Appearance of green tea compounds in plasma following acute green tea consumption is modulated by the gut microbiome in mice

John D. Sterrett,[1,2] Kevin D. Quinn,[3] Katrina A. Doenges,[3] Nichole M. Nusbacher,[4] Cassandra L. Levens,[5] Mike L. Armstrong,[3] Richard M. Reisdorph,[3] Harry Smith,[3] Laura M. Saba,[3] Kristine A. Kuhn,[5] Catherine A. Lozupone,[4] Nichole A. Reisdorph[3,#]

1. Department of Integrative Physiology, University of Colorado, Boulder, CO
2. Interdisciplinary Quantitative Biology, University of Colorado, Boulder, CO
3. Skaggs School of Pharmacy and Pharmaceutical Sciences, University of Colorado, Aurora, CO
4. Department of Biomedical Informatics, Anschutz Medical Campus, University of Colorado, Aurora, CO
5. Division of Rheumatology, Department of Medicine, University of Colorado, Aurora, CO

#: Address correspondence to Nichole Reisdorph, nichole.reisdorph@cuanschutz.edu

Running Head: Gut microbiome affects green tea compounds in plasma

Abstract

Studies have suggested that phytochemicals in green tea have systemic anti-inflammatory and neuroprotective effects. However, the mechanisms behind these effects are poorly understood, possibly due to differential metabolism of phytochemicals resulting from variation in gut microbiome composition. To unravel this complex relationship, our team utilized a novel combined microbiome analysis and metabolomics approach applied to low complexity microbiome (LCM) and human colonized (HU) gnotobiotic mice treated with an acute dose of powdered matcha green tea. A total of 20 LCM mice received 10 distinct human fecal slurries for an n=2 mice per human gut microbiome; 9 LCM mice remained un-colonized with human slurries throughout the experiment. We performed untargeted metabolomics on green tea and plasma to identify green tea compounds that were found in plasma of LCM and HU mice that had consumed green tea. 16S ribosomal RNA gene sequencing was performed on feces of all mice at study end to assess microbiome composition. We found multiple green tea compounds in plasma associated with microbiome presence and diversity (including acetylagmatine, lactiflorin, and aspartic acid negatively associated with diversity). Additionally, we detected strong associations between bioactive green tea compounds in plasma and specific gut bacteria, including associations between spiramycin and *Gemmiger*, and between wildforlide and *Anaerorhabdus*. Additionally, some of the physiologically relevant green tea compounds are likely derived from plant-associated microbes, highlighting the importance of considering foods and food products as meta-organisms. Overall, we describe a novel workflow for discovering relationships between individual food compounds and composition of the gut microbiome.

Importance

Foods contain thousands of unique and biologically important compounds beyond the macro- and micro-nutrients listed on nutrition facts labels. In mammals, many of these compounds are metabolized by the community of microbes in the colon. These microbes may impact the thousands of biologically important compounds we consume; therefore, understanding microbial metabolism of food compounds will be important for understanding how foods impact health.

We used metabolomics to track green tea compounds in plasma of mice with and without complex microbiomes. From this, we can start to recognize certain groups of green tea-derived compounds that are impacted by mammalian microbiomes. This research presents a novel technique for understanding microbial metabolism of food-derived compounds in the gut, which can be applied to other foods.

Introduction

Assessing the impact of food on health is challenging.[1] This is exacerbated by the unique co-metabolism of foods by the host and the gut microbiome across individuals with distinct microbiota. Owing to recent advances in 'omics technologies, determining the identity of microbial and microbial:host metabolites following consumption of specific foods is now possible.[1–3] For example, Wang, *et. al.* found that a product of microbiome:host co-metabolism of phosphatidylcholine, trimethylamine N-oxide (TMAO), predicted risk for cardiovascular disease (CVD).[4] In addition, integrated metabolomics and microbiome approaches are being applied with promising results.[5–12] However, challenges remain, including identifying metabolites that are specifically produced by the microbiome or by host:microbiome interactions.

Alteration of the gut microbiome has been associated with the development of multiple disorders,[13] including depression and anxiety,[14] metabolic syndrome,[15] and inflammation.[16]A broad understanding of how the microbiome affects the host metabolome has come from comparing germ-free mice to those colonized with the microbiota of various humans[17] and/or treated with antibiotics compared to non-antibiotic treated controls.[17–19] These studies found that microbial composition has a profound influence on the presence and relative abundances of many metabolites in various sites including the blood, urine, feces, and the gastrointestinal tract. Additionally, dietary differences clearly influence microbial community structure, as certain substrates favor specific taxa. For example, Wu *et al.*[20] reported greater relative abundance of *Bacteroidetes* with high animal protein consumption and greater relative abundance of *Prevotella* with plant consumption, which is supported by other research investigating the effects of a mediterranean diet intervention.[21] Together, these findings suggest that unique microbiome profiles may be bidirectionally linked to specific dietary components.

Even dietary components that are not calorically dense, such as green tea (GT), have effects on mammalian health. Several studies have suggested that various components of tea, including flavonoids and phenolic acids, have both anti-inflammatory and neuroprotective effects.[22–25] These polyphenols are metabolized by a combination of the host and microbiome, which means that variation in the functional capacity of the microbiome along the gastrointestinal tract will affect the downstream host metabolism of GT-derived compounds.[26] However, the exact mechanisms by which the microbiome alters GT metabolism and downstream molecular networks are not yet understood. To our knowledge, only a few studies have focused on the metabolism of tea in the context of the microbiome. For example, Axling *et. al.* found that supplementing mice with GT powder along with *Lactobacillus plantarum* promoted growth of *Lactobacillus* and attenuated high fat diet-induced inflammation.[27] Studies utilizing humanized mice is a logical next step for determining the metabolites responsible for these effects and which microbes might be involved in their metabolism.

Our approach begins to address this challenge through use of human colonized (HU) mice fed GT. It aims to determine how the gut microbiome affects which GT compounds are found in the plasma of HU and low complexity microbiome (LCM) mice. These scientific premises resulted in the formation of our ***hypotheses*** *that specific GT compounds in plasma will associate with specific bacterial genera or be altered in concentration by the presence of a gut microbiome*. We tested these hypotheses by colonizing gnotobiotic mice with microbiomes from 10 healthy humans and then treating both colonized and LCM mice with an acute dose of GT extract by oral gavage (**Figure 1**).

Because diet is among the most significant modifiers of human health, detailed understanding of the microbiome in the context of food metabolism is critical for disease modification and prevention.

Methods

Chemicals, standards, and reagents: Chemicals, standards, and reagents used for sample preparation and liquid chromatography-mass spectrometry (LCMS) analysis were of high-performance liquid chromatography (HPLC) or LCMS grade. These included water from Honeywell Burdick & Jackson (Muskegon, MI, USA), methanol, and J.T. Baker methyl tert-butyl ether (MTBE) from VWR (Radnor, PA, USA), formic acid from ThermoFisher Scientific (Waltham, MA, USA), Fisher Chemical acetonitrile and methanol from Fisher Scientific (Fair Lawn, NJ, USA), and Supelco 2-Propanol from Millipore Sigma (Burlington, MA, USA). Authentic standards for sample preparation were from Avanti Polar Lipids Inc. (Alabaster, AL, USA), Cambridge Isotope Laboratories (Tewksbury, MA, USA), Sigma-Aldrich (St. Louis, MO, USA), and CDN Isotopes (Pointe-Claire, Quebec, Canada).

Human fecal sample collection: Human fecal samples were collected from 10 healthy participants using a commode specimen collector. Fecal samples were then shipped cold within 48 hours, aliquoted, and stored at -80 °C prior to use in gnotobiotic mouse experiments. Collection of the human fecal samples used in these experiments was approved by the Colorado Multiple Institutional Review Board (COMIRB) COMIRB #14-1595.

Mouse Colonization: All mouse studies were approved by the University of Colorado Institutional Animal Care and Use Committee (IACUC). C57BL/6 germ-free mice from the colony maintained at the University of Colorado Gnotobiotics Core were placed on an irradiated, low polyphenol diet (TD.97184, Envigo, Indianapolis, IN) at the time of weaning (3 weeks of age) through 6 weeks of age. At 4 weeks of age, mice were orally gavaged with 200 µL of either human fecal slurry (100 mg stool homogenized in 1 mL reduced phosphate buffered saline (PBS) in an anaerobic chamber) or PBS for LCM controls (n=9). A total of 20 mice received 10 distinct human fecal slurries, for an n=2 mice per human gut microbiome. Due to limited availability of LCM mice, the experiment was conducted in three cohorts as described in the Supplemental Methods. Mice were housed individually in a sterile caging system to maintain their LCM or individualized colonization status for the remainder of the experiment. Mice were weighed three times weekly and on the day of GT dosing.

Mouse GT Gavage: Gavage solutions of GT were prepared under sterile conditions in a Nuair Biological Safety Cabinet (Plymouth, MN). LCMS water was heated to 70 °C and sterile filtered through a Steriflip disposable vacuum filtration tube (Merck KgaA, Darmstadt, Germany). Jade Leaf brand matcha green tea ("Culinary Grade Premium Second Harvest - Authentic Japanese Origin (8.8 Ounce Pouch)", Seattle, WA) was irradiated by Envigo at the same time as the low polyphenol diet and stored in a sterile 50 mL Falcon Tube (Corning, Glendale, AZ) at 4 °C. GT slurries were prepared at 10 mg/mL with the warm sterile water, vortexed for 15 seconds to mix and then stored at 4 °C until gavage. A total of 29 mice (20 HU, 9 LCM) were gavaged with 100 μL of GT slurry (50 mg/kg equivalent). Four leftover slurries were stored at -80 °C for GT metabolomics analysis.

Plasma and Tissue Collection: Mice were sacrificed approximately 2 hours after the GT gavage. Blood was collected in 1.3 mL K3 EDTA micro sample tubes (Sarstedt Inc., Nuembrecht, Germany) via submandibular bleeding, inverted 5 times and immediately placed on ice. Blood was centrifuged at 3,000 xg for 30 mins at 4 °C, within 30 minutes of collection. Plasma was aliquoted into 1.5 mL microcentrifuge tubes (Fisher Scientific) and stored at -80 °C until analysis. Fecal microbiome samples were collected in 1.5 mL microcentrifuge tubes, flash frozen in liquid nitrogen and stored at -80 °C until analysis. The 2 hour time point was selected following a small pilot utilizing wild type (WT) mice whereby GT compounds were detected in higher numbers and abundance compared to 4 and 24 hour timepoints (Supplemental Methods).

Metabolomics: Frozen plasma samples and four GT extracts were thawed on ice and prepared for analyses as previously described.[28–30] Briefly, proteins were precipitated and small molecules were extracted from supernatants using an organic liquid-liquid extraction technique with methyl tert-butyl ether (MTBE) and water. This resulted in hydrophilic (aqueous) and hydrophobic (lipid) fractions. Nine hydrophobic and 6 hydrophilic labeled spike-in standards were used to monitor instrument performance and sample preparation batch variability.[31] Spiked and un-spiked methanol preparatory blanks were prepared alongside study samples. Quality control (QC) included analysis of purchased plasma (BioIVT, Westbury, NY), spiked with authentic standards and prepared alongside study plasma samples (Supplemental Methods).

Aqueous and lipid fractions were analyzed by LCMS using published methods.[29,31,32] Briefly, the hydrophobic fraction was analyzed using reverse phase C18 chromatography and a quadrupole time-of-flight mass spectrometer (QTOF 6545, Agilent Technologies, Santa Clara, CA) in positive ionization mode. The hydrophilic fraction was analyzed using an SB-Aq column (Agilent Technologies) on the same instrument. LCMS method details are provided in the Supplemental Methods.

Metabolomics Data Processing: Metabolomics spectral data for both plasma and GT extracts were processed using a recursive workflow and area-to-height conversion with Agilent's MassHunter Profinder ver. 10.0 service pack 1 (Profinder) and Mass Profiler Professional Ver. 15.1 (MPP, Agilent Technologies).[10] Data from lipid and aqueous fractions were extracted

separately in Profinder using Batch Molecular Feature Extraction followed by Batch Targeted Feature Extraction. Compounds found in blanks were removed. Plasma samples were limited to compounds eluting before 10.4 min and 13 min for lipid and aqueous fractions, respectively. GT samples were limited to compounds eluting before 10.4 min and 11 min for lipid and aqueous fractions, respectively. Compounds eluting past these times had poor signal to noise ratios and were not used.

