nature metabolism

Article

https://doi.org/10.1038/s42255-026-01550-8

Host metabolism can produce many indoles and phenols independently of the microbiome

Received: 7 May 2025

Accepted: 17 May 2026

Published online: 23 June 2026

Check for updates

Jenna E. AbuSalim[1,2,3], Kellen Olszewski[2], Salma Youssef[1,2], Sarah J. Mitchell[2], Craig J. Hunter[3], Jessica Little[4], Ashley Sidebottom[4], Jacob A. Boyer[2], Steve D. Knutson[5], Laith Z. Samarah[2,3,5], Michael R. MacArthur[2,3,5], Alessa L. Henneberg[6], Rolf-Peter Ryseck[3], Christiane A. Opitz[6], David W. C. MacMillan[5], Mohamed S. Donia[1], Eric G. Pamer[7], Sabrina Imam[8], Christopher J. Lehmann[7], Olatoyosi Odenike[9] & Joshua D. Rabinowitz[1,2,3,5]

Indole and phenol metabolites are typically thought to be products of bacterial digestion of tryptophan (indoles) and phenylalanine or tyrosine (phenols). Interest in controlling gut microbial production of these metabolites has continually grown because they have important physiological impacts, with indoles agonizing aryl hydrocarbon receptor signalling and phenols being associated with healthy body weight. Although there is a growing body of research on which bacteria produce these metabolites, the host contribution to their circulating pools has not been characterized. Here, through stable isotope tracing in cell culture, mice and rats, we show that mammalian cells can make aryl-pyruvates, aryl-lactates, aryl-acetates and aryl-carboxylic acids independently of the microbiome. We demonstrate that circulating levels of these metabolites in mice and human patients are robust to perturbations of the microbiome. By contrast, bacterial metabolism is required to synthesize aryl-propionates and free indole, phenol and p-cresol. Overall, these results suggest that host metabolism is a major contributor to circulating indole and phenol metabolite pools.

Indole and phenol metabolites have major impacts on host physiology[1–3]. Indole metabolites are ligands for the aryl hydrocarbon receptor, which is a key mediator of gut immune tolerance and epithelial homoeostasis[2–4]. Phenylalanine-derived phenol metabolites, such as cinnamic acid, correlate with healthy body weight and can affect glucose and lipid handling[1], while tyrosine-derived metabolites, such as phenol sulfate or p-cresol sulfate, are uraemic toxins[5,6]. Supplementation with phenylalanine-derived or tryptophan-derived metabolites shows promise for the treatment of various diseases. In rodent models, providing indole metabolites alleviates the disease severity of inflammatory bowel disease[7–9], while supplementing with cinnamic acid has protective effects against obesity and non-alcoholic fatty liver disease[1,10].

As indole and phenol metabolites have important physiologic impacts, there is growing interest in identifying which bacteria are responsible for their production[11]. A recent study showed that *Clostridium* species have a metabolic pathway that can reduce aryl amino acids to their respective propionates[12]. Other species, such as *Lactobacillus johnsonii*, have also been shown to produce indole-3-pyruvate and phenylpyruvate as well as downstream metabolites[3,13,14]. Changes in

A full list of affiliations appears at the end of the paper. e-mail: joshr@princeton.edu

indole and phenol metabolite concentrations are often correlated with changes in microbiome composition[2,15].

Although this work highlights a potential role of gut microbial health in controlling indole and phenol metabolite concentrations, microbiome compositional changes are also known to affect host metabolism[16]. Recently, it was shown that kynurenine, a tryptophan metabolite made by the host, can be induced when mice are colonized with *Achromobacter pulmonis*[17]. Additionally, the gut microbiome can even affect glucose homoeostasis through various mechanisms, including by producing short-chain fatty acids that can fuel hepatic gluconeogenesis[16,18]. IL4I1, a monocyte-specific enzyme, has the capacity to produce indole-3-pyruvate and phenylpyruvate, the transamination products of tryptophan and phenylalanine[19–21]. This suggests that there may be some contribution of host metabolism to circulating indole and potentially phenol pools. Additionally, recent work showed comparable levels of some indole metabolites between germ-free and conventional mice[22]. Whether the presence of these metabolites in germ-free mice reflects host metabolism or uptake from diet remains to be determined.

