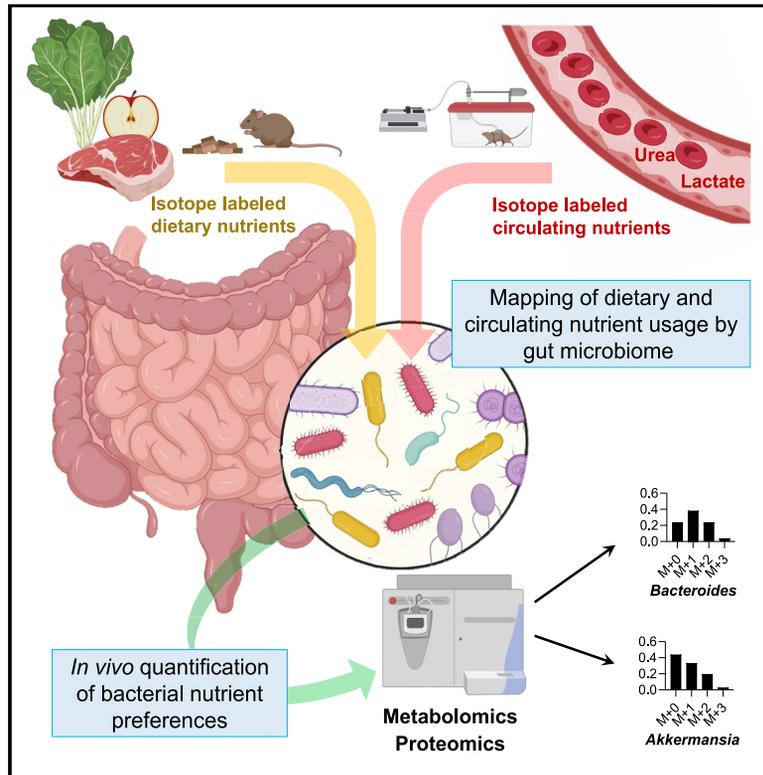


# Gut bacterial nutrient preferences quantified *in vivo*

## Graphical abstract



## Authors

Xianfeng Zeng, Xi Xing, Meera Gupta, ...,  
Martin Wühr, Cholsoon Jang,  
Joshua D. Rabinowitz

## Correspondence

wuhr@princeton.edu (M.W.),  
choljang@uci.edu (C.J.),  
josh@princeton.edu (J.D.R.)

## In brief

Isotope tracing into bacterial-specific protein sequences allows for a determination of nutrient preferences across gut microbes *in vivo*, and it reveals how diet alters microbiome composition.

## Highlights

- Gut microbiome feedstocks mapped by isotope tracing into bacteria-specific peptides
- Major contributors are dietary fiber and protein and host lactate, urea, and mucins
- Microbiome composition shifts toward bacteria that are fed their preferred nutrients
- Microbial metabolites' systemic levels reflect dietary precursors reaching microbiome



## Resource

# Gut bacterial nutrient preferences quantified *in vivo*

Xianfeng Zeng,<sup>1,4</sup> Xi Xing,<sup>1,4</sup> Meera Gupta,<sup>2,3,4</sup> Felix C. Keber,<sup>2,4</sup> Jaime G. Lopez,<sup>4</sup> Ying-Chiang J. Lee,<sup>2</sup> Asael Roichman,<sup>1,4</sup> Lin Wang,<sup>1,4,6</sup> Michael D. Neinst,<sup>1,4</sup> Mohamed S. Donia,<sup>2</sup> Martin Wühr,<sup>2,4,\*</sup> Cholsoon Jang,<sup>5,\*</sup> and Joshua D. Rabinowitz<sup>1,4,7,8,\*</sup>

<sup>1</sup>Department of Chemistry, Princeton University, Princeton, NJ 08544, USA

<sup>2</sup>Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA

<sup>3</sup>Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08544, USA

<sup>4</sup>Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA

<sup>5</sup>Department of Biological Chemistry, University of California, Irvine, Irvine, CA 92697, USA

<sup>6</sup>Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences & Peking Union Medical College, No. 5 Dong Dan San Tiao, Dongcheng District, Beijing 100005, China

<sup>7</sup>Ludwig Institute for Cancer Research, Princeton Branch, Princeton University, Princeton, NJ 08544, USA

<sup>8</sup>Lead contact

\*Correspondence: wuhr@princeton.edu (M.W.), choljang@uci.edu (C.J.), josh@princeton.edu (J.D.R.)

<https://doi.org/10.1016/j.cell.2022.07.020>

## SUMMARY

Great progress has been made in understanding gut microbiomes' products and their effects on health and disease. Less attention, however, has been given to the inputs that gut bacteria consume. Here, we quantitatively examine inputs and outputs of the mouse gut microbiome, using isotope tracing. The main input to microbial carbohydrate fermentation is dietary fiber and to branched-chain fatty acids and aromatic metabolites is dietary protein. In addition, circulating host lactate, 3-hydroxybutyrate, and urea (but not glucose or amino acids) feed the gut microbiome. To determine the nutrient preferences across bacteria, we traced into genus-specific bacterial protein sequences. We found systematic differences in nutrient use: most genera in the phylum Firmicutes prefer dietary protein, *Bacteroides* dietary fiber, and *Akkermansia* circulating host lactate. Such preferences correlate with microbiome composition changes in response to dietary modifications. Thus, diet shapes the microbiome by promoting the growth of bacteria that preferentially use the ingested nutrients.

## INTRODUCTION

The gut microbiome possesses an enormous diversity of enzymes, exceeding the number in mammals' genomes by more than 100-fold (Qin et al., 2010). This enzymatic capacity enables the processing of incoming dietary nutrients into a broad spectrum of microbial metabolites. Some of these reach the host circulation at substantial concentrations (Lai et al., 2021; Quinn et al., 2020). Microbial metabolites can play important roles in host pathophysiology. For example, short-chain fatty acids (SCFAs; acetate, propionate, and butyrate) (Dalile et al., 2019; Koh et al., 2016), trimethylamine *N*-oxide (Tang et al., 2013), secondary bile acids (Arab et al., 2017; Funabashi et al., 2020), indole-3-propionate (Wikoff et al., 2009), and imidazole propionate (Koh et al., 2018) affect immune maturation (Campbell et al., 2020; Hang et al., 2019), insulin sensitivity (Koh et al., 2018), cancer growth (Garrett, 2015; Yoshimoto et al., 2013), and cardiovascular disease (Nemet et al., 2020; Wang et al., 2011).

Both to replicate themselves and to release metabolic products, gut bacteria require nutrient inputs. These come in forms

including ingested food, host-synthesized gut mucus (Desai et al., 2016; Sicard et al., 2017), and host circulating metabolites (Scheiman et al., 2019). The availability of dietary nutrients to gut microbiota depends on the extent of host absorption: nutrients that are absorbed in the small intestine, like starch, are not available to the colonic microbiome. In contrast, nutrients that are poorly digested in the upper gastrointestinal tract, like fiber, can be key microbiome feedstocks (Lund et al., 2021; Wong and Jenkins, 2007).

Isotope tracing enables quantitative measurement of the inputs to metabolites and biomass. Studies employing radioactive tracers defined the basics of mammalian metabolism (Wolfe, 1984). Recent work has increasingly relied on stable isotope tracers coupled to mass spectrometry detection, which enables the measurement of labeling in specific downstream products (Fernández-García et al., 2020; McCabe and Previs, 2004). This approach has revealed fundamental features of host metabolism, such as circulating lactate being a major TCA fuel (Faubert et al., 2017; Hui et al., 2017). In addition, it has provided important insights into host-microbiome metabolic interplay.



For example, it revealed that dietary fructose is processed by the microbiome into acetate, which fuels hepatic lipogenesis (Jang et al., 2018; Zhao et al., 2020).

In principle, stable isotope tracing coupled to mass spectrometry can also be applied to determine the metabolic inputs to specific microbes, based on measuring labeling in bacteria-specific peptide sequences (Berry et al., 2015; Holmes et al., 2017; Oberbach et al., 2017; Reese et al., 2018; Zhang et al., 2016a, 2016b). By infusing nitrogen-labeled threonine to label host mucus, investigators were able to compare the contribution of dietary versus mucus protein to the gut microbiome and observed a shift toward more mucus contribution in mice that are fed a low-protein diet (Holmes et al., 2017).

Here, we perform a large-scale, quantitative assessment of the metabolic inputs to the gut microbiome and its products. We examine the contributions from dietary starch, fiber, and protein and the contribution from host mucus. We also examine most major circulating host nutrients, finding that lactate, 3-hydroxybutyrate, and urea stand out for passing from the host to the gut microbiome. Based on the measurement of bacteria-specific peptide sequences, we assess the nutrient preferences of different bacterial genera and show that these preferences align with microbiome composition changes in response to an altered diet.

## RESULTS

### Microbiome consumes less digestible dietary components

A major mechanism by which the microbiome may impact host physiology is via secreted metabolic products. We measured, in the portal and systemic circulation and the cecal contents, the absolute concentrations of more than 50 metabolites characterized in the literature as microbiome-derived (Campbell et al., 2020; De Vadder et al., 2014; Han et al., 2021; Hang et al., 2019; Koh et al., 2018; Mager et al., 2020; Ridlon et al., 2014; Wikoff et al., 2009) (Figures S1A and S1B; Tables 1 and S1). Most were elevated in the portal circulation relative to systemic blood, and all but two (inosine and *N*-acetyl-tryptophan, which are apparently mainly derived from the host) were depleted by antibiotics treatment.

The dominant excreted products on a molar basis (0.4–2 mM in the portal blood) are SCFAs. Other relatively abundant microbiome products (10–30  $\mu$ M) are aromatic amino acid fermentation products (phenol, indoxyl sulfate, and 3-phenylpropionate) and branched-chain fatty acids (valerate, isovalerate, 4-methylvalerate, isobutyrate, and 2-methylbutyrate). Primary bile acids, while present in the portal circulation at up to  $\sim$ 10  $\mu$ M concentration, are produced by the host and accordingly were not included in Table 1. Secondary bile acids, which are produced from primary bile acids by the microbiome, were lower in absolute concentration, with the most abundant being taurosoodeoxycholic acid (3  $\mu$ M in portal circulation).

To probe the dietary inputs to gut microbial products, we began by feeding mice via oral gavage, starch (readily digestible glucose polymer) and inulin (slowly digestible fructose polymer, i.e., soluble fiber) (Figure S1C). Following  $^{13}$ C-starch gavage, labeled glucose, lactate, and alanine quickly appeared in the

portal circulation and accounted for most starch carbons ( $\sim$ 75%) (Figures S1D–S1I) (Jang et al., 2018). In contrast, after  $^{13}$ C-inulin gavage, substantial labeled fructose, glucose, lactate, and alanine were not observed, and instead labeled portal metabolites slowly appeared in the form of SCFAs, with  $\sim$ 40% of inulin carbons becoming SCFAs and the remainder being undigested and excreted in the feces. Moreover, dietary inulin, but not starch, extensively labeled glycolytic and TCA intermediates and amino acids in the cecal content.

We next carried out similar experiments, comparing the gavage of a free amino acid mixture with algal protein, both uniformly  $^{13}$ C-labeled (Figure S1C). The free amino acids resulted in the rapid appearance of labeled amino acids in portal circulation, whereas the algal protein substantially labeled amino acids within the cecal contents (Figures S1J–S1M). Moreover, the algal protein copiously labeled microbiome-derived portal vein metabolites: SCFAs, branched-chain fatty acids, and aromatics (indole, indole-3-propionate, and 3-phenylpropionate). Thus, poorly digestible carbohydrates and protein feed the microbiome directly and the host indirectly via microbiome-derived products.

### Few circulating metabolites reach the microbiome

Next, we examined the possibility that nutrients in host circulation feed the gut microbiota. We infused deuterated water and eighteen major circulating nutrients ( $^{13}$ C-labeled) into the systemic circulation of pre-catheterized mice (Figure 1A). The infusion rates were selected to achieve modest but readily measurable labeling without substantially perturbing circulating concentrations. Circulating labeling reached a steady state by 2.5 h, at which time we collected serum and feces to quantitate the carbon contributions of each circulating nutrient to the corresponding fecal metabolites. Upon intravenous infusion of  $^{13}$ C-lactate, fecal lactate labeled rapidly (Figure 1B). Most infused circulating nutrients, however, did not penetrate the feces (Figures 1C and 1D). Indeed, while water fully exchanged with the feces, among abundant circulating carbon carriers, only lactate and 3-hydroxybutyrate penetrated. Glucose, amino acids, TCA intermediates, and fatty acids did not. Both lactate and 3-hydroxybutyrate are substrates of monocarboxylate transporters (MCTs), which are highly expressed in the colonic epithelium (Halestrap and Price, 1999, p. 1). Pharmacological MCT inhibition prevented lactate from penetrating the feces (Figure 1E). Thus, in contrast to most host circulating metabolites, which do not reach the colonic microbiome, MCTs render circulating lactate and 3-hydroxybutyrate accessible to gut microbes.

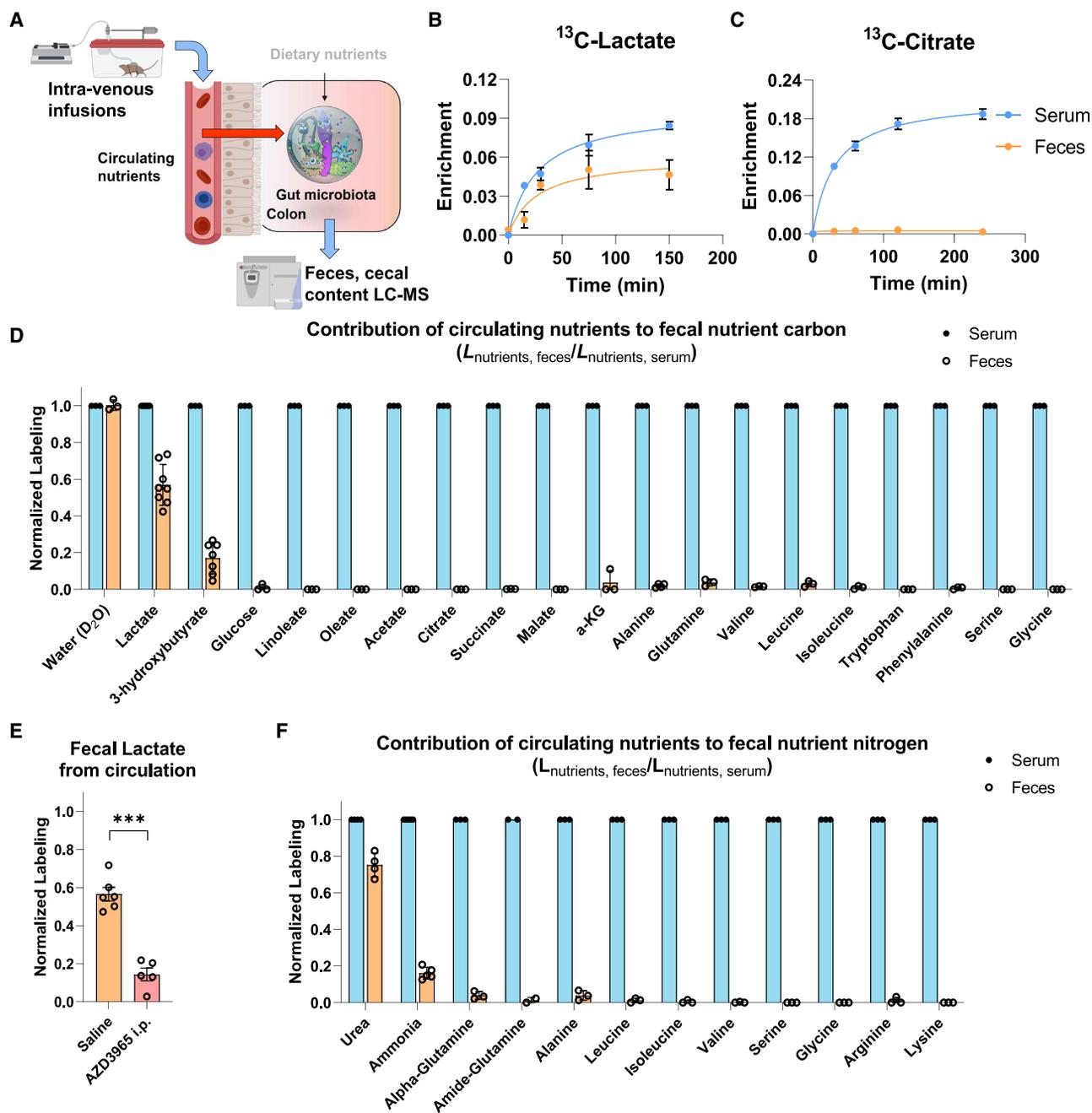
### Circulating urea is a microbiome nitrogen source

In addition to carbon, nitrogen is a fundamental constituent of all living cells. To assess nitrogen sources of the gut microbiome, we infused twelve abundant circulating nutrients in  $^{15}$ N-labeled form. Nitrogen from circulating urea and ammonia, but not amino acids, penetrates the feces and contributes to microbiome amino acids and ammonia (Figures 1F, S2A, and S2B). Urea usage by the microbiome involves its re-conversion to ammonia via the enzyme urease, which is expressed by a subset of gut microbes (Mora and Arioli, 2014; Ni et al., 2017), and gnotobiotic

**Table 1. Absolute concentrations and sources of microbiota-associated metabolites**

Metabolite	Systemic conc. (μM)	Portal/systemic (Log 2)	Abx/Conv (Log 2)	Source	Dietary fiber	Dietary protein	Circulating lactate	Source
Acetate	210.0	3.1	-1.9		41%	8%	7%	Fiber
Phenol	23.0	0.1	-4.3		0%	41%	0%	Protein
Indoxyl sulfate	22.0	-0.2	<-5.0		0%	13%	0%	Protein
Serotonin	12.0	-0.5	-1.4		0%	12%	0%	Protein
Propionate	9.9	>5.0	<-5.0		28%	4%	6%	Fiber
4-Methylvaleric acid	7.3	0.5	-0.1		6%	43%	2%	Protein
Indole-3-acetic acid	7.2	-0.3	-0.9		0%	23%	0%	Protein
Phenyl sulfate	6.6	-0.5	<-5.0		0%	42%	0%	Protein
Acetylglycine	5.4	0.2	-0.2		29%	22%	0%	Fiber
4-Ethylphenol	5.2	0.2	-0.9		0%	23%	0%	Protein
Indole-3-propionate	4.6	0.0	<-5.0		0%	26%	0%	Protein
Hippuric acid	4.2	0.3	-4.4		0%	33%	0%	Protein
4-Cresol	4.2	0.6	-0.4		0%	18%	0%	Protein
Indole-3-acetylaldehyde	4.1	-0.3	-0.2		0%	13%	0%	Protein
Trimethylamine-N-oxide	3.8	0.3	-3.8	Others	0%	0%	0%	Others
Butyrate	3.6	>5.0	<-5.0		54%	13%	7%	Fiber
Isobutyrate	2.8	2.8	-4.1		5%	35%	1%	Protein
2-Methylbutyrate	2.8	1.4	-3.0		5%	48%	1%	Protein
Catechol	2.8	0.0	<-5.0	Others	0%	0%	0%	Others
5-Aminovaleric acid	2.0	0.4	-0.3		0%	14%	0%	Protein
Indole-3-lactic acid	1.9	0.8	-0.7		0%	27%	0%	Protein
3-Hydroxycinnamic acid	1.6	-0.1	-0.6		0%	29%	0%	Protein
3-Phenylpropionate	1.5	2.4	-5.0		0%	22%	0%	Protein
Isovaleric acid	1.2	2.9	-3.7		1%	45%	1%	Protein
Benzoic acid	1.2	0.3	<-5.0		0%	32%	0%	Protein
Valerate	1.1	4.9	-3.8		6%	37%	5%	Protein
Cinnamoylglycine	1.1	0.7	-1.1		0%	39%	0%	Protein
Inosine	1.1	-3.3	0.4		12%	11%	0%	Fiber
Phenylacetic acid	0.79	0.6	-2.8		0%	50%	0%	Protein
4-Hydroxyphenylpropionate	0.76	0.2	-2.0		0%	56%	0%	Protein
Propionylglycine	0.75	1.2	-0.1		21%	24%	0%	Fiber
Phenylpropionylglycine	0.75	0.6	-0.3		2%	22%	0%	Protein
Heptanoic acid	0.74	0.6	-0.3		6%	11%	0%	Protein
Indole	0.53	0.2	-2.6		0%	49%	0%	Protein
Butyrylglycine	0.39	1.1	-1.4		37%	28%	0%	Fiber
4-Hydroxyphenylacetate	0.37	0.7	-1.0		0%	32%	0%	Protein
N-Acetyl-Tryptophan	0.37	0.1	0.5		0%	13%	0%	Protein
Tauroursodeoxycholic acid	0.33	2.9	-2.2		0%	0%	0%	Bile acids
4-Hydroxybenzoic acid	0.25	0.7	-1.0		0%	9%	0%	Protein
Cinnamic acid	0.20	2.5	-2.6		0%	41%	0%	Protein
Taurodeoxycholic acid	0.11	4.8	-3.1		0%	0%	0%	Bile acids
2-Hydroxyhippuric acid	0.062	1.5	-0.7		0%	22%	0%	Protein
Imidazole propionate	0.051	0.7	-0.7		0%	37%	0%	Protein
Deoxycholic acid	0.048	3.1	<-5.0		0%	0%	0%	Bile acids
Indole-3-ethanol	0.032	1.6	-0.3		0%	38%	0%	Protein
Ursodeoxycholic acid	0.026	2.5	<-5.0		0%	0%	0%	Bile acids
4-Cresol sulfate	0.026	-0.2	-2.8		0%	26%	0%	Protein
Phenylacetylglucosylglycine	0.017	0.2	-0.3		0%	32%	0%	Protein
Equol	0.013	3.7	<-5.0	Others	0%	0%	0%	Others
Hyodeoxycholic acid	0.009	3.0	-3.7		0%	0%	0%	Bile acids
Lithocholic acid	0.008	1.8	<-5.0		0%	0%	0%	Bile acids
Taroulithocholic acid	<0.001	>5.0	<-5.0		0%	0%	0%	Bile acids
Isoallothocholic acid	<0.001	>5.0	<-5.0		0%	0%	0%	Bile acids
Glycoursodeoxycholic acid	<0.001	>5.0	-4.5		0%	0%	0%	Bile acids
Glycodeoxycholic acid	<0.001	>5.0	-1.3		0%	0%	0%	Bile acids

Data are from *ad lib* fed state (ZT0); for *ad lib* fasted state (ZT12), see [Table S1](#). Absolute concentration is mean, n = 5 mice. Portal/systemic = fold change in concentration between the portal vein and tail vein (median, n = 5 mice). Abx/Conv refers to fold change in portal blood concentration between microbiome-depleted (antibiotics-treated) versus untreated (conventional) mice (median, n = 5 mice/group). Source bar indicates the relative contribution to the indicated metabolite from dietary inulin, algal protein, and circulating lactate (based on isotope tracing); bile acids are shown in gray. Percentages indicate quantitative relative contributions from those nutrients (median, n = 4). Numbers typically add up to less than 100%, as other sources (e.g., mucins) contribute. See also [Figures S1 and S4](#) and [Table S1](#).



