




# Baseline Gut Metagenomic Functional Gene Signature Associated with Variable Weight Loss Responses following a Healthy Lifestyle Intervention in Humans

Christian Diener,<sup>a</sup> Shizhen Qin,<sup>a</sup> Yong Zhou,<sup>a</sup> Sushmita Patwardhan,<sup>a</sup> Li Tang,<sup>a</sup> Jennifer C. Lovejoy,<sup>a,b</sup> Andrew T. Magis,<sup>a</sup> Nathan D. Price,<sup>a,c,d</sup> Leroy Hood,<sup>a,c</sup>  Sean M. Gibbons<sup>a,c,e</sup>

<sup>a</sup>Institute for Systems Biology, Seattle, Washington, USA

<sup>b</sup>Lifestyle Medicine Institute, Redlands, California, USA

<sup>c</sup>Department of Bioengineering, University of Washington, Seattle, Washington, USA

<sup>d</sup>Onegevity (a division of Thorne HealthTech), New York, New York, USA

<sup>e</sup>eScience Institute, University of Washington, Seattle, Washington, USA

**ABSTRACT** Recent human feeding studies have shown how the baseline taxonomic composition of the gut microbiome can determine responses to weight loss interventions. However, the functional determinants underlying this phenomenon remain unclear. We report a weight loss response analysis on a cohort of 105 individuals selected from a larger population enrolled in a commercial wellness program, which included healthy lifestyle coaching. Each individual in the cohort had baseline blood metabolomics, blood proteomics, clinical labs, dietary questionnaires, stool 16S rRNA gene sequencing data, and follow-up data on weight change. We generated additional targeted proteomics data on obesity-associated proteins in blood before and after intervention, along with baseline stool metagenomic data, for a subset of 25 individuals who showed the most extreme weight change phenotypes. We built regression models to identify baseline blood, stool, and dietary features associated with weight loss, independent of age, sex, and baseline body mass index (BMI). Many features were independently associated with baseline BMI, but few were independently associated with weight loss. Baseline diet was not associated with weight loss, and only one blood analyte was associated with changes in weight. However, 31 baseline stool metagenomic functional features, including complex polysaccharide and protein degradation genes, stress-response genes, respiration-related genes, and cell wall synthesis genes, along with gut bacterial replication rates, were associated with weight loss responses after controlling for age, sex, and baseline BMI. Together, these results provide a set of compelling hypotheses for how commensal gut microbiota influence weight loss outcomes in humans.

**IMPORTANCE** Recent human feeding studies have shown how the baseline taxonomic composition of the gut microbiome can determine responses to dietary interventions, but the exact functional determinants underlying this phenomenon remain unclear. In this study, we set out to better understand interactions between baseline BMI, metabolic health, diet, gut microbiome functional profiles, and subsequent weight changes in a human cohort that underwent a healthy lifestyle intervention. Overall, our results suggest that the microbiota may influence host weight loss responses through variable bacterial growth rates, dietary energy harvest efficiency, and immunomodulation.

**KEYWORDS** amylase, diet, health, metabolome, metagenome, microbiome, obesity, proteome, replication rate, weight loss


**Citation** Diener C, Qin S, Zhou Y, Patwardhan S, Tang L, Lovejoy JC, Magis AT, Price ND, Hood L, Gibbons SM. 2021. Baseline gut metagenomic functional gene signature associated with variable weight loss responses following a healthy lifestyle intervention in humans. *mSystems* 6:e00964-21. <https://doi.org/10.1128/mSystems.00964-21>.

**Editor** Danilo Ercolini, University of Naples Federico II

**Ad Hoc Peer Reviewer** Henrik Munch Roager, Technical University of Denmark

**Copyright** © 2021 Diener et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Christian Diener, [christian.diener@isbscience.org](mailto:christian.diener@isbscience.org), or Sean M. Gibbons, [sgibbons@isbscience.org](mailto:sgibbons@isbscience.org).

 Baseline gut microbiome functional signature associated with weight loss response following a healthy lifestyle intervention, independent of initial BMI or metabolic health state.

**Received** 23 July 2021

**Accepted** 23 August 2021

**Published** 14 September 2021

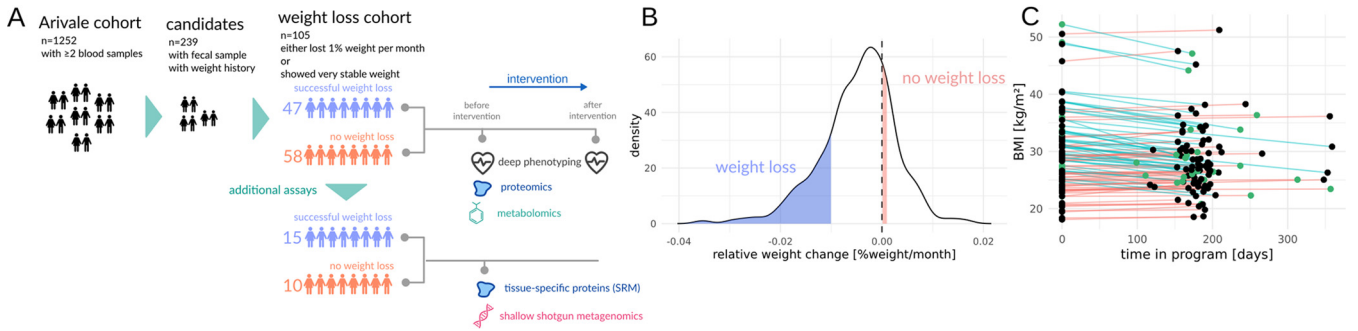
The question of whether or not the ecological composition of the gut microbiome may causally affect weight loss remains somewhat controversial (1). The composition and diversity of the gut microbiome have been correlated with body mass index (BMI) and metabolic health markers and have been shown to modulate weight gain in mice (2). There are many confounding variables with regard to obesity phenotypes, including genetics, prior health status, age, physical activity, and diet, which can modulate whether or not a person who is nominally “overweight” or “obese” is considered metabolically healthy (3–6). Recent work by our group has demonstrated that certain host-microbe metabolic associations are disrupted only in individuals experiencing severe obesity (BMI  $\geq$  35), but not in overweight or mildly obese individuals ( $25 < \text{BMI} < 35$ ), relative to individuals with a normal weight (BMI  $< 25$ ) (7). Thus, the cutoffs used to define obesity may not always match the underlying metabolism- and microbiome-associated phenotypic heterogeneity in the population.

Whether or not a consistent association between the microbiota and obesity phenotypes exists, another important, but unresolved, question is whether or not the human gut microbiome contributes directly to changes in weight after an intervention, independent of baseline BMI. While the gut microbiome has been shown to contribute to weight gain in mice (8, 9), it remains unclear whether similar factors might contribute to weight loss (10) and whether or not these results would translate well to humans (11, 12). Recent feeding studies have shed some light on this issue, demonstrating that humans with higher *Prevotella*-to-*Bacteroides* ratios tend to lose significantly more weight on a high-fiber diet, particularly individuals with low salivary amylase levels (13, 14). Similarly, high cecal *Prevotella* levels have also been shown to improve glucose homeostasis in animal models (15). In humans, it has been shown that a combination of baseline multi-omics features and microbiome data can predict postprandial glycemic responses to various foods (16, 17). Thus, while the exact mechanisms are unknown, the baseline taxonomic composition of the human gut microbiota appears to influence host responses to interventions. However, it is unclear whether or not associations between the gut microbiome and weight loss phenotypes are independent of associations between baseline BMI and the microbiome. For example, individuals with higher baseline BMIs tend to show larger-magnitude drops in BMI upon follow-up, which has been referred to as a “regression-to-the-mean” effect (18). In this study, we set out to understand the possible interactions between baseline BMI, dietary patterns, metabolic health, and gut microbiome profiles and how these factors may be associated with changes in weight and metabolic health following personalized, healthy lifestyle interventions.

## RESULTS

In this study, we leveraged existing data and biobanked samples from the Arivale cohort (see Materials and Methods). Briefly, participants enrolled in a commercial behavioral coaching program run by the former scientific wellness company Arivale, Inc., were paired with a registered dietitian or registered nurse coach. Personalized, telephonic behavioral coaching was provided to each participant on a monthly basis, with email or text communications between coaching calls. This service included longitudinal “deep phenotyping,” which involved collecting blood and stool for baseline single nucleotide polymorphism (SNP) genotyping or whole-genome sequencing (blood) and longitudinal clinical labs (blood), metabolomics (blood), proteomics (blood), and 16S amplicon sequencing of the gut microbiome (stool), along with lifestyle questionnaires, body weight measurements, and additional activity-tracking data from wearable devices. Arivale participants undergoing these personalized interventions showed broad improvements across a number of validated health markers, including an average reduction in BMI (19, 20).

