Supplemental Methods

Study design. The study protocol was approved by the Institutional Review Board of the University of Texas MD Anderson Cancer Center and is registered under the *ClinicalTrials.gov identifier* **NCT02843425**. Participants were fully informed of the voluntary nature and risks of the study and signed consent forms before engaging in any study procedures. This dietary study was considered low risk and Data Safety Monitoring Committee exempt and overseen by the PI and study physicians. Following initiation of a pilot in July 2016 among patients with a high-risk BMI and/or waist circumference[1] and history of precancerous colorectal polyps, the protocol was expanded to patients with a history of colorectal cancer in May 2017. Shortly thereafter, Tropical Storm Harvey hit the study site Houston, TX in August 2017.[2] Recruitment closed in January 2020. Full details of the study design, including recruitment setting, full inclusion and exclusion criteria of the BE GONE trial can be found in our published protocol[3] and under NCT02843425. Briefly, to establish the basal diet and microbiome and to track compliance study procedures prior to randomization, eligible and consented individuals were asked to provide a stool sample and to complete diet, physical activity and anthropometric assessments during a 4-week run-in and equilibration period. Participants who completed the run-in were randomized to either begin the intervention diet (usual diet, add study beans) or to continue the control diet (usual diet, avoid dry beans) for 8 weeks according to blocks of no use vs. regular use of statins and/or metformin, as they were commonly prescribed in our target population. Participants were free-living and able to choose and prepare their own meals (some recipes provided with close follow-up and counseling support from the study dietitian. The intervention diet consisted of ½ cup (1 serving) of pressure-cooked, canned navy beans (**see embedded table**) over a 2-week ramp-up period followed by 1 cup (2 servings) per day for an additional 6 weeks. Organic canned navy beans (from a common

Nutrient profile (values>0) of 100 grams of pressure-cooked navy beans in water and sea salt: standard reference from the Nutrition Data System for Research (NDSR)	
Food Name: vegetables, beans, navy beans, canned - drained, low sodium	
Gram Amount of Food	100
FDA Serving Size (Grams)	90
% Calories from Carbohydrate	75.731
% Calories from Fat	3.707
% Calories from MUFA	0.849
% Calories from Protein	20.399
% Calories from PUFA	2.929
% Calories from SFA	0.586
Alanine (g)	0.369
Arginine (g)	0.415
Ash (g)	1.3
Aspartic Acid (g)	1.056
Available Carbohydrate (g)	15.55
Betaine (mg)	0.1
Calcium (mg)	69
Cholesterol to Saturated Fatty Acid Index	0.099
Choline (mg)	44.7
Copper (mg)	0.21
Cystine (g)	0.076
Daidzein (mg)	0.004
Delta-Tocopherol (mg)	0.09
Dietary Folate Equivalents (mcg)	140
Energy (kcal)	140
Energy (kj)	585.76
Gamma-Tocopherol (mg)	1.28
Genistein (mg)	0.084
Glutamic Acid (g)	1.259
Glycemic Index (bread reference)	44.33
Glycemic Index (glucose reference)	31
Glycemic Load (bread reference)	6.893
Glycemic Load (glucose reference)	4.821
Glycine (g)	0.326
Histidine (g)	0.206
Inositol (g)	0.065
Insoluble Dietary Fiber (g)	9.82
Iron (mg)	2.36
Isoleucine (g)	0.387
Lariciresinol (mcg)	2.8
Leucine (g)	0.7
Lysine (g)	0.52
Magnesium (mg)	53
Manganese (mg)	0.527
Matairesinol (mcg)	0.2

generic/store brand) in water and sea salt were purchased in two large batches. Participants logged their bean intake daily, including the frequency, amount and manner in which they were consumed. Between one set of visits, 2 of the 48 intent-to-treat participants required navy bean crackers made by our research kitchen in lieu of the canned beans (for travel/other reasons). At week 8, all participants who completed the intervention diet immediately crossed over to the control diet and vice versa. Participants were asked to follow the control diet during the equilibration period and during one of two crossover sequences. During the control period, participants were instructed to follow their usual diet without any dry beans, including pinto beans, black beans, kidney beans and others. Similarly, during the intervention, participants were instructed not to consume non-study beans. For vast majority of participants, reported *pre-study* intake of dry beans and other legumes was quite low (Table S2). Side effects/adverse events, medications, health status, usual dietary intake, physical activity, body weight and blood pressure were monitored throughout the trial. Study visits: Randomized participants attended a total of 5 in-person clinic visits every 4 weeks (Fig. S1). For each visit, participants provided a stool sample and fasting blood sample. Weight, waist circumference and blood pressure were measured at each clinic visit in duplicate (or triplicate if the first two measures were not within precise agreement). To establish baseline status and to monitor deviations throughout the study period, medications and changes in health status, along with data provided on usual dietary habits and physical activity levels, were carefully reviewed at each visit. Usual diet and lifestyle assessment: Participant's diet was assessed and monitored throughout the study via biweekly NCI-Automated Self-Administered 24HR (ASA-24)[4] in conjunction with a "past month" web-based NCI DHQ (every 4 weeks)[5]. A snapshot of the usual diet during the month prior to enrollment (pre-visit) was assessed via the DHQ and used to compare the pre-study diet of enrolled individuals who withdrew pre or post-randomization versus those who completed the study (Table S2). Parameters of usual diet at randomization (V0; Table S1) were obtained via multiple 24HR collected during the 4-week equilibration, except for intake of alcohol and any legumes, which were obtained from the DHQ due to the episodic nature of their consumption. Similarly, 24HR methods were the basis for the longitudinal analysis of macronutrients and key dietary variables monitored over time (V0 to V4). Dietary variables were standardized for total energy intake and the Healthy Eating Index 2015 was evaluated as a measure of overall diet quality[6]. Physical activity levels were assessed and monitored monthly via validated long (enrollment/pre-visit) and short (V0 to V4) versions of the IPAQ[7,8] administered via REDCap. The intensity (dose and duration) of physical activity was estimated and defined as the following: "vigorous" >300 mins moderate intensity activity per week or >150 mins vigorous intensity activity per week or a combination; "moderate" 150-300 mins moderate intensity per week; and "inactive" <150 mins of moderate intensity activity per week.

Stool and fasting blood sample collection. For V0 to V4, an in-home, fresh-frozen stool sample collection kit, similar to that used in the Human Microbiome Project (HMP)[9] and refined in our previous studies[10], was provided to participants with instructions to collect a stool sample as close as possible to each scheduled clinic visit (with reminder emails and calls). Following the pre-visit and any other times the patient was unable to come to the clinic, an OMNIgene GUT (OMR-200, DNA Genotek, Ottawa, Canada) mailable kit was provided. A total of 6 (2.5%) of the 240 samples (5 x 48 individuals, intent-to-treat) in the primary analysis were provided via OMNIgene kits and similar to prior studies extensively comparing these two sample collection methods at an individual and group level[11-13], we did not observe systematic differences in

candidate taxa or other measures. Fasting blood was collected by trained phlebotomists at each morning in-clinic visit and processed on the day of collection by the MD Anderson Clinical and Translational Research Center [a total of 11 (4.6%) of 240 scheduled blood draws were missed during the study due to weather and other factors]. A fresh serum aliquot was immediately sent to LabCorp, a CLIA-certified laboratory, for lipoprotein analysis. Stool and plasma aliquots were stored at -80 °C until microbiome sequencing and metabolomic analysis.

Microbiome sequencing and data processing. The Alkek Center for Metagenomics and Microbiome Research at Baylor College of Medicine performed 16S rRNA gene sequencing and data processing via their established pipeline[14-16]. Briefly, total genomic DNA was extracted using the Qiagen MagAttract PowerSoil DNA kit. The 16S rRNA v4 region was amplified via polymerase chain reaction with primers 515F (GTGCCAGCMGCCGCGGTAA) and 806R (GGACTACHVGGGTWTCTAAT) that additionally contained sequencing adapters and a single-end barcode (reverse primer), allowing for pooling and direct sequencing of resulting amplicons. Libraries were sequenced on the Illumina MiSeq platform using the 2 × 250 paired-end protocol, yielding reads that overlapped almost completely, targeting at least 10,000 reads per sample[17]. Sequence reads were demultiplexed, quality-filtered, and subsequently merged using USEARCH (version 7.0.1090). Merge parameters required minimum merged length of 252 bp, overlap of at least 50 bp, truncation quality >5, and zero differences in the overlapping region. The merged files were filtered further allowing for a maximum expected error of 0.05. The resulting readset was dereplicated and iteratively clustered using UPARSE and bundled into operational taxonomic units at a similarity cutoff value of 97%[18]. Taxonomy assignment was performed against a modified version of the SILVA v128 database that contained only the V4 region[19]. Abundances were recovered by mapping the demultiplexed reads to the UPARSE OTUs. Alpha- and beta-diversity were calculated using the Agile Toolkit for Incisive Microbiome Analyses (ATIMA2)[20].

The same DNA extracts used for 16S sequencing (above) were used for metagenomic shotgun sequencing; however, 4 samples from the intent-to-treat 16S group did not pass QC standards for library preparation. Individual libraries constructed from each sample's extracted total gDNA were sequenced on the NovaSeq 6000 platform (Illumina) using the 2x150 bp pair-end read protocol. Raw data files in binary base call (BCL) format are converted into FASTQs and demultiplexed based on the dual-index barcodes using the Illumina 'bcl2fastq' software. Demultiplexed raw FASTQ sequences are processed using BBDuk (sourceforge.net/projects/bbmap/; BBMap version 38.82) to quality trim, remove Illumina adapters and filter PhiX reads. Trimming parameters are set to a k-mer length of 19 and a minimum Phred quality score of 25. Reads with a minimum average Phred quality score below 23 and length shorter than 50 bp after trimming are discarded. The trimmed FASTQS are mapped to a combined PhiX (standard Illumina spike in) and host reference genome database using a two-step BBTools approach (sourceforge.net/projects/bbmap/; BBMap version 38.82). Briefly, the trimmed reads are first processed through the bloomfilter script, with a strict k=31 to remove reads identified as human. The remaining reads are mapped to the reference genome with BBMap using a k-mer length of 15, the bloom filter enabled, and fast search settings in order to determine and remove host/PhiX reads. Taxonomic profiling of the sequenced samples is done using MetaPhlAn3[21]. Processed fastq reads are first mapped against the MetaPhlAn3 marker gene database (mpa_v30_CHOCOPhlAn_201901) using bbmap (3) with the bloom filter

enabled and fast search settings. Each sample is run through the metaphlan.py script to generate the kingdom-specific taxonomic profile per sample, using the flag to generate relative abundances and estimated read counts. The MetaPhlAn3 utility scripts are employed to merge the output for all samples into a single sample per taxon table for each kingdom and relative abundance and estimated read count output. Finally, the tables are converted into biom-format for further statistical analysis.

Functional profiling of the microbial community was done using HUMAnN3[22]. The standard recommended workflow was followed with modifications to the nucleotide and translated alignment steps. Briefly, nucleotide alignment is performed using bbmap with the bloom filter enabled (bloomk=22) and fast search settings generating a HUMAnN3 compatible SAM file output. The translated alignment step is performed using diamond (6; version 0.9.26). This creates the default pathway abundance and coverage tables, as well as gene family abundance output files per sample. Post-processing of the per-sample tables is done using a combination of HUMAnN3 utility scripts and in-house code designed to clean up the tables for better readability. The three default outputs are each merged across samples using 'humann_join_table' script. Merged pathway abundance and genefamilies tables are also normalized to relative abundances using 'humann_renorm_table' script. All tables are split into stratified (by Taxa) tables and unstratified (metagenome) tables. Additional tables are generated by regrouping the UniRef90 genefamilies into other functional categories using the 'humann_regroup_table' script with 'uniref90 to ko' and 'uniref90 to ec' utility mapping databases. The output tables for KEGG Orthogroups (KOs), molecular functions represented in terms of functional orthologs, and Level-4 enzyme commission (EC), categories of numerical nomenclature that classifies enzymes based on the overall reaction catalyzed, are merged across samples as described above. Using the legacy KEGG databases included with HUMAnN[23], genefamilies outputs are converted to KEGG Pathways (a collection of manually drawn pathway maps representing our knowledge of the molecular interaction, reaction and relation networks) and KEGG Modules (manually defined functional units of gene sets and reaction sets). This is done by processing the genefamilies default output tables through the 'humann' script specifying the 'pathways-database' for the KEGG Pathways and KEGG Modules. The abundance and coverage output tables are merged as described above.

Metabolomic sample handling and data processing

Sample Extraction of Primary Metabolites and Biogenic Amines. Plasma metabolites were extracted from pre-aliquoted biospecimens (15µL) with 45µL of LCMS grade methanol (ThermoFisher) in a 96-well microplate (Eppendorf). Plates were heat sealed, vortexed for 5 min at 750 rpm, and centrifuged at 2000 × g for 10 minutes at room temperature. The supernatant (30µL) was carefully transferred to a 96-well plate, leaving behind the precipitated protein. The supernatant was further diluted with 60µL of 100mM ammonium formate, pH3 (Fisher Scientific). For Hydrophilic Interaction Liquid Chromatography (HILIC) positive ion analysis, 15µL of the supernatant and ammonium formate mix were diluted with 195µL of 1:3:8:144 water (GenPure ultrapure water system, Thermofisher): LCMS grade methanol (ThermoFisher): 100mM ammonium formate, pH3 (Fisher Scientific): LCMS grade acetonitrile (ThermoFisher). For the HILIC negative ion analysis, 15µL of the supernatant and ammonium formate mix were diluted with 90µL of LCMS grade acetonitrile (ThermoFisher). For C18 analysis, 15µL of the

supernatant and ammonium formate mix were diluted with 90µL water (GenPure ultrapure water system, ThermoFisher) for positive and negative ion modes, respectively. Each sample solution was transferred to 384-well microplate (Eppendorf) for LCMS analysis.

Sample Extraction of Complex Lipids. Pre-aliquoted plasma samples (10µL) were extracted with 30µL of LCMS grade 2-propanol (ThermoFisher) in a 96-well microplate (Eppendorf). Plates were heat sealed, vortexed for 5min at 750 rpm, and centrifuged at 2000 x g for 10 minutes at room temperature. The supernatant (10µL) was carefully transferred to a 96-well plate, leaving behind the precipitated protein. The supernatant was further diluted with 90µL of 1:3:2 100mM ammonium formate, pH3 (Fischer Scientific): LCMS grade acetonitrile (ThermoFisher): LCMS grade 2-propanol (ThermoFisher) and transferred to a 384-well microplate (Eppendorf) for lipids analysis using LCMS.

Untargeted Analysis of Primary Metabolites and Biogenic Amines. Untargeted metabolomics analysis was conducted on Waters Acquity™ UPLC system with 2D column regeneration configuration (I-class and H-class) coupled to a Xevo G2-XS quadrupole time-of-flight (qTOF) mass spectrometer. Chromatographic separation was performed using HILIC (Acquity™ UPLC BEH amide, 100 Å, 1.7 µm 2.1× 100mm, Waters Corporation, Milford, U.S.A) and C18 (Acquity™ UPLC HSS T3, 100 Å, 1.8 µm, 2.1×100mm, Water Corporation, Milford, U.S.A) columns at 45°C.

Quaternary solvent system mobile phases were (A) 0.1% formic acid in water, (B) 0.1% formic acid in acetonitrile and (D) 100mM ammonium formate, pH 3. Samples were separated using the following gradient profile: for the HILIC separation a starting gradient of 95% B and 5% D was linearly changed to 70% A, 25% B and 5% D over a 5 min period at 0.4mL/min flow rate, and to 100% A over 1 min, followed by another 1 min isocratic gradient at 100 % A at 0.4mL/min flow rate to initiate the starting gradient for the next C18 run. For C18 separation, the chromatography gradient was as follows: starting conditions, 100% A, with a linear change to 5% A, 95% B over a 5 min period at 0.4 mL/min flow rate, reverted back to 95% B, 5% D over 1 min, and then followed by 1 min isocratic gradient at 95% B, 5% D at 0.4 mL/min for the next HILIC run.

A binary pump was used for column regeneration and equilibration. The solvent system mobile phases were (A1) 100mM ammonium formate, pH 3, (A2) 0.1% formic in 2-propanol and (B1) 0.1% formic acid in acetonitrile. The HILIC column was stripped using 90% A2 for 5 min at 0.25 mL/min flow rate, followed by a 2 min equilibration using 100% B1 at 0.3mL/min flow rate. Reverse phase C18 column regeneration was performed using 95% A1, 5% B1 for 2 min followed by column equilibration using 5% A1, 95% B1 for 5 min at 0.4mL/min flow rate.

Untargeted Analysis of Complex Lipids. For the lipidomic assay, untargeted metabolomics analysis was conducted on a Waters Acquity™ UPLC system coupled to a Xevo G2-XS quadrupole time-of-flight (qTOF) mass spectrometer. Chromatographic separation was performed using a C18 (Acquity™ UPLC HSS T3, 100 Å, 1.8 µm, 2.1×100mm, Water Corporation, Milford, U.S.A) column at 55°C. The mobile phases were (A) water, (B) acetonitrile, (C) 2-propanol and (D) 500mM ammonium formate, pH 3. A starting elution gradient of 20% A, 30% B, 49% C and 1% D was linearly changed to 4% A, 14% B, 81% C and

1 % D for 4.5 min, followed by isocratic elution at 4% A,14% B, 81% C and 1%D for 2.1 min and column equilibration with initial conditions for 1.4 min.