**Figure 1. Circulating lactate, 3-hydroxybutyrate, and urea feed the gut microbiome**

(A) Schematic of intravenous infusion of isotope-labeled nutrients to identify circulating metabolites that feed gut microbiome.

(B) Circulating lactate rapidly enters the feces. Mice were infused with  $^{13}\text{C}$ -lactate, and serum and fresh feces enrichment was compared. Mean  $\pm$  SE,  $n = 3$ .

(C) Circulating citrate does not enter the feces. As in (B), for  $^{13}\text{C}$ -citrate. Mean  $\pm$  SE,  $n = 3$ .

(D) Passage of circulating  $^{13}\text{C}$ -labeled nutrients into the feces. Mice were infused with labeled nutrients for 2.5 h, and labeling fraction in feces was normalized to labeling fraction in serum. Blue, serum labeling; orange, fecal labeling. Mean  $\pm$  SE,  $n = 3$  except for lactate ( $n = 8$ ) and 3-hydroxybutyrate ( $n = 7$ ). a-KG,  $\alpha$ -ketoglutarate.

(E) Pharmacological inhibition of MCT1 transporter decreases the passage of circulating lactate to feces. Mice were injected i.p. with saline or 100 mg/kg AZD3965, and fresh feces lactate enrichment was measured. Mean  $\pm$  SE,  $n = 6$  for saline and  $n = 5$  for AZD3965. \*\*\*  $p < 0.001$  by two-sided Student's  $t$  test.

(F) Passage of circulating  $^{15}\text{N}$ -labeled nutrients into the feces. As in (D), for  $^{15}\text{N}$ -labeling. Mean  $\pm$  SE,  $n = 3$  except for urea ( $n = 4$ ) and ammonia ( $n = 5$ ).

See also [Figure S2](#).

mice that colonized intentionally with only urease-negative bacteria showed no labeling from urea (Figures S2C and S2D).

Urea, which is made from ammonia in the liver, was a quantitatively greater source of microbiome nitrogen than ammonia. Moreover, urea, but not ammonia, was more abundant in the host circulation than cecal lumen, consistent with only urea being able to passively flow into the gut lumen (Figures S2E and S2F). We hypothesized that circulating host ammonia might be feeding the microbiome, mainly indirectly, after being converted by the host liver into circulating urea (Figure S2G) (Bartman et al., 2021). This indirect contribution was calculated by multiplying circulating urea's contribution to fecal amino acids ( $L_{AAs \leftarrow urea}$ ) by the circulating urea fraction that comes from circulating ammonia ( $L_{urea \leftarrow NH_3} = 33\%$ ). It fully explained the observed microbiome labeling from circulating ammonia (Figure S2H). Further supporting the indirect pathway, antibiotics treatment blocked both circulating urea and ammonia from becoming cecal ammonia (Figures S2I and S2J), which makes sense if flux of ammonia into the cecal contents goes through host urea and microbial urease (Figure S2K).

### Microbiota synthesize amino acids from fiber and urea

To determine quantitatively the sources of microbiome metabolites, we measured their labeling after *ad libitum* feeding of isotopically enriched food. To this end, we fed mice standard chow with a portion of the fiber, fat, or protein  $^{13}\text{C}$ -labeled, with cecal labeling reaching steady state within 12 h (Figure S3A). To account for circulating nutrient inputs, we also infused  $^{13}\text{C}$ -lactate or 3-hydroxybutyrate (Figure 2A). These studies identified a majority of the carbon feeding into most microbiome central metabolites, with glycolytic and pentose phosphate metabolites labeling almost exclusively coming from dietary fiber (inulin), whereas pyruvate and TCA metabolites are also labeled from dietary protein and circulating lactate (Figures 2B and S4A).

We next examined inputs to microbiome free amino acids, tracing also with  $^{15}\text{N}$ -labeled dietary protein and infused urea. Unlike mammals, most gut bacteria have the biosynthetic capacity to make all 20 proteogenic amino acids. Nevertheless, we observed that "essential amino acids", which cannot be made by mammals and require the expression of extensive biosynthetic pathways in bacteria, are derived mainly from dietary proteins (Figure 2C). In contrast, "non-essential amino acids" are primarily synthesized within the gut microbiome, using dietary inulin and circulating lactate as carbon sources. Microbiota depletion with antibiotics or in germ-free mice favored cecal accumulation of those amino acids coming (based on our isotope-tracing studies) largely from dietary protein and depletion of those being synthesized by the microbes (Figures S3B–S3G).

Dietary protein was the main nitrogen source for both essential and non-essential amino acids, with host urea also contributing substantially to the non-essential amino acids (Figure 2D). Dietary protein provides nitrogen to cecal amino acids mainly directly, not through circulating urea (Figures S3H and S3I). Consistent with the gut microbiome synthesizing amino acids from fiber carbon and urea nitrogen, across amino acids, urea's

nitrogen contribution correlated with inulin's carbon contribution (Figure 2E).

Amino acid labeling from inulin was typically partial (i.e., one or a few of the amino acid's carbon atoms were labeled), reflecting inulin's carbons being scrambled with other inputs into central metabolism (Figure S3J). In contrast, labeling from  $^{13}\text{C}$ ,  $^{15}\text{N}$ -proteins was typically complete (or complete except for the nitrogen label; Figure S3K), indicating direct usage of intact amino acids after proteolysis (sometimes after a cycle of deamination and re-amination). Consistent with such re-amination, the combination of  $^{15}\text{N}$ -urea infusion and  $^{13}\text{C}$ -protein feeding produced some double-labeled ( $^{13}\text{C}$ ,  $^{15}\text{N}$ -labeled) amino acids (Figure S3L).

Lastly, the amino acids synthesized by the microbiome stay in the microbiome: we do not observe discernible labeling of these amino acids in the host (Figure S3M). Taken together, we found the following: (1) essential amino acids, although capable of being synthesized by the microbiome, come mainly from the diet and do not go through any carbon rearrangements; (2) the most closely TCA-linked non-essential amino acids are substantially synthesized by the microbiome, using carbon from fiber scrambled with other carbon via central metabolic reactions; and (3) transamination reactions partially mix nitrogen from diet-derived amino acids with nitrogen from host urea.

### Diverse microbiome products come from dietary protein

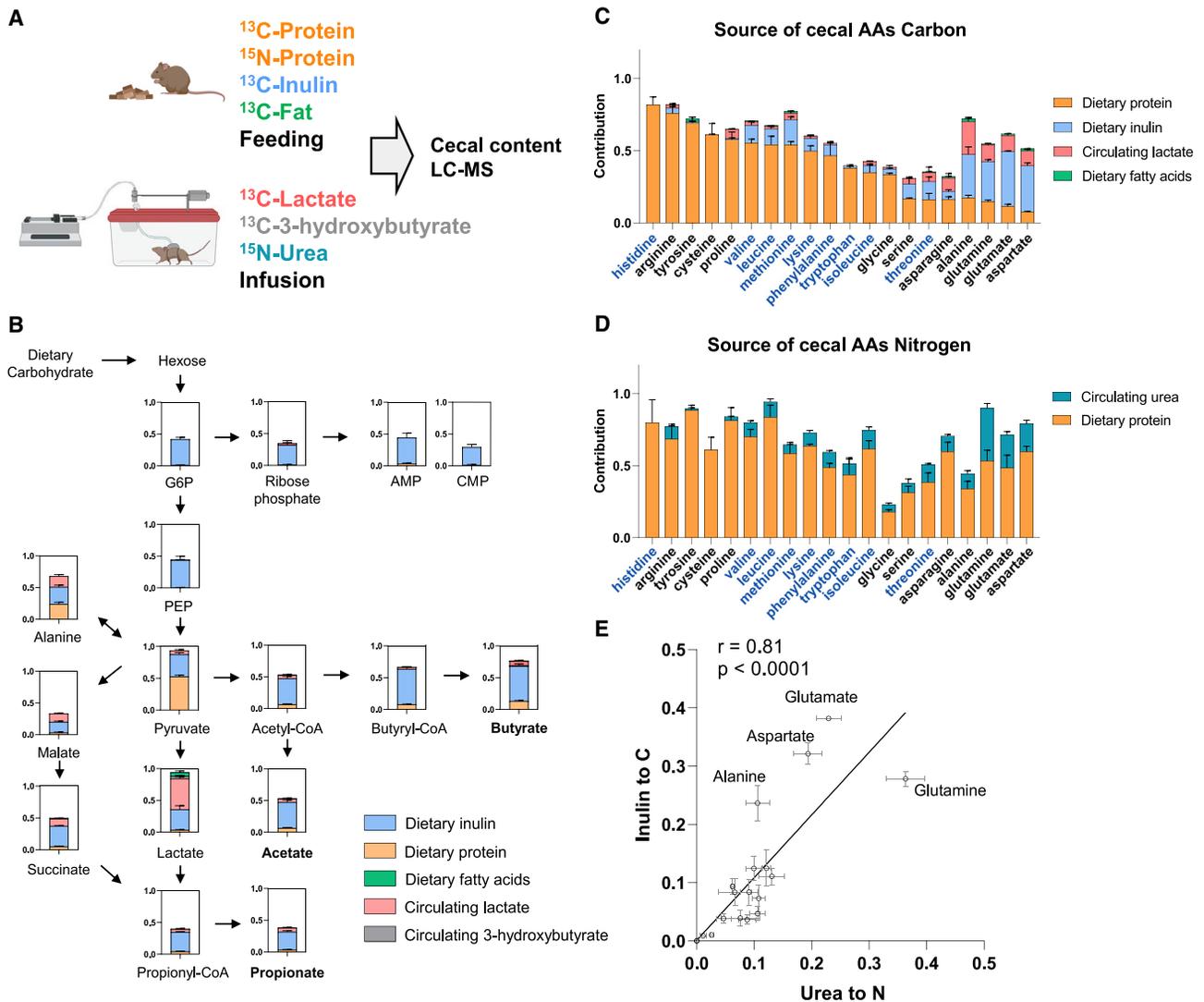
We next examined the carbon inputs to the other major microbiome products, especially the ones excreted into the portal circulation (Table 1). As expected, SCFAs, the most abundant microbial metabolites, come mainly from dietary fiber. Many less abundant ones, however, are mainly derived from dietary protein.

In addition to classical microbiome products, we also observed metabolites that are made in a collaborative manner, with the host carrying out the final synthesis using microbiome-derived inputs. For example, a wide range of microbiome-derived carboxylic acids are conjugated to glycine in the liver and kidneys to make different acyl-glycines (Figures S4B–S4E) (Wikoff et al., 2009).

We also examined the host clearance mechanisms of microbiome metabolites, based on arterial-venous gradients across the liver and kidney and levels in the urine. SCFAs and branched-chain fatty acids were avidly consumed by the liver. Most microbiome-derived metabolites were excreted by the kidney into the urine, with the notable exception of SCFAs, which are actively reabsorbed (Table S1A) (Jang et al., 2019; Ullrich et al., 1982). Thus, we establish dietary protein as a major precursor to many microbiome metabolites and identify host-microbiome interplay in the metabolism of SCFAs, including their renal reabsorption and use by the liver and kidney for the synthesis of acyl-glycines.

### Circulating levels of microbiota metabolites are controlled by protein digestibility

We found that many microbiome-derived metabolites are derived from unabsorbed dietary protein that reaches the colon. We hypothesized that the circulating levels of such metabolites would depend on the extent of dietary protein reaching the



**Figure 2. Quantitative analysis of dietary and circulating nutrient contributions to gut microbiome**

(A) Experimental design. Mice were fed chow containing  $^{13}\text{C}$ -protein,  $^{13}\text{C}$ -inulin,  $^{13}\text{C}$ -fatty acids, or  $^{15}\text{N}$ -protein for 24 h. Alternatively, mice were intravenously infused with  $^{13}\text{C}$ -lactate,  $^{13}\text{C}$ -3-hydroxybutyrate, or  $^{15}\text{N}$ -urea for 24 h. The labeling of cecal-content metabolites was analyzed by LC-MS.

(B) Contribution of dietary and circulating nutrients to carbohydrate fermentation pathways in gut microbiome. Mean  $\pm$  SE, n = 4. G6P, glucose-6-phosphate; PEP, phosphoenolpyruvate.

(C) Contribution of dietary and circulating nutrients to cecal amino acid carbon. The names of essential amino acids (EAA) are written in blue and non-essential amino acids (NEAA) in black. Mean  $\pm$  SE, n = 4.

(D) Contribution of dietary and circulating nutrients to cecal amino acid nitrogen. As in (C), for nitrogen.

(E) Positive correlation, across amino acids in the cecal contents, of carbon contribution from dietary inulin and nitrogen contribution from circulating urea. Mean  $\pm$  SE, n = 4.

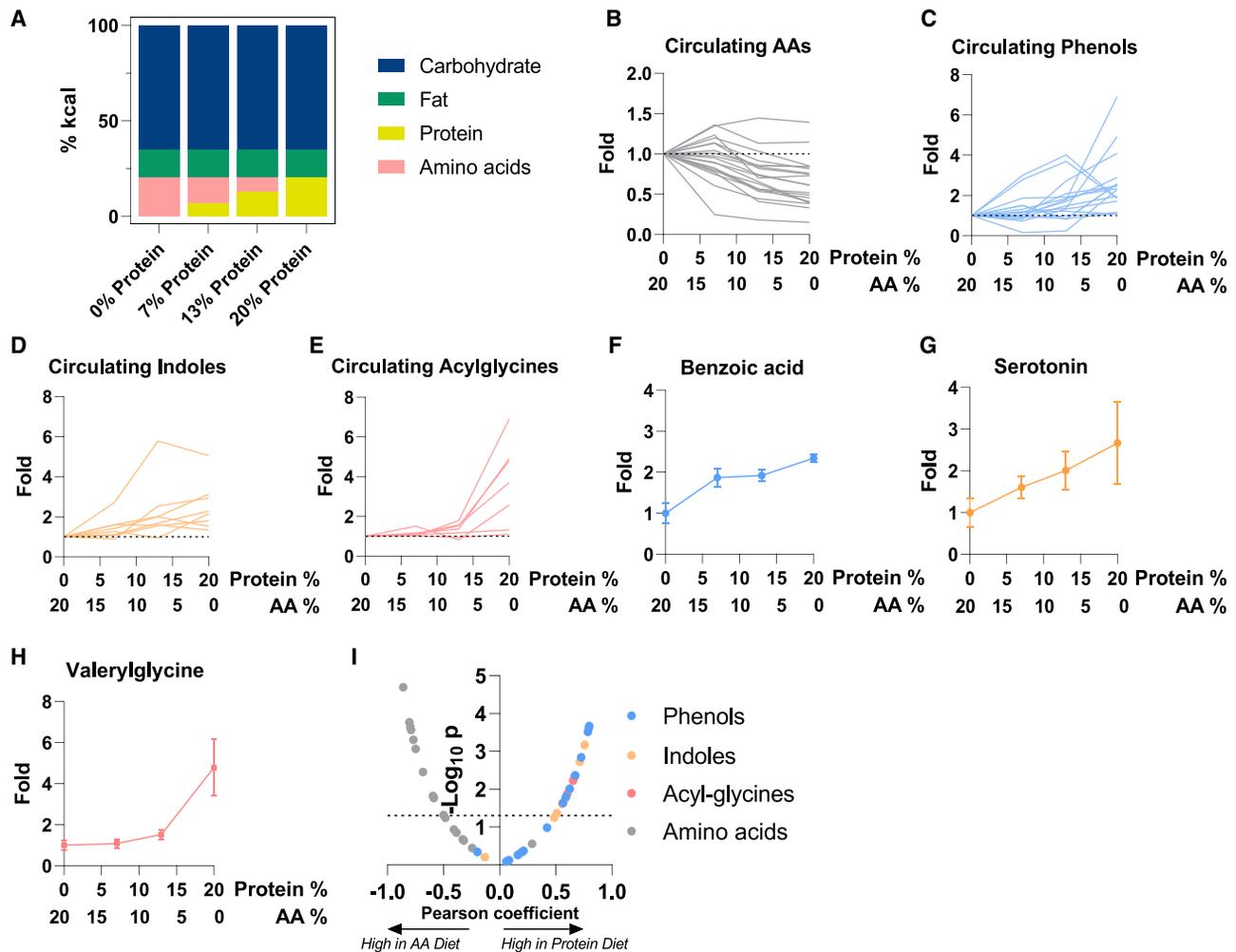
See also [Figures S3](#) and [S4](#).

colonic microbiome. To manipulate this, we fed mice diets in which a portion of the protein (casein, which in part reaches the colonic microbiome) was replaced with free amino acids (which are essentially fully absorbed in the small intestine) ([Figure 3A](#)). After 2 weeks, we performed metabolomics on the systemic blood. As expected, diets with less intact protein and more free amino acids tended to increase circulating amino acid levels ([Figure 3B](#)). Importantly, protein-derived circulating microbial

metabolites (phenols, indoles, and acyl-glycines) fell in tandem ([Figures 3C–3I](#)). Thus, knowledge of the nutrient sources of microbiome metabolites can be applied to manipulate their systemic levels.

**Gut bacterial growth is synchronized with host feeding**

Thus far, we have reported inputs and outputs of the gut microbiome as a whole. We now shift to examining the growth and



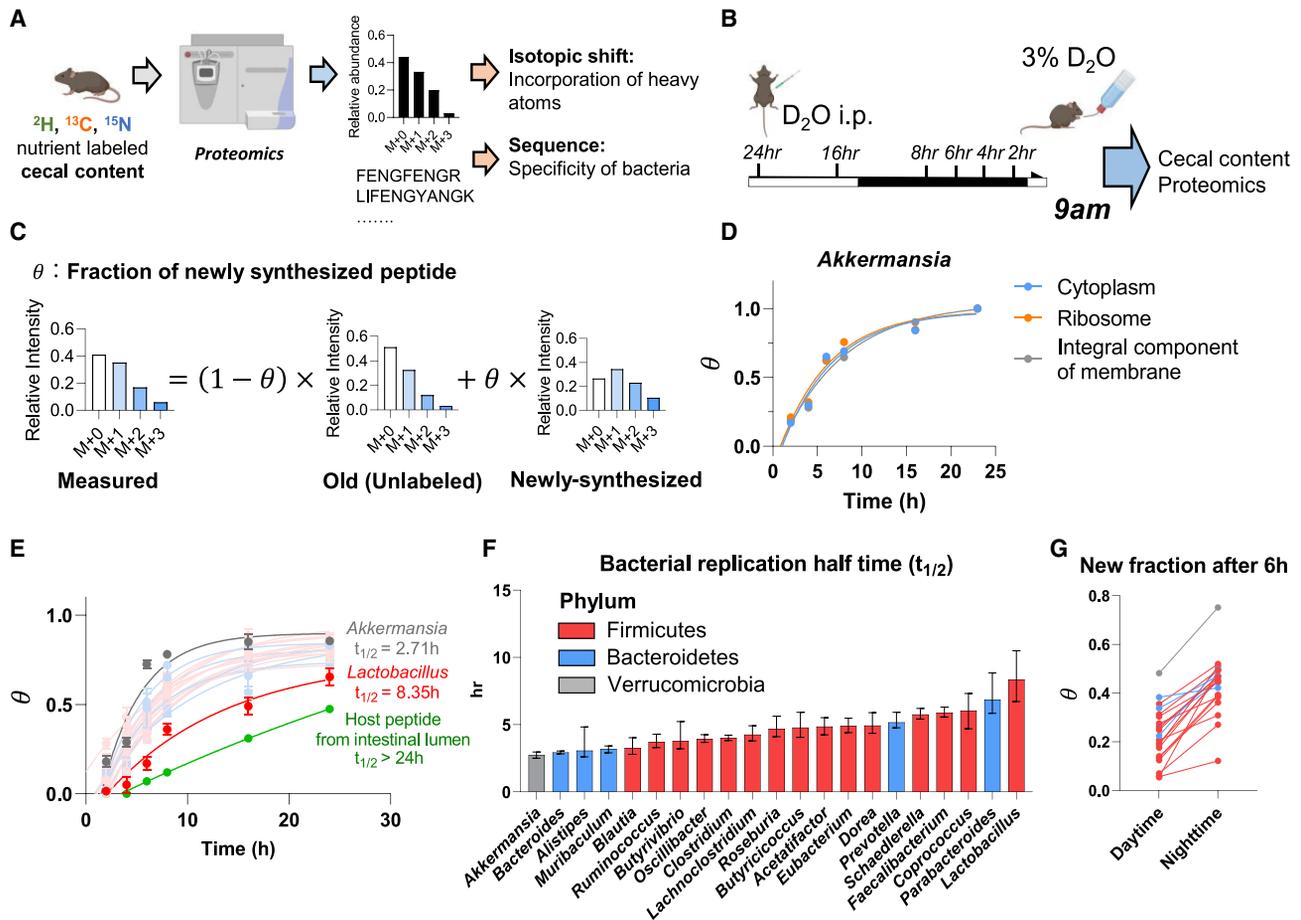
**Figure 3. Circulating levels of microbiota metabolites depend on protein reaching the microbiome**

(A) Compositions of diets used in the figure. "Protein" is casein. "Amino acids" are composition-matched free amino acids.  
 (B) Concentration of circulating amino acids in systemic circulation after 2 weeks test diet relative to free amino acids diet. Serum was taken at *ad lib* fed state. Each metabolite is a line. Mean,  $n = 4$  mice.  
 (C) As in (B), for phenols. Mean,  $n = 4$  mice.  
 (D) As in (B), for indoles. Mean,  $n = 4$  mice.  
 (E) As in (B), for acyl-glycines. Mean,  $n = 4$  mice.  
 (F) As in (B) for benzoic acid. Mean  $\pm$  SE,  $n = 4$  mice.  
 (G) As in (F), for serotonin. Mean  $\pm$  SE,  $n = 4$  mice.  
 (H) As in (F), for valerylglycine. Mean  $\pm$  SE,  $n = 4$  mice.  
 (I) Correlation between dietary protein (as opposed to free amino acid) fraction in diet and metabolite abundances (relative to amino acid diet). The volcano plot shows Pearson coefficient and p value of correlation between metabolite levels to casein abundance in diet.

metabolism of specific bacterial genera. To this end, we deployed proteomics to measure gut microbial peptides and their labeling, focusing on peptide sequences specific to a single bacterial genus (Figure 4A).