Plasma compounds were annotated using MassHunter ID Browser ver. 10.0 (Agilent Technologies) by searching custom in-house databases and public databases consisting of compounds from METLIN, Lipid Maps, Kyoto Encyclopedia of Genes and Genomes (KEGG), and Human Metabolome Database (HMDB) using accurate mass and isotope ratios. An initial database search was conducted using H+ as the primary charge carrier. Unannotated compounds were re-searched using the same databases with loss of water included. These compounds were designated Metabolomics Standards Initiative (MSI) level three and are listed by software-assigned compound number (e.g. C1287 is compound number 1287).[11] Plasma compound lists were exported for statistical analysis and to determine which GT compounds were detected in plasma.

Green Tea compounds in Plasma: Data from the four GT extracts were processed as above with the following adjustments: GT compounds had to be present in 100% of GT samples and the compounds were restricted to those eluting before 10 min for both the aqueous and lipid fractions. In addition to the databases used to annotate plasma, the GT samples were also searched against Phenol Explorer, FooDB, and a Traditional Chinese Medicine natural products database (Agilent Technologies). The GT compound lists were exported as .cef files and imported into Quantitative Analysis (Agilent Technologies).

To verify that GT compounds originated from GT, data from a separate experiment using mice not fed green tea was used to determine baseline differences (Supplemental Methods). Briefly, 10 mice were gavaged with the same 10 distinct human fecal slurries as described above (i.e. 1 microbiome per mouse), and 6 mice were gavaged with PBS. At two weeks post-gavage, plasma was collected as described in the main text. LCMS data for plasma samples was processed as described above and imported into Quantitative Analysis. Extracted ion chromatograms were generated for each GT compound from plasma LCMS data using a “targeted data extraction” strategy. The method parameters included: left and right extraction window of 10 ppm, retention time extraction window $\pm$ 0.2 minutes, and retention time outlier >0.3 minutes. For each GT compound, the monoisotopic peak (M1) and C13 peak (M2) were used for qualification. The M1 peak was used as the quantifier ion, while M1 and M2 peaks (or only M1 when no M2 peak was detected) were used as qualifier ions. When a second charge carrier such as sodium was detected, its monoisotopic peak was used as a qualifier. Qualifier relative uncertainty was set to 20%. Analysis results were imported into MPP and filtered to retain only targets that were present in at least 3 plasma samples with an area ≥20,000 counts for the aqueous fraction and ≥ 55,000 counts for the lipid fraction. These compound lists, including area under the curve (AOC) for relative quantitative comparison, were exported for statistical and informatic analysis.

Whole metabolomics datasets were visualized using principal components analysis (PCA) for the purpose of quality control (Supplemental Methods).

Microbiome - DNA extraction and sequencing: DNA was extracted from the fecal samples using the Power Soil Pro Kit protocol (Qiagen, Germantown, MD). Barcoded primers targeting the V4 region of the16S rRNA gene were used to PCR amplify the extracted bacterial DNA using the Earth Microbiome Project (EMP) standard protocols (http://www.earthmicrobiome.org).[33] PCR product quantification was completed using PicoGreen (Invitrogen, Carlsbad, CA) and used to pool equal amounts of amplified DNA from each sample. The QIAquick PCR Purification Kit (Qiagen) was used to clean the pooled libraries. Three runs were used to generate sequences using the Illumina MiSeq platform (San Diego, CA).

Microbiome - Sequence data processing: Microbiome sequencing data were processed in QIIME2-2021.8.[34] Sequences were demultiplexed using the q2-demux plugin, then denoised using q2-dada2[35] with a truncation length of 230bp. A phylogenetic tree was created using Saté-enabled phylogenetic placement (SEPP) via the q2-fragment-insertion plugin.[36] Taxonomy was assigned to reads using a naïve-Bayes classifier trained on the latest Greengenes 1 database as of January 2022 (Greengenes 13.8).[37] Reads with no taxonomy below the kingdom level and reads classified as mitochondria and chloroplast were removed from further consideration. For diversity analyses, samples were rarefied to 40,625 reads per sample, which was the highest rarefaction depth that did not exclude any samples. A taxa bar plot was created using Microshades.[38]

Differential abundance testing: Analysis of Compositions of Microbiomes with Bias Correction (ANCOM-BC)[39] was used to assess differentially abundant taxa between LCM and HU mice. ANCOM-BC was performed at the family and genus level, and a Benjamini-Hochberg FDR correction was applied to p values.

Diversity metrics: Alpha diversity was calculated using Faith's phylogenetic diversity,[40] and microbiome beta diversity was calculated using unweighted UniFrac distances.[41] Metabolome beta diversity was calculated using the Bray-Curtis distance metric.

Ordination: Microbiome principal coordinate analysis (PCoA) plots were generated using the unweighted UniFrac distance matrix via the Python package scikit-bio (v0.5.6),[42] and metabolomics PCoA plots were generated using Bray-Curtis distances.

Calculation of centroids for distance-based statistics: To avoid pseudoreplication[43] affecting precision of estimates in permutation-based tests where random effects or cluster-robust standard errors cannot be used, we collapsed microbiome UniFrac distances to centroids within humanized microbiome donors, or LCM centroids within each experimental cohort, using the dist_multi_centroids function from the usedist R package (v0.4.0).[44] This reduced sample size from n=29 (20 mice with microbiomes humanized in pairs from 10 unique donors + 9 LCM across 3 experimental cohorts) to n=13 (10 averaged HU centroids corresponding to each

donor microbiome + 3 LCM centroids corresponding to each cohort). Corresponding non-distance data (such as compound abundance) were averaged (arithmetic mean) within groups.

Procrustes randomization test: Procrustes randomization tests were performed on the metabolomics and microbiome PCoA coordinates from centroid distances using the protocol described by Peres-Neto and Jackson.[45] Specifically, we first performed a Procrustes transformation on the observed datasets, and then permuted the GT lipid or GT aqueous compound Bray-Curtis distance PCoA coordinates 10^4 times and performed a Procrustes transformation on the permuted datasets. The *p* value was calculated based on the portion of permuted Procrustes-transformed datasets with resulting m^2 (Gower's statistic, also referred to as disparity) scores lower than the Procrustes m^2 score of the observed datasets.

Compounds associated with microbiome composition: To identify compounds and neurochemicals associated with microbiome community composition, PERMANOVA tests were performed using Adonis2 from the R package vegan (v2.6-4)[46] to assess correlations between plasma abundances of each individual compound and the microbiome unweighted UniFrac distance matrix (using 10^4 permutations). *p* values were adjusted with a Benjamini-Hochberg correction to reduce false discovery rates. Due to reduced power from using centroids, we considered uncorrected $p < 0.01$ to be associated with microbiome composition. For any compounds associated with microbiome composition, we used the R package Selbal (v0.1.0)[47] with 10-fold cross validation to identify genus-level balances that explained compound abundance.

Compounds associated with microbiome diversity: We performed linear regression with cluster robust standard errors on each individual compound and Faith's phylogenetic diversity to assess which compounds associated with microbiome diversity. Standard errors were clustered by humanized microbiome donor, or by experiment for LCM mice, and *p* values were calculated based on clustered standard errors using the R package fixest (v0.11.2).[48] *p* values were adjusted with a Benjamini-Hochberg correction to reduce false discovery rates (FDR). Regressions with any data points that had a dffit (difference in model fit if the point was removed) absolute value greater than 1 were removed from further consideration.

Mixed-effects multi-omics modeling: Associations between metabolites and microbes were assessed by pairwise linear mixed effects models between each microbe and each metabolite, with the formula:

$$z\ score(metabolite) \sim \beta * arcsinh(microbe\ *\ 100)\ +\ (1|humanized\ id)\ +\ \epsilon \qquad (eqn.1)$$

Metabolite concentrations were z-score transformed so that β coefficients could be compared across compounds. Microbe relative abundances were multiplied by 100 to transform to percent relative abundance and then arcsinh transformed ($arcsinh(x) = ln\sqrt{x^2+1}$) to alter the distributions of relative abundances to minimize violations of linear modeling assumptions across features. (1|*humanized id*) indicates a random effect for the humanized microbiome ID or LCM status. The maximum influence of any datapoint was calculated for each regression using

the mdffits function from the R package HLMdiag (v0.5.0).[49] Any regressions with any data points that had an influence (difference in model fit if the point was removed) absolute value greater than or equal to 4 were removed from further consideration. Models were run using the R package lme4 (v1.1-29),[50] and *p* values were calculated based on Satterthwaite's degrees of freedom using the R package LmerTest (v3.1-3).[51] *p* values were adjusted with a Benjamini-Hochberg correction to reduce FDR.

Data and code availability: Data and code for reproducing these analyses can be accessed at https://github.com/sterrettJD/GT-micro-metabo. Raw 16S sequencing data is available in the European Nucleotide Archive using the accession IDs PRJEB77100 and ERP161582.

Results

Metabolomics of mouse plasma and GT: Metabolomics of the mouse plasma resulted in 4,282 lipid and aqueous compounds while metabolomics of GT extract resulted in 4,415 compounds (**Figure 1**). Of those, 624 GT compounds were detected in at least 3 plasma samples following GT gavage. 432 were found in the lipid-rich extract (GT lipids) while 192 were found in the aqueous extract (GT aqueous).

We removed from consideration any plasma GT compounds that were also found in the plasma of mice not gavaged with GT. This step removed 86 lipid compounds and 59 aqueous compounds, resulting in a total of 145 compounds removed. After all filtering steps, 479 (346 lipids, 133 aqueous) compounds remained, which were evaluated for their relationships with microbiome composition and diversity. These compounds are listed in the supplemental file denoted "Green Tea Compounds Table".

Contamination of germ-free controls: Following 16S rRNA gene amplicon sequencing of LCM and HU fecal microbiomes, it was discovered that the LCM mice were not germ-free as intended, and instead had microbiomes (**Figure 2**). Additionally, the LCM controls had small cecums, which are uncharacteristic of germ-free mice (data not shown). Specifically, the LCM controls were dominated by 6 Amplicon Sequence Variants (ASVs) in the Firmicutes phylum, which made up over 90% of the detected reads. **Supplemental Table 1** shows the taxonomic assignment and relative abundances of these ASVs. Following extensive investigation, it was determined that the contamination resulted from the irradiated low polyphenol food that was fed to mice. For example, bacterial colonies grew when irradiated food was plated and PCR analysis showed a range of 3.5 ng/µL-4.6 ng/µL in the 3 irradiated samples tested. No growth was seen when irradiated GT was plated. Discussions with the vendor were largely inconclusive, but it was surmised that the source of bacteria may have been casein used in the diet.[52] While not ideal, because all mice in the study received the same food, it was determined that all mice were exposed to this contamination, and therefore we continued with data analysis, focusing on statistical methods that consider microbiome community composition or diversity rather than solely group status (LCM vs HU).

Taxonomic composition of fecal microbiomes: LCM and HU mouse microbiomes differed in community composition (pseudo-F = 5.9, p = 0.005). LCM mice were colonized almost entirely

by Firmicutes, with six of the nine mice having high levels of unassigned ASVs in the family Peptostreptococcaceae **(Figure 2**, p = 0.005**)**. Additionally, LCM mice had higher relative abundances of *Epulopiscium* (ANCOM-BC; p = 0.015), and *Turicibacter* (ANCOM-BC; p = 0.002) in their microbiomes compared to HU mice **(Figure 2)**.

Microbiome composition was very similar within HU mouse replicates (UniFrac distance within pairs mean [95% confidence interval (CI)] = 0.25 [0.21 – 0.30]), and variable between mice colonized with different human donors (UniFrac distances between centroids mean [95% CI] = 0.50 [0.49 – 0.52]) (**Figure 2**). This is supported by a PERMANOVA demonstrating larger variance in phylogenetic composition between pairs than within pairs (R^2 = 0.87, psuedo-F = 7.6, p < 0.001). The microbiome of all but two pairs of HU mice had *Akkermansia* present at high relative abundances (≥20%). All HU pairs had representation of Bacteroidetes, primarily including *Bacteroides*, *Parabacteroides,* and *Alistipes*. Proteobacteria were present in the microbiome of all HU but not LCM mice, and this phylum predominantly consisted of ASVs belonging to *Bilophila* and *Sutterella*. Some HU pairs also had low relative abundances (<5%) of *Fusobacteria* in their fecal microbiome.