Isotope tracing, such as infusing ^{13}C-labelled substrates, provides a gold-standard method for determining metabolite sources[23]. Here, we systematically evaluate the sources of circulating indole and phenol metabolites based on: (1) concentrations in germ-free mice; (2) production from a 2 h infusion of ^{13}C-labelled amino acids (this short infusion feeds host cells but not the gut microbiome); (3) production from a 36 h infusion of ^{13}C-labelled amino acids (this long infusion feeds both the host directly and the microbiome indirectly by secreted host protein); (4) modulation by antibiotics of labelling from the long infusions; (5) production from microbial digestion of a ^{13}C-protein diet; (6) ^{13}C-tracing in mammalian cell culture; and (7) cell and mouse-tissue lysate biochemistry. Through these studies, we trace mammalian cell metabolism of aryl amino acids to their respective aryl-pyruvates, aryl-lactates, aryl-acetates and aryl-carboxylic acids (but not aryl-propionates and indole, phenol and p-cresol). We find that the microbiome-derived, but not host-derived, indole and phenol metabolites decline in human patients treated with antibiotics. Therefore, many important indole and phenol metabolites that were thought to be microbiome-derived are substantially synthesized by the host.

Results

Germ-free mice have indole metabolites

We began by measuring indole metabolites in serum from germ-free or control (specific-pathogen-free) mice (Fig. 1a). Standard chow (PicoLab Rodent Diet 20, 5053) has an abundance of microbiome metabolites, including many of the indole species of interest. Compared to chow, purified diets had lower levels of the indole metabolites (Extended Data Fig. 1a). Consistent with increase dietary intake of these metabolites, chow-fed mice had significantly higher levels of indole-3-lactate and indole-3-propionate than mice fed with a purified diet (Extended Data Fig. 1b). Accordingly, to identify endogenous sources of these metabolites, all mice were fed a doubly irradiated purified diet (Research Diets, D11112201-1.5Vii).

Faeces were collected 10 days after initiation of the purified diet to confirm sterility of the germ-free mice, and serum was collected and analysed by liquid chromatography–mass spectrometry (LC–MS). Sterility was confirmed through anaerobic and aerobic cultures of dietary and faecal pellets. Some indole metabolites, such as indole-acrylic acid and free indole, were below the detection limit in serum from all mice. Others were readily detected. In serum from germ-free mice, indole-3-propionate was significantly depleted in comparison to control mice, and indoxyl sulfate was absent (Fig. 1b). Indole-3-propionic acid was present in the purified diet (Extended Data Fig. 1a), which may explain why there is a depletion of circulating concentrations but not total ablation as seen with indoxyl sulfate. Notably, other measured indole metabolites (indole-3-pyruvate, indole-3-acetate, indole-3-carboxylic acid and indole-3-lactate) were not changed.

Many circulating indole metabolites are host-derived

To trace production of the indole metabolites, we infused minimally perturbative quantities of ^{13}C-tryptophan into the systemic circulation of jugular-vein-catheterized mice for 2 h (Fig. 1c). As shown in prior work by our group[24], circulating amino acids do not readily cross into the gut lumen, making the infused ^{13}C-tryptophan inaccessible over this duration to gut microbes[24]. Consistent with this expectation, there was no detectable ^{13}C-tryptophan in the faeces at 2 h despite 20% enrichment in circulating tryptophan (Fig. 1d). There was also no labelling observed in circulating indole-3-propionate or indoxyl sulfate (Fig. 1d). The remaining indole metabolites, however, were all substantially labelled by the infused ^{13}C-tryptophan in both the circulation and tissues, suggesting that they are rapidly produced from circulating tryptophan by host cells (Fig. 1d and Extended Data Fig. 1c,d). Based on the combination of these tracing data and the concentration data in germ-free mice, we propose that indole-3-pyruvate, indole-3-lactate, indole-3-acetate and indole-3-carboxylic acid are substantially host-derived, while indole-3-propionate and indoxyl sulfate production requires microbiome metabolism.