To quantify protein synthesis in different gut microbial genera, we used deuterated water ( $D_2O$ ) tracing (Holmes et al., 2015; O'Brien et al., 2020). To achieve steady-state labeling of body water, we gave mice  $D_2O$  by bolus injection followed by mixing it into drinking water. Peptide labeling in the cecal contents was then measured by proteomics (Figure 4B).

A key technical challenge in using proteomics to read out metabolic activity is the complexity, arising from natural isotope abundances, of peptide mass spectra. We used liquid chromatography-high-resolution mass spectrometry to obtain the full scan (MS1) mass isotope distribution for each peptide of interest, with MS/MS analysis of the unlabeled form used to determine the peptide's identity. We then calculated, based on the mass isotope distribution, the fraction of peptide that was newly synthesized ( $\theta$ ). To this end, first, we calculated the mass isotope distribution of unlabeled peptides based on natural isotope



**Figure 4. Growth rate of different gut bacterial genera quantified by isotope tracing**

(A) Experimental approach for isotope tracing into specific gut bacteria. Only peptides that are specific to a particular bacterial genus were examined.

(B) Growth rate quantification using  $\text{D}_2\text{O}$ . Mice received  $\text{D}_2\text{O}$  by i.p. injection followed by  $\text{D}_2\text{O}$  drinking water, and cecal-content labeling was measured over time by proteomics and metabolomics. Mice were fed *ad lib*; tissues were harvested at 9 a.m.

(C) Calculation of newly synthesized peptide fraction ( $\theta$ ). The experimentally observed peptide mass isotope distribution was fit to a linear combination of unlabeled peptide (“old,” heavy forms from natural isotope abundance) and newly synthesized peptide (“new,” heavy forms from isotope labeling pattern of free cecal amino acids and from natural isotope abundance).

(D) Different cellular compartments from the same bacterial genus show similar labeling rate. Mean,  $n = 5$  mice for each time point.

(E) Genus-specific growth rates were determined by a single exponential fitting, as a function of time, of  $\theta$  (mean across both different peptides measured from that genus and replicate mice). Mean  $\pm$  SE,  $n = 5$  mice for each time point.

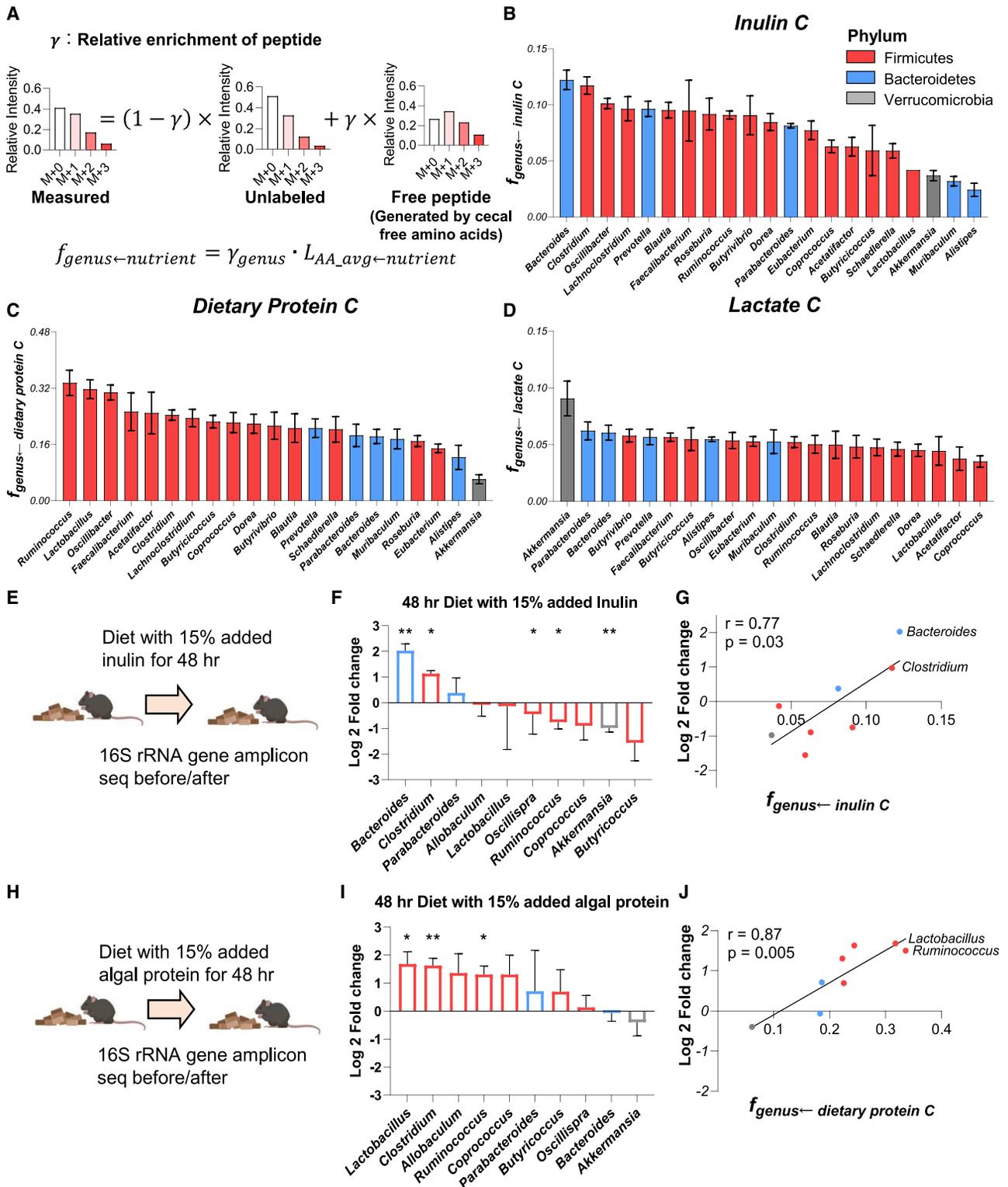
(F) Bacterial replication half-time of different gut bacteria. Data are exponential fits  $\pm$  SE

(G) The gut bacteria synthesize protein in sync with the physiological feeding patterns of the host. The figure shows the average newly synthesized peptide fraction ( $\theta$ ) for different gut bacterial genera after  $\text{D}_2\text{O}$  labeling during daytime versus nighttime. Each line connects the daytime and nighttime measurements for one genus. Mean,  $n = 10$  mice for daytime and for nighttime.

See also [Figure S5](#).

abundances (“old”). Second, we calculated the expected mass isotope distribution of a newly synthesized peptide generated from cecal free amino acids, whose labeling we experimentally measured by metabolomics. Then, we determined the fraction of newly synthesized ( $\theta$ ) by linear interpolation between the “old” and “newly synthesized” spectra (Figure 4C). To verify this approach *in vitro*, we cultured *Clostridium sporogenes* and *Bacteroides dorei* in media enriched with  $\text{D}_2\text{O}$  and measured growth rate as is typically done (based on  $\text{OD}_{600}$ ) and as above (using media in place of cecal amino acid labeling), finding good agreement (Figures S5A–S5C).

We then measured the newly synthesized fraction ( $\theta$ ) for a minimum of 5 peptides for each bacterial genus *in vivo*, with abundant gut bacteria yielding  $\theta$  for over 100 characteristic peptides. Irrespective of their intracellular location, different peptides from the same bacterial genus tended to label at a similar rate (Figures 4D, S5D, and S5E). Labeling rate varied across bacterial genera, with a half doubling time ranging from 2.5 h for *Akkermansia* to 8 h for *Lactobacillus*, which still markedly exceeded the labeling rate of host intestinal proteins (>24-h half doubling time) (Figures 4E, 4F, and S5F).



**Figure 5. Preferred carbon sources differ across gut bacteria**

(A) Calculation of peptide's relative  $^{13}\text{C}$ -enrichment ( $\gamma$ ) and carbon contribution from the tracer to a bacterial genus ( $f_{genus \leftarrow nutrient}$ ). First, the experimentally observed peptide mass isotope distribution was fit to a linear combination of an unlabeled peptide (heavy forms from natural isotope abundance) and a peptide made from free cecal amino acids (heavy forms from isotope labeling pattern of free cecal amino acids and from natural isotope abundance), yielding  $\gamma$ . Then,  $f_{genus \leftarrow nutrient}$  was determined by correcting for the fractional contribution of that tracer to the cecal free amino acid pools.

(legend continued on next page)

Our prior analyses revealed that the microbiome is fed substantially by dietary components. Accordingly, we hypothesized that microbial growth synchronizes with physiological feeding, which in mice occurs mainly during the nighttime. To assess the diurnal rhythm of gut bacterial protein synthesis, mice were given D<sub>2</sub>O for 6-h intervals throughout the diurnal cycle, followed by proteomic analysis of their cecal contents. Every measured bacterial genus showed greater protein synthesis during nighttime than daytime (Figure 4G). Thus, gut bacteria grow in sync with the physiological feeding patterns of the host.

### Preferred carbon sources differ across gut bacteria

Next, we quantitated the carbon feedstocks of different microbes by combining <sup>13</sup>C-nutrient labeling and proteomics. Each <sup>13</sup>C-labeled nutrient (dietary inulin, dietary algal protein, or circulating lactate) was provided for 24 h, which is sufficient to achieve steady-state labeling in the gut bacteria. Our analysis strategy involved two steps: first, we calculated, based on each genus-specific peptide's observed mass isotope distribution, its relative <sup>13</sup>C-enrichment ( $\gamma$ ) compared with that of cecal free amino acids (Figure 5A). Mathematically, this calculation is identical to the calculation of  $\theta$  in the D<sub>2</sub>O case, except here, the tracer is a particular <sup>13</sup>C-labeled nutrient, which unlike D<sub>2</sub>O is used preferentially by certain bacterial genera. The observed peptide's relative <sup>13</sup>C-enrichment multiplied by the average contribution of that <sup>13</sup>C-tracer to the gut microbial amino acid pool ( $L_{AA\_avg} \leftarrow nutrient$ ) gives a quantitative measure of the tracer's contribution to the observed genus-specific peptide. Averaging across such peptides gives a fractional contribution of the <sup>13</sup>C-labeled nutrient to protein synthesis in a bacterial genus.

Using this method, we measured feedstocks of the bacterial genera that were detected in every proteomics experiment. We were also able to make species-specific measurements in some cases (Figures S6A–S6F). We observed marked differences in nutrient preferences across microbiota. For example, *Bacteroides* and *Clostridium* use over 4-fold more inulin than *Akkermansia*, *Muribaculum*, or *Alistipes* (Figures 5B and S6A). Overall, bacteria from the phylum Firmicutes used more dietary protein than Bacteroidetes (Firmicutes  $0.237 \pm 0.052$ ; Bacteroidetes  $0.175 \pm 0.031$ ,  $p = 0.02$ ). *Akkermansia*, which is generally considered a health-promoting gut microbe, used among the least dietary inulin and protein (Figures 5B, 5C, S6A, and S6B). In contrast, it used by far the most circulating lactate from the host (Figures 5D and S6C).

We were curious whether these bacterial nutrient preferences predict microbiome composition changes upon dietary changes. To explore this possibility, we fed mice an inulin-enriched or algal protein-enriched diet for 2 days and measured microbiome

composition by 16S rRNA gene amplicon sequencing. *Bacteroides*, the top consumer of <sup>13</sup>C-inulin, increased by 4-fold after a high-inulin diet (Figures 5E–5G). *Clostridium*, another high-inulin consumer, also increased by 2-fold. Other genera that use less inulin carbon were either unchanged or slightly decreased. A similar consistency between microbes' nutrient preference and abundance changes was observed in mice fed the algal protein-enriched diet (Figures 5H–5J). Carbon-source preference measured by proteomics ( $f_{genus \leftarrow nutrient}$ ) correlates with abundance change, following a diet shift measured by 16S rRNA gene amplicon sequencing, for both the inulin and algal protein conditions (Figures 5G and 5J). Thus, the nutrient preferences of different gut bacteria help explain microbiome compositional changes following dietary manipulations (David et al., 2014).

### Firmicutes consume dietary protein, while Bacteroidetes consume secreted host protein

Lastly, we turned to the nitrogen-source preferences of different gut bacteria, comparing <sup>15</sup>N-labeled dietary protein feeding with <sup>15</sup>N-urea infusion. The analytical approach was identical to that employed above for carbon-source preferences. Bacterial genera that highly use carbon from dietary protein also highly use nitrogen from dietary protein, consistent with amino acids from dietary protein being assimilated intact in bacterial proteomes (Figures 6A, S6D, and S6G).

Conversely, among members of the phylum Firmicutes, genera preferring urea nitrogen tended to be avid inulin users, i.e., to synthesize their own amino acids using inulin and urea (Figures 6B, S6E, and S6H). This includes some urease-negative genera, which presumably acquire urea nitrogen via cross-feeding. Moreover, again among Firmicutes, we also saw the expected trade-off where some genera prefer nitrogen from dietary protein and others from circulating urea (Figure S6I). Following intravenous urea infusions to raise circulating urea concentrations, abundance of those Firmicutes preferring urea, along with *Akkermansia*, increased substantially (Figures 6C–6F).

Compared with Firmicutes, the lower use of both dietary protein and circulating urea nitrogen by Bacteroidetes raised a key question: how do Bacteroidetes get nitrogen? Some members of gut microbiome (e.g., *Bacteroides* and *Akkermansia*) are capable of digesting host-secreted proteins such as mucins (Berry et al., 2013; Reese et al., 2018). We hypothesized that host-secreted proteins are a key source of Bacteroidetes nitrogen. To probe this possibility, we performed long-term <sup>15</sup>N-labeled lysine and arginine infusions (12, 18, and 36 h) to label host proteins in the colon (Figures 6G and S7A–S7E). Despite not directly feeding the microbiome (Figures 1F and S7E), lysine

(B) Carbon contribution of dietary inulin across bacterial genera. Mean  $\pm$  SE,  $n = 4$  mice.

(C) Carbon contribution of dietary algal protein across bacterial genera. Mean  $\pm$  SE,  $n = 6$  mice.

(D) Carbon contribution of circulating lactate across bacterial genera. Mean  $\pm$  SE,  $n = 7$  mice.

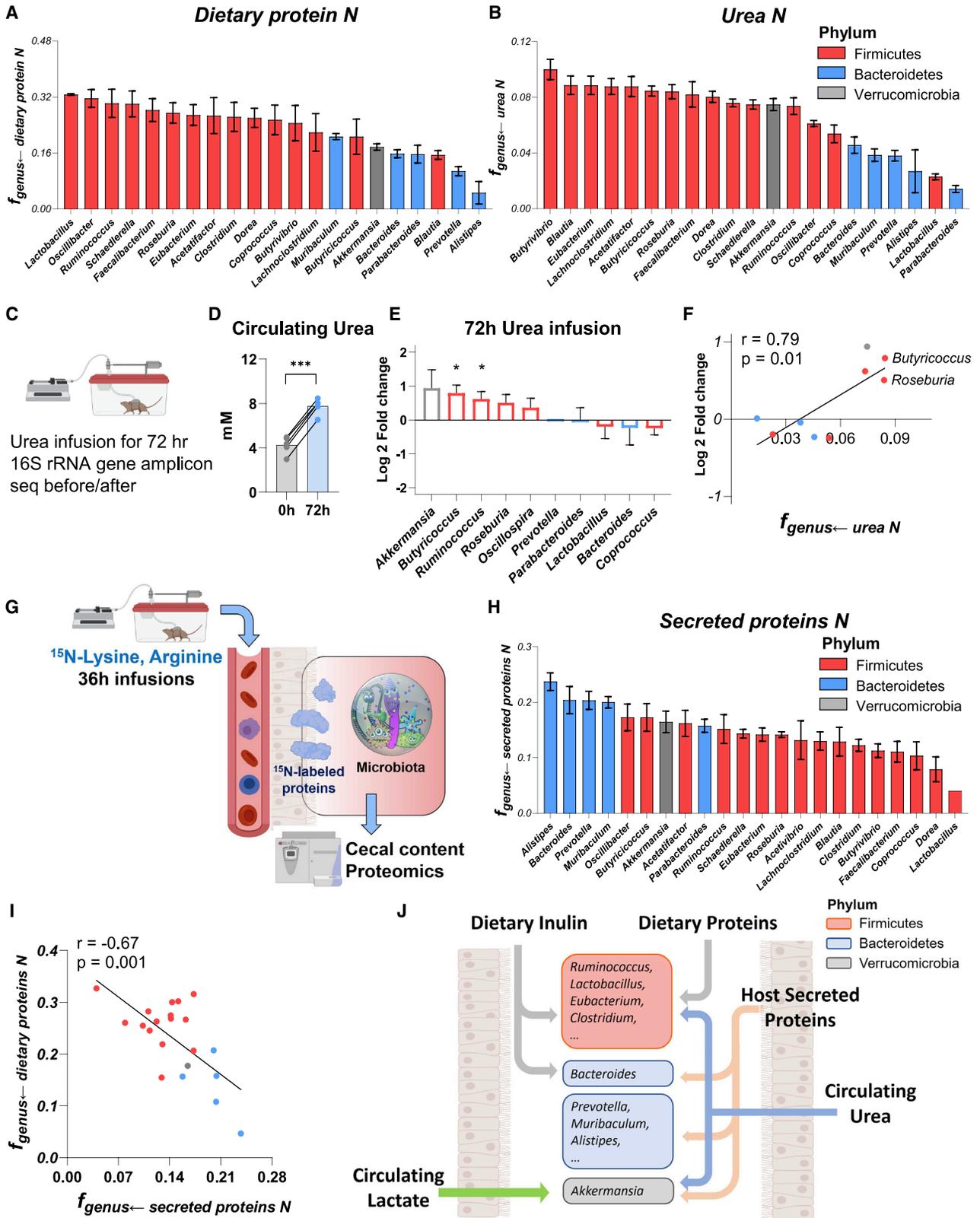
(E) Experimental scheme of high-inulin diet feeding followed by 16S rRNA gene amplicon sequencing.

(F) Genus-level microbiota composition changes after high-inulin diet. The genera increased after high-inulin diet prefer inulin in (B). Mean  $\pm$  SE,  $n = 3$  mice. \*  $p < 0.05$  and \*\*  $p < 0.01$  by two-sided Student's  $t$  test.

(G) Correlation between genera abundance changes and carbon-source preference.

(H–J) As in (E)–(G), for algal protein-supplemented diet.

See also Figure S6.



(legend on next page)

and arginine did contribute after 36-h infusion, consistent with the labeling occurring via host proteins. Such labeling occurred preferentially in Bacteroidetes and *Akkermansia* (Figures 6H and S6F). The nitrogen contributions from dietary and secreted host proteins were anti-correlated, consistent with some gut bacteria preferentially consuming dietary protein and others host protein (Figure 6I). Bacterial genera with a greater preference for dietary protein, whose availability depends on host feeding, grow more differently between daytime and nighttime (Figures S6J and S6K). Thus, dietary proteins and circulating urea are the major nitrogen feedstock of Firmicutes, whereas secreted host proteins provide nitrogen to Bacteroidetes.

## DISCUSSION

As for most microbial communities, the composition of the gut microbiome is shaped by nutrient availability. Here, we developed quantitative isotope-tracing approaches to measure the nutrient preferences of gut bacteria. In addition to dietary fiber and secreted host proteins, we establish dietary protein and circulating host lactate, 3-hydroxybutyrate, and urea as important nutrients feeding gut bacteria. Importantly, we rule out direct contributions from other circulating host nutrients, like glucose and amino acids, to the colonic microbiome.