Diversity, composition and relationships between the fecal microbiome, plasma metabolome, and plasma lipidome: Both alpha and beta diversity of the fecal microbiome differed between HU and LCM mice. Specifically, microbiome phylogenetic composition (UniFrac distance) strongly differed between LCM and HU mice (pseudo-F = 5.9, p = 0.005), and clear separation between LCM and HU can be seen in the unweighted UniFrac PCoA (**Figure 3A**). Additionally, Faith's phylogenetic diversity was significantly higher in HU mice (mean [95% CI] = 12.0 [12.3 – 13.3]) than LCM mice (mean [95% CI] = 4.1 [2.3 – 5.9]; regression β = 8.7, p < 0.001; **Figure 3B**).

The overall composition of GT-specific lipids in the plasma was not different between LCM and HU mice (Bray-Curtis distance PERMANOVA on centroids R^2 = 0.12, pseudo-F = 1.5, p = 0.14; **Figure 3C**), nor was the composition of aqueous GT-specific (Bray-Curtis distance PERMANOVA on centroids R^2 = 0.16, pseudo-F = 0.17, p = 0.96; **Figure 3D**). This can be seen in PCoA plots, where the GT compounds in plasma do not cluster according to LCM or HU group (**Figure 3C and 3D**). The portion of shared GT compounds present in plasma was high in both the lipid and aqueous datasets (Jaccard similarity lipid mean [min – max] = 1.0 [0.99 – 1.0]; aqueous mean [min – max] = 0.95 [0.88 – 0.98]. Furthermore, a Procrustes randomization test did not reveal an overall significant relationship between microbiome composition (UniFrac) and GT lipids in plasma (**Figure 4A**, m^2 = 0.39, p = 0.085) or aqueous GT compounds in plasma (using Procrustes-transformed Bray-Curtis PCoA, **Figure 4B**, m^2 = 0.66, p = 0.79).

Green tea compounds associated with microbiome composition: One goal of the study was to determine if GT compounds in plasma associate with microbiome diversity and composition. Understanding these associations are a first step towards determining which bacterial species may be responsible for the metabolism of specific GT compounds. PERMANOVA tests on centroids assessing associations between the concentration of each GT compound and

microbiome composition (unweighted UniFrac) found no significant relationships after FDR correction. For reference, **Supplemental Table 2** shows compounds with uncorrected $p < 0.05$**.**

The lack of significance looking while comparing multivariate similarity between samples may be too unfocused. It is very likely that many compounds would be unaffected by most members of the microbiome, as many food-derived compounds are absorbed in the small intestine. We thus looked for finer-grained relationships between individual compounds and smaller sets of microbes.

Though whole community-level associations were not found, for the 12 compounds most associated with microbiome composition, Selbal-identified balances were able to explain 45% to 70% of the variation in the abundance of these compounds, and all had a regression slope adjusted $p < 0.05$. Gamma-glutamyl-alanine having the most variation explained by a positive correlation with a balance of the genera *Alistipes* to *Butyricimonas* **(Figure 5)**. Balances identified by Selbal are ratios of taxa, meaning that as gamma-glutamyl-alanine increased, *Alistipes* increased relative to *Butyricimonas,* or *Butyricimonas* decreased relative to *Alistipes*.

Green tea compounds associated with microbiome phylogenetic diversity: When testing GT compound associations with microbiome alpha diversity (Faith's phylogenetic diversity), 8 compounds had FDR-corrected $p < 0.05$ **(Figure 6 and Supplemental Table 3)**, including lactiflorin, acetylagmatine, aspartic acid, 7,8- dihydroparasiloxanthin, montecristin, and CL(82:16) **(Figure 7).** Most of these compounds had negative associations with microbiome diversity, except for CL(82:16) and an unannotated compound.

Relationships between individual microbes and GT compounds: When controlling for the humanized microbiome with which mice were colonized with a random effect and applying a Benjamini-Hochberg FDR correction, across all mice (including LCM mice) we identified 161 significant relationships between microbes and GT compounds in plasma out of the potential 33,433 relationships tested. Of these, the strongest associations were between an ASV in the genus *Bifidobacterium* and an unannotated compound, an ASV in the genus *Allobacterium* and montecristin ($\beta = 277.8$, $p < 0.001$, data not shown), and an ASV in the family Peptostreptococcaceae and acetylagmatine ($\beta = 0.5$, $p < 0.001$, data not shown).

Despite including LCM status as a covariate in each compound-taxon regression, there was still a large signal related to the high relative abundance taxa in the LCM mice, such as Peptostreptococcaceae. Thus, we also performed this analysis in only HU mice, as differences in overall architecture of the LCM vs HU mice fecal microbiomes may have biased these individual metabolite-taxon relationships and violated the random effects assumption. When mixed effects regressions were performed for only HU mice, several GT compounds were significantly associated with specific bacterial ASVs. Of 30,594 potential relationships tested, 22 were significant after Benjamini-Hochberg FDR correction (**Figure 8A**). For example, the GT compound wilforlide was positively associated with an ASV in the genus *Anaerorhabdus* ($\beta = 15.1$, $p < 0.001$). Spiramycin was positively associated with taxa in the genera *Gemmiger* ($\beta = 13.7$, $p < 0.001$) and *Lactobacillus* ($\beta = 12.9$, $p < 0.001$, **Figure 8B**). While multiple unannotated

compounds had multiple significantly associated microbes, spiramycin was the annotated compound with more than one significant relationship with individual taxa (after FDR correction). The genus *Gemmiger* had the highest number of significant relationships with GT compounds, though 3 of these 4 relationships were primarily with unannotated compounds, and the relationships between compounds are not known.

Discussion:

We show that GT compounds significantly associated with specific gut bacterial genera following the acute feeding of GT extract to mice. Our analysis strategy included an important data reduction step, whereby analysis focused on only GT compounds that were found in the plasma of mice who had consumed GT. This greatly increased power (via decreasing the need for as stringent FDR correction) by limiting analysis from 4,282 plasma compounds to 479 compounds that were specific to the intervention. Importantly, these are compounds detected in dietary formulations of green tea, not solely compounds produced by the plant from which green tea is derived, *Camellia sinensis.* Much like the mice in this study with associated microbiomes, *C. sinensis* plants are meta-organisms that coexist with microbial communities.[53,54] Microbial metabolites and even excreted metabolites from other nearby plants or pollinators may be present in the leaves used to create the tea used in this study, so the GT compounds discussed may come from a variety of sources across multiple kingdoms of life.

Sequencing of LCM and HU microbiomes from fecal pellets allowed us to perform analyses considering microbiome composition and diversity, rather than solely LCM status. This was particularly useful because the LCM mice in this study were populated with a low diversity microbiome due to contamination of the low polyphenol food used. Despite using phylogenetic composition of the microbiome in our analyses, we did not see multivariate-multivariate relationships between overall composition of green tea compounds in plasma and the composition of the gut microbiome (per a Procrustes randomization test on PCoA coordinates from green tea compound Bray-Curtis distance and Unweighted UniFrac microbiome distance), which could be due to the absorption of many green tea compounds in the small intestine, before reaching the majority of microbial biomass in the colon.

Multiple compounds found in GT extracts were individually associated with microbiome composition and diversity. Importantly, these compounds are known to be present in plants and are relevant to mammalian physiology. For example, an annotated GT compound in plasma associated with microbiome diversity, montecristin, is an annonaceous acetogenin (a group of compounds that are waxy derivatives of fatty acids and are reported to have an array of health effects, including antimalarial, antiparasitic, and anti-cancer activities).[55] Another plant compound from GT that differed as a function of the microbiome was lactiflorin, a compound originally characterized in the plant genus *Paeonia*, that activates the antioxidant-controlling transcription factor nuclear factor erythroid 2-related factor (Nrf2) in rats.[56] Lactiflorin, a monoterpene glycoside, has been used in the treatment of rheumatoid arthritis via inhibiting leukocyte recruitment and angiogenesis via, potentially, the VEGFR and PI3K-Akt signaling pathways, interleukin signaling, and platelet activation.[57]

Previous studies have focused on the health effects of green tea catechins, a subgroup of flavonoids with known antioxidant properties.[58] While several catechins, including epigallocatechin 3-p-coumarate, catechin 3-gallate, gallocatechin 3'-gallate, and epigallocatechin 3-cinnamate were found in GT extract, these were not found at detectable levels in at least 3 mouse plasma samples and hence were not included in microbiome analyses. This could be due to the concentration of extract used, the sensitivity of the assay, or the timing of blood draw.

Most of the compounds associated with microbiome presence, diversity, or composition were found in lower abundance in the plasma of HU mice, relative to LCM mice. For example, lactiflorin was nearly 7 times less abundant in HU mice. Our data suggest that the presence of a gut microbiome either (1) inhibits the absorption of lactiflorin, (2) increases metabolism of lactiflorin by the host, or (3) allows for bacterial degradation of lactiflorin. In fact, most compounds associated with microbiome presence, diversity, or composition were less abundant in HU relative to LCM mice, suggesting potential co-metabolism of these compounds by the microbiome.

Another compound identified to be associated with microbiome composition was gamma-glutamyl-alanine. Notably, many plants and mammals possess gamma-glutamyl transpeptidase (GGT),[59] which is required for synthesis of gamma-glutamyl amino acids and often involved in glutathione metabolism; however, most bacteria also encode GGT,[60] making the microbiome a candidate to alter concentrations of gamma-glutamyl-alanine either *via* direct metabolism or *via* an unknown effect on host metabolism. We did not detect gamma-glutamyl-alanine in the plasma of mice not gavaged with GT, so we consider this to be gamma-glutamyl-alanine above baseline concentrations. Seventy percent of the variation in plasma gamma-glutamyl-alanine was explainable by the ratio of *Alistipes* to *Butyricimonas*. Notably, human and murine gut colonizers within the genus *Butyricimonas* contain the gene gamma-D-glutamyl-L-lysine dipeptidyl-peptidase (K20742),[61] which releases gamma-glutamyl-alanine from cell walls, and *Alistipes* is negatively correlated with plasma GGT in humans,[62] supporting the directionality of the relationship of this ratio with plasma gamma-glutamyl alanine abundance. However, this balance may just be a proxy for overall microbiome functional composition, as other members of the microbial community that interact with *Alistipes* and *Butyricimonas* may be more directly responsible for differences in gamma-glutamyl alanine.

We also found many significant relationships (after FDR correction) between GT compounds and individual taxa when only HU mice were included in analysis **(Figure 8)**. In contrast to the multivariate associations where most compounds were lower in HU mice, we found primarily positive associations between individual taxa and GT compounds in plasma. The annotated GT compound with the highest number of significant relationships with individual taxa was spiramycin, an antimicrobial produced by *Streptomyces ambofaciens*, a member of the soil microbiome/rhizosphere.[63] However, the mechanism of the positive association between spiramycin and ASVs within the genera *Gemmiger* and *Lactobacillus* is unclear. We found positive relationships with another bioactive plant compound, wilforlide B (a norditerpenoid) and

Anaerohabdus. While the data are compelling in that the majority of compounds have been found in plants, the mechanistic relationships between these compounds and individual taxa are unclear, especially given that metabolism pathways are poorly understood. Future work could track closely related compounds or apply machine learning-based approaches to identify potential metabolic pathways for these natural products and assess microbial genomes for genes responsible for these reactions.

