uniqueness, MICOM-simulated butyrate and propionate fluxes, and modeling covariates. Significant (FDR-controlled) differences in protein levels across samples were computed using LIMMA in R, and simulated butyrate and propionate flux values were determined using MICOM. Elements are grouped dendritically by similarity.

2.7 METHODS

Data and code availability

- Code used to analyze 16S rRNA gene amplicon sequencing data can be found at https://github.com/gibbons-lab/mbtools.
- Code used to run the statistical analyses described in this paper is available at https://github.com/Gibbons-Lab/JPJ_Frailty_Analysis .
- Qualified researchers can access the full Arivale deidentified dataset, including all raw data, supporting the findings in this study for research purposes through signing a Data Use Agreement (DUA). Inquiries to access the data can be made at data-access@isbscience.org and will be responded to within 7 business days.

Institutional review board approval for the study

The procedures for this study were reviewed and approved by the Western Institutional Review Board, under the institutional review board study number 20170658 for the Institute for Systems Biology and 1178906 for Arivale, Inc. (see Chapter 1.11).

Frailty cohort selection

All study participants were subscribers in the Arivale Scientific Wellness program (2015-2019, see Chapter 1.7-1.8) and provided informed consent for the use of their anonymized, de-identified data for research purposes. Participants were community-dwelling, residents of Washington State and California (which are slightly leaner and healthier than other parts of the USA), over the age of 19, non-pregnant, but were not screened for the presence or absence of any particular disease. Participants yielded blood and stool samples that were used to generate blood plasma metabolomics, proteomics, clinical laboratory chemistries, and gut microbiome data (**Fig. 10-12**). 3,090 individuals with frailty data available to be computed from the Arivale program were initially considered for this analysis. Their data were merged in Python (Jupyter

Notebooks) with those with gender, age, BMI, vendor (for 16s rRNA gene amplicon sequencing data) as covariates. Finally, the clinical blood chemistries data related to calculating FIS (see below for details) were obtained as well as uniqueness and butyrate and propionate fluxes, resulting in dataframes consisting of the following: 582 genera ASVs (N = 1,204), proteomics with 274 proteins (N = 761), and metabolomics with 1,292 metabolites (N = 539) data samples to yield dataframes representative of the respective cohorts at the intersections of these merges (see the similar procedure in Chapter 1.7-1.8). Only baseline time point samples were used for each participant for the baseline ‘omics analyses. A 50% prevalence filter was implemented for the blood plasma proteomics analysis to compensate for sparsity (see Chapter 1.7-1.8), meanwhile no prevalence filter was used for any of the other cohorts. NAN values were imputed to be the median or dropped, depending on the ‘omic analysis. The final dataframes for the regressions were used for downstream regression analyses (**Fig 9-12**).

2.8 QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analyses

Using a GLM (See Chapter 1.7-1.8), each of the following features was regressed against frailty and its covariates to determine if significant linear associations were found between each covariate, frailty, and the feature: butyrate flux, propionate flux, Bray-Curtis uniqueness, Weighted-UniFrac uniqueness, and the alpha-diversity metric: Shannon diversity. For example: butyrate flux ~ FIS + gender + age + BMI + PC1 + PC2 + PC3 + PC4 (See **Fig. 8**). These were computed in RStudio using R [73] and base R package functions like “lm( )”.

Using Python (Jupyter Notebooks) [94] and R (RStudio) [73] as well as bioinformatics packages like Phyloseq and Bioconductor [74,85] and tidyverse [75], multiple linear regressions were performed to characterize these variables. The response variables were either: log-transformed during LIMMA computation or raw ASV counts (for CORNCOB) bacterial genus data, depending on the analysis (see Chapter 1.7-1.8). For the blood proteomics, plasma chemistries, and metabolite associations, generalized linear regression models were run using the LIMMA package in R [76] .

FIS

To begin characterizing the main variables in the cohorts: frailty, the covariates, and the ‘omic dataset under scrutiny, regression was performed on each multi-omic data set versus FIS in light of covariates: gender, age, BMI, and the first four principal components of genetic ancestry (see Chapter 1), PC1-PC4, for these cohorts. FIS is defined here by the list of items for FIS calculation that was originally derived from Blodgett et al. and modified here due to the absence of several items [183] . It is composed of 35 self-reported features (there are 15 items for the disease category, 6 items for the satisfaction category, 9 items for the activity category, 3 items for the medication category and 2 items for the digestion category) and 35 laboratory test features (30 clinical blood plasma chemistries and 5 blood pressure-related items) determined for each individual to score their frailty. For each participant with available data, a normalized score was derived as follows: (1) a reference range was established from the Arivale dataset, accounting for inter-vendor variability; (2) categorical variables were transformed into numerical values ranging from 0 to 1, proportional to the number of levels. For instance, a 5-level variable was assigned values of 0, 0.25, 0.5, 0.75, and 1, corresponding to the inherent order of the categories. For each disease category, a binary score (0 or 1) was assigned, indicating the absence or presence of any disease within that category. FIS was defined as the overall score divided by the number of available items. Details for the FIS calculation can be found here: https://docs.google.com/spreadsheets/d/1oMdaFsTxLUviiUIrfQ9alRHg9_BuXB4Qdrd_LvUfOHc/edit?gid=177969499#gid=177969499 .

CONCLUSION

In Chapter 1, the effects of BMF on the human body were analyzed across hundreds of relatively healthy individuals. It was observed an enrichment of pro-inflammatory taxa and uremic toxins alongside a depletion in SCFA-producers associated with low BMF (mild or worse constipation). One of the toxins, 3-IS, mediated the effects of BMF on kidney function (eGFR). Clinical laboratory chemistries indicated a stressful state on the liver and kidneys as detected in the blood plasma of the host. Additionally, in Chapter 2, the interplay between aging, frailty, and the human gut microbiota was scrutinized from the same patient population. Several inflammatory proteins (e.g. tumor necrosis factors, immunoglobulins, and interleukins) metabolites (e.g. cholesterol, cholesteryl esters, glycerophospholipids), and taxa (SCFA-producers and pro-inflammatory genera) were revealed to be significant drivers of frailty independent of age and other covariates. All these results point to a strong inflammatory phenotype of those with aberrant bowel movements and the frail, with evidence displaying multiple types of inflammation including neuroinflammation and nephrotoxicity. Finally, diet and lifestyle factors showing a significant likelihood association with BMF effects underscore the importance of dietary fruit, vegetables, fiber, and diets that promote anti-inflammation (e.g. Mediterranean diet) [184–187] .

The human gut microbiome is increasingly recognized as a key modulator of both health and disease, and it appears this holds true even in relatively healthy individuals, prior to the emergence of disease. Elucidating the intricate bidirectional relationships between the host and its resident microbiota remains a fertile area of scientific inquiry. Leveraging recent advances in systems ecology, molecular biology, bioengineering, computational sciences, and data science, researchers are employing multi-omics data integration to unravel the complex ecosystem within the human gastrointestinal tract and its impact on host phenotypes.

