Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites

William R. Wikoff[a], Andrew T. Anfora[b], Jun Liu[b], Peter G. Schultz[b,1], Scott A. Lesley[b], Eric C. Peters[b], and Gary Siuzdak[a,1]

[a]Department of Molecular Biology and Center for Mass Spectrometry, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037; and [b]Genomics Institute of the Novartis Research Foundation, San Diego, CA 92121

Communicated by Steve A. Kay, University of California at San Diego, La Jolla, CA, December 19, 2008 (received for review December 12, 2008)

Although it has long been recognized that the enteric community of bacteria that inhabit the human distal intestinal track broadly impacts human health, the biochemical details that underlie these effects remain largely undefined. Here, we report a broad MS-based metabolomics study that demonstrates a surprisingly large effect of the gut "microbiome" on mammalian blood metabolites. Plasma extracts from germ-free mice were compared with samples from conventional (conv) animals by using various MS-based methods. Hundreds of features were detected in only 1 sample set, with the majority of these being unique to the conv animals, whereas ≈10% of all features observed in both sample sets showed significant changes in their relative signal intensity. Amino acid metabolites were particularly affected. For example, the bacterial-mediated production of bioactive indole-containing metabolites derived from tryptophan such as indoxyl sulfate and the antioxidant indole-3-propionic acid (IPA) was impacted. Production of IPA was shown to be completely dependent on the presence of gut microflora and could be established by colonization with the bacterium *Clostridium sporogenes*. Multiple organic acids containing phenyl groups were also greatly increased in the presence of gut microbes. A broad, drug-like phase II metabolic response of the host to metabolites generated by the microbiome was observed, suggesting that the gut microflora has a direct impact on the drug metabolism capacity of the host. Together, these results suggest a significant interplay between bacterial and mammalian metabolism.

The human body is colonized by hundreds of trillions of microbes, which collectively possess hundreds of times as many genes as coded in the human genome. The combined genetic potential of the endogenous flora is referred to as the "microbiome" (1), and typically results in a mutualistic relationship between microbe and host. Normal host activities, including the processing of nutrients and the regulation of the immune system, are affected by the intestinal microbiome (2, 3), and the microbiome has been implicated in the pathogenesis of diseases such as nonalcoholic steatohepatitis (4), allergy (5), the formation of gallstones (6), and inflammatory bowel disease (7, 8). The composition of the gut microbiome is highly variable (9), and its diversity can be significantly affected by alterations in diet (2, 10) or antibiotic use (11). Alternatively, probiotic therapy is the attempt to alter the extant gut microbial environment through the ingestion of live consumable cultures of beneficial bacteria (12). Differences in commensal microflora are likely to impact human health and disease through any number of ways reflective of the complex nature of the microbiome itself.

Historically, classical microbiology methods including the isolation and culture of individual bacterial species associated with the gut were used to study microbial colonization of higher organisms. More recently, metagenomic techniques have been used to characterize both the composition and the potential physiological effects of entire microbial communities without having to culture individual community members. For example, genes for specific metabolic pathways such as amino acid and glycan metabolism were found to be overrepresented in the microbiome of the distal gut, reinforcing the notion that human metabolism is an amalgamation of microbial and human attributes (13). Also, obese mice were found to extract energy from their food more efficiently compared with lean counterparts due to alterations in the composition of their gut microflora that resulted in an increased complement of genes for polysaccharide metabolism (10). It has also been observed that bile salt hydrolase encoding genes are enriched in the gut microbiome, and that enteric bacteria carry out a wide range of bile acid modifications (6, 14). These metagenomic studies suggest that the metabolites derived from this diverse microbial community can have a direct role in human health and disease. To date, metabolomics-based investigations of aspects of the impact of the microbiome on mammalian biochemistry have detailed changes in the levels of well-documented metabolites based primarily on NMR-based analysis and subsequent multivariate statistics of unfractionated samples, such as urine, gut tissue, or cecum extracts (15–17). Recently, this same group reported the multicompartmental effects of the microbiome on murine metabolism by using NMR-based analysis of urine and tissue extracts from both conventional (conv) and germ-free (GF) mice (18). Although extremely powerful, these studies provide only limited opportunity for the discovery of differences in unexpected or lower level metabolites.

Here, we demonstrate the large effect of the microbiome on mammalian plasma biochemistry. Specifically, a broad, untargeted, mass spectrometry-based profiling of serum from GF and conv mice demonstrates that a significantly large number of chemical species found in systemic circulation arise because of the presence of the microbiome, whereas at least 10% of all detectable endogenous circulating serum metabolites vary in concentration by at least 50% between the 2 mouse lines. Several microbiome-affected molecules identified in the serum of conv mice are either potentially harmful (uremic toxins) or beneficial (antioxidant) to the host. Also, we observed a broad, drug-like phase II metabolic response of the host to species generated by the microbiome, as evidenced by the exclusive presence of numerous sulfated, glycine-conjugated, and glucuronide adducts in the serum of conv mice.

Results and Discussion

Mass Spectrometry Reveals That the Microbiome Has a Broad Effect on Mouse Plasma Biochemistry. Mass spectrometry has become an increasing powerful tool for metabolomics studies due to its wide dynamic range, reproducible quantitative capabilities, and its ability to analyze samples of significant molecular complexity (19). However, the vast chemical diversity of the known metabolome creates numerous challenges for any MS-based "global" profiling effort, including analytical biases introduced by sample preparation, separation, and ionization techniques. Given these challenges, we sought to obtain an initial assessment of the extent of the effect of

Author contributions: W.R.W., A.T.A., P.G.S., S.A.L., E.C.P., and G.S. designed research; W.R.W., A.T.A., and J.L. performed research; W.R.W., A.T.A., J.L., and E.C.P. analyzed data; and W.R.W., P.G.S., S.A.L., E.C.P., and G.S. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence may be addressed. E-mail: schultz@gnf.org or siuzdak@scripps.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0812874106/DCSupplemental.

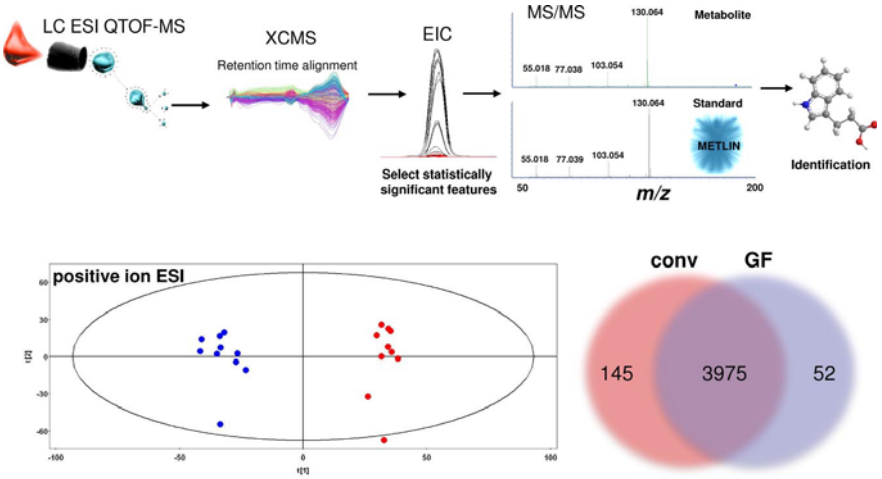

Fig. 1. The microbiome exerts a broad effect on mammalian biochemistry. (*Upper*) The experimental workflow for the metabolomics studies performed. (*Lower*) Major differences are observed in the profiles of GF and conv plasma obtained employing + ESI. PCA analysis of the profiling data from XCMS (*Left*) shows an excellent separation between sample groups. As shown in the Venn diagram (*Right*), most species are observed in both GF and conv samples, but a significant number are unique to a given sample set, with ≈3-fold more ions unique to the conv sample.

the microbiome on mammalian metabolism by analyzing uniformly processed samples employing various ionization methods, because studies have shown that such multiple ionization mode strategies increase both the number and types of analytes observed (20). We also chose to analyze plasma samples, because this biofluid provides an excellent opportunity to assess the extent of interplay between bacterial metabolic and systemic human pathways.