Mass Spectrometry Data Acquisition. Mass spectrometry data was acquired using 'sensitivity' mode in positive and negative electrospray ionization mode within 50-800 Da range for primary metabolites and 100-2000 Da for complex lipids. For the electrospray acquisition, the capillary voltage was set at 1.5kV (positive), 3.0kV (negative), sample cone voltage 30V, source temperature at 120°C, cone gas flow 50L/h and desolvation gas flow rate of 800L/h with scan time of 0.5 sec in continuum mode. Leucine Enkephalin; 556.2771 Da (positive) and 554.2615 Da (negative) was used for lockspray correction and scans were performed at 0.5sec. The injection volume for each sample was 3µL for complex lipids, and 6µL for primary metabolites. The acquisition was carried out with instrument auto gain control to optimize instrument sensitivity over the samples acquisition time.

Metabolomics data processing. Untargeted metabolomic analyses were conducted on a Waters Acquity™ UPLC system with 2D column regeneration (I-class and H-class) coupled to a Xevo G2-XS quadrupole time-of-flight (qTOF) mass spectrometer as previously described.[24-27]

Data were processed using Progenesis QI (Nonlinear, Waters). Peak picking and retention time alignment of LC-MS and MSe data were performed using Progenesis QI software (Nonlinear, Waters). Data processing and peak annotations were performed using an in-house automated pipeline as previously described.[24-27] Annotations were determined by matching accurate mass and retention times using customized libraries created from authentic standards and by matching experimental tandem mass spectrometry data against the NIST MSMS, LipidBlast or HMDB v3 theoretical fragmentations; for complex lipids retention time patterns characteristic of lipid subclasses was also considered. To correct for injection order drift, each feature was normalized using data from repeat injections of quality control samples collected every 10 injections throughout the run sequence. Measurement data were smoothed by Locally Weighted Scatterplot Smoothing (LOESS) signal correction (QC-RLSC) as previously described. Values are reported as ratios relative to the median of historical quality control reference samples run with every analytical batch for the given analyte.[24-27]

Immune and inflammatory proteomic markers. Plasma specimens were analyzed via the Olink® Target 96 Inflammation and Immune Response panels, with results reported as units standardized to each individual assay. The Olink proximity extension immunoassay technology uses a dual recognition DNA-coupled immunoassay that rapidly allows protein identification and relative quantification with high sensitivity and specificity. Olink-generated proteomics data were background corrected and normalized to a Normalized Protein Expression (NPX) scale. NPX values represent relative quantification, meaning that protein values can be compared for the same protein across samples[28].

Statistical analysis. Characteristics of study participants were described with mean and standard deviation for continuous variables and frequencies and percentages for categorical variables. We examined the differences between groups at enrollment (intent-to-treat vs. withdrew) and at randomization (intervention first vs. second), by Pearson Chi-Square, Fisher's exact or ANOVA test, as appropriate. Bacterial alpha diversity was assessed via the inverse

Simpson index, an indicator of the richness in a community with uniform evenness; and the Shannon index, another community diversity measure representing both the richness and the evenness of the different taxa within a sample. Differences in alpha-diversity by enrollment and randomization status were assessed via the Mann-Whitney U test. Beta diversity was analyzed via permutational multivariate analysis of variance (PERMANOVA implemented in "adonis" function in statistical software R package "vegan") with principal coordinate analysis (PCoA) plots to visualize the dissimilarity of the community composition using the weighted Jaccard distance as the distance matrix[29].

To quantify microbiome composition changes during the study, we conducted a longitudinal analysis of alpha diversity (including Shannon diversity and inverse Simpson index) and the relative abundance of individual taxa on the natural log scale (slope=0 under the null hypothesis) across the trial sequence using generalized linear mixed models (SAS, 'Proc Glimmix') with random intercept. As the usual diet control preceded the intervention diet in both arms, the primary outcome focused on changes during the on-intervention period in the full trial cohort [n=48; intervention first V0 to V2 (n=28) combined with intervention second V2 to V4 (n=20)]. We further characterized these findings across the crossover study sequence in each of the two randomization groups separately to assess the stability of these measures during the "return to control" [V2 to V4 (n=28)] and "remain on control" [V0 to V2 (n=20)] period, as well as the consistency of on-intervention findings in each group. The full trial sequence includes for the bean intervention first: 4 weeks on-intervention (mid-point), 8 weeks on-intervention (primary outcome), 4 weeks return to control, 8 weeks return to control; and for the bean intervention second: 4 weeks remain on control, 8 weeks remain on control, 4 weeks on-intervention (mid-point), 8 weeks on-intervention (primary outcome). These methods were also applied to other study variables such as dietary fiber intake, total energy intake, circulating profiles and biometric variables.

To maximize translation and to address issues with zero-inflated compositional data we applied a two-pronged method of analysis with regard to the relative abundance of individual taxa. First, we conducted analysis restricted to taxa that met an 80% prevalence threshold at baseline. Then we applied Multivariable Association Discovery in Population-scale Meta-omics Studies (MaAsLin2), https://huttenhower.sph.harvard.edu/maaslin, a linear mixed model with random intercept of normalized relative abundance (natural log scale; half the minimum relative abundance as pseudo count, min prevalence=0.1), which considers the sparsity of compositional data observed within microbial communities[30]. Taxa were selected with an unadjusted $p<0.05$ at either 4- or 8-weeks on-intervention in the full trial cohort. To quantify whether the microbiome community shifted during the intervention period and the control period over time, we evaluated patterns of beta diversity across the full trial cohort and by intervention order using the weighted Jaccard distance. We also explored the influence of community-level shifts by dietary factors, including usual dietary fiber intake or diet quality score (above vs. below median), metabolite "response to beans" defined as pipecolic acid fold change over 1.0 from baseline to 8 weeks on intervention vs. non-response fold change under 1.0, and chronic medication use (yes vs. no) using Kruskal-Wallis test. These methods were also applied to evaluate the change in LDL by diet and chronic medication use.

To quantify within-person change in circulating metabolites, we conducted a longitudinal analysis using the same methods as microbiome analysis (generalized linear mixed models (SAS, 'Proc Glimmix') on circulating metabolites during the intervention period and the control period over time. Final analysis was restricted to circulating metabolites that met the criteria: Benjamini-Hochberg false discovery rate (FDR) adjusted two-sided p-value (or q-value) less than 0.20, as well as the relevance in previous literature, which included PA, S-(5'-Adenosyl)-L-methionine (SAM), trigonelline, an indole-derivative (putatively cinnamoylglycine) and theophylline. Similar methods were used to assess proteomic markers between baseline and 8-weeks post-intervention with an initial evaluation for potential candidates meeting a one-sided $p<0.05$.

Using the 'diffcoexp' package (version 1.20.0) in R (version 4.2.0)[31], correlation coefficients of selected features (i.e., intervention-responsive biomarkers/key study variables) were estimated at baseline (usual diet control) and 8-weeks post intervention using Fisher's Z-transformation of a Pearson correlation coefficient. Pairs with corresponding differences in correlation values of greater than 0.25 or less than -0.25 were pre-selected and the P-value for the difference (under the null hypothesis of zero difference between the two time-points) were estimated. Features with q-value (using Benjamini-Hochberg method) of less than 0.25 were selected for visualization in the heatmap.

For all analyses in which the multiplicity of tests was an issue, we used the Benjamini-Hochberg FDR to report appropriately adjusted significance levels. Statistical analyses were performed using ATIMA2, R, Python, or SAS 9.4 (SAS Institute INC), as appropriate. We considered FDR-adjusted P value or q-value < 0.20, two-sided, as statistically significant.

Generalized linear mixed effects model with random intercept was followed in this study.

Briefly, let Y_{ij} be the response of subject i at time j. The standard mixed effect model with a random-intercept term in our study is:

$$\log(Y_{ij}) = (\beta_0 + \beta_i) + (\beta_1)\, j + \epsilon_{ij}$$

Here, $\beta_0 + \beta_i$ is the random intercept with $\beta_i \sim N(0,\sigma_0^2)$ where σ_0^2 is the variance of the random intercept. The parameter of primary interest **(Effect estimate)** is β_1 which is the shift in mean between time points and $\epsilon_{ij} \sim N(0,\sigma^2)$ is the error term (and is independent across i and j).

In this equation, we did not adjust for any covariates since confounders were balanced across the randomization sequence or arms.

Reference

1. The IDF consensus worldwide definition of metabolic syndrome. Brussels: International Diabetes Federation, 2006.
2. Oluyomi AO, Panthagani K, Sotelo J, et al. Houston hurricane Harvey health (Houston-3H) study: assessment of allergic symptoms and stress after hurricane Harvey flooding. *Environmental Health* 2021; **20**(1): 9.
3. Zhang X, Browman G, Siu W, et al. The BE GONE trial study protocol: a randomized crossover dietary intervention of dry beans targeting the gut microbiome of overweight and obese patients with a history of colorectal polyps or cancer. *BMC Cancer* 2019; **19**(1): 1233.
4. Subar AF, Kirkpatrick SI, Mittl B, et al. The Automated Self-Administered 24-hour dietary recall (ASA24): a resource for researchers, clinicians, and educators from the National Cancer Institute. *Journal of the Academy of Nutrition and Dietetics* 2012; **112**(8): 1134-7.
5. Millen AE, Midthune D, Thompson FE, Kipnis V, Subar AF. The National Cancer Institute diet history questionnaire: validation of pyramid food servings. *Am J Epidemiol* 2006; **163**(3): 279-88.
6. Krebs-Smith SM, Pannucci TE, Subar AF, et al. Update of the Healthy Eating Index: HEI-2015. *Journal of the Academy of Nutrition and Dietetics* 2018; **118**(9): 1591-602.
7. Hagströmer M, Oja P, Sjöström M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. *Public health nutrition* 2006; **9**(6): 755-62.
8. Rosenberg DE, Bull FC, Marshall AL, Sallis JF, Bauman AE. Assessment of sedentary behavior with the International Physical Activity Questionnaire. *Journal of physical activity & health* 2008; **5 Suppl 1**: S30-44.
9. Aagaard K, Petrosino J, Keitel W, et al. The Human Microbiome Project strategy for comprehensive sampling of the human microbiome and why it matters. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* 2012.
10. Zheng J, Hoffman KL, Chen J-S, et al. Dietary inflammatory potential in relation to the gut microbiome: results from a cross-sectional study. *The British journal of nutrition* 2020; **124**(9): 931-42.
11. Wang Z, Zolnik CP, Qiu Y, et al. Comparison of Fecal Collection Methods for Microbiome and Metabolomics Studies. *Front Cell Infect Microbiol* 2018; **8**: 301.
12. Song SJ, Amir A, Metcalf JL, et al. Preservation Methods Differ in Fecal Microbiome Stability, Affecting Suitability for Field Studies. *mSystems* 2016; **1**(3).
13. Park C, Yun KE, Chu JM, et al. Performance comparison of fecal preservative and stock solutions for gut microbiome storage at room temperature. *J Microbiol* 2020; **58**(8): 703-10.
14. Sinha R, Abu-Ali G, Vogtmann E, et al. Assessment of variation in microbial community amplicon sequencing by the Microbiome Quality Control (MBQC) project consortium. *Nat Biotechnol* 2017; **35**(11): 1077-86.
15. Stewart CJ, Ajami NJ, O'Brien JL, et al. Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* 2018; **562**(7728): 583-8.
16. Methé BA, Nelson KE, Pop M, et al. A framework for human microbiome research. *Nature* 2012; **486**(7402): 215-21.
17. Caporaso JG, Lauber CL, Walters WA, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME journal* 2012; **6**(8): 1621-4.
18. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* 2013; **10**(10): 996-8.

19. Quast C, Pruesse E, Yilmaz P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* 2013; **41**(D1): D590-D6.
20. Paradis E, Claude J, Strimmer K. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics (Oxford, England)* 2004; **20**(2): 289-90.
21. Truong DT, Franzosa EA, Tickle TL, et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nature methods* 2015; **12**(10): 902-3.
22. Beghini F, McIver LJ, Blanco-Míguez A, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *eLife* 2021; **10**: e65088.
23. Abubucker S, Segata N, Goll J, et al. Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput Biol* 2012; **8**(6): e1002358.
24. Fahrmann JF, Bantis LE, Capello M, et al. A Plasma-Derived Protein-Metabolite Multiplexed Panel for Early-Stage Pancreatic Cancer. *J Natl Cancer Inst* 2019; **111**(4): 372-9.
25. Fahrmann JF, Irajizad E, Kobayashi M, et al. A MYC-Driven Plasma Polyamine Signature for Early Detection of Ovarian Cancer. *Cancers (Basel)* 2021; **13**(4).
26. Fahrmann JF, Vykoukal J, Fleury A, et al. Association Between Plasma Diacetylspermine and Tumor Spermine Synthase With Outcome in Triple-Negative Breast Cancer. *J Natl Cancer Inst* 2020; **112**(6): 607-16.
27. Vykoukal J, Fahrmann JF, Gregg JR, et al. Caveolin-1-mediated sphingolipid oncometabolism underlies a metabolic vulnerability of prostate cancer. *Nat Commun* 2020; **11**(1): 4279.
28. Assarsson E, Lundberg M, Holmquist G, et al. Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PloS one* 2014; **9**(4): e95192.
29. Chung NC, Miasojedow B, Startek M, Gambin A. Jaccard/Tanimoto similarity test and estimation methods for biological presence-absence data. *BMC Bioinformatics* 2019; **20**(Suppl 15): 644.
30. Mallick H, Rahnavard A, McIver LJ, et al. Multivariable association discovery in population-scale meta-omics studies. *PLOS Computational Biology* 2021; **17**(11): e1009442.
31. Tesson BM, Breitling R, Jansen RC. DiffCoEx: a simple and sensitive method to find differentially coexpressed gene modules. *BMC Bioinformatics* 2010; **11**(1): 497.

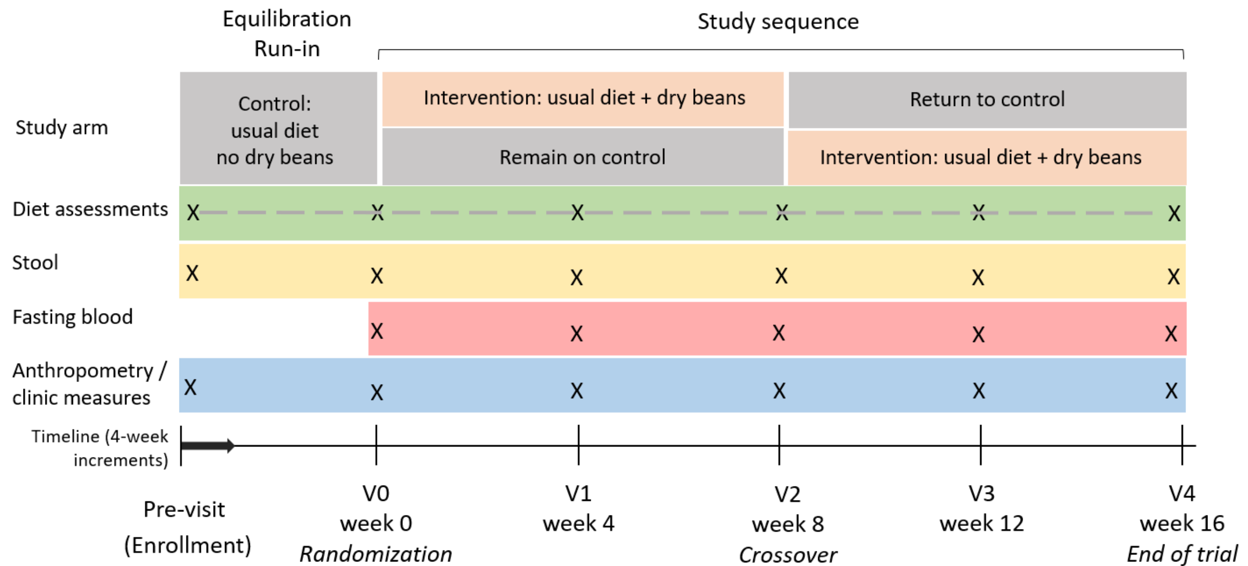

Fig. S1 | Study design

Schematic of study design, intervention sequence and assessments. *Pre-visit, equilibration and run-in:* To establish the basal diet and microbiome and to track compliance with study procedures prior to randomization, eligible and consented individuals were asked to provide a stool sample and to complete dietary and anthropometric assessments. *Randomization and follow-up visits (V0-V4):* Participants who completed the run-in were randomized to begin the intervention diet or to continue the control diet for 8 weeks according to blocks of no use vs. regular use of statins and/or metformin, as they are commonly prescribed in our target population of obese patients with a history of colorectal neoplasia. The intervention diet consisted of ½ cup (1 serving) of study beans over a 2-week ramp-up period followed by 1 cup (2 servings) per day for an additional 6 weeks. At week 8, participants who completed the intervention diet immediately crossed over to the control diet and vice versa. Randomized participants attended a total of 5 in-person visits every 4 weeks. For each visit, participants provided a stool sample, fasting blood sample, anthropometry and blood pressure with review of any changes in medications or health status. Usual dietary habits and physical activity levels were assessed and monitored throughout the study.