Historically underappreciated in clinical medicine, the gut microbiota is now emerging as a focal point for precision healthcare. The vast metabolic and immunomodulatory capacity of the gut microbiota offers a promising avenue for the development of novel diagnostics and therapeutics [188–191] . Microbiome-mediated interventions will likely play a pivotal role in combating the global rise in chronic diseases and in extending the human healthspan.

BIBLIOGRAPHY

1. Kasarello, K., Cudnoch-Jedrzejewska, A., and Czarzasta, K. (2023). Communication of gut microbiota and brain via immune and neuroendocrine signaling. Front. Microbiol. *14*, 1118529.

2. Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. Nature *464*, 59–65.

3. Ley, R.E., Peterson, D.A., and Gordon, J.I. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell *124*, 837–848.

4. Huang, Y., Xin, W., Xiong, J., Yao, M., Zhang, B., and Zhao, J. (2022). The Intestinal Microbiota and Metabolites in the Gut-Kidney-Heart Axis of Chronic Kidney Disease. Front. Pharmacol. *13*, 837500.

5. Sumida, K., and Kovesdy, C.P. (2019). The gut-kidney-heart axis in chronic kidney disease. Physiology international *106*. https://doi.org/10.1556/2060.106.2019.19.

6. Johnson-Martínez, J.P., Diener, C., Levine, A.E., Wilmanski, T., Suskind, D.L., Ralevski, A., Hadlock, J., Magis, A.T., Hood, L., Rappaport, N., et al. (2024). Aberrant bowel movement frequencies coincide with increased microbe-derived blood metabolites associated with reduced organ function. Cron. Med. *5*. https://doi.org/10.1016/j.xcrm.2024.101646.

7. Magalhães-Guedes, K.T. (2022). Psychobiotic Therapy: Method to Reinforce the Immune System. Clin. Psychopharmacol. Neurosci. *20*, 17–25.

8. Kim, J., and Lee, H.K. (2021). Potential Role of the Gut Microbiome In Colorectal Cancer Progression. Front. Immunol. *12*, 807648.

9. Hughes, R.L. (2020). A Review of the Role of the Gut Microbiome in Personalized Sports Nutrition. Preprint, https://doi.org/10.3389/fnut.2019.00191 https://doi.org/10.3389/fnut.2019.00191.

10. Ceballos, D., Hernández-Camba, A., and Ramos, L. (2021). Diet and microbiome in the beginning of the sequence of gut inflammation. World J Clin Cases *9*, 11122–11147.

11. Cryan, J.F., O'Riordan, K.J., Cowan, C.S.M., Sandhu, K.V., Bastiaanssen, T.F.S., Boehme, M., Codagnone, M.G., Cussotto, S., Fulling, C., Golubeva, A.V., et al. (2019). The Microbiota-Gut-Brain Axis. Physiol. Rev. *99*, 1877–2013.

12. Kim, D.S., Zhang, T., and Park, S. (2022). Protective effects of and water extract against memory deficits through the gut-microbiome-brain axis in an Alzheimer's disease model. Pharm. Biol. *60*, 212–224.

13. 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. (2021). Blue poo: impact of gut transit time on the gut microbiome using a novel marker. Gut *70*, 1665–1674.

14. Müller, M., Canfora, E.E., and Blaak, E.E. (2018). Gastrointestinal Transit Time, Glucose Homeostasis and Metabolic Health: Modulation by Dietary Fibers. Nutrients *10*.

https://doi.org/10.3390/nu10030275.

15. Vanuytsel, T., Tack, J., and Farre, R. (2021). The Role of Intestinal Permeability in Gastrointestinal Disorders and Current Methods of Evaluation. Front Nutr *8*, 717925.

16. Roager, H.M., Hansen, L.B.S., Bahl, M.I., Frandsen, H.L., Carvalho, V., Gøbel, R.J., Dalgaard, M.D., Plichta, D.R., Sparholt, M.H., Vestergaard, H., et al. (2016). Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. Nature Microbiology *1*, 1–9.

17. Wiskur, B., and Meerveld, B.G.-V. (2010). The Aging Colon: The Role of Enteric Neurodegeneration in Constipation. Current Gastroenterology Reports *12*, 507–512.

18. Stirpe, P., Hoffman, M., Badiali, D., and Colosimo, C. (2016). Constipation: an emerging risk factor for Parkinson's disease? European Journal of Neurology *23*, 1606–1613.

19. Wilmanski, T., Rappaport, N., Earls, J.C., Magis, A.T., Manor, O., Lovejoy, J., Omenn, G.S., Hood, L., Gibbons, S.M., and Price, N.D. (2019). Blood metabolome predicts gut microbiome α-diversity in humans. Preprint, https://doi.org/10.1038/s41587-019-0233-9 https://doi.org/10.1038/s41587-019-0233-9.

20. Duvallet, C., Gibbons, S.M., Gurry, T., Irizarry, R.A., and Alm, E.J. (2017). Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. Nat. Commun. *8*, 1784.

21. McDonald, D., Hyde, E., Debelius, J.W., Morton, J.T., Gonzalez, A., Ackermann, G., Aksenov, A.A., Behsaz, B., Brennan, C., Chen, Y., et al. (2018). American Gut: an Open Platform for Citizen Science Microbiome Research. mSystems *3*. https://doi.org/10.1128/mSystems.00031-18.

22. Tomkovich, S., Taylor, A., King, J., Colovas, J., Bishop, L., McBride, K., Royzenblat, S., Lesniak, N.A., Bergin, I.L., and Schloss, P.D. (2021). An Osmotic Laxative Renders Mice Susceptible to Prolonged Clostridioides difficile Colonization and Hinders Clearance. mSphere *6*. https://doi.org/10.1128/mSphere.00629-21.

23. Falony, G., Joossens, M., Vieira-Silva, S., Wang, J., Darzi, Y., Faust, K., Kurilshikov, A., Bonder, M.J., Valles-Colomer, M., Vandeputte, D., et al. (2016). Population-level analysis of gut microbiome variation. Science *352*, 560–564.

24. Vandeputte, D., Falony, G., Vieira-Silva, S., Tito, R.Y., Joossens, M., and Raes, J. (2016). Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. Preprint, https://doi.org/10.1136/gutjnl-2015-309618 https://doi.org/10.1136/gutjnl-2015-309618.

25. Adams-Carr, K.L., Bestwick, J.P., Shribman, S., Lees, A., Schrag, A., and Noyce, A.J. (2016). Constipation preceding Parkinson's disease: a systematic review and meta-analysis. J. Neurol. Neurosurg. Psychiatry *87*, 710–716.

26. Ramos, C.I., Nerbass, F.B., and Cuppari, L. (2022). Constipation in Chronic Kidney Disease: It Is Time to Bridge the Gap. Preprint, https://doi.org/10.3390/kidneydial2020023 https://doi.org/10.3390/kidneydial2020023.

27. Sumida, K., Yamagata, K., and Kovesdy, C.P. (2020). Constipation in CKD. Kidney Int Rep *5*, 121–134.

28. Ikee, R., Sasaki, N., Yasuda, T., and Fukazawa, S. (2020). Chronic Kidney Disease, Gut Dysbiosis, and Constipation: A Burdensome Triplet. Preprint, https://doi.org/10.3390/microorganisms8121862.