A key technical achievement is enabling tracing from different carbon and nitrogen sources into bacteria-specific peptides, thereby revealing the nutrient preferences of different bacteria within the complex and competitive gut lumen environment. We find that Firmicutes and Bacteroidetes differ systematically in their utilization of host-secreted protein versus dietary protein: Firmicutes tend to acquire amino acids from dietary protein, whereas Bacteroidetes rely more on secreted host protein (Figure 6J). This may relate to different localization of bacteria within the colon, either in terms of central versus peripheral (closer to host mucus) or distal versus proximal (closer to incoming food remnants) (Albenberg et al., 2014; Li et al., 2015; Yasuda et al., 2015).

Within these two major families of gut bacteria, we found marked disparities in the use of dietary fiber as a carbon source. The most abundant Bacteroidetes genus is *Bacteroides*, and it was the most avid assimilator of fiber (inulin). In contrast, other types of bacteria in the same phylum hardly consumed inulin.

Likewise, some Firmicutes like *Clostridium* avidly used fiber, whereas others did not. Strikingly, feeding a fiber-enriched diet led to an increased abundance of *Bacteroides* and *Clostridium*, the precise genera that most actively assimilate fiber based on isotope tracing.

A similar trend was observed in the case of dietary supplementation with algal protein: Firmicutes, which actively use such protein, tended to increase in abundance. Algal protein (the only type commercially available in bulk in  $^{13}\text{C}$ -labeled form) may be particularly hard for mammals to digest. This is reflected in the  $^{13}\text{C}$ -labeled amino acids from algal protein appearing limitedly in the portal circulation and instead extensive passing from the intestine into the colon. This influx of dietary protein to the microbiome was a major contributor to secreted microbiome metabolites. As shown by replacing intact dietary protein with more absorbable (and thus less microbiome-accessible) free amino acids, the production and hence systemic concentration of these products depend on dietary protein reaching the colonic microbiome. An important future question is whether the nature of dietary protein (e.g., plant- or animal-based) impacts passage through the small intestine to the colonic microbiome and thereby shapes microbiome composition or metabolite secretion (Madsen et al., 2017; Wali et al., 2021).

Host circulating metabolite levels may also impact microbiome nutrient access and ultimately composition. Here, we show such effects are likely limited to the few host metabolites that meaningfully penetrate the microbiome: urea, 3-hydroxybutyrate, and lactate. Among them, lactate was recently shown to feed the gut microbiome in human marathon runners (Scheiman et al., 2019). Among gut bacteria, *Akkermansia* most avidly use circulating lactate. *Akkermansia* are mucin degraders, and their proximity to the gut epithelial wall may augment their access to lactate from the host circulation. *Akkermansia* are more abundant in athletes, and exercise increases their levels in mice and human (Liu et al., 2017; Munukka et al., 2018). A possible mechanism involves increased circulating lactate levels following exercise directly feeding *Akkermansia*. Whether lactate-induced *Akkermansia* growth in part mediates beneficial effects of exercise is an important open question. Consistent with their urea preferences measured by isotope tracing, *Akkermansia* and certain genera within Firmicutes (e.g., *Roseburia*, *Butyrivococcus*, and *Ruminococcus*) also increase in abundance upon experimental elevation of circulating host urea.

### Figure 6. Firmicutes favor dietary protein, while Bacteroidetes prefer secreted host protein

- (A) Nitrogen contribution of dietary algal protein across bacterial genera. Mean  $\pm$  SE, n = 6 mice.  
 (B) Nitrogen contribution of circulating urea across bacterial genera. Mean  $\pm$  SE, n = 6 mice.  
 (C) Experimental scheme of 72 h urea infusion followed by 16S rRNA gene amplicon sequencing.  
 (D) Urea infusions increased urea concentration in systemic circulation. n = 5 mice. \*\*\* p < 0.001 by two-sided Student's t test.  
 (E) Genus-level microbiota composition changes after urea infusion. The genera increased after urea infusion prefer urea in (B). Mean  $\pm$  SE, n = 5 mice. \* p < 0.05 by two-sided Student's t test.  
 (F) Correlation between genera abundance changes and nitrogen-source preference.  
 (G) Experimental schematic of long-term  $^{15}\text{N}$ -lysine and  $^{15}\text{N}$ -arginine infusion to probe the contribution of secreted host proteins to different bacterial genera.  
 (H) Nitrogen contribution of secreted host proteins across bacterial genera. Mean  $\pm$  SE, n = 5 mice.  
 (I) Negative correlation between dietary protein nitrogen and secreted proteins nitrogen contribution.  
 (J) Summary of carbon and nitrogen inputs to different gut bacteria. Firmicutes prefer dietary carbon sources (fiber and protein) and nitrogen from host circulating urea. Bacteroidetes heavily use dietary fiber, while using host-secreted proteins for nitrogen. Verrucomicrobia prefer host-secreted nutrients, both protein and circulating small molecules (lactate and urea).

See also Figures S6 and S7.

Ultimately, manipulating the microbiome requires understanding which nutrients different bacteria consume and how such consumption impacts microbiome composition and product secretion. Through isotope tracing, including proteomic measurements that offer bacterial genus specificity, we provide foundational knowledge about which nutrients feed the gut microbiome and which bacteria prefer which nutrients. The methodologies developed here are poised for broader application, which could eventually contribute to the holistic and quantitative understanding of the diet-microbiome-health connection.

### Limitations of the study

Our investigation focuses solely on healthy mice that are fed standard chow (in some cases with specific fiber or protein supplements). Measurements of microbiome feedstocks are limited to isotope tracing and mass spectrometry. Feedstocks of different bacteria are determined based on the isotopic signatures of bacteria-specific peptides. Peptide identification involves a 2% false discovery rate. Taxonomic assignment is based on bacterial proteome sequences available on Uniprot (Gurdeep Singh et al., 2019). Orthogonal approaches, which could provide measurement validation or complementary information, such as fluorescence-activated cell sorting of bacteria, were not explored (Batani et al., 2019). In most cases, taxonomic assignment was limited to the genus level due to the lack of sufficient specificity of the detected peptide sequences. In the future, improved sensitivity may enable species- or strain-specific peptide sequence measurements.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
  - Mouse studies
  - Bacterial culture studies
- **METHOD DETAILS**
  - Mouse gavage and nutrient feeding
  - Intravenous infusions
  - Antibiotics treatment
  - Sample collection
  - 16S rRNA gene amplicon sequencing and analysis
  - Bacterial culture studies
  - Bacterial colonization in mice
  - Metabolite extraction
  - Measurements of metabolites, protein, and polysaccharides
  - Proteomics sample preparation
  - Proteomics peptide measurement
  - Proteomics data analysis
  - Quantification of newly-synthesized fraction of peptide

- Quantification of contribution of labeled nutrient to peptide
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cell.2022.07.020>.

### ACKNOWLEDGMENTS

This work is supported by the NIH Pioneer award IDP1DK113643 (J.D.R.), Stand Up to Cancer Convergence Award 3.14.16 probing the cancer-microbiome connection (J.D.R.), Ludwig Cancer Research (J.D.R.), NIH grant R35GM128813 (M.W.), Princeton Catalysis Initiative (M.W.), Eric and Wendy Schmidt Transformative Technology Fund (M.W.), NIH grant 1R01AA029124 (C.J.), and the Pew Biomedical Scholars Program (M.S.D.). M.G. is funded by a Harold W. Dodds Fellowship. F.C.K. is funded by EMBO ALTF 601-2018. J.G.L. is funded by a National Science Foundation Graduate Research Fellowship (2017249408). Y.-C.J.L. is funded by the High Meadows Environmental Institute at Princeton University through the generous support of the William Clay Ford, Jr. '79 and Lisa Vanderzee Ford '82 Graduate Fellowship fund and by a training grant from the National Institute of General Medicine Sciences (T32GM007388). M.D.N. is supported by NIH 5T32CA257957. We thank Sheng Hui, Michel Nofal, Yihui Shen, Lingfan Liang, Won Dong Lee, and other Rabinowitz lab members for their input and advice. We thank Seema Chatterjee for the assistance with DNA extraction. We thank Michael Stadlmeier, Thao Nguyen, and Alex Johnson for the assistance with proteomics. We thank Wei Wang and the Lewis-Sigler Institute sequencing core facility for assistance with HT sequencing.

### AUTHOR CONTRIBUTIONS

X.Z., and J.D.R. came up with the general approach. X.Z. performed most of the experiment and data analysis. C.J. worked intensively with X.Z. to develop the experimental strategy. M.W. designed and enabled the proteomic measurements. X.X. wrote the MATLAB code. M.G., F.C.K., and M.D.N. contributed to proteomics method development. J.G.L. and M.S.D. provided microbiome expertise and performed 16S rRNA gene amplicon sequencing. Y.-C.J.L. performed *in vitro* bacterial culture studies. A.R. assisted with isotope tracing. L.W. performed ammonia measurement. X.Z., C.J., and J.D.R. wrote the paper. All the authors discussed the results and commented on the paper.

### DECLARATION OF INTERESTS

J.D.R. is a member of the Rutgers Cancer Institute of New Jersey and the University of Pennsylvania Diabetes Research Center; a co-founder and stockholder in Empress Therapeutics and Serien Therapeutics; and an advisor and stockholder in Agios Pharmaceuticals, Bantam Pharmaceuticals, Colorado Research Partners, Rafael Pharmaceuticals, Barer Institute, and L.E.A.F. Pharmaceuticals. M.S.D. is a member of the scientific advisory boards of DeepBiome Therapeutics and VastBiome.

Received: January 16, 2022

Revised: June 2, 2022

Accepted: July 21, 2022

Published: September 1, 2022

### REFERENCES

Albenberg, L., Esipova, T.V., Judge, C.P., Bittinger, K., Chen, J., Laughlin, A., Grunberg, S., Baldassano, R.N., Lewis, J.D., Li, H., et al. (2014). Correlation between intraluminal oxygen gradient and radial partitioning of intestinal microbiota. *Gastroenterology* 147, 1055–1063.e8. <https://doi.org/10.1053/j.gastro.2014.07.020>.

- Arab, J.P., Karpen, S.J., Dawson, P.A., Arrese, M., and Trauner, M. (2017). Bile acids and nonalcoholic fatty liver disease: molecular insights and therapeutic perspectives. *Hepatology* 65, 350–362. <https://doi.org/10.1002/hep.28709>.
- Bartman, C.R., TeSlaa, T., and Rabinowitz, J.D. (2021). Quantitative flux analysis in mammals. *Nat. Metab.* 3, 896–908. <https://doi.org/10.1038/s42255-021-00419-2>.
- Batani, G., Bayer, K., Böge, J., Hentschel, U., and Thomas, T. (2019). Fluorescence in situ hybridization (FISH) and cell sorting of living bacteria. *Sci. Rep.* 9, 18618. <https://doi.org/10.1038/s41598-019-55049-2>.
- Berry, D., Mader, E., Lee, T.K., Wobcken, D., Wang, Y., Zhu, D., Palatinszky, M., Schintlmeister, A., Schmid, M.C., Hanson, B.T., et al. (2015). Tracking heavy water (D<sub>2</sub>O) incorporation for identifying and sorting active microbial cells. *Proc. Natl. Acad. Sci. USA* 112, E194–E203. <https://doi.org/10.1073/pnas.1420406112>.
- Berry, D., Stecher, B., Schintlmeister, A., Reichert, J., Brugiroux, S., Wild, B., Wanek, W., Richter, A., Rauch, I., Decker, T., et al. (2013). Host-compound foraging by intestinal microbiota revealed by single-cell stable isotope probing. *Proc. Natl. Acad. Sci. USA* 110, 4720–4725. <https://doi.org/10.1073/pnas.1219247110>.
- Bokulich, N.A., Kaehler, B.D., Rideout, J.R., Dillon, M., Bolyen, E., Knight, R., Huttley, G.A., and Gregory Caporaso, J. (2018). Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* 6, 90. <https://doi.org/10.1186/s40168-018-0470-z>.
- Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857. <https://doi.org/10.1038/s41587-019-0209-9>.
- Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. <https://doi.org/10.1038/nmeth.3869>.
- Campbell, C., McKenney, P.T., Konstantinovskiy, D., Isaeva, O.I., Schizas, M., Verter, J., Mai, C., Jin, W.-B., Guo, C.-J., Violante, S., et al. (2020). Bacterial metabolism of bile acids promotes generation of peripheral regulatory T cells. *Nature* 581, 475–479. <https://doi.org/10.1038/s41586-020-2193-0>.
- Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M., et al. (2012). Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6, 1621–1624. <https://doi.org/10.1038/ismej.2012.8>.
- Dalile, B., Van Oudenhove, L., Vervliet, B., and Verbeke, K. (2019). The role of short-chain fatty acids in microbiota–gut–brain communication. *Nat. Rev. Gastroenterol. Hepatol.* 16, 461–478. <https://doi.org/10.1038/s41575-019-0157-3>.
- David, L.A., Maurice, C.F., Carmody, R.N., Gootenberg, D.B., Button, J.E., Wolfe, B.E., Ling, A.V., Devlin, A.S., Varna, Y., Fischbach, M.A., et al. (2014). Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 505, 559–563. <https://doi.org/10.1038/nature12820>.
- De Vadder, F., Kovatcheva-Datchary, P., Goncalves, D., Vinera, J., Zitoun, C., Duchamp, A., Bäckhed, F., and Mithieux, G. (2014). Microbiota-generated metabolites promote metabolic benefits via gut–brain neural circuits. *Cell* 156, 84–96. <https://doi.org/10.1016/j.cell.2013.12.016>.
- Desai, M.S., Seekatz, A.M., Koropatkin, N.M., Kamada, N., Hickey, C.A., Wolter, M., Pudlo, N.A., Kitamoto, S., Terrapon, N., Muller, A., et al. (2016). A dietary fiber-deprived gut microbiota degrades the colonic mucus barrier and enhances pathogen susceptibility. *Cell* 167, 1339–1353.e21. <https://doi.org/10.1016/j.cell.2016.10.043>.
- Elias, J.E., and Gygi, S.P. (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* 4, 207–214. <https://doi.org/10.1038/nmeth1019>.
- Faubert, B., Li, K.Y., Cai, L., Hensley, C.T., Kim, J., Zacharias, L.G., Yang, C., Do, Q.N., Doucette, S., Burguete, D., et al. (2017). Lactate metabolism in human lung tumors. *Cell* 171, 358–371.e9. <https://doi.org/10.1016/j.cell.2017.09.019>.
- Fernández-García, J., Altea-Manzano, P., Pranzini, E., and Fendt, S.-M. (2020). Stable isotopes for tracing mammalian-cell metabolism in vivo. *Trends Biochem. Sci.* 45, 185–201. <https://doi.org/10.1016/j.tibs.2019.12.002>.
- Funabashi, M., Grove, T.L., Wang, M., Varma, Y., McFadden, M.E., Brown, L.C., Guo, C., Higginbottom, S., Almo, S.C., and Fischbach, M.A. (2020). A metabolic pathway for bile acid dehydroxylation by the gut microbiome. *Nature* 582, 566–570. <https://doi.org/10.1038/s41586-020-2396-4>.
- Garrett, W.S. (2015). Cancer and the microbiota. *Science* 348, 80–86. <https://doi.org/10.1126/science.aaa4972>.
- Gupta, M., Sonnett, M., Ryazanova, L., Presler, M., and Wühr, M. (2018). Quantitative proteomics of *Xenopus* embryos I, sample preparation. In *Xenopus: Methods and Protocols, Methods in Molecular Biology*, K. Vlemingck, ed. (Springer), pp. 175–194. [https://doi.org/10.1007/978-1-4939-8784-9\\_13](https://doi.org/10.1007/978-1-4939-8784-9_13).
- Guurdeep Singh, R., Tanca, A., Palomba, A., Van der Jeugt, F., Verschaffelt, P., Uzzau, S., Martens, L., Dawyndt, P., and Mesuere, B. (2019). Unipept 4.0: functional analysis of metaproteome data. *J. Proteome Res.* 18, 606–615. <https://doi.org/10.1021/acs.jproteome.8b00716>.
- Halestrap, A.P., and Price, N.T. (1999). The proton-linked monocarboxylate transporter (MCT) family: structure, function and regulation. *Biochem. J.* 343, 281–299.
- Han, S., Van Treuren, W., Fischer, C.R., Merrill, B.D., DeFelice, B.C., Sanchez, J.M., Higginbottom, S.K., Guthrie, L., Fall, L.A., Dodd, D., et al. (2021). A metabolomics pipeline for the mechanistic interrogation of the gut microbiome. *Nature* 595, 415–420. <https://doi.org/10.1038/s41586-021-03707-9>.
- Hang, S., Paik, D., Yao, L., Kim, E., Trinath, J., Lu, J., Ha, S., Nelson, B.N., Kelly, S.P., Wu, L., et al. (2019). Bile acid metabolites control T H 17 and T reg cell differentiation. *Nature* 576, 143–148. <https://doi.org/10.1038/s41586-019-1785-z>.
- Holmes, A.J., Chew, Y.V., Colakoglu, F., Cliff, J.B., Klaassens, E., Read, M.N., Solon-Biet, S.M., McMahon, A.C., Cogger, V.C., Ruohonen, K., et al. (2017). Diet–microbiome interactions in health are controlled by intestinal nitrogen source constraints. *Cell Metab* 25, 140–151. <https://doi.org/10.1016/j.cmet.2016.10.021>.
- Holmes, W.E., Angel, T.E., Li, K.W., and Hellerstein, M.K. (2015). Chapter seven – dynamic proteomics: in vivo proteome-wide measurement of protein kinetics using metabolic labeling. In *Methods in Enzymology, Metabolic Analysis Using Stable Isotopes* (Academic Press), pp. 219–276. <https://doi.org/10.1016/bs.mie.2015.05.018>.
- Hui, S., Ghergurovich, J.M., Morscher, R.J., Jang, C., Teng, X., Lu, W., Esparza, L.A., Reya, T., Le Zhan, Yanxiang Guo, J., et al. (2017). Glucose feeds the TCA cycle via circulating lactate. *Nature* 551, 115–118. <https://doi.org/10.1038/nature24057>.
- Hunter, J.D. (2007). Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* 9, 90–95. <https://doi.org/10.1109/MCSE.2007.55>.
- Jang, C., Hui, S., Lu, W., Cowan, A.J., Morscher, R.J., Lee, G., Liu, W., Tesz, G.J., Birnbaum, M.J., and Rabinowitz, J.D. (2018). The small intestine converts dietary fructose into glucose and organic acids. *Cell Metab* 27, 351–361.e3. <https://doi.org/10.1016/j.cmet.2017.12.016>.
- Jang, C., Hui, S., Zeng, X., Cowan, A.J., Wang, L., Chen, L., Morscher, R.J., Reyes, J., Frezza, C., Hwang, H.Y., et al. (2019). Metabolite exchange between mammalian organs quantified in pigs. *Cell Metab* 30, 594–606.e3. <https://doi.org/10.1016/j.cmet.2019.06.002>.
- Koh, A., De Vadder, F.D., Kovatcheva-Datchary, P., and Bäckhed, F. (2016). From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. *Cell* 165, 1332–1345. <https://doi.org/10.1016/j.cell.2016.05.041>.
- Koh, A., Molinaro, A., Ståhlman, M., Khan, M.T., Schmidt, C., Mannerås-Holm, L., Wu, H., Carreras, A., Jeong, H., Olofsson, L.E., et al. (2018). Microbially produced imidazole propionate impairs insulin signaling through mTORC1. *Cell* 175, 947–961.e17. <https://doi.org/10.1016/j.cell.2018.09.055>.
- Lai, Y., Liu, C.-W., Yang, Y., Hsiao, Y.-C., Ru, H., and Lu, K. (2021). High-coverage metabolomics uncovers microbiota-driven biochemical landscape