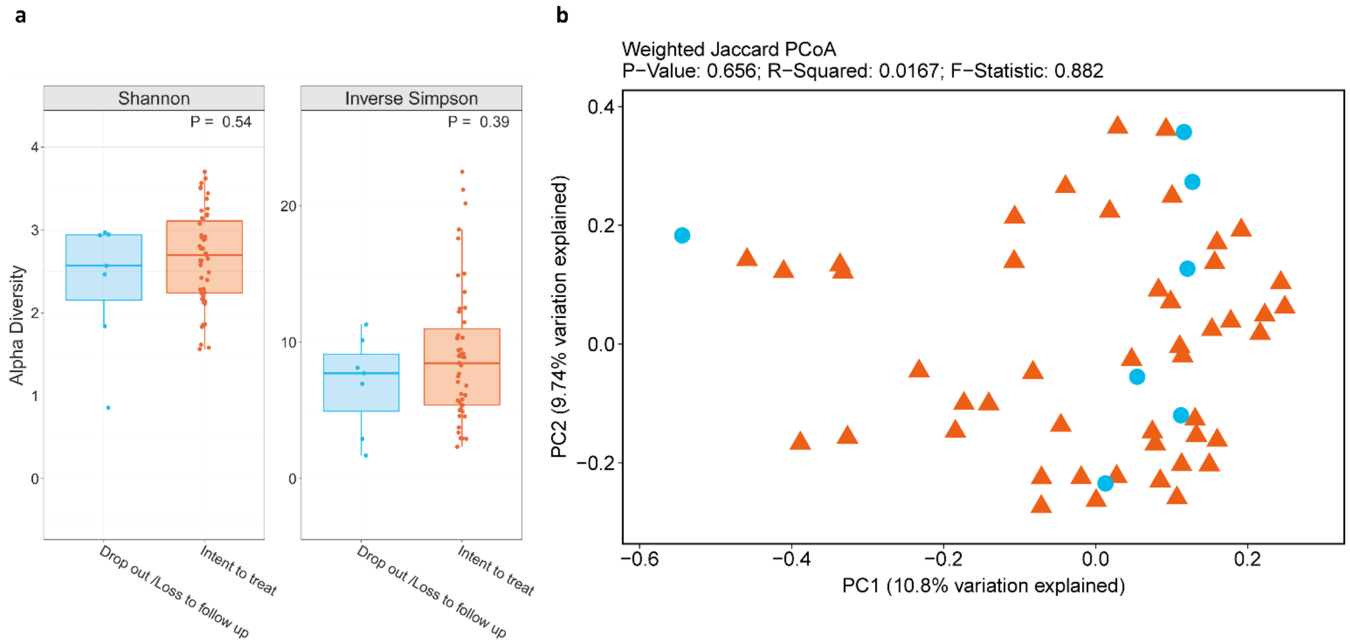

Fig. S2 | Microbiome differences among intent-to-treat and withdrawal groups

a. Alpha diversity comparison of participants in the intent-to-treat group (n=48; orange), as compared to those who withdrew or were lost to follow-up post-randomization (n=7; blue), such that each analysis included 55 total participants. Alpha diversity (left) was quantified by the Shannon diversity index and the Inverse Simpson index. P-for linear trend was derived from a general linear model with each index as the dependent variable and intent-to-treat or withdrawal as independent variables. The significance test was two-sided Mann-Whitney U test. Box plot centers show medians of each alpha diversity index with boxes indicating their inter-quartile ranges (IQRs); upper and lower whiskers indicate 1.5 times the IQR from above the upper quartile and below the lower quartile, respectively. No notable differences in either alpha diversity index within the basal microbiome was observed between these two groups (all $p \geq 0.39$). **b**. Proportion of variation in taxonomy (beta diversity) by intent-to-treat or withdrawal as quantified by two-sided permutational multivariate analysis of variance (PERMANOVA implemented in "adonis" function in statistical software R package "vegan") based on weighted Jaccard distance, with principal coordinate analysis (PCoA) plots to visualize the dissimilarity of the community composition. No notable differences in weighted Jaccard distance within the basal microbiome were observed between these two groups (p=0.66).

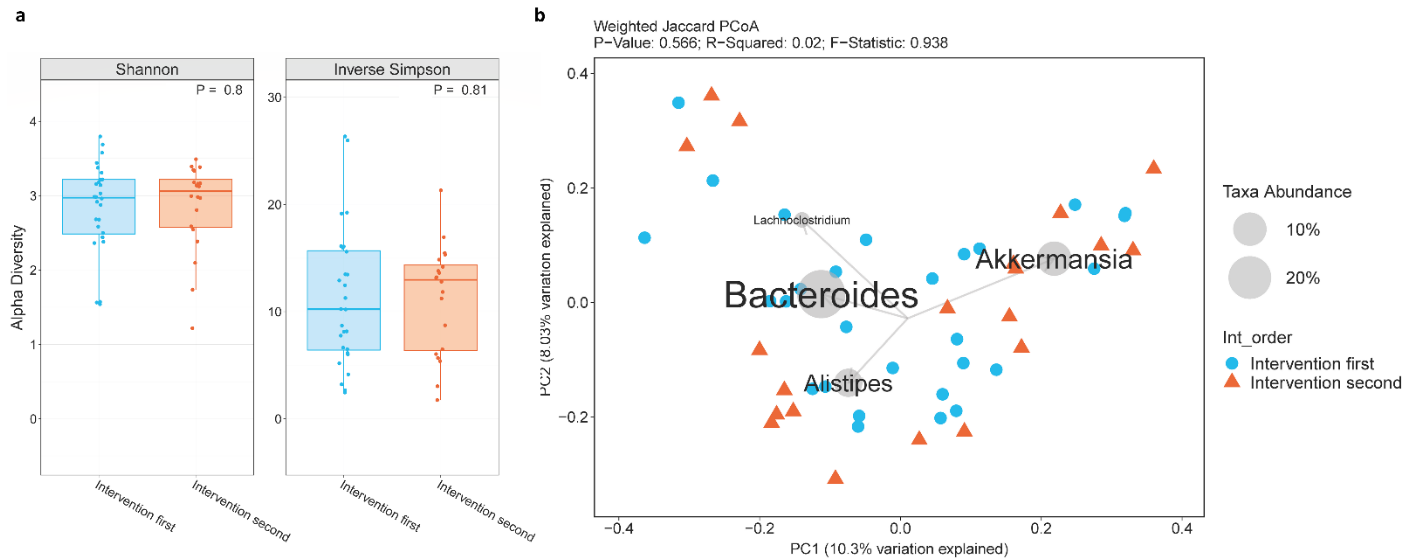

Fig. S3 | Microbiome differences by intervention order (randomization group) within the intent-to treat group

a. Alpha diversity comparison of participants randomized to the intervention first (n=28; blue) versus second (n=20; orange) in the intent-to-treat group (total n=48). Alpha diversity was quantified by the Shannon diversity index and the inverse Simpson index. P for linear trend was derived from a general linear model with each index as the dependent variable and randomization order as independent variables. The significance test was two-sided Mann-Whitney U test. Box plot centers show medians of each alpha diversity index with boxes indicating their inter-quartile ranges (IQRs); upper and lower whiskers indicate 1.5 times the IQR from above the upper quartile and below the lower quartile, respectively. No notable differences in either alpha diversity index within the basal microbiome was observed between these two groups (all p≥0.80). **b.** Proportion of variation in taxonomy (beta diversity) by randomization order as quantified by two-sided permutational multivariate analysis of variance (PERMANOVA implemented in "adonis" function in statistical software R package "vegan") based on weighted Jaccard distance, with principal coordinate analysis (PCoA) biplots to visualize the dissimilarity of the community composition with the abundance of genus-level taxa. No notable differences in weighted Jaccard distance within the basal microbiome were observed between randomization groups (p=0.58).

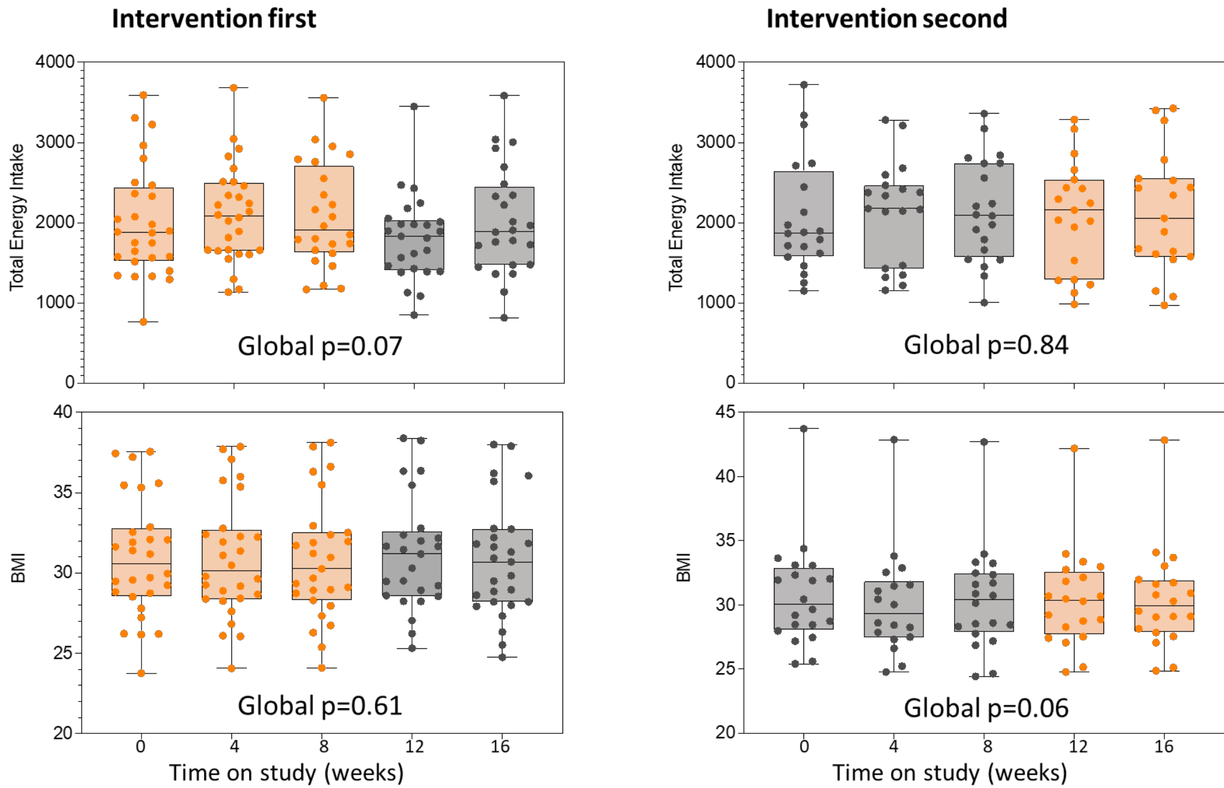

Fig. S4 | Overall trend of total energy intake and BMI by the intervention sequence

Overall trends in total energy intake (top) and BMI (bottom) remained largely stable over the study period. Within person change across the 8-week intervention period for the full trial cohort for total energy intake and BMI, respectively: 0.01 (-0.07, 0.09), two-sided p=0.85 and 0.001 (-0.004, 0.006) two-sided p=0.70 by generalized linear mixed models (Proc Glimmix, SAS 9.4) with random intercept on the natural log scale.

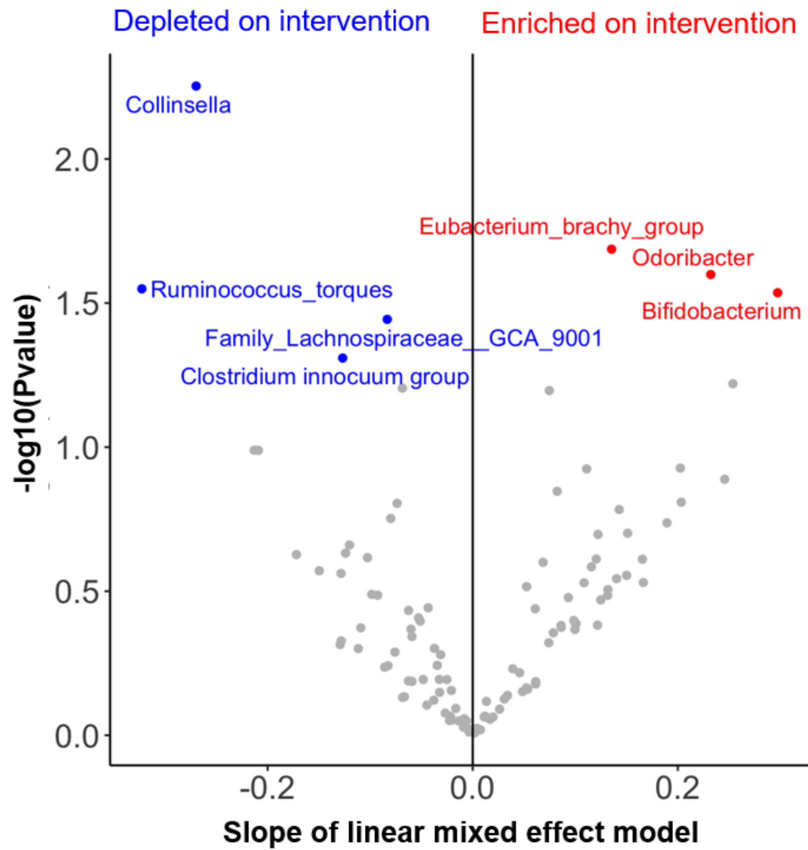

Fig. S5 | Intra-individual shift in the relative abundance of bacteria across the three on-intervention time-points in the full trial cohort: 16S genus-level analysis

Volcano plot visualizing the results of MaAsLin2 analysis characterizing the slope across the three on-intervention time points (144 stool samples in 48 patients). Maximum estimated false discovery rate when calling all p-values (one-sided) ≤ 0.05 is 0.35

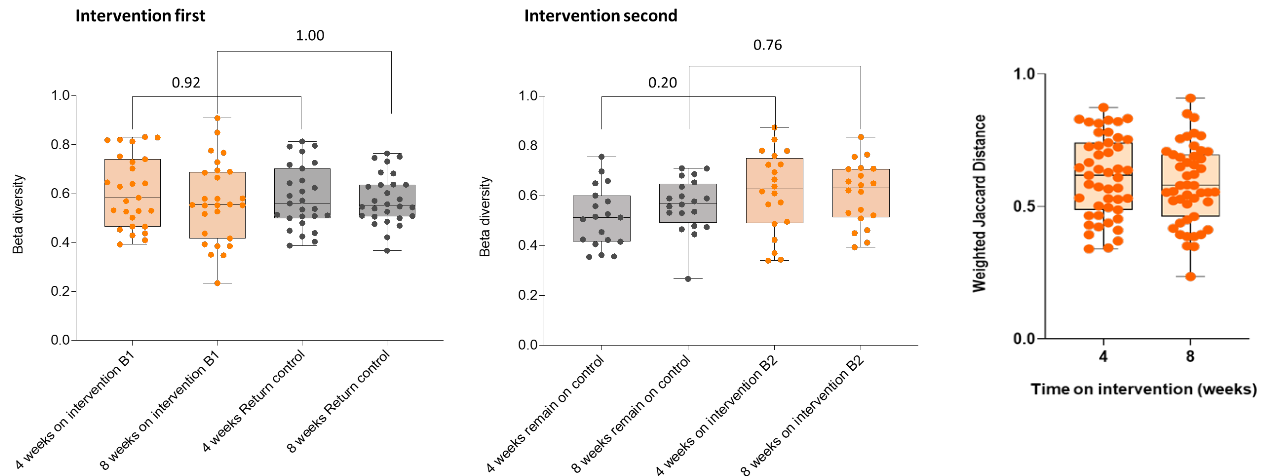

Fig. S6 | Intra-individual shifts in the gut microbiota community

Left to right Beta diversity shifts by weighted Jaccard distances across the study sequence (n=28 randomized to the intervention first; n=20 randomized to intervention second; among total n for intent-to-treat=48). No difference was detected between the overall shifts achieved in the first 4 weeks as compared to the full 8 weeks on-intervention (Kruskal Wallis test, two-sided p>0.05; not shown above). Both groups completed a 4-week equilibration on the control diet before randomization to one of two sequences. At left, participants randomized to the intervention first, who returned to the usual diet without dry beans (control) in the second 8 weeks of the trial, experienced nearly equivocal shifts in weighted Jaccard distance at both 4 weeks and 8 weeks post-intervention vs. post-control (p=0.92 for 4 weeks and p=1.00 for 8 weeks), which is consistent with the reversal of several individual taxa changes. At right, participants randomized to the intervention second remained on the control diet for another 8 weeks before beginning the intervention. In this group, we observed that the weighted Jaccard distance shifts at both 4 weeks and 8 weeks post-intervention were not significantly different than what was observed during the same period on the control (two -sided p=0.20 for 4 weeks and two-sided p=0.76 for 8 weeks). Beta diversity shifts by weighted Jaccard distances [mean (SD] from baseline to 4 weeks [0.61 (0.15)] and baseline to 8 weeks [0.58 (0.15)] post-intervention among the full trial cohort (n=48). No difference was detected between the overall shifts achieved in the first 4 weeks as compared to the full 8 weeks on-intervention (Kruskal-Wallis Test two-sided p=0.43).