29. Gr, B.A.P.J. (2013). American Gastroenterological Association technical review on constipation. Gastroenterology *144*, 218–238.

30. Sumida, K., Molnar, M.Z., Potukuchi, P.K., Thomas, F., Lu, J.L., Matsushita, K., Yamagata, K., Kalantar-Zadeh, K., and Kovesdy, C.P. (2017). Constipation and Incident CKD. Preprint, https://doi.org/10.1681/asn.2016060656 https://doi.org/10.1681/asn.2016060656.

31. Brocca, A., Virzì, G.M., de Cal, M., Cantaluppi, V., and Ronco, C. (2013). Cytotoxic effects of p-cresol in renal epithelial tubular cells. Blood Purif. *36*, 219–225.

32. Poesen, R., Claes, K., Evenepoel, P., de Loor, H., Augustijns, P., Kuypers, D., and Meijers, B. (2016). Microbiota-Derived Phenylacetylglutamine Associates with Overall Mortality and Cardiovascular Disease in Patients with CKD. J. Am. Soc. Nephrol. *27*, 3479–3487.

33. Barrios, C., Beaumont, M., Pallister, T., Villar, J., Goodrich, J.K., Clark, A., Pascual, J., Ley, R.E., Spector, T.D., Bell, J.T., et al. (2015). Gut-Microbiota-Metabolite Axis in Early Renal Function Decline. PLoS One *10*, e0134311.

34. Ikee, R., Yano, K., and Tsuru, T. (2019). Constipation in chronic kidney disease: it is time to reconsider. Preprint, https://doi.org/10.1186/s41100-019-0246-3 https://doi.org/10.1186/s41100-019-0246-3.

35. Venables, W. N., & Ripley, B. D. (2024). MASS (version 7.3-6): Support Functions and Datasets for Venables and Ripley's MASS. Available at https://www.rdocumentation.org/packages/MASS/versions/7.3-6

36. Martin, B.D., Witten, D., and Willis, A.D. (2020). MODELING MICROBIAL ABUNDANCES AND DYSBIOSIS WITH BETA-BINOMIAL REGRESSION. Ann. Appl. Stat. *14*, 94–115.

37. Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2013). mediation: Package for Causal Mediation Analysis. R package version 4.4.2. Available at CRAN.R-project.org/package=mediation.

38. Wan, M., King, L., Baugh, N., Arslan, Z., Snauwaert, E., Paglialonga, F., and Shroff, R. (2023). Gutted: constipation in children with chronic kidney disease and on dialysis. Pediatr. Nephrol. https://doi.org/10.1007/s00467-022-05849-y.

39. Kim, J.E., Park, J.J., Lee, M.R., Choi, J.Y., Song, B.R., Park, J.W., Kang, M.J., Son, H.J., Hong, J.T., and Hwang, D.Y. (2019). Constipation in Tg2576 mice model for Alzheimer's disease associated with dysregulation of mechanism involving the mAChR signaling pathway and ER stress response. PLoS One *14*, e0215205.

40. Baldini, F., Hertel, J., Sandt, E., Thinnes, C.C., Neuberger-Castillo, L., Pavelka, L., Betsou, F., Krüger, R., Thiele, I., and NCER-PD Consortium (2020). Parkinson's disease-associated alterations of the gut microbiome predict disease-relevant changes in metabolic functions.

BMC Biol. *18*, 62.

41. Neugarten, J., and Golestaneh, L. (2013). Gender and the Prevalence and Progression of Renal Disease. Preprint, https://doi.org/10.1053/j.ackd.2013.05.004 https://doi.org/10.1053/j.ackd.2013.05.004.

42. Werth, B.L., and Christopher, S.-A. (2021). Potential risk factors for constipation in the community. World J. Gastroenterol. *27*, 2795–2817.

43. Chen, H.B., Huang, Y., Song, H.W., Li, X.L., He, S., Xie, J.T., Huang, C., Zhang, S.J., Liu, J., and Zou, Y. (2010). Clinical Research on the Relation Between Body Mass Index, Motilin and Slow Transit Constipation. Gastroenterol. Res. Pract. *3*, 19–24.

44. Vermorken, A.J.M., Andrès, E., and Cui, Y. (2016). Bowel movement frequency, oxidative stress and disease prevention. Mol Clin Oncol *5*, 339–342.

45. Vermorken, A.J.M., Cui, Y., Kleerebezem, R., and Andrès, E. (2016). Bowel movement frequency and cardiovascular mortality, a matter of fibers and oxidative stress? Atherosclerosis *253*, 278–280.

46. Jankipersadsing, S.A., Hadizadeh, F., Bonder, M.J., Tigchelaar, E.F., Deelen, P., Fu, J., Andreasson, A., Agreus, L., Walter, S., Wijmenga, C., et al. (2017). A GWAS meta-analysis suggests roles for xenobiotic metabolism and ion channel activity in the biology of stool frequency. Gut *66*, 756–758.

47. Romano, S., Savva, G.M., Bedarf, J.R., Charles, I.G., Hildebrand, F., and Narbad, A. (2021). Meta-analysis of the Parkinson's disease gut microbiome suggests alterations linked to intestinal inflammation. NPJ Parkinson's Disease *7*. https://doi.org/10.1038/s41531-021-00156-z.

48. Singh, S.B., Carroll-Portillo, A., and Lin, H.C. (2023). Desulfovibrio in the Gut: The Enemy within? Microorganisms *11*. https://doi.org/10.3390/microorganisms11071772.

49. Boutard, M., Cerisy, T., Nogue, P.-Y., Alberti, A., Weissenbach, J., Salanoubat, M., and Tolonen, A.C. (2014). Functional Diversity of Carbohydrate-Active Enzymes Enabling a Bacterium to Ferment Plant Biomass. PLoS Genet. *10*. https://doi.org/10.1371/journal.pgen.1004773.

50. Song, Y., Chen, K., Lv, L., Xiang, Y., Du, X., Zhang, X., Zhao, G., and Xiao, Y. (2022). Uncovering the biogeography of the microbial commmunity and its association with nutrient metabolism in the intestinal tract using a pig model. Front. Nutr. *9*, 1003763.

51. van den Bogert, B., Meijerink, M., Zoetendal, E.G., Wells, J.M., and Kleerebezem, M. (2014). Immunomodulatory Properties of Streptococcus and Veillonella Isolates from the Human Small Intestine Microbiota. PLoS One *9*, e114277.

52. Silva, Y.P., Bernardi, A., and Frozza, R.L. (2020). The Role of Short-Chain Fatty Acids From Gut Microbiota in Gut-Brain Communication. Front. Endocrinol. *11*, 508738.

53. Martin-Gallausiaux, C., Marinelli, L., Blottière, H.M., Larraufie, P., and Lapaque, N. (2021). SCFA: mechanisms and functional importance in the gut. Proc. Nutr. Soc. *80*, 37–49.

54. Zhao, Y., Liu, Q., Hou, Y., and Zhao, Y. (2022). Alleviating effects of gut micro-ecologically

regulatory treatments on mice with constipation. Front. Microbiol. *13*, 956438.