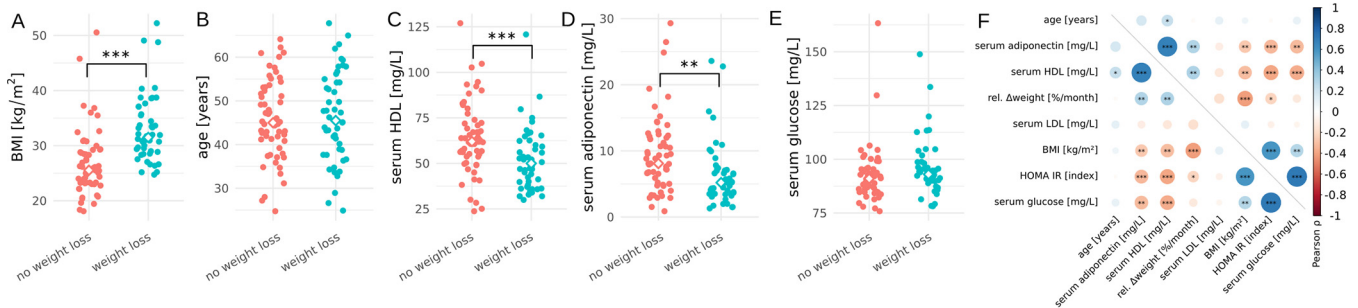
We targeted a subset of the ~5,000 Arivale participants to look specifically at weight loss phenotypes during this lifestyle intervention period (Fig. 1A). Briefly, there were 1,252 individuals with blood collected at two time points over the course of a



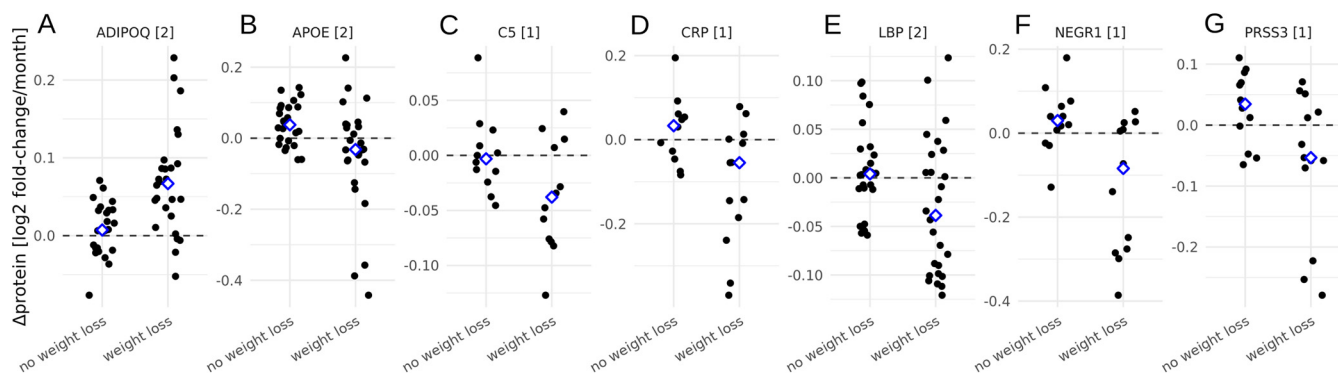
**FIG 1** Study design and cohorts. Schematic showing number of individuals within the Arivale wellness intervention cohort who match our selection criteria for data completeness and “weight loss” and “no weight loss” groups (A). In panels A to C, individuals in the “weight loss” group lost >1% of body weight per month during the program whereas the “no weight loss” group maintained a stable weight, changing less than 0.1%. Distribution of relative weight change for the 239 candidate individuals, with blue area showing individuals who lost >1% of their body weight per month ( $n = 48$ ) and red area showing individuals who showed no change in weight ( $n = 57$ ) over the same intervention period (B). Baseline and follow-up BMI values (points from the same individual connected by lines colored by weight loss group: cyan lines denote “weight loss” and red lines denote “no weight loss”) for our “weight loss” and “no weight loss” cohorts ( $n = 105$ ) (C). Green dots in panel C denote individuals with additional proteomic and metagenomic data ( $n = 25$ ).

year, 239 of whom had a paired stool sample at baseline and longitudinal data on BMI (Fig. 1A to C). We further subdivided these 239 participants by selecting individuals who lost >1% of their body weight per month over a 6- to 12-month period ( $n = 48$ ) and those who maintained a stable BMI ( $n = 57$ ) over the same period (Fig. 1B). From this 105-person cohort, another subset of 25 individuals (15 “weight loss” and 10 with “no weight loss”) were subselected for additional assays to evaluate whether weight loss responses were associated with (i) concomitant improvements in protein markers of metabolic health and (ii) baseline metagenomic functional gene and taxonomic profiles (Fig. 1A and C). Biobanked fecal samples from this 25-person cohort were used to generate shallow shotgun metagenomes (>2 million reads per sample), in order to obtain gut microbiome functional and taxonomic profiles. Two biobanked plasma samples (taken before and after intervention) taken from each of these 25 individuals were used to generate additional proteomic data on a broad set of obesity and cardiometabolic health markers (see Table S1 in the supplemental material).

In the full cohort of 105 individuals there were no significant differences in age and glucose levels between the “weight loss” and “no weight loss” groups at baseline, but the “weight loss” group had a significantly higher baseline BMI, lower baseline serum high-density lipoprotein (HDL) levels, and slightly lower baseline serum adiponectin levels (Fig. 2A to E). All individuals in the “weight loss” group were considered either overweight or obese (BMI > 25 and 30, respectively), while half of the “no weight loss” group were overweight and the other half were considered normal weight (BMI > 25 and < 25, respectively; Fig. 2A). Across the cohort baseline BMI was significantly



**FIG 2** Baseline metabolic markers in the full cohort ( $n = 105$ ). Dot plots showing baseline BMI (A), age (B), baseline serum HDL (C), baseline serum adiponectin (D), and baseline serum glucose (E) in the “weight loss” and “no weight loss” groups. In panels A to E, asterisks denote significance under a Welch  $t$  test and diamonds denote group medians. Correlation matrix showing Pearson’s correlation coefficients between baseline BMI, weight loss, and clinical markers of metabolic health across the entire cohort (F). Asterisks in panel F denote significance under a Pearson moment-product correlation test. In panels A to F, \*\*\*,  $P < 0.001$ ; \*\*,  $P < 0.01$ ; \*,  $P < 0.05$ .