One limitation of the study is the small number of replicate mice that were colonized with a specific human microbiome (n=2), and that mice were colonized with only 10 different human microbiomes. While significant relationships between specific bacteria and GT compounds were found, increasing the sample size would increase power to identify relationships, especially in the permutation-based tests, where we needed to reduce our effective sample size to control for pseudoreplication. More samples would potentially allow for using within-estimator (within each humanized microbiome source) to estimate the effects of small changes to individual taxa within a similar microbiome community structure subtype; this is akin to estimating the effects of increasing one species of bird within different terrestrial biomes like grassland and savannah. Another weakness is the lack of plasma and microbiome sampling before green tea gavage. We removed any compounds that were detectable in (HU and LCM) mice that did not receive a GT gavage; however, having samples from mice before and after gavage would also increase our ability to identify metabolism of GT compounds. Our focus on only GT compounds found in plasma does reduce the possibility of assessing changes in metabolic products of GT compounds.

Moreover, we chose the two-hour timepoint to maximize the GT compound signal in plasma, though multiple plasma samples after feeding would allow for more detailed resolution of GT compound abundance trajectories based on microbiome composition, including time lags to assess causal directionality. While the two-hour timepoint is when the most GT compounds were detectable in plasma in a small pilot study, it is not necessarily when the microbiome may have the strongest effects. Stronger doses of GT or more variety in timepoints may alter the effects of the microbiome on GT compounds in plasma. It should be noted that the dose of GT, if scaling linearly, is comparable to 1-3 servings in an adult human, or if scaling allometrically, is equivalent to a small serving of GT in an adult human.

An additional limitation of the study is the presence of gut microbiomes, albeit of low diversity, in the LCM mouse group. Although it was assumed that all mice were exposed to the same contaminated low polyphenol mouse food, the contaminating taxa were able to dominate the fecal microbiome of the LCM mice, such that there were no mice free of microbial metabolism of GT compounds. Currently underway studies utilize mice unexposed to GT as a control in a similarly designed acute feeding study, as well as a two-week feeding of GT to compare the effects of the microbiome on sub-chronic versus acute metabolism of GT.

One strength of our experimental design lies in the use of 10 different human microbiomes to enable the evaluation of relationships between individual microbial species and GT compounds. While many humanized microbiome studies are pseudoreplicated,[43] our use of multiple donor

microbiomes while controlling for replication covers a broad variety of possible human-like microbiomes. Together, these results suggest that specific bacterial species may affect the metabolism of these bioactive compounds, thereby influencing their health effects.

Overall, this study represents a successful workflow for discovering relationships between food compounds and composition of the gut microbiome, as well as individual gut bacterial genera. In addition, this methodology has allowed us to track GT compounds from the tea to plasma. Though our study was limited in size and timepoints, we believe this methodology can and should be applied to study the effects on the gut microbiome on food metabolism. We identified multiple relationships between microbiome composition and GT compounds in plasma after GT consumption, supporting that bacterial taxa affect the absorption and metabolism of GT compounds, thereby possibly influencing their positive or negative health effects.

Acknowledgements
This project was funded through NIH/NIDDK R01DK113957 (Reisdorph, Campbell, and Krebs) and an ALSAM Therapeutic Innovation Grant through the Skaggs School of Pharmacy and Pharmaceutical Sciences to Dr. Reisdorph. Support for this project was also provided by the National Science Foundation-sponsored Interdisciplinary Quantitative Biology PhD program, the Integrated Data Science (Int dS) Graduate Training Fellowship, the William J. Freytag Fellowship.

Figure 1. Compounds from green tea were detected in plasma after green tea gavage in mice. Flowchart shows study design. A total of 29 mice (20 humanized microbiomes, 9 low complexity microbiomes) were gavaged with 100 µL of green tea then sacrificed 2 hours later, and plasma, whole brain, and fecal microbiome samples were collected. Liquid chromatography mass spectrometry-based metabolomics was performed on green tea samples and plasma. Plasma metabolomics data was mined to determine which green tea compounds could be found in plasma. Relationships between the fecal microbiome and green tea compounds found in plasma were assessed. Abbreviations: LCMS, liquid chromatography-mass spectrometry; MeOH, methanol; rRNA, ribosomal ribonucleic acid.

Figure 2. Taxa bar plot shows differences in phylum and genus level composition between low complexity microbiome and humanized microbiome mice. Bars are colored by phylum and microshaded by genus, and height of each colored bar represents the relative abundance of the corresponding taxon. One stacked bar represents the composition of one mouse's microbiome, and humanized microbiome IDs are shown along the x-axis. Low complexity microbiome mice are labeled as "Control" and have vastly different composition compared to humanized mice.

Figure 3. Microbiome PCoA and alpha diversity show clear separation between low complexity microbiome and humanized mice, whereas metabolome PCoA plots do not separate as clearly. A) Microbiome unweighted UniFrac PCoA is colored by humanized vs low complexity microbiome status. B) Boxplot shows higher Faith's phylogenetic diversity in humanized vs low complexity microbiome mice. C) PCoA of lipid fraction of green tea compounds in plasma is colored by humanized vs low complexity microbiome status. D) PCoA of aqueous fraction of green tea compounds in plasma is colored by humanized vs low complexity microbiome. Abbreviation: PCoA, principal coordinates analysis; PCo, principal coordinates axis.

Figure 4. Procrustes analysis does not indicate multivariate-multivariate relationships between green tea lipids or compounds in plasma and phylogenetic composition of the microbiome. Panel A shows the connection between microbiome composition and the composition of aqueous GT compounds detected in plasma, where each triangle represents the composition of one microbiome sample, and a line connects to that sample's corresponding aqueous GT compounds in plasma. Panel B shows the connection between microbiome composition and the composition of lipid GT compounds detected in plasma, where each triangle represents the composition of one microbiome sample, and a line connects to that sample's corresponding lipid GT compounds in plasma. Procrustes randomization test did not reveal an overall significant relationship between microbiome composition (UniFrac) and GT lipids in plasma ($m^2 = 0.39$, $p = 0.085$) or aqueous GT compounds in plasma ($m^2 = 0.66$, $p = 0.79$). In both panels, unweighted UniFrac microbiome PCoA coordinates were left

untransformed, and metabolomics sample Bray-Curtis PCoA coordinates were transformed with a Procrustes function. Samples of the same type positioned close together were compositionally similar, whereas samples far apart were dissimilar. Shorter lines between samples indicate better overlap between the microbiome and metabolomics datasets. For statistical testing, samples belonging to mice from the same donor microbiome or low complexity microbiome mice from the same experimental cohort were collapsed into one centroid per treatment unit. Abbreviation: PCoA, principal coordinates analysis; PCo, principal coordinates axis.

Figure 5. Compositional balances in microbial relative abundances explain green tea compound abundance in plasma. Selbal with cross validation was used to identify balances (log-ratios) of taxa that predict compounds that were associated with microbiome composition. The x-axis of each plot shows the microbiome balance, with the numerator and denominator labeled on the plot as <family>_<genus>. The y-axis shows the abundance of each green tea compound, which is labeled at the top of each plot, along with the regression R^2. The line on each plot shows the regression line of best fit and the shaded region indicates the 95% confidence interval for that regression line. Each point represents the average balance for each donor microbiome or experimental control.

Figure 6. Several green tea metabolites are associated with phylogenetic diversity of the gut microbiome. Each point represents one green tea compound. The y-axis represents p-values from a regression between that compound and microbiome phylogenetic diversity, and the x-axis represents the effect size of that compound's association with microbiome diversity. Compounds abundances were z-score transformed to allow for comparison of effect sizes, and p-values were calculated using cluster robust standard errors, where cluster ID was the donor microbiome source, or experimental cohort for low complexity microbiome mice. The horizontal line represents $p = 0.01$, and vertical lines represent an effect size of 1, indicating that a 1 standard deviation increase in compound abundance was associated with a 1 unit increase in Faith's phylogenetic diversity. Compounds with $p < 0.01$ and fold change absolute value > 2 are labeled with their annotation.

Figure 7. Phylogenetic diversity of the gut microbiome is associated with green tea compound abundance in plasma. Fixed effects regression with cluster robust standard errors (clustered by humanized microbiome donor ID or low complexity microbiome experimental cohort) was used to identify green tea compounds that were associated with Faith's phylogenetic diversity of the gut microbiome, and compounds that had significant (Benjamini-Hochberg $p < 0.05$) relationship with Faith's phylogenetic diversity are shown. The x-axis of each plot shows Faith's phylogenetic diversity, and the y-axis shows the abundance of each green tea compound, which is labeled at the top of each plot. The line on each plot shows the regression line of best fit and the shaded region indicates the 95% confidence interval for that regression line. Each point represents one sample.

Figure 8. Individual green tea compounds are associated with specific genera when controlling for overall microbiome composition. A) A stacked bar plot shows a summary of relationships between green tea compounds and individual genus-level relative abundances, indicating the significance and direction of compound-microbe relationships. B) A shaped heatmap shows the direction and significance of any microbes and compounds that had significant relationships after Benjamini-Hochberg p-value correction. To generate these results, 39,860 pairwise linear mixed effects regressions were run on z-score(metabolite) ~ arcsinh(microbe) + (1|humanized_id), where (1|humanized_id) indicates the microbiome mice were humanized with, or if mice were in the low complexity microbiome group. Metabolites were transformed to have a mean of 0 and standard deviation of 1 to allow for comparison of effect sizes across metabolites. Each point represents the regression run between the metabolite on the x-axis and the microbe shown on the y-axis. The color of each point represents the beta extracted from the model (coefficient of the microbe), and the size of points represents the p-value of that coefficient. Circles represent $p > 0.05$, whereas triangles represent $p < 0.05$.

References

1. Jones DP, Park Y, Ziegler TR. Nutritional metabolomics: Progress in addressing complexity in diet and health. *Annu Rev Nutr*. 2012;32:183-202. doi:10.1146/annurev-nutr-072610-145159

2. Bouchard-Mercier A, Rudkowska I, Lemieux S, Couture P, Vohl MC. The metabolic signature associated with the Western dietary pattern: a cross-sectional study. *Nutr J*. 2013;12:158. doi:10.1186/1475-2891-12-158

3. Rangel-Huerta OD, Gil A. Nutrimetabolomics: An Update on Analytical Approaches to Investigate the Role of Plant-Based Foods and Their Bioactive Compounds in Non-Communicable Chronic Diseases. *Int J Mol Sci*. 2016;17(12):2072. doi:10.3390/ijms17122072

4. Wang Z, Klipfell E, Bennett BJ, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature*. 2011;472(7341):57-63. doi:10.1038/nature09922

5. Dopkins N, Nagarkatti PS, Nagarkatti M. The role of gut microbiome and associated metabolome in the regulation of neuroinflammation in multiple sclerosis and its implications in attenuating chronic inflammation in other inflammatory and autoimmune disorders. *Immunology*. 2018;154(2):178-185. doi:10.1111/imm.12903

6. Shaffer M, Armstrong AJS, Phelan VV, Reisdorph N, Lozupone CA. Microbiome and metabolome data integration provides insight into health and disease. *Transl Res J Lab Clin Med*. 2017;189:51-64. doi:10.1016/j.trsl.2017.07.001

7. Markle JGM, Frank DN, Mortin-Toth S, et al. Sex Differences in the Gut Microbiome Drive Hormone-Dependent Regulation of Autoimmunity. *Science*. 2013;339(6123):1084-1088. doi:10.1126/science.1233521

8. Markle JGM, Frank DN, Adeli K, von Bergen M, Danska JS. Microbiome manipulation modifies sex-specific risk for autoimmunity. *Gut Microbes*. 2014;5(4):485-493. doi:10.4161/gmic.29795

9. Bowler RP, Jacobson S, Cruickshank C, et al. Plasma sphingolipids associated with chronic obstructive pulmonary disease phenotypes. *Am J Respir Crit Care Med*. 2015;191(3):275-284. doi:10.1164/rccm.201410-1771OC

10. Hughes G, Cruickshank-Quinn C, Reisdorph R, et al. MSPrep--summarization, normalization and diagnostics for processing of mass spectrometry-based metabolomic data. *Bioinforma Oxf Engl*. 2014;30(1):133-134. doi:10.1093/bioinformatics/btt589

11. Quinn KD, Schedel M, Nkrumah-Elie Y, et al. Dysregulation of metabolic pathways in a mouse model of allergic asthma. *Allergy*. 2017;72(9):1327-1337. doi:10.1111/all.13144