- of interorgan transport and gut-brain communication in mice. *Nat. Commun.* 12, 6000. <https://doi.org/10.1038/s41467-021-26209-8>.
- Li, H., Limenitakis, J.P., Fuhrer, T., Geuking, M.B., Lawson, M.A., Wyss, M., Brugiroux, S., Keller, I., Macpherson, J.A., Rupp, S., et al. (2015). The outer mucus layer hosts a distinct intestinal microbial niche. *Nat. Commun.* 6, 8292. <https://doi.org/10.1038/ncomms9292>.
- Liu, Z., Liu, H.-Y., Zhou, H., Zhan, Q., Lai, W., Zeng, Q., Ren, H., and Xu, D. (2017). Moderate-intensity exercise affects gut microbiome composition and influences cardiac function in myocardial infarction mice. *Front. Microbiol.* 8, 1687. <https://doi.org/10.3389/fmicb.2017.01687>.
- Lund, P.J., Gates, L.A., Leboeuf, M., Smith, S.A., Chau, L., Friedman, E.S., Lopes, M., Saiman, Y., Kim, M.S., Petucci, C., et al. (2021). Stable isotope tracing in vivo reveals a metabolic bridge linking the microbiota to host histone acetylation. *bioRxiv*. <https://doi.org/10.1101/2021.07.05.450926>.
- Madsen, L., Myrmel, L.S., Fjære, E., Liaset, B., and Kristiansen, K. (2017). Links between dietary protein sources, the gut microbiota, and obesity. *Front. Physiol.* 8, 1047. <https://doi.org/10.3389/fphys.2017.01047>.
- Mager, L.F., Burkhard, R., Pett, N., Cooke, N.C.A., Brown, K., Ramay, H., Paik, S., Stagg, J., Groves, R.A., Gallo, M., et al. (2020). Microbiome-derived inosine modulates response to checkpoint inhibitor immunotherapy. *Science* 369, 1481–1489. <https://doi.org/10.1126/science.abc3421>.
- McCabe, B.J., and Previs, S.F. (2004). Using isotope tracers to study metabolism: application in mouse models. *Metab. Eng.* 6, 25–35. <https://doi.org/10.1016/j.ymben.2003.09.003>.
- McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A., Andersen, G.L., Knight, R., and Hugenholtz, P. (2012). An improved green genes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J* 6, 610–618. <https://doi.org/10.1038/ismej.2011.139>.
- McKinney, W. (2010). Data structures for statistical computing in python. In *Proceedings of the 9th Python Science Conference*, pp. 56–61. <https://doi.org/10.25080/Majora-92bf1922-00a>.
- Mora, D., and Arioli, S. (2014). Microbial urease in health and disease. *PLOS Pathog* 10, e1004472. <https://doi.org/10.1371/journal.ppat.1004472>.
- Munukka, E., Ahtiaainen, J.P., Puigbó, P., Jalkanen, S., Pahkala, K., Keskitalo, A., Kujala, U.M., Pietilä, S., Hollmén, M., Elo, L., et al. (2018). Six-week endurance exercise alters gut metagenome that is not reflected in systemic metabolism in over-weight women. *Front. Microbiol.* 9, 2323. <https://doi.org/10.3389/fmicb.2018.02323>.
- Nemet, I., Saha, P.P., Gupta, N., Zhu, W., Romano, K.A., Skye, S.M., Cajka, T., Mohan, M.L., Li, L., Wu, Y., et al. (2020). A cardiovascular disease-linked gut microbial metabolite acts via adrenergic receptors. *Cell* 180, 862–877.e22. <https://doi.org/10.1016/j.cell.2020.02.016>.
- Ni, J., Shen, T.-C.D., Chen, E.Z., Bittinger, K., Bailey, A., Roggiani, M., Sirota-Madi, A., Friedman, E.S., Chau, L., Lin, A., et al. (2017). A role for bacterial urease in gut dysbiosis and Crohn's disease. *Sci. Transl. Med.* 9. <https://doi.org/10.1126/scitranslmed.aah6888>.
- Nusinow, D.P., Szpyt, J., Ghandi, M., Rose, C.M., McDonald, E.R., Kalocsay, M., Jané-Valbuena, J., Gelfand, E., Schweppe, D.K., Jedrychowski, M., et al. (2020). Quantitative proteomics of the cancer cell line encyclopedia. *Cell* 180, 387–402.e16. <https://doi.org/10.1016/j.cell.2019.12.023>.
- O'Brien, J.J., Narayan, V., Wong, Y., Seitzer, P., Sandoval, C.M., Haste, N., Smith, M., Rad, R., Gaun, A., Baker, A., et al. (2020). Precise estimation of in vivo protein turnover rates. *bioRxiv*. 2020.11.10.377440. <https://doi.org/10.1101/2020.11.10.377440>.
- Oberbach, A., Haange, S.-B., Schlichting, N., Heinrich, M., Lehmann, S., Till, H., Hugenholtz, F., Kullnick, Y., Smidt, H., Frank, K., et al. (2017). Metabolic *in vivo* labeling highlights differences of metabolically active microbes from the mucosal gastrointestinal microbiome between high-fat and normal chow diet. *J. Proteome Res.* 16, 1593–1604. <https://doi.org/10.1021/acs.jproteome.6b00973>.
- Purser, D.B., and Buechler, S.M. (1966). Amino acid composition of rumen organisms. *J. Dairy Sci.* 49, 81–84. [https://doi.org/10.3168/jds.S0022-0302\(66\)87791-3](https://doi.org/10.3168/jds.S0022-0302(66)87791-3).
- Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65. <https://doi.org/10.1038/nature08821>.
- Quinn, R.A., Melnik, A.V., Vrbancac, A., Fu, T., Patras, K.A., Christy, M.P., Bodai, Z., Belda-Ferre, P., Tripathi, A., Chung, L.K., et al. (2020). Global chemical effects of the microbiome include new bile-acid conjugations. *Nature* 579, 123–129. <https://doi.org/10.1038/s41586-020-2047-9>.
- Reese, A.T., Pereira, F.C., Schintlmeister, A., Berry, D., Wagner, M., Hale, L.P., Wu, A., Jiang, S., Durand, H.K., Zhou, X., et al. (2018). Microbial nitrogen limitation in the mammalian large intestine. *Nat. Microbiol.* 3, 1441–1450. <https://doi.org/10.1038/s41564-018-0267-7>.
- Ridlon, J.M., Kang, D.J., Hylemon, P.B., and Bajaj, J.S. (2014). Bile acids and the gut microbiome. *Curr. Opin. Gastroenterol.* 30, 332–338. <https://doi.org/10.1097/MOG.0000000000000057>.
- Savitski, M.M., Wilhelm, M., Hahne, H., Kuster, B., and Bantscheff, M. (2015). A scalable approach for protein false discovery rate estimation in large proteomic data sets. *Mol. Cell. Proteomics* 14, 2394–2404. <https://doi.org/10.1074/mcp.M114.046995>.
- Scheiman, J., Luber, J.M., Chavkin, T.A., MacDonald, T., Tung, A., Pham, L.-D., Wibowo, M.C., Wurth, R.C., Punthambaker, S., Tierney, B.T., et al. (2019). Meta-omics analysis of elite athletes identifies a performance-enhancing microbe that functions via lactate metabolism. *Nat. Med.* 25, 1104–1109. <https://doi.org/10.1038/s41591-019-0485-4>.
- Sicard, J.-F., Le Bihan, G., Vogeleeer, P., Jacques, M., and Harel, J. (2017). Interactions of intestinal bacteria with components of the intestinal mucus. *Front. Cell. Infect. Microbiol.* 7, 387. <https://doi.org/10.3389/fcimb.2017.00387>.
- Spinelli, J.B., Kelley, L.P., and Haigis, M.C. (2017a). An LC-MS approach to quantitative measurement of ammonia isotopologues. *Sci. Rep.* 7, 10304. <https://doi.org/10.1038/s41598-017-09993-6>.
- Spinelli, J.B., Yoon, H., Ringel, A.E., Jeanfavre, S., Clish, C.B., and Haigis, M.C. (2017b). Metabolic recycling of ammonia via glutamate dehydrogenase supports breast cancer biomass. *Science* 358, 941–946. <https://doi.org/10.1126/science.aam9305>.
- Tang, W.H.W., Wang, Z., Levison, B.S., Koeth, R.A., Britt, E.B., Fu, X., Wu, Y., and Hazen, S.L. (2013). Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk [WWW Document]. *N. Engl. J. Med.* 368, 1575–1584. <https://doi.org/10.1056/NEJMoa1109400>. [10.1056/NEJMoa1109400](https://doi.org/10.1056/NEJMoa1109400).
- Ullrich, K.J., Rumrich, G., and Klöss, S. (1982). Reabsorption of monocarboxylic acids in the proximal tubule of the rat kidney. II. Specificity for aliphatic compounds. *Pflüg. Arch. Eur. J. Physiol.* 395, 220–226. <https://doi.org/10.1007/BF00584813>.
- Wali, J.A., Milner, A.J., Luk, A.W.S., Pulpitel, T.J., Dodgson, T., Facey, H.J.W., Wahl, D., Kebede, M.A., Senior, A.M., Sullivan, M.A., et al. (2021). Impact of dietary carbohydrate type and protein-carbohydrate interaction on metabolic health. *Nat. Metab.* 3, 810–828. <https://doi.org/10.1038/s42255-021-00393-9>.
- Wang, Z., Klipfell, E., Bennett, B.J., Koeth, R., Levison, B.S., DuGar, B., Feldstein, A.E., Britt, E.B., Fu, X., Chung, Y.-M., et al. (2011). Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* 472, 57–63. <https://doi.org/10.1038/nature09922>.
- Wikoff, W.R., Anfora, A.T., Liu, J., Schultz, P.G., Lesley, S.A., Peters, E.C., and Siuzdak, G. (2009). Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc. Natl. Acad. Sci. USA* 106, 3698–3703. <https://doi.org/10.1073/pnas.0812874106>.
- Wolfe, R.R. (1984). Tracers in metabolic research: radioisotope and stable isotope/mass spectrometry methods. *Lab. Res. Methods Biol. Med.* 9, 1–287.
- Wong, J.M.W., and Jenkins, D.J.A. (2007). Carbohydrate digestibility and metabolic effects. *J. Nutr.* 137, 2539S–2546S. <https://doi.org/10.1093/jn/137.11.2539S>.

- Wühr, M., Freeman, R.M., Presler, M., Horb, M.E., Peshkin, L., Gygi, S.P., and Kirschner, M.W. (2014). Deep proteomics of the *Xenopus laevis* egg using an mRNA-derived reference database. *Curr. Biol.* *24*, 1467–1475. <https://doi.org/10.1016/j.cub.2014.05.044>.
- Yasuda, K., Oh, K., Ren, B., Tickle, T.L., Franzosa, E.A., Wachtman, L.M., Miller, A.D., Westmoreland, S.V., Mansfield, K.G., Vallender, E.J., et al. (2015). Biogeography of the intestinal mucosal and luminal microbiome in the rhesus macaque. *Cell Host Microbe* *17*, 385–391. <https://doi.org/10.1016/j.chom.2015.01.015>.
- Yoshimoto, S., Loo, T.M., Atarashi, K., Kanda, H., Sato, S., Oyadomari, S., Iwakura, Y., Oshima, K., Morita, H., Hattori, M., et al. (2013). Obesity-induced gut microbial metabolite promotes liver cancer through senescence secretome. *Nature* *499*, 97–101. <https://doi.org/10.1038/nature12347>.
- Zhang, X., Ning, Z., Mayne, J., Deeke, S.A., Li, J., Starr, A.E., Chen, R., Singleton, R., Butcher, J., Mack, D.R., et al. (2016a). *In vitro* metabolic labeling of intestinal microbiota for quantitative metaproteomics. *Anal. Chem.* *88*, 6120–6125. <https://doi.org/10.1021/acs.analchem.6b01412>.
- Zhang, X., Ning, Z., Mayne, J., Moore, J.I., Li, J., Butcher, J., Deeke, S.A., Chen, R., Chiang, C.-K., Wen, M., et al. (2016b). MetaPro-IQ: a universal metaproteomic approach to studying human and mouse gut microbiota. *Microbiome* *4*, 31. <https://doi.org/10.1186/s40168-016-0176-z>.
- Zhao, S., Jang, C., Liu, J., Uehara, K., Gilbert, M., Izzo, L., Zeng, X., Trefely, S., Fernandez, S., Carrer, A., et al. (2020). Dietary fructose feeds hepatic lipogenesis via microbiota-derived acetate. *Nature* *579*, 586–591. <https://doi.org/10.1038/s41586-020-2101-7>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
DEUTERIUM OXIDE (D, 99.9%)	Cambridge Isotope Laboratories	Cat# DLM-4-PK
D-GLUCOSE (U-13C6, 99%)	Cambridge Isotope Laboratories	Cat# CLM-1396-PK
SODIUM L-LACTATE (13C3, 98%) 20% W/W in H2O	Cambridge Isotope Laboratories	Cat# CLM-1579-PK
L-GLUTAMINE (13C5, 99%)	Cambridge Isotope Laboratories	Cat# CLM-1822-H-PK
L-GLUTAMINE (ALPHA-15N, 98%)	Cambridge Isotope Laboratories	Cat# NLM-1016-PK
L-GLUTAMINE (AMIDE-15N, 98%+)	Cambridge Isotope Laboratories	Cat# NLM-557-PK
SODIUM D-3-HYDROXYBUTYRATE (13C4, 99%) 97% CHEMICAL PURITY	Cambridge Isotope Laboratories	Cat# CLM-3853-PK
LINOLEIC ACID, POTASSIUM SALT (U-13C18, 98%)	Cambridge Isotope Laboratories	Cat# CLM-8835-PK
OLEIC ACID, SODIUM SALT (U-13C18, 98%)	Cambridge Isotope Laboratories	Cat# CLM-8763-PK
SODIUM PALMITATE (U-13C16, 98%+)	Cambridge Isotope Laboratories	Cat# CLM-6059-PK
SODIUM ACETATE (1,2-13C2, 99%)	Cambridge Isotope Laboratories	Cat# CLM-440-PK
CITRIC ACID (13C6, 99%)	Cambridge Isotope Laboratories	Cat# CLM-9021-PK
SUCCINIC ACID (13C4, 99%)	Cambridge Isotope Laboratories	Cat# CLM-1571-PK
L-MALIC ACID (13C4, 99%)	Cambridge Isotope Laboratories	Cat# CLM-8065-PK
ALPHA-KETOGLUTARIC ACID, DISODIUM SALT (1,2,3,4-13C4, 99%)	Cambridge Isotope Laboratories	Cat# CLM-4442-PK
L-ALANINE (13C3, 99%)	Cambridge Isotope Laboratories	Cat# CLM-2184-H-PK
L-ALANINE (15N, 98%)	Cambridge Isotope Laboratories	Cat# NLM-454-PK
L-VALINE (13C5, 99%)	Cambridge Isotope Laboratories	Cat# CLM-2249-H-PK
L-VALINE (15N, 98%)	Cambridge Isotope Laboratories	Cat# NLM-316-PK
L-LEUCINE (13C6, 99%)	Cambridge Isotope Laboratories	Cat# CLM-2262-H-PK
L-LEUCINE (15N, 98%)	Cambridge Isotope Laboratories	Cat# NLM-142-PK
L-ISOLEUCINE (13C6, 99%)	Cambridge Isotope Laboratories	Cat# CLM-2248-H-PK
L-ISOLEUCINE (15N, 98%)	Cambridge Isotope Laboratories	Cat# NLM-292-PK
L-SERINE (13C3, 99%; 15N, 99%)	Cambridge Isotope Laboratories	Cat# CNLM-474-H-PK
GLYCINE (13C2, 97-99%)	Cambridge Isotope Laboratories	Cat# CLM-1017-PK
GLYCINE (13C2, 99%; 15N, 99%)	Cambridge Isotope Laboratories	Cat# CNLM-1673-H-PK
L-TYROSINE (13C9, 99%)	Cambridge Isotope Laboratories	Cat# CLM-2263-H-PK
UREA (15N2, 98%+)	Cambridge Isotope Laboratories	Cat# NLM-233-PK
AMMONIUM CHLORIDE (15N, 99%)	Cambridge Isotope Laboratories	Cat# NLM-467-PK
L-LYSINE:2HCL (ALPHA-15N, 98%)	Cambridge Isotope Laboratories	Cat# NLM-143-PK
L-ARGININE:HCL (ALPHA-15N, 98%+)	Cambridge Isotope Laboratories	Cat# NLM-1267-PK
ALGAL STARCH (U-13C, 98%+)	Cambridge Isotope Laboratories	Cat# CLM-1699-PK
INULIN (FROM CHICORY) (U-13C, 97%+) 97% CHEMICAL PURITY	Cambridge Isotope Laboratories	Cat# CLM-9181-PK
Algal crude protein extract-13C	Sigma-Aldrich	Cat# 642878
Algal crude protein extract-15N	Sigma-Aldrich	Cat# 586773
Algal crude protein extract-13C,15N	Sigma-Aldrich	Cat# 608254
Algal amino acid mixture-13C	Sigma-Aldrich	Cat# 426199
Algal fatty acid mixture-13C	Sigma-Aldrich	Cat# 487937
AZD 3965	AstraZeneca	Cat# AZD3965
Ampicillin	Sigma-Aldrich	Cat# A0166

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Neomycin trisulfate salt hydrate	Sigma-Aldrich	Cat# N6386
Metronidazole	Sigma-Aldrich	Cat# M1547
Vancomycin hydrochloride from <i>Streptomyces orientalis</i>	Sigma-Aldrich	Cat# V1130
Trypsin	Promega	Cat# V5113
Lysyl endopeptidase R	Wako Chemicals USA	Cat# 12902541
Aspartame	Sigma-Aldrich	Cat# 47135
GAM Broth Modified	HyServe	Cat# 5433
LB Broth (Miller, Luria Broth)	Sigma	Cat# L3522
MRS Broth	Sigma	Cat# 69966
TSB (Tryptic Soy Broth)	Bacto	Cat# 211825

**Experimental models: Organisms/strains**

Mouse: C57BL/6	Charles River Laboratories	Cat #027
Strain: <i>Bacteroides dorei</i> (CL02T00C15)	BEI	#HM-717
Strain: <i>Clostridium sporogenes</i> (ATCC 15579)	ATCC	15579
Strain: <i>Escherichia coli</i> (ATCC 25922)	ATCC	25922
Strain: <i>Lactobacillus reuteri</i> (CF-48-34A)	BEI	#HM-102
Strain: <i>Staphylococcus aureus</i> subsp. <i>Aureus</i> Rosenbach	ATCC	29213

**Software and algorithms**

El-MAVEN software	Elucidata	<a href="https://www.elucidata.io/el-maven">https://www.elucidata.io/el-maven</a>
AccuCor	GitHub	<a href="https://github.com/XiaoyangSu/AccuCor">https://github.com/XiaoyangSu/AccuCor</a>
PepMID	GitHub	<a href="https://github.com/xxing9703/pepMID_simul">https://github.com/xxing9703/pepMID_simul</a>
MATLAB R2021b	MathWorks	N/A

**Deposited data**

Proteomics Data	N/A	PXD031015
-----------------	-----	-----------

**Others**

PicoLab Rodent Diet 20	LabDiet	Cat# 5053
20% Diet premix	Research Diets	Cat# D11112201Npx2i
20% Amino acids diet	Research Diets	Cat# A11112201Bi

**RESOURCE AVAILABILITY****Lead contact**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Professor Joshua D. Rabinowitz ([joshr@princeton.edu](mailto:joshr@princeton.edu)).

**Materials availability**

This study did not generate new unique reagents or new mouse lines.

**Data and code availability**

The proteomics datasets generated during this study are deposited in ProteomeXchange: PXD031015. The isotope tracing data are included in [Table S2](#). The taxonomic assignment of the detected tryptic peptides in the study are included in [Table S3](#). Composition of the diet used in the study are included in [Table S4](#). The 16S rRNA gene amplicon sequencing datasets generated during this study are available in [Table S5](#). The code for peptide enrichment calculations generated during this study is available at GitHub: ([https://github.com/xxing9703/pepMID\\_simul](https://github.com/xxing9703/pepMID_simul)). Any additional information required to re-analyze the data reported in this work is available from the [lead contact](#) upon request.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Mouse studies

Mouse studies followed protocols approved by the Princeton University Animal Care and Use Committee. Unless otherwise indicated, 7-9-week-old male C57BL/6NCrl mice (strain 027; Charles River Laboratories) fed PicoLab Rodent Diet 20 were group-housed on a normal light-dark cycle (8:00-20:00) with free access to water and chow.

### Bacterial culture studies

*B. dorei*, *C. sporogenes*, *E. coli*, *S. aureus* and *L. reuteri* glycerol stocks were brought into an anaerobic chamber (70% N<sub>2</sub>, 25% CO<sub>2</sub>, 5% H<sub>2</sub>) and grown in liquid media: *L. reuteri* was grown on MRS (MRS Broth, Sigma); *E. coli* was grown on LB (Luria Broth, Sigma); *S. aureus* was grown in TSB (Tryptic Soy Broth, Bacto) and *C. sporogenes* and *B. dorei* were grown in GAM (GAM Broth Modified, HyServe).

## METHOD DETAILS

### Mouse gavage and nutrient feeding

For the <sup>13</sup>C-nutrient gavage experiments, mice were fasted at 9 am and received a 1:2:4 mixture of inulin, protein/amino acids, and starch (0.5 g kg<sup>-1</sup> inulin, 1 g kg<sup>-1</sup> protein/amino acids 2g kg<sup>-1</sup> starch dissolved in water) at 3 pm via oral gavage with a plastic feeding tube (Instech Laboratories). Food was given back at 8 pm.

For the mouse experiments involving labeled nutrient feeding, the labeled diet was prepared by adding <sup>13</sup>C/<sup>15</sup>N-nutrients to a diet mixture premix (modified from normal diet with reduced protein, inulin, and starch content, Research diets Inc, D20030303). The final enrichment for each labeled dietary nutrient was 10% - 25% (with observed labeling corrected by dividing by the fraction dietary nutrient labeled). The contribution of each dietary nutrient to metabolites is calculated by the metabolite labeling enrichment normalized to the final enrichment of each labeled dietary nutrient. All diets shared the same final macronutrient composition (40% starch, 20% protein or amino acids, 7.5% inulin and 2.5% cellulose). Mice were first adapted to a non-labeled diet (of identical composition to the subsequent labeled diet) for 10 days, and then fed labeled diet for 24 h prior to sacrifice.

For the deuterium water drinking experiment, mice were administered a bolus intraperitoneal injected of D<sub>2</sub>O (1.26 % w/w relative to body weight), followed by having ad lib access to 3% D<sub>2</sub>O drinking water.

For the protein and amino acids diet feeding experiment, mice were fed on casein or compositional matched amino acids diet (20% casein, 13% casein + 7% amino acids, 7% casein +13% amino acids, and 20% amino acids as protein/amino acids sources; [Table S4](#)) for 2 weeks. Serum was sampled by tail-bleed at 9 am *ad lib*.

### Intravenous infusions

To quantify contribution of circulating nutrients to microbiota metabolism, 9-11-week-old C57BL/6 mice were catheterized in house in the right jugular vein. The mice were infused with carbon or nitrogen-labeled tracer starting at 3:30 pm without any fasting. Infusion rate was 0.1 ul/min/g. Infusion solutions are described in [Table S2A](#). Overnight (24 h) infusions both started and finished around 9 am. The contribution of circulating nutrient to each metabolite is calculated by the metabolite labeling enrichment normalized to the average tracer serum enrichment throughout 24 hr.