55. Vicentini, F.A., Keenan, C.M., Wallace, L.E., Woods, C., Cavin, J.-B., Flockton, A.R., Macklin, W.B., Belkind-Gerson, J., Hirota, S.A., and Sharkey, K.A. (2021). Intestinal microbiota shapes gut physiology and regulates enteric neurons and glia. Microbiome *9*, 210.

56. Chen, Y.-Y., Chen, D.-Q., Chen, L., Liu, J.-R., Vaziri, N.D., Guo, Y., and Zhao, Y.-Y. (2019). Microbiome–metabolome reveals the contribution of gut–kidney axis on kidney disease. J. Transl. Med. *17*, 1–11.

57. Hobson, S., Qureshi, A.R., Ripswedan, J., Wennberg, L., de Loor, H., Ebert, T., Söderberg, M., Evenepoel, P., Stenvinkel, P., and Kublickiene, K. (2023). Phenylacetylglutamine and trimethylamine N-oxide: Two uremic players, different actions. Eur. J. Clin. Invest. *53*, e14074.

58. Sun, C.-Y., Li, J.-R., Wang, Y.-Y., Lin, S.-Y., Ou, Y.-C., Lin, C.-J., Wang, J.-D., Liao, S.-L., and Chen, C.-J. (2020). p-Cresol Sulfate Caused Behavior Disorders and Neurodegeneration in Mice with Unilateral Nephrectomy Involving Oxidative Stress and Neuroinflammation. Int. J. Mol. Sci. *21*. https://doi.org/10.3390/ijms21186687.

59. Levi, I., Gurevich, M., Perlman, G., Magalashvili, D., Menascu, S., Bar, N., Godneva, A., Zahavi, L., Chermon, D., Kosower, N., et al. (2021). Potential role of indolelactate and butyrate in multiple sclerosis revealed by integrated microbiome-metabolome analysis. Cell Rep Med *2*, 100246.

60. Sun, C.-Y., Hsu, H.-H., and Wu, M.-S. (2012). p-Cresol sulfate and indoxyl sulfate induce similar cellular inflammatory gene expressions in cultured proximal renal tubular cells. Nephrol. Dial. Transplant *28*, 70–78.

61. Barreto, F.C., Barreto, D.V., Liabeuf, S., Meert, N., Glorieux, G., Temmar, M., Choukroun, G., Vanholder, R., Massy, Z.A., and European Uremic Toxin Work Group (EUTox) (2009). Serum indoxyl sulfate is associated with vascular disease and mortality in chronic kidney disease patients. Clin. J. Am. Soc. Nephrol. *4*, 1551–1558.

62. Tsai, M.-T., and Tarng, D.-C. (2018). Beyond a Measure of Liver Function—Bilirubin Acts as a Potential Cardiovascular Protector in Chronic Kidney Disease Patients. Int. J. Mol. Sci. *20*, 117.

63. Mayo Clinic. (2024). Creatinine test. https://www.mayoclinic.org/tests-procedures/creatinine-test/about/pac-20384646.

64. National Kidney Foundation. (2023). Creatinine. Available at https://www.kidney.org/atoz/content/serum-blood-creatinine.

65. - HealthMatters.io. (2024). Linoleic Acid - Lab Results explained. Available at https://healthmatters.io/understand-blood-test-results/linoleic-acid.

66. Sherrell, Z., and Labedzki, M. (2022). What are normal MCH levels? The Checkup. https://www.singlecare.com/blog/mch-blood-test/.

67. Hosseinzadeh, S.T., Poorsaadati, S., Radkani, B., and Forootan, M. (2011). Psychological

disorders in patients with chronic constipation. Gastroenterol Hepatol Bed Bench *4*, 159–163.

68. Blanc, F., Bouteloup, V., Paquet, C., Chupin, M., Pasquier, F., Gabelle, A., Ceccaldi, M., de Sousa, P.L., Krolak-Salmon, P., David, R., et al. (2022). Prodromal characteristics of dementia with Lewy bodies: baseline results of the MEMENTO memory clinics nationwide cohort. Alzheimers. Res. Ther. *14*, 96.

69. Nedelec, T., Couvy-Duchesne, B., Monnet, F., Daly, T., Ansart, M., Gantzer, L., Lekens, B., Epelbaum, S., Dufouil, C., and Durrleman, S. (2022). Identifying health conditions associated with Alzheimer's disease up to 15 years before diagnosis: an agnostic study of French and British health records. Lancet Digit Health *4*, e169–e178.

70. Kuang, R., and Binion, D.G. (2022). Should high-fiber diets be recommended for patients with inflammatory bowel disease? Curr. Opin. Gastroenterol. *38*, 168–172.

71. Eating right for chronic kidney disease (2022). National Institute of Diabetes and Digestive and Kidney Diseases. https://www.niddk.nih.gov/health-information/kidney-disease/chronic-kidney-disease-ckd/eating-nutrition.

72. O'Rourke, H.P., and MacKinnon, D.P. (2018). Reasons for Testing Mediation in the Absence of an Intervention Effect: A Research Imperative in Prevention and Intervention Research. J. Stud. Alcohol Drugs *79*, 171.

73. VWR. (2024). Mo Bio Laboratories PowerMag™ Soil DNA Isolation Kit, epMotion®. Available at https://us.vwr.com/store/product/12783066/mo-bio-laboratories-powermagtm-soil-dna-isolation-kit-epmotion.

74. Bioconductor. (2024). Install Bioconductor. Available at http://bioconductor.org.

75. Tidyverse. (2024). Tidyverse packages. Available at https://www.tidyverse.org/packages/.

76. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47–e47.

77. Hansen, K.D. (2024). limma. Available at https://kasperdanielhansen.github.io/genbioconductor/html/limma.html.

78. Cleveland Clinic. (2024). Frequent Bowel Movements: Causes, Symptoms & Treatment. Available at https://my.clevelandclinic.org/health/symptoms/17791-frequent-bowel-movements.

79. Mo Bio Laboratories PowerMag^{TM} Soil DNA Isolation Kit, epMotion® VWR. https://us.vwr.com/store/product/12783066/mo-bio-laboratories-powermagtm-soil-dna-isolation-kit-epmotion.

80. Illumina. (2024). MiSeq System. Available at https://www.illumina.com/systems/sequencing-platforms.html.

81. Zubair, N., Conomos, M.P., Hood, L., Omenn, G.S., Price, N.D., Spring, B.J., Magis, A.T.,

and Lovejoy, J.C. (2019). Genetic Predisposition Impacts Clinical Changes in a Lifestyle Coaching Program. Sci. Rep. *9*. https://doi.org/10.1038/s41598-019-43058-0.

82. Manor, O., Zubair, N., Conomos, M.P., Xu, X., Rohwer, J.E., Krafft, C.E., Lovejoy, J.C., and Magis, A.T. (2018). A Multi-omic Association Study of Trimethylamine N-Oxide. Cell Rep. *24*, 935–946.

83. Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods *13*, 581–583.

84. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. *41*, D590–D596.