12. Reisdorph N, Wechsler ME. Utilizing metabolomics to distinguish asthma phenotypes: strategies and clinical implications. *Allergy*. 2013;68(8):959-962. doi:10.1111/all.12238

13. Frank DN, Zhu W, Sartor RB, Li E. Investigating the biological and clinical significance of human dysbioses. *Trends Microbiol*. 2011;19(9):427-434. doi:10.1016/j.tim.2011.06.005

14. Vlainić JV, Šuran J, Vlainić T, Vukorep AL. Probiotics as an Adjuvant Therapy in Major Depressive Disorder. *Curr Neuropharmacol*. 2016;14(8):952-958. doi:10.2174/1570159x14666160526120928

15. Weiss R, Dziura J, Burgert TS, et al. Obesity and the metabolic syndrome in children and adolescents. *N Engl J Med*. 2004;350(23):2362-2374. doi:10.1056/NEJMoa031049

16. Strober W. Impact of the gut microbiome on mucosal inflammation. *Trends Immunol*. 2013;34(9):423-430. doi:10.1016/j.it.2013.07.001

17. Marcobal A, Kashyap PC, Nelson TA, et al. A metabolomic view of how the human gut microbiota impacts the host metabolome using humanized and gnotobiotic mice. *ISME J*. 2013;7(10):1933-1943. doi:10.1038/ismej.2013.89

18. Wikoff WR, Anfora AT, Liu J, et al. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc Natl Acad Sci*. 2009;106(10):3698-3703. doi:10.1073/pnas.0812874106

19. Theriot CM, Bowman AA, Young VB. Antibiotic-Induced Alterations of the Gut Microbiota Alter Secondary Bile Acid Production and Allow for Clostridium difficile Spore Germination and Outgrowth in the Large Intestine. *mSphere*. 2016;1(1):10.1128/msphere.00045-15. doi:10.1128/msphere.00045-15

20. Wu GD, Chen J, Hoffmann C, et al. Linking long-term dietary patterns with gut microbial enterotypes. *Science*. 2011;334(6052):105-108. doi:10.1126/science.1208344

21. Filippis FD, Pellegrini N, Vannini L, et al. High-level adherence to a Mediterranean diet beneficially impacts the gut microbiota and associated metabolome. *Gut*. 2016;65(11):1812-1821. doi:10.1136/gutjnl-2015-309957

22. Naveed M, BiBi J, Kamboh AA, et al. Pharmacological values and therapeutic properties of black tea (Camellia sinensis): A comprehensive overview. *Biomed Pharmacother Biomedecine Pharmacother*. 2018;100:521-531. doi:10.1016/j.biopha.2018.02.048

23. Cao J, Han J, Xiao H, Qiao J, Han M. Effect of Tea Polyphenol Compounds on Anticancer Drugs in Terms of Anti-Tumor Activity, Toxicology, and Pharmacokinetics. *Nutrients*. 2016;8(12):762. doi:10.3390/nu8120762

24. de Mejia EG, Ramirez-Mares MV, Puangpraphant S. Bioactive components of tea: Cancer, inflammation and behavior. *Brain Behav Immun*. 2009;23(6):721-731. doi:10.1016/j.bbi.2009.02.013

25. Jochmann N, Baumann G, Stangl V. Green tea and cardiovascular disease: from molecular targets towards human health. *Curr Opin Clin Nutr Metab Care*. 2008;11(6):758-765. doi:10.1097/MCO.0b013e328314b68b

26. van Duynhoven J, Vaughan EE, Jacobs DM, et al. Metabolic fate of polyphenols in the human superorganism. *Proc Natl Acad Sci*. 2011;108(supplement_1):4531-4538. doi:10.1073/pnas.1000098107

27. Axling U, Olsson C, Xu J, et al. Green tea powder and Lactobacillus plantarum affect gut microbiota, lipid metabolism and inflammation in high-fat fed C57BL/6J mice. *Nutr Metab*. 2012;9(1):105. doi:10.1186/1743-7075-9-105

28. Uffelman CN, Doenges KA, Armstrong ML, et al. Metabolomics Profiling of White Button, Crimini, Portabella, Lion's Mane, Maitake, Oyster, and Shiitake Mushrooms Using Untargeted Metabolomics and Targeted Amino Acid Analysis. *Foods*. 2023;12(16):2985. doi:10.3390/foods12162985

29. Turi KN, Michel CR, Manke J, Doenges KA, Reisdorph N, Bauer AK. Multi-Omics Analysis of Lung Tissue Demonstrates Changes to Lipid Metabolism during Allergic Sensitization in Mice. *Metabolites*. 2023;13(3):406. doi:10.3390/metabo13030406

30. Yang Y, Cruickshank C, Armstrong M, Mahaffey S, Reisdorph R, Reisdorph N. New sample preparation approach for mass spectrometry-based profiling of plasma results in improved coverage of metabolome. *J Chromatogr A*. 2013;1300:217-226. doi:10.1016/j.chroma.2013.04.030

31. Reisdorph NA, Hendricks AE, Tang M, et al. Nutrimetabolomics reveals food-specific compounds in urine of adults consuming a DASH-style diet. *Sci Rep*. 2020;10(1):1157. doi:10.1038/s41598-020-57979-8

32. Sutliff AK, Saint-Cyr M, Hendricks AE, et al. Lipidomics-Based Comparison of Molecular Compositions of Green, Yellow, and Red Bell Peppers. *Metabolites*. 2021;11(4):241. doi:10.3390/metabo11040241

33. Gilbert JA, Meyer F, Jansson J, et al. The Earth Microbiome Project: Meeting report of the "1st EMP meeting on sample selection and acquisition" at Argonne National Laboratory October 6th 2010. *Stand Genomic Sci*. 2010;3(3):249-253. doi:10.4056/aigs.1443528

34. Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol*. 2019;37(8):852-857. doi:10.1038/s41587-019-0209-9

35. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016;13(7):581-583. doi:10.1038/nmeth.3869

36. Janssen S, McDonald D, Gonzalez A, et al. Phylogenetic Placement of Exact Amplicon Sequences Improves Associations with Clinical Information. *mSystems*. 2018;3(3):e00021-18. doi:10.1128/mSystems.00021-18

37. McDonald D, Price MN, Goodrich J, et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J*. 2012;6(3):610-618. doi:10.1038/ismej.2011.139

38. Dahl EM, Neer E, Bowie KR, Leung ET, Karstens L. microshades: An R Package for Improving Color Accessibility and Organization of Microbiome Data. *Microbiol Resour Announc*. 2022;11(11):e00795-22. doi:10.1128/mra.00795-22

39. Lin H, Peddada SD. Analysis of compositions of microbiomes with bias correction. *Nat Commun*. 2020;11(1):3514. doi:10.1038/s41467-020-17041-7

40. Faith DP. Conservation evaluation and phylogenetic diversity. *Biol Conserv*. 1992;61(1):1-10. doi:10.1016/0006-3207(92)91201-3

41. Lozupone C, Knight R. UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. *Appl Environ Microbiol*. 2005;71(12):8228-8235. doi:10.1128/AEM.71.12.8228-8235.2005

42. Rideout JR, Caporaso G, Bolyen E, et al. scikit-bio. Published online August 2, 2023. doi:10.5281/ZENODO.8209901

43. Walter J, Armet AM, Finlay BB, Shanahan F. Establishing or Exaggerating Causality for the Gut Microbiome: Lessons from Human Microbiota-Associated Rodents. *Cell*. 2020;180(2):221-232. doi:10.1016/j.cell.2019.12.025

44. Apostol TM, Mnatsakanian MA. Sums of Squares of Distances in m-Space. *Am Math Mon*. 2003;110(6):516-526. doi:10.2307/3647907

45. Peres-Neto PR, Jackson DA. How well do multivariate data sets match? The advantages of a Procrustean superimposition approach over the Mantel test. *Oecologia*. 2001;129(2):169-178. doi:10.1007/s004420100720

46. Oksanen J, Simpson GL, Blanchet FG, et al. *Vegan: Community Ecology Package*.; 2024. https://vegandevs.github.io/vegan/

47. Rivera-Pinto J, Egozcue JJ, Pawlowsky-Glahn V, Paredes R, Noguera-Julian M, Calle ML. Balances: a New Perspective for Microbiome Analysis. *mSystems*. 2018;3(4):10.1128/msystems.00053-18. doi:10.1128/msystems.00053-18

48. Bergé L. Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm. *CREA Discuss Pap*. 2018;(13).

49. Loy A, Hofmann H. HLMdiag: A Suite of Diagnostics for Hierarchical Linear Models in R. *J Stat Softw*. 2014;56(5):1-28.

50. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw*. 2015;67(1):1-48. doi:10.18637/jss.v067.i01

51. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models. *J Stat Softw*. 2017;82(13):1-26. doi:10.18637/jss.v082.i13

52. Bisanz JE, Upadhyay V, Turnbaugh JA, Ly K, Turnbaugh PJ. Meta-Analysis Reveals Reproducible Gut Microbiome Alterations in Response to a High-Fat Diet. *Cell Host Microbe*. 2019;26(2):265-272.e4. doi:10.1016/j.chom.2019.06.013

53. Bag S, Mondal A, Banik A. Exploring tea (*Camellia sinensis*) microbiome: Insights into the functional characteristics and their impact on tea growth promotion. *Microbiol Res*. 2022;254:126890. doi:10.1016/j.micres.2021.126890

54. Hassani MA, Özkurt E, Seybold H, Dagan T, Stukenbrock EH. Interactions and Coadaptation in Plant Metaorganisms. *Annu Rev Phytopathol*. 2019;57:483-503. doi:10.1146/annurev-phyto-082718-100008

55. Moghadamtousi SZ, Fadaeinasab M, Nikzad S, Mohan G, Ali HM, Kadir HA. Annona muricata (Annonaceae): A Review of Its Traditional Uses, Isolated Acetogenins and Biological Activities. *Int J Mol Sci*. 2015;16(7):15625-15658. doi:10.3390/ijms160715625

56. Yu J, Elix JA, Iskander MN. Lactiflorin, a monoterpene glycoside from paeony root. *Phytochemistry*. 1990;29(12):3859-3863. doi:10.1016/0031-9422(90)85347-I

57. Hao J, Qi F, Wang H, et al. Network pharmacology-based prediction of inhibiting leukocyte recruitment and angiogenesis of total glucosides of peony against rheumatoid arthritis. *Ann Palliat Med*. 2022;11(10):3085-3101. doi:10.21037/apm-21-2203

58. Musial C, Kuban-Jankowska A, Gorska-Ponikowska M. Beneficial Properties of Green Tea Catechins. *Int J Mol Sci*. 2020;21(5):1744. doi:10.3390/ijms21051744