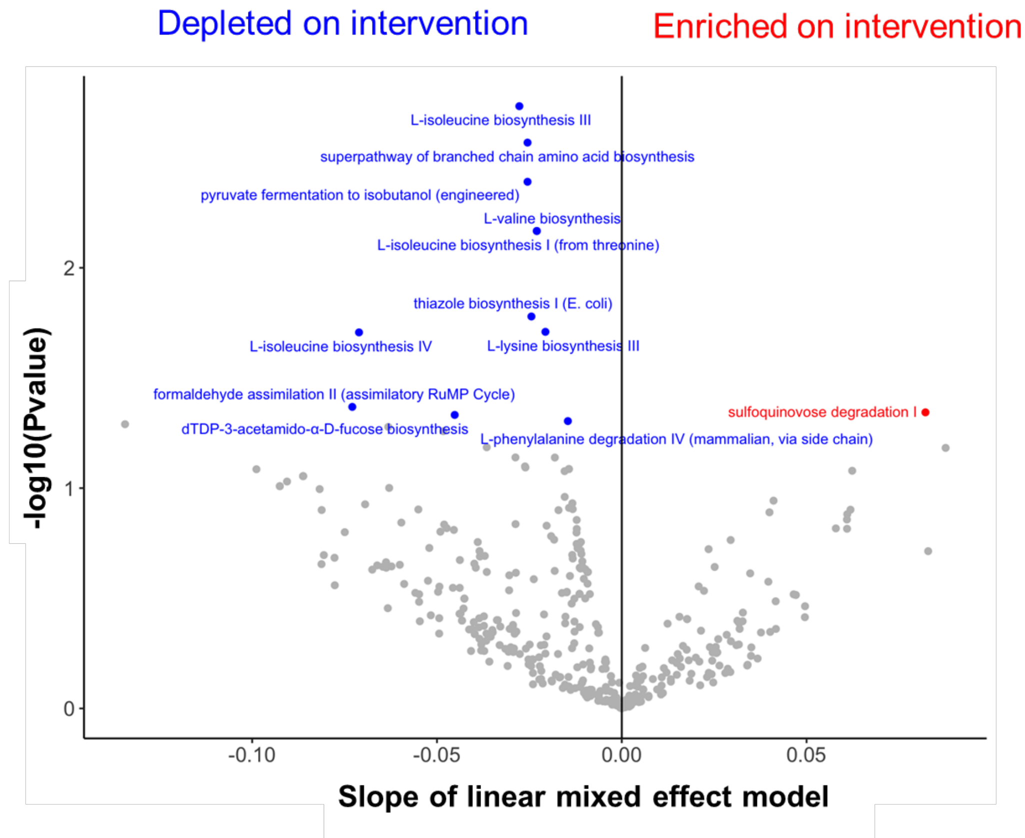

Fig. S7 | Intra-individual shifts in microbial gene content pathways across the three on-intervention time points in the full trial cohort

Volcano plot visualizing the results of metagenomic pathways via MaAsLin2 analysis applied to the metagenomic subset (140 stool samples in 48 patients) characterizing the slope across the three on-intervention time points. Maximum estimated false discovery rate when calling all p-values (one-sided) ≤ 0.05 is 0.46

Table S1. Characteristics of the BE GONE Trial participants by randomization group (intent-to-treat, n=48)

		Intervention first	Intervention second	*P-value*[c]
Number of subjects (n)		28	20	
Block randomization factors				
Statin and/or metformin use				*0.88*
	Yes	12 (42.9)	9 (45.0)	
	No	16 (57.1)	11 (55.0)	
Colorectal history				*0.50*
	Precancer	6 (21.4)	6 (30.0)	
	Cancer	22 (78.6)	14 (70.0)	
Demographic characteristics				
Age (years)	Mean (SD)	63.5 (10.3)	61.3 (8.3)	*0.43*
Gender				*0.73*
	Male	14 (50.0)	11 (55.0)	
	Female	14 (50.0)	9 (45.0)	
Race				*0.57*
	White, NH	20 (71.4)	16 (80.0)	
	Black, NH	3 (10.7)	1 (5.0)	
	Hispanic	4 (14.3)	1 (5.0)	
	Asian	1 (3.6)	2 (10.0)	
Education				*0.27*
	High school (HS)	4 (14.3)	0 (0.0)	
	post HS training/ some college	7 (25.0)	4 (20.0)	
	College graduate	11 (39.3)	8 (40.0)	
	Postgraduate	6 (21.4)	8 (40.0)	
Marital status				*0.70*
	Married/living as married	21 (75.0)	14 (70.0)	
	Widowed/divorced/Never married	7 (25.0)	6 (31.0)	
Biometrics				
BMI (kg/m2)	Mean (SD)	30.9 (3.7)	30.7 (4.1)	*0.87*
Waist (cm)	Mean (SD)	101.8 (10.3)	104.2 (10.7)	*0.42*
Blood pressure	Mean (SD)			
	Systolic (mmHg)	126.4 (11.9)	127.1 (13.9)	*0.84*
	Diastolic (mmHg)	76.2 (8.6)	76.2 (8.1)	*0.99*
Usual diet and lifestyle				
Dietary intake[a]	Mean (SD)			
HEI-2015 score		57.6 (13.1)	60.9 (10.2)	*0.35*
Total energy (kcal/day)		2006 (689)	2078 (745)	*0.73*
Fat (% E)		37.7 (7.1)	39.5 (5.7)	*0.34*
Carbohydrate (% E)		43.1 (9.9)	42.5 (8.5)	*0.83*
Protein (% E)		18.1 (5.1)	17.2 (4.2)	*0.50*
Alcohol (% E)		3.4 (5.4)	2.6 (2.9)	*0.55*
Fiber (g/day)		18.2 (8.0)	19.8 (11.3)	*0.56*
Fiber (g/1000 kcal)		9.6 (3.6)	9.8 (4.1)	0.85
Legumes (cup eq/day)		0.09 (0.10)	0.12 (0.16)	*0.49*
Legumes (cup eq/1000 kcal)		0.06 (0.06)	0.07 (0.09)	*0.63*
Physical activity[b]				*0.07*
	Inactive	11 (44.0)	15 (79.0)	
	Moderate	9 (36.0)	2 (10.5)	
	Vigorous	5 (20.0)	2 (10.5)	
Total METs	Mean (SD)	3626 (3518)	2263 (2451)	*0.16*

[a] Dietary intake was obtained from baseline dietary recalls during the equilibration period except for usual intake of legumes and alcohol, which were obtained from diet history questionnaire

[b] "Inactive" is defined as <150 mins of moderate intensity activity/wk; "Moderate" defined as 150-300 mins of moderate intensity activity/wk; and "Vigorous" defined as >300 mins of moderate intensity activity/wk, or 150 mins of vigorous intensity activity/wk, or a combination. Excludes 4 participants who provided insufficient data

[c] P-values derived from comparison of frequencies or means by Pearson Chi-Square, Fisher's exact or ANOVA, as appropriate

Table S2. Comparison of patient characteristics among the intent-to-treat group versus those who withdrew pre- or post-randomization (total n=62)

Characteristic		A	B	C	*Global*	*B vs. C*
		Run-in / Pre-randomization withdrawal[a]	Post-randomization withdrawal / loss to follow-up	Completed trial	*P-value*[b]	*P-value*[b]
Number of subjects (n)		7	7	48		
Demographic factors						
Age (years)	Mean, SD	63.4 (9.0)	60.1 (13.8)	62.6 (9.5)	*0.80*	*0.55*
Gender					*0.67*	*1.00*
	Male	2 (28.6)	4 (57.1)	25 (52.1)		
	Female	5 (71.4)	3 (42.9)	23 (47.9)		
Race					*0.94*	*0.87*
	White, Not Hispanic	6 (85.7)	5 (71.4)	36 (75.0)		
	Black, Not Hispanic	1 (14.3)	1 (14.3)	4 (8.3)		
	Hispanic	0 (0.0)	1 (14.3)	5 (10.4)		
	Asian	0 (0.0)	0 (0.0)	3 (6.3)		
Education					*0.55*	*0.19*
	12 years or completed high school	0 (0.0)	1 (14.3)	4 (8.3)		
	post high school training/some college	2 (28.6)	4 (57.1)	11 (23.0)		
	College graduate	3 (42.9)	1 (14.3)	19 (39.6)		
	Postgraduate	2 (28.6)	1 (14.3)	14 (29.2)		
Marital status					*0.69*	*1.00*
	Married or living as married	4 (57.1)	6 (85.7)	35 (72.9)		
	Widowed/divorced/Separated	3 (42.9)	1 (14.3)	12 (25.0)		
	Never married	0 (0.0)	0 (0.0)	1 (2.1)		
Dietary factors						
HEI-2015 score	Mean, SD	60.6 (5.7)	64.8 (8.5)	68.1 (9.8)	*0.20*	*0.41*
Fiber (g/1000 kcal)	Mean, SD	8.5 (1.4)	10.3 (3.3)	10.3 (3.3)	*0.50*	*0.99*
Legumes (cup eq/1000 kcal)	Mean, SD	0.05 (0.05)	0.04 (0.04)	0.06 (0.07)	*0.87*	*0.61*
Randomization factors						
Colorectal history					*NA*	*0.02*
	Precancer	NA	5 (71.4)	12 (25.0)		
	Cancer	NA	2 (28.6)	36 (75.0)		

Statin and/or metformin use					*NA*	*0.04*
	Yes	NA	7 (100.0)	21 (43.8)		
	No	NA	0 (0.0)	27 (56.3)		
Intervention order					*NA*	*0.04*
	Intervention first	NA	1 (14.3)	28 (58.3)		
	Intervention second	NA	6 (85.7)	20 (41.7)		

[a] Group A refers to withdrawals among participants who completed the pre-visit and began the run-in but does not include consented individuals who dropped out before providing any study data; n=2 of group A did not sufficiently complete the dietary assessment to estimate intake.
[b] P-values derived from comparison of frequencies or means by Pearson Chi-Square, Fisher's exact or ANOVA, as appropriate

Table S3. Adverse effects reported by patients enrolled in the BE GONE trial (randomized, n=55)

Participant	Randomization Order	Completed (C) versus withdrew (W)	Description (summary code)*	Related/Attributable
A	Bean second	C	Other Illness/Injury	No
			Gas/Flatulence	Possible
			Constipation/Diarrhea/Other change in bowel habits	Possible
B	Bean first	C	None	
C	Bean second	W	Other Illness/Injury	No
D	Bean second	W	Other Illness/Injury	No
E	Bean second	C	Other Illness/Injury	No
F	Bean second	W	Other Illness/Injury	No
G	Bean first	C	Other Illness/Injury	No
H	Bean second	C	Gas/Flatulence	Yes
			Other Illness/Injury	No
I	Bean first	C	None	
J	Bean second	C	None	
K	Bean first	C	None	
L	Bean second	C	Constipation/Diarrhea/Other change in bowel habits	Unlikely
			Gas/Flatulence	Possible
			Other Illness/Injury	No
M	Bean second	C	Other Illness/Injury	No
N	Bean second	C	Other Illness/Injury	No
O	Bean first	C	Other Illness/Injury	No
P	Bean second	W	None	
Q	Bean second	C	None	
R	Bean first	C	Constipation/Diarrhea/Other change in bowel habits	Possible
			Reflux/Heartburn	No
			Other Illness/Injury	No
S	Bean first	C	Constipation/Diarrhea/Other change in bowel habits	No
T	Bean second	C	Gas/Flatulence	Probable
U	Bean first	C	Constipation/Diarrhea/Other change in bowel habits	Unlikely
V	Bean first	C	Constipation/Diarrhea/Other change in bowel habits	Unlikely
			Other Illness/Injury	No
			Gas/Flatulence	Possible
W	Bean second	C	None	
X	Bean second	C	Other Illness/Injury	No
			Other Illness/Injury	No
			Gas/Flatulence	Probable
			Other Systemic Effects	No
Y	Bean second	C	Gas/Flatulence	Probable
Z	Bean first	C	Other Illness/Injury	No

			Other Systemic Effects	No
AA	Bean first	C	None	
BB	Bean first	C	Gas/Flatulence	Possible
			Other Illness/Injury	No
			Constipation/Diarrhea/Other change in bowel habits	Unlikely
CC	Bean first	C	None	
DD	Bean first	C	Other Illness/Injury	No
			Constipation/Diarrhea/Other change in bowel habits	Possible
EE	Bean first	C	None	
FF	Bean second	C	Other Illness/Injury	No
			Constipation/Diarrhea/Other change in bowel habits	Yes
			Gas/Flatulence	Yes
GG	Bean first	C	Gas/Flatulence	Yes
HH	Bean second	C	Flatulence	Yes
II	Bean first	C	Other Illness/Injury	No
JJ	Bean first	C	Other Illness/Injury	No
		C	Gas/Flatulence	Yes
KK	Bean first	C	Gas/Flatulence	Possible
		C	Other Illness/Injury	No
		C	Constipation/Diarrhea/Other change in bowel habits	Unlikely
		C	Reflux/Heartburn	No
LL	Bean first	C	None	
MM	Bean second	C	None	
NN	Bean second	C	Gas/Flatulence	Yes
			Other Illness/Injury	No
			Constipation/Diarrhea/Other change in bowel habits	Unlikely
OO	Bean first	C	Other Illness/Injury	No
PP	Bean second	W	None	
QQ	Bean first	C	Gas/Flatulence	Yes
RR	Bean second	C	Other Systemic Effects	Possible
SS	Bean second	C	Other Illness/Injury	No
TT	Bean second	C	Other Illness/Injury	No
UU	Bean first	C	Gas/Flatulence	Yes
			Constipation/Diarrhea/Other change in bowel habits	Possible
			Other Illness/Injury	No
VV	Bean second	C	None	
WW	Bean first	W	Other Systemic Effects	No
XX	Bean first	C	Other Illness/Injury	No
YY	Bean first	C	Other Illness/Injury	No
			Constipation/Diarrhea/Other change in bowel habits	No
ZZ	Bean first	C	Other Illness/Injury	No
AAA	Bean first	C	Other Illness/Injury	No
BBB	Bean second	W	None	

CCC	Bean first	C	None	

*GI side effects defined as above. Other illness/injury refers to a range of reported changes in health status, such as respiratory viruses, muscle injuries or surgeries that were not related to the intervention. Other Systemic effects refers to reported joint pain or results of non-trial labs, some of which were possibly related to the intervention.

Table S4. Change in biometrics following 4-week and 8-week increase in dry bean intake (intent-to-treat, n=48)

Markers	Baseline to week 4 on intervention			Baseline to week 8 on intervention		
	Fold change[a]	n	Effect estimate[b] (log) (95%CI)	Fold change[a]	n	Effect estimate[b] (log) (95%CI)
Blood pressure						
Systolic	1.02	48	0.01 (-0.02, 0.04)	1.01	47	0.01 (-0.02, 0.03)
Diastolic	1.01	48	0.01 (-0.02, 0.03)	0.99	47	-0.01 (-0.04, 0.01)
Anthropometric						
BMI	1.00	48	0.0005 (-0.004, 0.005)	1.00	47	0.001 (-0.004, 0.006)
Waist Circumference	1.01	47	0.01 (0.0004, 0.02)	1.01	47	0.01 (0.0002, 0.02)
Dietary intake						
HEI-2015 score	1.11	47	0.09 (0.03, 0.15)	1.14	43	0.11 (0.05, 0.17)
Total energy (kcal)	1.05	47	0.01 (-0.07, 0.10)	1.04	43	0.01 (-0.07, 0.09)
Fat (% E)	0.91	47	-0.12 (-0.18, -0.06)	0.88	43	-0.16 (-0.24, -0.09)
Carbohydrate (% E)	1.09	47	0.06 (0.01, 0.12)	1.14	43	0.11 (0.04, 0.18)
Protein (% E)	1.03	47	0.01 (-0.04, 0.07)	1.00	43	-0.02 (-0.10, 0.05)
Alcohol (% E)[c]	0.87	25	-0.34 (-0.78, 0.10)	1.10	24	-0.11 (-0.67, 0.44)
Fiber (g/1000 kcal)	1.27	47	0.19 (0.08, 0.29)	1.36	43	0.22 (0.10, 0.35)
Physical activity						
Total METs	1.66	41	0.03 (-0.32, 0.37)	1.69	40	-0.09 (-0.38, 0.20)

[a] Fold change as a ratio between follow up and baseline value; the sample size for each analysis might not equal to 48 due to missing at any time point

[b] Generalized linear mixed models with random intercept, natural log transformation applied.

