


RESEARCH

Open Access



# Gut microbiome variation modulates the effects of dietary fiber on host metabolism

Sofia M. Murga-Garrido<sup>1,2†</sup>, Qilin Hong<sup>3†</sup>, Tzu-Wen L. Cross<sup>1,4†</sup>, Evan R. Hutchison<sup>1</sup>, Jessica Han<sup>5</sup>, Sydney P. Thomas<sup>5</sup>, Eugenio I. Vivas<sup>1</sup>, John Denu<sup>5</sup>, Danilo G. Ceschin<sup>6</sup>, Zheng-Zheng Tang<sup>3,5\*</sup> and Federico E. Rey<sup>1\*</sup> 

## Abstract

**Background:** There is general consensus that consumption of dietary fermentable fiber improves cardiometabolic health, in part by promoting mutualistic microbes and by increasing production of beneficial metabolites in the distal gut. However, human studies have reported variations in the observed benefits among individuals consuming the same fiber. Several factors likely contribute to this variation, including host genetic and gut microbial differences. We hypothesized that gut microbial metabolism of dietary fiber represents an important and differential factor that modulates how dietary fiber impacts the host.

**Results:** We examined genetically identical gnotobiotic mice harboring two distinct complex gut microbial communities and exposed to four isocaloric diets, each containing different fibers: (i) cellulose, (ii) inulin, (iii) pectin, (iv) a mix of 5 fermentable fibers (assorted fiber). Gut microbiome analysis showed that each transplanted community preserved a core of common taxa across diets that differentiated it from the other community, but there were variations in richness and bacterial taxa abundance within each community among the different diet treatments. Host epigenetic, transcriptional, and metabolomic analyses revealed diet-directed differences between animals colonized with the two communities, including variation in amino acids and lipid pathways that were associated with divergent health outcomes.

**Conclusion:** This study demonstrates that interindividual variation in the gut microbiome is causally linked to differential effects of dietary fiber on host metabolic phenotypes and suggests that a one-fits-all fiber supplementation approach to promote health is unlikely to elicit consistent effects across individuals. Overall, the presented results underscore the importance of microbe-diet interactions on host metabolism and suggest that gut microbes modulate dietary fiber efficacy.

## Introduction

Humans harbor diverse and dynamic microbial communities in their intestines that span the three domains of life [1, 2]. These microbes play key roles on host biology, including breakdown of complex dietary components, vitamin production, energy harvesting, immune system

maturation, and protection against pathogens [2–4]. While many microbial functions are shared among gut communities from unrelated individuals, large interpersonal differences have also been reported [1]. Factors such as genetics, environment, and lifestyles contribute to these differences [5, 6]. Identifying the consequences of this variation as it relates to host immune responses, drug effectiveness, and metabolism is key to better understand how microbes modulate human biology and for successful implementation of precision medicine and personalized nutritional strategies.

Metabolic disease represents a major health challenge worldwide, with an estimate prevalence of 20–25% of

\* Correspondence: [ztang2@wisc.edu](mailto:ztang2@wisc.edu); [ferey@wisc.edu](mailto:ferey@wisc.edu)

<sup>†</sup>Sofia M. Murga-Garrido, Qilin Hong, and Tzu-Wen L. Cross are co-first authors; order determined by relative overall contributions.

<sup>3</sup>Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, 600 Highland Avenue, Madison, WI 53792, USA

<sup>1</sup>Department of Bacteriology, University of Wisconsin-Madison, 1550 Linden Dr., Madison, WI 53706, USA

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

the world's adult population [7, 8]. A large number of studies indicate that the gut microbiota influences the development of metabolic syndrome [9–13]. Gut microbes exacerbate metabolic disease in part by activating inflammatory pathways, and by producing compounds from diet that dysregulate host signaling and metabolism [14–17]. Microbes can also play protective roles against metabolic disease. A large body of evidence suggests that microbes and microbial metabolites derived from dietary fiber, including short-chain fatty acids (SCFAs), mediate some of the beneficial effects associated with dietary fiber consumption [18–20].

Dietary fiber are edible carbohydrate polymers with at least three monomeric units that are resistant to host digestive enzymes and not broken down or absorbed in the small intestine [21]. The chemical structure of a fiber determines important physicochemical properties including its solubility and viscosity. Dietary fibers can be divided into soluble and insoluble forms [22, 23]. Insoluble forms such as cellulose have a fecal bulking effect, and resist metabolization by gut microbes, particularly in monogastric hosts. Dietary fibers that microbes can use for carbon and energy are also referred as microbiota-accessible carbohydrates (MACs) [24]. MACs, such as inulin, pectin, and resistant starches, are broken down and metabolized through complex mechanisms by different gut bacteria [25]. Multiple lines of evidence suggest that dietary MACs have important effects on the ecology of the gut ecosystem [26]. MACs can support the growth of beneficial bacteria, promote intestinal barrier function, lower systemic inflammation [27], and prevent some of the detrimental effects caused by high-fat diet [28]. Microbial metabolism of MACs also promotes hepatic fatty acid metabolism at least in part via production of acetate, which serves as precursor for hepatic synthesis of fatty acids and related glycerophospholipid species [29].

While epidemiological studies support the notion that consumption of dietary fiber is generally beneficial for metabolic and cardiovascular health [30], results of interventions in humans vary widely. Several studies indicate that there is a large degree of interpersonal variation in the benefits attained among individuals receiving the same dietary fiber intervention [31–36]. In some cases, these inconsistent effects of fiber on host metabolism across subjects have been linked to differences in the gut microbiota of the consumers [32, 36] and potential microbial biomarkers for responsiveness to specific dietary interventions have been identified [34, 35, 37]. One study identified a higher *Prevotella/Bacteroides* ratio associated with improved glucose homeostasis in response to barley kernels in humans and demonstrated that *Prevotella copri* modulates glucose homeostasis in mice [35].

The studies described above suggest that interindividual differences in the gut microbiota may influence host metabolic responses to dietary fiber in humans. We sought to further examine this hypothesis using a tractable animal model and defined dietary fiber interventions. We colonized genetically identical germ-free (GF) mice with two distinct human fecal communities and fed them isocaloric diets containing different types of fiber. We found that the two transplanted communities elicited divergent metabolic epigenetic and transcriptional responses to the same dietary fiber. Furthermore, differences between mice colonized with these two communities varied depending on the type of fiber the animals consumed. Lastly, we identified candidate taxa and metabolites associated with these host phenotypes.

## Results and discussion

### Identifying fecal microbiomes with distinct metabolic potential

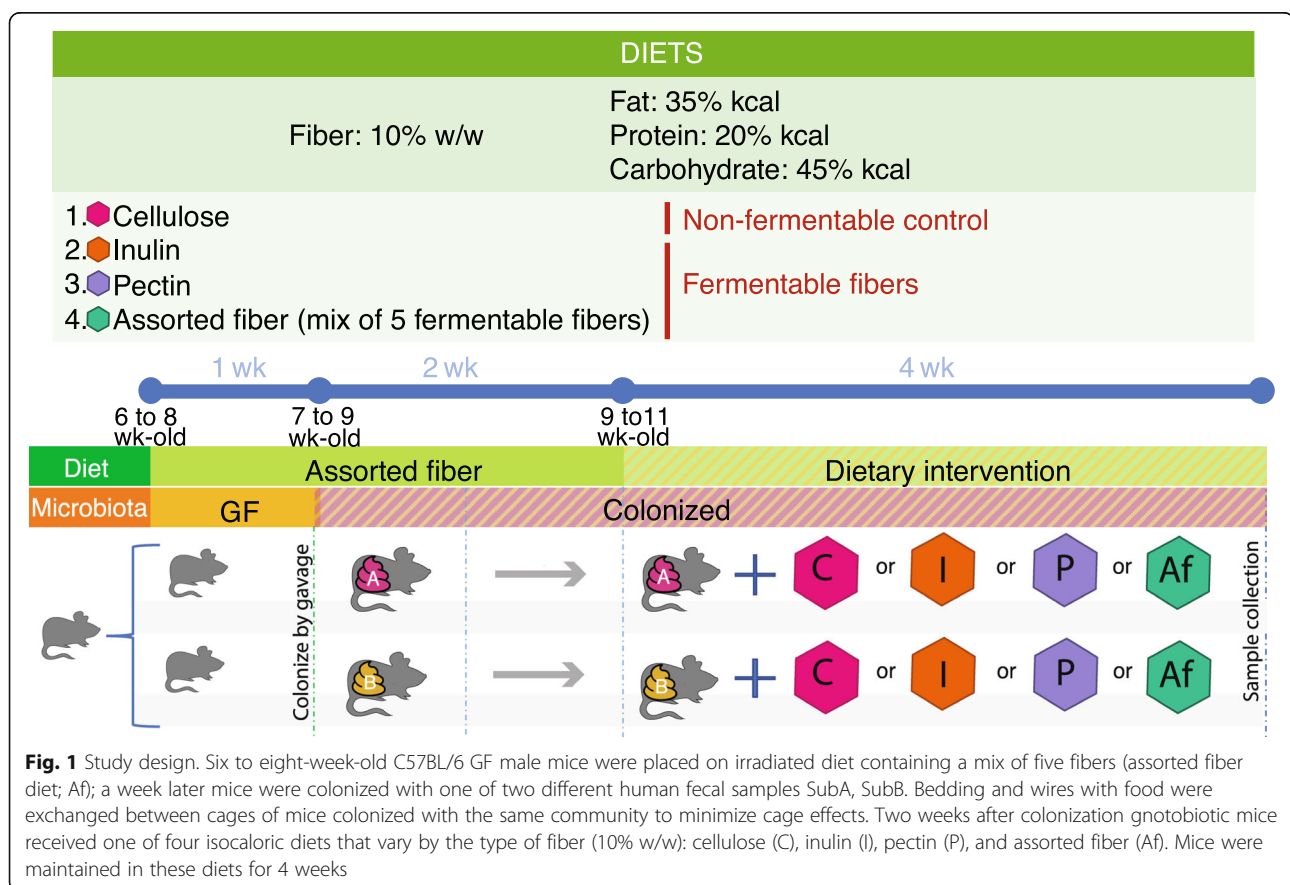
We sought to identify two human gut communities that upon engraftment in mice exhibit significantly distinct metabolic capacities. We used fecal specimens from a cohort of previously analyzed samples obtained from adults in their mid-seventies [38]. We initially selected eight fecal samples that showed significant compositional differences (Fig. S1A, B) and used them to colonize eight groups of adult male GF C57BL/6 mice ( $n = 2\text{--}4/\text{fecal sample}$ ). Mice were fed a semi-purified diet containing an assortment of diverse, commonly consumed, commercially available fibers (i.e., assorted fiber) that included resistant starch (RS) type 2 and 4, short-chain fructo-oligosaccharides (scFOS), inulin, and pectin (total fiber content 10% w/w; Table S1). This diet—as well as all the other diets used in this study, were formulated to mimic human consumption while maintaining a defined and reproducible composition. We used 10% w/w dietary fiber and 35% kcal derived from fat as it is comparable to the intake level of dietary fiber in US adults from 2001 to 2010 based on the National Health and Nutrition Examination Survey (NHANES) data [39, 40]. This was calculated by gram of fiber per 1000 kcal intake estimated for humans and adjusted to typical mouse caloric consumption. Animals were placed on this diet 1 week prior to colonization and maintained for two additional weeks after inoculation. Comparison of 16S rRNA gene analysis of cecal samples obtained from transplanted animals clustered by donor (Fig. S1C). Transplanted bacterial communities were more similar to that of their donor than to any other human sample in the dataset (Fig. S1D). Recovery of genera of at least 0.02% relative abundance in at least one of the samples associated with each human subject (i.e., donor and mouse fecal pellets) was  $59.2\% \pm 10.8$  (Fig. S1E) and accounted for  $85 \pm 14\%$  of the relative abundance of the

community from the donor (Fig. S1F). The transplanted communities also exhibited differences in alpha diversity (Fig. S1G). Furthermore, as expected from this variation, transplanted communities also differed in their capacity to produce SCFAs (Fig. S2). There was a ~ 4-fold range in the levels of cecal butyrate among the eight groups despite all animals consuming the same diet (Fig. S2A). Butyrate is known to vary widely among humans and has been linked with beneficial health effects on the host [41, 42]. Additionally, we used PICRUSt2 to predict the functional profiles of the 8 transplanted communities using 16S rRNA gene data [43, 44]. Principal coordinates analysis (PCoA) using Bray Curtis dissimilarity (Fig. S3) shows clear separation among most communities suggesting distinct functional capabilities of engrafted microbiomes. Following these analyses, we selected two markedly different communities: samples 1 and 8, from here on referred as SubA (i.e., fecal community from subject A) and SubB (i.e., fecal community from subject B) respectively, based on differences in alpha diversity, predicted metabolic properties, and capacities to produce butyrate, to examine how variation in gut community composition modulate host responses to different types of fiber. The donors of these samples are overweight (BMI = 30), have no history of type II diabetes,

cancer, or heart disease, and have self-reported consumption of a standard western-type diet.

#### Effects of dietary fiber on host metabolic outcomes is influenced by gut microbial community

Six- to eight-week-old male GF C57BL/6 mice were placed on the assorted fiber diet described above for 1 week, and subsequently colonized via oral gavage with fecal communities SubA or SubB ( $n = 30\text{--}36$  mice/community,  $n = 66$ ). Mice colonized with these two communities were maintained on the same diet for 2 weeks to allow the engrafted microbiomes to stabilize. After this stabilization period, mice colonized with each community were divided into four treatments (Fig. 1) and received one of four isocaloric diets ( $n = 7\text{--}10$  mice/community/diet) that differed on the type of fiber they contained: (i) cellulose (non-fermentable fiber), (ii) inulin, (iii) pectin, or (iv) assorted fiber described above (Fig. 1). We selected cellulose as a source fiber that is mostly not accessible to microbes in the mouse gut. This diet served as a control, to define baseline differences between mice colonized with the two communities that were independent of interactions with MACs. We chose inulin and pectin as the former is commonly used as a prebiotic in the USA, while the latter has been proven to



support growth of a wide variety of gut microbes [45], and it is commonly used as a dietary supplement. We also chose these two dietary fibers due to their distinct structures, including differences in basic units, linkages, and degree of polymerization. We colonized all mice in the same diet (assorted fiber diet) to favor consistent engraftment of mice inoculated with the same community. Furthermore, the assorted fiber diet has the same total amount of dietary fiber as the rest of the treatment groups used in this study, but with more diversity in fermentable substrates, which we reasoned would support engraftment of taxa relevant to all dietary treatments. Inclusion of this group in the experimental phase of the study also served as a control to inform whether this diet used during the colonization period drives major differences between mice colonized with the two communities used in the study.

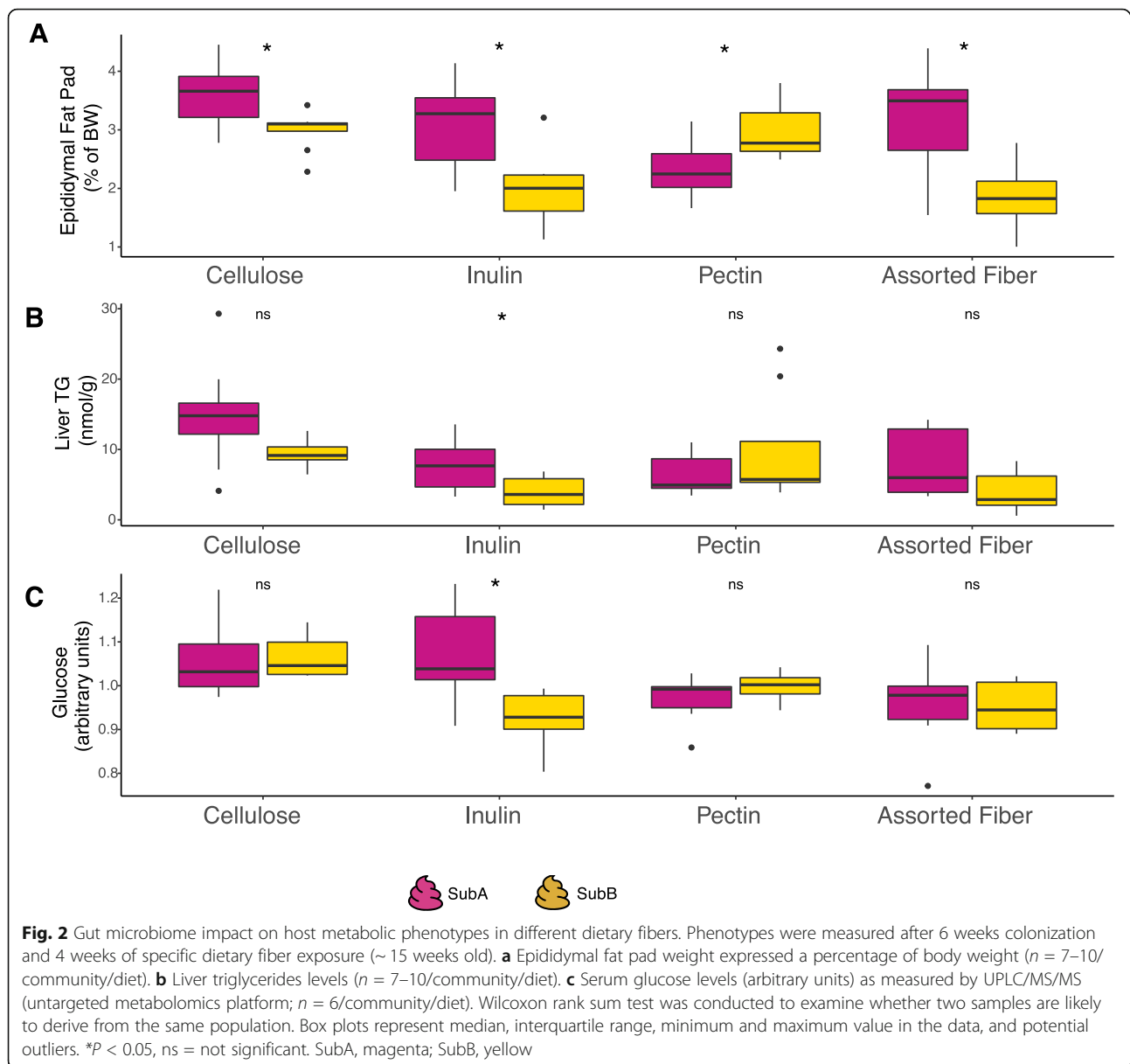
We compared metabolic phenotypes of animals colonized with the two different communities consuming each of the diets described above (Table S2). We found significant effects of diet, gut community, and their interaction on host adiposity as determined by epididymal fat pad weight (normalized by body weight) (two-way ANOVA  $P < 0.01$ ). Diet and community, but not their interaction, also showed significant effect on liver triglycerides (TG) (two-way ANOVA  $P < 0.05$ ), whereas diet and its interaction with gut community showed a significant effect on serum glucose levels (two-way ANOVA  $P < 0.05$ ; Table S2). Remarkably, while the pectin diet had an overall beneficial effect on metabolic phenotypes relative to non-fermentable cellulose (i.e., reduced adiposity and liver TG) for SubA-colonized mice, this diet was less favorable for SubB-colonized animals, which showed the strongest benefits on the inulin fiber (Fig. 2, Table S2). We also assessed whether there were significant differences in these phenotypes between mice colonized with the two different communities that were exposed to the same diet through pairwise comparisons using Wilcoxon rank sum test. In the cellulose diet, SubB-colonized mice showed lower levels of adiposity compared to mice colonized with SubA, whereas there were no statistical differences in the levels of liver TG and fasting serum glucose between these groups (Fig. 2). In the inulin diet, mice inoculated with SubB showed decreased adiposity, decreased liver TG, and lower serum levels of fasting glucose relative to animals colonized with SubA. In contrast, pectin-fed mice colonized with SubB accumulated more fat mass relative to SubA-colonized counterparts (Fig. 2a), whereas serum glucose and liver TG were comparable between the two community groups. Lastly, mice colonized with SubB showed significantly lower levels of adiposity than those colonized with SubA in the assorted fiber diet, whereas serum glucose and liver TG were comparable between

the two groups (Fig. 2). Altogether, these results underscore the importance of microbe by dietary fiber interactions on host metabolism and suggest that gut microbes modulate responses to dietary fiber.