85. McMurdie, P.J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One *8*, e61217.

86. Precision for Medicine. (2024). Olink Proteomics. Available at https://www.precisionformedicine.com/specialty-lab-services/cytokine-and-protein-analysis/olink-proteomics.

87. Illumina. (2024). Olink Proteomics Tech Note. Available at https://www.illumina.com/content/dam/illumina/gcs/assembled-assets/marketing-literature/olink-proteomics-tech-note-m-us-00196/olink-proteomics-tech-note-m-us-00196.pdf.

88. Olink. (2024). Assay validation. Available at https://olink.com/our-platform/assay-validation/.

89. Wik, L., Nordberg, N., Broberg, J., Björkesten, J., Assarsson, E., Henriksson, S., Grundberg, I., Pettersson, E., Westerberg, C., Liljeroth, E., et al. (2021). Proximity Extension Assay in Combination with Next-Generation Sequencing for High-throughput Proteome-wide Analysis. Mol. Cell. Proteomics *20*, 100168.

90. Metabolon – Small Molecules, Big Insights (2022). Metabolon. https://www.metabolon.com/.

91. Diener, C., Dai, C.L., Wilmanski, T., Baloni, P., Smith, B., Rappaport, N., Hood, L., Magis, A.T., and Gibbons, S.M. (2022). Genome-microbiome interplay provides insight into the determinants of the human blood metabolome. Nat Metab *4*, 1560–1572.

92. Quest Diagnostics. (2024). Home. Available at https://www.questdiagnostics.com/.

93. Delgado, C., Baweja, M., Crews, D.C., Eneanya, N.D., Gadegbeku, C.A., Inker, L.A., Mendu, M.L., Miller, W.G., Moxey-Mims, M.M., Roberts, G.V., et al. (2022). A Unifying Approach for GFR Estimation: Recommendations of the NKF-ASN Task Force on Reassessing the Inclusion of Race in Diagnosing Kidney Disease. Am. J. Kidney Dis. *79*, 268–288.e1.

94. Python Software Foundation. (2024). Python. Available at https://www.python.org/downloads/.

95. Dattani, S., Rodés-Guirao, L., Ritchie, H., Ortiz-Ospina, E., and Roser, M. (2023). Life Expectancy. Our World in Data.

96. Li, X., Li, C., Zhang, W., Wang, Y., Qian, P., and Huang, H. (2023). Inflammation and aging: signaling pathways and intervention therapies. Signal Transduction and Targeted Therapy *8*, 1–29.

97. Rossiello, F., Jurk, D., Passos, J.F., and d'Adda di Fagagna, F. (2022). Telomere dysfunction in ageing and age-related diseases. Nat. Cell Biol. *24*, 135–147.

98. Song, J., Xiao, L., Zhang, Z., Wang, Y., Kouis, P., Rasmussen, L.J., and Dai, F. (2024). Effects of reactive oxygen species and mitochondrial dysfunction on reproductive aging. Front. Cell Dev. Biol. *12*, 1347286.

99. Tenchov, R., Sasso, J.M., Wang, X., and Zhou, Q.A. (2023). Aging Hallmarks and Progression and Age-Related Diseases: A Landscape View of Research Advancement. ACS Chem. Neurosci. https://doi.org/10.1021/acschemneuro.3c00531.

100. Ageing and health https://www.who.int/news-room/fact-sheets/detail/ageing-and-health#:~:text=Ageing%20explained,and%20cellular%20damage%20over%20time.

101. Kaushik, S., and Cuervo, A.M. (2015). Proteostasis and aging. Nat. Med. *21*, 1406–1415.

102. Seegren, P.V., Harper, L.R., Downs, T.K., Zhao, X.-Y., Viswanathan, S.B., Stremska, M.E., Olson, R.J., Kennedy, J., Ewald, S.E., Kumar, P., et al. (2023). Reduced mitochondrial calcium uptake in macrophages is a major driver of inflammaging. Nature Aging *3*, 796–812.

103. Michaud, M., Balardy, L., Moulis, G., Gaudin, C., Peyrot, C., Vellas, B., Cesari, M., and Nourhashemi, F. (2013). Proinflammatory cytokines, aging, and age-related diseases. J. Am. Med. Dir. Assoc. *14*. https://doi.org/10.1016/j.jamda.2013.05.009.

104. Ershler, W.B., and Keller, E.T. (2000). Age-Associated Increased Interleukin-6 Gene Expression, Late-Life Diseases, and Frailty. Annu. Rev. Med. *51*, 245–270.

105. Adaptive mechanisms to oxidative stress during aging (2006). Mech. Ageing Dev. *127*, 436–443.

106. Franceschi, C., and Campisi, J. (2014). Chronic Inflammation (Inflammaging) and Its Potential Contribution to Age-Associated Diseases. J. Gerontol. A Biol. Sci. Med. Sci. *69*, S4–S9.

107. Khansari, N., Shakiba, Y., and Mahmoudi, M. Chronic Inflammation and Oxidative Stress as a Major Cause of Age- Related Diseases and Cancer. Recent Pat. Inflamm. Allergy Drug Discov. *3*, 73–80.

108. Yoshimura, A. (2024). Fibrosis: from mechanisms to novel treatments. Inflamm. Regen. *44*, 1–2.

109. Pothier, K., Gana, W., Bailly, N., and Fougère, B. (2022). Associations Between Frailty and Inflammation, Physical, and Psycho-Social Health in Older Adults: A Systematic Review. Front. Psychol. *13*. https://doi.org/10.3389/fpsyg.2022.805501.

110. Nagpal, R., Mainali, R., Ahmadi, S., Wang, S., Singh, R., Kavanagh, K., Kitzman, D.W., Kushugulova, A., Marotta, F., and Yadav, H. (2018). Gut microbiome and aging:

Physiological and mechanistic insights. Nutrition and Healthy Aging *4*, 267.

111. Kim, S., and Jazwinski, S.M. (2018). The Gut Microbiota and Healthy Aging: A Mini-Review. Gerontology *64*. https://doi.org/10.1159/000490615.

112. Salazar, J., Durán, P., Díaz, M.P., Chacín, M., Santeliz, R., Mengual, E., Gutiérrez, E., León, X., Díaz, A., Bernal, M., et al. (2023). Exploring the Relationship between the Gut Microbiota and Ageing: A Possible Age Modulator. Int. J. Environ. Res. Public Health *20*. https://doi.org/10.3390/ijerph20105845.

113. Ragonnaud, E., and Biragyn, A. (2021). Gut microbiota as the key controllers of “healthy” aging of elderly people. Immun. Ageing *18*, 1–11.

114. 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., and Nelson, K.E. (2006). Metagenomic Analysis of the Human Distal Gut Microbiome. Science *312*, 1355.

115. Honda, K., and Littman, D.R. (2016). The microbiota in adaptive immune homeostasis and disease. Nature *535*, 75–84.

116. Wilmanski, T., Diener, C., Rappaport, N., Patwardhan, S., Wiedrick, J., Lapidus, J., Earls, J.C., Zimmer, A., Glusman, G., Robinson, M., et al. (2021). Gut microbiome pattern reflects healthy ageing and predicts survival in humans. Nature Metabolism *3*, 274–286.