[c] Only among non-zeros (e.g., baseline and regular drinkers)

Table S5. Change in circulating lipoproteins across the full trial sequence (intent-to-treat, n=48)

	Baseline to week 4 to week 8 on intervention (full trial cohort across 3 time points)	
	Effect estimate[a] (log) (95%CI)	**p-value[b]**
Circulating lipoprotein panel		
LDL	-0.01 (-0.02, 0.02)	0.56
HDL	0.01 (-0.02, 0.01)	0.44
LDL/HDL	-0.01 (-0.03, 0.02)	0.63
Total cholesterol	-0.01 (-0.02, 0.01)	0.23
Triglycerides	-0.01 (-0.04, 0.03)	0.58

	Baseline to week 4 on intervention			**Baseline to week 8 on intervention**		
	Fold change[c]	**n**	**Effect estimate[a] (log) (95%CI)**	**Fold change[a]**	**n**	**Effect estimate[a] (log) (95%CI)**
LDL	1.02	47	0.002 (-0.05, 0.05)	0.99	46	-0.02 (-0.06, 0.02)
HDL	0.97	47	-0.03 (-0.05, -0.01)	0.99	46	-0.01 (-0.04, 0.12)
LDL/HDL	1.04	47	0.03 (-0.02, 0.08)	0.99	46	-0.02 (-0.06, 0.02)
Total cholesterol	0.99	47	-0.02 (-0.04, 0.01)	0.98	46	-0.02 (-0.05, 0.01)
Triglycerides	1.00	47	-0.02 (-0.08, 0.03)	1.02	46	-0.02 (-0.10, 0.06)
Beans first, n=28						
	Baseline to week 4 on intervention			**Baseline to week 8 on intervention**		
LDL	1.06	27	0.03 (-0.05, 0.12)	0.99	26	-0.01 (-0.0, 0.04)
HDL	0.98	27	-0.03 (-0.07, 0.01)	0.98	26	-0.03 (-0.07, 0.01)
LDL/HDL	1.08	27	0.05 (-0.02, 0.13)	1.00	26	-0.002 (-0.05, 0.04)
	Week 8 to week 12 on control			**Week 8 to week 16 on control**		
LDL	1.06	22	0.04 (-0.4, 0.13)	1.02	25	0.01 (-0.05, 0.06)
HDL	1.04	22	0.03 (0.002, 0.07)	1.03	25	0.03 (-0.01, 0.06)
LDL/HDL	1.03	22	0.01 (-0.08, 0.10)	0.99	25	-0.02 (-0.08, 0.04)
Beans second, n=20						
	Baseline to week 4 on control			**Baseline to week 8 on control**		
LDL	1.06	18	-0.06 (-0.14, 0.63)	1.05	20	0.04 (-0.03, 0.10)
HDL	1.00	18	-0.004 (-0.05, 0.04)	1.00	20	-0.01 (-0.05, 0.04)
LDL/HDL	0.97	18	-0.05 (-0.15, 0.04)	1.06	20	0.05 (-0.04, 0.13)
	Week 8 to week 12 on intervention			**Week 8 to week 16 on intervention**		
LDL	0.97	20	-0.04 (-0.09, 0.01)	0.98	20	-0.03 (-0.10, 0.4)
HDL	0.97	20	-0.03 (-0.06, -0.004)	1.02	20	0.01 (-0.05, 0.07)
LDL/HDL	1.00	20	-0.01 (-0.06, 0.04)	0.98	20	-0.04 (-0.13, 0.04)

[a] Generalized linear mixed model with random intercept, natural log transformation applied. All 48 patients were able to contribute to the 3 time point analysis, although the total number of blood samples from 48 patients across 3 time points does not sum to 144 (n=140) because some blood samples were not collected due to missed clinic visits (whereas every stool sample was recovered). All other patient n's as shown.

[b] P-value by Wald test

[c] Fold change as a ratio between follow up and baseline value; the sample size for each analysis might not equal to 28 or 20 due to missing at any of the time-points

Table S6. Evaluation of diet, medication use and other factors in modulating the effect of the intervention on circulating LDL and microbial alpha and beta diversity (intent-to-treat, n=48)

		LDL[a]		Alpha Diversity (Inverse Simpson Index)[a]		Beta diversity[a]	
		Effect estimate[b], 95%CI		Effect estimateb, 95%CI		Weighted Jaccard Distance	
		4 weeks on intervention	8 weeks on intervention	4 weeks on intervention	8 weeks on intervention	4weeks with baseline	8weeks with baseline
Overall		0.002 (-0.05, 0.05)	-0.02 (-0.06, 0.02)	-0.02 (-0.18, 0.14)	0.16 (0.02, 0.30)	0.61 (0.15)	0.58 (0.15)
Gender							
	Male (n=25)	-0.02 (-0.07, 0.03)	-0.04 (-0.10, 0.27)	-0.10 (-0.27, 0.09)	0.12 (-0.10, 0.34)	0.61 (0.14)	0.58 (0.14)
	Female (n=23)	0.02 (-0.07, 0.12)	-0.01 (-0.06, 0.05)	0.06 (-0.22, 0.35)	0.23 (0.04, 0.43)	0.62 (0.17)	0.59 (0.17)
	p for interaction[c]	0.42	0.51	0.34	0.43		
	Kruskal-Wallis Test p-value[d]					0.85	0.73
Age							
	Below median, baseline (n=24)	-0.003(-0.09, 0.08)	-0.04 (-0.11, 0.03)	-0.04 (-0.26, 0.17)	0.08 (-0.10, 0.24)	0.56 (0.34)	0.55 (0.14)
	Median and above, baseline (n=24)	0.01 (-0.06, 0.07)	-0.01 (-0.06, 0.05)	0.01 (-0.25, 0.26)	0.25 (0.02, 0.48)	0.66 (0.14)	0.61 (0.16)
	p for interaction[c]	0.86	0.32	0.75	0.22		
	Kruskal-Wallis Test p-value[d]					0.02	0.15
Dietary fiber							
	Below median, baseline (n=24)	-0.01 (-0.07, 0.05)	-0.04 (-0.10, 0.02)	0.01 (-0.20, 0.22）	0.19 (-0.01, 0.40)	0.64 (0.16)	0.59 (0.16)
	Median and above, baseline (n=24)	0.01 (-0.08, 0.10)	0.001 (-0.06, 0.07)	-0.04 (-0.30, 0.21)	0.13 (-0.07, 0.34)	0.58 (0.14)	0.58 (0.14)
	p for interaction[c]	0.63	0.34	0.75	0.68		
	Kruskal-Wallis Test p-value[d]					0.18	0.78
Dietary fiber fold change from baseline to 8 weeks							
	1.0 or less (n=11)	-0.01 (-0.12, 0.09)	-0.09 (-0.18, 0.02)	0.16 (-0.31, 0.63)	0.32 (-0.04, 0.69)	0.65 (0.17)	0.63 (0.17)
	Above 1.0 (n=32)	-0.03 (-0.08, 0.02)	-0.01 (-0.06, 0.04)	-0.08 (-0.25, 0.10)	0.12 (-0.04, 0.29)	0.59 (0.15)	0.56 (0.13)
	p for interaction[c]	0.83	0.15	0.21	0.24		
	Kruskal-Wallis Test p-value[d]					0.32	0.1
Diet quality (HEI2015)							
	Below median, baseline (n=24)	0.01 (-0.05, 0.07)	-0.02 (-0.08, 0.05)	-0.04 (-0.29, 0.20)	0.12 (-0.13, 0.36)	0.63 (0.16)	0.60 (0.16)
	Median and above, baseline (n=24)	-0.01 (-0.10, 0.08)	-0.03 (-0.09, 0.04)	0.01 (-0.21, 0.24)	0.21 (0.05, 0.37)	0.59 (0.14)	0.57 (0.14)
	p for interaction[c]	0.79	0.78	0.76	0.52		
	Kruskal-Wallis Test p-value[d]					0.31	0.63
Dietary quality (HEI2015) fold change from baseline to 8 weeks							
	1.0 or less (n=9)	-0.03 (-0.14, 0.08)	-0.05 (-0.20, 0.10)	-0.17 (-0.59, 0.25)	0.20 (-0.10, 0.50)	0.65 (0.14)	0.59 (0.10)
	Above 1.0 (n=34)	-0.02 (-0.07, 0.03)	-0.03 (-0.08, 0.02)	0.03 (-0.17, 0.22)	0.17 (-0.01, 0.35)	0.59 (0.16)	0.58 (0.16)
	p for interaction[c]	0.81	0.69	0.34	0.87		
	Kruskal-Wallis Test p-value[d]					0.35	0.93
Chronic medication use							
	Yes (n=21)	0.02 (-0.08, 0.13)	-0.02 (-0.09, 0.06)	-0.07 (-0.31, 0.17)	0.11 (-0.07, 0.29)	0.63 (0.15)	0.58 (0.16)
	No (n=27)	-0.02 (-0.07, 0.03)	-0.03 (-0.08, 0.03)	0.03 (-0.21, 0.26)	0.21 (-0.01, 0.42)	0.59 (0.15)	0.58 (0.15)
	p for interaction[c]	0.44	0.84	0.57	0.49		
	Kruskal-Wallis Test p-value[d]					0.42	0.89

Pipecolic acid fold change from baseline to 8 weeks						
1.0 or less (n=7)	-0.04 (-0.18, 0.09)	-0.07 (-0.20, 0.07)	-0.08 (-0.31, 0.14)	0.11 (-0.06, 0.27)	0.61 (0.18)	0.61 (0.15)
Above 1.0 (n=38)	-0.01 (-0.05, 0.04)	-0.01 (-0.05, 0.04)	0.01 (-0.25, 0.28)	0.24 (-0.01, 0.50)	0.59 (0.13)	0.58 (0.15)
p for interaction[c]	0.54	0.32	0.58	0.35		
Kruskal-Wallis Test p-value[d]					0.95	0.99
Gas/flatulence						
Yes (n=16)	-0.02 (0-0.08, 0.05)	-0.03 (-0.09, 0.04)	0.07 (-0.15, 0.29)	0.13 (-0.07, 0.32)	0.64 (0.15)	0.59 (0.17)
No (n=32)	0.01 (-0.06, 0.08)	-0.02 (-0.07, 0.04)	-0.06 (-0.28, 0.16)	0.18 (-0.02, 0.38)	0.60 (0.15)	0.58 (0.14)
p for interaction[c]	0.62	0.89	0.46	0.73		
Kruskal-Wallis Test p-value[d]					0.4	0.62
Constipation/Diarrhea/Other change in bowel habits						
Yes (n=5)	0.06 (-0.06, 0.18)	0.05 (-0.09, 0.18)	0.28 (-0.56, 1.11)	0.61 (0.17, 1.04)	0.72 (0.09)	0.64 (0.13)
No (n=43)	-0.01 (-0.06, 0.05)	-0.03 (-0.07, 0.02)	-0.05 (-0.22, 0.12)	0.11 (-0.03, 0.26)	0.60 (0.09)	0.58 (0.15)
p for interaction[c]	0.44	0.26	0.22	0.03		
Kruskal-Wallis Test p-value[d]					0.07	0.33

[a] The sample size for each analysis might not equal to 48 due to missing at any time-point
[b] Generalized linear mixed model with random intercept, natural log transformation applied.
[c] P-value by Wald test
[d] P-value by Kruskal-Wallis Test applies to beta-diversity (weighted Jaccard distance) only

Table S7. Changes in alpha diversity among the full trial cohort (intent-to-treat n=48)

	Baseline to week 4 on intervention				Baseline to week 8 on intervention			
	Fold change[a]	n	Effect estimate[b] (log) (95%CI)	p-value[c]	Fold change[b]	n	Effect estimate[b] (log) (95% CI)	p-value[c]
Alpha diversity								
Shannon Diversity Index (16S)	1.01	48	-0.003 (-0.06, 0.05)	0.90	1.06	48	0.04 (-0.01, 0.09)	0.10
Inverse Simpson Index (16S)	1.14	48	-0.02 (-0.18, 0.14)	0.84	1.32	48	0.16 (0.02, 0.30)	0.02
Shannon Diversity Index (WGS)	1.10	45	0.02 (-0.04, 0.07)	0.57	1.13	45	0.05 (-0.005, 0.11)	0.07
Inverse Simpson Index (WGS)	1.14	45	0.04 (-0.14, 0.22)	0.68	1.19	45	0.16 (-0.02, 0.34)	0.07
	Beans first, n=28							
	Baseline to week 4 on intervention				**Baseline to week 8 on intervention**			
Shannon Diversity Index (16S)	1.06	28	0.04 (-0.04, 0.12)	0.32	1.08	28	0.06 (-0.02, 0.14)	0.15
Inverse Simpson Index (16S)	1.25	28	0.08 (-0.14, 0.30)	0.47	1.35	28	0.15 (-0.06, 0.37)	0.17
Shannon Diversity Index (WGS)	1.12	25	0.04 (-0.03, 0.12)	0.26	1.13	25	0.06 (-0.02, 0.13)	0.15
Inverse Simpson Index (WGS)	1.16	25	0.10 (-0.13, 0.33)	0.39	1.15	25	0.12 (-0.11, 0.35)	0.30
	Week 8 to week 12 return to control				**Week 8 to week 16 return to control**			
Shannon Diversity Index (16S)	1.06	27	0.05 (-0.01, 0.11)	0.08	1.03	28	0.03 (-0.04, 0.08)	0.50
Inverse Simpson Index (16S)	1.32	27	0.17 (-0.02, 0.37)	0.08	1.19	28	0.04 (-0.17, 0.25)	0.69
Shannon Diversity Index (WGS)	1.07	23	0.04 (-0.03, 0.10)	0.25	1.01	27	-0.01 (-0.07, 0.05)	0.76
Inverse Simpson Index (WGS)	1.08	23	0.04 (-0.17, 0.26)	0.68	1.01	27	-0.06 (-0.26, 0.14)	0.54
	Beans second, n=20							
	Baseline to week 4 remain on control				**Baseline to week 8 remain on control**			
Shannon Diversity Index (16S)	1.02	19	0.01 (-0.06, 0.07)	0.79	1.07	20	0.05 (-0.04, 0.13)	0.25
Inverse Simpson Index (16S)	0.99	19	-0.08 (-0.28, 0.11)	0.37	1.15	20	0.06 (-0.14, 0.26)	0.54
Shannon Diversity Index (WGS)	0.84	18	-0.01 (-0.15, 0.13)	0.90	0.84	20	0.03 (-0.11, 0.17)	0.62
Inverse Simpson Index (WGS)	1.05	18	-0.14 (-0.45, 0.17)	0.38	1.26	20	-0.06 (-0.36, 0.24)	0.68
	Week 8 to Week 12 on intervention				**Week 8 to Week 16 on intervention**			
Shannon Diversity Index (16S)	0.95	20	-0.06 (-0.14, 0.01)	0.10	1.02	20	0.02 (-0.03, 0.06)	0.45
Inverse Simpson Index (16S)	0.99	20	-0.15 (-0.39, 0.10)	0.22	1.27	20	0.18 (0.02, 0.35)	0.03
Shannon Diversity Index (WGS)	1.07	20	-0.02 (-0.11, 0.07)	0.68	1.02	20	0.05 (-0.04, 0.14)	0.28
Inverse Simpson Index (WGS)	1.13	20	-0.05 (-0.34, 0.24)	0.75	1.23	20	0.22 (-0.07, 0.51)	0.13

[a] Fold change as a ratio between follow up and baseline value, the sample size for each analysis might not equal to 48 or 28 or 20 due to missing at any time point
[b] Generalized linear mixed models with random intercept, natural log transformation applied
[c] P-value by Wald test.