#### Dietary fibers cause significant restructuring of SubA and SubB communities

16S rRNA gene sequences were generated from cecal samples collected from the mice described above (Tables S3 and S4). PCoA of unweighted UniFrac distances—a metric sensitive to taxonomic phylogenetic distances that does not consider abundance—of these samples show a clear clustering by donor community (Fig. 3a) and less by diet, supporting the notion that all four diets support colonization of similar assemblages for both communities. In all diets, mice colonized with the same donor preserved a core of common species that differentiated it from mice colonized with the other community (Fig. S4A). Nevertheless, there were concomitant subtle richness and pronounced abundance variations within each donor community across the different fiber treatments (Fig. S4). PCoA of weighted UniFrac distances between both communities in the four different diet treatments (Fig. 3b) shows that SubA and SubB post-intervention microbiomes (i.e., microbiome in each diet) cluster separately; however, the inulin diet appeared to be the treatment that separated the two communities the most (Fig. S5). Furthermore, the two communities shifted consistently in response to the same fibers (Fig. 3b), suggesting that related taxa from both communities are responding similarly to a given diet and that the different fibers have distinct effects on abundance of the taxa. PERMANOVA on weighted and unweighted UniFrac distances between engrafted microbiomes derived from the two communities for each dietary fiber intervention showed that these are different in all four diet comparisons ( $P < 0.05$ ). Each diet resulted in a unique set of differences between the two microbial communities and included phyla-level variations. Remarkably, the Firmicutes:Bacteroidetes (F/B) ratio was significantly increased in SubB-colonized mice consuming inulin and assorted fiber relative to SubA-colonized animals in these diets respectively, whereas the F/B ratio was higher in animals colonized with SubA consuming the cellulose diet relative to SubB counterparts, and there was no difference in this ratio between the two communities for mice consuming the pectin diet (Fig. S6A, B).

We used linear discriminant analysis effect size (LEfSe) analysis to identify bacterial taxa significantly contributing to the differences observed between post-intervention microbiomes for each diet [46] (Fig. 3c; Fig. S7). In the cellulose diet, SubA-colonized mice showed higher overall abundance of Firmicutes, Bacteroidetes, and Actinobacteria, whereas in SubB-colonized animals

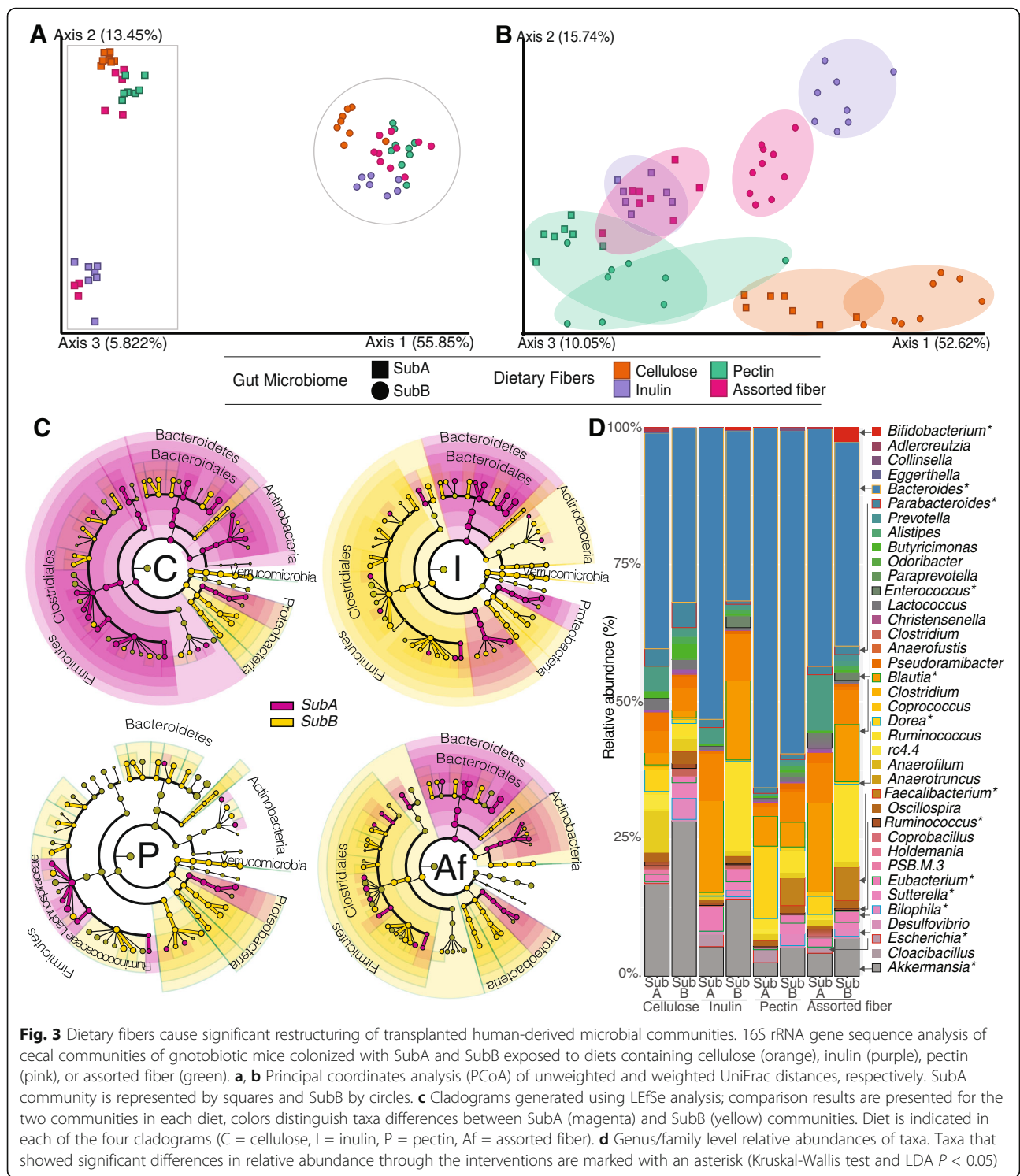


Verrucomicrobia and Proteobacteria exhibited increased abundance. Interestingly, phylum level comparisons between the two communities yielded similar differences for the inulin and assorted fiber-fed mice. Animals colonized with the SubB community showed higher levels of Firmicutes, Actinobacteria, and Proteobacteria, whereas mice colonized with SubA had higher relative abundance of Bacteroidetes in both diets. At the genus level, SubB-colonized animals showed higher relative abundance of *Bifidobacterium* in these two diets, whereas mice harboring the SubA microbiome showed increased levels of *Eubacterium*, *Bacteroides*, *Butyricimonas*, *Lactococcus*, and *Ruminococcus* (Ruminococcaceae) relative to SubB-colonized animals ( $P < 0.05$ ). In the pectin diet, mice colonized with the SubB microbiome showed higher

levels of the Proteobacteria phylum and higher levels of the *Akkermansia*, *Faecalibacterium*, *Eubacterium*, and *Clostridium* genera whereas SubA-colonized mice showed higher levels of the Lachnospiraceae family and *Dorea* genus ( $P < 0.05$ ); some of the changes observed in pectin (e.g., *Holdemania*, Lachnospiraceae, and *Eubacterium*) exhibit the opposite pattern seen in mice fed inulin (Figs. S7 and S8).

It is also important to note that some of the genera that showed diet-specific differences between communities were only detected in one of the communities; these included *Bifidobacterium*, *Clostridium*, and *Faecalibacterium*, all detected only in SubB-colonized mice. Interestingly, while the bifidogenic effect of inulin has been well documented [47], in our study the assorted fiber





**Fig. 3** Dietary fibers cause significant restructuring of transplanted human-derived microbial communities. 16S rRNA gene sequence analysis of cecal communities of gnotobiotic mice colonized with SubA and SubB exposed to diets containing cellulose (orange), inulin (purple), pectin (pink), or assorted fiber (green). **a, b** Principal coordinates analysis (PCoA) of unweighted and weighted UniFrac distances, respectively. SubA community is represented by squares and SubB by circles. **c** Cladograms generated using LEfSe analysis; comparison results are presented for the two communities in each diet, colors distinguish taxa differences between SubA (magenta) and SubB (yellow) communities. Diet is indicated in each of the four cladograms (C = cellulose, I = inulin, P = pectin, Af = assorted fiber). **d** Genus/family level relative abundances of taxa. Taxa that showed significant differences in relative abundance through the interventions are marked with an asterisk (Kruskal-Wallis test and LDA  $P < 0.05$ )

diet, which has only 2% inulin, supported significantly higher levels of this genus than the inulin diet, which contains 10% w/w inulin. However, assorted fiber diet also contains scFOS and resistant starch—both known to have bifidogenic effects [48]. It is also possible that the inclusion of additional substrates opens niche

opportunities for taxa that would otherwise compete with Bifidobacteria for inulin.

Gas chromatography-based quantification analyses revealed dietary fiber-specific differences in cecal levels of the SCFA acid butyrate and valerate, whereas levels of acetate and propionate were comparable between

communities for each diet (Fig. S9). Butyrate was significantly increased in the cecum of SubB-colonized mice consuming cellulose, pectin, and assorted fiber relative to the SubA-colonized animals consuming the same diets (Fig. S9B). Valerate was significantly higher in the gut of SubA-colonized animals in all diets except assorted fiber, whereas cecal levels of the branched-chain fatty acids (BCFA) isovalerate and isobutyrate were higher in SubA-colonized mice relative to SubB counterparts in the cellulose diet ( $P < 0.05$ ; Fig. S9E). Altogether, these results illustrate the divergent effects of dietary fibers on microbial-derived SCFA and BCFA production, which is in consistent with previous work [49, 50].

SubA and SubB were selected based on their distinct abilities to produce butyrate when transplanted to mice fed the assorted fiber diet (Figs. S2 and S9B). While previous work suggests that butyrate has a protective role against metabolic disease [40, 51], comparisons of cecal butyrate levels between mice colonized with the two communities for each diet tested did not fully explain differences in the observed metabolic phenotypes, suggesting that butyrate is not a major contributor to the phenotypic differences observed between SubA- and SubB-colonized mice (Fig. S9B, Fig. 2). However, Pearson correlation analysis across all mice in the study revealed that cecal levels of butyrate were negatively associated with adiposity ( $r = -0.4682$   $P = 0.003$ ); similar results were seen for cecal levels of acetate ( $r = -0.3595$   $P = 0.0266$ ) and total SCFA ( $r = -0.3226$   $P = 0.0483$ ). No significant associations were detected between any of the SCFAs and the other phenotypes measured. These results suggest that while butyrate and potentially acetate may influence adiposity, there are likely other microbial, community-specific, fiber type-dependent metabolites or microbial signals that modulate their effects on this phenotype.

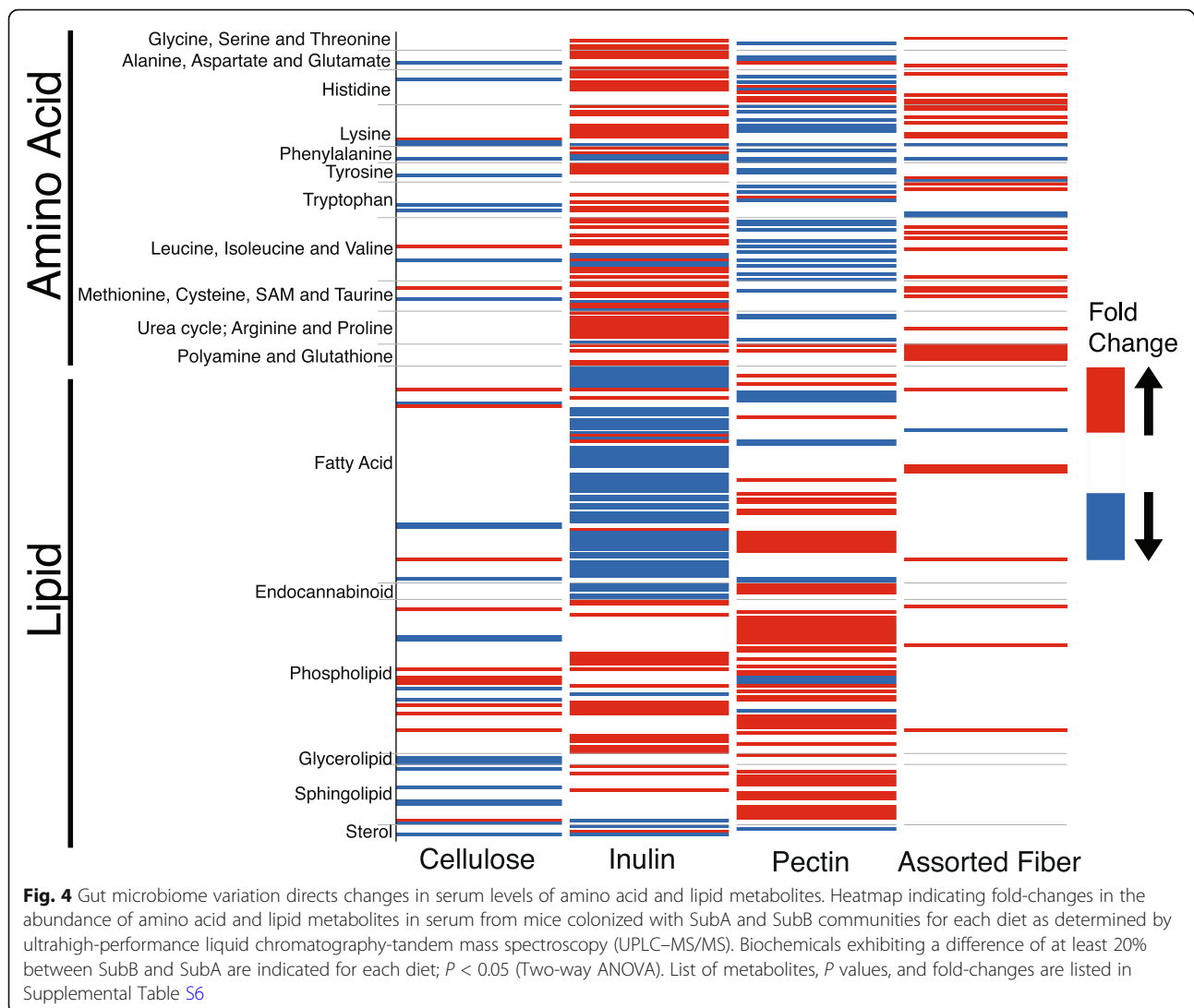
#### Post-intervention microbiomes mediate community effects on metabolic outcomes

We sought to evaluate the potential mediating role of post-intervention microbiomes in the SubA/B community effects on metabolic outcomes (adiposity, liver TG, and serum glucose) within each diet group. We focused this analysis on the cellulose, inulin, and pectin dietary fiber interventions, as these included a single type of fiber, as opposed to the assorted fiber which included a mixture of several fibers (RS type 2 and 4, scFOS, inulin, and pectin). We first performed association tests between post-intervention microbiomes (beta-diversity) and host metabolic phenotypes using PERMANOVA. In the inulin diet, post-intervention microbiomes showed differences in beta-diversity associated with changes in liver TG ( $P < 0.05$ ) and glucose phenotypes ( $P < 0.01$ ), whereas differences between the two microbiomes in

pectin-fed mice were associated with variation in the adiposity phenotype ( $P < 0.05$ ). We then assessed the mediation effect of the gut microbiome in the relationship between communities and host metabolic phenotypes. The distance-based mediation test showed significant mediation effect of overall diversity of microbiome for the glucose outcome in the inulin-fed group (unweighted UniFrac and Jaccard  $P < 0.05$ ). To identify mediator taxa, we applied a causal mediation model on each internal node of the taxonomy tree (Table S5). For the adiposity outcome, we found the Proteobacteria phylum, the Bacteroidales order, and the Christensenellaceae family as mediators in pectin-fed mice ( $P < 0.05$ ). For the serum glucose outcome, we found Clostridiales order in inulin-fed mice ( $P < 0.05$ ). Altogether, these analyses further support the notion that dietary fiber-microbe interactions modulate metabolic phenotypes and highlight potentially relevant taxa.