117. Diener, C., Gibbons, S.M., and Resendis-Antonio, O. (2020). MICOM: Metagenome-Scale Modeling To Infer Metabolic Interactions in the Gut Microbiota. mSystems. https://doi.org/10.1128/msystems.00606-19.

118. Quinn-Bohmann, N., Wilmanski, T., Sarmiento, K.R., Levy, L., Lampe, J.W., Gurry, T., Rappaport, N., Ostrem, E.M., Venturelli, O.S., Diener, C., et al. (2024). Microbial community-scale metabolic modelling predicts personalized short-chain fatty acid production profiles in the human gut. Nat Microbiol *9*, 1700–1712.

119. Ellermann, M., and Balfour Sartor, R. (2018). Intestinal bacterial biofilms modulate mucosal immune responses. Journal of Immunological Sciences *2*.

120. Datorre, J.G., de Carvalho, A.C., Guimarães, D.P., and Reis, R.M. (2020). The Role of Fusobacterium nucleatum in Colorectal Carcinogenesis. Pathobiology *88*, 127–140.

121. Martin, R.M., and Bachman, M.A. (2018). Colonization, Infection, and the Accessory Genome of. Front. Cell. Infect. Microbiol. *8*, 4.

122. Natividad, J.M., Lamas, B., Pham, H.P., Michel, M.-L., Rainteau, D., Bridonneau, C., da Costa, G., van Hylckama Vlieg, J., Sovran, B., Chamignon, C., et al. (2018). Bilophila wadsworthia aggravates high fat diet induced metabolic dysfunctions in mice. Nat. Commun. *9*, 1–15.

123. Haiser, H.J., Gootenberg, D.B., Chatman, K., Sirasani, G., Balskus, E.P., and Turnbaugh, P.J. (2013). Predicting and Manipulating Cardiac Drug Inactivation by the Human Gut Bacterium Eggerthella lenta. Science. https://doi.org/10.1126/science.1235872.

124. Dong, X., Guthrie, B.G.H., Alexander, M., Noecker, C., Ramirez, L., Glasser, N.R., Turnbaugh, P.J., and Balskus, E.P. (2022). Genetic manipulation of the human gut

bacterium Eggerthella lenta reveals a widespread family of transcriptional regulators. Nat. Commun. *13*, 1–20.

125. Huang, Z., Wang, C., Huang, Q., Yan, Z., and Yin, Z. (2023). Hungatella hathewayi impairs the sensitivity of colorectal cancer cells to 5-FU through decreasing CDX2 expression. Hum. Cell *36*. https://doi.org/10.1007/s13577-023-00938-y.

126. Louis, P., and Flint, H.J. (2017). Formation of propionate and butyrate by the human colonic microbiota. Environ. Microbiol. *19*. https://doi.org/10.1111/1462-2920.13589.

127. Vital, M., Howe, A.C., and Tiedje, J.M. (2014). Revealing the Bacterial Butyrate Synthesis Pathways by Analyzing (Meta)genomic Data. MBio. https://doi.org/10.1128/mbio.00889-14.

128. Pu, Y., Sun, Z., Zhang, H., Huang, Q., Wang, Z., Mei, Z., Wang, P., Kong, M., Yang, W., Lin, C., et al. (2024). Gut microbial features and circulating metabolomic signatures of frailty in older adults. Nat Aging. https://doi.org/10.1038/s43587-024-00678-0.

129. Ghosh, T.S., Shanahan, F., and O'Toole, P.W. (2022). The gut microbiome as a modulator of healthy ageing. Nat. Rev. Gastroenterol. Hepatol. *19*, 565–584.

130. Evenepoel, P., Meijers, B.K., Bammens, B.R., and Verbeke, K. (2009). Uremic toxins originating from colonic microbial metabolism. Kidney Int. Suppl. https://doi.org/10.1038/ki.2009.402.

131. Sheldon, A.L., and Robinson, M.B. (2007). The Role of Glutamate Transporters in Neurodegenerative Diseases and Potential Opportunities for Intervention. Neurochem. Int. *51*, 333.

132. Serum metabolomics reveals γ-glutamyl dipeptides as biomarkers for discrimination among different forms of liver disease (2011). J. Hepatol. *55*, 896–905.

133. Zhang, H., Kovacs-Nolan, J., Kodera, T., Eto, Y., and Mine, Y. (2015). γ-Glutamyl cysteine and γ-glutamyl valine inhibit TNF-α signaling in intestinal epithelial cells and reduce inflammation in a mouse model of colitis via allosteric activation of the calcium-sensing receptor. Biochim. Biophys. Acta *1852*. https://doi.org/10.1016/j.bbadis.2014.12.023.

134. Salama, S.A., Arab, H.H., Hassan, M.H., Al Robaian, M.M., and Maghrabi, I.A. (2019). Cadmium-induced hepatocellular injury: Modulatory effects of γ-glutamyl cysteine on the biomarkers of inflammation, DNA damage, and apoptotic cell death. J. Trace Elem. Med. Biol. *52*. https://doi.org/10.1016/j.jtemb.2018.12.003.

135. Geng, P., Qin, W., and Xu, G. (2021). Proline metabolism in cancer. Amino Acids *53*. https://doi.org/10.1007/s00726-021-03060-1.

136. Field, B.C., Gordillo, R., and Scherer, P.E. (2020). The Role of Ceramides in Diabetes and Cardiovascular Disease Regulation of Ceramides by Adipokines. Front. Endocrinol. *11*. https://doi.org/10.3389/fendo.2020.569250.

137. Xia, Z., Zhou, X., Li, J., Li, L., Ma, Y., Wu, Y., Huang, Z., Li, X., Xu, P., and Xue, M. (2019). Multiple-Omics Techniques Reveal the Role of Glycerophospholipid Metabolic Pathway in the Response of Saccharomyces cerevisiae Against Hypoxic Stress. Front. Microbiol. *10*, 457285.

138. Akyol, S., Ugur, Z., Yilmaz, A., Ustun, I., Gorti, S.K.K., Oh, K., McGuinness, B., Passmore, P., Kehoe, P.G., Maddens, M.E., et al. (2021). Lipid Profiling of Alzheimer's Disease Brain Highlights Enrichment in Glycerol(phospho)lipid, and Sphingolipid Metabolism. Cells *10*, 2591.

139. Cai, T., and Yang, F. (2018). Phospholipid and Phospholipidomics in Health and Diseases. Lipidomics in Health & Disease, 177–202.

140. Qian, X., Hai, W., Chen, S., Zhang, M., Jiang, X., and Tang, H. (2023). Multi-omics data reveals aberrant gut microbiota-host glycerophospholipid metabolism in association with neuroinflammation in APP/PS1 mice. Gut Microbes. https://doi.org/10.1080/19490976.2023.2282790.

141. Owei, I., Umekwe, N., Stentz, F., Wan, J., and Dagogo-Jack, S. (2021). Association of plasma acylcarnitines with insulin sensitivity, insulin secretion, and prediabetes in a biracial cohort. Exp. Biol. Med. *246*, 1698.

142. Epinephrine (Adrenaline) Cleveland Clinic. https://my.clevelandclinic.org/health/articles/22611-epinephrine-adrenaline.