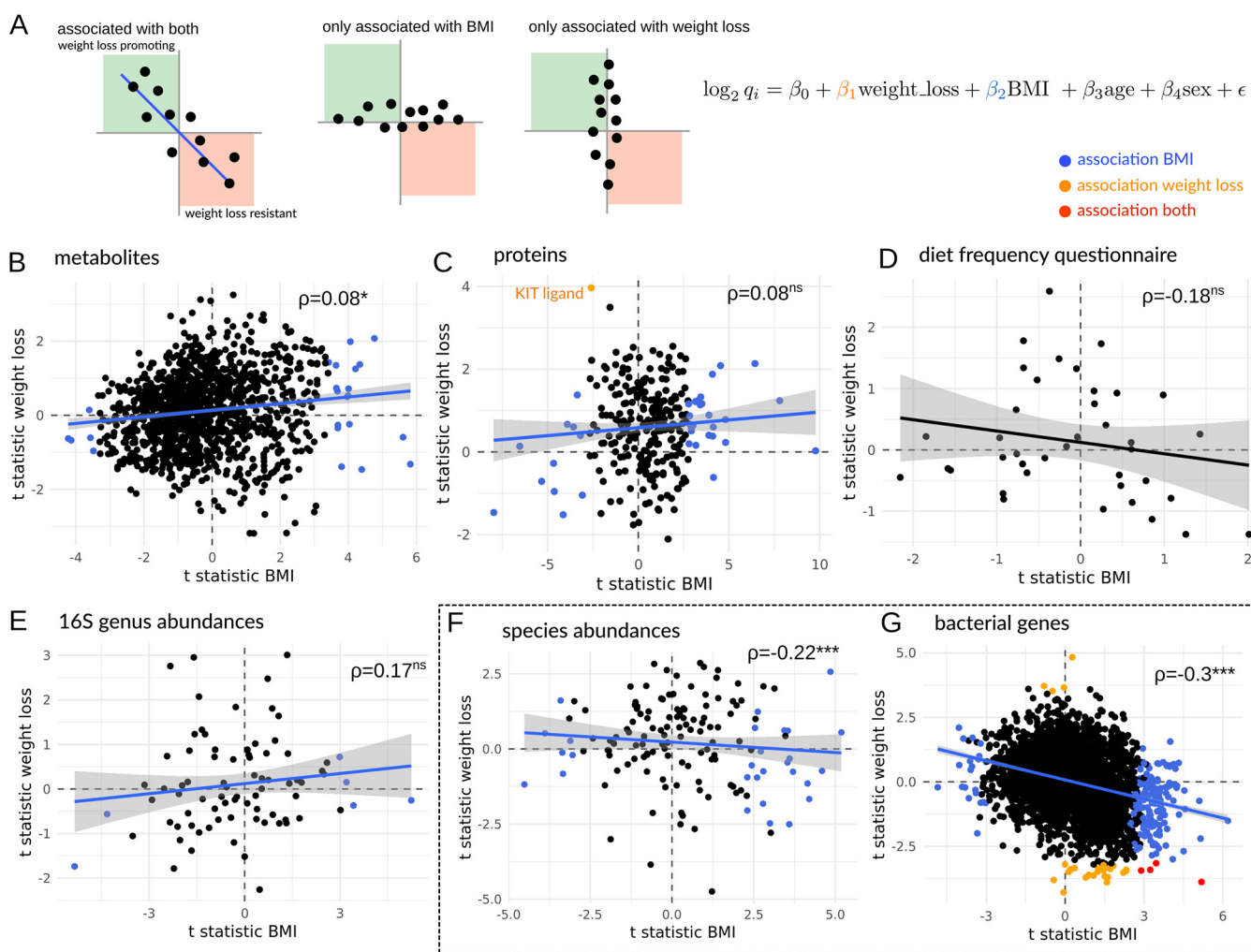
**FIG 3** Additional proteomic markers of metabolic health in the subcohort of 25 individuals. Each plot shows blood proteins that changed significantly in abundance (FDR-corrected  $P < 0.1$  when corrected for baseline BMI) between baseline and follow-up sampling in the “weight loss” group, independent of baseline BMI (A to G). Dashed line denotes no change in protein abundance over time, blue diamonds denote the means for the two groups, and numbers in brackets denote how many unique peptide fragments were used to quantify each protein (A to G).

correlated with several baseline metabolic health markers, such as adiponectin, HDL, insulin resistance, and glucose levels (Fig. 2F). We also saw the expected correlation between weight loss and baseline BMI, where those with higher baseline BMIs tended to lose more weight (Fig. 2F), which is often termed the “regression-to-the-mean” effect in weight loss studies (21). Because baseline BMI could potentially mask independent measures associated with weight loss via this regression-to-the-mean effect, we decided to correct all weight loss associations for baseline BMI.

To evaluate whether metabolic health improved in the weight loss group independently of baseline BMI, we used a panel of 22 serum protein markers associated with obesity and metabolic health, measured before and after intervention in our 25-person subcohort. On average, only individuals in the “weight loss” group showed broad improvements in seven blood proteomic markers of metabolic health following the intervention (false-discovery rate [FDR]-corrected analysis of variance [ANOVA]  $P < 0.1$ , Fig. 3A to G). Specifically, the “weight loss” group showed a marked increase in ADIPOQ (adiponectin) levels, which have previously been negatively associated with BMI and positively associated with fasting (22). The “weight loss” group also showed decreased levels of APOE, C5, CRP, LBP, NEGR1, and PRSS3, which have all been positively associated with obesity, inflammation, and metabolic disorders (Fig. 3B to G) (22–26). Thus, not only did the “weight loss” group reduce their BMI during the intervention period, but they became metabolically and immunologically healthier as well.

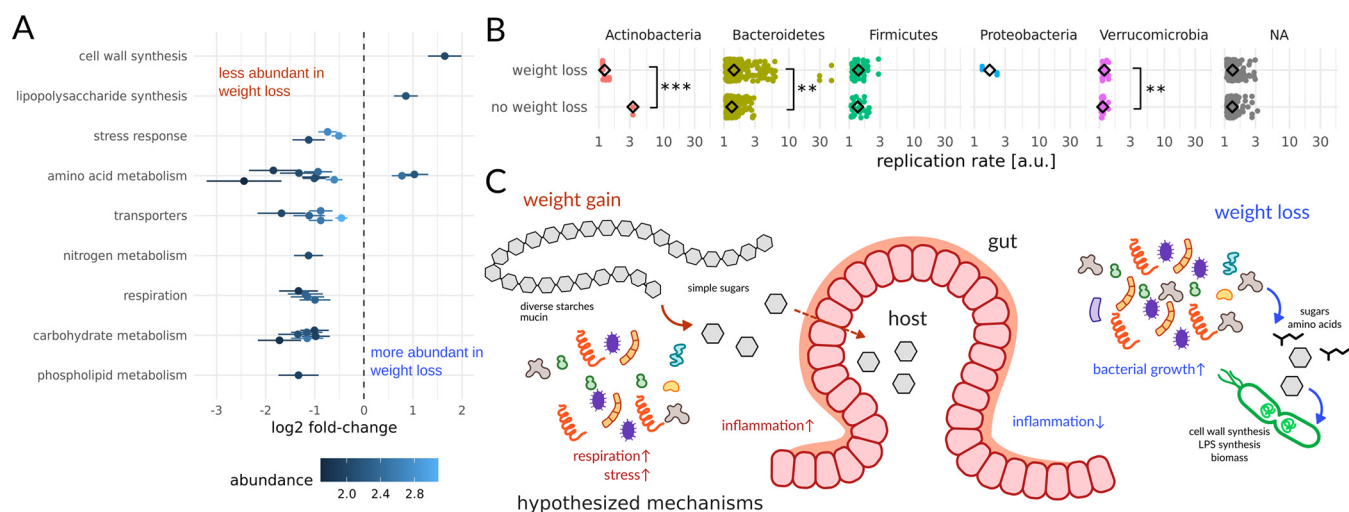
We tested for associations between baseline features and weight loss that were independent of baseline BMI, age, and sex (Fig. 4A). Although one might expect baseline phenotypic and dietary factors associated with baseline BMI to have similar associations with changes in BMI (Fig. 4A), we found that these associations were largely independent for blood metabolomics, blood proteomics, 16S genus-level abundances, and dietary patterns in the 105-person cohort (all Pearson’s coefficients  $> 0$  or nonsignificant, Fig. 4B to E) and only weakly correlated for metagenome-derived gut microbial species abundances (Pearson  $\rho = -0.22$ , Fig. 4F) and functional gene abundances (Pearson  $\rho = -0.3$ , Fig. 4G) in the 25-person subcohort. Thus, phenotypic associations with BMI and weight loss were largely orthogonal (Fig. 4B to G). None of the food frequency measures collected from this cohort were significantly associated with BMI or weight loss (Fig. 4D). Sixty baseline blood- and stool-derived features were independently associated with baseline BMI in the 105-person cohort, including known markers of weight loss and weight gain such as leptin and insulin-like growth factor (Fig. 4B to E) (27, 28). There were no baseline blood metabolites significantly associated with weight change, independent of baseline BMI (Fig. 4B). Only a single protein (KIT ligand) out of 268 baseline proteins tested was independently associated with weight loss resistance (Fig. 4C). The KIT ligand has been reported previously to be associated with obesity and energy expenditure in mice and humans (18, 29, 30). While 6 baseline





**FIG 4** Associations between baseline multi-omic features, BMI, and weight loss. Biplots show  $t$  statistic for features' independent associations with BMI or weight loss, controlling for age and sex (A). Analyses were run separately for baseline blood metabolites (B), baseline blood proteins (C), baseline dietary features (D), baseline 16S gut bacterial genera (E), baseline metagenome-derived gut bacterial species (F), and baseline metagenome-derived gut bacterial functional genes (G). Blue dots denote features significantly associated with BMI only (i.e., independent of weight loss, age, and sex), orange dots denote features significantly associated with weight loss only (independent of BMI, age, and sex), and red dots denote features independently associated with both BMI and weight loss (independent of age and sex). In panels B to G, asterisks denote significance under a Pearson correlation test and  $\rho$  denotes the Pearson correlation coefficient between the  $t$  statistics for BMI and weight loss (\*\*\*,  $P < 0.001$ ; \*\*,  $P < 0.01$ ; \*,  $P < 0.05$ ; ns,  $P > 0.05$ ). The dashed box around panels F and G denotes metagenomic results from the subcohort of 25 individuals, while the results in panels B to E are from the larger cohort of 105 individuals.