### Antibiotics treatment

To deplete the mouse resident microbiome, an antibiotic drinking water protocol was used. In brief, mice were treated with a cocktail of antibiotics (1 g/L ampicillin, 1 g/L neomycin, 1 g/L metronidazole, and 1 g/L vancomycin) in both their drinking water 14 days. To make the drinking water more palatable, 5% aspartame was added. The effectiveness of antibiotics treatments was verified by observing much lower SCFAs in the feces by LC-MS.

### Sample collection

Systemic blood samples were collected by tail bleeding. For sampling from tissue-specific draining veins, a mouse was put under anesthesia and different tissue veins were exposed, and blood samples were pulled with an insulin syringe (BD insulin syringes, # SY8290328291) insertion into the vein. Successful isolation of portal vein was confirmed by much higher (> 10x) concentrations of SCFAs and secondary bile acids (deoxycholic acid and lithocholic acid) than systemic vein; hepatic vein was confirmed by much lower secondary bile acids, SCFAs and higher glucose, 3-hydroxybutyrate levels compared to portal vein. Mouse urine was collected from the urinary bladder using a syringe. All serum samples were placed on ice without anticoagulant for 15 min, and centrifuged at 16,000 x g for 15 min at 4 C.

Tissues were harvested by quick dissection and snap freezing (< 5 sec) in liquid nitrogen with a pre-cooled Wollenberger clamp; intestinal contents were removed before clamping. For cecal content sampling, the mouse cecum was first removed and cut on the surface, then the cecal content was squeezed out using a tweezer followed by freeze clamping. Whole liver, intestine, and intestinal contents were collected and ground to homogenous powder. To sample fresh feces, the mouse belly was gently massaged to induce defecation and fresh feces were freeze clamped. For long-term feces collection, a mouse was transferred to a new cage and mouse

fecal pellets on the bedding were collected every 1~2 h and freeze clamped. Serum, tissue, and feces samples were kept at -80 °C until further analysis.

### 16S rRNA gene amplicon sequencing and analysis

Extraction of Bacterial DNA from cecal or fecal samples was performed using the Power Soil DNA Isolation kit (QIAGEN). A section of the 16S rRNA gene (~250 bp, V4 region) was amplified, and Illumina sequencing libraries were prepared from these amplicons according to a previously published protocol and primers (Caporaso et al., 2012). Libraries were further pooled together at equal molar ratios and sequenced on an Illumina HiSeq 2500 Rapid Flowcell or MiSeq as paired-end reads. These reads were 2x150 bp with an average depth of ~20,000 reads. Also included were 8 bp index reads, following the manufacturer's protocol (Illumina, USA). Pass-Filter reads were generated from raw sequencing reads using Illumina HiSeq Control Software. Samples were de-multiplexed using the index reads. The DADA2 plugin within QIIME2 version 2018.6 was used to infer Amplicon sequencing variants (ASVs) from the unmerged paired-end sequences (Bolyen et al., 2019; Callahan et al., 2016). The forward reads were trimmed at 150 bp and the reverse reads trimmed at 140 bp, with all other DADA2 as default. Taxonomy was assigned to the resulting ASVs with a naïve Bayes classifier trained on the Greengenes database version 13.8, with only the target region of the 16S rRNA gene used to train the classifier (Bokulich et al., 2018; McDonald et al., 2012). Downstream analyses were performed MATLAB (Hunter, 2007; McKinney, 2010).

### Bacterial culture studies

For the D<sub>2</sub>O experiment, 250 - 1000 µl D<sub>2</sub>O was added into the media (5-10 mL, to reach a final enrichment of 5-10%) with either *B. dorei* or *C. sporogenes*, and OD<sub>600</sub> was recorded at the addition. After every 25-30 min, OD<sub>600</sub> was recorded and 100-200 µl bacterial solution was taken for metabolomics and proteomics analysis. The newly synthesized fraction of bacteria was calculated by  $(OD_{600} - OD_{600, 0min})/OD_{600}$ .

### Bacterial colonization in mice

Mice were treated with antibiotics in drinking water for 10 days. On day 11, no antibiotics were administered, and mice were gavaged with 250 µl of bacterial consortia consisting of urease-negative bacteria (*B. dorei*, *C. sporogenes* and *E. coli*) or a combination of urease-negative and urease-positive bacteria (*B. dorei*, *C. sporogenes*, *E. coli*, *S. aureus* and *L. reuteri*).

### Metabolite extraction

For serum samples, 3 µl serum was added to 90 µl methanol and incubated on ice for 10 min, followed by centrifugation at 17,000 × g for 10 min at 4°C. The supernatant was transferred to an MS vial until further analysis. For tissues and feces samples, frozen samples were first ground at liquid nitrogen temperature with a cryomill (Restch, Newtown, PA). The resulting tissue powder was extracted with 40:40:20 methanol: acetonitrile: water (40 µl extraction solvent per 1 mg tissue) for 10 min on ice, followed by centrifugation at 17,000 × g for 10 min, and the supernatant was transferred to a MS vial until further analysis.

### Measurements of metabolites, protein, and polysaccharides

To measure metabolites in serum, tissue and feces samples, a quadrupole orbitrap mass spectrometer (Q Exactive; Thermo Fisher Scientific) was coupled to a Vanquish UHPLC system (Thermo Fisher Scientific) with electrospray ionization and scan range m/z from 60 to 1000 at 1 Hz, with a 140,000 resolution. LC separation was performed on an XBridge BEH Amide column (2.1 × 150 mm, 2.5 µm particle size, 130 Å pore size; Waters Corporation) using a gradient of solvent A (95:5 water: acetonitrile with 20 mM of ammonium acetate and 20 mM of ammonium hydroxide, pH 9.45) and solvent B (acetonitrile). Flow rate was 150 µl/min. The LC gradient was: 0 min, 85% B; 2 min, 85% B; 3 min, 80% B; 5 min, 80% B; 6 min, 75% B; 7 min, 75% B; 8 min, 70% B; 9 min, 70% B; 10 min, 50% B; 12 min, 50% B; 13 min, 25% B; 16 min, 25% B; 18 min, 0% B; 23 min, 0% B; 24 min, 85% B; and 30 min, 85% B. Injection volume was 5-10 µl and autosampler temperature was set at 4°C. For cysteine measurement, samples were derivatized before measurement as follows: Serum, cecal content or feces samples were extracted and centrifuged. To the supernatant, 2 mM N-ethylmaleimide was added and incubated at room temperature for 20 min. The resulting mixture was transferred to a MS vial. Derivatized cysteine has a m/z at 245.06015 in negative mode.

To quantify the metabolite concentration in serum and tissue samples, either isotope spike-in or standard spike-in was performed. For isotope spike-in, known concentrations of isotope-labeled standard were added to the serum or tissues extraction solution, then the concentration was calculated by the ratio of labeled and unlabeled metabolites. When isotope standard is not available, a serially diluted non-labeled standard was added, and a linear fitting between measured total ion count and added concentration of standard was generated. Then, the concentration of endogenous metabolite was determined by the x intercept of the fitting line.

Starch and inulin were measured by acid hydrolysis and LC-MS. In brief, 5-10 mg sample was mixed with 10 µl 2 M hydrochloric acid, and samples were incubated at 80°C for 2 h. After cooling down, the resulting mixture was neutralized with 12 µl saturated sodium bicarbonate, followed with 88 µl 1:1 acetonitrile: methanol solution. After centrifugation at 17,000 × g for 10 min at 4°C, the supernatant was transferred to a MS vial. Inulin and starch concentration in samples was inferred from total ion count of fructose and glucose, respectively.

SCFAs and BCFAs were derivatized and measured by LC-MS. Serum (5 µl) or tissue samples (~10 mg) were added to 100 µl derivatizing reagents containing 12 mM 1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide, 21 mM 3-Nitrophenylhydrazine hydrochloride

acid and pyridine (2.4% v/v) in methanol. The reaction was incubated at 4°C for 1 h. Then, the reaction mixture was centrifuged at 17,000 g for 10 min. 20 µl supernatant was quenched with 200 µl 0.5 mM beta-mercaptoethanol in 0.1% formic acid water. After centrifugation at 17,000 g for 10 min, the supernatant was transferred to MS vials until further analysis. The measurement of SCFAs and BCFAs are performed using the same Q Exactive PLUS hybrid quadrupole-orbitrap mass spectrometer with different column and LC setup. LC separation was on Acquity UPLC BEH C18 column (2.1 mm x 100 mm, 1.7 µm particle size, 130 Å pore size, Waters, Milford, MA) using a gradient of solvent A (water) and solvent B (methanol). Flow rate was 200 µL/min. The LC gradient was: 0 min, 10% B; 1 min, 10% B; 5 min, 30% B; 11 min 100% B; 14 min, 100% B; 14.5 min 10% B; 22 min 10% B. Autosampler temperature was 5 °C, column temperature was 60 °C and injection volume was 10 µl. Ion masses for derivatized acetate, propionate, butyrate, iso-butyrate, valeric acid, isovaleric acid, 2-methylbutyrate, 4-methylvaleric acid were 194.0571, 208.0728, 222.0884, 222.0884, 236.1041, 236.1041, 236.1041, 250.1197 in negative mode, respectively.

The ammonia derivatization method was modified from the previous reported Berthelot reaction assay (Spinelli et al., 2017b, 2017a). In brief, 20 mg tissue or 10 µl serum was extracted by using 200 µl 80% methanol. 100 µl of the metabolite extract was mixed with 100 µl Solution #1 (100 mM Phenol, 50 mg/L sodium nitroprusside) and 100 µl Solution #2 (0.38 M dibasic sodium phosphate, 125 mM NaOH, 1% sodium hypochlorite, available chlorine 10-15%). The mixture was incubated at 40°C for 30 min. Then, 100 µl reaction solution was mixed with 200 µl methanol to oversaturate the inorganic salt to quench the reaction. The final solution was centrifuged for 30 min. Then the supernatant was loaded to LC-MS for analysis. Ion mass for derivatized ammonia is 198.05605 in negative ion mode.

Protein amino acid composition was measured by acid hydrolysis. Approximately 10 mg of protein was extracted with 400 µl methanol, 200 µl chloroform and 300 µl water, followed by centrifugation at 20,000 × g for 10 min at 4 °C. The upper layer was removed. The resulting mixture was further extracted with 600 µl methanol twice and supernatant was discarded. The resulting precipitate was dried under nitrogen gas and then hydrolyzed with 250 µl 6 M hydrochloric acid incubated overnight at 115°C. After incubation, the samples were dried under nitrogen gas and reconstituted in 1 mL methanol, and the supernatant was transferred to a MS vial for analysis. Amino acid composition of the proteins used in the study are shown in Figure S4F, and such differences in protein amino acid composition do not correlate with the quantified dietary protein contribution to cecal amino acids (Figure S4G).

### Proteomics sample preparation

Proteomics samples were prepared mostly as previously described (Gupta et al., 2018; Wühr et al., 2014). Mouse cecal contents samples (10 mg each) were dissolved in 400 µl lysis buffer (6M guanidium chloride, 2% cetrimonium bromide, 5 mM dithiothreitol, 50 mM (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) (HEPES), pH 7.2). Then the sample mixture was put on ice and sonicated for 10 cycles (30 s on and 30 s off cycle, amplitude 50%) by a sonicator (Qsonica), followed by centrifugation at 20,000 × g for 20 min at 4 °C. The supernatant was taken and alkylated with 20 mM N-ethylmaleimide for 20 min at room temperature, 5 mM dithiothreitol was added to quench the excessive alkylating reagents. Proteins were purified by methanol-chloroform precipitation. The dried protein pellet was resuspended in 10 mM EPPS (N-(2-Hydroxyethyl) piperazine-N'-(3-propanesulfonic acid)) at pH 8.5 with 6 M guanidine hydrochloride. Samples were heated at 60°C for 15 min and protein concentration was determined by BCA assay (Pierce BCA Protein Assay Kit, Thermo Scientific). The protein mixture (30~50 µg) was diluted with 10 mM EPPS pH 8.5 to 2 M GuaCl and digested with 10 ng/µL LysC (Wako) at room temperature overnight. Samples were further diluted to 0.5 M GuaCl with 10 M EPPS pH 8.5 and digested with an additional 10 ng/µL LysC and 20 ng/µL sequencing grade Trypsin (Promega) at 37°C for 16 h. Samples were desalted using a SepPak cartridges (Waters) and then vacuum-dried and resuspended in 1% formic acid before mass spectrometry analysis.

### Proteomics peptide measurement

Samples were analyzed on an EASY-nLC 1200 (Thermo Fisher Scientific) HPLC coupled to an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific) with Tune version 3.3. Peptides were separated on an Aurora Series emitter column (25 cm × 75 µm ID, 1.6 µm C18) (Ionopticks, Australia) and held at 60°C during separation using an in-house built column oven over 180 min, applying nonlinear acetonitrile gradients at a constant flow rate of 350 nL/min. The Fusion Lumos was operated in data dependent mode. The survey scan was performed at a resolution setting of 120k in orbitrap, followed by MS2 duty cycle of 1.5 s. The normalized collision energy for CID MS2 experiments was set to 30%.

Solvent A consisted of 2% DMSO (LC-MS-grade, Life Technologies), 0.125% formic acid (98%+, TCI America) in water (LC-MS-grade, OmniSolv, VWR), solvent B of 80% acetonitrile (LC-MS-grade, OmniSolv, Millipore Sigma), 2% DMSO and 0.125% formic acid in water. The following 120 min-gradient with percentage of solvent B was applied at a constant flow rate of 350 nL/min after thorough equilibration of the column to 0% B: 0% – 6% in 5 min; 6 – 25% in 160 min; 25% – 100% in 10 min; 100% for 5 min. For electrospray ionization, 2.6 kV were applied between minutes 1 and 113 (or minutes 1 and 83 for fractionated samples) of the gradient through the column. To avoid carry-over of peptides, 2,2,2-trifluoroethanol (> 99% Reagent plus, Millipore Sigma) was injected in a 30 min wash between each sample.

### Proteomics data analysis

The data were analyzed using GFY software licensed from Harvard (Nusinow et al., 2020). Thermo Fisher Scientific. raw files were converted to mzXML using ReAdW.exe. MS2 spectra assignment was performed using the SEQUEST algorithm v.28 (rev. 12) by

searching the data against the combined reference proteomes for *Mus Musculus*, *Bos Taurus*, and all the abundant bacterial families detected in 16S rRNA gene amplicon sequencing (*Bacteroidaceae*, *Porphyromonadaceae*, *Prevotellaceae*, *Rikenellaceae*, *Muribaculaceae*, *Lachnospiraceae*, *Ruminococcaceae*, *Erysipelotrichaceae*, *Oscillospiraceae*, *Clostridiaceae*, *Eubacteriaceae*, *Lactobacillaceae* and *Verrucomicrobiaceae*) acquired from Uniprot on Jan 2021 (SwissProt + Trembl) along with common contaminants such as human keratins and trypsin. The target-decoy strategy was used to construct a second database of reverse sequences that were used to estimate the peptide false discovery rate (Elias and Gygi, 2007). A 20-ppm precursor ion tolerance with the requirement that both N- and C- terminal peptide ends are consistent with the protease specificities of LysC and Trypsin was used for SEQUEST searches, two missed cleavages were allowed. NEM was set as a static modification of cysteine residues (+125.047679 Da). An MS2 spectral assignment false discovery rate of 0.5% was achieved by applying the target decoy database search strategy. Linear Discriminant analysis was used for filtering with the following features: SEQUEST parameters XCorr and unique ΔXCorr, absolute peptide ion mass accuracy, peptide length and charge state. Forward peptides within three standard deviations of the theoretical m/z of the precursor were used as positive training set. All reverse peptides were used as negative training set. Linear Discriminant scores were used to sort peptides with at least seven residues and to filter with the desired cutoff. Furthermore, we performed a filtering step on the protein level by the “picked” protein FDR approach (Savitski et al., 2015). Protein redundancy was removed by assigning peptides to the minimal number of proteins which can explain all observed peptide, with above-described filtering criteria.

To quantify the intensities of all the isotopic peaks of the peptides, we used raw intensity. Missed cleavage peptides (more than one K or R in the peptide) and low signal to FT-noise peptides ( $M_0$  S/N < 20) were removed. Peptide phylogenetic assignment was performed using Unipept 4.0 (Gurdeep Singh et al., 2019), ‘Equate I and L’ and ‘Advanced missed cleavage handling’ were not selected. Only peptides that are specific at a genus level were used for further analysis.

### Quantification of newly-synthesized fraction of peptide

To determine the newly synthesized fraction of a bacterial peptide in D<sub>2</sub>O drinking water experiment, we first measured the cecal content free amino acids deuterium labeling pattern using metabolomics. Then, for each peptide, we simulated the expected isotope envelope pattern if the peptide were old, i.e., unlabeled with deuterium ( $I_{old}$ ), versus if it were newly synthesized by taking up free amino acids from the cecal content ( $I_{new}$ ).  $I_{old}$  was calculated based on the peptide’s molecular formula and <sup>13</sup>C, <sup>15</sup>N, <sup>2</sup>H, <sup>17</sup>O, <sup>18</sup>O, <sup>32</sup>S, <sup>33</sup>S and <sup>36</sup>S natural abundance.  $I_{new}$  was calculated based on the peptide’s sequence and experimentally observed labeling of the corresponding cecal free amino acids (after natural isotope correction), and the natural isotope abundance of the unlabeled atoms in the peptide’s formula. The simulation of expected peptide isotope distribution and fitting was performed using a MATLAB code: [https://github.com/xxing9703/pepMID\\_simul](https://github.com/xxing9703/pepMID_simul). Exact mass isotopic peaks with appreciable abundances were bundled by nominal mass into fraction M+0, M+1,...M+n, constituting the final simulated spectrum. A least square fit was used to find the scalar  $\theta$  that best fit the measured peptide isotopic distribution ( $I_{measured}$ ) to a linear combination of  $I_{old}$  and  $I_{new}$ :

$$I_{measured} = I_{old} \times (1 - \theta) + I_{new} \times \theta$$

The root mean square error was determined for each peptide fitting, and any fitting with a root mean square error > 1% was removed. For genus-level turnover quantification, only genera with more than two measurements were kept in the analysis, with the median value across peptides reported.

### Quantification of contribution of labeled nutrient to peptide

To determine the contribution of a <sup>13</sup>C- or <sup>15</sup>N-labeled nutrient to a bacterial peptide, similar to the above approach, we first measured the cecal content free amino acids <sup>13</sup>C- or <sup>15</sup>N-labeling using metabolomics. Then, for each peptide, we simulated the expected isotope envelope pattern if the peptide were unlabeled ( $I_{unlabeled}$ ) versus if it were synthesized from free cecal amino acids ( $I_{free}$ ). A scalar  $\gamma$  (analogous to  $\theta$  above) can then be determined by fitting the measured peptide isotope distribution ( $I_{measured}$ ) to a linear combination of  $I_{unlabeled}$  and  $I_{free}$ . Note that  $\gamma$  will exceed 1 when a bacterial genus uses a particular nutrient in excess of that nutrient contribution’s to cecal free amino acids. Because the <sup>13</sup>C- and <sup>15</sup>N-labeling patterns are simpler than the D<sub>2</sub>O labeling patterns, in lieu of carrying out this fitting, we instead determined  $\gamma$  (with the same conceptual and mathematical meaning) using simple algebraic equations.

Specifically, we measured  $\gamma$  for each peptide as follows:

$$\gamma = \frac{\varphi_{measured} - \varphi_{unlabeled}}{\varphi_{free} - \varphi_{unlabeled}}$$

where (with the exception of <sup>13</sup>C-protein feeding data, discussed immediately below)  $\varphi$  is the average number of extra neutrons in a given peptide (or simulated peptide), relative to the M+0 form. This was calculated based on the experimentally observed (or simulated, as above) fraction of M+0, M+1, M+2, and M+3, which account for > 90% of the isotopes for each peptide (with more heavily labeled forms too low abundance and noisy to contribute productively to the measurements):

$$\varphi = \frac{\sum_{i=0}^3 i \cdot M_i}{\sum_{i=0}^3 M_i}$$

For the  $^{13}\text{C}$ -protein feeding experiments, the most readily detected labeled forms involve incorporation of a single midsized U- $^{13}\text{C}$ -amino acid, which manifests as M+5 or M+6 peptide labeling. Other isotopic forms were sufficiently noisier, as to render their inclusion unhelpful. Accordingly, we calculated  $\gamma$  based on  $\varphi'$ :

$$\varphi' = \frac{M_5 + M_6}{M_0 + M_5 + M_6}$$

The above equations give nearly identical values for  $\gamma$  as fitting (as done to determine  $\theta$ ).

For genus-level measurements of feedstock contributions, only genera with more than 3 peptides measured per mouse was kept in the analysis, with the median value across peptides reported as  $\gamma_{genus}$ . Only genera that were consistently detected in proteomics, and the family of that genera detected ( $> 0.5\%$ ) in 16S rRNA gene amplicon sequencing were analyzed. The product of  $\gamma_{genus}$  and the contribution of each nutrient to cecal free amino acids ( $L_{AA\_avg \leftarrow nutrient}$ ) was used to determine the contribution of each nutrient to bacterial genus ( $f_{genus \leftarrow nutrient}$ ):

$$f_{genus \leftarrow nutrient} = \gamma_{genus} \times L_{AA\_avg \leftarrow nutrient}$$

where the contribution of each nutrient to bacterial protein pool ( $L_{AA\_avg \leftarrow Nutrient}$ ) was calculated as the average labeling across amino acids, weighted based on their abundance in that genus' protein and corrected for fraction of the nutrient interest labeled ( $T$ ):

$$L_{AA\_avg \leftarrow nutrient} = \sum f_{Cecal\_AA \leftarrow nutrient} \times w\%_{AA, bacteria} / T$$

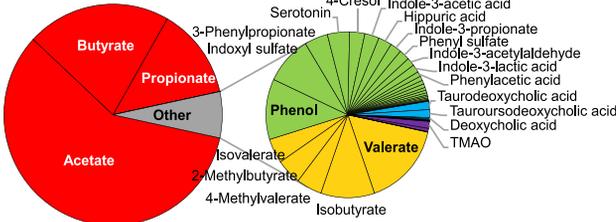
with  $w\%_{AA, bacteria}$  taken from literature (Purser and Buechler, 1966).