143. Magistretti, P., and Allaman, I. (2018). Lactate in the brain: from metabolic end-product to signalling molecule. Nat. Rev. Neurosci. https://doi.org/10.1038/nrn.2018.19.

144. Wu, A., Lee, D., and Xiong, W.-C. (2023). Lactate Metabolism, Signaling, and Function in Brain Development, Synaptic Plasticity, Angiogenesis, and Neurodegenerative Diseases. Int. J. Mol. Sci. *24*. https://doi.org/10.3390/ijms241713398.

145. Shi, Z., Chen, G., Cao, Z., Wu, F., Lei, H., Chen, C., Song, Y., Liu, C., Li, J., Zhou, J., et al. (2021). Gut Microbiota and Its Metabolite Deoxycholic Acid Contribute to Sucralose Consumption-Induced Nonalcoholic Fatty Liver Disease. J. Agric. Food Chem. https://doi.org/10.1021/acs.jafc.0c07467.

146. Zhang, L., Zheng, Z., Huang, H., Fu, Y., Chen, T., Liu, C., Yi, Q., Lin, C., Zeng, Y., Ou, Q., et al. (2024). Multi-omics reveals deoxycholic acid modulates bile acid metabolism via the gut microbiota to antagonize carbon tetrachloride-induced chronic liver injury. Gut Microbes *16*. https://doi.org/10.1080/19490976.2024.2323236.

147. Calder, P.C. (2017). Omega-3 fatty acids and inflammatory processes: from molecules to man. Biochem. Soc. Trans. *45*. https://doi.org/10.1042/BST20160474.

148. Schmitz, G., and Ecker, J. (2008). The opposing effects of n-3 and n-6 fatty acids. Prog. Lipid Res. *47*. https://doi.org/10.1016/j.plipres.2007.12.004.

149. Xia, Y., Zhang, Y., Li, Y., Li, X., Wu, Y., and Yao, Q. (2024). Omega-3 polyunsaturated fatty acids play a protective role in a mouse model of Parkinson's disease by increasing intestinal inducible Treg cells. Cell. Mol. Biol. *70*, 107–112.

150. Cheah, I.K., and Halliwell, B. (2012). Ergothioneine; antioxidant potential, physiological function and role in disease. Biochim. Biophys. Acta *1822*. https://doi.org/10.1016/j.bbadis.2011.09.017.

151. Fu, T.-T., and Shen, L. (2022). Ergothioneine as a Natural Antioxidant Against Oxidative

Stress-Related Diseases. Front. Pharmacol. *13*, 850813.

152. Marcinkiewicz, J., and Kontny, E. (2014). Taurine and inflammatory diseases. Amino Acids *46*. https://doi.org/10.1007/s00726-012-1361-4.

153. Schaffer, S.W., Jong, C.J., Ramila, K.C., and Azuma, J. (2010). Physiological roles of taurine in heart and muscle. J. Biomed. Sci. *17 Suppl 1*, S2.

154. Chen, S., Liu, Y., Wang, X., Wang, H., Li, S., Shi, H., Zhu, H., Zhang, J., Pi, D., Hu, C.A., et al. (2016). Asparagine improves intestinal integrity, inhibits TLR4 and NOD signaling, and differently regulates p38 and ERK1/2 signaling in weanling piglets after LPS challenge. Innate Immun. *22*. https://doi.org/10.1177/1753425916664124.

155. Ma, C., Tu, D., Gu, J., Xu, Q., Hou, P., Wu, H., Guo, Z., Bai, Y., Zhao, X., and Li, P. (2021). The Predictive Value of Myoglobin for COVID-19-Related Adverse Outcomes: A Systematic Review and Meta-Analysis. Frontiers in Cardiovascular Medicine *8*. https://doi.org/10.3389/fcvm.2021.757799.

156. Marullo, A.L., and O'Halloran, K.D. (2023). Microbes, metabolites and muscle: Is the gut–muscle axis a plausible therapeutic target in Duchenne muscular dystrophy? Exp. Physiol. *108*, 1132.

157. Chew, W., Lim, Y.P., Lim, W.S., Chambers, E.S., Frost, G., Wong, S.H., and Ali, Y. (2022). Gut-muscle crosstalk. A perspective on influence of microbes on muscle function. Frontiers in Medicine *9*. https://doi.org/10.3389/fmed.2022.1065365.

158. Ueland, T., Otterdal, K., Lekva, T., Halvorsen, B., Gabrielsen, A., Sandberg, W.J., Paulsson-Berne, G., Pedersen, T.M., Folkersen, L., Gullestad, L., et al. (2009). Dickkopf-1 Enhances Inflammatory Interaction Between Platelets and Endothelial Cells and Shows Increased Expression in Atherosclerosis. Arterioscler. Thromb. Vasc. Biol. https://doi.org/10.1161/ATVBAHA.109.189761.

159. Son, M.-Y., and Cho, H.-S. (2023). Anticancer Effects of Gut Microbiota-Derived Short-Chain Fatty Acids in Cancers. J. Microbiol. Biotechnol. *33*, 849.

160. Feitelson, M.A., Arzumanyan, A., Medhat, A., and Spector, I. (2023). Short-chain fatty acids in cancer pathogenesis. Cancer Metastasis Rev. *42*, 677.

161. Špehar Uroić, A., Filipović, M., Šućur, A., Kelava, T., Kovačić, N., and Grčević, D. (2024). Distinct association patterns of chemokine profile and cardiometabolic status in children and adolescents with type 1 diabetes and obesity. Front. Endocrinol. *15*, 1335371.

162. Huang, Y., Xuan, L., Xu, Q., Wang, J., and Liu, J. (2024). Differences in T cell-associated serum markers between ischemic cardiomyopathy and dilated cardiomyopathy. J. Thorac. Dis. *16*, 4655–4665.

163. Umamoto, K., Bouchi, R., Soeda, K., Satake, S., Hosoyama, T., Ohsugi, M., Ueki, K., and Kajio, H. (2024). Association of biomarkers and Barthel Index with occurrence of age-related adverse health outcomes in individuals with diabetes. J. Diabetes Investig. https://doi.org/10.1111/jdi.14286.

164. Greenman, R., Segal-Salto, M., Barashi, N., Hay, O., Katav, A., Levi, O., Vaknin, I., Aricha,

R., Aharoni, S., Snir, T., et al. (2023). CCL24 regulates biliary inflammation and fibrosis in primary sclerosing cholangitis. JCI insight *8*. https://doi.org/10.1172/jci.insight.162270.

165. Mor, A., Friedman, S., Hashmueli, S., Peled, A., Pinzani, M., Frankel, M., and Safadi, R. (2024). Targeting CCL24 in Inflammatory and Fibrotic Diseases: Rationale and Results from Three CM-101 Phase 1 Studies. Drug Saf. *47*, 869–881.

166. Heiman, M., Gupta, S., Lewandowska, M., and Shapiro, A.D. (2023). Complete Plasminogen Activator Inhibitor 1 Deficiency. In GeneReviews® [Internet] (University of Washington, Seattle).