In Fig. 1 we show the experimental workflow for the metabolomics studies performed. Ten plasma samples each from both individual GF and conv mice were subjected to methanol protein precipitation and subsequent centrifugation, because this technique was previously reported to be the most effective, straightforward, and reproducible extraction method for such samples (21). The resulting supernatants were each analyzed by 4 different methods; reversed-phase liquid chromatography coupled to (*i*) electrospray ionization (ESI) in the positive mode (+), (*ii*) ESI in the negative mode (−), (*iii*) atmospheric pressure chemical ionization (APCI) in the + mode, and (*iv*) gas chromatography after derivatization with trimethylsilyl chloride. The raw data from the individual analyses were subjected to nonlinear data alignment (22), and the results were analyzed for both global changes by using multivariate statistics to determine group separation, as well as univariate statistics to evaluate the number and percentage of features that are either unique or vary significantly between the 2 sample sets. As seen in Fig. 1*B*, principle component analysis (plotting principle component 1 vs. principle component 2) shows an excellent separation of the GF and conv sample groups analyzed under + ESI conditions. Similar separations were also seen for the − ESI, + APCI, and GC-based analyses (Fig. S1). The majority of individual features detected were present in both the GF and conv samples. However, hundreds of features were detected in only 1 sample set, with the majority of these being unique to the conv samples. Also, as summarized in Table 1, ≈10% of all features detected in both sample sets under any of the analysis conditions showed significant changes in their relative signal intensity (defined as a ≥1.5-fold change with a $P \leq 0.0001$). It should be noted that a given molecule may be represented by several different features, such as the naturally occurring components of its isotopic cluster or nonspecific adduct ions. Nevertheless, these comparative results show that the microbiome exerts a dramatic effect on mammalian plasma biochemistry.

We next turned our focus toward the identification of those features that showed significant differences in level between the 2 sample sets. This task is by far the most laborious part of any untargeted MS-based metabolomics profiling study, involving the integrated analysis of accurate mass measurements, tandem MS fragmentation patterns, and literature/database searches to produce candidate structures that still require experimental verification. For the purposes of this article, an observed feature was not designated as unequivocally identified unless an authentic sample of the proposed species was independently analyzed in our laboratories and found to exhibit the same chromatographic and mass spectrometric properties as the feature in question. By comparison, highly probable identification status was assigned when the experimentally determined properties of a given feature matched those reported in the literature for a particular molecule, or were fully consistent with those reported for a close analog. Although this process remains a work in progress, the identifications produced to date as listed in Table 2 (and further detailed in Table S1) already exhibit several striking features. For example, despite the use of a single sample preparation method, we identified various chemically distinct compounds ranging from fatty acids to highly polar charged molecules, along with several groups of functionally related molecules.

The Microbiome Affects the Diversity of Indole-Containing Compounds in Serum. Multiple features that differed significantly between conv and GF mice correspond to compounds bearing indole moieties. For example, the plasma concentrations of tryptophan and *N*-acetyltryptophan in conv mice were 40 and 60% lower than in their respective GF counterparts. A subset of enteric bacteria express tryptophanase, which converts tryptophan to indole, pyru-

Table 1. Summary of comparative metabolomics profiles using different MS-based methodologies

	Total peaks	Total significant changed peaks*	Significant change conv > GF	Significant change GF > conv	Unique to conv	Unique to GF
LC ESI +	4,172	466, 11.2%	318, 68%	148, 32%	145	52
LC ESI −	4,132	271, 6.6%	168, 62%	103, 38%	100	61
LC APCI +	1,584	152, 9.5%	96, 63%	56, 37%	27	5
GCMS +	1,803	253, 14.0%	225, 88.9%	28, 11.1%	28	12

*Defined as having a fold-change ≥ 1.5 and $P \leq 0.0001$.

Table 2. Summary of metabolites identified that showed significant differences in concentration between the two sample sets

Metabolite	Fold change	*P* value	Compound and metabolism class
Indole derivatives			
Tryptophan	1.7, GF	8.42×10^{-12}	—
N-acetyltryptophan	2.4, GF	3.56×10^{-4}	—
Indoxyl sulfate	conv*	1.34×10^{-7}	Phase I: hydroxyl; phase II: sulfate
Serotonin	2.8, conv	1.27×10^{-10}	—
IPA	conv*	7.69×10^{-7}	Bacterial conjugation
Phenyl derivatives			
Phenylalanine	1.05, GF	0.3	—
Tyrosine	1.44, GF	1.14×10^{-4}	—
Hippuric acid	17.4, conv	1.98×10^{-9}	Phase II: glycine
Phenylacetylglycine	3.8, conv	4.70×10^{-8}	Phase II: glycine
Phenyl sulfate	conv*	9.85×10^{-7}	Phase II: sulfate
p-Cresol sulfate	conv*	0.002	Phase II: sulfate
Phenylpropionylglycine	conv*	3.07×10^{-7}	Phase II: glycine
Cinnamoylglycine	conv*	2.93×10^{-7}	Phase II: glycine
Flavones			
Equol sulfate	conv*	1.44×10^{-5}	Phase II: sulfate
Methyl equol sulfate	conv*	2.18×10^{-6}	Phase II: sulfate
Others			
Urate	1.99, conv	1.51×10^{-6}	—
Creatinine	1.08, conv	0.071	—
Dihydroxyquinoline glucuronide	conv*	7.64×10^{-6}	Phase II: glucuronide
12-Hydroxy-5*Z*,8*Z*,10*E*,14*Z*,17*Z*-eicosapentaenoic acid	4.0, conv	8.20×10^{-5}	Fatty acid
3-Carboxy-4-methyl-5-pentyl-2-furanpropionic acid glucuronide	3.4, conv	1.37×10^{-6}	Phase II: glucuronide

Fold change equals the fold difference in concentration observed between conv and GF samples, with the group indicating which is higher. Names in italics refer to compound identifications that are highly probable, as defined in the text.
*Compound observed only in a single group (conv or GF).

vate, and ammonia. Given that tryptophanase activity in conv mice has been shown to increase nearly 2-fold after dosing with tryptophan (23), it is highly likely that the decreased concentrations of tryptophan and *N*-acetyltryptophan observed in conv mouse serum resulted from the metabolism of dietary tryptophan by the direct action of enteric bacteria.

Interestingly, serotonin plasma levels were 2.8-fold higher in conv animals. However, this increase in serotonin is difficult to attribute to a direct metabolic transformation mediated by gut bacteria, because the production of serotonin by characterized enteric bacterial species has not been described. Despite the well-documented role of serotonin as a neurotransmitter, Enterochromaffin cells in the gut are the largest source of serotonin production in the body, and this molecule has been implicated in gastrointestinal pathologies such as irritable bowel syndrome and Crohn's disease (24). These findings suggest that the increased plasma serotonin levels observed could indirectly result from an as yet undefined host microbe interaction.

As seen in Fig. 2, other indole-containing molecules were also affected by the microflora. For example, indoxyl sulfate (indican) was identified only in the serum of conv mice. This molecule, a known nephrotoxin that accumulates in the blood of patients suffering from chronic kidney failure (25), arises from hepatic transformation of the bacterial metabolite indole. Because tryptophanase activity derives from only a subset of enteric bacteria, non-indole-producing bacteria, such as various *Bifidobacterium* species, have been administered as a test probiotic to dialysis patients to decrease their plasma levels of indoxyl sulfate (26). Conversely, a different set of enteric bacteria has been implicated in the metabolic transformation of indole to indole-3-propionic acid (IPA) (27). IPA, also identified only in the plasma of conv mice, has been shown to be a powerful antioxidant (28), and is currently being investigated as a possible treatment for Alzheimer's disease (29, 30).