## QUANTIFICATION AND STATISTICAL ANALYSIS

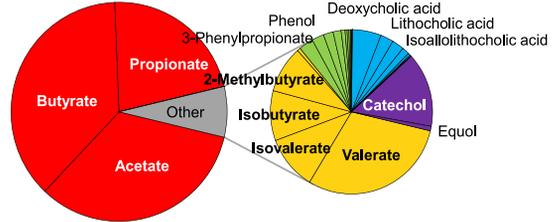
A two-tailed, unpaired student's t-test was used to calculate P values, with  $P < 0.05$  used to determine statistical significance.

# Supplemental figures

**A Microbiome metabolites in Portal Vein**



**B Microbiome metabolites in Cecal Content**



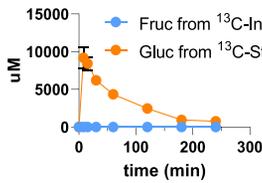
**C**

Meal	<sup>13</sup> C-Source	Source	Label
Meal 1	<sup>13</sup> C-Starch	Inulin	Protein
Meal 2	Starch	<sup>13</sup> C-Inulin	Protein
Meal 3	Starch	Inulin	<sup>13</sup> C-AA
Meal 4	Starch	Inulin	<sup>13</sup> C-protein

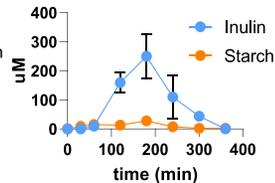


Portal blood, cecal content  
time course LC-MS

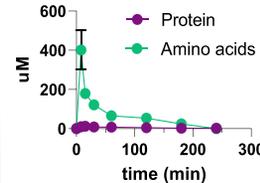
**D Portal Hexose**



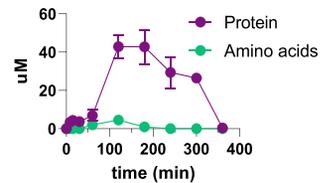
**E Portal Acetate**



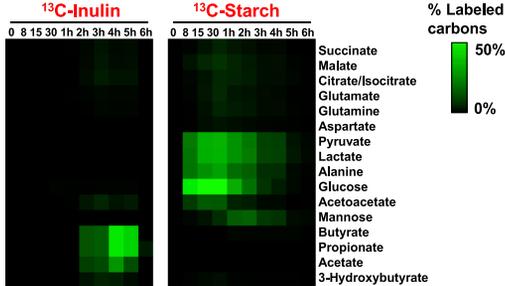
**J Portal Valine**



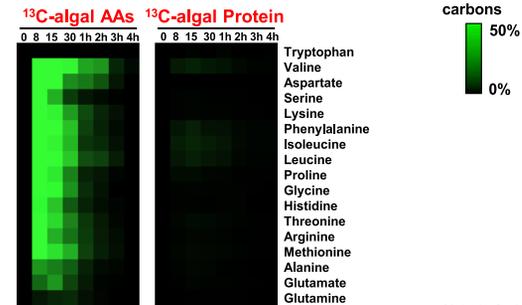
**K Portal Acetate**



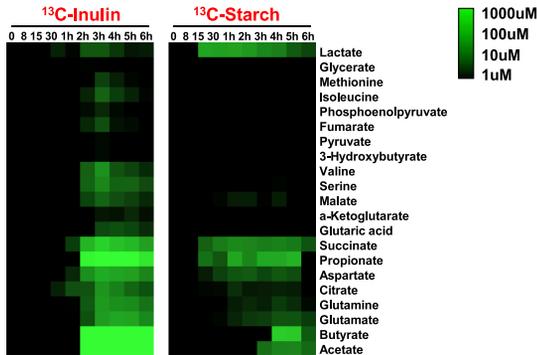
**F Portal**



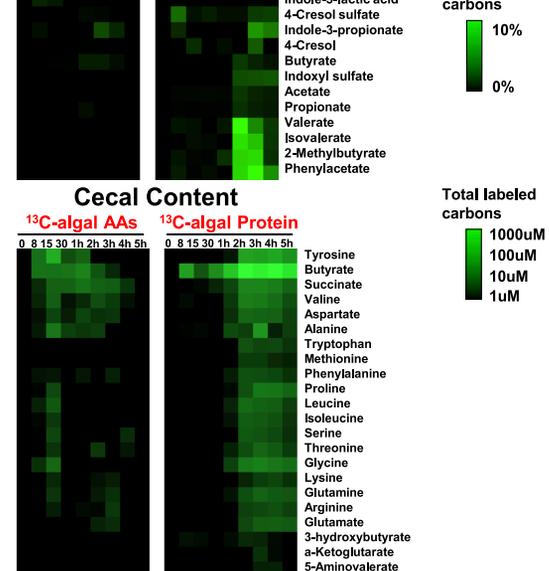
**L Portal**



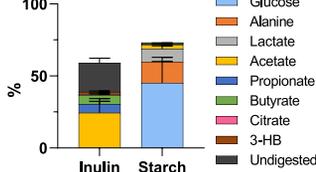
**G Cecal Content**



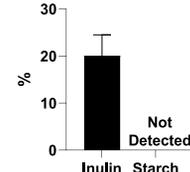
**M Cecal Content**



**H**



**I Feces**



(legend on next page)

---

**Figure S1. Microbiome consumes dietary fiber and protein, related to Table 1**

(A) Composition of the measured portal microbial metabolome. The pie charts show the relative molar abundance of different gut microbiota-associated metabolites in mice (n = 6 mice).

(B) As in (A), for cecal-content microbial metabolome (n = 6 mice).

(C) Experimental scheme. Mice received an oral gavage of 4:2:1 starch: protein (or free amino acids): inulin by weight. In each dietary condition, one component was <sup>13</sup>C-labeled. After gavage of the labeled meal, tissue and serum metabolite labeling were measured over time by LC-MS.

(D) Dietary starch feeds the host, whereas dietary inulin feeds the microbiome. The data show concentrations of labeled carbons in hexose in portal circulation (mean ± SE, n = 3 mice).

(E) As in (D), for acetate (mean ± SE, n = 3 mice).

(F) Heatmap showing the percentage of labeled carbon atoms in the indicated metabolites in portal circulation, following gavage of 4:2:1 starch: protein: inulin, with the indicated nutrient labeled. Each data point is median of n = 3 mice.

(G) Heatmap showing the molarity of total labeled carbon atoms in the indicated metabolites in cecal content, following gavage of 4:2:1 starch: protein: inulin, with the indicated nutrient labeled. Each data point is the median of n = 3 mice.

(H) Metabolic fate of inulin and starch. Stacked bars show the fraction of gavaged inulin and starch that is converted into each of the indicated metabolic products, with the undigested fraction being excreted in the feces (mean ± SE, n = 3 mice).

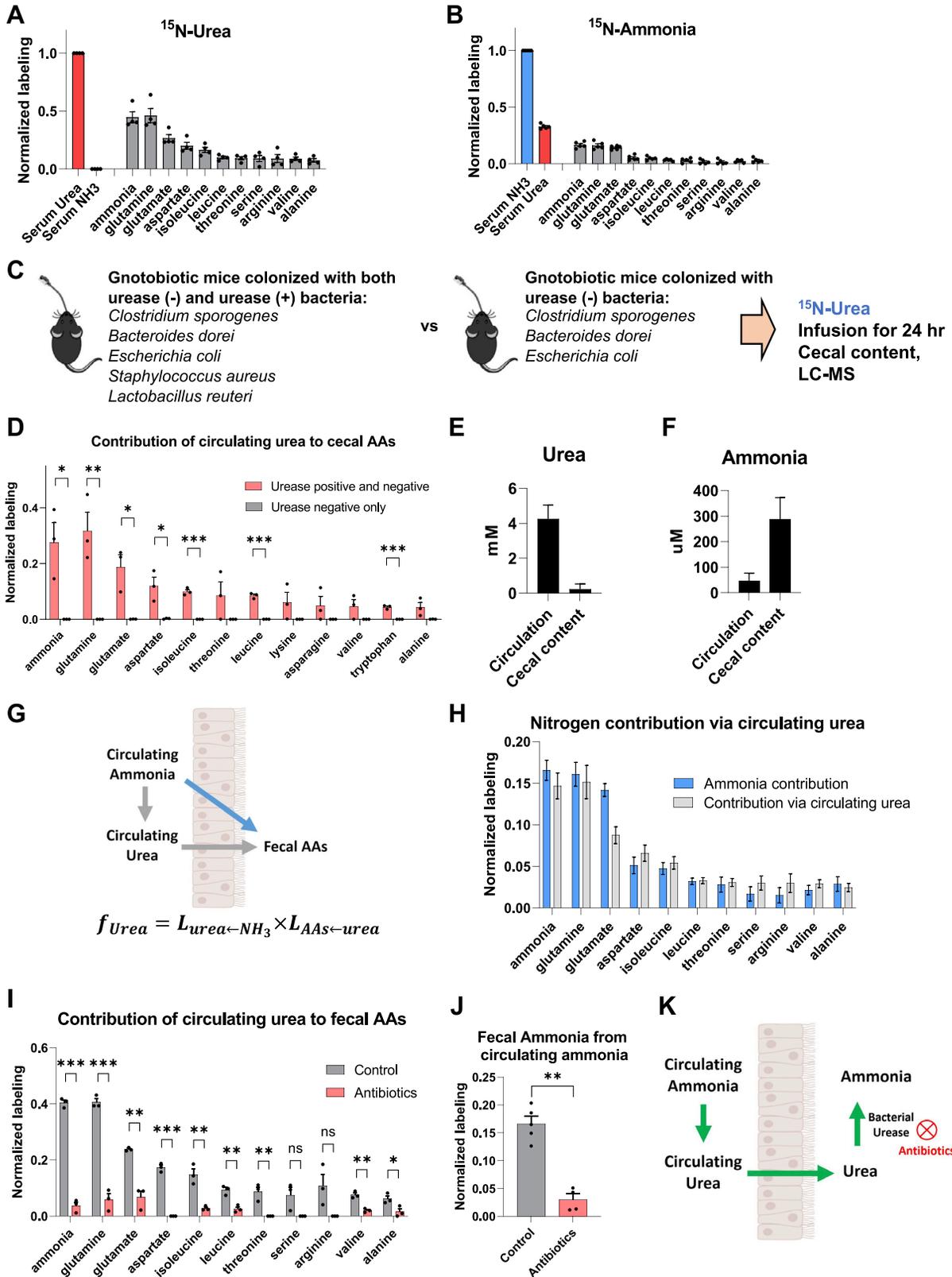
(I) Undigested carbohydrate in the feces. Graph shows the fraction of labeled hexose after cecal-content hydrolysis, 12 h following gavage as above (mean ± SE, n = 3 mice).

(J) Dietary amino acids feed the host, whereas dietary algal protein feeds the microbiome. The data show concentrations of labeled carbons in valine in portal circulation (mean ± SE, n = 3 mice).

(K) As in (J), for acetate (mean ± SE, n = 3 mice).

(L) Heatmap showing the percentage of labeled carbon atoms in the indicated amino acids in portal circulation, following gavage of 4:2:1 starch: protein (or free amino acids): inulin, with the indicated nutrient labeled. Each data point is median of n = 3 mice.

(M) Heatmap showing the molarity of total labeled carbon atoms in the indicated amino acids in cecal content, following gavage of 4:2:1 starch: protein (or free amino acids): inulin, with the indicated nutrient labeled. Each data point is median of n = 3 mice.



---

**Figure S2. Circulating ammonia contributes to microbiota metabolism via circulating urea, related to Figure 1**

(A) Normalized labeling of serum urea, ammonia, and fecal amino acids after  $^{15}\text{N}$ -urea infusion (mean  $\pm$  SE, n = 4 mice). Feces labeling fraction is normalized to serum-infused tracer (urea) labeling fraction.

(B) As in (A), for  $^{15}\text{N}$ -ammonia infusion (mean  $\pm$  SE, n = 5 mice).

(C) Experimental design. Gnotobiotic mice colonized with urease-negative bacteria (*C. sporogenes*, *B. Dorei*, and *E. coli*) or a combination of urease-negative and -positive bacteria (*C. sporogenes*, *B. dorei*, *E. coli*, *L. reuteri*, and *S. aureus*) were intravenously infused with  $^{15}\text{N}$ -urea for 24 h. Labeling of cecal-content amino acids was measured by LC-MS.

(D) Normalized labeling of cecal ammonia and amino acids after  $^{15}\text{N}$ -urea infusion in urease-negative bacteria colonized gnotobiotic mice and urease-positive and -negative bacteria colonized gnotobiotic mice (mean  $\pm$  SE, n = 3 mice). \* p < 0.05, \*\* p < 0.01, and \*\*\* p < 0.001 by two-sided Student's t test.

(E) Ammonia concentration in systemic circulation and cecal content (mean  $\pm$  SE, n = 5 mice).

(F) As in (E), for urea (mean  $\pm$  SE, n = 5 mice).

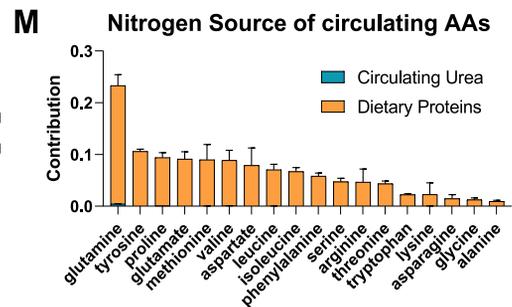
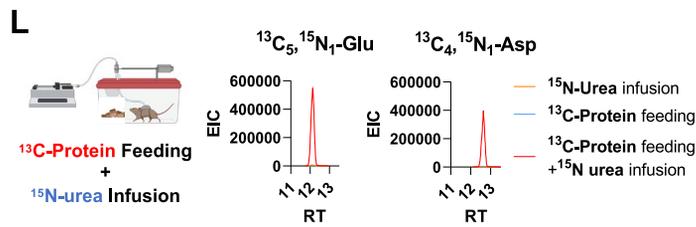
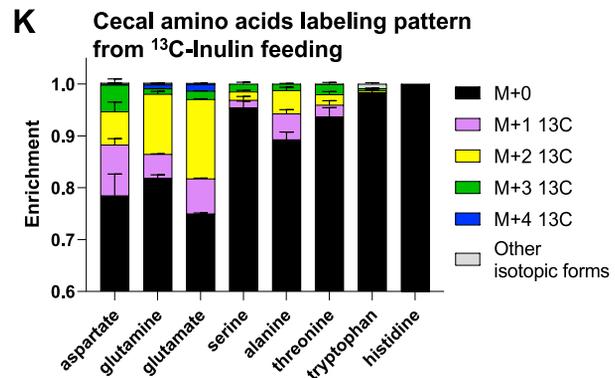
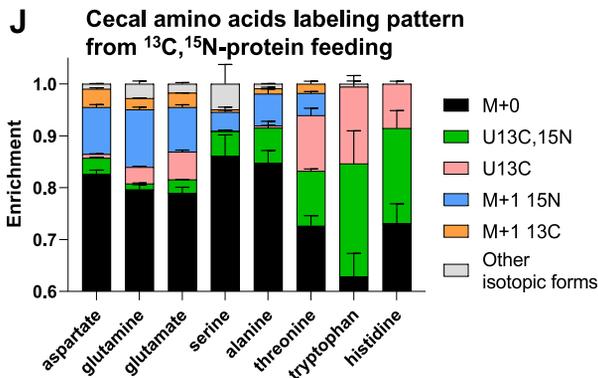
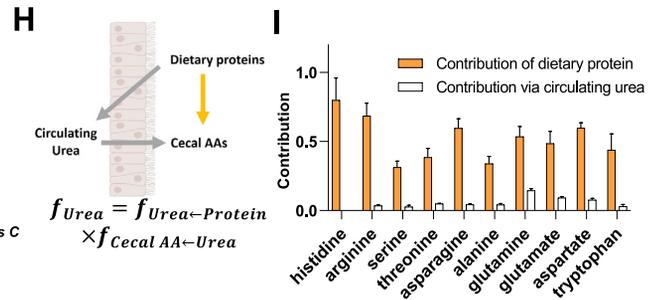
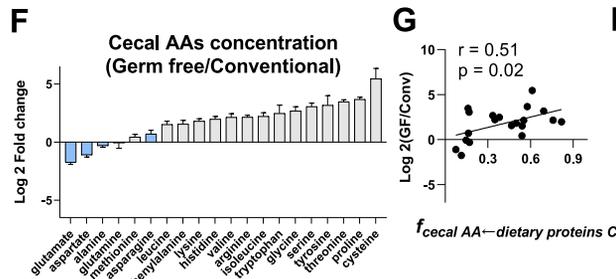
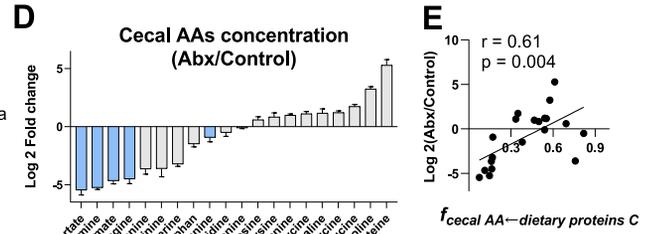
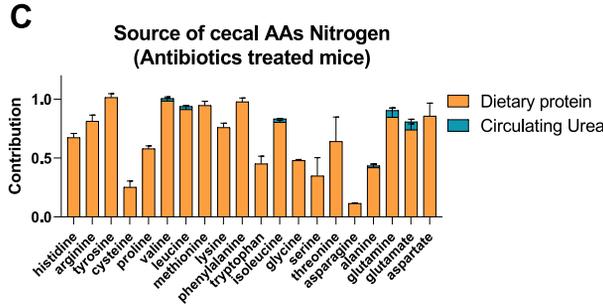
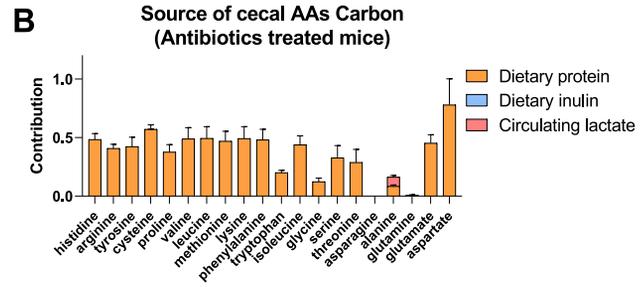
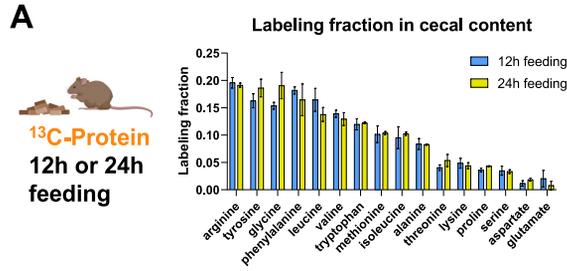
(G) Calculation of the indirect contribution from circulating ammonia to fecal amino acids via circulating urea.

(H) Contribution of circulating ammonia to fecal amino acids is quantitatively explained by ammonia's labeling of circulating urea; i.e., occurs via the indirect route shown in gray in (G) (mean  $\pm$  SE, n = 4 mice).

(I) Normalized labeling of fecal ammonia and amino acids after  $^{15}\text{N}$ -urea infusion in control and antibiotics-treated mice (mean  $\pm$  SE, n = 3 mice). \* p < 0.05, \*\* p < 0.01, and \*\*\* p < 0.001 by two-sided Student's t test.

(J) Normalized labeling of fecal ammonia after  $^{15}\text{N}$ -ammonia infusion in control and antibiotics-treated mice (mean  $\pm$  SE, n = 5 for control mice, n = 4 for antibiotics-treated mice). \*\* p < 0.01 by two-sided Student's t test.