We next sought to observe ^{13}C-labelling in the microbiome-derived indoles. For this purpose, we again infused ^{13}C-tryptophan, but for a much longer duration (Fig. 1e). Previous work by our group showed that bacteria in the gut microbiome use labelled amino acids that are infused for 36 h[24], as the infused amino acid reaches the gastrointestinal lumen through host secretion of proteins such as mucins. Faecal tryptophan labelling was indeed observed at 36 h (Extended Data Fig. 1e). Serum measurements at 36 h revealed ^{13}C-labelling into all indole metabolites, including indole-3-propionate and indoxyl sulfate, supporting a source of the microbiome-derived indoles being bacterial processing of tryptophan derived from secreted host protein (Extended Data Fig. 1f). In mice treated for 14 days with broad-spectrum antibiotics (ampicillin, neomycin, vancomycin and metronidazole), luminal tryptophan labelling was retained (Extended Data Fig. 1e), but labelling of indole-3-propionate and indoxyl sulfate was largely ablated (Fig. 1f). Antibiotic efficacy was confirmed by a decrease in both alpha diversity of the microbial composition (Extended Data Fig. 1g) and overall bacterial load (Extended Data Fig. 1h). Crucially, there was no change in labelling for the remaining indoles, supporting their production by the host.

In addition to secreted host proteins, dietary protein is a major source of amino acids feeding the gut microbiota. To assess whether there is a contribution to circulating pools of indoles from dietary amino acids, we gavaged mice with ^{13}C-tryptophan. As tryptophan is a free amino acid, it was readily absorbed by the small intestine into the systemic circulation by 15 min, while the caecum did not show detectable ^{13}C-enrichment until 30 min following the gavage (Extended Data Fig. 1i). Consistent with our ^{13}C-tryptophan infusion experiments, all host-derived indoles were highly enriched by 15 min after gavage (Extended Data Fig. 1i). By contrast, there was very little enrichment of the microbiome-derived indole-3-propionate or indoxyl sulfate, even 2 h following the gavage. To more robustly provide the gut microbiome with ^{13}C-amino acids from diet, we fed mice a purified diet containing ^{13}C-algal protein, which is resistant to small intestinal digestion[25] (Fig. 1g). Serum and caecal contents were collected 24 h after initiation of the ^{13}C-enriched diet. Enrichment of the indole metabolites in serum was normalized to ^{13}C-tryptophan enrichment in the caecal content (13%; Extended Data Fig. 1j). After 24 h, there was substantial labelling into indole-3-propionate and indoxyl sulfate (80% and 100% normalized labelling, respectively; Fig. 1h). Opposingly, none of the host-derived indole metabolites were enriched to a greater extent than circulating tryptophan (all less than 3% enriched), consistent with production by host tissues instead of the gut microbiome (Fig. 1h).