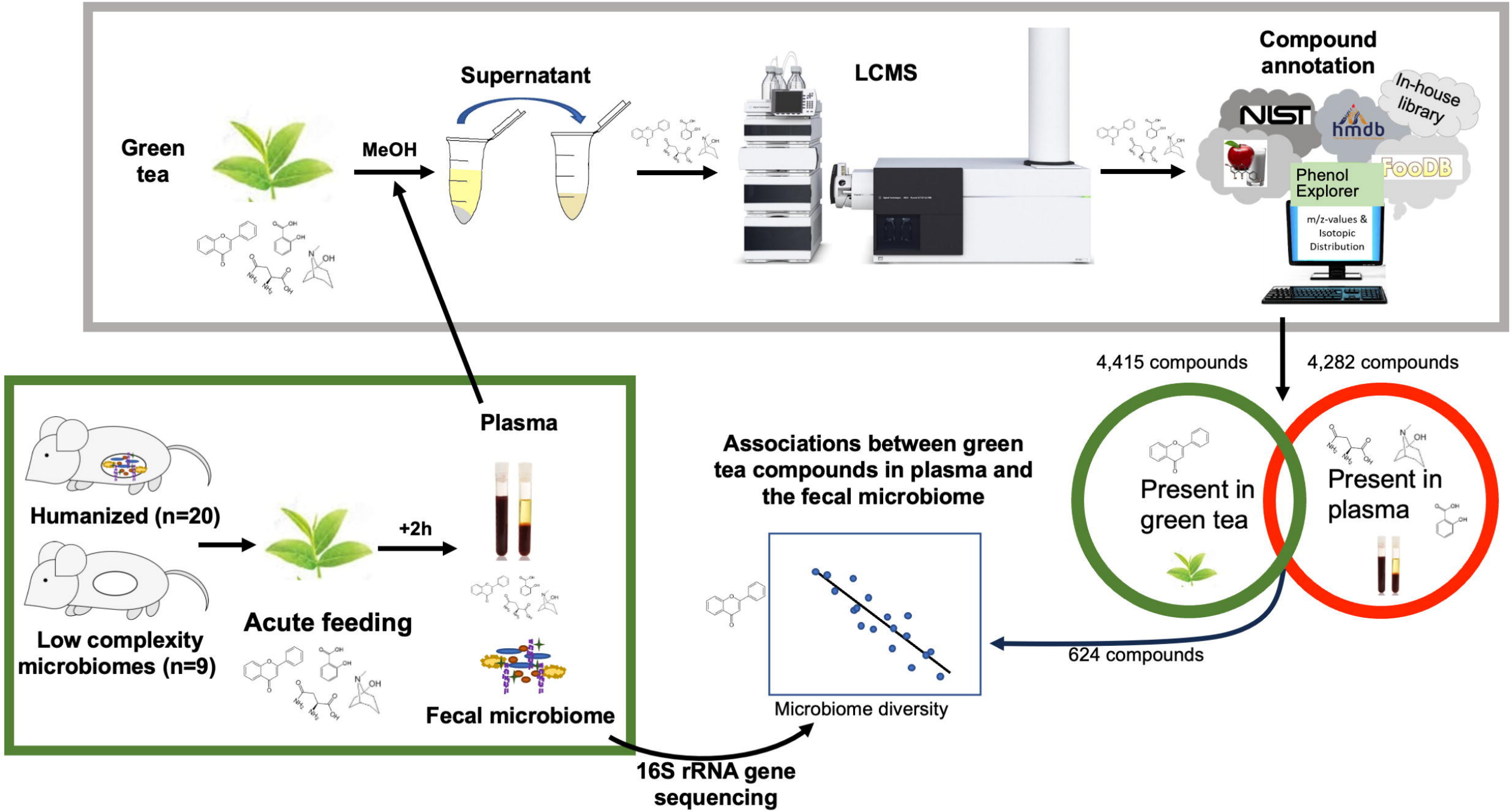
59. Masi A, Trentin AR, Agrawal GK, Rakwal R. Gamma-glutamyl cycle in plants: a bridge connecting the environment to the plant cell? *Front Plant Sci*. 2015;6. doi:10.3389/fpls.2015.00252

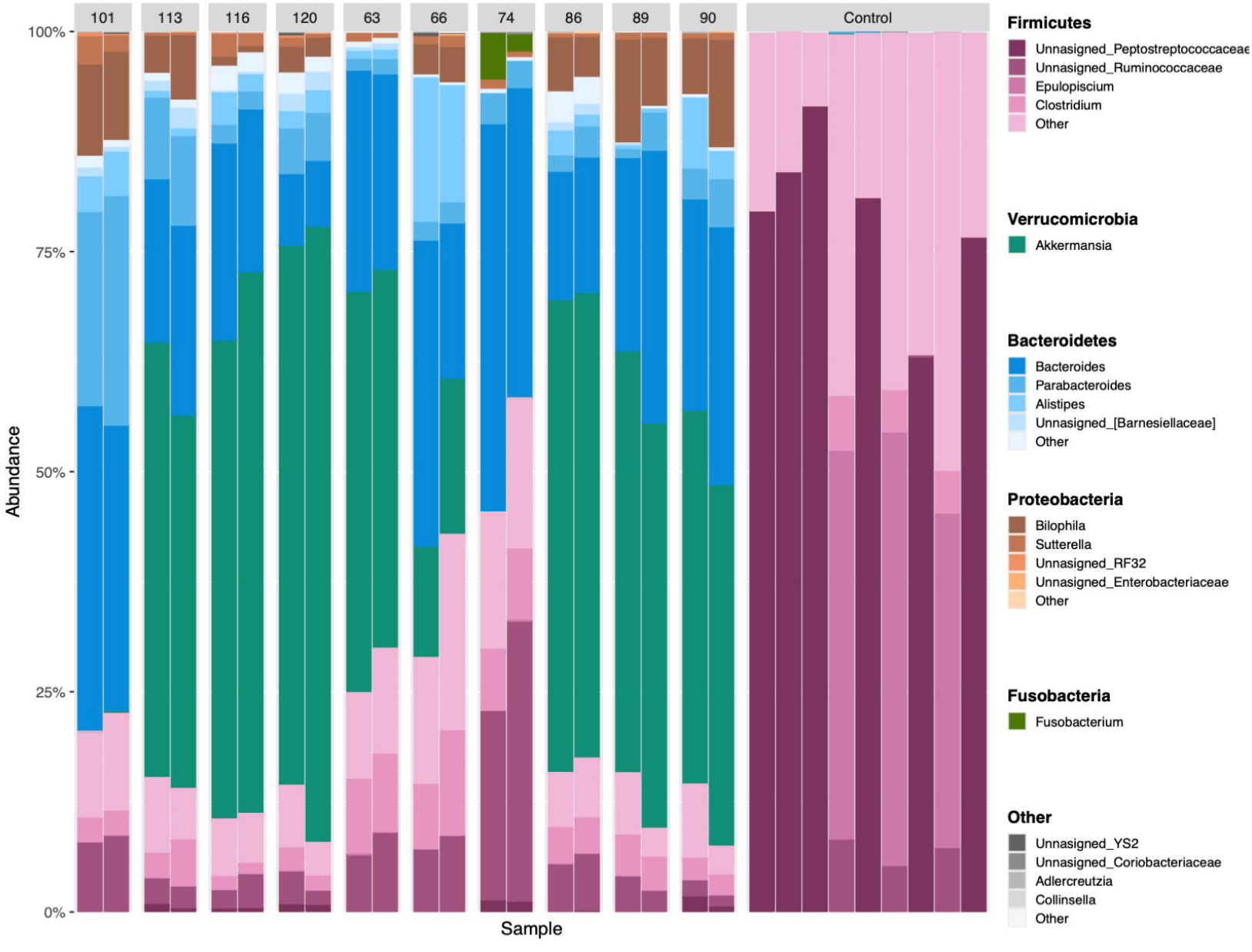
60. Saini M, Kashyap A, Bindal S, Saini K, Gupta R. Bacterial Gamma-Glutamyl Transpeptidase, an Emerging Biocatalyst: Insights Into Structure–Function Relationship and Its Biotechnological Applications. *Front Microbiol*. 2021;12. doi:10.3389/fmicb.2021.641251

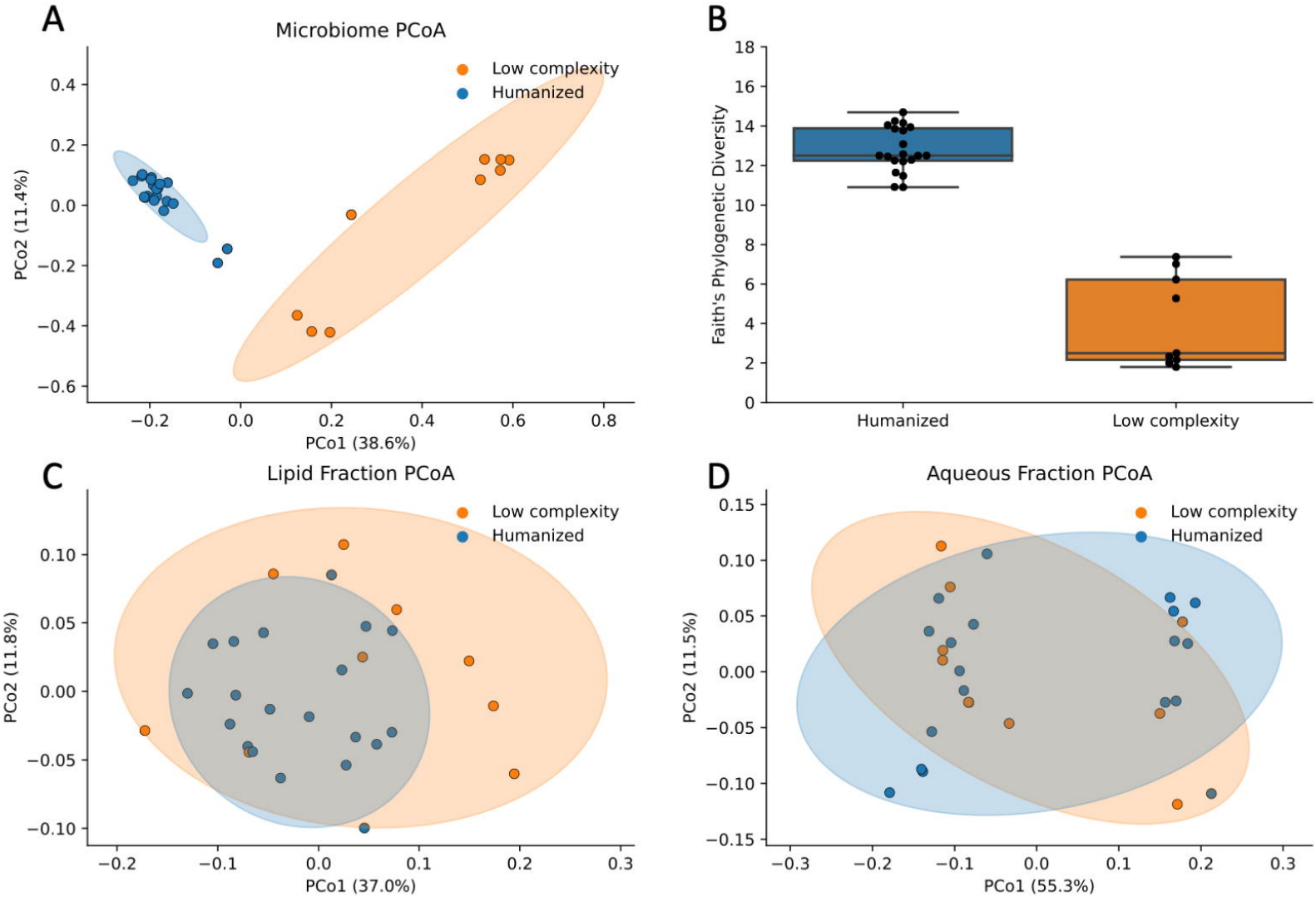
61. Le Roy T, Van der Smissen P, Paquot A, et al. Butyricimonas faecalis sp. nov., isolated from human faeces and emended description of the genus Butyricimonas. *Int J Syst Evol Microbiol*. 2019;69(3):833-838. doi:10.1099/ijsem.0.003249

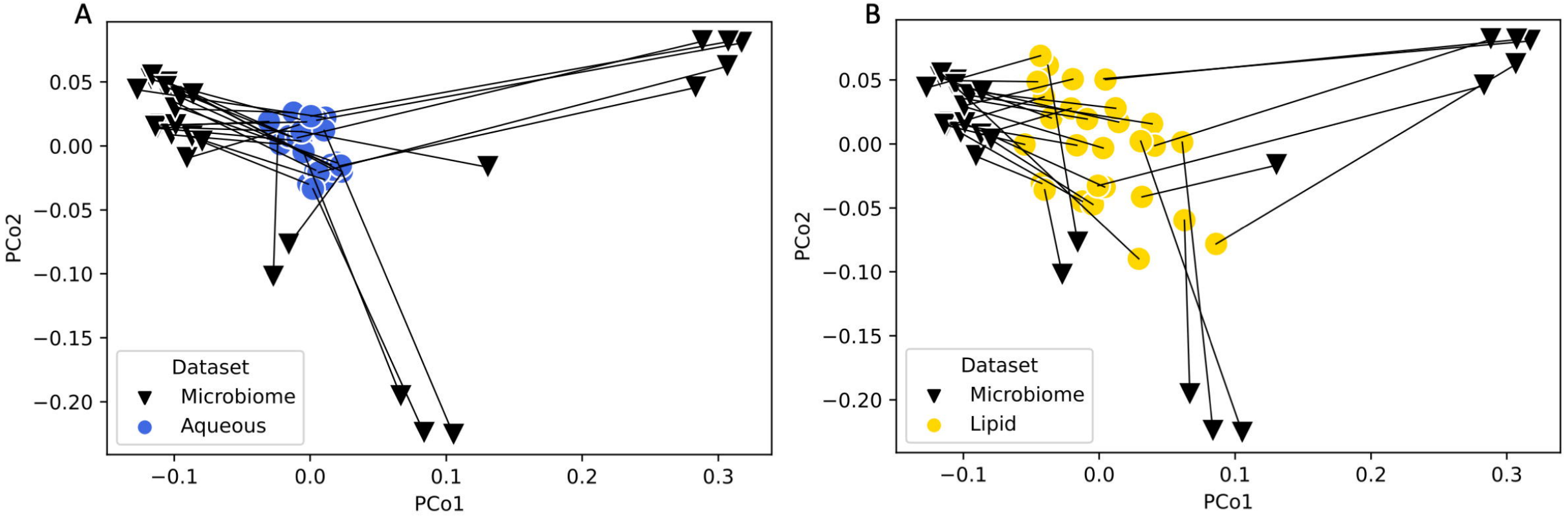
62. Yang C, Xu J, Xu X, et al. Characteristics of gut microbiota in patients with metabolic associated fatty liver disease. *Sci Rep*. 2023;13(1):9988. doi:10.1038/s41598-023-37163-4