#### Gut microbiome-dietary fiber interactions modulate blood metabolites

Gut microbes influence host metabolism at least in part by modulating nutrient availability and through the production of a myriad of metabolites, many of which reach systemic circulation [14–17]. We sought to examine whether the gut microbial differences described above between mice colonized with SubA and SubB in the four diets resulted in changes in host metabolism as assessed by metabolomics of serum samples. We applied ultrahigh-performance liquid chromatography-tandem mass spectroscopy (UPLC–MS/MS) to quantify 774 compounds in serum from the 8 groups of mice described above (6 samples/community/diet) (Table S6). Two-way ANOVA analysis between SubB- and SubA-transplanted mice, identified 1 serum metabolite in the cellulose diet, 235 in the inulin diet, 160 in the pectin diet, and 19 in the assorted fiber diet that showed significant differences between the two communities ( $P < 0.05$ , false discovery rate (FDR) adjusted- $P < 0.1$ ). A large fraction of these metabolites also showed a significant diet by community interaction ( $P < 0.05$ ; Table S6). For the purpose of generating hypotheses, all metabolites showing  $P < 0.05$  were used for further analyses. Animals consuming inulin and pectin diets showed the largest number of significant changes in biochemicals between the two communities. Inspection of the differentially abundant biochemicals within each diet revealed diet-specific changes in metabolites related to amino acids and lipid metabolic pathways, including alanine-aspartate, histidine, lysine, tyrosine, leucine-isoleucine-valine, urea cycle: arginine and proline, methionine-cysteine-taurine, fatty acids, endocannabinoid, and sphingolipid metabolic pathways (Fig. 4, Table S6). These results support the notion that consumption of



dietary fiber elicits individual responses on blood metabolites that are influenced by differences in the gut microbiome.

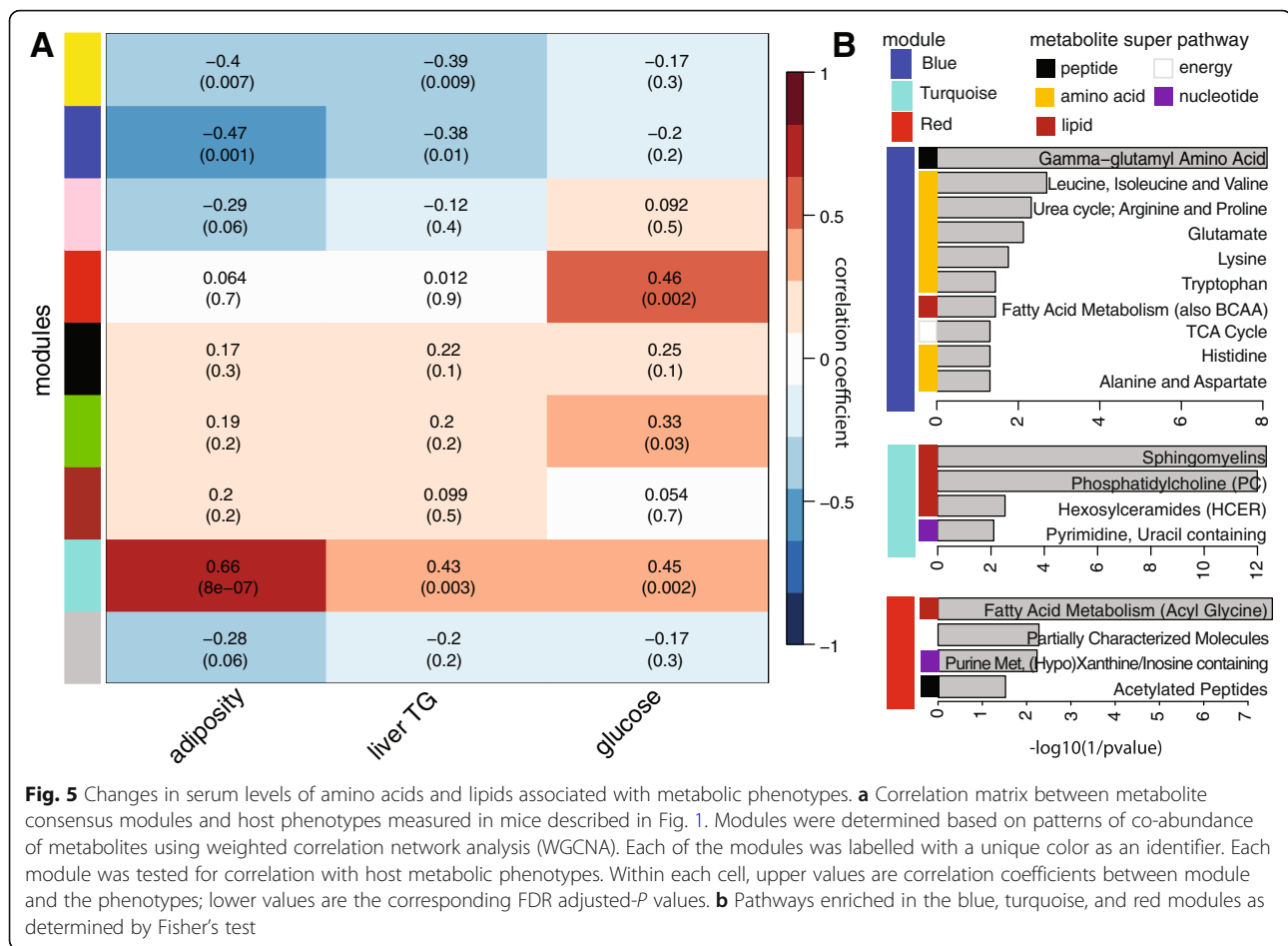
#### Amino acids and lipid metabolic pathways associate with host phenotypes

We tested whether levels of serum metabolites were associated with host metabolic outcomes. We used the dynamic tree cut method in weighted correlation network analysis (WGCNA, Fig. S10; Table S7) to define modules of tightly correlated serum metabolites. The associations between these modules and host metabolic phenotypes are shown in Fig. 5a. Candidate biochemicals that belong to these enriched pathways are listed in Table S6. The turquoise module is positively correlated with adiposity ( $r = 0.66$ ,  $P = 8e-07$ ), liver TG ( $r = 0.43$ ,  $P = 0.003$ ), and serum glucose ( $r = 0.45$ ,  $P = 0.002$ ), whereas the blue module yielded negative correlations with adiposity ( $r =$

$-0.47$   $P = 0.001$ ) and liver TG ( $r = -0.38$   $P = 0.01$ ). The red module showed highest association with the glucose phenotype ( $r = 0.46$   $P = 0.002$ ). Pathway enrichment analysis revealed significant over-representation of metabolites related to sphingomyelins, phosphatidylcholine (PC), and hexosylceramides pathways in the turquoise module whereas the blue module was enriched for metabolites in the gamma-glutamyl amino acid, branched-chain amino acid (BCAA), urea cycle, glutamate, lysine and tryptophan pathways and the red module in fatty acid metabolism (acyl glycine), purine metabolism ((Hypo)Xanthine/Inosine containing), and acetylated peptides (Fig. 5b).

Sphingomyelins (turquoise module) are strongly associated with adiposity, TG, and glucose levels. These metabolites are increased in SubB-colonized mice in the pectin diet, which exhibited increased adiposity relative to SubA-colonized mice. Previous work indicates that





sphingolipids mediate cellular processes involved in apoptosis, cell differentiation, and inflammation [52]. Furthermore, serum levels of sphingomyelins and ceramides have been associated with the development of obesity [53]. Sphingomyelins can be hydrolyzed by sphingomyelinases releasing phosphocholines and ceramides, leading to metabolic impairment [53]. Furthermore, most of the identified sphingomyelin-lipid metabolism changes are within the pectin-fed group (Tables S6 and S8). These results support the notion that microbiome-fiber interactions modulate host levels of sphingolipids and ceramides. This is consistent with recent work suggesting that *Bacteroides*-derived sphingolipids in the intestine provide an endogenous source of sphingolipids to the host [54].

As mentioned above, amino acids are enriched in the blue module and largely contribute to the negative association observed between this module and adiposity, liver TG, and glucose levels. These include histidine, which was increased in SubB-colonized mice consuming the inulin diet compared to SubA-colonized counterparts (Table S8). SubB-colonized animals in this diet exhibited lower glucose and adiposity relative to the SubA-

colonized animals (Fig. 2). This is consistent with previous work showing that histidine supplementation results in an improvement in insulin sensitivity and lower body fat [55]. Several bacterial metabolites derived from histidine were also increased in SubB-colonized mice in the inulin diet compared to SubA-colonized animals, including imidazole propionate (Table S6), which has been linked to impaired insulin signaling and type II diabetes in humans [15, 56].

BCAAs were also enriched in the blue module and detected at higher levels in SubB-colonized mice in the inulin diet compared to SubA-colonized animals. Previous reports indicated that BCAAs upregulate glucose transporters and activate insulin secretion [57, 58]. However, there is also evidence that leucine and isoleucine have a negative impact on metabolic health [59, 60], and several studies have suggested that excessive intake of amino acids could lead to inhibition of insulin signaling [61]. Additionally, the bacterial metabolite derived of tryptophan, indolepropionate [62], which we detected at significantly higher levels in mice colonized with the SubB community consuming inulin relative to SubA counterparts, has been previously associated with

increased dietary fiber intake and linked to reduced risk of low-grade inflammation [63] and improved glucose homeostasis [64]. Mice fed the cellulose diet showed virtually no differences in the metabolites enriched in this module. Gamma-glutamyl amino acids were also highly enriched in the blue module. Gamma-glutamyl dipeptides are also involved in glutathione (GSH) metabolism, which plays an important role in antioxidant defense, and are produced when gamma-glutamyl transpeptidase catalyzes the transfer of the gamma-glutamyl moiety of glutathione to amino acids. Gamma-glutamyl transpeptidase is expressed in several mammalian tissues and in bacteria [65, 66]. While the role of these metabolites on metabolic health remains poorly understood, a recent study showed gamma-glutamyl amino acids  $\gamma$ -glutamyl cysteine and  $\gamma$ -glutamyl valine inhibit TNF- $\alpha$  signaling in intestinal epithelial cells and reduce inflammation [67]. Overall, these results suggest that circulating levels of sphingolipids, amino acids, and gamma-glutamyl dipeptides are impacted by interactions between microbes and dietary fibers. Interestingly, these circulating metabolites also showed significant correlations with cecal levels of acetate and butyrate (Fig. S11). While further studies are needed to assess the role of these two SCFAs in modulating the abundance of systemic metabolites, both acetate and butyrate are known to impact many facets of metabolism via interactions with G-protein coupled receptors present in the gut and in the periphery [20].

#### Connecting bacterial taxa, serum metabolites, and host metabolic phenotypes

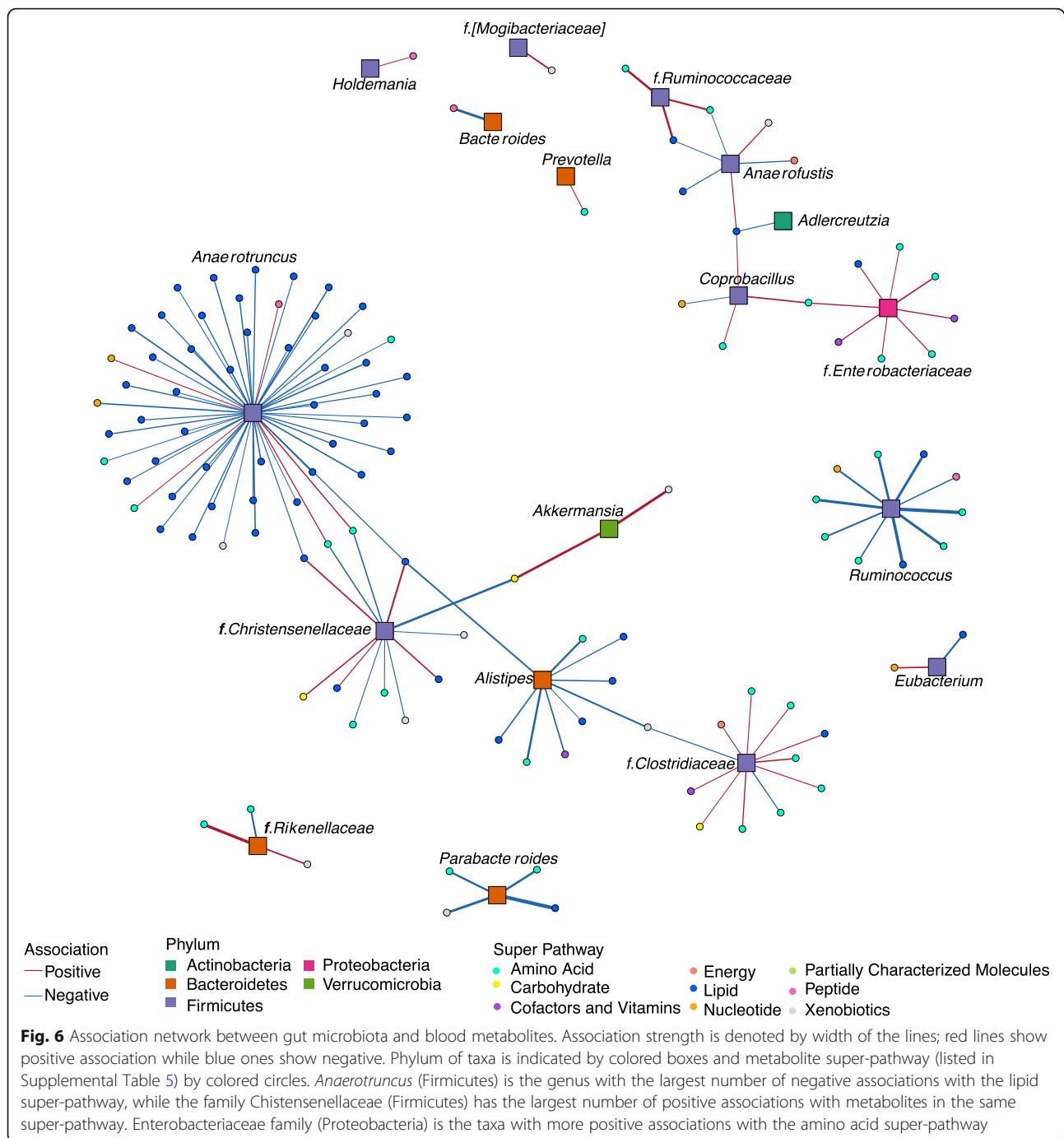
To identify bacterial taxa associated with metabolites, we applied a log-contrast model with metabolites as the response. The association analysis linking taxa and metabolite super-pathways show that the Firmicutes phylum is involved in most associations, followed by Bacteroidetes (Fig. 6). Interestingly, the genus *Anaerotruncus* is negatively associated with metabolites of the lipid super-pathway that includes fatty acids (long-chain saturated and unsaturated, and branched), lysophospholipids, and monacylglycerol. This genus is also positively associated with metabolites of the amino acid and the nucleotide super-pathways, including lysine, glycine, arginine-proline metabolism, and the purine and pyrimidine. *Ruminococcus* (Ruminococcaceae) is negatively associated with metabolites of the amino acid and the lipid super-pathways, including BCAA, glutamate and tryptophan metabolism, fatty acids, purine, and gamma-glutamyl amino acid, whereas *Parabacteroides* is negatively associated with metabolites in the arginine-proline and dihydroxy fatty acid pathways. Interestingly, *Anaerotruncus* and *Ruminococcus* (Ruminococcaceae), along with *Parabacteroides*, showed significantly higher

relative abundance in mice colonized with the SubA community in the inulin diet, which exhibit higher adiposity, liver TG, and glucose compared to mice in the same diet colonized with the SubB community ( $P < 0.01$ ; Fig. S8).

The Rikenellaceae family showed a strong positive association with tryptophan and tyrosine metabolism and a negative association with BCAA metabolism. Interestingly, Rikenellaceae and Ruminococcaceae were also positively correlated with cecal levels of BCFAs—the end products of bacterial BCAA catabolism (Fig. S12). The relative abundance of Rikenellaceae was significantly increased in SubA-colonized mice consuming cellulose, inulin, and assorted fiber and in SubB-colonized mice consuming pectin, which exhibited worse metabolic outcomes relative to the other community in their respective diets. This is consistent with previous reports describing increased abundance of Rikenellaceae in leptin-resistant obese and diabetic mice [68, 69]. Moreover, we found this family was positively associated with genes in the glycerophospholipid metabolic process, which generates lipids that can be packed in very low-density lipoproteins [70]. The *Bacteroides* genus showed a strong negative correlation with gamma-glutamyl amino acids, and it was detected at significantly higher levels in SubA-colonized animals except when animals were fed pectin ( $P < 0.05$ ). *Eubacterium* was negatively associated with plasmalogens and positively associated to purine metabolism. As discussed above, this genus was also more abundant in SubA-colonized mice consuming inulin diet relative to SubB-colonized counterparts, and in SubB-colonized mice consuming pectin diet relative to SubA-colonized animals in the same diet. In both cases, the increased levels of *Eubacterium* were associated with less favorable metabolic outcomes (Fig. 2 and Fig. S8). Altogether, these results link changes in levels of bacterial taxa differentially represented in the two communities across the different diets with alterations in systemic levels of metabolites and host metabolic phenotypes.