167. Smith, K.G.C., and Clatworthy, M.R. (2010). FcγRIIB in autoimmunity and infection: evolutionary and therapeutic implications. Nat. Rev. Immunol. *10*, 328.

168. Barlev, A.N., Malkiel, S., Kurata-Sato, I., Dorjée, A.L., Suurmond, J., and Diamond, B. (2022). FcγRIIB regulates autoantibody responses by limiting marginal zone B cell activation. J. Clin. Invest. *132*. https://doi.org/10.1172/JCI157250.

169. Luckheeram, R.V., Zhou, R., Verma, A.D., and Xia, B. (2012). CD4+T Cells: Differentiation and Functions. Clin. Dev. Immunol. *2012*. https://doi.org/10.1155/2012/925135.

170. Sun, K., Johansen, F. E., Eckmann, L., & Metzger, D. W. (2004). An Important Role for Polymeric Ig Receptor-Mediated Transport of IgA in Protection against Streptococcus pneumoniae Nasopharyngeal Carriage1. J Immunol, 173(7), 4576–4581. https://doi.org/10.4049/jimmunol.173.7.4576.

171. Harris, W.S. (2010). The omega-3 index: clinical utility for therapeutic intervention. Curr. Cardiol. Rep. *12*. https://doi.org/10.1007/s11886-010-0141-6.

172. Du, F., Li, Y., Shen, J., Zhao, Y., Kaboli, P.J., Xiang, S., Wu, X., Li, M., Zhou, J., Zheng, Y., et al. (2019). Glyoxalase 1 gene improves the antistress capacity and reduces the immune inflammatory response. BMC Genet. *20*. https://doi.org/10.1186/s12863-019-0795-z.

173. Sukocheva, O.A., Neganova, M.E., Aleksandrova, Y., Burcher, J.T., Chugunova, E., Fan,R., Tse, E., Sethi, G., Bishayee, A., and Liu, J. (2024). Signaling controversy and future therapeutical perspectives of targeting sphingolipid network in cancer immune editing and resistance to tumor necrosis factor-α immunotherapy. Cell Commun. Signal. *22*, 251.

174. Bender, M., Reichart, B., Figueiredo, C., Burgmann, J.M., Leuschen, M., Wall, F., Radan, J., Neumann, E., Mokelke, M., Buttgereit, I., et al. (2024). An Approach to Controlling Inflammation and Coagulation in Pig-to-Baboon Cardiac Xenotransplantation. Xenotransplantation *31*, e12877.

175. Xing, Z., Li, X., He, J., Chen, Y., Zhu, L., Zhang, X., Huang, Z., Tang, J., Guo, Y., and He, Y. (2024). OLFM4 modulates intestinal inflammation by promoting IL-22ILC3 in the gut. Commun Biol *7*, 914.

176. Liu, L., Lu, X., Fan, Z., Deng, J., Zhang, S., Zhang, L., and Zha, X. (2024). TPCA-1 compound, inhibiting testis-specific serine/threonine protein kinase 3 for potential male sterile in Bombyx mori. Pest Manag. Sci. https://doi.org/10.1002/ps.8347.

177. Zhang, Q., and Li, C.S. (2019). Risk stratification and prognostic evaluation of endothelial

cell-specific molecule1, von Willebrand factor, and a disintegrin-like and metalloprotease with thrombospondin type 1 motif for sepsis in the emergency department: An observational study. Exp. Ther. Med. *17*. https://doi.org/10.3892/etm.2019.7485.

178. Xu, H., Diolintzi, A., and Storch, J. (2019). Fatty acid-binding proteins: functional understanding and diagnostic implications. Curr. Opin. Clin. Nutr. Metab. Care *22*, 407.

179. Jiménez-Cortegana, C., López-Enríquez, S., Alba, G., Santa-María, C., Martín-Núñez, G.M., Moreno-Ruiz, F.J., Valdés, S., García-Serrano, S., Rodríguez-Díaz, C., Ho-Plágaro, A., et al. (2024). The Expression of Genes Related to Reverse Cholesterol Transport and Leptin Receptor Pathways in Peripheral Blood Mononuclear Cells Are Decreased in Morbid Obesity and Related to Liver Function. Int. J. Mol. Sci. *25*. https://doi.org/10.3390/ijms25147549.

180. Serpine1 serine (or cysteine) peptidase inhibitor, clade E, member 1 [Mus musculus (house mouse)] - Gene - NCBI Serpine1 (https://www.ncbi.nlm.nih.gov/gene/18787).

181. Tahir, R.A., Sehgal, S.A., Khattak, N.A., Khan Khattak, J.Z., and Mir, A. (2013). Tumor necrosis factor receptor superfamily 10B (TNFRSF10B): an insight from structure modeling to virtual screening for designing drug against head and neck cancer. Theor. Biol. Med. Model. *10*, 1–14.

182. Dietary linoleic acid and risk of coronary heart disease (2014). The Nutrition Source. https://nutritionsource.hsph.harvard.edu/2014/11/05/dietary-linoleic-acid-and-risk-of-coronary-heart-disease/.

183. Blodgett, J.M., Theou, O., Howlett, S.E., and Rockwood, K. (2017). A frailty index from common clinical and laboratory tests predicts increased risk of death across the life course. GeroScience *39*. https://doi.org/10.1007/s11357-017-9993-7.

184. Ticinesi, A., Nouvenne, A., Cerundolo, N., Parise, A., Mena, P., and Meschi, T. (2024). The interaction between Mediterranean diet and intestinal microbiome: relevance for preventive strategies against frailty in older individuals. Aging Clin. Exp. Res. *36*, 58.

185. O'Mahony, L., O'Shea, E., O'Connor, E.M., Tierney, A., Harkin, M., Harrington, J., Kennelly, S., Arendt, E., O'Toole, P.W., and Timmons, S. (2023). Older adults and healthcare professionals have limited awareness of the link between the Mediterranean diet and the gut microbiome for healthy aging. Front Nutr *10*, 1104238.

186. Andreo-López, M.C., Contreras-Bolívar, V., Muñoz-Torres, M., García-Fontana, B., and García-Fontana, C. (2023). Influence of the Mediterranean Diet on Healthy Aging. Int. J. Mol. Sci. *24*. https://doi.org/10.3390/ijms24054491.

187. Ecarnot, F., and Maggi, S. (2024). The impact of the Mediterranean diet on immune function in older adults. Aging Clin. Exp. Res. *36*, 117.

188. Lam, K.N., Alexander, M., and Turnbaugh, P.J. (2019). Precision Medicine Goes Microscopic: Engineering the Microbiome to Improve Drug Outcomes. Cell Host Microbe *26*, 22–34.

189. D, F., Barone, M., Tavella, T., Rampelli, S., Brigidi, P., and Turroni, S. (2022). Host Microbiomes in Tumor Precision Medicine: How far are we? Curr. Med. Chem. *29*,

3202–3230.

190. Chan, I.S., and Ginsburg, G.S. (2011). Personalized medicine: progress and promise. Annu. Rev. Genomics Hum. Genet. *12*, 217–244.

191. Jobin, C. (2018). Precision medicine using microbiota. Science *359*, 32–34.