https://doi.org/10.1038/s42255-026-01550-8

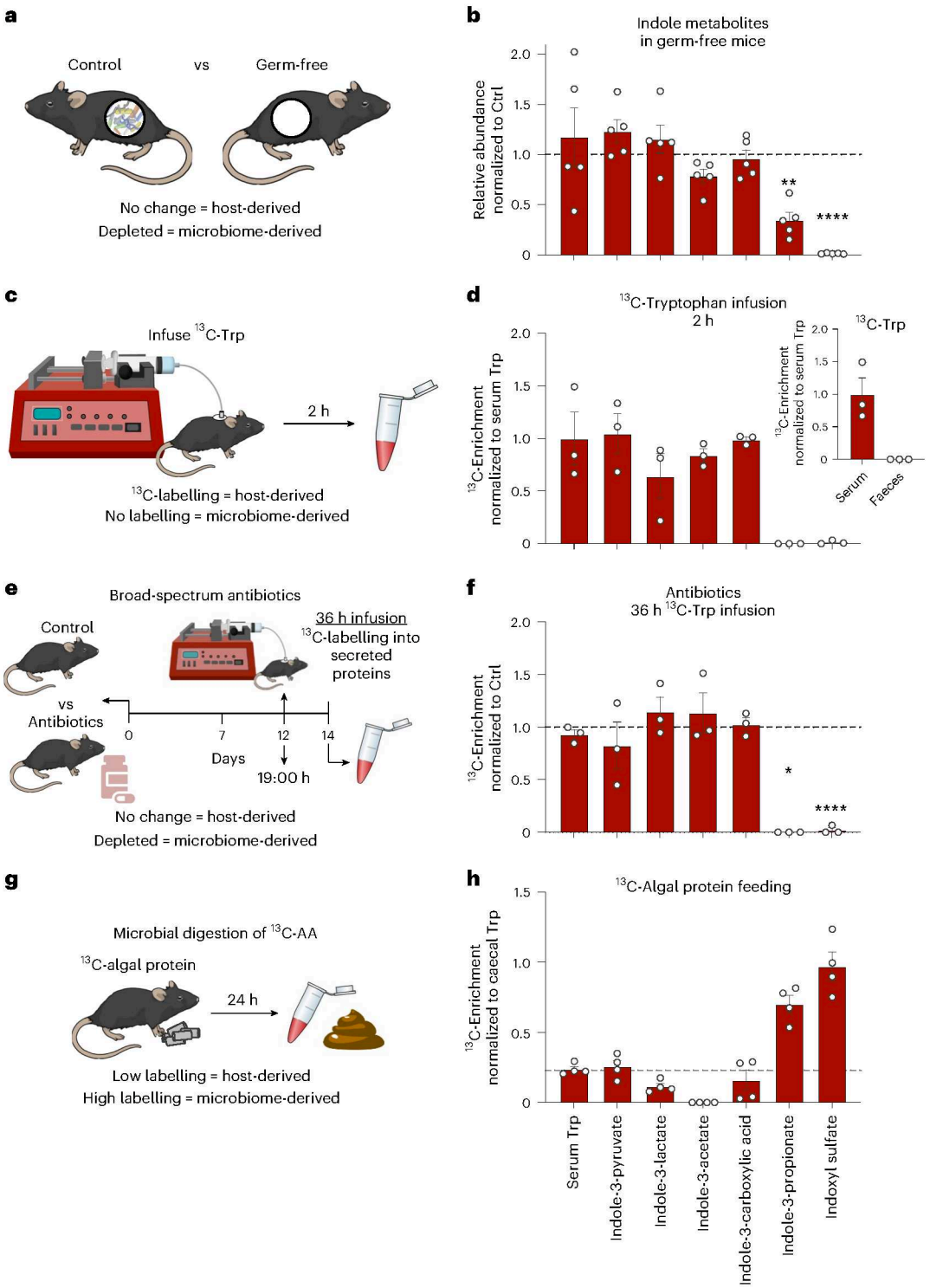


Fig. 1 | The host produces many circulating indole metabolites. a,**b**, Scheme (**a**) and measurement (**b**) of indole metabolites in serum of germ-free mice ($n = 5$ female mice) compared to control (Ctrl) mice ($n = 4$ female mice). **c**,**d**, Scheme describing the infusion of ^{13}C-tryptophan for 2 h (**c**) and ^{13}C-enrichment of indole metabolites in serum normalized to serum ^{13}C-tryptophan enrichment ($n = 3$ male mice) (**d**). Inset shows ^{13}C-tryptophan in serum and faeces at 2 h. **e**,**f**, Scheme showing a 14 day antibiotic treatment with ^{13}C-tryptophan infused for the last 36 h (**e**) and ^{13}C-indole enrichment in antibiotic-treated male mice ($n = 4$) compared to Ctrl male mice ($n = 5$) (**f**). **g**,**h**, Scheme for tracing circulating indoles coming from microbial digestion of dietary protein (**g**) and ^{13}C-indole enrichment in portal vein from male mice fed a ^{13}C-algal protein-containing diet for 24 h ($n = 4$) (**h**). Metabolite labels for **b**, **d** and **f** are the same as **h**. All data are graphed as mean values; error bars, s.e.m. *P* values were calculated using a two-tailed *t*-test; $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$; $^{****}P < 0.0001$.

Our classification of metabolites as host-derived does not negate the known capacity of the microbiome to make these biochemicals. Rather, it reflects a capacity of the host, independent of the microbiome, to produce the metabolites in physiologically relevant quantities. Based on this definition, we conclude that four of the measured indole metabolites made from tryptophan are host-derived (indole-3-pyruvate, indole-3-lactate, indole-3-acetate and indole-3-carboxylic acid) and two are microbiome-dependent (indole-3-propionate and indoxyl sulfate).