bacterial genera (16S) were associated with baseline BMI in the 105-person cohort, none were independently associated with weight loss (Fig. 4E), and this was consistent with metagenomic species-level results from the 25-person cohort (Fig. 4F). In concordance with a previous study (31), we observed that individuals with higher baseline BMI showed slightly lower metagenomic gene richness, where an increase in 1 BMI unit was associated with a loss of approximately 19 genes ( $P = 0.02$ , ANOVA corrected for baseline sex and age). Baseline gene richness was not predictive of future weight loss success, when adjusting for baseline BMI (ANOVA  $P = 0.93$ ). However, several of the 2,975 gut bacterial gene clusters included in this analysis showed independent associations with either BMI (177, FDR-corrected  $P < 0.05$ ; see Table S3) or weight loss (27, FDR-corrected  $P < 0.05$ ; see Table S3) in the 25-person cohort, and a few showed independent associations with both baseline BMI and weight loss (4, FDR-corrected  $P < 0.05$ , Fig. 3G and Table S3). Prior work in a larger cohort of several hundred individuals showed how blood analytes could predict glycemic responders during a clinical weight loss program (5). However, that study did not look at baseline gut microbiomes.



**FIG 5** Metagenomic markers of weight loss in a subcohort of 25 individuals. Metagenomic gene cluster abundances significantly associated with weight loss (independent of baseline BMI, age, and sex), binned into high-level functional categories (A). Average phylum-specific bacterial replication rates estimated from metagenomes show significant differences across weight loss groups (B). In panel B, “NA” denotes contigs without a phylum-level classification (i.e., not enough single-copy phylogenetic marker genes within those contigs to obtain a phylum-level classification) and asterisks denote significance under ANOVAs while correcting for age and baseline BMI (\*\*\*,  $P < 0.001$ ; \*\*,  $P < 0.01$ ). Schematic of the proposed microbiome-mediated mechanisms involved in weight loss promotion or resistance based on specific metagenomic functions from panel A that were positively or negatively associated with weight loss (C).

Here, we find that the baseline stool metagenomic functional genes show a much larger number of associations with weight loss phenotypes than baseline gut taxonomic, blood proteomic, blood metabolomic, or dietary features (Fig. 4B to G).

In total, 31 baseline gut microbiome functional genes were associated with weight loss, independent of baseline BMI (Fig. 5A). Cell wall and lipopolysaccharide (LPS) synthesis were positively associated with weight loss, which suggested that cell division, biomass production, and Gram-negative bacterial growth potential might be important. To explore this further, we calculated baseline bacterial replication rates directly from metagenome-assembled contigs (32) and found that average replication rates were indeed significantly higher in the “weight loss” group (ANOVA  $P = 0.001$ , corrected for age and baseline BMI), with Gram-negative *Bacteroidetes* contigs contributing most to this effect (ANOVA  $P = 0.002$ , Fig. 5B). Most contigs could not be annotated beyond the phylum level, but the fastest-replicating contigs (replication rates  $> 3$ ) with genus-level annotations belonged to *Prevotella* and were observed only in the weight loss group. Most functional genes were associated with resistance to weight loss: specifically, functions involved in glycan (e.g., glycosyl hydrolases) and protein catabolism, response to stress, peptide antibiotic synthesis, and respiration (Fig. 5A).

## DISCUSSION

Based on our rather preliminary results from this modest-size cohort, we propose a tentative set of hypotheses for how human gut commensals modulate the host’s absorption of calories from the diet and potentially impact intestinal inflammation (Fig. 5C). Specifically, we know that the gut microbiota help break down complex, extracellular polysaccharides into simpler sugars that are more readily absorbed by the host. Indeed, we saw that certain CAZy enzyme classes (e.g., GH13, which includes the starch-degrading amylases [see Fig. S3 in the supplemental material]) were enriched in individuals who were resistant to weight loss, independent of baseline BMI. A similar metagenomic increase in bacterial amylase gene frequency has been associated with increased weight gain in mice (33). Furthermore, gut bacterial replication rates were reduced in those who were resistant to weight loss, independent of baseline BMI. Similarly, prior cross-sectional work identified associations between the gut bacterial replication rates from a few taxa and BMI (34). We hypothesize that lower commensal

growth rates may allow the host epithelium to absorb a larger fraction of extracellular polysaccharide breakdown products in the lumen before they can be transformed into less-energy-dense fermentation by-products, like short-chain fatty acids (SCFAs), and bacterial biomass. SCFA production itself can reduce intestinal inflammation (35), which in turn may help to improve metabolic health and better facilitate weight loss (36). Concordantly, we saw reduced levels of circulating inflammation-related proteins in participants who lost weight (Fig. 3A to G). Finally, reduced inflammation could itself promote fermentative metabolism and redox homeostasis in the gut, minimizing oxidative stress to strict anaerobes and suppressing respiratory pathways that favor facultative anaerobes (Fig. 3F).

We suggest that dietary energy harvest, host-microbe substrate competition, and modulation of host inflammation by commensal bacteria may be, in part, responsible for determining host responses to weight loss interventions, independent of baseline BMI or metabolic health state. Gut ecosystems optimized for fermentative metabolism and higher bacterial growth rates appear to be conducive to weight loss. Prior work has shown that the higher baseline levels of *Prevotella* can improve weight loss responses to a standardized high-fiber diet (13), and here we found higher baseline *Bacteroidetes* growth rates, driven in part by the genus *Prevotella*, in individuals who lost weight in a commercial wellness program, which often involved suggested increases in dietary fiber and exercise (see Materials and Methods). Recent studies have suggested that one can predict weight loss outcomes following an intervention from baseline 16S rRNA gene community profiles (21, 37). However, these weight loss studies did not correct for baseline BMI, which can act as a significant confounder (e.g., Fig. 2F) due to the regression-to-the-mean effect described above (21, 37). The putative microbiome-centric weight loss mechanisms identified in this study are largely consistent with prior work in nonhuman animal models and in human observational studies, indicating that energy harvest, abundances of glycosyl hydrolase genes, and inflammation are relevant to weight gain and obesity (8, 33, 38, 39).

In summary, our results represent a preliminary set of baseline gut microbiome functional features that are associated with future changes in weight following an intervention, independent of baseline BMI (Fig. 5C). It remains to be seen whether these results will replicate in other cohorts and whether the funneling of dietary starches and fibers into simple sugars accessible to the host, rather than toward conversion into SCFAs, does indeed lead to weight loss resistance. In particular, the current study only looked at baseline dietary patterns and did not track detailed dietary records throughout the full duration of this personalized intervention study. Future studies should capture this longitudinal dietary data in order to better delineate between the influence of dietary variation and baseline gut microbiomes in predicting weight loss responses. Finally, larger interventional trials are needed in humans to further advance our understanding of how our commensal gut microbiota and our lifestyle interact to causally contribute to weight loss. By combining these emerging insights with recently developed models for predicting personalized gut microbiome metabolic outputs (40, 41), we can begin to engineer the functional capacity of our microbiota to optimize the outcomes of dietary and lifestyle interventions.

## MATERIALS AND METHODS

**Arivale cohort and subcohort selection criteria.** Procedures for this study were run under the Western Institutional Review Board (WIRB) with Institutional Review Board (IRB) study number 20170658 at the Institute for Systems Biology and 1178906 at Arivale. The research was performed entirely using deidentified and aggregated data of individuals who had signed a research authorization allowing the use of their anonymized data in research. Per current U.S. regulations for use of deidentified data, informed consent was not required. To be eligible to join the program, participants had to be over 18 years of age, not pregnant, and a resident of any U.S. state except New York. The participants analyzed in this study are the 92% of participants who agreed to research use as of 19 June 2018 and enrolled in the program between July 2015 and March 2018.