Table S8. Changes in the relative abundance of selected prevalent[a] genus-level taxa among the full trial cohort (intent-to-treat n =48)

	Baseline to week 4 on intervention (full trial cohort)					Baseline to week 8 on intervention (full trial cohort)				
	Fold change[b]	n	Effect estimate[c] (log) (95%CI)	p-value[d]	q-value[e]	Fold change[b]	n	Effect estimate[c] (log) (95% CI)	p-value[d]	q-value[e]
Taxa										
Faecalibacterium	1.42	42	0.002 (-0.29, 0.30)	0.99	0.99	3.06	42	0.40 (0.02, 0.78)	0.04	0.12
Roseburia	1.14	39	-0.49 (-0.86, -0.13)	0.01	0.03	1.49	39	-0.18 (-0.50, 0.14)	0.27	0.41
Streptococcus	0.86	39	-0.48 (-0.92, -0.04)	0.03	0.045	5.93	39	0.05 (-0.53, 0.64)	0.85	0.55
	Beans first, n=28									
	Baseline to week 4 on intervention					**Baseline to week 8 on intervention**				
Faecalibacterium	1.45	25	0.06 (-0.29, 0.41)	0.72	0.72	2.41	25	0.28 (-0.18, 0.75)	0.22	0.66
Roseburia	1.20	25	-0.48 (-0.92, -0.04)	0.03	0.09	1.86	25	-0.06 (-0.58, 0.45)	0.80	0.80
Streptococcus	1.01	20	-0.34 (-1.06, 0.39)	0.34	0.51	3.25	20	0.18 (-0.58, 0.94)	0.62	0.80
	Week 8 to week 12 return to control					**Week 8 to week 16 return to control**				
Faecalibacterium	1.70	25	-0.52 (-1.10, 0.06)	0.08	0.24	1.67	25	-0.69 (-1.30, -0.08)	0.03	0.09
Roseburia	1.79	23	-0.20 (-0.83, 0.44)	0.53	0.53	2.04	22	-0.03 (-0.58, 0.52)	0.91	0.91
Streptococcus	4.08	17	0.42 (-0.36, 1.21)	0.27	0.41	2.87	18	-0.14 (-1.01, 0.72)	0.73	0.91
	Beans second, n=20									
	Baseline to week 4 remain on control					**Baseline to week 8 remain on control**				
Faecalibacterium	1.90	15	0.49 (-0.01, 0.99)	0.05	0.15	1.76	16	0.16 (-0.34, 0.67)	0.50	0.50
Roseburia	1.57	15	-0.25 (-1.08, 0.59)	0.54	0.94	1.21	16	0.20 (-0.24, 0.63)	0.35	0.50
Streptococcus	2.45	16	0.12 (-0.56, 0.79)	0.71	0.94	6.31	17	0.32 (-0.44, 1.09)	0.39	0.50
	Week 8 to Week 12 on intervention					**Week 8 to Week 16 on intervention**				
Faecalibacterium	1.38	17	-0.10 (-0.69, 0.49)	0.72	0.72	4.02	17	0.58 (-0.12, 1.27)	0.10	0.15
Roseburia	1.03	14	-0.51 (-1.25, 0.23)	0.16	0.24	0.83	14	-0.30 (-0.58, -0.02)	0.04	0.12
Streptococcus	0.70	19	-0.61 (-1.21, -0.02)	0.04	0.12	8.75	19	-0.04 (-1.03, 0.94)	0.93	0.93

[a] Selected results ($p<0.05$) of an analysis of 17 taxa meeting 80% prevalence at baseline.
[b] Fold change as a ratio between follow up and baseline value, the sample size for each analysis might not equal to 48 or 28 or 20 due to missing at any time point
[c] Generalized linear mixed models with random intercept, natural log transformation applied
[d] P-value by Wald test.
[e] Multiple correction using False Discovery Rate Benjamini–Hochberg adjustment

Table S9. Change in alpha diversity and the relative abundance of selected individual genus-level taxa across three on-intervention time points in the full trial cohort (intent-to-treat, n=48)

	Baseline to week 4 to week 8 on intervention	
	Effect estimate[a] (log) (95%CI)	**p-value[b]**
Shannon Diversity Index (16S)	0.02 (-0.004, 0.05)	0.10
Inverse Simpson Index (16S)	**0.08 (0.01, 0.16)**	**0.03**
Shannon Diversity Index (WGS)	0.03 (-0.002,0.05)	0.07
Inverse Simpson Index (WGS)	0.08 (-0.01, 0.17)	0.07
Faecalibacterium	**0.20 (0.03, 0.37)**	**0.02**
Roseburia	-0.07 (-0.25, 0.11)	0.44
Streptococcus	0.02 (-0.27, 0.30)	0.90

[a] Generalized linear mixed model with random intercept, natural log transformation applied
[b] P-value by Wald test

Table S10. On-intervention changes in the relative abundance of genus-level taxa among the full trial cohort using MaAsLin2[a] (intent-to-treat, n=48)

	Baseline to week 4 on-intervention					Baseline to week 8 on-intervention				
Genus-level taxa[b]	**N≠0[c]**	**β**	**SD**	**pval[d]**	**qval[e]**	**N≠0[c]**	**β**	**SD**	**pval[d]**	**qval[e]**
Ruminococcus	111	0.190	0.084	0.026	0.405	111	0.114	0.084	0.178	0.751
Lachnospiraceae_NK4A136_group	103	0.174	0.077	0.025	0.405	103	0.122	0.077	0.115	0.681
Eubacterium_siraeum_group	56	0.152	0.075	0.045	0.484	56	0.060	0.075	0.423	0.874
Odoribacter	65	0.168	0.061	0.007	0.405	65	0.140	0.061	0.023	0.405
Bifidobacterium	99	0.104	0.081	0.205	0.805	99	0.179	0.081	0.030	0.405
Eubacterium_brachy_group	42	0.025	0.035	0.472	0.888	42	0.082	0.035	0.021	0.405
Collinsella	72	-0.025	0.057	0.662	0.950	72	-0.162	0.057	0.006	0.405
Fournierella	25	-0.075	0.031	0.017	0.405	25	-0.044	0.031	0.151	0.751
Streptococcus	109	-0.235	0.109	0.033	0.405	109	-0.078	0.109	0.477	0.888
Oscillospiraceae(NK4A214_group)	84	-0.189	0.080	0.021	0.405	84	-0.066	0.080	0.414	0.874
Ruminococcus_torques	105	-0.213	0.087	0.016	0.405	105	-0.194	0.087	0.027	0.405
Oscillibacter	115	-0.172	0.086	0.049	0.495	115	-0.104	0.086	0.232	0.823
Escherichia_Shigella	74	-0.291	0.119	0.017	0.405	74	-0.041	0.119	0.731	0.953
Eubacterium_eligens_group	66	0.157	0.072	0.032	0.405	66	0.060	0.072	0.403	0.874

[a] MaAsLin2, a linear mixed model with random intercept of normalized relative abundance (natural log scale; half the minimum relative abundance as pseudo count) for use in sparse, compositional microbial communities

[b] Selected taxa shown with p-value <0.05 in either baseline to 4 weeks on intervention or baseline to 8 weeks on intervention for full trial cohort

[c] Refers to number of samples with non-zero values for feature among N=144 total samples (3 samples for each of the 48 patients)

[d] P-value by Wald test.

[e] Multiple correction using False Discovery Rate Benjamini–Hochberg adjustment

Table S11. On-intervention changes in the relative abundance of individual genus-level taxa among participants randomized to the intervention first using MaAsLin2[a] (intent-to-treat, n=28)

	On-intervention Intervention first (n=28)									
	Baseline to week 4 on-intervention					Baseline to week 8 on-intervention				
	N≠0[c]	β	SD	pval[d]	qval[e]	N≠0[c]	β	SD	pval[d]	qval[e]
Genus-level taxa[b]										
Ruminococcus	41	0.25	0.10	0.02	0.57	42	0.11	0.11	0.33	0.87
Lachnospiraceae_NK4A136_group	42	0.17	0.10	0.10	0.64	42	0.12	0.12	0.30	0.87
Eubacterium_siraeum_group	23	0.11	0.07	0.15	0.64	23	0.11	0.11	0.34	0.87
Odoribacter	24	0.10	0.08	0.18	0.64	26	0.11	0.08	0.19	0.87
Bifidobacterium	35	0.08	0.12	0.50	0.78	37	0.19	0.12	0.13	0.87
Eubacterium_brachy_group	10	-0.02	0.04	0.59	0.85	17	0.10	0.05	0.05	0.87
Collinsella	29	-0.07	0.06	0.26	0.66	27	-0.12	0.07	0.11	0.87
Fournierella	9	-0.11	0.05	0.04	0.57	11	-0.07	0.04	0.10	0.87
Streptococcus	39	-0.14	0.13	0.29	0.68	38	-0.05	0.16	0.74	0.98
Oscillospiraceae (NK4A214_group)	36	-0.16	0.11	0.16	0.64	37	-0.07	0.09	0.44	0.98
Ruminococcus_torques	40	-0.19	0.13	0.15	0.64	36	-0.09	0.12	0.46	0.98
Oscillibacter	49	-0.22	0.11	0.06	0.57	48	-0.07	0.13	0.60	0.98
Escherichia_Shigella	30	-0.32	0.15	0.05	0.57	32	-0.23	0.18	0.20	0.87
Eubacterium_eligens_group	28	0.24	0.11	0.04	0.57	24	0.03	0.07	0.73	0.98

[a] MaAsLin2, a linear mixed model with random intercept of normalized relative abundance (natural log scale; half the minimum relative abundance as pseudo count) for use in sparse, compositional microbial communities
[b] Selected taxa shown with p-value <0.05 in either baseline to 4 weeks on intervention or baseline to 8 weeks on intervention for full trial cohort
[c] Refers to number of samples with non-zero values for feature at the designated time-points (2 samples for each of the 28 patients, N=56 total samples)
[d] P-value by Wald test
[e] Multiple correction using False Discovery Rate Benjamini–Hochberg adjustment

Table S12. On-control changes in the relative abundance of individual taxa among participants randomized to the intervention first using MaAsLin2[a] (intent-to-treat, n=28)

	Return to control Intervention first (n=28)									
	Week 8 to Week 12 on-control					Week 8 to Week 16 on-control				
	N≠0[c]	β	SD	pval[d]	qval[e]	N≠0[c]	β	SD	pval[d]	qval[e]
Genus-level taxa[b]										
Ruminococcus	42	0.07	0.08	0.37	0.70	43	0.04	0.08	0.59	0.95
Lachnospiraceae_NK4A136_group	40	-0.17	0.13	0.21	0.57	43	-0.03	0.13	0.81	0.97
Eubacterium_siraeum_group	24	-0.08	0.08	0.34	0.69	21	-0.07	0.13	0.59	0.95
Odoribacter	26	0.09	0.11	0.45	0.72	25	0.01	0.08	0.91	0.99
Bifidobacterium	38	0.00	0.14	1.00	1.00	40	-0.13	0.12	0.27	0.95
Eubacterium_brachy_group	20	0.02	0.04	0.60	0.83	20	-0.05	0.05	0.39	0.95
Collinsella	26	0.24	0.09	0.01	0.28	27	0.13	0.10	0.20	0.95
Fournierella	13	0.17	0.07	0.02	0.29	9	0.07	0.05	0.19	0.95
Streptococcus	38	0.28	0.14	0.06	0.39	39	0.07	0.12	0.57	0.95
Oscillospiraceae(NK4A214_group)	37	0.22	0.09	0.02	0.29	37	0.05	0.11	0.69	0.95
Ruminococcus_torques	35	0.08	0.12	0.51	0.77	36	0.20	0.13	0.13	0.92
Oscillibacter	49	0.05	0.11	0.61	0.83	44	-0.04	0.12	0.73	0.96
Escherichia_Shigella	29	0.07	0.19	0.71	0.88	32	0.33	0.18	0.08	0.87
Eubacterium_eligens_group	27	-0.02	0.09	0.86	0.95	24	-0.04	0.09	0.63	0.95

[a] MaAsLin2, a linear mixed model with random intercept of normalized relative abundance (natural log scale; half the minimum relative abundance as pseudo count) for use in sparse, compositional microbial communities
[b] Selected taxa shown with p-value <0.05 in either baseline to 4 weeks on intervention or baseline to 8 weeks on intervention for full trial cohort
[c] Refers to number of samples with non-zero values for feature at the designated time-points (2 samples for each of the 28 patients, N=56 total samples)
[d] P-value by Wald test
[e] Multiple correction using False Discovery Rate Benjamini–Hochberg adjustment

Table S13. On-control changes in the relative abundance of individual taxa among participants randomized to the intervention second using MaAsLin2[a] (intent-treat, n=20)

	Remain on control Intervention second (n=20)									
	Week 0 to Week 4 on-control					Week 0 to Week 8 on-control				
	N≠0[c]	β	SD	pval[d]	qval[e]	N≠0[c]	β	SD	pval[d]	qval[e]
Genus-level taxa[b]										
Ruminococcus	33	0.10	0.13	0.42	0.92	32	-0.11	0.09	0.24	0.76
Lachnospiraceae_NK4A136_group	25	-0.04	0.09	0.63	0.92	30	-0.07	0.11	0.55	0.90
Eubacterium_siraeum_group	16	0.05	0.10	0.65	0.92	14	-0.03	0.07	0.65	0.90
Odoribacter	19	0.09	0.11	0.44	0.92	16	-0.15	0.10	0.15	0.76
Bifidobacterium	29	-0.12	0.12	0.36	0.92	29	-0.05	0.14	0.74	0.90
Eubacterium_brachy_group	8	0.03	0.06	0.58	0.92	8	0.00	0.05	0.93	0.96
Collinsella	21	0.06	0.06	0.33	0.91	23	0.09	0.12	0.45	0.90
Fournierella	5	0.04	0.04	0.27	0.85	5	0.05	0.04	0.20	0.76
Streptococcus	32	0.01	0.14	0.95	1.00	36	0.24	0.18	0.20	0.76
Oscillospiraceae(NK4A214_group)	21	0.12	0.09	0.19	0.84	21	0.16	0.10	0.13	0.76
Ruminococcus_torques	35	0.09	0.12	0.49	0.92	37	0.17	0.12	0.17	0.76
Oscillibacter	26	0.20	0.11	0.09	0.68	29	0.27	0.12	0.04	0.63
Escherichia_Shigella	25	0.10	0.18	0.59	0.92	23	-0.03	0.12	0.79	0.90
Eubacterium_eligens_group	17	0.14	0.12	0.25	0.85	15	-0.03	0.13	0.81	0.90

[a] MaAsLin2, a linear mixed model with random intercept of normalized relative abundance (natural log scale; half the minimum relative abundance as pseudo count) for use in sparse, compositional microbial communities
[b] Selected taxa shown with p-value <0.05 in either baseline to 4 weeks on intervention or baseline to 8 weeks on intervention for full trial cohort
[c] Refers to number of samples with non-zero values for feature at the designated time-points (2 samples for each of the 20 patients, N=40 total samples)
[d] P-value by Wald test
[e] Multiple correction using False Discovery Rate Benjamini–Hochberg adjustment

Table S14. On-intervention changes in the relative abundance of individual taxa among participants randomized to the intervention second using MaAsLin2[a] (intent-to-treat, n=20

	On intervention Intervention second (n=20)									
	Baseline to week 4 on-intervention					Baseline to week 8 on-intervention				
	N≠0[c]	β	SD	pval[d]	qval[e]	N≠0[c]	β	SD	pval[d]	qval[e]
Genus-level taxa[b]										
Ruminococcus	32	0.10	0.14	0.49	0.79	32	0.12	0.16	0.47	0.88
Lachnospiraceae_NK4A136_group	27	0.19	0.12	0.12	0.56	27	0.12	0.12	0.34	0.79
Eubacterium_siraeum_group	14	0.21	0.11	0.07	0.53	14	-0.01	0.10	0.93	0.99
Odoribacter	17	0.26	0.12	0.05	0.53	16	0.19	0.09	0.06	0.76
Bifidobacterium	28	0.13	0.13	0.32	0.72	30	0.17	0.10	0.11	0.79
Eubacterium_brachy_group	12	0.09	0.05	0.08	0.53	13	0.06	0.04	0.16	0.79
Collinsella	23	0.04	0.09	0.67	0.82	20	-0.22	0.09	0.02	0.76
Fournierella	7	-0.02	0.04	0.61	0.80	8	-0.00	0.04	0.91	0.97
Streptococcus	35	-0.37	0.12	0.01	0.45	36	-0.11	0.18	0.53	0.91
Oscillospiraceae (NK4A214_group)	20	-0.19	0.10	0.06	0.53	21	-0.04	0.13	0.78	0.97
Ruminococcus_torques	34	-0.25	0.09	0.01	0.45	34	-0.35	0.13	0.01	0.76
Oscillibacter	31	-0.10	0.10	0.31	0.72	29	-0.15	0.14	0.27	0.79
Escherichia_Shigella	19	-0.25	0.15	0.11	0.53	21	0.22	0.15	0.14	0.79
Eubacterium_eligens_group	23	0.04	0.09	0.67	0.82	17	0.11	0.12	0.34	0.79

[a] MaAsLin2, a linear mixed model with random intercept of normalized relative abundance (natural log scale; half the minimum relative abundance as pseudo count) for use in sparse, compositional microbial communities
[b] Selected taxa shown with p-value <0.05 in either baseline to 4 weeks on intervention or baseline to 8 weeks on intervention for full trial cohort
[c] Refers to number of samples with non-zero values for feature at the designated time-points (2 samples for each of the 20 patients, N=40 total samples)
[d] P-value by Wald test
[e] Multiple correction using False Discovery Rate Benjamini–Hochberg adjustment

Table S15. On-intervention changes across three time points in in full trial cohort using MaAsLin2[a] (intent-to-treat, n=48)

	Baseline to week 4 to week 8 on intervention (n=48)				
Genus-level taxa[b]	**N≠0[c]**	**β**	**SD**	**pval[d]**	**qval[e]**
Collinsella	72	-0.27	0.10	0.01	0.63
Eubacterium_brachy_group	42	0.14	0.06	0.02	0.66
Odoribacter	65	0.23	0.10	0.03	0.66
Ruminococcus_torques	105	-0.32	0.14	0.03	0.66
Bifidobacterium	99	0.30	0.13	0.03	0.66
Family_Lachnospiraceae__GCA_9001	17	-0.08	0.04	0.04	0.68
Clostridium innocuum group	32	-0.13	0.06	0.05	0.72

[a] MaAsLin2, a linear mixed model with random intercept of normalized relative abundance (natural log scale; half the minimum relative abundance as pseudo count) for use in sparse, compositional microbial communities

[b] Selected genus shown with p-value <0.05

[c] The total number of stool samples from 48 patients across 3 time points (n=144)

[d] P-value by Wald test

[e] Multiple correction using False Discovery Rate Benjamini–Hochberg adjustment

Table S16. On-intervention changes in the relative abundance of individual species among the full trial subcohort with WGS sequencing using MaAsLin2[a] (intent-to-treat, n=48)