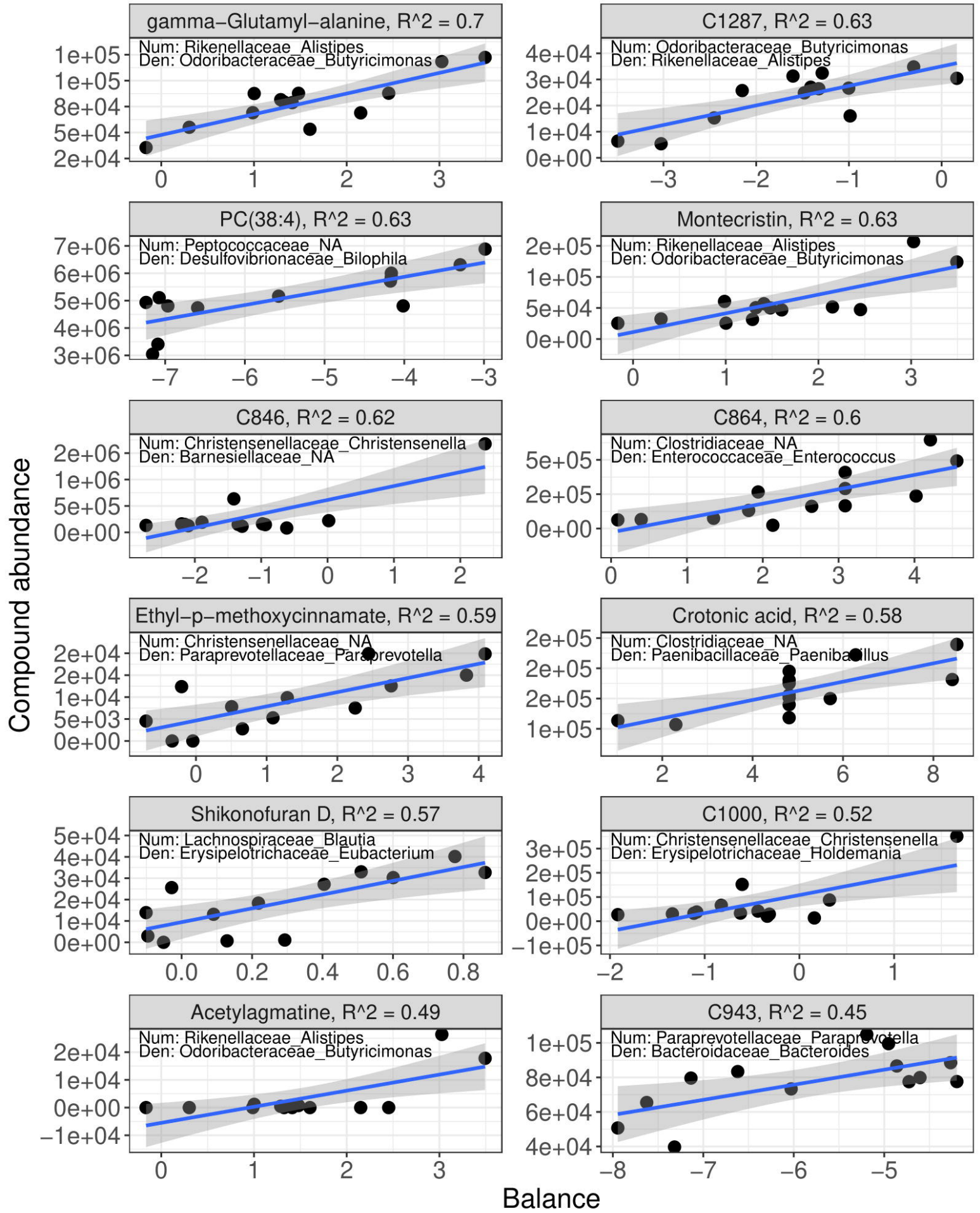
63. Vurukonda SSKP, Giovanardi D, Stefani E. Plant Growth Promoting and Biocontrol Activity of Streptomyces spp. as Endophytes. *Int J Mol Sci*. 2018;19(4):952. doi:10.3390/ijms19040952

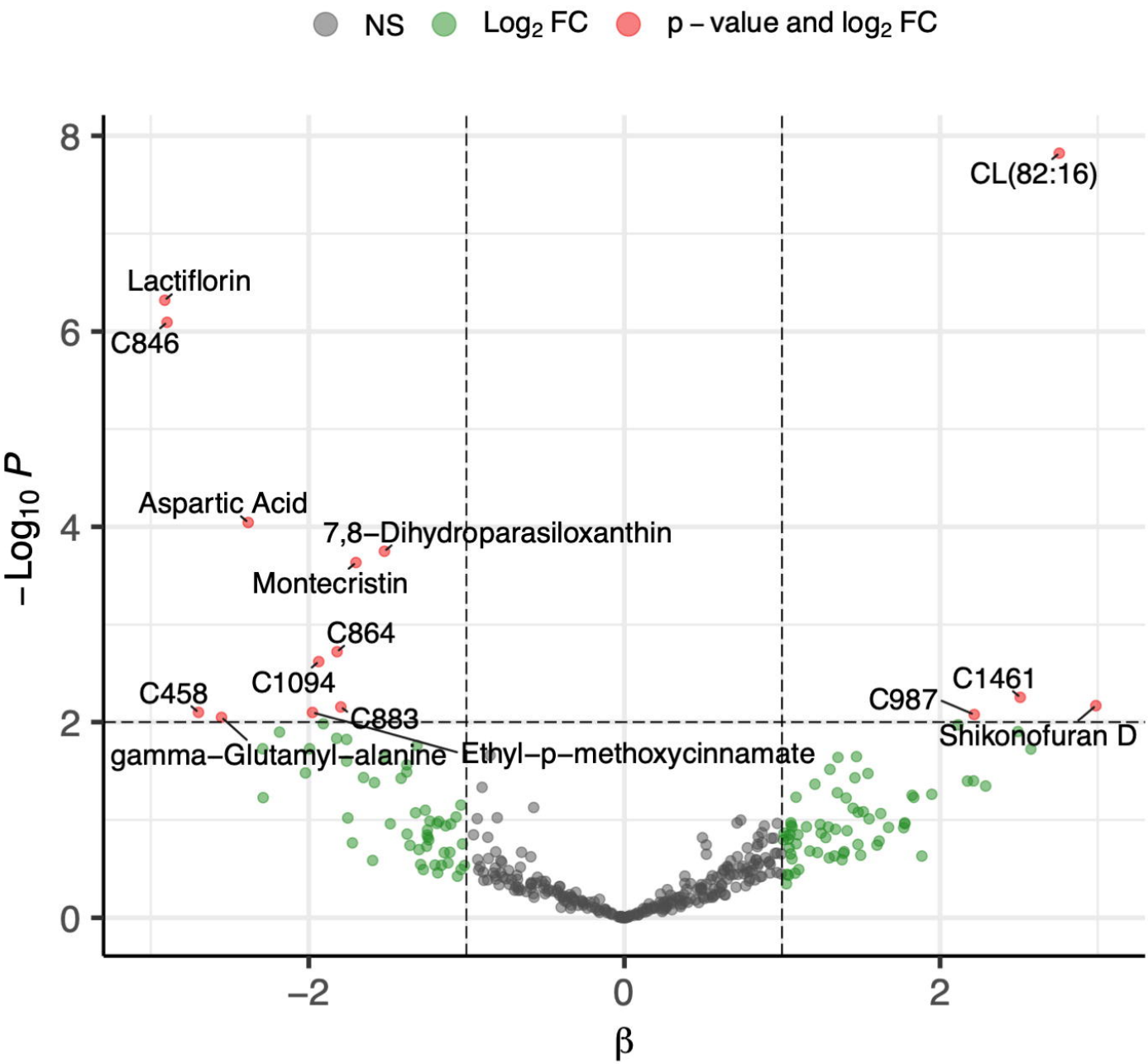

Green tea
MeOH
Supernatant
LCMS
Compound annotation
NIST
hmdb
In-house library
FooDB
Phenol Explorer
m/z-values & Isotopic Distribution
4,415 compounds
4,282 compounds
Present in green tea
Present in plasma
624 compounds
Humanized (n=20)
Low complexity microbiomes (n=9)
Acute feeding
+2h
Plasma
Fecal microbiome
16S rRNA gene sequencing
Associations between green tea compounds in plasma and the fecal microbiome
Microbiome diversity