#### Dietary modulation of gut microbiome influences impact of gut communities on hepatic gene expression

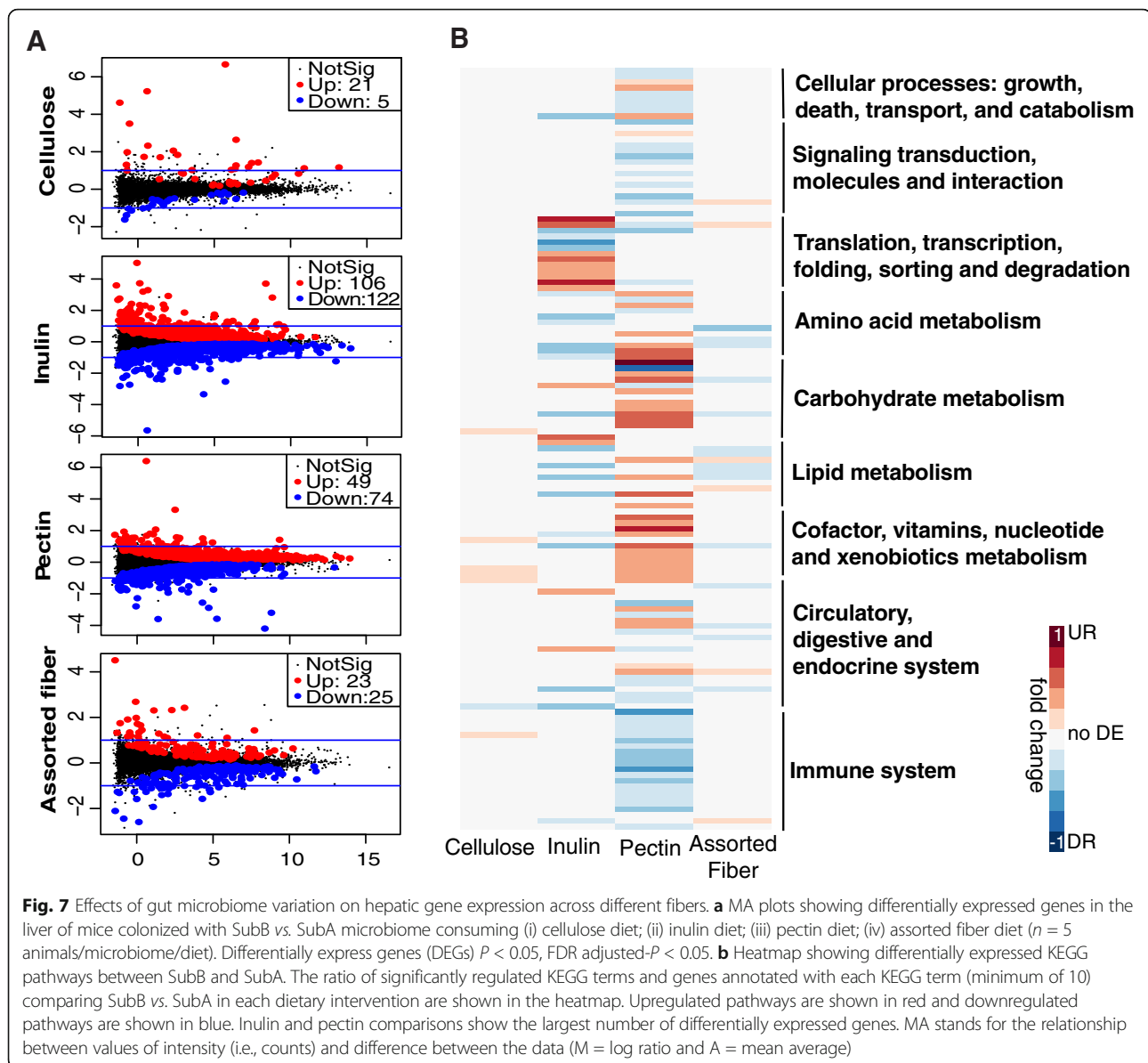
The liver receives a large fraction of its blood supply through the portal circulation, which is the direct venous outflow of the intestine. As such, the liver is continuously exposed to gut microbial-derived products, including SCFAs and bacterial toxins [71]. Differences in abundance of microbial metabolites, including SCFAs, have been linked to changes in global host epigenetic states and gene expression [72]. Epigenetic states of chromatin are reflected in the covalent post-translational modifications (PTMs) on histone proteins. We found that colonization of mice with SubA and SubB



communities affected covalent post-translational modifications (PTMs) on histone proteins (Supporting information, Figs. S13 and S14). To test directly whether these differences were linked to changes in hepatic gene expression, we performed RNA sequencing (RNA-seq) analysis. Tests of differential gene expression yielded larger number of significantly regulated genes between the two communities for the inulin and pectin dietary interventions (Fig. 7), while cellulose-fed mice showed the

least number of differences. Only 26 genes were differentially expressed between the two communities in cellulose-fed mice, compared to 228 genes in inulin, 123 in pectin, and 48 in the assorted fiber diets ( $P < 0.05$  and FDR adjusted- $P < 0.05$ ; Fig. 7a, Table S9).

Linear model fit test for over-representation of gene ontology (GO) among differentially expressed genes in biological process (BP), molecular function (MF), and cellular component (CC) categories as well as



enrichment analysis on Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways revealed divergent results between dietary interventions (Table S10). Among the transcripts expressed at lower levels in SubB relative to SubA-colonized mice in the cellulose diet, there were genes associated with regulation of phosphorylation and protein modification. SubB-colonized animals consuming inulin diet showed lower levels of expression of genes associated with amino acid metabolism, fatty acid metabolic process, peroxisome components, oxidoreductase activity, and peroxisome proliferator-activated receptor (PPAR) signaling pathway relative SubA-colonized counterparts. Upregulated genes in SubB-colonized mice were associated with ribosome

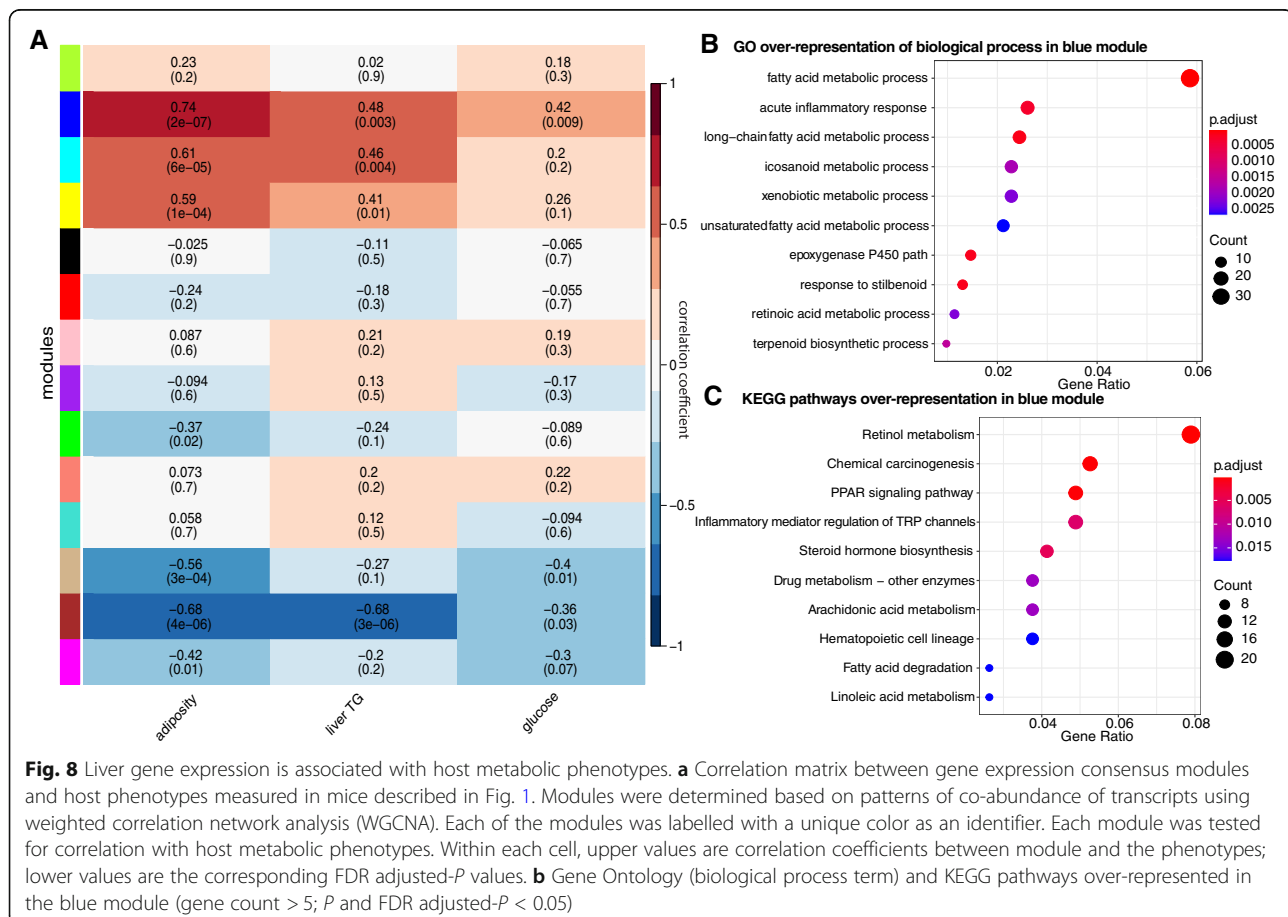
biogenesis, RNA metabolic process, and protein processing in endoplasmic reticulum. Mice consuming assorted fiber diet showed some similarities in the enrichment of differentially expressed genes as those in the inulin diet. Genes downregulated in SubB relative to SubA-colonized mice fed inulin and assorted fiber diets showed overlapping GO terms including oxidative stress, regulation of cellular ketone metabolic process, lipid metabolic process, fatty acid metabolic-catabolic process and lipid modification, fatty acid oxidation, long-chain fatty acid metabolic process, and oxidoreductase activity.

In the pectin diet, mice colonized with SubB expressed higher levels of genes involved in cofactor, vitamins, and nucleotide metabolism; fatty acid metabolism,

oxidoreductase activity, carbohydrate and valine, leucine and isoleucine metabolism (Fig. 7b, Table S10). These include six genes encoding cytochrome P450, some of which are known to play important roles in the synthesis of steroid hormones (*Cyp2a4*, *Cyp17a1*) and xenobiotic metabolism (*Cyp2a5*) [73, 74]. Surprisingly, SubA-colonized animals in the pectin diet—which show lower levels of adiposity (Fig. 2a) relative to SuB-colonized mice—exhibited higher levels of expression of genes involved in immune system process, inflammatory response, phagocytic cup and vesicle cellular component, cytokine receptor activity and binding, and infection pathways. Previous work has shown that the overall health effects of consumption of MACs such as inulin and pectin can be context dependent. Long-term consumption of pectin or inulin by Toll-like receptor 5 (TLR5) KO mice, which show innate immune deficiencies and gut microbiome alterations, results in liver inflammation and hepatocellular carcinoma [75]. Altogether, these results suggested that the gut microbiome is a differential factor that modulates hepatic gene expression. Furthermore, data suggests that the impact of the gut microbiome on liver gene expression is influenced by the type of dietary fiber consumed.

### Liver gene expression is associated with host metabolic phenotypes

WGCNA was applied to the normalized read count data obtained from RNA-Seq analysis. Fourteen gene modules, each clustering highly co-expressed genes, were identified (each module was assigned a different color; Fig. S15, Table S11). We performed correlation analysis between the phenotypic data and the calculated eigen-gene—defined as the first principal component of the expression matrix of the corresponding module—for each module identified (Fig. 8a). The blue module and adiposity showed the strongest association ( $r = 0.74$ ,  $P = 2e-07$ ). We narrowed down possible relevant genes contributing to this module by performing biological process GO and KEGG pathway analysis enrichment (Fig. 8b, c). In total, 614 out of the 677 transcripts in the blue module mapped into biological process gene ontology. These showed significant enrichment in several processes including the fatty acid metabolic process, acute inflammatory response, long-chain fatty acid, and unsaturated fatty acid metabolic process ( $P < 0.05$  and FDR adjusted- $P < 0.05$ ). Pathway enrichment analysis of 266 genes with KEGG annotation in this module yielded significant enrichment in pathways that included PPAR





signaling, inflammatory mediator regulation of transient receptor potential (TRP) channels, fatty acid degradation, and linoleic acid metabolism ( $P < 0.05$  and FDR adjusted- $P < 0.05$ ).

When only considering genes in the blue module with significant positive association to each phenotype, we observed 209, 151, and 107 genes for adiposity, liver TG, and glucose, respectively ( $P < 0.05$ ). The GO biological process analysis for this set of genes in the blue module showed that throughout the measured phenotypes, the fatty acid metabolic process was the gene ontology with more gene counts, followed by long-chain fatty acid metabolic process, xenobiotic metabolic process, and epoxygenase P450 pathway. The statistically significant ( $P < 0.05$  and FDR adjusted- $P < 0.05$ ) enriched KEGG pathways associated with the measured phenotypes were retinol metabolism, chemical carcinogenesis, PPAR signaling pathway, steroid hormone biosynthesis, linoleic acid metabolism, biosynthesis of unsaturated fatty acids, and fatty acid elongation. Nucleoside, ribonucleoside and purine nucleoside bisphosphate metabolic process, and fatty acid degradation are the GO terms and KEGG pathways that were enriched in the association with adiposity and liver TG (Fig. S16), whereas inflammatory mediator regulation of TRP channels and pyruvate metabolism pathways are enriched among the genes associated with glucose levels.

#### Gene expression in liver is associated with abundance of gut bacterial taxa

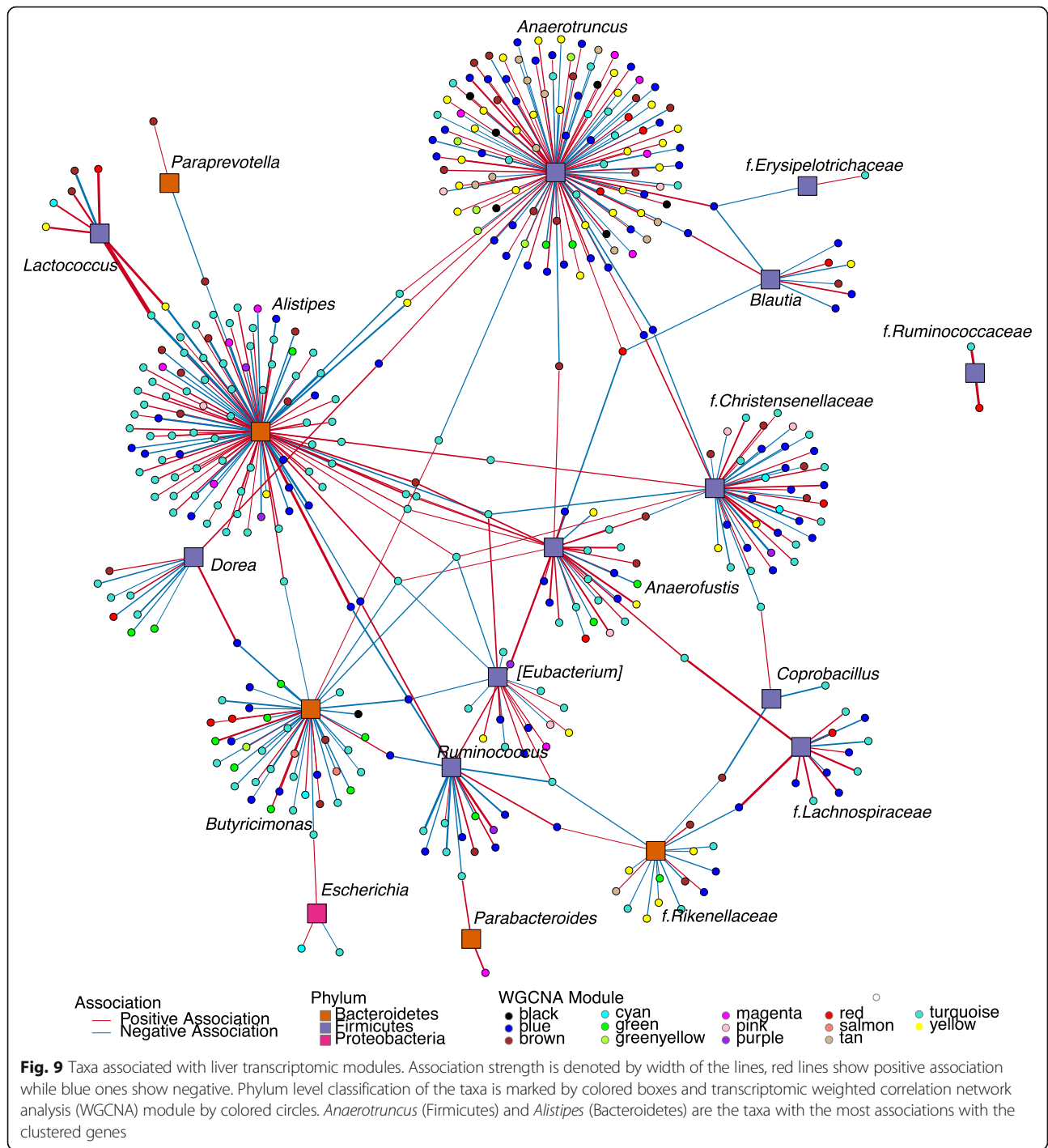
The association analysis linking taxa and liver gene expression modules showed that the Firmicutes phylum is involved in most associations, followed by Bacteroidetes (Fig. 9). Interestingly, *Anaerotruncus* cluster has the highest number of gene associations (334), followed by *Alistipes* (105), *Butyricimonas* (58), and Christensenellaceae (49) (Table S12). This association network revealed an overall positive association between *Anaerotruncus* with genes belonging to the lipid metabolic process, and a negative association with genes belonging to purine ribonucleotide and nucleoside metabolic process, cholesterol biosynthetic process, cellular lipid, fatty acid and phospholipid biosynthetic process, and response to cytokines. Interestingly, serum metabolomics data shows an overall negative association of this genus with metabolites of the lipid super-pathway (Fig. 6), suggesting that these taxa may play a role in host lipid metabolism. *Alistipes* was positively associated with genes in the lipoprotein biosynthetic process clustered in the turquoise module, whereas *Butyricimonas* was overall negatively associated with immune system processes. Previous work has found the *Anaerotruncus* and *Alistipes* genera enriched in genetically obese (*ob/ob*) mice exhibiting severe glycolipid metabolism disorders relative to wild type and *ob/ob* mice

consuming inulin with improved metabolic parameters [76, 77]. Abundance of these two genera was also linked with consumption of high-fat diets in humans [78–80]. Altogether, these studies suggest that *Anaerotruncus* and *Alistipes* may directly or indirectly contribute to the metabolism of dietary fats and fiber types and community context influence this dynamic. Furthermore, the strong correlations between intestinal microbes, hepatic gene expression, and host metabolic phenotypes provide potential connections between dietary fiber-microbes-health that warrant further examination.

#### Conclusion

Human and mouse studies encompassing genetically diverse populations have shown that host's genetic variation impacts all facets of physiology including responses to diet [81]. However, it is now clear that these populations also contain a significant amount of genetic variation derived from their largely individual associated microbiomes. Previous work in humans showed that dietary supplementation of resistant starch increases fecal butyrate levels, but with remarkable interindividual variation [50]. Furthermore, a recent study suggested that individual gut microbiota differences can be used to predict post-prandial glycemic responses to specific foods [82]. While these studies provide strong support to the notion that the gut microbiome is a major source of variability that influences responses to diet, dissecting the effects of microbial vs. host genetic variation while controlling environmental exposure is practically impossible to achieve in human studies. Modeling this variation in GF mice provides the opportunity to unravel the effects of host genetics from environmental and microbial exposures allowing the discovery of causal relationships between microbes and host phenotypes [76, 77]. Using this approach, we demonstrate that consumption of the same dietary fiber by genetically identical mice harboring different human-derived microbiomes can lead to different host metabolic outcomes (Fig. 2). Remarkably, these phenotypic responses varied as a function of the type of fiber present in the diet and were associated with changes in cecal and blood metabolites, and hepatic gene expression.