Although the presence of IPA in mammals has long been ascribed in the literature to bacterial metabolic processes, this conclusion was based on either the production of IPA in ex vivo cultures of individual bacterial species (31) or observed decreases in IPA levels in animals after administration of antibiotics (32). In our own survey of IPA production by representative members of the intestinal flora, only *Clostridium sporogenes* was found to produce IPA in culture (Table S2). Based on these results, individual GF mice were intentionally colonized with *C. sporogenes* strain ATCC 15579, and blood samples were taken at several intervals after colonization. IPA was undetectable in the samples taken shortly after introduction of the microbes, and was first observed in the serum 5 days after colonization, reaching plateau values comparable with conv mice by day 10. These colonization studies demonstrate that the introduction of enteric bacteria capable of IPA production in vivo into the gastrointestinal tract is sufficient to introduce IPA into the bloodstream of the host. Also, other GF animals were injected i.p. with either IPA (at 10, 20, or 40 mg/kg) or sterile PBS vehicle, and their serum concentrations of IPA were measured over time. As seen in Table S3, the high serum levels of IPA observed 1 h after injection decreased more than 90% within 5 h, showing that IPA is rapidly cleared from the blood, and that its presence in the serum of conv animals must result from continuous production from 1 or more bacterial species associated with the mammalian gut. The ability to measure the metabolic profiles of animals selectively colonized with individual species or even complex communities of defined bacterial populations promises to be a powerful new tool for deconvoluting the contributions of the various species comprising the microbiome and for dissecting complex microbial-mammalian metabolic interactions

Sulfation Is Observed as a Prevalent Host Response to Bacterial Metabolites. Serum metabolites from GF and conv mice differed significantly in their level of sulfation. For example, both phenyl and *p*-cresol sulfate were present only in conv animals (Fig. S2). These molecules likely arose from the sulfation of direct bacterial metabolites of tyrosine (33, 34), the plasma concentration of which was 1.4-fold higher in GF animals.

Sulfation is a well-studied phase II drug metabolism mechanism used by the body to facilitate the removal of hydrophobic endogenous species and xenobiotics. Given the identification of several sulfated molecules found only in conv animals, we decided to investigate the extent to which this human chemical detoxification mechanism targets microbiome affected vs. endogenous metabolites. One MS-based methodology for the detection of putative sulfated molecules employs a triple quadrupole mass spectrometer to scan for species that exhibit a neutral loss of 80 *m/z* when analyzed under negative ionization conditions (35). Thus, pooled serum samples from GF and conv mice were subjected to this more targeted analysis. As seen in Fig. 3, the conv sample resulted in nearly double the number of peaks (36) compared with its GF counterpart (23). Importantly, several of the peaks unique to the conv sample were later confirmed to be the already identified phenyl (peak a), indoxyl (peak b), and *p*-cresol (peak c) sulfates; thus, demonstrating the potential of this targeted scanning technique to quickly and effectively detect potential sulfated bacterial metabolites.

Next, we sought to confirm whether other unique features resulting from the targeted scanning technique were also sulfated molecules. Thus, equol sulfate was identified as being present only in the conv sample, whereas a second conv-specific feature was identified with high confidence as methyl equol sulfate (Fig. S3).

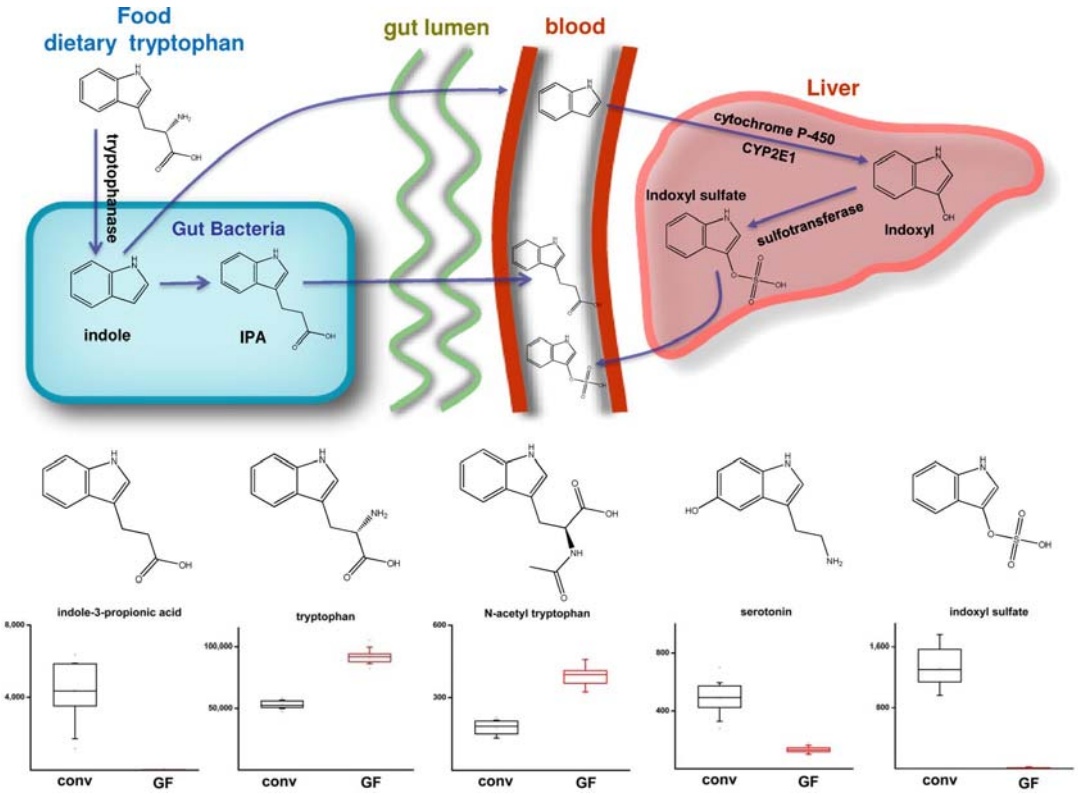

Fig. 2. The diversity of indole-containing compounds in serum is affected by the microbiome. (*Upper*) Various indole-containing molecules arise only in the presence of the microbiome and ultimately enter the plasma by means of different routes. (*Lower*) Differences in the plasma levels of indole-containing compounds attributed to the action of the microbiome. The integrated signal intensities plotted on the *y* axes are reduced by a factor of 1,000.

Equol is itself a reported metabolite of the dietary isoflavone dihydrodaidzein. Purported to have the potential to lower the incidence of prostate cancer (37), equol has been shown to arise exclusively as a result of bacterial metabolic processes (38), although only a small percentage of humans harbor the gut bacteria that affect this biotransformation (39). It is important to note that neither equol sulfate nor its methyl ester was listed in any metabolite database, again emphasizing the degree of complexity involving the microbiome that remains to be explored.

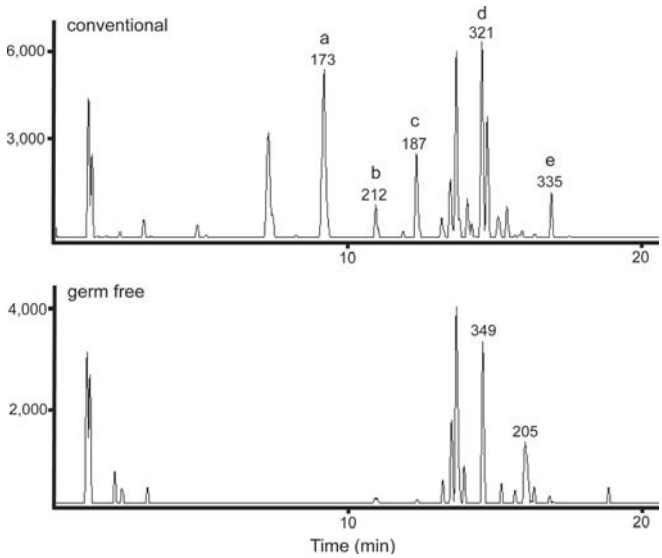

Fig. 3. Sulfate profiling based on constant neutral loss scanning of 80 *m/z* in − ESI mode for conv (*Upper*) and GF (*Lower*) pooled plasma samples. The *m/z* values for several species are listed above their respective peak. Identified sulfates include: phenyl sulfate (a), indoxyl sulfate (b), *p*-cresol sulfate (c), equol sulfate (d), and methyl equol sulfate (e).

Other Phase II Metabolic Products Are also Greatly Elevated in Response to the Microbiome. We next sought to determine whether other bacterial metabolites were modified by mammalian chemical detoxification processes. Indeed, numerous products of another prevalent phase II drug metabolism mechanism, glycine conjugation, were found to be either exclusively present or present at much higher concentrations in conv animals. As shown in Fig. 4, the plasma concentration of hippuric acid was 17-fold higher in conv samples compared with their GF counterparts. Hippuric acid is produced by the conjugation of benzoic acid with glycine, a reaction that occurs not only in the liver (40) but also directly in the intestine (41) and kidney (42). Also, it has been reported that hippuric acid levels increase when animals are fed increased levels of phenylalanine (43). However, the concentration of circulating phenylalanine was not observed to differ significantly between the 2 sample sets, suggesting that unlike the presence of indican or IPA in conv mice, the observed increased levels of hippuric acid may not simply result from the removal of some of the dietary supply of an essential amino acid by the direct action of enteric bacteria. These findings again illustrate the broad interplay of mammalian and enteric bacterial metabolism.