Of the ~5,000 Arivale participants who agreed to research use of their data, 1,252 had blood draws at two time points (i.e., a baseline sample and then a follow-up sample at ~6 to 12 months). Of these 1,252 individuals, 239 had follow-up BMI data within the year after the first blood draw and had

biobanked serum and fecal samples available which were sampled within 30 days of each other. We removed individuals with zero variance in weight measurements, which results from digital scales when regular weighing is not performed and the prior weight is reported repeatedly. Relative weight change was calculated as (follow up weight – baseline weight)/months between measurements. The study cohort was then assembled by selecting all individuals who either lost more than 1% of body weight per month ( $n = 48$ , “weight loss” group) or retained a very stable body weight (gained less than 0.1% of their body weight,  $n = 57$ , “no weight loss” group) during the lifestyle intervention. A subcohort of 25 of these 105 individuals was selected for additional proteomic and metagenomic assays. Specifically, 15 individuals with the largest declines in weight were used as the “weight loss” group, whereas 10 individuals with the 20 smallest positive weight change values were chosen as the “no weight loss” group (10-person subset was selected to ensure a balanced representation of sexes across groups).

**Arivale behavioral intervention.** Participants who enrolled in the yearlong commercial behavioral coaching program were paired with a registered dietitian or registered nurse coach. Personalized, telephonic behavioral coaching was provided to each participant on a monthly basis, with email or text communications between coaching calls. Each participant’s clinical and genetic data were available to them via a web dashboard and mobile app, which they could also use to communicate with their coach and schedule calls or blood draws. Coaches provided specific recommendations to address out-of-range clinical results based on clinical practice guidelines, published scientific evidence, or professional society guidelines. Examples of the evidence behind the coaching recommendations include guidelines from the American Heart Association or American Diabetes Association (42), comprehensive lifestyle interventions such as those developed for the Diabetes Prevention Program (DPP) (43), nutrition recommendations such as those based on the DASH dietary pattern (44) or MIND diet (45), and exercise recommendations from the American College of Sports Medicine (46).

**Blood collection and multi-omic data generation.** Blood draws for all assays were performed by trained phlebotomists at LabCorp or Quest service centers and were scheduled every 6 months, but actual collection times varied. Metabolon conducted their Global Metabolomics high-performance liquid chromatography (HPLC)–mass spectrometry (MS) assays on participant plasma samples. Sample handling, quality control, and data extraction along with biochemical identification, data curation, quantification, and data normalizations have been previously described (47). For analysis, the raw metabolomics data were median scaled within each batch, such that the median value for each metabolite was 1. To adjust for possible batch effects, further normalization across batches was performed by dividing the median-scaled value of each metabolite by the corresponding average value for the same metabolite in quality control samples of the same batch. Missing values for metabolites were imputed to be the minimum observed value for that metabolite. Values for each metabolite were log transformed. Plasma protein levels were measured using three ProSeek proximity extension assay (PEA) panels (cardiovascular II, cardiovascular III, and inflammation arrays) from Olink Biosciences (Uppsala, Sweden), processed, and batch corrected as described previously (47). For analysis, a threshold of less than 5% missing values was set for each protein, which was passed by 263 different analytes. Missing values for the proteins were imputed to be the minimum observed value for that protein.

**Dietary food frequency questionnaires.** Upon sign-up to the Arivale program, individuals filled out extensive questionnaires online. Consumption frequencies for a set of 39 different food entities were presented on an ordinal scale ranging from 0 (no consumption) up to 8 (very frequent consumption). The interpretation of individual consumption levels for each food group can be found in Table S2 in the supplemental material.

**Stool collection and metagenomic data generation.** At-home stool collection kits (DNA Genotek; OMR-200) were shipped directly to participants and then shipped back to DNA Genotek for processing. Microbial DNA was isolated from 200  $\mu$ l of homogenized fecal material using the DNeasy PowerSoil Pro extraction kit (Qiagen, Germany) with bead beating in Qiagen Powerbead Pro plates (catalog no. 19311; Qiagen, Germany). Extracted DNA was quantified using the Quant-iT PicoGreen double-stranded DNA (dsDNA) assay kit (Invitrogen, USA), and all samples passed the quality threshold of 1 ng/ $\mu$ l (range, 8 to 101 ng/ $\mu$ l).

16S amplicon sequencing was performed as described previously in reference 7. In brief, the 16S V3-V4 region was amplified and sequenced with 300-bp paired-end libraries on an Illumina MiSeq. Samples were demultiplexed using Illumina Basespace (San Diego, CA), yielding the FASTQ files used in this study.

Shallow shotgun sequencing was performed with the BoosterShot service (Corebiome, USA). In brief, single-stranded 100-bp libraries were prepared using an optimized proprietary protocol of the provider (Corebiome, USA) based on the Nextera library prep kit (Illumina, USA) and sequenced on a NovaSeq (Illumina, USA) to a minimum of 2.6 million (2.6M) reads per sample (mean 3.5M, ranging from 2.6M to 4M). Demultiplexing was performed on Basespace (Illumina, USA), yielding the final FASTQ files.

**Anthropometric data.** Height, weight, and waist circumference either were measured at the blood draws (45%) or were self-reported via an online assessment, or through the Fitbit Aria scale. Reference ranges for anthropometric data were defined by U.S. public health guidelines (48).

**SRM of obesity-related proteins.** Serum samples were processed following a previously published protocol that ensured maximum yield of signal (49). We targeted a curated selection of 22 mostly organ-specific proteins with known genetic variants associated with obesity or metabolic syndrome (Table S1). Prepared samples, along with spiked-in heavy-isotope-labeled synthetic standard peptides, were quantified using a triple-quadrupole mass spectrometer (Agilent 6490; Agilent, Santa Clara, CA) with a nano-spray ion source and Chip Cube nano-HPLC. Three to four transitions were monitored for each target peptide (see Table S1). Two micrograms of tryptically digested Mars-14 (Agilent, Santa Clara, CA)



depleted serum was eluted from a high-capacity nano-HPLC chip (160 nl, 150 mm by 75  $\mu$ m inside diameter [i.d.]; Agilent, Santa Clara, CA) with a 30-min gradient of 3 to 40% acetonitrile as described previously (49, 50). Raw selective reaction monitoring (SRM) mass spectrometry data were analyzed with the Skyline targeted proteomics environment (51). Each detected peptide was quantified by the light/heavy (L/H) ratio of monitored transitions, after adjusting for the volume of the original serum sample.

**16S amplicon sequencing data processing.** The samples were processed using a customized open-source pipeline previously described in reference 52. Here, individual samples were processed using DADA2 in order to yield individual amplicon sequence variants (ASVs). After merging forward and reverse ASVs, chimeras were removed using the *de novo* algorithm in DADA2, which removed about 17% of all reads as chimeric. Taxonomic names were assigned using the RDP method and using the Silva 16S reference database (version 132). Eighty-nine percent of the total reads could be mapped to at least the genus level this way. The resulting ASV abundance tables, taxonomy assignments, and sample metadata were finally merged into a single phyloseq object that was used for further analysis (53).

**Metagenomics data processing.** Trimming and filtering of the raw sequencing data were performed using FASTP v0.20.1 (54). The first five bases on the 5' end were trimmed from each read to avoid leftover PCR primers, and each read was furthermore trimmed on the 3' by the sliding window method with a minimum quality threshold of 20. Abundances of species were obtained using KRAKEN v2.0.9 and BRACKEN v2.6.0 using the default KRAKEN database (55, 56). Contigs were assembled *de novo* with MEGAHIT v1.2.9 with a cross-assembly across all samples. Open reading frames (ORFs) in the resulting contigs were then identified with PRODIGAL v2.6.3 (57). Reads from each sample were then aligned to each contig using MINIMAP2 v2.17, and gene abundances for each sample were quantified with the Expectation-Maximization algorithm from SALMON v3.1.3 (58, 59). The identified ORFs were annotated using the EGGNOG EMAPPER v2.0.1.

Replication rates were inferred using the iRep approach (32). Here, we first aligned the reads for each sample to the full assembled contigs using MINIMAP2 v2.17. Coverage profiles were extracted for all contigs larger than 5,000 bp across bins of a 100-bp width, but only contigs with a minimum length of 11,000 bp and a mean coverage of  $2\times$  were used for the iRep inference. Coverage profiles were smoothed using a sliding window mean over 50 bins (5,000-bp window width) before calculating the replication rates using the iRep implementation in mbtools v0.44.14 (<https://gibbons-lab.github.io/mbtools>). Taxonomic classifications of individual contigs were obtained using CAT v5.1.2 with the default database of single-copy marker genes (60).

**Statistical analyses. (i) SRM data.** Raw SRM abundances from the 25-person cohort were log-transformed, which yielded data that appeared to be normally distributed (as validated by QQ plots). Change in protein abundance across the intervention was then quantified as the difference of protein abundance after intervention and the baseline abundance, yielding log ratios of postintervention versus baseline abundances. Associations with weight loss were obtained by linear regression of the obtained log ratios using the design shown in Fig. 2. Here, assignment to the “weight loss” group was the target covariate, correcting for baseline BMI, age, and sex. Due to the low sample size in the metagenomics cohort, we did not fit interaction terms between sex and weight loss groups as this would have led some coefficients to be estimated from very small cohorts ( $n < 6$ ). False-discovery rates were controlled by adjusting  $P$  values using the Benjamini-Hochberg correction.