A recent study using antibiotics and PEG treatment to deplete the mouse indigenous gut microbiota, followed by transplant of human fecal microbiota from four obese donors into these animals which were fed a high-fat diet, showed that host metabolic effects in response to inulin supplementation varied as a function of the human community used to colonize the mice [83]. This work also identified several genera including *Barnesiella*, *Bilophila*, *Butyricimonas*, and several *Alistipes* amplicon sequence variants (ASVs) positively correlated with adiposity and/or hepatic steatosis whereas *Akkermansia*, *Raoultella*, and *Blautia* were negatively



correlated with at least one of these metabolic outcomes in the transplanted mice. Some of these findings are consistent with our results, e.g., *Alistipes* and *Butyrivimonas* were detected at lower levels in SubB-colonized subjects in the inulin diet relative to SubA-colonized mice in the same diet, whereas *Akkermansia* was detected at higher levels in SubB-colonized in this diet, which showed an overall healthier metabolic profile than SubA-colonized counterparts.

Interestingly, higher basal levels of *Akkermansia* were detected in the gut microbiomes of obese individuals that benefited from a dietary inulin intervention, (i.e., responders) relative to non-responders of the same treatment [83]. Furthermore, higher baseline levels of *Akkermansia* were also associated with better clinical outcomes among individuals with metabolic syndrome subjected to a calorie-restricted diet [84]. Further studies are needed to establish the causal

relationship between *Akkermansia* and enhanced response to this MAC and to identify potential molecular/microbial players involved.

The results presented here, combined with the work discussed above, suggest that the efficacy of dietary interventions such as prebiotics depends on the gut microbiota of the consumer and one-fits-all approaches to promote health are unlikely to elicit consistent effects across individuals. Identifying gut microbial biomarkers associated with beneficial responses to common interventions may help to stratify subjects into more effective personalized treatments.

While ascertaining causal mechanisms explaining differences in host metabolic phenotypes between microbiomes for each diet is beyond the goal of the current study, our results illustrate how introduction of MACs promote divergent host phenotypes caused by the two gut communities. This result has implications not only for personalized nutrition approaches but is also relevant for animal studies in the nutrition field. The use of AIN93-based purified diets [85] that contain cellulose as the sole source of fiber is common practice in nutrition studies. These diets are useful tools for the field, but their lack/low levels of dietary MACs might not support gut bacterial taxa that are relevant to the process studied. While there is not sufficient evidence to prescribe the use of a particular fiber or a specific combination of different fibers, the presented results describe some of the consequences of their inclusion or lack thereof. Furthermore, inclusion of some fermentable fibers in rodent diets will enable conditions that are more representative of human and rodent nutrition.

A major limitation of the study is that only two communities were compared. While these showed clear organismal and functional differences (Figs. S1 and S3) and elicited distinct host metabolic phenotypes, they likely capture a small fraction of the variability observed among human gut microbiomes. Expanding this study to include a wider range of communities, including samples from subjects with extreme diets, markedly different lifestyles, or different health status may expand the range of responses to dietary fiber mediated by gut microbes. Additionally, our study did not characterize the degree to which each of the fibers was metabolized by the two communities. Identifying gut signatures associated with desired health outcomes in response to specific fibers may reveal biomarkers of beneficial diet-microbiome interactions that guide personalized nutrition approaches. In conclusion, the present study underscores the importance of the gut microbiome as a differential factor that contributes to individual variation in metabolic responses to dietary fiber.

## Methods

### Dietary formulation

The four diets used during this study contained 35% kcal fat, 20% kcal protein, 45% kcal carbohydrate, and

different fiber type which equals 10% weight. A non-fermentable fiber, cellulose (Solka Floc), was used as control while inulin [Oligo-Fiber Inulin instant (100010911); Cargill, Minneapolis, MN], pectin (PE1006; Gojira Fine Chemicals, LLC, Bedford Height, OH), and a formulation of assorted fibers which contained 23.4 g/kg inulin, 21.5 g/kg short-chain fructooligosaccharide (i.e., scFOS, NUTRAFLORA<sup>®</sup>, Ingredion Inc., Westchester, IL), 33.3 g/kg resistant starch type 2 (HI-MAIZE<sup>®</sup> 260, Ingredion Inc., Westchester, IL), 23.5 g/kg resistant starch type 4 (Fibersym<sup>®</sup>, MPG, Atchison, KS), 23.5 pectin, as sources of fermentable fibers (Table S1). Experimental diets were manufactured and sterilized via irradiation by Envigo (10% cellulose diet; TD.170720, 10% inulin diet; TD.170721, 10% pectin diet; TD.170725, and assorted fiber diet; TD.170726).

### Gnotobiotic husbandry

All experiments involving gnotobiotic mice were performed under protocols approved by the University of Wisconsin-Madison Animal Care and Use Committee. All germ-free (GF) C57BL/6 mice were maintained in a controlled environment in plastic flexible film gnotobiotic isolators under a strict 12-h light/dark cycle and received sterilized water and chow (LabDiet 5021; LabDiet, St. Louis, MO) ad libitum. A week prior to colonization, mice were switched to the assorted fiber diet. GF status of mice was confirmed prior to starting the experiment using culture-dependent methods. No turbidity was observed when feces collected from the mice were inoculated in a panel of rich media and incubated at 37 °C aerobically and anaerobically for a week.

### Colonization of germ-free mice and dietary fiber interventions

#### Screening phase protocol

Richness and diversity metrics (i.e., alpha and beta) obtained from previous publication [38] plus Faith's phylogenetic diversity (calculated using QIIME 2 [86]) were used to describe 8 Wisconsin Longitudinal Study (WLS) samples selected for this study [87]. All donors were adults between 70 and 82 years old, (6 men, 2 women) with reported body mass index interval was between 23 and 38. Fecal suspensions were prepared under anaerobic conditions in Hungate tubes. A ~ 0.5-cm piece of frozen fecal aliquot straw technique (FAST) straw material were resuspended in 5 ml mega media as previously described [16, 88]. Adult 6–8-week-old male C57BL/6 GF mice (total  $n = 23$ ) were inoculated by oral gavage with ~ 200  $\mu$ l of fecal suspension ( $n = 2$ –4 mice/sample), after being fed with an assorted fiber diet (see below) for a week in sealed positive pressure individually ventilated cages (IVCs; Allentown). Each cage contained 2–4 mice. Upon inoculation, mice were maintained on

the same assorted fiber diet for 2 more weeks. Cecal contents were collected for SCFAs and 16S rRNA gene sequencing analysis from each of the 8 gnotobiotic groups.

### Study design

The two fecal communities exhibiting the highest (SubB) vs. lowest (SubA) butyrate-producing activity were then selected for subsequent experiments. Both samples belonged to individuals of the same sex (males), similar age ( $76.5 \pm 1$  years old), BMI = 30, and with no history of diabetes, cancer, or heart disease. Three-day diet recall indicates that both subjects consumed a typical western diet. Fecal samples from SubA and SubB were inoculated into 7–9-week-old GF male mice ( $n = 30$ – $36$  mice/community) consuming assorted fiber diet as described above. All gnotobiotic mice continued with the same irradiated assorted fiber diet for two more weeks (stabilization period). Bedding and wires with food were exchanged between cages of mice colonized with the same community to minimize cage effects. Mice were then switched to one of the four diets described above: cellulose, inulin, pectin, and assorted fiber ( $n = 7$ – $10$  mice/community/diet), and maintained in these diets for another 4 weeks. Animals were euthanized after 4 h of fasting.

### Measurements of short-chain fatty acids (SCFA)

#### SCFA analysis of mouse samples

Cecal levels of SCFAs were measured as previously described [19]. Briefly, a mixture of 10  $\mu$ l of internal standards (200 mM for mice and 20 mM for human each; acetic acid-D4, Sigma-Aldrich no. 233315; propionic acid-D6, Sigma-Aldrich no. 490644; and butyric acid-D7, CDN isotopes no. D-171) was subsequently added, followed by 20  $\mu$ l of 33% HCl and 1 ml diethyl ether. The vials were sealed, vortexed vigorously for 3 min, and then centrifuged (4000g, 10 min). The upper organic layer was transferred to another vial and a second diethyl ether extraction was performed. After combining the 2 ether extracts, a 60  $\mu$ l-aliquot was removed, combined with 2  $\mu$ l *N*-tert-butyltrimethylsilyl-*N*-methyltrifluoroacetamide (Sigma-Aldrich no. 394882) in a GC auto-sampler vial with a 200  $\mu$ l glass insert, and incubated for 2 h at room temperature. Derivatized samples (1  $\mu$ l) were injected onto an Agilent 7890B/5977A GC/MSD instrument with an Agilent DB1-ms 0.25 mm  $\times$  60 m column with a 0.25- $\mu$ m bonded phase. A discontinuous oven program was used starting at 40 °C for 2.25 min, then ramping at 20 °C min<sup>-1</sup> to 200 °C, then ramping at 100 °C min<sup>-1</sup> to 300 °C and holding for 7 min. The total run time was 18.25 min. Linear column flow was maintained at 1.26 ml min<sup>-1</sup>. The inlet temperature was set to 250 °C with an injection split ratio of 15:1. Acquisition B.07.02.1938. The m/z values of monitored ions in mice cecal measurements were as follows: 117 (acetic

acid), 120 (acetic acid-D4), 131 (propionic acid), 136 (propionic acid-D6), 145 (butyric acid), and 152 (butyric acid-D7). Concentrations were normalized to milligrams of cecal contents.

### Measurements of branched-chain fatty acids (BCFA)

Levels of BCFA were quantified in cecal samples collected from mice exposed to the four different diets used in the study ( $n = 66$ ;  $n = 7$ – $10$  per community/dietary fiber intervention) using the headspace GC analysis method. Frozen, weighed samples (~70 mg) were added to chilled 20-ml headspace vials (Restek, Bellefonte, PA) containing 2.0 g NaHSO<sub>4</sub>, distilled water (300  $\mu$ l—sample weight), and 1.0 ml of 60  $\mu$ M 2-butanol (internal standard; added just prior to the sample). Vials were crimp sealed immediately after sample addition and vortexed periodically to disperse and mix the contents. Headspace GC analyses were performed using a Shimadzu (Columbia, MD) HS-20 headspace sampler connected to a Shimadzu GC-2010 Plus GC equipped with a SH-Stabilwax column (30 m, 0.25 mm ID, 0.10  $\mu$ m df) linked to a FID. Samples were equilibrated with shaking to 80 °C for 20 min, pressurized to 80 kPa for 3 min prior to column injection (2 ml injection loop, load time 0.2 min, sample and transfer line temperature 150 °C, 1:15 split ratio, N<sub>2</sub> column flow 1.2 ml/min), with a column temperature program starting at 40 °C/2 min, increased to 200 °C (20 °C/min), held 2 min, decreased to 120 °C (20 °C/min), decreased to 40 °C (40 °C/min), and stabilized 1 min prior to the subsequent injection. The GC cycle time was approximately 23 min. Standard mixtures were prepared and analyzed by the same method, and peak areas determined using Shimadzu Lab Solution software (version 5.92), with adjustment for fecal sample size.

### Tissue collection and analysis

Blood was collected via cardiac puncture of anesthetized mice following a 4-h fast. Serum was obtained by centrifugation and stored at -80 °C. Cecal contents, liver, and gonadal fat pads were collected at the time of euthanasia, snap frozen in liquid nitrogen, and stored at -80 °C until analysis. Glucose was measured using ultrahigh-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS) by Metabolon.

### Liver triglyceride measurements

Liver triglycerides (TG) were quantified as previously described [19, 89]. Briefly, between 30 and 40 mg of frozen liver tissue was homogenized in 30 ml of 2:1 chloroform:methanol and disrupted using a bead beater (BioSpec Products, Barlesville, OK; maximum setting for 6 min at room temperature). Samples were incubated overnight at 4 °C with gentle agitation and 1 ml of 4 mM MgCl was added for phase separation. The organic solvent (500  $\mu$ l) was left to evaporate overnight and the dried



lipids were reconstituted in 200  $\mu$ l butanol:triton-  $\times$  114 mix (3:2 vol:vol). TG content was determined by colorimetric assay from Sigma (Sigma, F6428), according to the manufacturer's instructions and expressed in nanomole per gram of wet tissue for final concentration.

### Statistical analysis of mouse phenotypes

To assess differences on metabolic phenotypes measured between microbiota communities within each dietary intervention and between the same microbiota across different diets, we performed a nonparametric test and use permutation approach to obtain the  $P$  value for two-group comparison through Wilcoxon rank sum test. Homogeneity of variance was tested using Levene's test ( $P > 0.05$ ) previous to performing a two-way ANOVA to investigate the effect of dietary intervention, transplanted microbiota community, and their interaction on each phenotype ( $P < 0.05$ ).

### 16S rRNA gene sequencing

Genomic DNA was extracted from cecal contents using a bead-beating protocol [76]. Briefly,  $\sim$  50 mg of fecal pellet sample were resuspended in a solution containing 500  $\mu$ l of extraction buffer [200 mM Tris:HCl (pH 8.0), 200 mM NaCl, 20 mM EDTA], 210  $\mu$ l of 20% SDS, 500  $\mu$ l phenol:chloroform:isoamyl alcohol (pH 7.9, 25:24:1) (Invitrogen 15593-049), and 500  $\mu$ l of 0.1-mm diameter zirconia/silica beads. Samples were mechanically disrupted using a bead beater (BioSpec Products, Barlesville, OK; maximum setting for 3 min at room temperature), followed by centrifugation, recovery of the aqueous phase with 60  $\mu$ l 3 M NaAcetate, and precipitation with isopropanol. QIAquick 96-well PCR Purification Kit was used to remove contaminants. Isolated DNA was eluted in 10 mM Tris (pH 8.0) buffer and was stored at  $-20^\circ\text{C}$  until further use.

Amplification of 16S rRNA genes (V4) was done from DNA by PCR using unique 8-bp barcodes on the forward and reverse primers and fused with Illumina sequencing adapters [90]. Each sample was amplified in duplicate in a reaction volume of 12.5  $\mu$ l using KAPA HiFi HotStart DNA polymerase (KAPA Biosystems, Wilmington, MA, cat. # KK2602), 10  $\mu$ M of each primer, and  $\sim$  12.5 ng of genomic DNA. PCR was carried out under the following conditions: initial denaturation for 3 min at  $95^\circ\text{C}$ , followed by 25 cycles of denaturation for 30 s at  $95^\circ\text{C}$ , annealing for 30 s at  $55^\circ\text{C}$  and elongation for 30 s at  $72^\circ\text{C}$ , and a final elongation step for 5 min at  $72^\circ\text{C}$ . PCR products were purified with the QIAquick 96-well PCR Purification Kit and then quantified using Qubit dsDNA BR Assay kit (Invitrogen, Oregon, USA). Samples were equimolar pooled and sequenced on the Illumina MiSeq 2  $\times$  250 bp platform.

Sequences were processed using QIIME 2 pipeline [86]. Demultiplexed 250 bases paired-end sequences were imported using Casava 1.8 format and denoised using DADA2 [91, 92] to obtain amplicon sequence variant (ASV) table. Singletons (ASV present  $< 2$  times) and ASVs that are present in less than 10% of the samples were discarded. Greengenes [93] reference sequences (clustered at 99% similarity) were used to train a naïve Bayes taxonomy classifier to further annotate ASVs taxonomically. ASVs were then collapsed based on genus or lowest-level (i.e., family, order, class, phylum) taxonomy possible. An even sampling depth of 5795 and 33,714 sequences per sample was used for assessing alpha- and beta-diversity measures in the screening and study phase, respectively. Shannon diversity Index and Faith's phylogenetic diversity (PD) was used to measure alpha diversity. Beta-diversity was calculated using principal coordinates analysis (PCoA), Jaccard, and weighted and unweighted UniFrac metrics [94]. Weighted UniFrac distances between microbiota communities were tested by pairwise PERMANOVA using Qiime2 beta-group-significance command with the  $-p$ -pairwise parameter [95]. Also, linear discriminant analysis (LDA) effect size (LEfSe Galaxy Version 1.0) was performed to each microbiota pair (SubA and SubB) for each dietary fiber intervention to elucidate significantly different abundances of bacterial taxa. The parameters used for these analyses were set with default  $P$  value ( $\alpha = 0.05$ ) and LDA score of 2.0 [46].

PICRUSt2 was used to predict functional content or microbiome 16S rRNA genes [43, 44] using QIIME2 generated data. An even sampling depth of 3,919,286 gene counts was used to rarefy all samples to further analyze diversity using Bray Curtis analysis.

### Untargeted metabolomics of serum samples

Untargeted mass spectrometry data was collected at Metabolon Inc from 100  $\mu$ l serum samples of 6 randomly selected mice in each treatment. The 48 samples were prepared using the automated MicroLab STAR system (Hamilton Company). Recovery standards were added and protein, dissociate small molecules bound to protein or trapped in the precipitated protein matrix were removed. To recover chemically diverse metabolites, proteins were precipitated with methanol under vigorous shaking for 2 min (Glen Mills GenoGrinder 2000) followed by centrifugation. The resulting extract was divided into five fractions: two for analysis by two separate reverse-phase UPLC-MS/MS methods with positive ion mode electrospray ionization (ESI), one for analysis by reverse-phase UPLC-MS/MS with negative ion mode ESI, one for analysis by HILIC/UPLC-MS/MS with negative ion mode ESI, and one sample was reserved for backup. Samples were placed briefly on a



TurboVap (Zymark) to remove the organic solvent. The sample extracts were stored overnight under nitrogen before preparation for analysis.