In Fig. 4 we indicate several additional glycine conjugates identified exclusively in the serum of conv animals. For example, cinnamoylglycine was identified as being present only in conv samples despite not being listed in any metabolite database, illustrating the fact that our knowledge of the metabolome even in biofluids as well-studied as plasma is far from complete. Also, phenylpropionylglycine was also detected exclusively in conv sam-

BIOCHEMISTRY

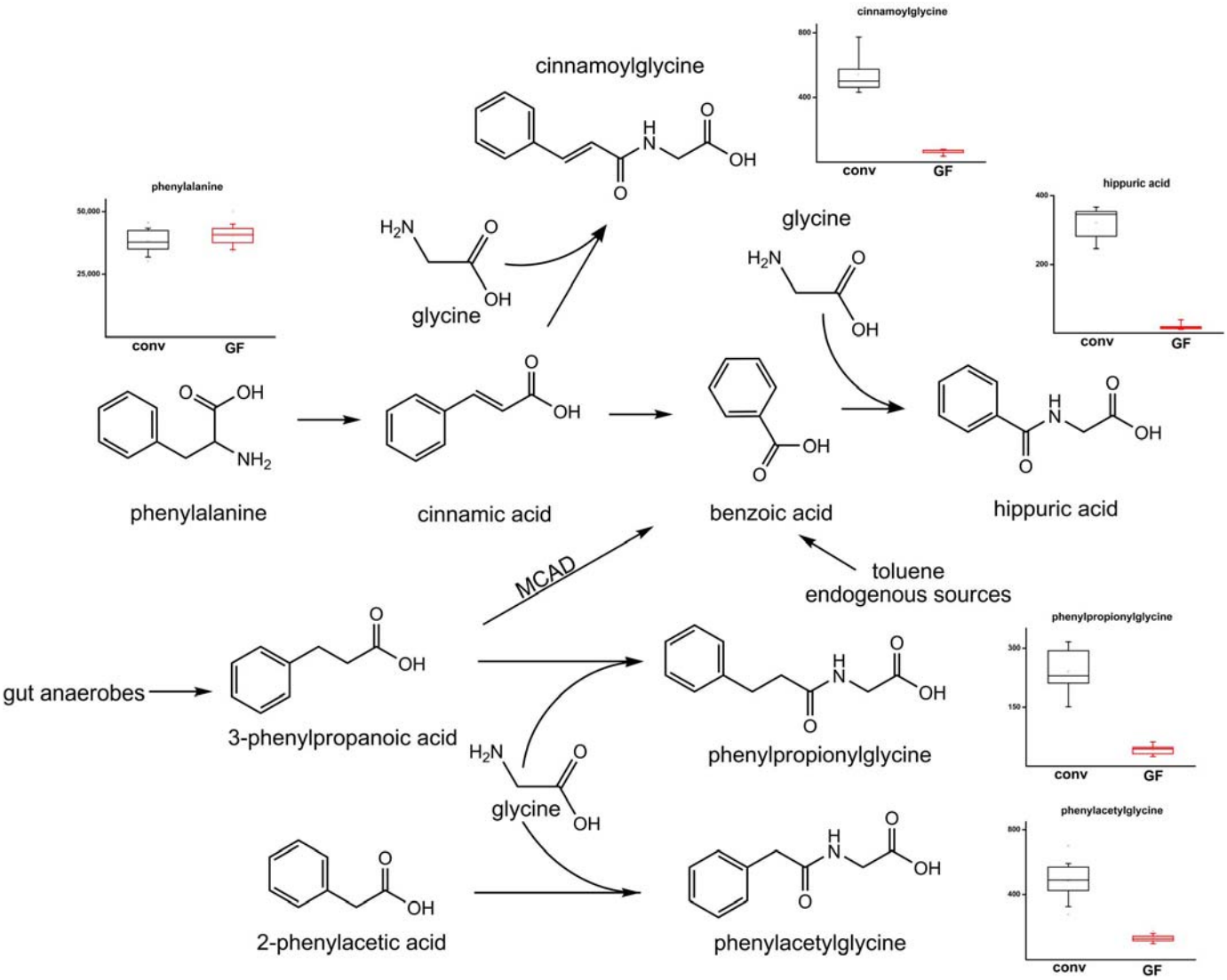

Fig. 4. Differences in the plasma levels of various glycine-conjugated compounds attributed to the action of the microbiome. Potential metabolic pathways leading to the formation of hippuric acid as well as 3 other glycine conjugates arising from the presence of the microbiome are provided. The integrated signal intensities plotted on the *y* axes are reduced by a factor of 1,000.

ples. This molecule almost certainly arose from the direct conjugation of glycine to phenylpropionic acid, a known metabolic product of anaerobic bacteria (36). However, in humans, phenylpropionic acid is instead normally converted to benzoic acid by the action of the β-oxidizing enzyme medium chain acyl-CoA dehydrogenase (MCAD) (44). Thus, phenylpropionylglycine is detected only in the urine of MCAD-deficient patients, and has been used as a biomarker for the diagnosis of this condition (45). These results illustrate the specificity of host–bacteria mutualism, and suggest the likelihood of coevolution between symbionts and their hosts.

Several products of yet another mammalian chemical detoxification mechanism, glucuronidation, were also elevated in the presence of bacterial metabolism. For example, a glucuronide of 3-carboxy-4-methyl-5-pentyl-2-furanpropionic acid, a known uremic retention solute shown to inhibit the binding of salicylate and 5,5-diphenylhydantoin to albumin (46), was identified exclusively in the serum of conv animals. The greatly increased concentrations of this furan dicarboxylic acid, as well as other classic uremic retention solutes such as indoxyl sulfate, *p*-cresol sulfate, urate, and hippuric acid in the blood of these animals, point to the large burden that the metabolic processes of enteric bacteria impose on the chemical detoxification mechanism of their hosts.

Conclusion

In summary, we have demonstrated a significant effect of the microbiome on mammalian blood metabolites by applying an untargeted MS-based metabolomics approach to a germ-free mouse model system. Numerous circulating molecules were determined to arise exclusively in the presence of gut microflora, whereas ≈10% of commonly observed features differed significantly in concentration between GF and conv animals. Several pathways including the metabolic processing of indole-containing molecules were seen to particularly interact with the microbiome. Numerous molecules apparently resulting from phase II drug-like chemical processing of microbial metabolites were also significantly elevated. Such broad metabolomics profiling studies combined with the selective colonization of germ-free animals with defined bacterial population promise to be a powerful tool for the deconvolution of the complex interplay between mammalian and bacterial metabolic processes.

Materials and Methods

Serum Sample Collection and Preperation. Blood from 10 conv and 10 GF male Swiss Webster mice age 8–10 weeks was ordered from Taconic Farms. Both sets of animals were maintained on autoclaved NIH-31 chow. Blood was collected by

retro-orbital bleed into heparinized tubes, spun down at 3,000 g for 10 min, and the entire serum content of each animal was frozen immediately at −80 °C and shipped on dry ice. Metabolites were extracted from plasma with methanol. We added 4 volumes of cold methanol to 50 μL of plasma, vortexed, and incubated at −20 °C for 1 h. Samples were centrifuged 10 min at 14,000 × *g*, the supernatant was collected, and the centrifugation was repeated. The supernatant was dried in a SpeedVac and resuspended in 50 μL of 95:5 water:acetonitrile, and clarified for 5 min at 14,000 × *g*.

Metabolomics Profiling with ESI. The HPLC system consisted of a degasser, capillary pump, binary pump, and autosampler (1100 series, Agilent); 3 μL of extracted plasma maintained at 4 °C in a thermostated autosampler was applied to a capillary reversed-phase column (Zorbax C18 SB-300, Agilent) with dimensions 300-μm ID × 150-mm length. The flow rate was 4 μL/min with solvent A composed of water + 0.1% formic acid and solvent B composed of acetonitrile + 0.1% formic acid in positive ion mode. For negative ion mode data, 5 mM ammonium acetate was substituted for the formic acid in solvents A and B. The gradient consisted of 5% B for 5 min, followed by a gradient to 95% B over 45 min, hold at 95% B for 5 min, and a reequilibration at 5% B for 10 min.