The host also produces phenylalanine catabolites

Similar to indole metabolites, phenol metabolites from phenylalanine and tyrosine are commonly considered microbiome-derived[15]. Other than changes in the aryl side group between tryptophan, phenylalanine and tyrosine, the enzymatic transformations needed to produce the phenols from phenylalanine or tyrosine mimic the transformations needed to produce indoles from tryptophan. Chow contained high levels of several phenols, which were present at lower levels in purified diet (phenylpyruvate, phenyllactate and benzoic acid; Extended Data Fig. 2a). Chow-fed mice also had significantly higher circulating levels of 3-phenylpropionate, consistent with its increased concentration in the chow diet (Extended Data Fig. 2a,b). Additional phenol metabolites such as phenylacetylglycine and hippuric acid are readily detected in both chow and serum from chow-fed mice but were depleted to near the limit of detection in mice fed a purified diet (Extended Data Fig. 2a,b), suggesting that diet is a major feedstock for these metabolites.

To explore whether there is host production of the other phenylalanine-derived metabolites, we analysed serum by LC–MS from purified diet-fed germ-free and control mice (Fig. 2a). Phenylpyruvate (the transamination product of phenylalanine), phenyllactate (the subsequent reduced product) and cinnamic acid (the associated alpha–beta unsaturated metabolite) were not changed between germ-free and control mice (Fig. 2b). Phenylacetate was below detectable levels in all the serum samples. By contrast, benzoic acid and 3-phenylpropionate were significantly decreased in germ-free mice.

Similar to tracing indole metabolite production using ^{13}C-tryptophan, we used stable isotope labelling to measure the production of the above phenols from ^{13}C-phenylalanine (Fig. 2c). In a minimally perturbative jugular vein infusion of ^{13}C-phenylalanine for 2 h, there was no detectable ^{13}C-phenylalanine in the faeces. Examining circulating and tissue labelling, the phenylalanine-derived phenols whose levels were maintained in germ-free mice (phenylpyruvate, phenyllactate and cinnamic acid) were robustly labelled (Fig. 2d and Extended Data Fig. 2c,d). Phenol metabolites that were significantly decreased in the germ-free mice, by contrast, were labelled modestly (benzoic acid, 20%) or not at all (3-phenylpropionate) (Fig. 2d and Extended Data Fig. 2d).

To more fully map the sources of these metabolites, we infused the ^{13}C-phenylalanine tracer for 36 h. By 36 h into the infusion, the circulating ^{13}C-phenylalanine enrichment reached 25% and the faecal enrichment reached 7% (Extended Data Fig. 2e). As with the 2 h infusion, there was full enrichment of phenylpyruvate and 80% enrichment of phenyllactate and cinnamic acid (Extended Data Fig. 2f). Even after 36 h of a ^{13}C-phenylalanine infusion, however, benzoic acid was still only enriched by 30% and no labelled 3-phenylpropionate was detected. Consistent with phenylpyruvate, phenyllactate, cinnamic acid and benzoic acid being substantially host-derived, their labelling from 36 h ^{13}C-phenylalanine infusion was maintained in antibiotic-treated mice (Fig. 2e,f and Extended Data Fig. 2e,g,h).

The missing labelling for benzoic acid and 3-phenylpropionate implies that there is another upstream precursor for these metabolites, possibly from dietary contamination or bacterial digestion of dietary protein. To assess the contribution from orally delivered amino acids, we gavaged ^{13}C-phenylalanine. At 15 min after the gavage, there was substantial labelling of phenylalanine in serum (55%) but minimal enrichment in the caecum (3%). Consistent with being host-derived, phenylpyruvate, phenyllactate, cinnamic acid and benzoic acid all had circulating enrichment higher than the caecal phenylalanine enrichment (Extended Data Fig. 2i), but there was still no labelling into 3-phenylpropionate. To deliver the ^{13}C-phenylalanine more robustly to the caecum, we fed mice diets with ^{13}C-algal protein for 24 h (Fig. 2g). Caecal enrichment of phenylalanine reached 42%, while serum phenylalanine enrichment was 12% (Extended Data Fig. 2j). The 3-phenylpropionate was extensively labelled, equivalent to caecal phenylalanine (Fig. 2h). The remaining phenylalanine-derived metabolites (phenylpyruvate, phenyllactate, cinnamic acid and benzoic acid) were labelled less than serum phenylalanine, further supporting their systemic concentrations being substantially maintained by host metabolism independent of the microbiome.