#### **Ultrahigh-performance liquid chromatography-tandem mass spectrometry (UPLC-MS/MS)**

All methods utilized a Waters ACQUITY ultraperformance liquid chromatography system and a Thermo Scientific Q-Exactive high-resolution/accurate mass spectrometer interfaced with a heated ESI source and an Orbitrap mass analyzer operated at 35,000 mass resolution. The sample extract was dried and then reconstituted in solvents compatible to each of the four methods. Each reconstitution solvent contained a series of standards at fixed concentrations to ensure injection and chromatographic consistency. One aliquot was analyzed using acidic positive ion conditions, chromatographically optimized for more hydrophilic compounds. In this method, the extract was gradient eluted from a C18 column (Waters UPLC BEH C18-2.1 × 100 mm, 1.7 μm) using water and methanol, containing 0.05% perfluoropentanoic acid and 0.1% formic acid. Another aliquot was also analyzed using acidic positive ion conditions; however, it was chromatographically optimized for more hydrophobic compounds. In this method, the extract was gradient eluted from the same aforementioned C18 column using methanol, acetonitrile, water, 0.05% perfluoropentanoic acid, and 0.01% formic acid and was operated at an overall higher organic content. Another aliquot was analyzed using basic negative ion optimized conditions using a separate dedicated C18 column. The basic extracts were gradient eluted from the column using methanol and water, amended with 6.5 mM ammonium bicarbonate at pH 8. The fourth aliquot was analyzed via negative ionization following elution from a HILIC column (Waters UPLC BEH Amide 2.1 × 150 mm, 1.7 μm) using a gradient consisting of water and acetonitrile with 10 mM ammonium formate, pH 10.8. Raw data was extracted, peak-identified, and QC processed using Metabolon's hardware and software. Compounds were identified by comparison to library entries of purified standards or recurrent unknown entities based on retention time/index, mass to charge ratio ( $m/z$ ), and chromatographic data, and peaks were quantified using area-under-the-curve.

#### **Statistical analysis**

The dataset comprises a total of 774 biochemicals. Metabolic profiles were quantified in terms of relative abundance and median scaled to 1. Following log transformation and imputation of missing values, if any, with the minimum observed value for each compound, two-way ANOVA contrast were used to identify biochemicals that differed significantly between experimental groups. An FDR adjusted- $P$  value (i.e.,  $q$ -

value) is calculated to take into account the multiple comparisons that normally occur in metabolomic-based studies, and all metabolites with  $q$ -value < 0.05 were included.

#### **RNA-seq analysis**

Mouse liver tissue samples were submitted to the University of Wisconsin Biotechnology Center (UWBC) Gene Expression Center for total RNA extraction. In a 96-well format, tissue samples were lysed using QIAzol Lysis Reagent (Qiagen, Hilden, Germany) and the TissueLyser. Following phase separation by centrifugation, the aqueous phase was recovered, ethanol was added, and the solution was added to an RNeasy 96 Universal Tissue plate. Plate was processed following the RNeasy 96 Universal Tissue protocol. An on-column DNase treatment step was included. RNA was eluted in nuclease-free water. Each sample was quantified and analyzed on a NanoDrop One Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and Agilent 2100 Bioanalyzer (Santa Clara, CA, USA) for purity and integrity, respectively.

Total RNA samples that met the Illumina sample input guidelines were prepared according the TruSeq® Stranded mRNA Sample Preparation Guide (Rev. E) using the Illumina® TruSeq® Stranded mRNA Sample Preparation kit (Illumina Inc., San Diego, CA, USA). For each library preparation, mRNA was purified from 1000 ng total RNA using poly-T oligo-attached magnetic beads. Subsequently, each poly-A-enriched sample was fragmented using divalent cations under elevated temperature. The mRNA fragments were converted to double-stranded cDNA (ds cDNA) using SuperScript II (Invitrogen, Carlsbad, CA, USA), RNaseH, and DNA Pol I, primed by random primers. The ds cDNA was purified with AMPure XP beads (Agencourt, Beckman Coulter). The cDNA products were incubated with Klenow DNA Polymerase to add an "A" base (Adenine) to the 3' end of the blunt DNA fragments. DNA fragments were ligated to unique dual index (UDI) adapters (IDT for Illumina- TruSeq RNA UD Index- catalog 20022371, IDT for Illumina - Nextera DNA Unique Dual Indexes, Set A and custom synthesized UDIs), which have a single "T" base (Thymine) overhang at their 3' end. The adapter-ligated DNA products were purified with AMPure XP beads. Adapter-ligated DNA was amplified in a Linker Mediated PCR reaction (LM-PCR) for 10 cycles using Phusion™ DNA Polymerase and Illumina's PE genomic DNA primer set followed by purification with AMPure XP beads. Finally, the quality and quantity of the finished libraries were assessed using an Agilent DNA1000 chip (Agilent Technologies, Inc., Santa Clara, CA, USA) and Qubit® dsDNA HS Assay Kit (Invitrogen, Carlsbad, CA, USA), respectively. Libraries were standardized to 2 nM. Paired-end 2 × 150 bp sequencing

was performed, using standard SBS chemistry on an Illumina NovaSeq6000 sequencer. Images were analyzed using the standard Illumina Pipeline, version 1.8.2.

High-quality reads were obtained after removal of adaptor sequences and high content of unknown bases. Filtered reads were mapped to mouse genome reference GRCm38.p6 from Ensemble (release 97) database using Spliced Transcripts Alignment to Reference tool (STAR, v2.7.2a). Further quantification of mapped transcript reads to the reference was performed using featureCounts tool (v1.6.4). Readcount data was analyzed for differential gene expression using edgeR Bioconductor package (version 3.28.0) performing the generalized linear model quasi-likelihood F-test from empirical Bayes methods to estimate gene-specific biological variation [96]. Quasi-likelihood method accounts for uncertainty in dispersion estimation, therefore, gives stricter error rate control which makes it ideal for differential expression analyses of our RNA-seq data [96]. Filtering of low expressed genes was performed by keeping counts in a minimum number of samples computed through the built-in function “filterByExpr”. Normalization for RNA composition effect was performed in order to compare relative changes in expression levels between conditions using trimmed mean of  $M$ -values (TMM) between each pair of samples [97]. Biological and technical variability estimation was performed by analyzing sample replicates. Negative binomial generalized linear models were fitted and tagwise dispersion estimates were obtained using Cox-Reid profile-adjusted likelihood method in order to determine differential expression. Differentially expressed genes (DEGs) were obtained after filtering by logarithmic fold-change 1 and  $-1$ ,  $P$  value 0.05, and false discovery rate (FDR adjusted- $P$ ) 0.05, yielding 55,573 genes in total. Gene Ontology (GO) functional enrichment analysis for biological process (BP), cellular component (CC), molecular function (MF), and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways was obtained as part of the downstream procedure to interpret the differential expression of genes. We performed Gene Ontology and KEGG pathways enrichment analyses for the differentially expressed genes. The resulting gene representation data was then filtered to include only GO terms or KEGG pathways with at least 10 annotated genes; a minimum of 5 genes that were significantly regulated and a  $P$  value ( $\geq 5$ ).

#### **Mass spectrometry analysis of post-translational modification (PTM) of histones from liver**

##### ***Tissue fractionation and histone extraction and label-free chemical derivatization from liver***

Tissue fractionation and histone acid extraction was performed using previously published protocols [72, 98].

Briefly, 50–100 mg of frozen mice liver tissue ( $n = 32$ , 4 samples per treatment) was dounce-homogenized on ice in a hypotonic lysis buffer containing histone deacetylase and protease inhibitors, followed by the centrifugation to pellet nuclei. Histones were acid extracted, and a Bradford assay was performed to quantify protein yield. In total, 5  $\mu\text{g}$  of dried histone extract was then subjected to hybrid chemical derivatization [99] using heavy acetic anhydride. This procedure was followed by trypsinization for 4 h and derivatization of newly generated peptide N-termini with phenylisocyanate (PIC). Finally, labelled histones were desalted using C18 stage tips.

##### ***Nano-liquid chromatography and electrospray ionization tandem MS***

For each sample, derivatized histone peptides were injected onto a Dionex Ultimate3000 nanoflow HPLC with a Waters nanoAcquity UPLC C18 column (100 m  $\times$  150 mm, 3 m) coupled to a Thermo Fisher Q-Exactive mass spectrometer at 700 nL/min. Mobile phase consisted of water + 0.1% formic acid (A) and acetonitrile + 0.1% formic acid (B). Histone peptides were resolved with a 2-step linear gradient of 2 to 25% mobile phase B over 60 min followed by 25 to 40% mobile phase B over 15 min. Data was acquired using data-independent acquisition (DIA) mode. The mass spectrometer was operated with a MS1 scan at resolution = 35,000, automatic gain control target =  $1 \times 10^6$ , and scan range = 390–910 m/z, followed by a DIA scan with a loop count of 10. DIA settings were as follows: window size = 10 m/z, resolution = 17,500, automatic gain control target =  $1 \times 10^6$ , DIA maximum fill time = AUTO, and normalized collision energy = 30. For each cycle, one full MS1 scan was followed by 10 MS2 scans using an isolation window size of 10 m/z.

##### ***Histone PTM quantification***

EpiProfile 2.0 was used for quantification of histone PTMs [100]. The R script provided by Denu lab (<https://github.com/DenuLab/HistoneAnalysisWorkflow>) was used to perform data cleaning, normalization, statistical analysis, and visualization.

#### **Data processing for microbiome, metabolome, and transcriptome**

##### ***Normalization/transformation/filtering***

For the taxa that were unclassified at the genus level, their identities at higher levels were used. We combined all ASVs belonging to the same genus and filtered out the genera that appear in fewer than 20% of total samples, leaving 5 phyla, 11 classes, 11 orders, 25 families, and 45 genera. For the metabolome data, we normalized biochemicals by using inverse normal transformation and transformed variables that did not follow a normal

distribution (Shapiro-Wilk test  $P < 0.05$ ) were removed, resulting in 712 biochemicals. The RNA-seq data were normalized with voom methodology [101]. The voom method estimates the mean-variance relationship and computes appropriate observation-level weights to transform count data to log<sub>2</sub>-counts per million. We used median absolute deviation to measure the variability of each gene because two genes without notable variance between samples will be highly correlated. As a heuristic cut-off, the top 5000 most variant genes had been used in the downstream analysis.

#### **Weighted correlation network analysis (WGCNA) for metabolome and transcriptome**

In order to group the biochemicals that were highly correlated, we built the co-expression network using WGCNA [102]. The WGCNA is an efficient and robust method in grouping metabolomic and transcriptomic data [103, 104] and allowed us to summarize each module by its module eigenvalue. A one-sided Fisher test was used to determine if a pathway was enriched within the turquoise and blue modules in metabolomic data.  $P$  values were then adjusted using Benjamini-Hochberg method, and a cut-off of  $P < 0.05$  and FDR adjusted- $P < 0.05$  were chosen to determine if a pathway was significantly enriched. We used Pearson's correlation between expression profile of each gene and module eigenvalue to identify module membership. Using the module eigenvalue, the module-traits relationships were estimated by calculating Pearson's correlations between the module eigenvalue and the traits of interest. We considered 0.90 as a correlation cut-off to choose soft-thresholding power and set the minimal module size as 20. For metabolome, the metabolites were clustered into 8 modules plus 43 unclustered metabolites. The transformed values of the unclustered metabolites were combined with standardized module eigenvalues in the following analysis. For the transcriptome data, 14 modules (defined as clusters of highly interconnected genes) were identified by using DynamicTree Cut algorithm. WGCNA led to 14 different modules by using DynamicTree Cut algorithm. Over-representation of genes in the blue module was characterized based on gene ontology biological process and KEGG pathways using clusterProfiler [105].

#### **Microbiome association with host -omics**

We applied the sparse linear log-contrast model [106] to pinpoint important genera that are associated with individual metabolite/gene. In this model, the host omics variable is the response and the genus-level microbial taxa are compositional covariates. The sparse linear log-contrast model respects the compositional nature of the microbiome data and avoids choosing an arbitrary

reference taxon, in which the unit-sum constraint on the compositional vector is translated into the zero-sum constraint on the association coefficients across taxa in log ratio scale. In our analysis, we used 10-fold cross validation to choose the tuning parameter. To obtain stable selection results, we generated 100 bootstrap samples and used the same cross validation procedure to select the genera. We also followed the stability selection approach [107] to assess the stability of the selected genera, where 100 subsamples of half sample size were taken to compute the selection probabilities. In the association network (Figs. 6 and 9), we kept genera with stability selection probability larger than 0.85 and filter out the lowly associated genera which the absolute value of coefficient is less than 0.1.

#### **Causal mediation analysis**

We performed mediation analyses to investigate how the post-intervention gut microbiome/metabolites/gene expressions may mediate the effect of fecal colonization on various phenotypes (adiposity, liver TG, glucose) under a given diet. We employed the causal mediation model with batch effects as confounder, SubA/B microbial communities as exposure, and different phenotypes as outcome. All the mediation effect hypotheses were tested by using the resampling method. Below we described two approaches for testing microbiome mediation effects on the global community level and on the subcompositions defined on the taxonomy tree. For the metabolites and genes, we applied mediation analysis to the corresponding WGCNA modules.

#### **Beta-diversity mediation analysis for microbiome**

For beta-diversity distance matrices, we performed the distance-based mediation test by using the MedTest package in R language [108]. The Jaccard and unweighted UniFrac distance matrices were calculated based on the rarefied genus-level abundance matrix (rarefied to the minimum sequence depth) to reduce potential sequence depth-dependent bias.

#### **Tree-based mediation analysis microbiome**

We used maximum round error 0.5 to replace 0 [109] in full-composition abundance matrix, then calculated the sub-composition relative abundance matrix for each high-rank internal node. We removed extremely rare taxa (only detected in 10% of the observations or less). We selected the most abundant taxa in the relative abundance matrix as the reference taxa and took additive logarithm transformation on the compositional data so that the transformed data could be considered as multivariate variables. Finally, we applied the causal mediation model on each high-rank internal node in the taxonomy tree.

### Metabolite module/gene module mediation analysis

In order to identify the microbial community status effect on different phenotypes which were transmitted through the metabolite modules/gene modules, we applied causal mediation model by considering the corresponding WGCNA modules as multiple independent mediators.

#### Abbreviations

ANOVA: Analysis of variance; Af: Assorted fiber; ASV: Amplicon sequence variant; BCAA: Branched-chain amino acid; BCFA: Branched-chain fatty acids; BP: Biological process; C: Cellulose; CC: Cellular component; DEGs: Differentially expressed genes; ds cDNA: Double-stranded cDNA; F/B: Firmicutes:Bacteroidetes ratio; FAST: Fecal aliquot straw technique; FC: Fold-change; FDR: False discovery rate; GF: Germ-free; GO: Gene ontology; GSH: Glutathione; I: Inulin; KEGG: Kyoto Encyclopedia of Genes and Genomes; LEfSe: Linear discriminant analysis effect size; MACs: Microbiota-accessible carbohydrates; MF: Molecular function; *ob/ob*: Genetically obese; P: Pectin; PC: Phosphatidylcholine; PCoA: Principal coordinates analysis; PD: Phylogenetic diversity; PERMANOVA: Permutational multivariate analysis of variance; PPAR: Peroxisome proliferator-activated receptor; PTMs: Post-translational modifications; RNA-seq: RNA sequencing; RS: Resistant starch; SCFAs: Short-chain fatty acids; scFOS: Short-chain fructo-oligosaccharides; SubA: Subject A; SubB: Subject B; TG: Liver triglycerides; TMM: Trimmed mean of M-values; TNF- $\alpha$ : Tumor necrosis factor alpha; TRP: Transient receptor potential; UPLC-MS/MS: Ultrahigh-performance liquid chromatography-tandem mass spectroscopy; uwUF: Unweighted UniFrac; WGCNA: Weighted correlation network analysis; WLS: Wisconsin Longitudinal Study; wUF: Weighted UniFrac

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40168-021-01061-6>.

**Additional file 1: Fig. S1.** Screening phase. 16S rRNA gene sequence analysis of fecal samples from human donors and recipient mice. A. Principal Coordinates Analysis (PCoA) of unweighted UniFrac (uwUF) distances from eight human fecal samples collected for Wisconsin Longevity Study. B. Bacterial relative abundance summarized at the phylum level. C. Percentage of genera shared between each donor fecal sample and its corresponding recipient mice cecal samples. Number of mice colonized for each fecal donor is reported under each bar. D. Percent relative abundance of the fecal donor community captured in the mouse cecal samples. E. uwUF distances between donor fecal samples and each engrafted cecal community. Averages of distances between corresponding human donor-mouse engrafted community are indicated as DONOR. Average of uwUF distances between non-matched donor-mouse community are indicated as OTHER. F. PCoA of uwUF distances of the eight human fecal samples engrafted in the mouse cecum. Circles with the same colors indicate biological replicates colonized with the same community. G. Alpha-diversity as determined by Faith's phylogenetic diversity of each of the eight engrafted communities. \*\*\*\* $P < 0.001$ .

**Additional file 2: Fig. S2.** Variation in cecal short-chain fatty acids among transplanted communities. Cecal levels of (A) butyrate; (B) acetate; and (C) propionate ( $\mu\text{moles/g}$  wet weight) for each the eight transplanted groups of mice described in Fig. S1.

**Additional file 3: Fig. S3.** Variation in predicted metabolic capacity among engrafted gut communities. Principal Coordinates Analysis (PCoA) of Bray Curtis dissimilarity using the PICRUSt2 predicted metabolic functions of the eight transplanted human microbiota samples used in this study. Circles with the same colors indicate biological replicates colonized with the same community.

**Additional file 4: Fig. S4.** Characterization of transplanted communities in mice. 16S rRNA gene sequence analysis of engrafted cecal communities. Germ-free mice were colonized with SubA or SubB and

exposed to one of four diets containing a different type of fiber; (i) Cellulose; (ii) Inulin; (iii) Pectin; or (iv) Assorted fiber. A. Heatmap showing presence/absence of bacterial taxa in the gut of transplanted animals across the four different diets. Red indicates presence and black absence. Each column represents an individual mouse. B. Alpha diversity (Shannon Index) of SubA and SubB communities after dietary intervention. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ .

**Additional file 5: Fig. S5.** Differences in gut microbiota between SubA- and SubB-colonized animals across the different diets used. Weighted UniFrac distances between fecal microbiomes of SubA and SubB colonized mice. The UniFrac matrix was permuted 999 times;  $n = 7$ -10 animals/microbiome/diet.

**Additional file 6: Fig. S6.** Individual effect of dietary fibers on Firmicutes to Bacteroidetes ratio. Comparison of Firmicutes and Bacteroidetes between engrafted SubA and SubB communities across different diets. A. Relative abundance of Bacteroidetes (white) and Firmicutes (black) in SubA (magenta) and SubB (yellow) colonized mice for each dietary fiber intervention. B. Firmicutes:Bacteroidetes ratio in SubA (magenta) and SubB (yellow) for the four dietary interventions. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ .

**Additional file 7: Fig. S7.** Linear discriminant analysis Effect Size (LEfSe) summary. List of taxa differentially abundant between gut community SubA (magenta) and SubB (yellow) in the four diets. LDA score (log 10) is indicated at the bottom of each graph.