**(ii) Metabolomic and proteomic data.** For the 105-person cohort, mass spectrometry data from untargeted metabolomics and proteomics data were log-transformed, as this yielded near-normal distributions on QQ plots. Log-abundance values were then used for linear regressions using the design formula shown in Fig. 3A. For each metabolite and protein, we also performed a regression without the “weight loss” group and using the baseline BMI as the target covariate to yield the association strength with BMI. Linear regressions were run using the LIMMA R package without Bayesian shrinkage as this is specific to gene expression data (61). False-discovery rates were controlled by adjusting  $P$  values using the Benjamini-Hochberg correction. T-values for each association coefficient were calculated as the ratio of coefficient and estimated coefficient standard deviation obtained from the Fisher matrix of the regression.

**(iii) Diet data.** Responses to food frequency questions were extracted from the 105-person cohort and covered a set of 39 food groups on an ordinal scale ranging from 0 (no consumption) up to 8 (very frequent consumption). The numeric frequency values were used in univariate regression models with the food frequency measure as dependent variable and the same independent variables used in the metabolomic and proteomic data analysis (and shown in Fig. 3A). For each food group we also performed a regression without the “weight loss” group and using the baseline BMI as the target covariate to yield the association strength with BMI. T-values for each association coefficient were again calculated as the ratio of coefficient and estimated coefficient standard deviation obtained from the Fisher matrix of the regression.

**(iv) Metagenomic and 16S data.** Like 16S genus abundances (105 samples), metagenomic species abundances (25 samples) and gene abundances (25 samples) were both obtained from sequencing count data as described above. Each data type was stored as its own phyloseq object. We analyzed both data types (taxon and gene abundances) using negative binomial regressions, which have been shown to fit metagenomic and amplicon sequencing data well (62). This again used the design shown in Fig. 2. However, this time the regressions were performed with negative binomial regression using DESeq2 and using a prior normalization (“poscounts” method in DESeq2) (63). For each microbiome feature (genus, species, or gene) we also performed a regression without the “weight loss” group and using the baseline BMI as the target covariate to yield the association strength with BMI. False-discovery rates were controlled by adjusting  $P$  values using the Benjamini-Hochberg correction within each data type.

Pseudo-T-values were calculated as the ratio of coefficient and estimated coefficient standard error obtained from DESeq2. For gene richness estimates all samples were first downsampled to 100,000 total reads with assigned gene clusters. Gene richness was then calculated as the number of observed unique KO term gene clusters in each sample. Regressions were performed with the formulation shown in Fig. 4A and using the gene richness as the response variable.

**Data availability.** Raw metagenomic sequencing data have been deposited on the NCBI Sequence Read Archive (SRA) under BioProject no. [PRJNA748449](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA748449). SRM data can be found on the GitHub repository associated with this study ([https://github.com/gibbons-lab/weight\\_loss\\_2019](https://github.com/gibbons-lab/weight_loss_2019)). The Institute for Systems Biology manages all Arivale data requests for nonprofit research purposes and will grant access to qualified researchers. Data requests should be sent to A.T.M. ([andrew.magis@isbscience.org](mailto:andrew.magis@isbscience.org)). The full workflow used to process the metagenomic data is provided as a Nextflow pipeline at [https://github.com/Gibbons-Lab/pipelines/tree/master/shallow\\_shotgun](https://github.com/Gibbons-Lab/pipelines/tree/master/shallow_shotgun). All analyses can be found in Rmarkdown notebooks, which allow the reproduction of all analyses and figures in this paper ([https://github.com/gibbons-lab/weight\\_loss\\_2019](https://github.com/gibbons-lab/weight_loss_2019)). Specialized functions, such as the specific implementation for calculating replication rates or association analyses, can be found in a dedicated R package along with documentation at <https://github.com/gibbons-lab/mbtools>.

## SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

**FIG S1**, PDF file, 1 MB.

**FIG S2**, PDF file, 0.8 MB.

**FIG S3**, PDF file, 0.05 MB.

**TABLE S1**, CSV file, 0 MB.

**TABLE S2**, CSV file, 0.01 MB.

**TABLE S3**, CSV file, 0.2 MB.

## ACKNOWLEDGMENTS

This work was supported by an Institute for Systems Biology Innovator Award (principal investigators [PIs] C.D. and S.Q.). S.M.G. and C.D. were supported by the Washington Research Foundation Distinguished Investigator Award and startup funds from the Institute for Systems Biology. The funders had no role in designing, carrying out, or interpreting the work presented in the manuscript.

Contributions: C.D., S.Q., S.M.G., and L.H. conceptualized the study; C.D. and S.P. acquired and managed samples for metagenomics; S.Q., Y.Z., and L.T. performed the SRM experiments; J.C.L., A.T.M., and N.D.P. ran and supervised the healthy lifestyle intervention and obtained the deep phenotyping data from the Arivale cohort; C.D., S.P., and S.M.G. analyzed the data; S.M.G. and C.D. wrote the original draft. All authors reviewed and edited the manuscript.

## REFERENCES

- Sze MA, Schloss PD. 2016. Looking for a signal in the noise: revisiting obesity and the microbiome. *mBio* 7:e01018-16. <https://doi.org/10.1128/mBio.01018-16>.
- Turnbaugh PJ, Hamady M, Yatsunenkov T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP, Egholm M, Henrissat B, Heath AC, Knight R, Gordon JL. 2009. A core gut microbiome in obese and lean twins. *Nature* 457:480–484. <https://doi.org/10.1038/nature07540>.
- Cirulli ET, Guo L, Leon Swisher C, Shah N, Huang L, Napier LA, Kirkness EF, Spector TD, Caskey CT, Thorens B, Venter JC, Telenti A. 2019. Profound perturbation of the metabolome in obesity is associated with health risk. *Cell Metab* 29:488–500.e2. <https://doi.org/10.1016/j.cmet.2018.09.022>.
- Vujkovic-Cvijin I, Sklar J, Jiang L, Natarajan L, Knight R, Belkaid Y. 2020. Host variables confound gut microbiota studies of human disease. *Nature* 587:448–454. <https://doi.org/10.1038/s41586-020-2881-9>.
- Valsesia A, Chakrabarti A, Hager J, Langin D, Saris WHM, Astrup A, Blaak EE, Viguerie N, Masoodi M. 2020. Integrative phenotyping of glycemic responders upon clinical weight loss using multi-omics. *Sci Rep* 10:9236. <https://doi.org/10.1038/s41598-020-65936-8>.
- Duvallet C, Gibbons SM, Gurry T, Irizarry RA, Alm EJ. 2017. Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. *Nat Commun* 8:1784. <https://doi.org/10.1038/s41467-017-01973-8>.
- Wilmanski T, Rappaport N, Earls JC, Magis AT, Manor O, Lovejoy J, Omenn GS, Hood L, Gibbons SM, Price ND. 2019. Blood metabolome predicts gut microbiome  $\alpha$ -diversity in humans. *Nat Biotechnol* 37:1217–1228. <https://doi.org/10.1038/s41587-019-0233-9>.
- Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JL. 2006. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444:1027–1031. <https://doi.org/10.1038/nature05414>.
- Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau AL, Griffin NW, Lombard V, Henrissat B, Bain JR, Muehlbauer MJ, Ilkayeva O, Semenkovich CF, Funai K, Hayashi DK, Lyle BJ, Martini MC, Ursell LK, Clemente JC, Van Treuren W, Walters WA, Knight R, Newgard CB, Heath AC, Gordon JL. 2013. Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science* 341:1241214. <https://doi.org/10.1126/science.1241214>.
- Maruvada P, Leone V, Kaplan LM, Chang EB. 2017. The human microbiome and obesity: moving beyond associations. *Cell Host Microbe* 22:589–599. <https://doi.org/10.1016/j.chom.2017.10.005>.
- Yu EW, Gao L, Stastka P, Cheney MC, Mahabamunuge J, Soto MT, Ford CB, Bryant JA, Henn MR, Hohmann EL. 2020. Fecal microbiota transplantation for the improvement of metabolism in obesity: the FMT-TRIM double-blind placebo-controlled pilot trial. *PLoS Med* 17:e1003051. <https://doi.org/10.1371/journal.pmed.1003051>.
- Mocanu V, Zhang Z, Deehan EC, Kao DH, Hotte N, Karmali S, Birch DW, Samarasinghe KK, Walter J, Madsen KL. 2021. Fecal microbial transplantation and fiber supplementation in patients with severe obesity and metabolic syndrome: a randomized double-blind, placebo-controlled phase 2 trial. *Nat Med* 27:1272–1279. <https://doi.org/10.1038/s41591-021-01399-2>.