To reduce systematic error associated with instrumental drift, samples were run in an order that alternated between wild type and knockout. Data were collected in + and − ESI mode in separate runs on a TOF (Agilent 6210) operated in full scan mode from 100 to 1,000 *m/z*. The capillary voltage was 3,500 (+) or 2,000 V (−) with a scan rate of 0.5 scan per second; the nebulizer flow rate was 12 L/hr.

Data Analysis and Statistics. Data in instrument specific format (.wiff) were converted to common data format (.cdf) files. The program XCMS (22) was used for nonlinear alignment of the data in the time domain and automatic integration and extraction of the peak intensities. Accurate masses of features representing significant differences were searched against the METLIN, KEGG, HMDB, and LIPIDMAPS databases. SIMCA-P (Umetrics) was used for multivariate statistical calculations and plotting. Differences between wild type and GF plasma were evaluated for individual metabolites by using a 2-tailed *t* test, assuming unequal variance (Welch's *t* test), calculated by using Excel. Statistical plots were calculated by using Origin version 6.1.

Compound Identification. The identity of compounds was confirmed by LC/MS/MS by using a QTOF (model 6510, Agilent). Pooled plasma extracts were made from the same set of conv or GF samples, and the experiment was repeated with the identical chromatography conditions on the TOF, with the exceptions that the column was 2.1-mm ID (SB-C18, Zorbax), the flow rate was 200 μL/min, and the nebulizer flow rate was 30 L/hr. Ions were targeted for collision-induced dissociation (CID) fragmentation on the fly based on the previously determined accurate mass and retention time. The exact retention time was determined at the higher flow rate by manually comparing the pooled extracted ion chromatograms from the 2 groups. The fragmentation patterns for plasma unknowns and authentic model compounds were compared after an initial optimization of the collision energy.

Other. See *SI Materials and Methods*.

ACKNOWLEDGMENTS. This work was supported by the U.S. Department of Energy Office of Science under contracts DE-FG02-07ER64325 and DE-AC02-05CH11231 through the Genomics: GTL Research Program (G.S.), National Institutes of Health Grant P30 MH062261 (G.S.), and the Novartis Research Foundation. This is manuscript 19923 of The Scripps Research Institute.

1. Backhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI (2005) Host-bacterial mutualism in the human intestine. *Science* 307:1915–1920.
2. Turnbaugh PJ, Backhed F, Fulton L, Gordon JI (2008) Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host Microbe* 3:213–223.
3. Mazmanian SK, Liu CH, Tzianabos AO, Kasper DL (2005) An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* 122:107–118.
4. Dumas ME, *et al.* (2006) Metabolic profiling reveals a contribution of gut microbiota to fatty liver phenotype in insulin-resistant mice. *Proc Natl Acad Sci USA* 103:12511–12516.
5. Noverr MC, Noggle RM, Toews GB, Huffnagle GB (2004) Role of antibiotics and fungal microbiota in driving pulmonary allergic responses. *Infect Immun* 72:4996–5003.
6. Ridlon JM, Kang DJ, Hylemon PB (2006) Bile salt biotransformations by human intestinal bacteria. *J Lipid Res* 47:241–259.
7. Alverdy JC, Chang EB (2008) The re-emerging role of the intestinal microflora in critical illness and inflammation: Why the gut hypothesis of sepsis syndrome will not go away. *J Leukoc Biol* 83:461–466.
8. Frank DN, *et al.* (2007) Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci USA* 104:13780–13785.
9. Turnbaugh PJ, *et al.* (2007) The human microbiome project. *Nature* 449:804–810.
10. Turnbaugh PJ, *et al.* (2006) An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444:1027–1031.
11. Jernberg C, Lofmark S, Edlund C, Jansson JK (2007) Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. *Isme J* 1:56–66.
12. Hord NG (2008) Eukaryotic-microbiota crosstalk: Potential mechanisms for health benefits of prebiotics and probiotics. *Annu Rev Nutr* 28:215–231.
13. Gill SR, *et al.* (2006) Metagenomic analysis of the human distal gut microbiome. *Science* 312:1355–1359.
14. Jones BV, Begley M, Hill C, Gahan CG, Marchesi JR (2008) Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome. *Proc Natl Acad Sci USA* 105:13580–13585.
15. Wang Y, *et al.* (2004) Metabonomic investigations in mice infected with Schistosoma mansoni: An approach for biomarker identification. *Proc Natl Acad Sci USA* 101:12676–12681.
16. Martin FP, *et al.* (2007) Effects of probiotic Lactobacillus paracasei treatment on the host gut tissue metabolic profiles probed via magic-angle-spinning NMR spectroscopy. *J Proteome Res* 6:1471–1481.
17. Yap IK, *et al.* (2008) Metabonomic and microbiological analysis of the dynamic effect of vancomycin-induced gut microbiota modification in the mouse. *J Proteome Res* 7:3718–3728.
18. Claus SP, *et al.* (2008) Systemic multicompartmental effects of the gut microbiome on mouse metabolic phenotypes. *Mol Syst Biol* 4:219.
19. Want EJ, Nordstrom A, Morita H, Siuzdak G (2007) From exogenous to endogenous: The inevitable imprint of mass spectrometry in metabolomics. *J Proteome Res* 6:459–468.
20. Nordstrom A, Want E, Northen T, Lehtio J, Siuzdak G (2008) Multiple ionization mass spectrometry strategy used to reveal the complexity of metabolomics. *Anal Chem* 80:421–429.
21. Want EJ, *et al.* (2006) Solvent-dependent metabolite distribution, clustering, and protein extraction for serum profiling with mass spectrometry. *Anal Chem* 78:743–752.
22. Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G (2006) XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem* 78:779–787.
23. Botsford JL, Demoss RD (1972) Escherichia coli tryptophanase in the enteric environment. *J Bacteriol* 109:74–80.
24. Costedio MM, Hyman N, Mawe GM (2007) Serotonin and its role in colonic function and in gastrointestinal disorders. *Dis Colon Rectum* 50:376–388.
25. Deguchi T, *et al.* (2002) Major role of organic anion transporter 3 in the transport of indoxyl sulfate in the kidney. *Kidney Int* 61:1760–1768.
26. Takayama F, Taki K, Niwa T (2003) Bifidobacterium in gastro-resistant seamless capsule reduces serum levels of indoxyl sulfate in patients on hemodialysis. *Am J Kidney Dis* 41(Suppl 1):S142–S145.
27. Smith EA, Macfarlane GT (1997) Formation of phenolic and indolic compounds by anaerobic bacteria in the human large intestine. *Microb Ecol* 33:180–188.
28. Karbownik M, *et al.* (2001) Indole-3-propionic acid, a melatonin-related molecule, protects hepatic microsomal membranes from iron-induced oxidative damage: Relevance to cancer reduction. *J Cell Biochem* 81:507–513.
29. Chyan YJ, *et al.* (1999) Potent neuroprotective properties against the Alzheimer beta-amyloid by an endogenous melatonin-related indole structure, indole-3-propionic acid. *J Biol Chem* 274:21937–21942.
30. Bendheim PE, *et al* (2002) Development of indole-3-propionic acid (OXIGON) for Alzheimer's disease. *J Mol Neurosci* 19:213–217.
31. Elsden SR, Hilton MG, Waller JM (1976) The end products of the metabolism of aromatic amino acids by clostridia. *Arch Microbiol* 107:283–288.
32. Young SN, Anderson GM, Gauthier S, Purdy WC (1980) The origin of indoleacetic acid and indolepropionic acid in rat and human cerebrospinal fluid. *J Neurochem* 34:1087–1092.
33. Smith EA, Macfarlane GT (1996) Enumeration of human colonic bacteria producing phenolic and indolic compounds: Effects of pH, carbohydrate availability and retention time on dissimilatory aromatic amino acid metabolism. *J Appl Bacteriol* 81:288–302.
34. Bone E, Tamm A, Hill M (1976) The production of urinary phenols by gut bacteria and their possible role in the causation of large bowel cancer. *Am J Clin Nutr* 29:1448–1454.
35. Yi L, Dratter J, Wang C, Tunge JA, Desaire H (2006) Identification of sulfation sites of metabolites and prediction of the compounds' biological effects. *Anal Bioanal Chem* 386:666–674.
36. Lambert PW, Roos BA (1980) Paired-ion reversed-phase high-performance liquid chromatography of human and rat calcitonin. *J Chromatogr* 198:293–299.
37. Akaza H, *et al.* (2004) Comparisons of percent equol producers between prostate cancer patients and controls: Case-controlled studies of isoflavones in Japanese, Korean and American residents. *Jpn J Clin Oncol* 34:86–89.
38. Wang XL, Hur HG, Lee JH, Kim KT, Kim SI (2005) Enantioselective synthesis of S-equol from dihydrodaidzein by a newly isolated anaerobic human intestinal bacterium. *Appl Environ Microbiol* 71:214–219.
39. Frankenfeld CL, *et al.* (2005) High concordance of daidzein-metabolizing phenotypes in individuals measured 1 to 3 years apart. *Br J Nutr* 94:873–876.
40. Chiba M, Poon K, Hollands J, Pang KS (1994) Glycine conjugation activity of benzoic acid and its acinar localization in the perfused rat liver. *J Pharmacol Exp Ther* 268:409–416.
41. Strahl NR, Barr WH (1971) Intestinal drug absorption and metabolism. 3. Glycine conjugation and accumulation of benzoic acid in rat intestinal tissue. *J Pharm Sci* 60:278–281.
42. Poon K, Pang KS (1995) Benzoic acid glycine conjugation in the isolated perfused rat kidney. *Drug Metab Dispos* 23:255–260.
43. Armstrong MD, Chao FC, Parker VJ, Wall PE (1955) Endogenous formation of hippuric acid. *Proc Soc Exp Biol Med* 90:675–679.
44. Rinaldo P, O'Shea JJ, Welch RD, Tanaka K (1990) The enzymatic basis for the dehydrogenation of 3-phenylpropionic acid: In vitro reaction of 3-phenylpropionyl-CoA with various acyl-CoA dehydrogenases. *Pediatr Res* 27:501–507.
45. Rinaldo P, O'Shea JJ, Welch RD, Tanaka K (1990) Diagnosis of medium chain acyl-CoA dehydrogenase deficiency by stable isotope dilution analysis of urinary acylglycines: Retrospective and prospective studies, and comparison of its accuracy to acylcarnitine identification by FAB/mass spectrometry. *Prog Clin Biol Res* 321:411–418.
46. Niwa T, Takeda N, Maeda K, Shibata M, Tatematsu A (1988) Accumulation of furancarboxylic acids in uremic serum as inhibitors of drug binding. *Clin Chim Acta* 173:127–138.