**Additional file 8: Fig. S8.** Relative abundance of gut bacterial taxa for SubA and SubB. Box plots indicating relative abundance of taxa of interest relevant to the diversity, association, and mediation analyses. This figure shows relative abundance of taxa that has at least one significant difference between SubA and SubB within a dietary intervention. SubA is represented with the color magenta and SubB with the color yellow. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ .

**Additional file 9: Fig. S9.** Short Chain Fatty Acids (SCFA). Cecal levels of (A) acetate; (B) butyrate, (C) propionate and total SCFA ( $\mu\text{moles/g}$  wet wt) E. valerate and Branched-chain Fatty Acids (BCFA) Isobutyrate and Isovalerate of SubA (magenta) and SubB (yellow) by diet. Wilcoxon test comparison \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , \*\*\*\* $P < 0.0001$ .

**Additional file 10: Fig. S10.** Dendrogram of serum metabolites from transplanted mice. Clustering dendrograms of 712 serum metabolites with dissimilarity based on topological overlap, together with assigned module colors. There are 9 modules that cluster different numbers of metabolites.

**Additional file 11: Fig. S11.** Correlations between Short Chain Fatty Acids and blood metabolites. Correlation matrix between metabolites consensus modules and cecal SCFAs from mice described in Fig. 1. Modules were determined based on patterns of co-abundance of metabolites using weighted correlation network analysis (WGCNA). Each of the modules was labelled with a unique color as an identifier. Each module was tested for correlation with each cecal SCFA quantified. Within each cell, upper values are correlation coefficients between module and the phenotypes; lower values are the corresponding FDR adjusted- $P$  values. B. Pathways enriched in the yellow, blue and turquoise modules as determined by Fisher's test.

**Additional file 12: Fig. S12.** Branched-Chain Fatty Acids and taxa correlation. The heatmap shows all correlations ( $P < 0.05$ ) using Spearman method between BCFA and taxa in each dietary intervention.

**Additional file 13: Fig. S13.** Gut community-mediated epigenetic changes in liver are sensitive to dietary fiber. Abundance of histone Post-Translational Modifications (PTMs) on H3 lysines (K9, K14, K27, and K36). A. H3K9K14 peptide. B. H3K27K36 peptide ( $n = 4$ /community/diet). \* $P < 0.05$ , \*\* $P < 0.01$ ; ac, acetylated; unmod, unmodified; meth1,2,3, mono- di- and try- methylated respectively; pr, propionylated.

**Additional file 14: Fig. S14.** Effect of gut community on liver histone post-translational modifications (PTMs). Heatmap show relative difference in abundance for each histone PTM quantified in liver for mice colonized with the two communities (Log2 fold-change of SubB/SubA) in each diet. \* $P < 0.1$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  ( $n = 3$ -4/condition). ac, acetylated;



unmod, unmodified; meth1,2,3, mono- di- and try- methylated respectively; pr, propionylated.

**Additional file 15: Fig. S15.** Dendrogram of liver transcripts from transplanted mice. Clustering dendrograms of genes with dissimilarity based on topological overlap, together with assigned module colors. There are 14 modules that cluster different numbers of transcripts.

**Additional file 16: Fig. S16.** Gene Ontology and KEGG pathway enrichment of transcripts in the blue module associated with metabolic phenotypes. A. Biological Process GO and KEGG enrichment of blue module associated with adiposity, B. Association with liver triglycerides. C. Association with glucose. Gene counts and FDR adjusted-*P* values are indicated for each enrichment box.

**Additional file 17: Table S1.** Diets used in the study.

**Additional file 18: Table S2.** Statistical analysis of metabolic phenotypes in transplanted mice.

**Additional file 19: Table S3.** Relative abundance of bacterial ASVs detected in mice. 16S rRNA gene sequence analysis of cecal samples from gnotobiotic mice colonized with SubA or SubB communities consuming one of four diets containing different types of fiber; (i) Cellulose; (ii) Inulin; (iii) Pectin; or (iv) Assorted fiber for four weeks.

**Additional file 20: Table S4.** Relative abundance of bacterial genera detected in mice. 16S rRNA gene sequence analysis of cecal samples from gnotobiotic mice colonized with SubA or SubB communities consuming one of four diets containing different type of fiber; (i) Cellulose; (ii) Inulin; (iii) Pectin; or (iv) Assorted fiber for four weeks. Taxa summarized at the genus level.

**Additional file 21: Table S5.** Mediation analysis results. For each dietary fiber intervention associated phenotype a mediator type was evaluated.

**Additional file 22: Table S6.** List of serum metabolites measured in transplanted mice. Fold-change (FC) between SubB and SubA is indicated as follow: darker green indicates  $FC < 1$  with  $P \leq 0.05$ , lighter green indicates  $FC < 1$  with  $0.05 < P < 0.1$ , darker red indicates  $FC \geq 1$  with  $P \leq 0.05$ , and lighter red indicates  $FC \geq 1$  with  $0.05 < P < 0.1$ . Two-Way ANOVA main effects: Community, diet, and their interaction are indicated in blue shaded cells when significant ( $P \leq 0.05$ ) ANOVA effect; light blue shaded cells indicate  $0.05 < P < 0.10$ .

**Additional file 23: Table S7.** Weighted Correlation Network Analysis assignment of serum metabolites into modules. Abbreviations: MS, Metabolite Significance; p.MS., *P*-value of Metabolite Significance; MM, module membership; p.MM *P*-value of the module membership.

**Additional file 24: Table S8.** Serum metabolites contained in the blue and turquoise modules. Metabolites (biochemicals) are organized by pathways. Fold-change between SubB and SubA is indicated for the four diets as follow: darker green indicates  $FC < 1$  with  $P \leq 0.05$ , lighter green indicates  $FC < 1$  with  $0.05 < P < 0.1$ , darker red indicates  $FC \geq 1$  with  $P \leq 0.05$ , and lighter red indicates  $FC \geq 1$  with  $0.05 < P < 0.1$ . The last two columns show metabolite significance (MS) which reports the association of each metabolite with each phenotype and the corresponding *P*-value (p.MS).

**Additional file 25: Table S9.** Liver gene expression analysis. List of genes differentially expressed in the liver from mice colonized with SubB vs. SubA communities. (LogFC > 1 and < -1, *P* and FDR adjusted-*P* < 0.05).

**Additional file 26: Table S10.** Analysis of differentially expressed genes in liver. Number of differentially regulated genes contained in each listed KEGG pathway and Gene Ontology Term.

**Additional file 27: Table S11.** Weighted Correlation Network Analysis assignment of liver transcripts. Abbreviations: GS, Gene Significance; p.MS., *P*-value of Gene Significance; MM, module membership; p.MM for the *P*-value of the module membership.

**Additional file 28: Table S12.** List of liver transcripts associated with bacterial taxa. bcv.prob denotes the probability of specific taxa being selected in log-contrast model based on 100 bootstrap samples; stab.-prob denotes the probability of specific taxa being selected in log-contrast model based on 100 subsamples of half sample size. refitted.cof

denotes the coefficient estimation in the log-contrast model based on selected taxa listed in this table.

**Additional file 29. Supplemental results.** Effects of microbiota-fiber interactions on liver histone posttranslational modifications.

### Acknowledgements

The authors thank the University of Wisconsin Biotechnology Center DNA Sequencing Facility for providing sequencing and support services, and the University of Wisconsin Center for High Throughput Computing (CHTC) in the Department of Computer Sciences for providing computational resources, support, and assistance. Part of the bioinformatics analysis was performed in the Cluster Navira, CIDIE - CONICET - UCC (Plataforma Nacional de Bioinformática). The authors also thank Mr. Chris Pruessner and Dr. Ody Maningat from MPG for their generosity in donating Fibersym for our diet formulation.

### Authors' contributions

TWLC and FER conceived the study. TWLC, SMMG, ERH, and EIV performed mouse studies and collected phenotypic and transcriptomic data. JH, SPT, and JD conducted and analyzed epigenetic studies. QH, SMMG, DGC, and ZZT conducted statistical analyses. SMMG, QH, ZZT, and FER wrote the manuscript. All authors read and approved the final manuscript.

### Funding

This work was supported in part by grants NIH DK108259 (to F.E.R.), GM140464 (to Z.T) and by the Food Safety, Nutrition, and Health program under grant no. 2019-67017-29255 /project accession no. 1018638 and Hatch grant WIS03073 both from the USDA National Institute of Food and Agriculture (to F.E.R.). This work was also supported in part by a Transatlantic Networks of Excellence Award from the Leducq Foundation. T.-W.L.C was supported by the National Institutes of Health, under Ruth L. Kirschstein National Research Service Award T32 HL 007936 from the National Heart Lung and Blood Institute to the University of Wisconsin-Madison Cardiovascular Research Center. S.M.M.G. was supported by "El Consejo Nacional de Ciencia y Tecnología - CONACYT" (The National Council for Science and Technology) with register number 895733 and by "Programa 2018 para Actividades Especiales de Cooperación Interinstitucional con Fines de Internacionalización para alumnos y egresados de nivel licenciatura de la UNAM" (PAECI Scholarship) offered by the "Dirección General de Cooperación e Internacionalización (DGECI)" of the "Universidad Nacional Autónoma de México (UNAM)." Funding for this research was also provided by Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT) PICT-2017-4497. D.G.C is staff researcher of CONICET.

### Availability of data and materials

The data reported in this paper are accessible in the NCBI Short Read Archive (SRA) under accession ID PRJNA665643 and European Nucleotide Archive (ENA) accession number PRJEB40242. Original R scripts and data used for statistical analysis are available in GitHub (<https://github.com/KIRinHong/fiberNmicrobiome>).

### Declarations

#### Ethics approval and consent to participate

The use of Wisconsin Longitudinal Study fecal samples was approved by the Institutional Review Board at the University of Wisconsin-Madison.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Department of Bacteriology, University of Wisconsin-Madison, 1550 Linden Dr., Madison, WI 53706, USA. <sup>2</sup>PECEM (MD/PhD), Facultad de Medicina, Universidad Nacional Autónoma de México, Coyoacán, Ciudad de México, México. <sup>3</sup>Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison, 600 Highland Avenue, Madison, WI 53792, USA. <sup>4</sup>Present Address: Department of Nutrition Science, Purdue University, 700 W. State



Street, Stone Hall 205, West Lafayette, IN 47907, USA. <sup>5</sup>Wisconsin Institute for Discovery, Madison, WI, USA. <sup>6</sup>Unidad de Bioinformática Traslacional, Centro de Investigación en Medicina Traslacional Severo Amuchástegui, Instituto Universitario de Ciencias Biomédicas de Córdoba, Av. Naciones Unidas 420, 5000 Córdoba, CP, Argentina.

Received: 19 October 2020 Accepted: 24 March 2021

Published online: 20 May 2021

## References

- Li J, Jia H, Cai X, Zhong H, Feng Q, Sunagawa S, et al. An integrated catalog of reference genes in the human gut microbiome. *Nat Biotechnol*. 2014; 32(8):834–41. <https://doi.org/10.1038/nbt.2942>.
- Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. *Nature*. 2012;486(7402):222–7. <https://doi.org/10.1038/nature11053>.
- Kau AL, Ahern PP, Griffin NW, Goodman AL, Gordon JL. Human nutrition, the gut microbiome and the immune system. *Nature*. 2011;474(7351):327–36. <https://doi.org/10.1038/nature10213>.
- Thursby E, Juge N. Introduction to the human gut microbiota. *Biochem J*. 2017;474(11):1823–36. <https://doi.org/10.1042/BCJ20160510>.
- Deschasaux M, Bouter KE, Prodan A, Levin E, Groen AK, Herrema H, et al. Depicting the composition of gut microbiota in a population with varied ethnic origins but shared geography. *Nat Med*. 2018;24(10):1526–31. <https://doi.org/10.1038/s41591-018-0160-1>.
- Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, et al. Population-level analysis of gut microbiome variation. *Science*. 2016; 352(6285):560–4. <https://doi.org/10.1126/science.aad3503>.
- The GBD 2015 Obesity Collaborators. Health effects of overweight and obesity in 195 countries over 25 years. *N Engl J Med*. 2017;377:13–27.
- Wei B, Liu Y, Lin X, Fang Y, Cui J, Wan J. Dietary fiber intake and risk of metabolic syndrome: a meta-analysis of observational studies. *Clin Nutr*. 2018;37(6):1935–42. <https://doi.org/10.1016/j.clnu.2017.10.019>.
- Bäckhed F, Manchester JK, Semenkovich CF, Gordon JL. Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *PNAS*. 2007;104(3):979–84. <https://doi.org/10.1073/pnas.0605374104>.
- Cho I, Yamanishi S, Cox L, Methé BA, Zavadil J, Li K, et al. Antibiotics in early life alter the murine colonic microbiome and adiposity. *Nature*. 2012; 488(7413):621–6. <https://doi.org/10.1038/nature11400>.
- Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau AL, et al. Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science*. 2013;341(6150):1241214. <https://doi.org/10.1126/science.1241214>.
- Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JL. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature*. 2006;444(7122):1027–31. <https://doi.org/10.1038/nature05414>.
- Vrieze A, Van Nood E, Holleman F, Salojärvi J, Kootte RS, Bartsman JFWM, et al. Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. *Gastroenterology*. 2012; 143:913–916.e7.
- Koeth RA, Wang Z, Levison BS, Buffa JA, Org E, Sheehy BT, et al. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat Med*. 2013;19(5):576–85. <https://doi.org/10.1038/nm.3145>.
- Koh A, Molinaro A, Ståhlman M, Khan MT, Schmidt C, Mannerås-Holm L, et al. Microbially produced imidazole propionate impairs insulin signaling through mTORC1. *Cell*. 2018;175:947–961.e17.
- Romano KA, Vivas EI, Amador-Noguez D, Rey FE. Intestinal microbiota composition modulates choline bioavailability from diet and accumulation of the proatherogenic metabolite trimethylamine-N-oxide. *mBio*. 2015;6: e02481–14.
- Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, DuGar B, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature*. 2011;472(7341):57–63. <https://doi.org/10.1038/nature09922>.
- De Vadder F, Kovatcheva-Datchary P, Goncalves D, Vinera J, Zitoun C, Duchamp A, et al. Microbiota-generated metabolites promote metabolic benefits via gut-brain neural circuits. *Cell*. 2014;156(1-2):84–96. <https://doi.org/10.1016/j.cell.2013.12.016>.
- Kasahara K, Krautkramer KA, Org E, Romano KA, Kerby RL, Vivas EI, et al. Interactions between *Roseburia intestinalis* and diet modulate atherogenesis in a murine model. *Nat Microbiol*. 2018;3(12):1461–71. <https://doi.org/10.1038/s41564-018-0272-x>.
- Koh A, De Vadder F, Kovatcheva-Datchary P, Bäckhed F. From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. *Cell*. 2016;165(6):1332–45. <https://doi.org/10.1016/j.cell.2016.05.041>.
- Makki K, Deehan EC, Walter J, Bäckhed F. The impact of dietary fiber on gut microbiota in host health and disease. *Cell Host Microbe*. 2018;23(6):705–15. <https://doi.org/10.1016/j.chom.2018.05.012>.
- Deehan EC, Duar RM, Armet AM, Perez-Muñoz ME, Jin M, Walter J. Modulation of the gastrointestinal microbiome with nondigestible fermentable carbohydrates to improve human health. *Microbiol Spectr*. 2017;5:24.
- Wang H, Hong T, Li N, Zang B, Wu X. Soluble dietary fiber improves energy homeostasis in obese mice by remodeling the gut microbiota. *Biochem Biophys Res Commun*. 2018;498(1):146–51. <https://doi.org/10.1016/j.bbrc.2018.02.017>.
- Sonnenburg ED, Sonnenburg JL. Starving our microbial self: the deleterious consequences of a diet deficient in microbiota-accessible carbohydrates. *Cell Metab*. 2014;20(5):779–86. <https://doi.org/10.1016/j.cmet.2014.07.003>.
- Martens EC, Kelly AG, Tausin AS, Brumer H. The devil lies in the details: how variations in polysaccharide fine-structure impact the physiology and evolution of gut microbes. *J Mol Biol*. 2014;426(23):3851–65. <https://doi.org/10.1016/j.jmb.2014.06.022>.
- Gentile CL, Weir TL. The gut microbiota at the intersection of diet and human health. *Science*. 2018;362(6416):776–80. <https://doi.org/10.1126/science.aau5812>.
- Cani PD, Possemiers S, Van de Wiele T, Guiot Y, Everard A, Rottier O, et al. Changes in gut microbiota control inflammation in obese mice through a mechanism involving GLP-2-driven improvement of gut permeability. *Gut*. 2009;58(8):1091–103. <https://doi.org/10.1136/gut.2008.165886>.
- Chassaing B, Miles-Brown J, Pellizzon M, Ulman E, Ricci M, Zhang L, et al. Lack of soluble fiber drives diet-induced adiposity in mice. *Am J Physiol Gastrointest Liver Physiol*. 2015;309(7):G528–41. <https://doi.org/10.1152/ajpgi.00172.2015>.
- Kindt A, Liebisch G, Clavel T, Haller D, Hörmannspurger G, Yoon H, et al. The gut microbiota promotes hepatic fatty acid desaturation and elongation in mice. *Nat Commun*. 2018;9(1):3760. <https://doi.org/10.1038/s41467-018-05767-4>.
- Veronese N, Solmi M, Caruso MG, Giannelli G, Osella AR, Evangelou E, et al. Dietary fiber and health outcomes: an umbrella review of systematic reviews and meta-analyses. *Am J Clin Nutr*. 2018;107(3):436–44. <https://doi.org/10.1093/ajcn/nqx082>.
- Armet AM, Deehan EC, Thöne JV, Hewko SJ, Walter J. The effect of isolated and synthetic dietary fibers on markers of metabolic diseases in human intervention studies: a systematic review. *Adv Nutr*. 2020;11(2):420–38. <https://doi.org/10.1093/advances/nmz074>.
- Healey GR, Murphy R, Brough L, Butts CA, Coad J. Interindividual variability in gut microbiota and host response to dietary interventions. *Nutr Rev*. 2017;75(12):1059–80. <https://doi.org/10.1093/nutrit/nux062>.
- Hegele RA, Wolever TM, Story JA, Connelly PW, Jenkins DJ. Intestinal fatty acid-binding protein variation associated with variation in the response of plasma lipoproteins to dietary fibre. *Eur J Clin Invest*. 1997;27(10):857–62. <https://doi.org/10.1046/j.1365-2362.1997.2010748.x>.
- Korpela K, Flint HJ, Johnstone AM, Lappi J, Poutanen K, Dewulf E, et al. Gut microbiota signatures predict host and microbiota responses to dietary interventions in obese individuals. *PLoS One*. 2014;9(3):e90702. <https://doi.org/10.1371/journal.pone.0090702>.
- Kovatcheva-Datchary P, Nilsson A, Akrami R, Lee YS, De Vadder F, Arora T, et al. Dietary fiber-induced improvement in glucose metabolism is associated with increased abundance of Prevotella. *Cell Metab*. 2015;22(6): 971–82. <https://doi.org/10.1016/j.cmet.2015.10.001>.
- Martínez I, Lattimer JM, Hubach KL, Case JA, Yang J, Weber CG, et al. Gut microbiome composition is linked to whole grain-induced immunological improvements. *ISME J*. 2013;7(2):269–80. <https://doi.org/10.1038/ismej.2012.104>.
- Nguyen NK, Deehan EC, Zhang Z, Jin M, Baskota N, Perez-Muñoz ME, et al. Gut microbiota modulation with long-chain corn bran arabinoxylan in adults with overweight and obesity is linked to an individualized temporal increase in fecal propionate. *Microbiome*. 2020;8(1):118. <https://doi.org/10.1186/s40168-020-00887-w>.
- Dill-McFarland KA, Tang Z-Z, Kemis JH, Kerby RL, Chen G, Palloni A, et al. Close social relationships correlate with human gut microbiota composition. *Sci Rep*. 2019;9(1):703. <https://doi.org/10.1038/s41598-018-37298-9>.