13. Hjorth MF, Christensen L, Larsen TM, Roager HM, Krych L, Kot W, Nielsen DS, Ritz C, Astrup A. 2020. Pretreatment Prevotella-to-Bacteroides ratio and salivary amylase gene copy number as prognostic markers for dietary weight loss. *Am J Clin Nutr* 111:1079–1086. <https://doi.org/10.1093/ajcn/nqaa007>.
14. Christensen L, Vuholm S, Roager HM, Nielsen DS, Krych L, Kristensen M, Astrup A, Hjorth MF. 2019. Prevotella abundance predicts weight loss success in healthy, overweight adults consuming a whole-grain diet ad libitum: a post hoc analysis of a 6-wk randomized controlled trial. *J Nutr* 149: 2174–2181. <https://doi.org/10.1093/jn/nxz198>.
15. Péan N, Le Lay A, Brial F, Wasserscheid J, Rouch C, Vincent M, Myridakis A, Hedjazi L, Dumas M-E, Grundberg E, Lathrop M, Magnan C, Dewar K, Gauguier D. 2020. Dominant gut Prevotella copri in gastrectomised non-obese diabetic Goto-Kakizaki rats improves glucose homeostasis through enhanced FXR signalling. *Diabetologia* 63:1223–1235. <https://doi.org/10.1007/s00125-020-05122-7>.
16. Berry SE, Valdes AM, Drew DA, Asnicar F, Mazidi M, Wolf J, Capdevila J, HadjiGeorgeiou G, Davies R, Al Khatib H, Bonnett C, Ganesh S, Bakker E, Hart D, Mangino M, Merino J, Linenberg I, Wyatt P, Ordovas JM, Gardner CD, Delahanty LM, Chan AT, Segata N, Franks PW, Spector TD. 2020. Human postprandial responses to food and potential for precision nutrition. *Nat Med* 26:964–973. <https://doi.org/10.1038/s41591-020-0934-0>.
17. Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, Ben-Yacov O, Lador D, Avnit-Sagi T, Lotan-Pompan N, Suez J, Mahdi JA, Matot E, Malka G, Kosower N, Rein M, Zilberman-Schapira G, Dohnalová L, Pevsner-Fischer M, Bikovsky R, Halpern Z, Elinav E, Segal E. 2015. Personalized nutrition by prediction of glycemic responses. *Cell* 163:1079–1094. <https://doi.org/10.1016/j.cell.2015.11.001>.
18. George BJ, Beasley TM, Brown AW, Dawson J, Dimova R, Divers J, Goldsby TU, Heo M, Kaiser KA, Keith SW, Kim MY, Li P, Mehta T, Oakes JM, Skinner A, Stuart E, Allison DB. 2016. Common scientific and statistical errors in obesity research. *Obesity (Silver Spring)* 24:781–790. <https://doi.org/10.1002/oby.21449>.
19. Price ND, Magis AT, Earls JC, Glusman G, Levy R, Lausted C, McDonald DT, Kusebauch U, Moss CL, Zhou Y, Qin S, Moritz RL, Brogaard K, Omenn GS, Lovejoy JC, Hood L. 2017. A wellness study of 108 individuals using personal, dense, dynamic data clouds. *Nat Biotechnol* 35:747–756. <https://doi.org/10.1038/nbt.3870>.
20. Zubair N, Conomos MP, Hood L, Omenn GS, Price ND, Spring BJ, Magis AT, Lovejoy JC. 2019. Genetic predisposition impacts clinical changes in a lifestyle coaching program. *Sci Rep* 9:6805. <https://doi.org/10.1038/s41598-019-43058-0>.
21. Nielsen RL, Helenius M, Garcia SL, Roager HM, Aytan-Aktug D, Hansen LBS, Lind MV, Vogt JK, Dalgaard MD, Bahl MI, Jensen CB, Muktupavela R, Warinner C, Aaskov V, Gøbel R, Kristensen M, Frøkiær H, Sparholt MH, Christensen AF, Vestergaard H, Hansen T, Kristiansen K, Brix S, Petersen TN, Lauritzen L, Licht RR, Pedersen O, Gupta R. 2020. Data integration for prediction of weight loss in randomized controlled dietary trials. *Sci Rep* 10:20103. <https://doi.org/10.1038/s41598-020-76097-z>.
22. Nigro E, Scudiero O, Monaco ML, Palmieri A, Mazzarella G, Costagliola C, Bianco A, Daniele A. 2014. New insight into adiponectin role in obesity and obesity-related diseases. *Biomed Res Int* 2014:658913. <https://doi.org/10.1155/2014/658913>.
23. Long J-R, Liu P-Y, Liu Y-J, Lu Y, Xiong D-H, Elze L, Recker RR, Deng H-W. 2003. APOE and TGF-1 genes are associated with obesity phenotypes. *J Med Genet* 40:918–924. <https://doi.org/10.1136/jmg.40.12.918>.
24. Aronson D, Bartha P, Zinder O, Kerner A, Markiewicz W, Avizohar O, Brook GJ, Levy Y. 2004. Obesity is the major determinant of elevated C-reactive protein in subjects with the metabolic syndrome. *Int J Obes* 28:674–679. <https://doi.org/10.1038/sj.ijo.0802609>.
25. Bollineni RC, Fedorova M, Blüher M, Hoffmann R. 2014. Carbonylated plasma proteins as potential biomarkers of obesity induced type 2 diabetes mellitus. *J Proteome Res* 13:5081–5093. <https://doi.org/10.1021/pr500324y>.
26. Bandesh K, Bharadwaj D. 2020. Genetic variants entail type 2 diabetes as an innate immune disorder. *Biochim Biophys Acta Proteins Proteom* 1868:140458. <https://doi.org/10.1016/j.bbapap.2020.140458>.
27. Pardini E, Ferrer R, Baena-Fustegueras JA, Lecube A, Fort JM, Vargas V, Catalán R, Peinado-Onsurbe J. 2010. The relationships between IGF-1 and CRP, NO, leptin, and adiponectin during weight loss in the morbidly obese. *Obes Surg* 20:623–632. <https://doi.org/10.1007/s11695-010-0103-5>.
28. Heymsfield SB, Greenberg AS, Fujioka K, Dixon RM, Kushner R, Hunt T, Lubina JA, Patane J, Self B, Hunt P, McCamish M. 1999. Recombinant leptin for weight loss in obese and lean adults: a randomized, controlled, dose-escalation trial. *JAMA* 282:1568–1575. <https://doi.org/10.1001/jama.282.16.1568>.
29. Huang Z, Ruan H-B, Xian L, Chen W, Jiang S, Song A, Wang Q, Shi P, Gu X, Gao X. 2014. The stem cell factor/Kit signalling pathway regulates mitochondrial function and energy expenditure. *Nat Commun* 5:4282. <https://doi.org/10.1038/ncomms5282>.
30. Gutierrez DA, Muralidhar S, Feyerabend TB, Herzig S, Rodewald H. 2015. Hematopoietic Kit deficiency, rather than lack of mast cells, protects mice from obesity and insulin resistance. *Cell Metab* 21:678–691. <https://doi.org/10.1016/j.cmet.2015.04.013>.
31. Le Chatelier E, Nielsen T, Qin J, Prifti E, Hildebrand F, Falony G, Almeida M, Arumugam M, Batto J-M, Kennedy S, Leonard P, Li J, Burgdorf K, Grarp N, Jørgensen T, Brandslund I, Nielsen HB, Juncker AS, Bertalan M, Levenez F, Pons N, Rasmussen S, Sunagawa S, Tap J, Tims S, Zoetendal EG, Brunak S, Clément K, Doré J, Kleerebezem M, Kristiansen K, Renault P, Sicheritz-Ponten T, de Vos WM, Zucker J-D, Raes J, Hansen T, MetaHIT consortium, Bork P, Wang J, Ehrlich SD, Pedersen O. 2013. Richness of human gut microbiome correlates with metabolic markers. *Nature* 500:541–546. <https://doi.org/10.1038/nature12506>.
32. Brown CT, Olm MR, Thomas BC, Banfield JF. 2016. Measurement of bacterial replication rates in microbial communities. *Nat Biotechnol* 34: 1256–1263. <https://doi.org/10.1038/nbt.3704>.
33. Carmody RN, Bisanz JE, Bowen BP, Maurice CF, Lyalina S, Louie KB, Treen D, Chadaideh KS, Maini Redkal V, Bess EN, Spanogiannopoulos P, Ang QY, Bauer KC, Balon TW, Pollard KS, Northen TR, Turnbaugh PJ. 2019. Cooking shapes the structure and function of the gut microbiome. *Nat Microbiol* 4:2052–2063. <https://doi.org/10.1038/s41564-019-0569-4>.
34. Korem T, Zeevi D, Suez J, Weinberger A, Avnit-Sagi T, Pompan-Lotan M, Matot E, Jona G, Harmelin A, Cohen N, Sirota-Madi A, Thaiss CA, Pevsner-Fischer M, Sorek R, Xavier R, Elinav E, Segal E. 2015. Growth dynamics of gut microbiota in health and disease inferred from single metagenomic samples. *Science* 349:1101–1106. <https://doi.org/10.1126/science.aac4812>.
35. Kumari M, Kozyrskiy AL. 2017. Gut microbial metabolism defines host metabolism: an emerging perspective in obesity and allergic inflammation. *Obes Rev* 18:18–31. <https://doi.org/10.1111/obr.12484>.
36. Saltiel AR, Olefsky JM. 2017. Inflammatory mechanisms linking obesity and metabolic disease. *J Clin Invest* 127:1–4. <https://doi.org/10.1172/JCI92035>.
37. Jie Z, Yu X, Liu Y, Sun L, Chen P, Ding Q, Gao Y, Zhang X, Yu M, Liu Y, Zhang Y, Kristiansen K, Jia H, Brix S, Cai K. 2021. The baseline gut microbiota directs dieting-induced weight loss trajectories. *Gastroenterology* 160:2029–2042.e16. <https://doi.org/10.1053/j.gastro.2021.01.029>.
38. Jiao N, Baker SS, Nugent CA, Tsompana M, Cai L, Wang Y, Buck MJ, Genco RJ, Baker RD, Zhu R, Zhu L. 2018. Gut microbiome may contribute to insulin resistance and systemic inflammation in obese rodents: a meta-analysis. *Physiol Genomics* 50:244–254. <https://doi.org/10.1152/physiolgenomics.00114.2017>.
39. Engström G, Hedblad B, Stavenow L, Lind P, Janzon L, Lindgärde F. 2003. Inflammation-sensitive plasma proteins are associated with future weight gain. *Diabetes* 52:2097–2101. <https://doi.org/10.2337/diabetes.52.8.2097>.
40. Diener C, Gibbons SM, Resendis-Antonio O. 2020. MICOM: metagenome-scale modeling to infer metabolic interactions in the gut microbiota. *mSystems* 5:e00606-19. <https://doi.org/10.1128/mSystems.00606-19>.
41. Thiele I, Sahoo S, Heinken A, Hertel J, Heirendt L, Aurich MK, Fleming RMT. 2020. Personalized whole-body models integrate metabolism, physiology, and the gut microbiome. *Mol Syst Biol* 16:e8982. <https://doi.org/10.15252/msb.20198982>.
42. Whelton PK, Carey RM, Aronow WS. 2018. ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *American College of Cardiology, Washington, DC*.
43. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, Nathan DM, Diabetes Prevention Program Research Group. 2002. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 346:393–403. <https://doi.org/10.1056/NEJMoa012512>.
44. Appel LJ, Moore TJ, Obarzanek E, Vollmer WM, Svetkey LP, Sacks FM, Bray GA, Vogt TM, Cutler JA, Windhauser MM, Lin P-H, Karanja N, Simons-Morton D, McCullough M, Swain J, Steele P, Evans MA, Miller ER, Harsha DW. 1997. A clinical trial of the effects of dietary patterns on blood pressure. *N Engl J Med* 336:1117–1124. <https://doi.org/10.1056/NEJM199704173361601>.