Species[b]	Baseline to week 4 on-intervention					Baseline to week 8 on-intervention				
	N≠0[c]	β	SD	pval[d]	qval[e]	N≠0[c]	β	SD	pval[d]	qval[e]
Adlercreutzia_equolifaciens	130	-0.282	0.124	0.025	0.454	130	-0.230	0.124	0.066	0.525
Anaerofustis_stercorihominis	16	-0.244	0.098	0.015	0.454	16	-0.189	0.098	0.058	0.513
Asaccharobacter_celatus	137	-0.231	0.102	0.026	0.454	137	-0.137	0.102	0.184	0.677
Bacteroides_nordii	25	-0.004	0.081	0.965	0.983	25	-0.187	0.081	0.024	0.454
Bacteroides_salyersiae	22	0.123	0.060	0.042	0.483	22	0.082	0.060	0.176	0.677
Bifidobacterium_adolescentis	53	0.040	0.049	0.420	0.808	53	0.110	0.049	0.027	0.454
Bilophila_wadsworthia	33	-0.200	0.095	0.038	0.483	33	-0.261	0.095	0.007	0.454
Blautia_hydrogenotrophica	30	-0.223	0.090	0.015	0.454	30	-0.199	0.090	0.029	0.454
Butyricimonas_virosa	28	0.086	0.040	0.034	0.463	28	0.046	0.040	0.259	0.691
Christensenella_minuta	22	-0.209	0.097	0.035	0.463	22	-0.244	0.097	0.014	0.454
Clostridium_bolteae	53	-0.333	0.129	0.011	0.454	53	-0.247	0.129	0.058	0.513
Clostridium_innocuum	49	-0.285	0.117	0.017	0.454	49	-0.271	0.117	0.023	0.454
Denitrobacterium_detoxificans	100	-0.399	0.184	0.032	0.463	100	-0.119	0.184	0.518	0.837
Dialister_sp_CAG_357	18	0.011	0.033	0.731	0.920	18	0.076	0.033	0.025	0.454
Enterorhabdus_caecimuris	137	-0.245	0.105	0.022	0.454	137	-0.137	0.105	0.192	0.677
Eubacterium_rectale	119	0.271	0.130	0.039	0.483	119	0.398	0.130	0.003	0.454
Eubacterium_sp_CAG_38	41	0.266	0.135	0.053	0.513	41	0.303	0.135	0.028	0.454
Firmicutes_bacterium_CAG_145	38	-0.149	0.096	0.126	0.626	38	-0.191	0.096	0.050	0.513
Gordonibacter_pamelaeae	135	-0.310	0.128	0.018	0.454	135	-0.185	0.128	0.153	0.660
Oxalobacter_formigenes	18	0.028	0.091	0.762	0.920	18	0.210	0.091	0.023	0.454
Roseburia_sp_CAG_471	65	-0.089	0.082	0.281	0.704	65	-0.220	0.082	0.009	0.454
Ruminococcus_torques	81	-0.248	0.097	0.012	0.454	81	-0.156	0.097	0.111	0.599
Ruthenibacterium_lactatiformans	134	-0.258	0.099	0.011	0.454	134	-0.030	0.099	0.761	0.920
Sellimonas_intestinalis	41	-0.243	0.117	0.041	0.483	41	-0.180	0.117	0.129	0.626
Sharpea_azabuensis	65	-0.196	0.114	0.090	0.554	65	-0.281	0.114	0.016	0.454
Veillonella_atypica	28	0.076	0.106	0.474	0.832	28	0.231	0.106	0.032	0.463

[a] MaAsLin2, a linear mixed model with random intercept of normalized relative abundance (natural log scale; half the minimum relative abundance as pseudo count) for use in sparse, compositional microbial communities
[b] Taxonomic profiling via MetaPhlAn3. Selected species shown with p-value <0.05 in either baseline to 4 weeks on-intervention or baseline to 8 weeks on-intervention for full trial cohort
[c] Refers to number of samples with non-zero values for feature among the total number of stool samples from 48 patients across 3 time points does not sum to 144 (n=140) because some stool samples failed library preparation for WGS sequencing.
[d] P-value by Wald test.
[e] Multiple correction using False Discovery Rate Benjamini–Hochberg adjustment

Table S17. On-intervention changes across three time points in species among the full trial subcohort with WGS sequencing using MaAsLin2[a]

	Baseline to week 4 to week 8 on intervention (n=48[a])				
Species[b]	**N≠0[c]**	**β**	**SD**	**pval[d]**	**qval[e]**
Eubacterium_rectale	119	0.198	0.065	0.003	0.416
Bilophila_wadsworthia	33	-0.130	0.047	0.007	0.416
Roseburia_sp_CAG_471	65	-0.110	0.041	0.009	0.416
Christensenella_minuta	22	-0.122	0.049	0.014	0.416
Sharpea_azabuensis	65	-0.140	0.057	0.016	0.416
Oxalobacter_formigenes	18	0.105	0.045	0.022	0.416
Bacteroides_nordii	25	-0.094	0.041	0.024	0.416
Clostridium_innocuum	49	-0.135	0.059	0.024	0.416
Dialister_sp_CAG_357	18	0.038	0.017	0.025	0.416
Bifidobacterium_adolescentis	53	0.055	0.024	0.026	0.416
Eubacterium_sp_CAG_38	41	0.151	0.068	0.028	0.416
Veillonella_atypica	28	0.116	0.053	0.031	0.416
Blautia_hydrogenotrophica	30	-0.099	0.045	0.031	0.416
Firmicutes_bacterium_CAG_145	38	-0.095	0.048	0.050	0.600

[a] MaAsLin2, a linear mixed model with random intercept of normalized relative abundance (natural log scale; half the minimum relative abundance as pseudo count) for use in sparse, compositional microbial communities
[b] Taxonomic profiling via MetaPhlAn3. Selected species shown with p-value <0.05
[c] The total number of stool samples from 48 patients across 3 time points does not sum to 144 (n=140) because some stool samples failed library preparation for WGS sequencing; however all 48 patients were able to contribute to this analysis.
[d] P-value by Wald test
[e] Multiple correction using False Discovery Rate Benjamini–Hochberg adjustment

Table S18. On-intervention changes in the relative abundance of metagenomic pathways among the full trial subcohort with WGS sequencing using MaAsLin2[a]

	Baseline to week 4 on intervention					Baseline to week 8 on intervention				
Pathways[b]	**N≠0[c]**	**β**	**SD**	**pval[d]**	**qval[e]**	**N≠0[c]**	**β**	**stderr**	**pval[d]**	**qval[e]**
γ-glutamyl cycle	139	-0.446	0.178	0.014	0.335	139	-0.043	0.178	0.808	0.941
acetyl-CoA fermentation to butanoate II	136	-0.535	0.218	0.016	0.335	136	-0.150	0.218	0.492	0.829
acetylene degradation	105	-0.760	0.338	0.027	0.364	105	-0.277	0.338	0.415	0.802
adenosine nucleotides degradation II	140	-0.566	0.199	0.005	0.288	140	-0.247	0.199	0.217	0.695
allantoin degradation to glyoxylate II	26	-0.628	0.229	0.007	0.288	26	-0.078	0.229	0.734	0.907
aspartate superpathway	126	-0.762	0.348	0.031	0.392	126	-0.251	0.348	0.473	0.807
C4 photosynthetic carbon assimilation cycle, NADP-ME type	131	-0.822	0.307	0.009	0.288	131	-0.175	0.307	0.570	0.851
C4 photosynthetic carbon assimilation cycle, PEPCK type	131	-0.780	0.306	0.012	0.335	131	-0.206	0.306	0.502	0.840
D-galactarate degradation I	114	-0.862	0.347	0.015	0.335	114	-0.273	0.347	0.434	0.802
D-glucarate degradation I	103	-0.957	0.342	0.006	0.288	103	-0.506	0.342	0.142	0.598
fatty acid β-oxidation I	124	-0.678	0.295	0.024	0.352	124	-0.239	0.295	0.420	0.802
fatty acid β-oxidation II (peroxisome)	124	-0.670	0.301	0.029	0.379	124	-0.238	0.301	0.432	0.802
formaldehyde assimilation III (dihydroxyacetone cycle)	130	-0.623	0.265	0.021	0.349	130	-0.400	0.265	0.134	0.593
gluconeogenesis I	140	-0.314	0.103	0.003	0.276	140	-0.193	0.103	0.063	0.507
glucose and glucose-1-phosphate degradation	116	-0.960	0.426	0.027	0.364	116	-0.054	0.426	0.899	0.973
glutaryl-CoA degradation	138	-0.469	0.194	0.017	0.335	138	-0.119	0.194	0.541	0.844
glycogen degradation I (bacterial)	115	-0.837	0.334	0.014	0.335	115	-0.094	0.334	0.778	0.928
glycolysis VI (metazoan)	140	-0.329	0.128	0.012	0.335	140	-0.193	0.128	0.134	0.593
heme biosynthesis I (aerobic)	107	-0.926	0.345	0.009	0.288	107	-0.400	0.345	0.250	0.705
heme biosynthesis II (anaerobic)	116	-0.839	0.351	0.019	0.335	116	-0.607	0.351	0.087	0.551
hexitol fermentation to lactate, formate, ethanol and acetate	138	-0.861	0.223	0.000	0.162	138	-0.373	0.223	0.097	0.568
L-1,2-propanediol degradation	132	-0.750	0.288	0.011	0.319	132	-0.467	0.288	0.108	0.568
L-glutamine biosynthesis III	131	-0.626	0.277	0.026	0.364	131	-0.252	0.277	0.365	0.792
L-homoserine and L-methionine biosynthesis	140	-0.567	0.193	0.004	0.276	140	-0.259	0.193	0.183	0.661
L-isoleucine biosynthesis I (from threonine)	140	-0.108	0.055	0.053	0.464	140	-0.153	0.055	0.007	0.288
L-isoleucine biosynthesis III	140	-0.142	0.057	0.015	0.335	140	-0.185	0.057	0.002	0.276
L-isoleucine biosynthesis IV	140	-0.362	0.199	0.072	0.551	140	-0.474	0.199	0.020	0.339
L-lysine biosynthesis III	140	-0.085	0.058	0.145	0.598	140	-0.137	0.058	0.020	0.339
L-methionine biosynthesis I	140	-0.596	0.211	0.006	0.288	140	-0.267	0.211	0.209	0.688
L-valine biosynthesis	140	-0.108	0.055	0.053	0.464	140	-0.153	0.055	0.007	0.288
methylphosphonate degradation I	33	-0.863	0.358	0.018	0.335	33	0.318	0.358	0.377	0.800
mixed acid fermentation	130	-0.814	0.302	0.008	0.288	130	-0.354	0.302	0.243	0.700
NAD salvage pathway I	108	-1.339	0.408	0.001	0.276	108	-0.119	0.408	0.772	0.925
nitrate reduction V (assimilatory)	70	-1.022	0.440	0.022	0.352	70	-0.429	0.440	0.332	0.764
peptidoglycan maturation (meso-diaminopimelate containing)	140	-0.182	0.084	0.032	0.392	140	-0.086	0.084	0.304	0.738
phosphatidylcholine acyl editing	81	-0.908	0.371	0.016	0.335	81	-0.337	0.371	0.367	0.792
purine nucleobases degradation I (anaerobic)	136	-0.405	0.184	0.030	0.392	136	-0.137	0.184	0.456	0.802
purine nucleotides degradation II (aerobic)	140	-0.444	0.170	0.010	0.319	140	-0.205	0.170	0.230	0.699
pyruvate fermentation to isobutanol (engineered)	140	-0.101	0.058	0.084	0.551	140	-0.170	0.058	0.004	0.276
Rubisco shunt	70	-1.051	0.439	0.019	0.335	70	-0.281	0.439	0.524	0.844
stearate biosynthesis III (fungi)	46	-0.908	0.401	0.026	0.364	46	-0.524	0.401	0.194	0.674
superpathay of heme biosynthesis from glutamate	87	-1.037	0.348	0.004	0.276	87	-0.267	0.348	0.446	0.802
superpathway of (Kdo)2-lipid A biosynthesis	39	-0.786	0.360	0.032	0.392	39	0.167	0.360	0.644	0.868
superpathway of allantoin degradation in plants	26	-0.628	0.229	0.007	0.288	26	-0.078	0.229	0.734	0.907
superpathway of branched amino acid biosynthesis	140	-0.125	0.055	0.025	0.364	140	-0.170	0.055	0.003	0.276
superpathway of D-glucarate and D-galactarate degradation	114	-0.862	0.347	0.015	0.335	114	-0.273	0.347	0.434	0.802
superpathway of glucose and xylose degradation	139	-0.327	0.140	0.022	0.350	139	-0.092	0.140	0.513	0.844
superpathway of glycerol degradation to 1,3-propanediol	138	-0.511	0.212	0.018	0.335	138	-0.422	0.212	0.049	0.449

superpathway of hexitol degradation (bacteria)	139	-0.485	0.149	0.002	0.276	139	-0.144	0.149	0.337	0.771
superpathway of L-lysine, L-threonine and L-methionine biosynthesis I	127	-0.784	0.336	0.022	0.350	127	-0.153	0.336	0.650	0.868
superpathway of L-methionine biosynthesis (transsulfuration)	140	-0.550	0.186	0.004	0.276	140	-0.250	0.186	0.182	0.658
superpathway of purine deoxyribonucleosides degradation	135	-0.833	0.241	0.001	0.276	135	-0.424	0.241	0.082	0.551
superpathway of purine nucleotide salvage	129	-0.676	0.310	0.032	0.392	129	-0.334	0.310	0.283	0.726
superpathway of S-adenosyl-L-methionine biosynthesis	140	-0.555	0.189	0.004	0.276	140	-0.260	0.189	0.173	0.646
superpathway of sulfur amino acid biosynthesis (Saccharomyces cerevisiae)	89	-1.114	0.462	0.018	0.335	89	-0.525	0.462	0.260	0.717
TCA	135	-0.476	0.186	0.012	0.335	135	-0.262	0.186	0.163	0.631
TCA cycle II (plants and fungi)	137	-0.563	0.206	0.008	0.288	137	-0.305	0.206	0.142	0.598
TCA cycle VI (obligate autotrophs)	131	-0.886	0.325	0.008	0.288	131	-0.227	0.325	0.487	0.822
TCA cycle VII (acetate-producers)	73	-0.737	0.321	0.024	0.352	73	-0.405	0.321	0.210	0.690
tetrapyrrole biosynthesis II (from glycine)	101	-1.091	0.472	0.023	0.352	101	-0.151	0.472	0.749	0.914
thiamin salvage II	140	-0.148	0.068	0.033	0.392	140	-0.090	0.068	0.192	0.674
thiazole biosynthesis I (E. coli)	140	-0.160	0.066	0.018	0.335	140	-0.163	0.066	0.016	0.335

[a] MaAsLin2, a linear mixed model with random intercept of normalized relative abundance (natural log scale; half the minimum relative abundance as pseudo count) for use in sparse, compositional microbial communities

[b] Functional profiling via HUMAnN3. Selected pathways shown with p-value <0.05 in baseline to 4 weeks to 8 weeks on intervention for full trial cohort

[c] Refers to number of samples with non-zero values for feature among N=140 total samples. The total number of stool samples from 48 patients across 3 time points does not sum to 144 (n=140) because some stool samples failed library preparation for WGS sequencing.

[d] P-value by Wald test.