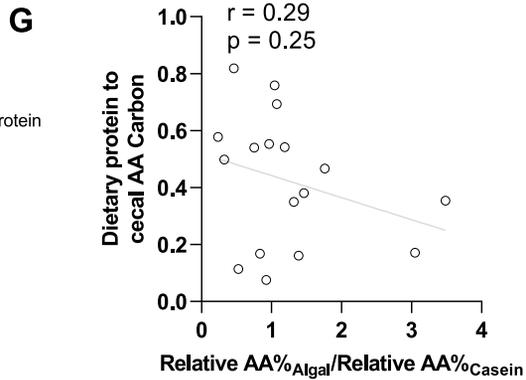
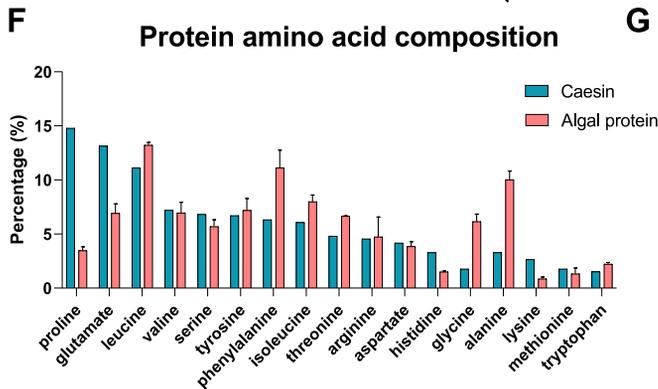
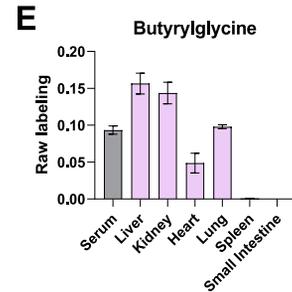
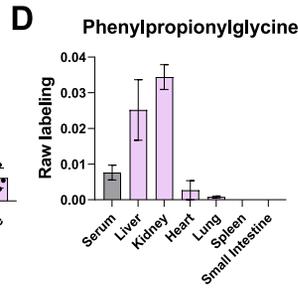
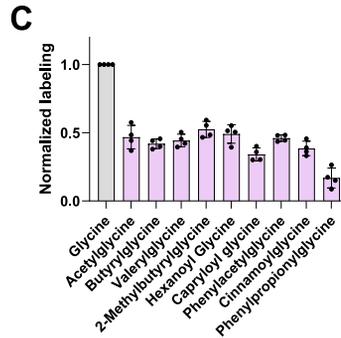
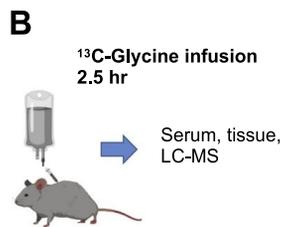
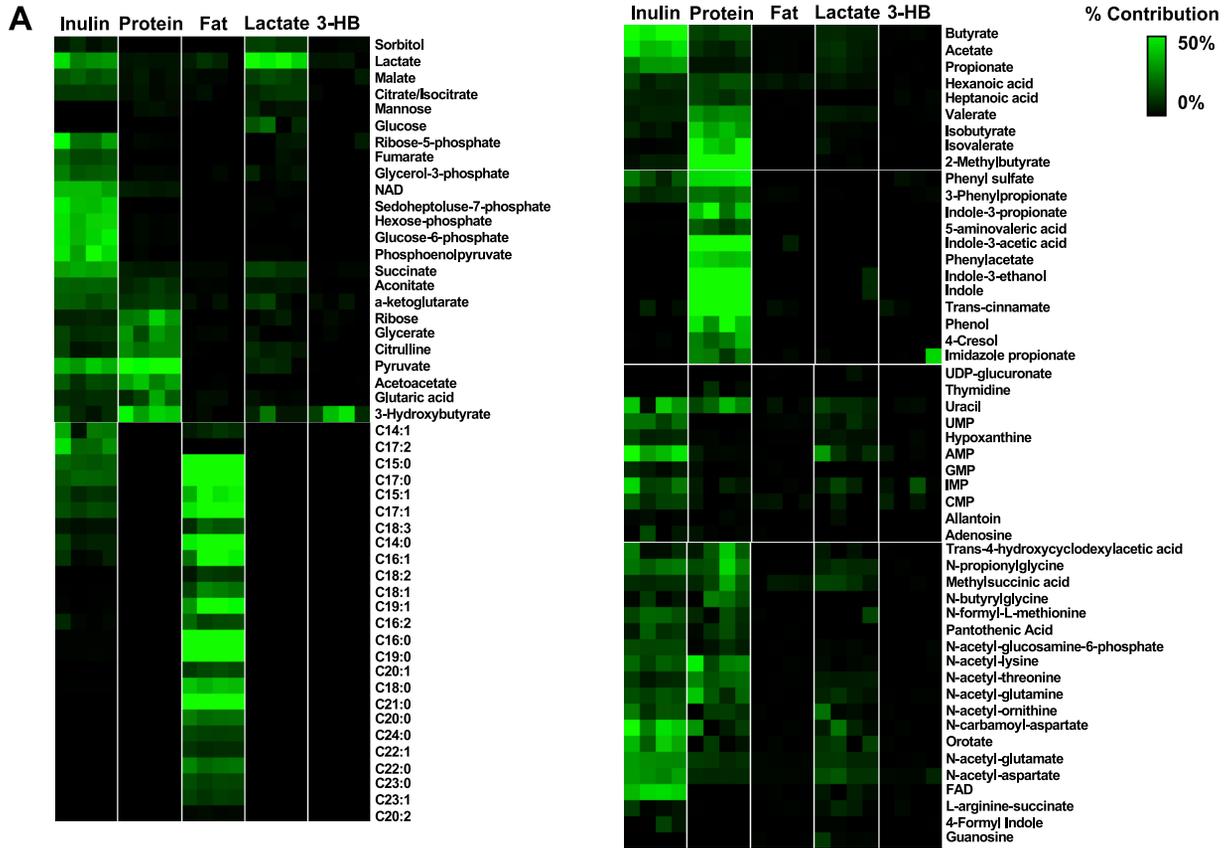
(K) Schematic of the pathway from circulating ammonia to luminal ammonia.



---

**Figure S3. Quantitative analysis of microbiome amino acid metabolism, related to Figure 2**

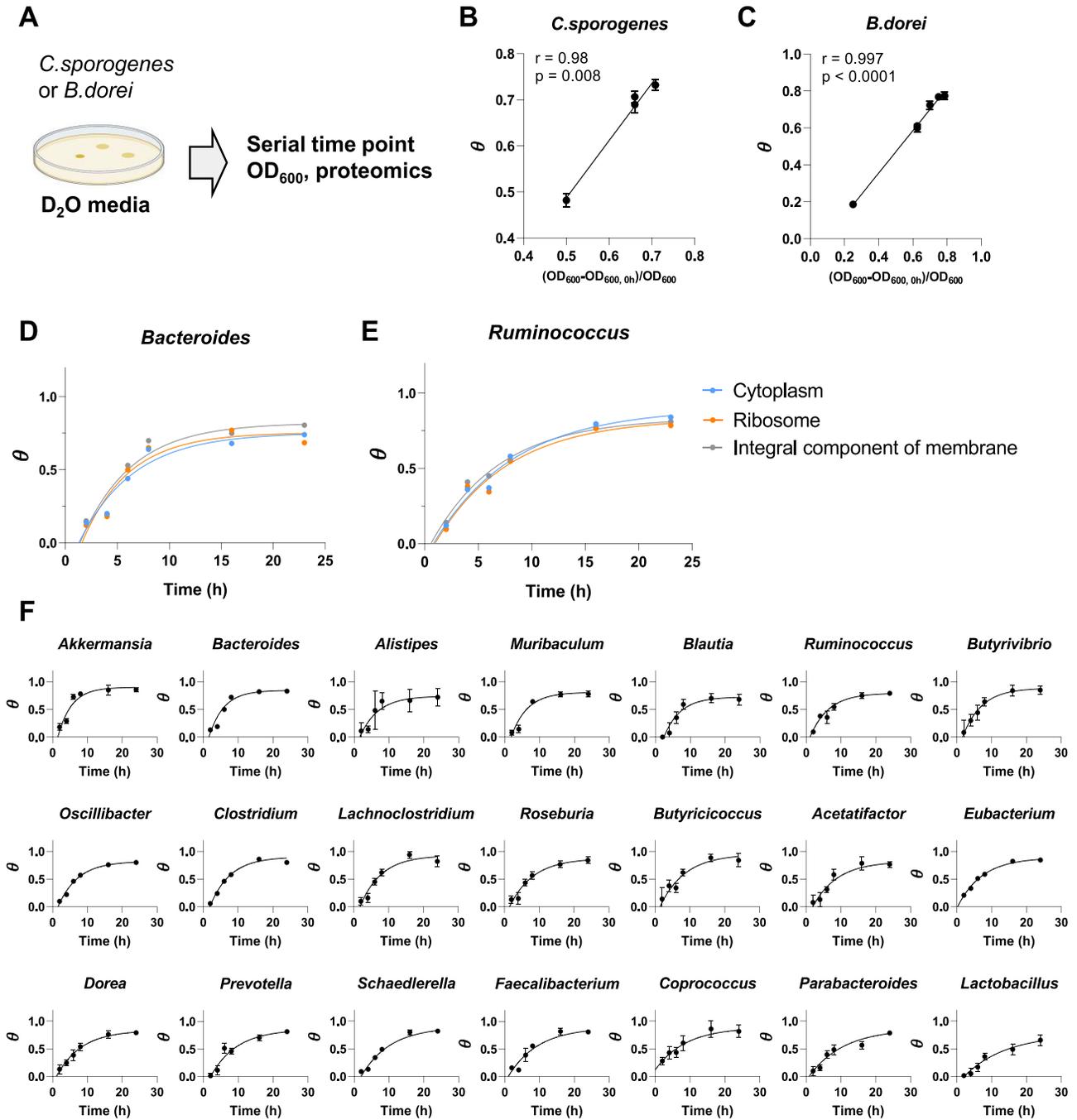
- (A) Labeling fraction of cecal amino acids from 12 or 24 h  $^{13}\text{C}$ -algal protein feeding. Mean  $\pm$  SE, n = 4 mice.
- (B) Contribution of dietary and circulating nutrients to cecal amino acid carbon in antibiotic-treated mice. Mean  $\pm$  SE, n = 3 mice.
- (C) As in (B), for nitrogen.
- (D) Cecal amino acids concentration fold change in antibiotic-treated mice relative to mock-treated mice. Mean  $\pm$  SE, n = 3. TCA-related amino acids are shown in blue bars and the others are shown in gray bars.
- (E) Cecal amino acids coming directly from diet tend to go up with antibiotic treatment. Plot shows the correlation, across amino acids (each point is an individual amino acid), between concentration fold change in antibiotic-treated mice and carbon contribution from dietary proteins based on isotope tracing (without any antibiotic treatment).
- (F) As in (D), for germ-free mice relative to conventional mice. Mean  $\pm$  SE, n = 3. TCA-related amino acids are shown in blue bars, and the others are shown in gray bars.
- (G) As in (E), for germ-free mice.
- (H) Schematic of the possibility for dietary protein nitrogen to contribute to cecal amino acid nitrogen via circulating urea.
- (I) Contribution from dietary proteins to cecal amino acid nitrogen is mostly independent of circulating urea. Mean  $\pm$  SE, n = 3.
- (J) Cecal amino acid isotopic labeling forms from  $^{13}\text{C}$ ,  $^{15}\text{N}$ -algal protein feeding. Mean  $\pm$  SE, n = 3.
- (K) Cecal amino acid isotopic labeling forms from  $^{13}\text{C}$ -inulin feeding. Mean  $\pm$  SE, n = 3.
- (L) Total ion chromatogram of [ $^{13}\text{C}_5$ ,  $^{15}\text{N}_1$ ]glutamate and [ $^{13}\text{C}_4$ ,  $^{15}\text{N}_1$ ]aspartate from simultaneous  $^{15}\text{N}$ -urea infusion and  $^{13}\text{C}$ -protein feeding (relative to the  $^{15}\text{N}$ -urea infusion or  $^{13}\text{C}$ -protein feeding, separately).
- (M) Amino acids synthesized in the gut microbiome stay in the microbiome, as urea contributes to microbiome amino acids but not host circulating amino acids. Mean  $\pm$  SE, n = 4 mice.



---

**Figure S4. Sources of cecal metabolites and acyl-glycines, related to Figure 2 and Table 1**

- (A) Heatmaps showing the contribution of dietary or circulating nutrients to cecal metabolites. For experimental design, see Figure 3. n = 4 mice.
- (B) Experimental design. Mice were intravenously infused with [U-<sup>13</sup>C] glycine for 2.5 h, and tissue and serum glycine and acyl-glycine labeling were measured.
- (C) Circulating acyl-glycines are made from circulating glycine. Mean ± SE, n = 4 mice.
- (D) Tissue phenylpropionyl-glycine labeling (normalized to circulating glycine labeling). Mean ± SE, n = 4 mice.
- (E) Tissue butyryl-glycine labeling (normalized to circulating glycine labeling). Mean ± SE, n = 4 mice.
- (F) Amino acid composition of algal protein and casein measured by acid hydrolysis. Mean ± SE, n = 3 for algal protein and n = 1 for casein.
- (G) Contribution of dietary algal protein versus casein to cecal amino acid carbon correlates poorly with relative amino acids abundances in algal protein versus casein.



**Figure S5. Single exponential fit of newly synthesized fraction of microbial peptides over time, related to Figure 4**

(A) Growth rate quantification using D<sub>2</sub>O *in vitro*. *Clostridium sporogenes* or *Bacteroides dorei* were cultured in media with 5% D<sub>2</sub>O, and bacteria growth over time was measured by either OD<sub>600</sub> or proteomics.

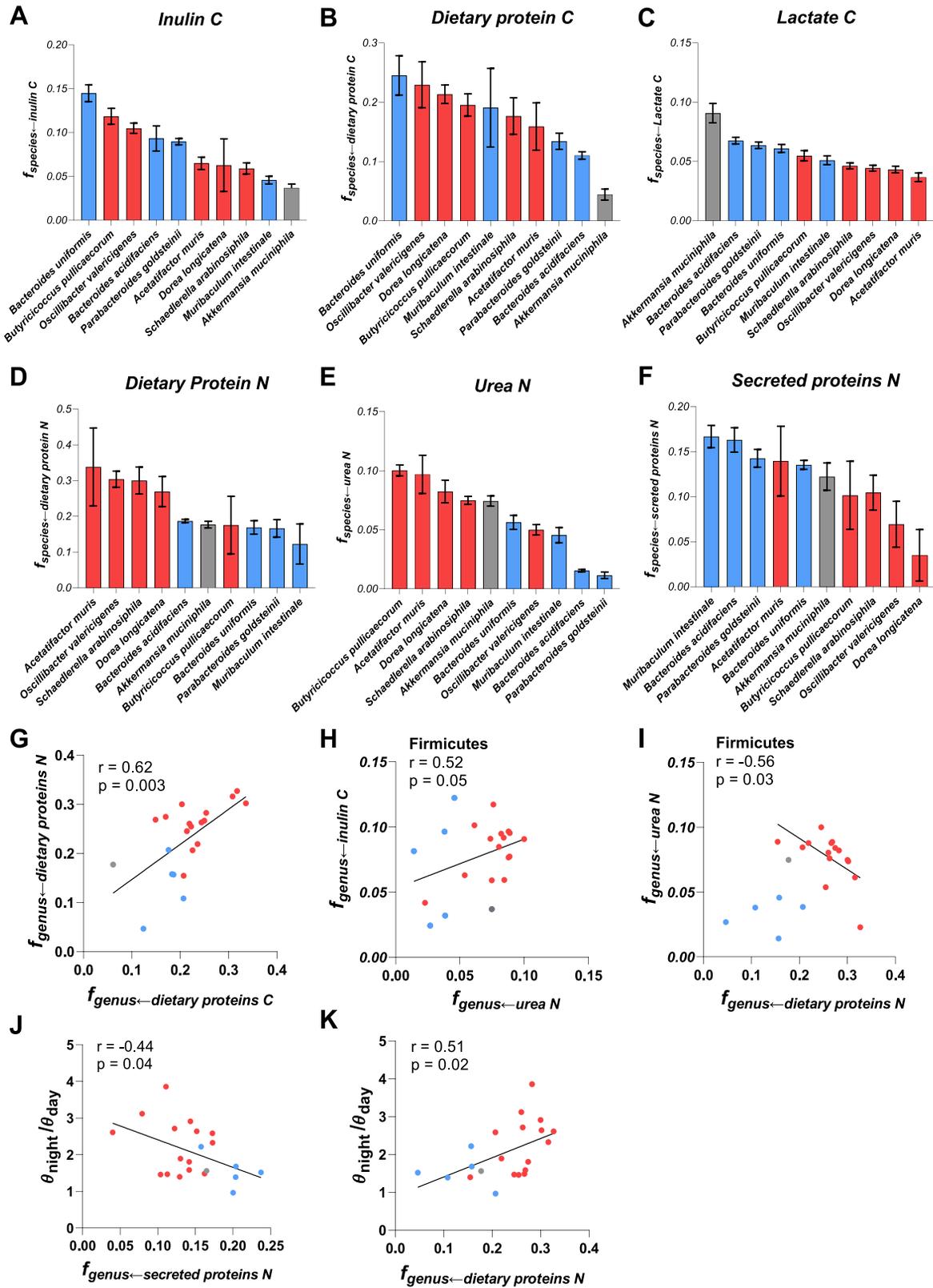
(B) Proteomics-measured growth rate mirrors the growth rate measured by OD<sub>600</sub>. Mean  $\pm$  SE,  $n > 200$  characteristic peptides for each time point.

(C) As in (B), for *Bacteroides dorei*.

(D) Different cellular compartments from the genus *Bacteroides* show similar labeling rate.

(E) As in (D), for genus *Ruminococcus*.

(F) Single exponential fit was applied to determine genus-level microbial turnover. Data are mean  $\pm$  SE,  $n = 5$  mice.

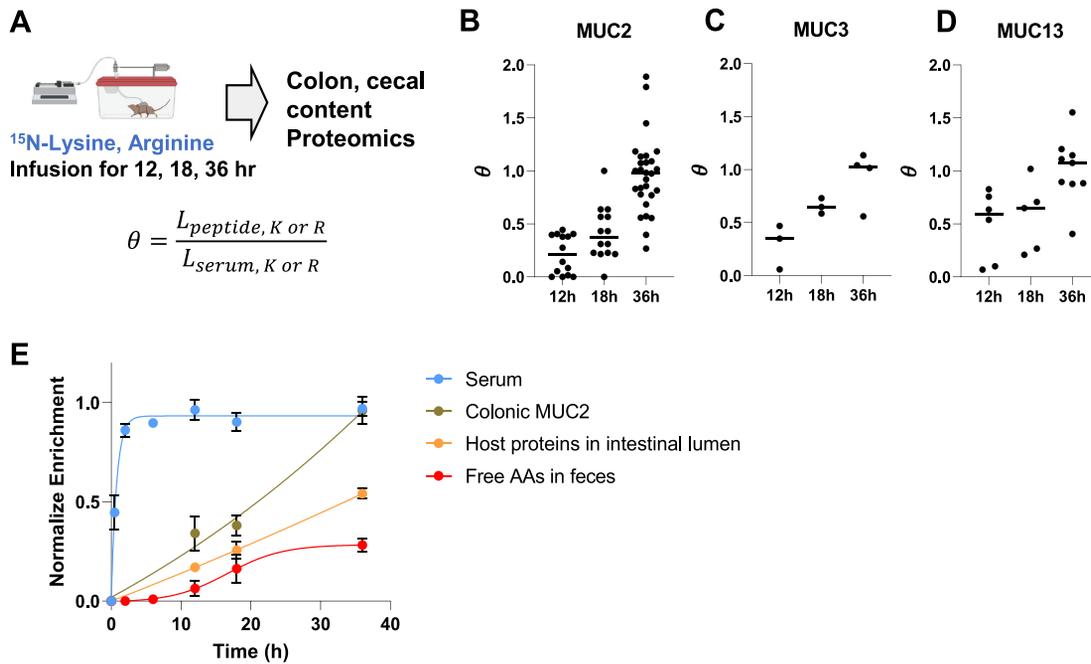


(legend on next page)

---

**Figure S6. Species-level and correlation analysis of nutrient preferences across different gut bacteria, related to Figures 5 and 6**

- (A) Carbon contribution of dietary inulin across bacterial species. Mean  $\pm$  SE, n = 4 mice.
- (B) Carbon contribution of dietary algal protein across bacterial species. Mean  $\pm$  SE, n = 6 mice.
- (C) Carbon contribution of circulating lactate across bacterial species. Mean  $\pm$  SE, n = 7 mice.
- (D) Nitrogen contribution of circulatory urea across bacterial species. Mean  $\pm$  SE, n = 6 mice.
- (E) Nitrogen contribution of dietary algal protein across bacterial species. Mean  $\pm$  SE, n = 6 mice.
- (F) Carbon contribution of secreted host proteins across bacterial species. Mean  $\pm$  SE, n = 5 mice.
- (G) Positive correlation between dietary protein nitrogen and carbon contributions across bacterial genera.
- (H) Positive correlation between inulin carbon contribution and urea nitrogen contribution in Firmicutes.
- (I) Negative correlation between urea nitrogen contribution and dietary nitrogen contribution in Firmicutes.
- (J) Negative correlation between diurnal growth rate variability and host secreted protein nitrogen contribution across bacterial genera.
- (K) Positive correlation between diurnal growth rate variability and dietary protein nitrogen contribution across bacterial genera.



**Figure S7. Host mucins are synthesized using circulating host amino acids and eventually contribute to fecal amino acids, related to Figure 6**

(A) Experimental design. Mice were infused with  $^{15}\text{N}$ -lysine and  $^{15}\text{N}$ -arginine, and mouse colon proteins were analyzed by proteomics.

(B) MUC2 labeling (multiple different MUC2 peptides).  $n = 2$  mice for 12 and 18 h,  $n = 3$  mice for 36 h.

(C) As in (B), for MUC3.

(D) As in (B), for MUC13.

(E) Passage of circulating amino acids from serum into host colonic protein and from there into host proteins and finally fecal free amino acids. Figure shows the enrichment of serum AA, colonic MUC2 proteins, host proteins in intestinal lumen, and fecal AAs normalized to the serum AA enrichment at 36 h following  $^{15}\text{N}$ -lysine and  $^{15}\text{N}$ -arginine infusions. Mean  $\pm$  SE,  $n = 2$  mice for 12 and 18 h,  $n = 3$  mice for 36 h for serum and fecal AAs.  $n = 14$  characteristic MUC2 peptides for 12 and 18 h,  $n = 28$  characteristic MUC2 peptides for 36 h.  $n > 50$  characteristic peptides for host proteins in intestinal lumen for 12, 18, and 36 h.