45. Morris MC, Tangney CC, Wang Y, Sacks FM, Barnes LL, Bennett DA, Aggarwal NT. 2015. MIND diet slows cognitive decline with aging. *Alzheimers Dement* 11:1015–1022. <https://doi.org/10.1016/j.jalz.2015.04.011>.
46. Haskell WL, Lee I-M, Pate RR, Powell KE, Blair SN, Franklin BA, Macera CA, Heath GW, Thompson PD, Bauman A. 2007. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Med Sci Sports Exerc* 39:1423–1434. <https://doi.org/10.1249/mss.0b013e3180616b27>.
47. Manor O, Zubair N, Conomos MP, Xu X, Rohwer JE, Krafft CE, Lovejoy JC, Magis AT. 2018. A multi-omic association study of trimethylamine N-oxide. *Cell Rep* 24:935–946. <https://doi.org/10.1016/j.celrep.2018.06.096>.
48. American College of Cardiology/American Heart Association Task Force on Practice Guidelines, Obesity Expert Panel. 2013. 2014. Executive summary: guidelines (2013) for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Obesity Society published by the Obesity Society and American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Based on a systematic review from the The Obesity Expert Panel, 2013. *Obesity* 22 (Suppl 2):S5–S39. <https://doi.org/10.1002/oby.20821>.
49. Zhou Y, Qin S, Sun M, Tang L, Yan X, Kim T-K, Caballero J, Glusman G, Brunkow ME, Soloski MJ, Rebman AW, Scavarda C, Cooper D, Omenn GS, Moritz RL, Wormser GP, Price ND, Aucott JN, Hood L. 2020. Measurement of organ-specific and acute-phase blood protein levels in early Lyme disease. *J Proteome Res* 19:346–359. <https://doi.org/10.1021/acs.jproteome.9b00569>.
50. Qin S, Zhou Y, Gray L, Kusebauch U, McEvoy L, Antoine DJ, Hampson L, Park KB, Campbell D, Caballero J, Glusman G, Yan X, Kim T-K, Yuan Y, Wang K, Rowen L, Moritz RL, Omenn GS, Pirmohamed M, Hood L. 2016. Identification of organ-enriched protein biomarkers of acute liver injury by targeted quantitative proteomics of blood in acetaminophen- and carbon-tetrachloride-treated mouse models and acetaminophen overdose patients. *J Proteome Res* 15:3724–3740. <https://doi.org/10.1021/acs.jproteome.6b00547>.
51. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ. 2010. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26:966–968. <https://doi.org/10.1093/bioinformatics/btq054>.
52. Wilmanski T, Diener C, Rappaport N, Patwardhan S, Wiedrick J, Lapidus J, Earls JC, Zimmer A, Glusman G, Robinson M, Yurkovich JT, Kado DM, Cauley JA, Zmuda J, Lane NE, Magis AT, Lovejoy JC, Hood L, Gibbons SM, Orwoll ES, Price ND. 2021. Gut microbiome pattern reflects healthy ageing and predicts survival in humans. *Nat Metab* 3:274–286. <https://doi.org/10.1038/s42255-021-00348-0>.
53. McMurdie PJ, Holmes S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8: e61217. <https://doi.org/10.1371/journal.pone.0061217>.
54. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
55. Lu J, Salzberg SL. 2020. Ultrafast and accurate 16S rRNA microbial community analysis using Kraken 2. *Microbiome* 8:124. <https://doi.org/10.1186/s40168-020-00900-2>.
56. Lu J, Breitwieser FP, Thielen P, Salzberg SL. 2017. Bracken: estimating species abundance in metagenomics data. *PeerJ Comput Sci* 3:e104. <https://doi.org/10.7717/peerj-cs.104>.
57. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>.
58. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>.
59. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* 14:417–419. <https://doi.org/10.1038/nmeth.4197>.
60. von Meijenfeldt FAB, Arkhipova K, Cambuy DD, Coutinho FH, Dutilh BE. 2019. Robust taxonomic classification of uncharted microbial sequences and bins with CAT and BAT. *Genome Biol* 20:217. <https://doi.org/10.1186/s13059-019-1817-x>.
61. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43:e47. <https://doi.org/10.1093/nar/gkv007>.
62. Calgaro M, Romualdi C, Waldron L, Risso D, Vitulo N. 2020. Assessment of statistical methods from single cell, bulk RNA-seq, and metagenomics applied to microbiome data. *Genome Biol* 21:191. <https://doi.org/10.1186/s13059-020-02104-1>.
63. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550. <https://doi.org/10.1186/s13059-014-0550-8>.