Supporting Information

Wikoff *et al.* 10.1073/pnas.0812874106

SI Materials and Methods

GC/MS Profiling. We dried 8 μL of methanol extracted plasma by using a SpeedVac in a glass vial, and 75 μL of BSTFA + 1% TMCS (Pierce) was added, capped, and incubated at 37 °C for 1 h. This mixture was transferred to a GC autosampler vial with a 100-μL glass insert; 2.5 μL was injected without a split into an Agilent 6890 GC/MS instrument equipped with a HP5-MS column (J&W Scientific/Agilent) with a length of 30 m, ID of 0.25 mm, and film thickness of 0.25 μm. The injector port temperature was 290 °C and the transfer line temperature was 280 °C. The flow rate was 1.2 mL/min with a total run time of 27.5 min. The temperature program was 50 °C for 5 min, followed by a gradient of 5 °C per minute to 300 °C, followed by a hold at 300 °C for 10 min for a total run time of 55 min per sample. A 20-min blank run was included between each sample. The mass spectrometer was scanned from m/z 50 to 700.

APCI Profiling. For APCI-TOF profiling, 5 μL of methanol-extracted plasma was injected onto a SB-C18 Zorbax column (Agilent) 2.1-mm ID × 150-mm length at a flow rate of 250 μL/min. Mobile phase A was H_2O + 0.1% formic acid, and B was MeOH + 0.1% formic acid, with a gradient from 10% to 95% B from 2 to 30 min, and to 10% B from 33 to 40 min. A standard APCI source was used with a capillary voltage of 4,000 V, 35 psig nebulizer, 6.0 L/min drying gas, a vaporizer temp of 325 °C with a gas temperature of 350 °C and a corona discharge of 6 μA, and fragmentor of 125 V. The TOF was scanned from m/z 50 to 500. The injection order was alternated between the 2 groups of samples, and a 20-min gradient wash was run between each sample. Samples were profiled by using APCI positive mode with the TOF (Agilent, model 6210).

Microbial Cultures to Assess Indole-3-Propionic Acid (IPA) Production. Microbial strains were grown under anaerobic conditions according to ATCC guidelines until the culture achieved an OD_{600} of ≈1.0 or maximal density if the microbe could not achieve OD_{600} of ≈1.0 in broth culture. See Table S2 for a list of strains used in the study; 1 mL of culture was filtered with a 0.2-μm filter and stored at −80 °C. Samples were brought to room temperature, and a 50-μL aliquot was methanol extracted and prepared for LC/MS as were the serum samples. Samples were examined for detectable quantities of IPA. IPA was detected in all *Clostridium sporogenes* isolates under all growth conditions screened. *C. sporogenes* ATCC 15579 was selected for further experimentation due to the availability of its genomic sequence.

Mouse Colonization Experiments. All experiments involving live animals were approved by the institutional review board (IACUC) and conducted in an AALAS-accredited facility.

We aseptically introduced 12 female GF Swiss Webster mice age 6–8 week (Taconic) into semirigid isolators (Charles River Labs), were maintained on irradiated Pico-Vac Lab Rodent Diet (Purina), and given autoclaved water ad libitum. Nine mice were given a 100-μL suspension of *C. sporogenes* (ATCC 15579) produced from an overnight TSB culture diluted 1:100 in LB containing 40% glycerol. Two animals were removed from the isolator at 0, 5, and 10 days postinoculation, and were anesthetized with isoflurane, 50 μL of blood was collected retro-orbitally into clay capped capillary tubes, and, while under anesthesia, animals were euthanized by cervical dislocation. Also, 3 animals were treated in a similar manner 24 h postinoculation. Three animals were maintained as GF for 10 days and had blood drawn as above to serve as controls. Serum was separated from blood cells, methanol extracted, and analyzed as described previously.

IPA Clearance Studies. All experiments involving live animals were approved by the institutional review board (IACUC) and conducted in an AALAS-accredited facility.

We ordered 20 GF Swiss Webster female mice aged 6–8 weeks from Taconic Farms. Animals were removed from the GF shipper 24 h after arrival and placed in a biosafety cabinet that had been decontaminated with 3% Nolvasan. The mice were randomly divided into four groups (n = 5 animals per group) and administered 10, 20, 40 mg/kg IPA (Acros Organics) or sterile PBS vehicle by i.p. injection. Animals were anesthetized with isoflurane (Henry Schein), and had 50 μL of blood drawn retro-orbitally at 0, 2, and 6 h postinjection. Serum was separated from blood cells by centrifugation of capillary tubes for 2 min in a BD Adams Autocrit Ultra 3 Centrifuge, tubes were broken above the blood cell-serum interface, and the serum layer was collected into capped tubes. The serum was then methanol extracted and analyzed for the presence of IPA. All experiments involving live animals were approved by the institutional review board (IACUC) and conducted in an AALAS-accredited facility.

IPA Quantitation in Mouse Serum. Samples were prepared as described above. LC/MS data were recorded by using an Agilent 6520 Accurate-Mass Q-TOF LC/MS system with dual ESI source and an Agilent 1200 HPLC system; 10 μL was injected onto column (Agilent Zorbax 300SB-C18, 5 μm, 150 × 2.0 mm). Eluent A was 0.1% (vol/vol) formic acid in water; eluent B was 0.1% (vol/vol) formic acid in acetonitrile. The gradient consisted of 5% B for 5 min, followed by a gradient to 95% B >45 min, hold at 95% B for 5 min, and a reequilibration at 5% B for 10 min. The flow rate was 200 μL/min. The mass spectrometric data were collected in profile mode by using − ESI. The gas temperature was 350 °C. The drying gas flow rate was 5 L/min, and the pressure of the nitrogen nebulizer gas was 30 psig. The capillary, fragmentor, skimmer, and OCT1 RF Vpp voltages were 3,500 V (−), 150 V, 65 V, and 750 V, respectively. A calibration curve was made by running IPA standard solutions of 50, 200, 500, and 1,000 ng/mL.