[e] Multiple correction using False Discovery Rate Benjamini–Hochberg adjustment

Table S19. On-intervention changes across three time points in metagenomic pathways among the full trial subcohort with WGS sequencing, using MaAsLin2[a], n=48

	Baseline to week 4 to week 8 on intervention (n=48)				
Pathways[b]	**N≠0[c]**	**β**	**SD**	**pval[d]**	**qval[e]**
L-isoleucine biosynthesis III	140	-0.028	0.009	0.002	0.531
superpathway of branched chain amino acid biosynthesis	140	-0.025	0.008	0.003	0.531
Engineered Pathway: pyruvate fermentation to isobutanol (engineered)	140	-0.025	0.009	0.004	0.533
L-isoleucine biosynthesis I (from threonine)	140	-0.023	0.008	0.007	0.535
L-valine biosynthesis	140	-0.023	0.008	0.007	0.535
thiazole biosynthesis I (E. coli)	140	-0.024	0.010	0.017	0.930
L-lysine biosynthesis III	140	-0.021	0.009	0.020	0.930
L-isoleucine biosynthesis IV	140	-0.071	0.030	0.020	0.930
formaldehyde assimilation II (assimilatory RuMP Cycle)	138	-0.073	0.035	0.043	0.930
sulfoquinovose degradation I	35	0.082	0.040	0.045	0.930
dTDP-3-acetamido-α-D-fucose biosynthesis	32	-0.045	0.022	0.047	0.930
L-phenylalanine degradation IV (mammalian, via side chain)	140	-0.015	0.007	0.050	0.930

[a] MaAsLin2, a linear mixed model with random intercept of normalized relative abundance (natural log scale; half the minimum relative abundance as pseudo count) for use in sparse, compositional microbial communities
[b] Functional profiling via HUMAnN3. Selected pathways shown with p-value <0.05
[c] The total number of stool samples from 48 patients across 3 time points does not sum to 144 (n=140) because some stool samples failed library preparation for WGS sequencing; however all 48 patients were able to contribute to this analysis.
[d] P-value by Wald test
[e] Multiple correction using False Discovery Rate Benjamini–Hochberg adjustment

Table S20. On-intervention changes across three time points in circulating metabolites among the full trial cohort, n=48

	Baseline to week 4 to week 8 on intervention (n=48[a])		
	Effect estimate[c] (log) (95%CI)	**p-value[d]**	**q-value[e]**
Metabolites[b]			
Pipecolic acid (PA)	0.18 (0.14, 0.23)	<0.001	<0.001
S-(5'-Adenosyl)-L-methionine (SAM)	0.15 (0.09, 0.21)	<0.001	0.001
Trigonelline	0.28 (0.17, 0.40)	<0.001	0.001
Theophylline	0.14 (0.06, 0.22)	0.001	0.10
Indole derivative	-0.09 (-0.14, -0.04)	0.001	0.09
Caffeine	0.11 (0.04, 0.18)	0.002	0.25
PARAXANTHINE; THEOBROMINE; THEOPHYLLINE	0.15 (0.04, 0.26)	0.01	0.69
Cholesterol Ester (22:4)	-0.04 (-0.07, -0.01)	0.01	0.69
LactosylCeramide (32:0)	-0.04 (-0.07, -0.01)	0.01	0.69
URIDINE-5-MONOPHOSPHATE	0.05 (0.01, 0.09)	0.01	0.69
D-PANTOTHENIC ACID	-0.03 (-0.05, -0.01)	0.01	0.69
Triacylglycerol (55:8)	-0.07 (-0.13, -0.01)	0.02	0.69
nicotinate beta-D-ribonucleotide	0.06 (0.01, 0.10)	0.02	0.69
Prostaglandin A1	0.04 (0.01, 0.08)	0.02	0.69
Lysophosphatidylcholine (20:3)	-0.05 (-0.10, -0.01)	0.02	0.69
6-PHOSPHOGLUCONIC ACID	0.04 (0.01, 0.07)	0.02	0.69
L-CYSTINE	-0.05 (-0.09, -0.01)	0.02	0.69
MELATONIN	-0.02 (-0.04, -0.004)	0.02	0.69
GALACTITOL	-0.09 (-0.17, -0.02)	0.02	0.69
2-Methylbutyroylcarnitine	0.04 (0.01, 0.07)	0.02	0.69
2-Acetylpyrrolidine_exogenous	-0.01 (-0.02, -0.002)	0.02	0.73
ALPHA-D-GALACTOSE 1-PHOSPHATE	-0.09 (-0.17, -0.01)	0.03	0.88
p-Aminobenzoic Acid	-0.03 (-0.06, -0.003)	0.03	0.88
2'-DEOXYADENOSINE 5'-TRIPHOSPHATE	-0.05 (-0.1, -0.004)	0.03	0.88
Phosphatidylcholine (40:5)	-0.04 (-0.08, -0.003)	0.03	0.88
HYDROQUINONE	0.08 (0.01, 0.16)	0.03	0.88
Triacylglycerol (58:4)	-0.02 (-0.04, -0.001)	0.03	0.88
Triacylglycerol (57:3)	-0.02 (-0.04, -0.001)	0.04	0.88
Ceramide (39:1)	-0.04 (-0.08, -0.002)	0.04	0.96
GLYCERATE; N-ACETYL-L-ALANINE	0.02 (0.001, 0.03)	0.04	0.96
INOSINE	-0.02 (-0.04, -0.001)	0.04	0.97
Ceramide (40:2)	-0.04 (-0.08, -0.001)	0.04	0.97
3-Dehydrocarnitine	-0.03 (-0.05, -0.0002)	0.048	0.97
6-PHOSPHOGLUCONIC ACID	-0.08 (-0.16, -0.0002)	0.049	0.97
ROSMARINIC ACID	0.04 (0.0001, 0.08)	0.0497	0.97

[a] The total number of blood samples from 48 patients across 3 time points does not sum to 144 (n=140) because some blood samples were not collected due to missed clinic visits (whereas every stool sample was recovered) though all 48 patients were able to contribute to this primary circulating metabolite analysis.
[b] Selected metabolites shown with p-value <0.05
[c] Generalized linear mixed models with random intercept, natural log transformation applied
[d] P-value by Wald test
[e] Multiple correction using False Discovery Rate Benjamini–Hochberg adjustment

Table S21. Changes in circulating metabolites among the full trial cohort (intent-to-treat n=48)

	Fold change[b]	N	Effect estimate[c] (log) (95%CI)	p-value[d]	q-value[e]	Fold change[b]	n	Effect estimate[c] (log) (95% CI)	p-value[d]	q-value[e]
	Baseline to week 4 on intervention (full trial cohort)					**Baseline to week 8 on intervention (full trial cohort)**				
Metabolites[a]										
Pipecolic acid (PA)	1.54	46	0.39 (0.31, 0.47)	<0.001	<0.001	1.54	45	0.37 (0.28, 0.46)	<0.001	<0.001
S-(5'-Adenosyl)-L-methionine (SAM)	1.53	46	0.40 (0.31, 0.49)	<0.001	<0.001	1.42	45	0.30 (0.19, 0.40)	<0.001	<0.001
Trigonelline	1.28	46	0.49 (0.30, 0.68)	<0.001	0.001	1.65	45	0.57 (0.33, 0.82)	<0.001	0.004
Theophylline	0.92	46	-0.04 (-0.19, 0.12)	0.63	0.89	1.23	45	0.29 (0.13, 0.46)	0.001	0.15
Indole derivative	0.92	46	-0.09 (-0.18, -0.0003)	0.049	0.68	0.82	45	-0.18 (-0.29, -0.07)	0.002	0.19
Beans first, n=28										
	Baseline to week 4 on intervention					**Baseline to week 8 on intervention**				
Pipecolic acid (PA)	1.48	27	0.37 (0.28, 0.45)	<0.001	<0.001	1.54	26	0.38 (0.27, 0.49)	<0.001	<0.001
S-(5'-Adenosyl)-L-methionine (SAM)	1.60	27	0.43 (0.29, 0.57)	<0.001	<0.001	1.50	26	0.34 (0.21, 0.48)	<0.001	0.006
Trigonelline	1.25	27	0.43 (0.18, 0.68)	0.002	0.36	1.94	26	0.59 (0.28, 0.9)	0.001	0.10
Theophylline	0.85	27	-0.15 (-0.32, 0.03)	0.10	0.79	1.25	26	0.26 (0.06, 0.46)	0.01	0.59
Indole derivative	0.89	27	-0.15 (-0.27, -0.04)	0.01	0.48	0.81	26	-0.25 (-0.39, -0.12)	0.001	0.12
	Week 8 to week 12 return to control					**Week 8 to week 16 return to control**				
Pipecolic acid (PA)	0.66	22	-0.35 (-0.49, -0.21)	<0.001	0.02	0.80	25	-0.20 (-0.41, 0.004)	0.054	0.77
S-(5'-Adenosyl)-L-methionine (SAM)	0.70	22	-0.30 (-0.45, -0.14)	0.001	0.22	0.79	25	-0.20 (-0.33, -0.07)	0.004	0.57
Trigonelline	1.06	22	-0.25 (-0.61, 0.11)	0.16	0.96	0.94	25	-0.17 (-0.50, 0.15)	0.28	0.90
Theophylline	1.00	22	-0.05 (-0.30, -0.21)	0.72	0.99	1.01	25	-0.03 (-0.22, 0.17)	0.79	0.99
Indole derivative	1.10	22	-0.02 (0.29, 0.09)	0.09	0.96	1.24	25	0.20 (0.01, 0.40)	0.04	0.73
Beans second, n=20										
	Baseline to week 4 remain on control					**Baseline to week 8 remain on control**				
Pipecolic acid (PA)	0.95	18	-0.05 (-0.11, 0.01)	0.08	0.89	0.98	19	-0.03 (-0.11, 0.4)	0.33	0.77
S-(5'-Adenosyl)-L-methionine (SAM)	0.91	18	-0.10 (-0.19, -0.01)	0.02	0.87	0.93	19	-0.06 (-0.18, 0.06)	0.28	0.72
Trigonelline	1.12	18	-0.08 (-0.40, 0.25)	0.63	0.96	1.12	19	-0.11 (-0.52, 0.30)	0.58	0.92
Theophylline	1.21	18	0.20 (-0.11, 0.50)	0.19	0.90	1.08	19	0.0003 (-0.23, 0.23)	0.998	0.999
Indole derivative	0.92	18	-0.09 (-0.25, 0.07)	0.27	0.90	0.79	19	-0.19 (-0.38, 0.01)	0.06	0.52
	Week 8 to week 12 on intervention					**Week 8 to week 16 on intervention**				
Pipecolic acid (PA)	1.64	19	0.43 (0.26, 0.59)	<0.001	0.012	1.53	19	0.36 (0.19, 0.53)	<0.001	0.20
S-(5'-Adenosyl)-L-methionine (SAM)	1.44	19	0.35 (0.26, 0.45)	<0.001	<0.001	1.32	19	0.23 (0.05, 0.40)	0.01	0.90
Trigonelline	1.32	19	0.58 (0.28, 0.88)	0.001	0.154	1.25	19	0.55 (0.15, 0.96)	0.01	0.90
Theophylline	1.02	19	0.12 (-0.15, 0.39)	0.360	0.848	1.19	19	0.33 (0.04, 0.62)	0.03	0.90
Indole derivative	0.97	19	-0.003 (0.15,0.14)	0.960	0.995	0.85	19	-0.09 (-0.28, 0.09)	0.31	0.90

[a] Selected metabolites shown with False Discovery Rate Benjamini–Hochberg adjusted p-value <0.20 in baseline to 4 weeks to 8 weeks on intervention for full trial cohort
[b] Fold change as a ratio between follow up and baseline value, the sample size for each analysis might not equal to 48 or 28 or 20 due to missing at any time point
[c] Generalized linear mixed models with random intercept. Natural log transformation applied such that an effect estimates the slope above zero
[d] P-value by Wald test.
[e] Multiple correction using False Discovery Rate Benjamini–Hochberg adjustment

Table S22. On-intervention changes in circulating proteins associated with intestinal and systemic inflammatory and immune response (Olink) among the full trial cohort

	Baseline to week 8 on intervention (n=45[a])		
	Effect estimate[b] (log) (95%CI)	p-value[c]	q-value[d]
Olink 96 inflammation panel[a]			
FGF-19	0.32 (0.07,0.57)	0.01	0.48
IL-10RA	-0.29 (-0.51,-0.06)	0.01	0.48
TRANCE	-0.17 (-0.34,-0.003)	0.02	0.48
CD8A	-0.22 (-0.45,0.01)	0.03	0.48
PD-L1	-0.12 (-0.26,0.02)	0.04	0.48
CXCL1	-0.23 (-0.50,0.04)	0.05	0.48
uPA	-0.09 (-0.19,0.02)	0.05	0.48
CD5	-0.10 (-0.23,0.03)	0.06	0.48
IL18	-0.14 (-0.33,0.06)	0.08	0.48
Beta-NGF	-0.02 (-0.06,0.01)	0.09	0.48
IL-24	-0.15 (-0.37,0.07)	0.09	0.48
IL-2RB	-0.15 (-0.37,0.07)	0.09	0.48
CD244	-0.08 (-0.22,0.05)	0.11	0.48
IL4	0.17 (-0.11,0.44)	0.11	0.48
IL7	-0.14 (-0.37,0.09)	0.11	0.48
CCL4	-0.12 (-0.33,0.08)	0.12	0.48
CD40	-0.09 (-0.26,0.07)	0.13	0.48
CXCL5	-0.19 (-0.53,0.15)	0.13	0.48
IL13	0.11 (-0.09,0.31)	0.13	0.48
NT-3	-0.11 (-0.29,0.08)	0.13	0.48
CCL3	-0.09 (-0.26,0.08)	0.14	0.48
EN-RAGE	-0.22 (-0.62,0.18)	0.14	0.48
IFN-gamma	-0.14 (-0.40,0.12)	0.14	0.48
CSF-1	-0.04 (-0.11,0.04)	0.15	0.48
IL-12B	-0.11 (-0.31,0.10)	0.15	0.48
TNFSF14	-0.14 (-0.40,0.12)	0.15	0.48
DNER	-0.04 (-0.11,0.04)	0.16	0.48
FGF-5	0.04 (-0.05,0.13)	0.16	0.48
MCP-1	-0.07 (-0.21,0.08)	0.17	0.48
TSLP	0.11 (-0.12,0.34)	0.17	0.48
IL-20	-0.05 (-0.16,0.06)	0.18	0.48
LAP TGF-beta-1	-0.07 (-0.24,0.09)	0.18	0.48
OPG	0.06 (-0.07,0.18)	0.18	0.48
STAMBP	-0.19 (-0.59,0.22)	0.18	0.48
ADA	-0.07 (-0.23,0.09)	0.19	0.48
CASP-8	-0.12 (-0.40,0.15)	0.19	0.48
IL-20RA	0.07 (-0.09,0.22)	0.2	0.48
Inc Ctrl 1	-0.01 (-0.04,0.02)	0.2	0.48
MCP-3	-0.11 (-0.37,0.15)	0.2	0.48
CXCL11	-0.23 (-0.80,0.33)	0.21	0.48
IL5	0.16 (-0.23,0.56)	0.21	0.48
MMP-10	-0.09 (-0.30,0.13)	0.21	0.48
CCL19	-0.07 (-0.23,0.10)	0.22	0.48
SIRT2	-0.2 (-0.73,0.33)	0.23	0.48
VEGFA	-0.04 (-0.15,0.07)	0.23	0.48
AXIN1	-0.16 (-0.60,0.29)	0.24	0.48
CD6	-0.05 (-0.20,0.10)	0.25	0.48
IL-17A	-0.05 (-0.19,0.10)	0.25	0.48
TNFB	-0.05 (-0.18,0.09)	0.25	0.48
CCL11	-0.05 (-0.21,0.11)	0.26	0.48
FGF-23	0.06 (-0.12,0.24)	0.26	0.48
Inc Ctrl 2	0.01 (-0.03,0.05)	0.26	0.48
TNF	-0.05 (-0.20,0.11)	0.27	0.48
TNFRSF9	-0.04 (-0.18,0.10)	0.27	0.48
CXCL10	-0.07 (-0.31,0.17)	0.28	0.48

IL-17C	-0.07 (-0.33,0.18)	0.28	0.48
IL-10RB	-0.03 (-0.13,0.07)	0.29	0.48
MCP-4	-0.12 (-0.55,0.31)	0.29	0.48
TWEAK	-0.03 (-0.13,0.08)	0.3	0.48
CCL23	0.04 (-0.11,0.19)	0.31	0.49
HGF	-0.03 (-0.16,0.10)	0.32	0.49
Det Ctrl	-0.01 (-0.03,0.02)	0.33	0.49
IL10	-0.04 (-0.21,0.13)	0.33	0.49
IL-18R1	-0.02 (-0.11,0.07)	0.33	0.49
CCL25	0.04 (-0.16,0.24)	0.34	0.49
IL33	0.02 (-0.07,0.10)	0.35	0.49
SLAMF1	-0.03 (-0.17,0.11)	0.36	0.49
TRAIL	-0.02 (-0.14,0.10)	0.36	0.49
LIF	0.04 (-0.20,0.28)	0.37	0.49
SCF	0.02 (-0.11,0.15)	0.37	0.49
CCL20	-0.05 (-0.36,0.26)	0.38	0.49
OSM	-0.05 (-0.35,0.25)	0.38	0.49
ARTN	-0.02 (-0.18,0.14)	0.39	0.49
MCP-2	-0.03 (-0.23,0.18)	0.39	0.49
CST5	-0.02 (-0.15,0.12)	0.4	0.49
CXCL6	-0.04 (-0.36,0.27)	0.4	0.49
Flt3L	0.02 (-0.13,0.17)	0.41	0.49
IL8	0.03 (-0.24,0.29)	0.42	0.49
IL-1 alpha	0.01 (-0.08,0.10)	0.43	0.49
ST1A1	-0.03 (-0.46,0.39)	0.43	0.49
4E-BP1	-0.03 (-0.44,0.39)	0.44	0.49
CCL28	0.01 (-0.10,0.11)	0.44	0.49
IL-15RA	0.01 (-0.08,0.10)	0.44	0.49
TGF-alpha	-0.01 (-0.11,0.10)	0.44	0.49
IL-22 RA1	-0.01 (-0.21,0.18)	0.45	0.49
NRTN	-0.02 (-0.29,0.25)	0.45	0.49
CX3CL1	-0.01 (-0.15,0.14)	0.46	0.49
CXCL9	0.01 (-0.24,0.26)	0.46	0.49
CDCP1	0.01 (-0.16,0.18)	0.47	0.49
FGF-21	0.02 (-0.39,0.43)	0.47	0.49
GDNF	0.003 (-0.13,0.13)	0.48	0.49
IL6	0.01 (-0.26,0.28)	0.48	0.49
LIF-R	-0.002 (-0.09,0.09)	0.48	0.49
IL2	0.001 (-0.08,0.09)	0.49	0.5
MMP-1	-0.0001 (-0.30,0.30)	0.5	0.5

[a] Total patient sample is 45 versus 48 due to missed clinic visit at either time point for fasting blood draw. We detected 75 immunologic proteins measured by proximity extension assay (PEA) using the Olink 96 inflammation panel

[b] Generalized linear mixed models with random intercept, natural log transformation applied.

[c] P-value by Wald test

[d] Multiple correction using False Discovery Rate Benjamini–Hochberg adjustment