39. McGill CR, Fulgoni VL, Devareddy L. Ten-year trends in fiber and whole grain intakes and food sources for the United States population: National Health and Nutrition Examination Survey 2001-2010. *Nutrients*. 2015;7(2):1119–30. <https://doi.org/10.3390/nu7021119>.
40. Shan Z, Rehm CD, Rogers G, Ruan M, Wang DD, Hu FB, et al. Trends in dietary carbohydrate, protein, and fat intake and diet quality among US adults, 1999-2016. *JAMA*. 2019;322(12):1178–87. <https://doi.org/10.1001/jama.2019.13771>.
41. Kasubuchi M, Hasegawa S, Hiramatsu T, Ichimura A, Kimura I. Dietary gut microbial metabolites, short-chain fatty acids, and host metabolic regulation. *Nutrients*. 2015;7(4):2839–49. <https://doi.org/10.3390/nu7042839>.
42. McOrist AL, Miller RB, Bird AR, Keogh JB, Noakes M, Topping DL, et al. Fecal butyrate levels vary widely among individuals but are usually increased by a diet high in resistant starch. *J Nutr*. 2011;141(5):883–9. <https://doi.org/10.3945/jn.110.128504>.
43. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, et al. PICRUSt2 for prediction of metagenome functions. *Nat Biotechnol*. 2020;38(6):685–8. <https://doi.org/10.1038/s41587-020-0548-6>.
44. Langille MGJ, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol*. 2013;31(9):814–21. <https://doi.org/10.1038/nbt.2676>.
45. Gibson GR, Hutkins R, Sanders ME, Prescott SL, Reimer RA, Salminen SJ, et al. Expert consensus document: The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics. *Nat Rev Gastroenterol Hepatol*. 2017;14(8):491–502. <https://doi.org/10.1038/nrgastro.2017.75>.
46. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. *Genome Biol*. 2011;12(6):R60. <https://doi.org/10.1186/gb-2011-12-6-r60>.
47. Ramirez-Farias C, Slezak K, Fuller Z, Duncan A, Holtrop G, Louis P. Effect of inulin on the human gut microbiota: stimulation of *Bifidobacterium adolescentis* and *Faecalibacterium prausnitzii*. *Br J Nutr*. 2021;101:541–50.
48. Bounhik Y, Raskine L, Simoneau G, Paineau D, Bornet F. The capacity of short-chain fructo-oligosaccharides to stimulate faecal bifidobacteria: a dose-response relationship study in healthy humans. *Nutr J*. 2006;5(1). <https://doi.org/10.1186/1475-2891-5-8>.
49. Cantu-Jungles TM, Hamaker BR. New view on dietary fiber selection for predictable shifts in gut microbiota. *mBio*. 2020;11:e02179–19.
50. Venkataraman A, Sieber JR, Schmidt AW, Waldron C, Theis KR, Schmidt TM. Variable responses of human microbiomes to dietary supplementation with resistant starch. *Microbiome*. 2016;4(1):33. <https://doi.org/10.1186/s40168-016-0178-x>.
51. Sanna S, van Zuydam NR, Mahajan A, Kurilshikov A, Vich Vila A, Vösa U, et al. Causal relationships among the gut microbiome, short-chain fatty acids and metabolic diseases. *Nat Genet*. 2019;51(4):600–5. <https://doi.org/10.1038/s41588-019-0350-x>.
52. Heaver SL, Johnson EL, Ley RE. Sphingolipids in host-microbial interactions. *Curr Opin Microbiol*. 2018;43:92–9. <https://doi.org/10.1016/j.mib.2017.12.011>.
53. Boini KM, Xia M, Koka S, Gehr TW, Li PL. Sphingolipids in obesity and related complications. *Front Biosci (Landmark Ed)*. 2017;22:96–116. <https://doi.org/10.2741/4474>.
54. Johnson EL, Heaver SL, Waters JL, Kim BI, Bretin A, Goodman AL, et al. Sphingolipids produced by gut bacteria enter host metabolic pathways impacting ceramide levels. *Nat Commun*. 2020;11(1):2471. <https://doi.org/10.1038/s41467-020-16274-w>.
55. DiNicolantonio JJ, McCarty MF, O'Keefe JH. Role of dietary histidine in the prevention of obesity and metabolic syndrome. *Open Heart*. 2018;5(2):e000676. <https://doi.org/10.1136/openhrt-2017-000676>.
56. Molinaro A, Bel Lassen P, Henricsson M, Wu H, Adriouch S, Belda E, et al. Imidazole propionate is increased in diabetes and associated with dietary patterns and altered microbial ecology. *Nat Commun*. 2020;11(1):1–0.
57. Floyd JC, Fajans SS, Conn JW, Knopf RF, Rull J. Stimulation of insulin secretion by amino acids. *J Clin Invest*. 1966;45(9):1487–502. <https://doi.org/10.1172/JCI105456>.
58. Nishitani S, Takehana K, Fujitani S, Sonaka I. Branched-chain amino acids improve glucose metabolism in rats with liver cirrhosis. *Am J Physiol Gastrointest Liver Physiol*. 2005;288(6):G1292–300. <https://doi.org/10.1152/ajpgi.00510.2003>.
59. Connelly MA, Wolak-Dinsmore J, Dullaart RPF. Branched chain amino acids are associated with insulin resistance independent of leptin and adiponectin in subjects with varying degrees of glucose tolerance. *Metab Syndr Relat Disord*. 2017;15(4):183–6. <https://doi.org/10.1089/met.2016.0145>.
60. Lynch CJ, Adams SH. Branched-chain amino acids in metabolic signalling and insulin resistance. *Nat Rev Endocrinol*. 2014;10(12):723–36. <https://doi.org/10.1038/nrendo.2014.171>.
61. Tremblay F, Lavigne C, Jacques H, Marette A. Role of dietary proteins and amino acids in the pathogenesis of insulin resistance. *Annu Rev Nutr*. 2007;27(1):293–310. <https://doi.org/10.1146/annurev.nutr.25.050304.092545>.
62. Elsden SR, Hilton MG, Waller JM. The end products of the metabolism of aromatic amino acids by clostridia. *Arch Microbiol*. 1976;107(3):283–8. <https://doi.org/10.1007/BF00425340>.
63. Brial F, Le Lay A, Dumas M-E, Gauguier D. Implication of gut microbiota metabolites in cardiovascular and metabolic diseases. *Cell Mol Life Sci*. 2018;75(21):3977–90. <https://doi.org/10.1007/s00018-018-2901-1>.
64. de Mello VD, Paananen J, Lindström J, Lankinen MA, Shi L, Kuusisto J, et al. Indolepropionic acid and novel lipid metabolites are associated with a lower risk of type 2 diabetes in the Finnish Diabetes Prevention Study. *Sci Rep*. 2017;7(1):46337. <https://doi.org/10.1038/srep46337>.
65. Jump RLP, Polinkovsky A, Hurlless K, Sitzlar B, Eckart K, Tomas M, et al. Metabolomics analysis identifies intestinal microbiota-derived biomarkers of colonization resistance in clindamycin-treated mice. *PLoS One*. 2014;9(7):e101267. <https://doi.org/10.1371/journal.pone.0101267>.
66. Olson CA, Vuong HE, Yano JM, Liang QY, Nusbaum DJ, Hsiao EY. The gut microbiota mediates the anti-seizure effects of the ketogenic diet. *Cell*. 2018;173:1728–1741.e13.
67. Zhang H, Kovacs-Nolan J, Kodera T, Eto Y, Mine Y.  $\gamma$ -Glutamyl cysteine and  $\gamma$ -glutamyl valine inhibit TNF- $\alpha$  signaling in intestinal epithelial cells and reduce inflammation in a mouse model of colitis via allosteric activation of the calcium-sensing receptor. *Biochim Biophys Acta*. 2015;1852(5):792–804. <https://doi.org/10.1016/j.bbdis.2014.12.023>.
68. Geurts L, Lazarevic V, Derrien M, Everard A, Van Roye M, Knauf C, et al. Altered gut microbiota and endocannabinoid system tone in obese and diabetic leptin-resistant mice: impact on apelin regulation in adipose tissue. *Front Microbiol*. 2011;13:2–149.
69. Kim K-A, Gu W, Lee I-A, Joh E-H, Kim D-H. High fat diet-induced gut microbiota exacerbates inflammation and obesity in mice via the TLR4 signaling pathway. *PLoS One*. 2012;7(10):e47713. <https://doi.org/10.1371/journal.pone.0047713>.
70. Rauschert S, Uhl O, Koletzko B, Kirchberg F, Mori TA, Huang R-C, et al. Lipidomics reveals associations of phospholipids with obesity and insulin resistance in young adults. *J Clin Endocrinol Metab*. 2016;101(3):871–9. <https://doi.org/10.1210/jc.2015-3525>.
71. Son G, Kremer M, Hines IN. Contribution of gut bacteria to liver pathobiology. *Gastroenterol Res Pract*. 2010;2010:1–13. <https://doi.org/10.1155/2010/453563>.
72. Krautkramer KA, Reiter L, Denu JM, Dowell JA. Quantification of SAHA-dependent changes in histone modifications using data-independent acquisition mass spectrometry. *J Proteome Res*. 2015;14(8):3252–62. <https://doi.org/10.1021/acs.jproteome.5b00245>.
73. Gilep AA, Sushko TA, Usanov SA. At the crossroads of steroid hormone biosynthesis: the role, substrate specificity and evolutionary development of CYP17. *Biochim Biophys Acta*. 2011;1814(1):200–9. <https://doi.org/10.1016/j.bbapap.2010.06.021>.
74. Lindberg R, Burkhardt B, Ichikawa T. The structure and characterization of type I P-45015, gene as major steroid 15 $\alpha$ -hydroxylase and its comparison with type II P-4501 $\sim$  gene. *J Biol Chem*. 1989;264(11):6465–71. [https://doi.org/10.1016/S0021-9258\(18\)83371-3](https://doi.org/10.1016/S0021-9258(18)83371-3).
75. Singh V, Yeoh BS, Chassaing B, Xiao X, Saha P, Aguilera Olvera R, et al. Dysregulated microbial fermentation of soluble fiber induces cholestatic liver cancer. *Cell*. 2018;175:679–694.e22.
76. Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R, Gordon JL. The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci Transl Med*. 2009;1:6ra14.
77. Smits SA, Marcobal A, Higginbottom S, Sonnenburg JL, Kashyap PC. Individualized responses of gut microbiota to dietary intervention modeled in humanized mice. *mSystems*. 2016;1:e00098–16.
78. Shikany JM, Demmer RT, Johnson AJ, Fino NF, Meyer K, Ensrud KE, et al. Association of dietary patterns with the gut microbiota in older, community-dwelling men. *Am J Clin Nutr*. 2019;110(4):1003–14. <https://doi.org/10.1093/ajcn/nqz174>.

79. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature*. 2014;505(7484):559–63. <https://doi.org/10.1038/nature12820>.
80. Bailén M, Bressa C, Martínez-López S, González-Soltero R, Montalvo Lominchar MG, San Juan C, et al. Microbiota features associated with a high-fat/low-fiber diet in healthy adults. *Front Nutr*. 2020;7:583608. <https://doi.org/10.3389/fnut.2020.583608>.
81. Meddens SFW, de Vlaming R, Bowers P, et al. Genomic analysis of diet composition finds novel loci and associations with health and lifestyle. *Mol Psychiatry*. 2020. <https://doi.org/10.1038/s41380-020-0697-5>.
82. Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, et al. Personalized nutrition by prediction of glycemic responses. *Cell*. 2015;163(5):1079–94. <https://doi.org/10.1016/j.cell.2015.11.001>.
83. Rodriguez J, Hiel S, Neyrinck AM, Le Roy T, Pötgens SA, Leyrolle Q, et al. Discovery of the gut microbial signature driving the efficacy of prebiotic intervention in obese patients. *Gut*. 2020;69(11):1975–87. <https://doi.org/10.1136/gutjnl-2019-319726>.
84. Dao MC, Everard A, Aron-Wisniewsky J, Sokolovska N, Prifti E, Verger EO, et al. Akkermansia muciniphila and improved metabolic health during a dietary intervention in obesity: relationship with gut microbiome richness and ecology. *Gut*. 2016;65(3):426–36. <https://doi.org/10.1136/gutjnl-2014-3-08778>.
85. Reeves PG, Nielsen FH, Fahey GC. AIN-93 purified diets for laboratory rodents: final report of the American Institute of Nutrition Ad Hoc Writing Committee on the Reformulation of the AIN-76A Rodent Diet. *J Nutr*. 1993;123(11):1939–51. <https://doi.org/10.1093/jn/123.11.1939>.
86. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol*. 2019;37(8):852–7. <https://doi.org/10.1038/s41587-019-0209-9>.
87. Herd P, Carr D, Roan C. Cohort Profile: Wisconsin Longitudinal Study (WLS). *Int J Epidemiol*. 2014;43(1):34–41. <https://doi.org/10.1093/ije/dys194>.
88. Romano KA, Dill-McFarland KA, Kasahara K, Kerby RL, Vivas EI, Amador-Noguez D, et al. Fecal aliquot straw technique (FAST) allows for easy and reproducible subsampling: assessing interpersonal variation in trimethylamine-N-oxide (TMAO) accumulation. *Microbiome*. 2018;6(1):91. <https://doi.org/10.1186/s40168-018-0458-8>.
89. Rector RS, Thyfault JP, Morris RT, Laye MJ, Borengasser SJ, Booth FW, et al. Daily exercise increases hepatic fatty acid oxidation and prevents steatosis in Otsuka Long-Evans Tokushima Fatty rats. *Am J Physiol Gastrointest Liver Physiol*. 2008;294(3):G619–26. <https://doi.org/10.1152/ajpgi.00428.2007>.
90. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol*. 2013;79(17):5112–20. <https://doi.org/10.1128/AEM.01043-13>.
91. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016;13(7):581–3. <https://doi.org/10.1038/nmeth.3869>.
92. Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J*. 2017;11(12):2639–43. <https://doi.org/10.1038/ismej.2017.119>.
93. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J*. 2012;6(3):610–8. <https://doi.org/10.1038/ismej.2011.139>.
94. Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities. *AEM*. 2005;71(12):8228–35. <https://doi.org/10.1128/AEM.71.12.8228-8235.2005>.
95. Anderson MJ. Permutational multivariate analysis of variance (PERMANOVA). *Wiley StatsRef: Statistics Reference Online*. Chichester: Wiley; 2017. p. 1–15.
96. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139–40. <https://doi.org/10.1093/bioinformatics/btp616>.
97. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. 2010;11(3):R25. <https://doi.org/10.1186/gb-2010-11-3-r25>.
98. Krautkramer KA, Kreznar JH, Romano KA, Vivas EI, Barrett-Wilt GA, Rabaglia ME, et al. Diet-microbiota interactions mediate global epigenetic programming in multiple host tissues. *Mol Cell*. 2016;64(5):982–92. <https://doi.org/10.1016/j.molcel.2016.10.025>.
99. Maile TM, Izrael-Tomasevic A, Cheung T, Guler GD, Tindell C, Masselot A, et al. Mass spectrometric quantification of histone post-translational modifications by a hybrid chemical labeling method. *Mol Cell Proteomics*. 2015;14(4):1148–58. <https://doi.org/10.1074/mcp.O114.046573>.
100. Thomas SP, Haws SA, Borth LE, Denu JM. A practical guide for analysis of histone post-translational modifications by mass spectrometry: best practices and pitfalls. *Methods*. 2020;184:53–60. <https://doi.org/10.1016/j.ymeth.2019.12.001>.
101. Law CW, Chen Y, Shi W, Smyth GK. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014;15:R29.
102. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9(1):559. <https://doi.org/10.1186/1471-2105-9-559>.
103. McHardy IH, Goudarzi M, Tong M, Ruegger PM, Schwager E, Weger JR, et al. Integrative analysis of the microbiome and metabolome of the human intestinal mucosal surface reveals exquisite inter-relationships. *Microbiome*. 2013;1(1):17. <https://doi.org/10.1186/2049-2618-1-17>.
104. Tang Z-Z, Chen G, Hong Q, Huang S, Smith HM, Shah RD, et al. Multi-omic analysis of the microbiome and metabolome in healthy subjects reveals microbiome-dependent relationships between diet and metabolites. *Front Genet*. 2019;10:454. <https://doi.org/10.3389/fgene.2019.00454>.
105. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284–7. <https://doi.org/10.1089/omi.2011.0118>.
106. Lin W, Shi P, Feng R, Li H. Variable selection in regression with compositional covariates. *Biometrika*. 2014;101(4):785–97. <https://doi.org/10.1093/biomet/asu031>.
107. Meinshausen N, Bühlmann P. Stability selection: stability selection. *J R Stat Soc Series B Stat Methodol*. 2010;72:417–73.
108. Zhang J, Wei Z, Chen J. A distance-based approach for testing the mediation effect of the human microbiome. *Bioinformatics*. 2018;34(11):1875–83. <https://doi.org/10.1093/bioinformatics/bty014>.
109. Aitchison J. The statistical analysis of compositional data. *J R Stat Soc*. 1982;44:139–77.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