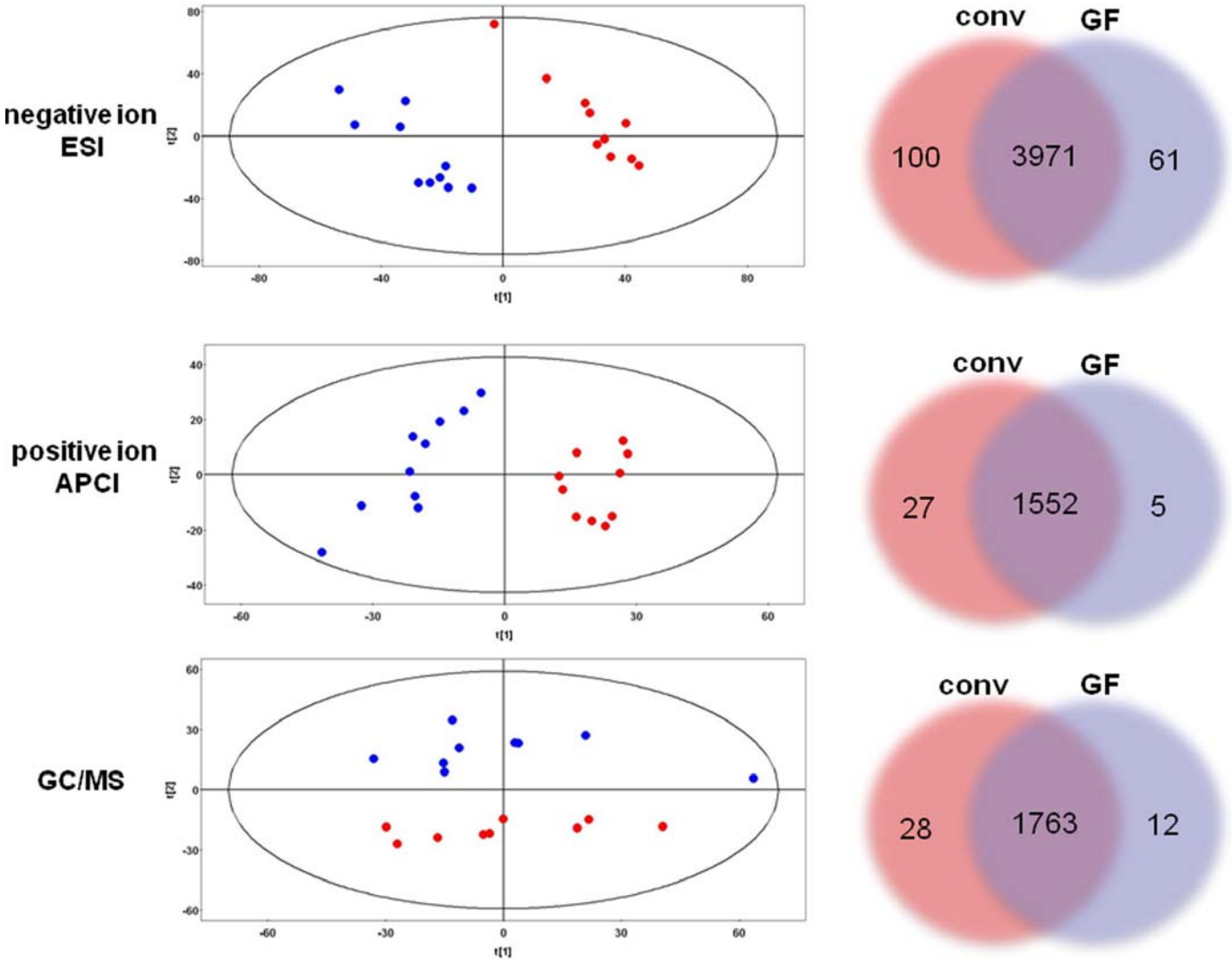

Fig. S1. PCA analysis of the profiling data and Venn diagrams for the individual features observed obtained employing − ESI (*Top*), + APCI (*Middle*), and GC/MS (*Bottom*).

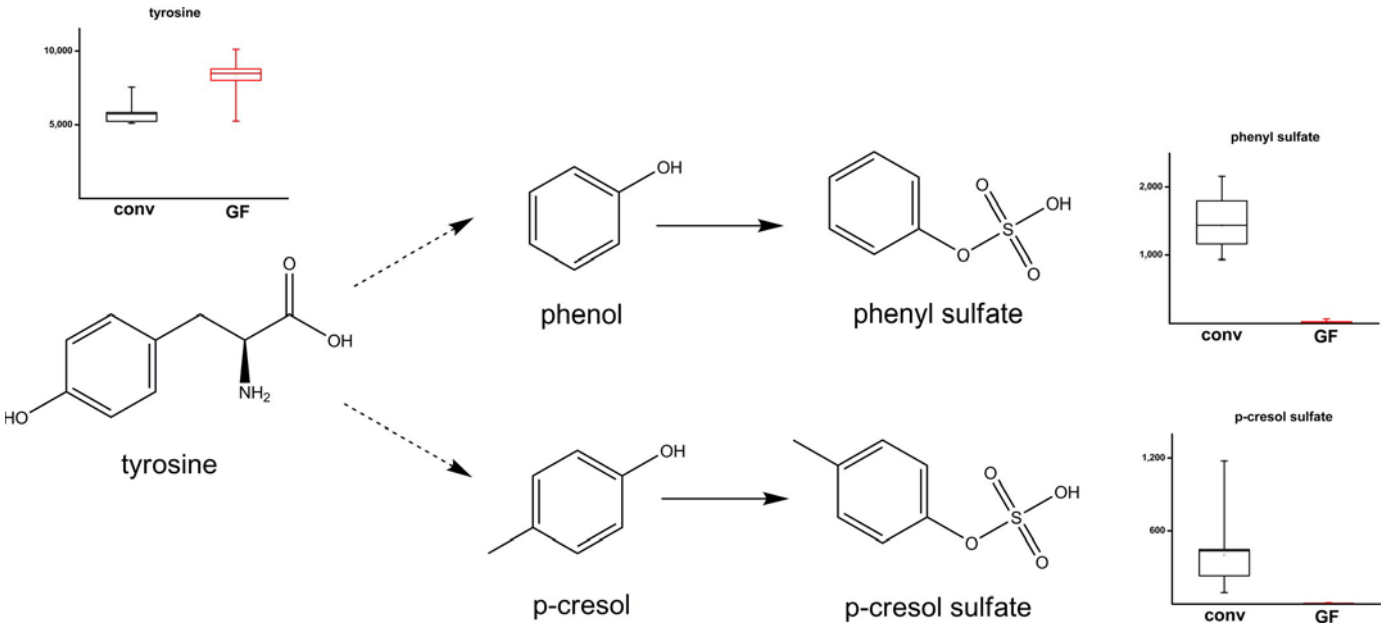

Fig. S2. Differences in metabolites that arise due to the action of the microbiome on tyrosine metabolism. The integrated signal intensities plotted on the *y* axes are reduced by a factor of 1,000.

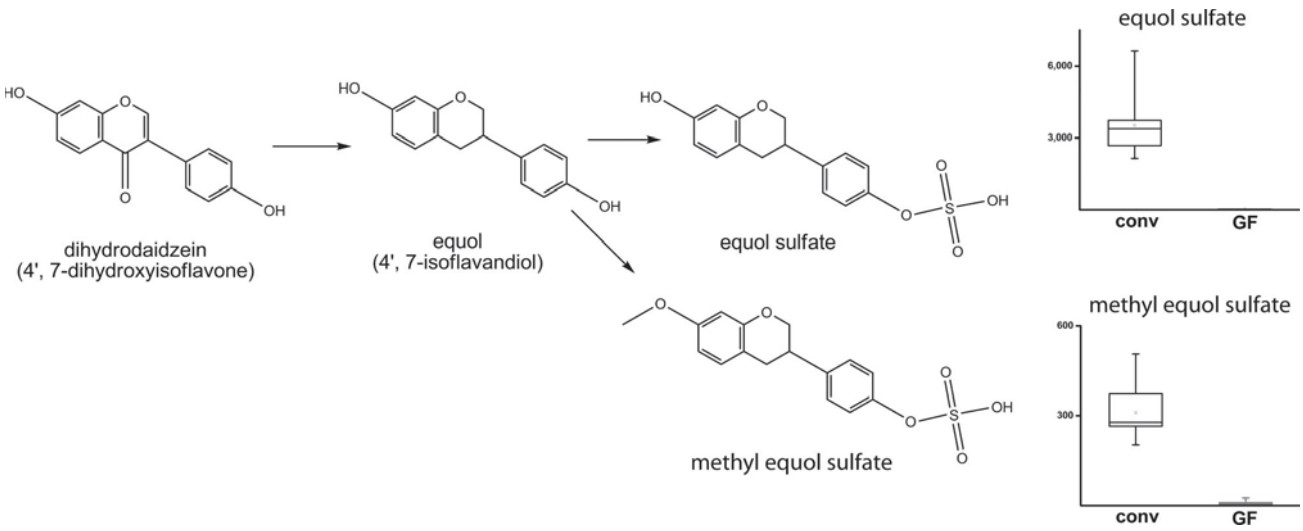

Fig. S3. Differences in metabolites that arise due to the action of the microbiome on the dietary isoflavone dihydrodaidzein. The integrated signal intensities plotted on the *y* axes are reduced by a factor of 1,000.