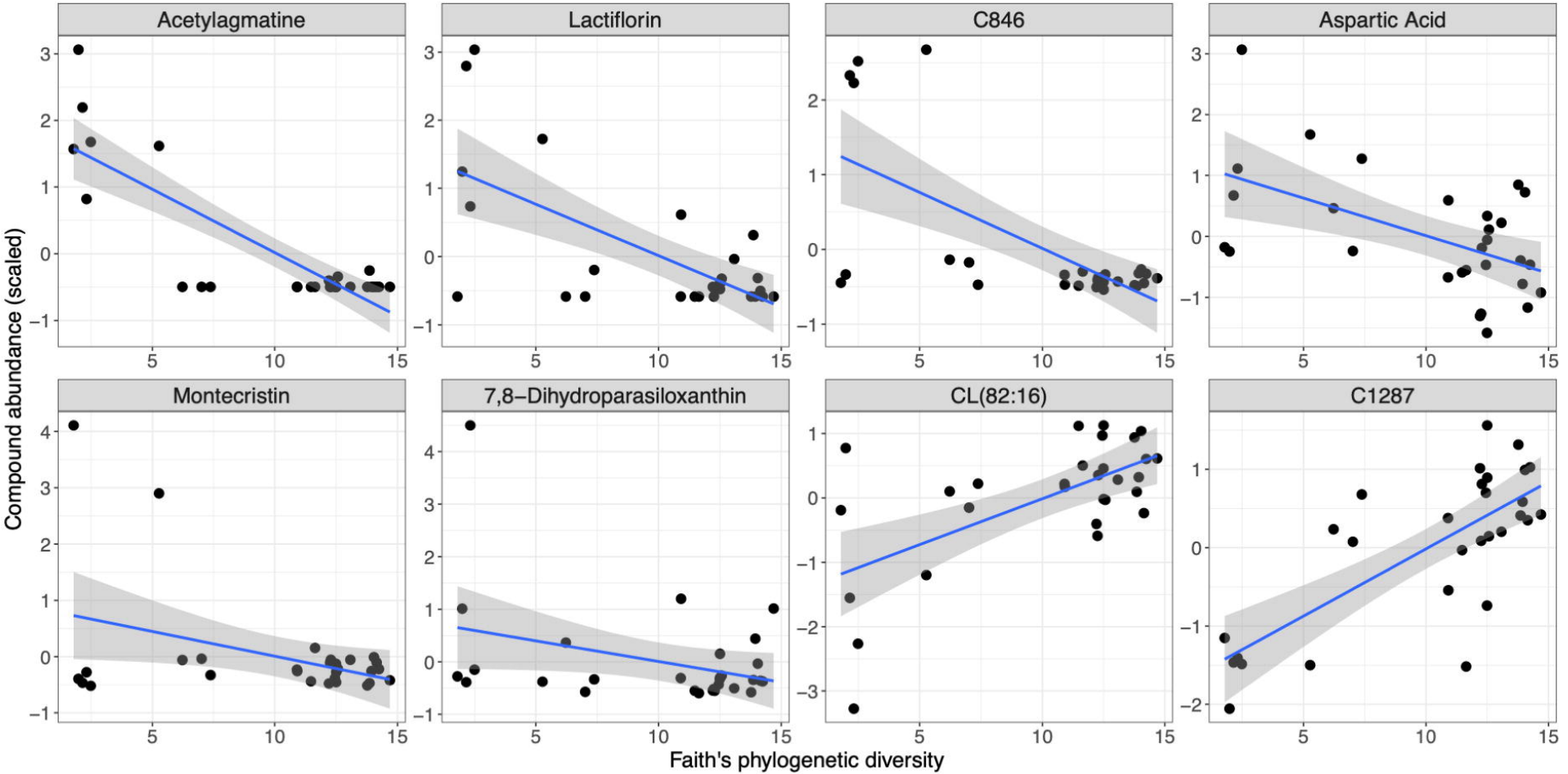

101
113
116
120
63
66
74
86
89
90
Control
100%
75%
50%
25%
0%
Abundance
Sample
Firmicutes
Unnasigned_Peptostreptococcaceae
Unnasigned_Ruminococcaceae
Epulopiscium
Clostridium
Other
Verrucomicrobia
Akkermansia
Bacteroidetes
Bacteroides
Parabacteroides
Alistipes
Unnasigned_[Barnesiellaceae]
Other
Proteobacteria
Bilophila
Sutterella
Unnasigned_RF32
Unnasigned_Enterobacteriaceae
Other
Fusobacteria
Fusobacterium
Other
Unnasigned_YS2
Unnasigned_Coriobacteriaceae
Adlercreutzia
Collinsella
Other

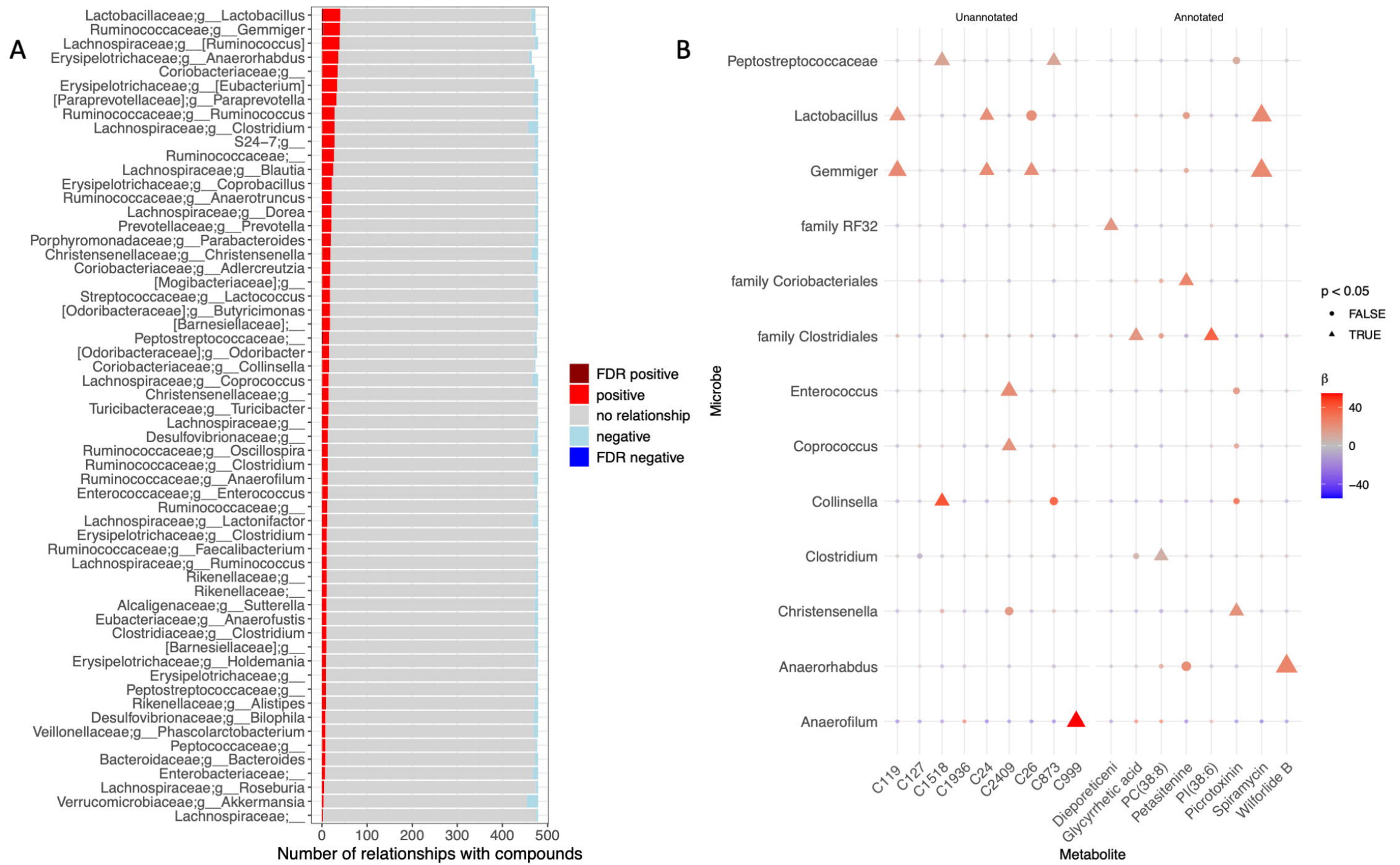

A
Microbiome PCoA
Low complexity
Humanized
PCo2 (11.4%)
PCo1 (38.6%)
B
Faith's Phylogenetic Diversity
Humanized
Low complexity
C
Lipid Fraction PCoA
Low complexity
Humanized
PCo2 (11.8%)
PCo1 (37.0%)
D
Aqueous Fraction PCoA
Low complexity
Humanized
PCo2 (11.5%)
PCo1 (55.3%)


A
B
PCo2
PCo1
0.05
0.00
−0.05
−0.10
−0.15
−0.20
−0.1
0.0
0.1
0.2
0.3
Dataset
Microbiome
Aqueous
Lipid


gamma−Glutamyl−alanine, R^2 = 0.7
Num: Rikenellaceae_Alistipes
Den: Odoribacteraceae_Butyricimonas
C1287, R^2 = 0.63
Num: Odoribacteraceae_Butyricimonas
Den: Rikenellaceae_Alistipes
PC(38:4), R^2 = 0.63
Num: Peptococcaceae_NA
Den: Desulfovibrionaceae_Bilophila
Montecristin, R^2 = 0.63
Num: Rikenellaceae_Alistipes
Den: Odoribacteraceae_Butyricimonas
C846, R^2 = 0.62
Num: Christensenellaceae_Christensenella
Den: Barnesiellaceae_NA
C864, R^2 = 0.6
Num: Clostridiaceae_NA
Den: Enterococcaceae_Enterococcus
Ethyl−p−methoxycinnamate, R^2 = 0.59
Num: Christensenellaceae_NA
Den: Paraprevotellaceae_Paraprevotella
Crotonic acid, R^2 = 0.58
Num: Clostridiaceae_NA
Den: Paenibacillaceae_Paenibacillus
Shikonofuran D, R^2 = 0.57
Num: Lachnospiraceae_Blautia
Den: Erysipelotrichaceae_Eubacterium
C1000, R^2 = 0.52
Num: Christensenellaceae_Christensenella
Den: Erysipelotrichaceae_Holdemania
Acetylagmatine, R^2 = 0.49
Num: Rikenellaceae_Alistipes
Den: Odoribacteraceae_Butyricimonas
C943, R^2 = 0.45
Num: Paraprevotellaceae_Paraprevotella
Den: Bacteroidaceae_Bacteroides
Compound abundance
Balance


NS
Log2 FC
p – value and log2 FC
CL(82:16)
Lactiflorin
C846
Aspartic Acid
7,8–Dihydroparasiloxanthin
Montecristin
C864
C1094
C458
C883
gamma–Glutamyl–alanine
Ethyl–p–methoxycinnamate
C987
C1461
Shikonofuran D
– Log10 P
β
8
6
4
2
0
–2
0
2


Acetylagmatine
Lactiflorin
C846
Aspartic Acid
Montecristin
7,8−Dihydroparasiloxanthin
CL(82:16)
C1287
Compound abundance (scaled)
Faith's phylogenetic diversity


A
Lactobacillaceae;g__Lactobacillus
Ruminococcaceae;g__Gemmiger
Lachnospiraceae;g__[Ruminococcus]
Erysipelotrichaceae;g__Anaerorhabdus
Coriobacteriaceae;g__
Erysipelotrichaceae;g__[Eubacterium]
[Paraprevotellaceae];g__Paraprevotella
Ruminococcaceae;g__Ruminococcus
Lachnospiraceae;g__Clostridium
S24–7;g__
Ruminococcaceae;__
Lachnospiraceae;g__Blautia
Erysipelotrichaceae;g__Coprobacillus
Ruminococcaceae;g__Anaerotruncus
Lachnospiraceae;g__Dorea
Prevotellaceae;g__Prevotella
Porphyromonadaceae;g__Parabacteroides
Christensenellaceae;g__Christensenella
Coriobacteriaceae;g__Adlercreutzia
[Mogibacteriaceae];g__
Streptococcaceae;g__Lactococcus
[Odoribacteraceae];g__Butyricimonas
[Barnesiellaceae];__
Peptostreptococcaceae;__
[Odoribacteraceae];g__Odoribacter
Coriobacteriaceae;g__Collinsella
Lachnospiraceae;g__Coprococcus
Christensenellaceae;g__
Turicibacteraceae;g__Turicibacter
Lachnospiraceae;g__
Desulfovibrionaceae;g__
Ruminococcaceae;g__Oscillospira
Ruminococcaceae;g__Clostridium
Ruminococcaceae;g__Anaerofilum
Enterococcaceae;g__Enterococcus
Ruminococcaceae;g__
Lachnospiraceae;g__Lactonifactor
Erysipelotrichaceae;g__Clostridium
Ruminococcaceae;g__Faecalibacterium
Lachnospiraceae;g__Ruminococcus
Rikenellaceae;g__
Rikenellaceae;__
Alcaligenaceae;g__Sutterella
Eubacteriaceae;g__Anaerofustis
Clostridiaceae;g__Clostridium
[Barnesiellaceae];g__
Erysipelotrichaceae;g__Holdemania
Erysipelotrichaceae;g__
Peptostreptococcaceae;g__
Rikenellaceae;g__Alistipes
Desulfovibrionaceae;g__Bilophila
Veillonellaceae;g__Phascolarctobacterium
Peptococcaceae;g__
Bacteroidaceae;g__Bacteroides
Enterobacteriaceae;__
Lachnospiraceae;g__Roseburia
Verrucomicrobiaceae;g__Akkermansia
Lachnospiraceae;__
0
100
200
300
400
500
Number of relationships with compounds
FDR positive
positive
no relationship
negative
FDR negative
B
Unannotated
Annotated
Peptostreptococcaceae
Lactobacillus
Gemmiger
family RF32
family Coriobacteriales
family Clostridiales
Enterococcus
Coprococcus
Collinsella
Clostridium
Christensenella
Anaerorhabdus
Anaerofilum
Microbe
C119
C127
C1518
C1936
C24
C2409
C26
C873
C999
Dieporeticeni
Glycyrrhetic acid
PC(38:8)
Petasitenine
PI(38:6)
Picrotoxinin
Spiramycin
Wilforlide B
Metabolite
p < 0.05
FALSE
TRUE
β
40
0
−40