Table S1. Summary of mass spectrometric measurements based on which molecular identifications were made

Compound ID	Molecular formula	Mode observed	M	*m/z* obs	*m/z* calc, neutral	ppm error	RT, min	CE	Major fragments, QTOF
Indole compounds									
Tryptophan	$C_{11}H_{12}N_2O_2$	ESI +, ESI −, AP +, GCMS	+	205.0958	204.0898	−9.28	7.8	35	118.069, 91.058, 143.077, 130.07
N-acetyltryptophan	$C_{13}H_{14}N_2O_3$	ESI +	+	247.1061	246.1004	−8.77	18.7	25	159.096, 146.064, 132.083, 170.064, 188.076, 118.069
Indoxyl sulfate (indican)	$C_8H_7NO_4S$	ESI −	−	212.0032	213.0095	—	5.5	—	+132.044, 79.959
Serotonin	$C_{10}H_{12}N_2O$	ESI +, AP +	+	177.1008	176.0949	−11.22	5.1	35	115.054, 159.068, 133.053, 105.068, 77.039
Indole-3-propionic acid	$C_{11}H_{11}NO_2$	ESI +, ESI −, AP +	+	190.0851	189.0789	−8.96	23.1	—	130.064, 103.054, 77.039, 55.018
Phenyl compounds									
Phenylalanine	$C_9H_{11}NO_2$	ESI +, ESI −, AP +	+	166.0849	165.0789	−11.46	4.8	—	120.0801, 103.0540, 91.0544, 77.0388
Tyrosine	$C_9H_{11}NO_3$	AP +, ESI +	+	182.0804	181.0739	−7.24	2.2	35	119.0500, 107.0502, 95.0501, 91.0556, 77.0389, 65.0386, 56.9438
Hippuric acid	$C_9H_9NO_3$	ESI −, AP +	−	178.0515	179.0582	6.08	4.5	10	134.061, 77.041
Phenylacetylglycine	$C_{10}H_{11}NO_3$	ESI +, AP +	−	194.0795	193.0738	−11.43	12.1	—	91.058, 65.042
Phenyl sulfate	$C_6H_6O_4S$	ESI −	−	172.9920	173.9986	6.62	5.5	—	+93.035, 79.957
p-Cresol sulfate	$C_7H_8O_4S$	ESI −	−	187.0080	188.0143	7.99	8.3	—	+80, 107
Phenylpropionylglycine	$C_{11}H_{13}NO_3$	ESI +	+	208.0952	207.0895	−10.4	17.4	35	—
Cinnamoylglycine	$C_{11}H_{11}NO_3$	ESI +, ESI −	+	206.0802	205.0738	−7.37	19.2	35	103.0542, 77.0394, 131.04803
Others									
Urate	$C_5H_4N_4O_3$	ESI −, AP +	−	167.0215	168.0283	5.90	3.7	25	69.0116, 124.0166, 96.0203
Creatinine	$C_4H_7N_3O$	AP +	+	114.0655	113.0589	−10.8	1.4	—	—
Equol sulfate	$C_{15}H_{14}SO_6$	ESI −	−	321.0426	322.0511	−2.13	21	25	+241.0875, 147.0442, 135.0447, 121.0302, 119.0504, 79.9582
Methyl equol sulfate	$C_{16}H_{16}SO_6$	ESI −	−	335.0621	336.0667	9.44	25.2	25	+255.1032, 121.0309, 79.9587, 59.0099

M refers to the ESI ionization mode used for peak integration and compound identification. RT, retention time. CE refers to the collision energy employed to obtain the tandem MS spectra. All compounds identified as sulfates were also verified to have an isotope pattern for which the M + 2 peak was higher in intensity than expected compared with an isobaric species not containing sulfur.

Table S2. Individual bacterial species and their culture conditions that were assayed for the production of IPA

Division	Strain	ATCC number(s)	Growth medium
Actinobacteria	*Actinomyces odontolyticus* Batty	17982	BBL Actinomyces broth
Actinobacteria	*Collinsella aerofaciens* Eggerth	25986	ATCC medium 1053
Archaea	*Methanobrevibacter smithii*	35061	ATCC medium 1340
Bacteroidetes	*Alistipes putredinis*	29800	ATCC medium 1490
Bacteroidetes	*Bacteroides caccae*	43185	BBL cooked meat medium
Bacteroidetes	*Bacteroides capillosus*	29799	BBL cooked meat medium
Bacteroidetes	*Bacteroides ovatus*	8483	BBL cooked meat medium
Bacteroidetes	*Bacteroides stercoris*	43183	BBL cooked meat medium
Bacteroidetes	*Bacteroides thetaiotaomicron*	29148	BBL cooked meat medium
Bacteroidetes	*Bacteroides uniformis*	8492	BBL cooked meat medium
Bacteroidetes	*Parabacteroides merdae*	43184	BBL cooked meat medium
Firmicutes	*Clostridium leptum*	29065	ATCC medium 1490
Firmicutes	*Clostridium ramosum*	25582	ATCC medium 38
Firmicutes	*Clostridium scindens*	35704	ATCC medium 1053
Firmicutes	*Clostridium sporogenes*	17888, 319, 15579	TSB broth, BBL cooked meat medium*
Firmicutes	*Clostridium symbiosum*	14940	ATCC medium 1490
Firmicutes	*Coprococcus eutactus*	27759	ATCC medium 735
Firmicutes	*Eubacterium dolichum*	29143	ATCC medium 593
Firmicutes	*Eubacterium ventriosum* Tissier	27560	ATCC medium 593
Firmicutes	*Micromonas micros*	33270	ATCC medium 1102
Firmicutes	*Ruminococcus gnavus*	29149	ATCC medium 1490
Firmicutes	*Ruminococcus obeum*	29174	ATCC medium 1490
Firmicutes	*Ruminococcus torques*	27756	ATCC medium 735
Proteobacteria	*Escherichia coli*	700926, 700891	TSB broth

*ATCC number 15579 was tested in both media and was able to grow and produce IPA in both media

Table S3. Serum concentration of IPA measured in plasma samples after intraperitoneal injection of IPA

Treatment	Animal no.	IPA post-injection conc., μg/mL		
		0 h	1 h	6 h
Vehicle	1	0.00	0.00	0.08
	2	0.00	0.00	0.00
	3	0.00	0.00	0.00
	4	0.00	0.00	0.00
	5	0.00	0.00	0.00
	Vehicle mean	0.00	0.00	0.02
	Vehicle SD	0.00	0.00	0.04
10 mg/kg IPA	1	0.00	18.24	0.78
	2	0.00	22.76	0.79
	3	0.00	20.70	0.89
	4	0.00	26.70	—
	5	0.00	—	—
	10 mg/kg mean	0.00	22.10	0.82
	10 mg/kg SD	0.00	3.58	0.06
20 mg/kg IPA	1	0.00	31.09	0.53
	2	0.00	33.24	1.42
	3	0.00	57.06	1.66
	4	0.00	31.16	0.09
	5	0.00	24.25	1.11
	20 mg/kg mean	0.00	35.36	0.96
	20 mg/kg SD	0.00	12.60	0.64
40 mg/kg IPA	1	0.00	35.43	0.91
	2	0.06	—	—
	3	0.00	53.81	3.87
	4	0.00	33.22	2.49
	5	1.60	55.03	4.31
	40 mg/kg mean	0.33	44.37	2.90
	40 mg/kg SD	0.71	11.65	1.53

All values are in μg/mL of serum, and the limit of detection was 0.053 μg/mL.