Host metabolism is responsible for circulating hydroxyphenylpyruvate and hydroxyphenyllactate

Tyrosine is structurally identical to phenylalanine except for a hydroxyl group on the aryl ring. As such, we reasoned that tyrosine undergoes host-mediated transformation to hydroxyphenylpyruvate and hydroxyphenyllactate. Additionally, phenol and p-cresol are metabolites that are understood to come from microbial digestion of tyrosine down to the aryl group or aryl group plus one remaining methyl group[11,26]. These metabolites are then rapidly sulfated by the host[26,27]. The dietary content of phenol sulfate and p-cresol sulfate was higher for the tested purified diet than for chow (Extended Data Fig. 3a), but mice fed a purified diet had similar or lower circulating concentrations of phenol or p-cresol sulfate to chow-fed mice, pointing to these metabolites being made (either by the host or microbiome) rather than simply coming from diet (Extended Data Fig. 3b).

To understand the host contribution to tyrosine-derived metabolites, we measured serum from purified diet-fed germ-free and control mice (Fig. 3a). Hydroxyphenylacetate, hydroxycinnamic acid, free phenol and free p-cresol were below detectable levels in all the serum samples. There was no change in circulating hydroxyphenylpyruvate and hydroxyphenyllactate in germ-free versus control mice (Fig. 3b), while phenol sulfate and p-cresol sulfate were significantly depleted in germ-free mice.

We next traced host production of tyrosine-derived metabolites. The insolubility of ^{13}C-tyrosine in saline at neutral pH limited the concentration at which tyrosine could be infused. At 2 h after infusion with ^{13}C-tyrosine, the serum enrichment reached only 1% (Extended Data Fig. 3c). As tyrosine can be made from phenylalanine, we measured ^{13}C-enrichment of tyrosine after a 2 h infusion of ^{13}C-phenylalanine. The ^{13}C-tyrosine enrichment from a phenylalanine infusion reached 2%. In both male and female mice, hydroxyphenylpyruvate and hydroxyphenyllactate were enriched by the tracer by 2 h into the infusion (Fig. 3d and Extended Data Fig. 3c,d). No labelled phenol sulfate or p-cresol sulfate was detected. This suggests that hydroxyphenylpyruvate and hydroxyphenyllactate are substantially host-derived, whereas the carbon skeletons of phenol sulfate and p-cresol sulfate are microbiome-derived (with the sulfation presumably done by the host).

To achieve labelling of phenol and p-cresol sulfate, we infused ^{13}C-tyrosine for 36 h. By 36 h after infusion, the circulating ^{13}C-tyrosine enrichment reached 4% and faecal enrichment was 2% (Extended Data Fig. 3d). All tyrosine-derived metabolites were enriched by 36 h (Extended Data Fig. 3e). To differentiate labelling by host metabolism versus microbiome metabolism, we infused ^{13}C-tyrosine in antibiotic-treated mice (Fig. 3e). Consistent with the assignment of hydroxyphenylpyruvate and hydroxyphenyllactate as host-derived, there was no change in enrichment between control and antibiotic-treated mice for these two metabolites (Fig. 3f). There was, however, a significant decrease with antibiotics in ^{13}C-enrichment of phenol sulfate and p-cresol sulfate, consistent with production of these metabolites requiring the microbiome. Oral gavage with ^{13}C-tyrosine aligned with these assignments (Extended Data Fig. 3i).

To deliver ^{13}C-tyrosine preferentially to the microbiome instead of host circulation, we fed mice a diet with ^{13}C-digestion-resistant algal protein for 24 h (Fig. 3g). When normalized to caecal ^{13}C-tyrosine labelling (Extended Data Fig. 3j), circulating phenol sulfate and p-cresol sulfate had enrichment reaching about 50% (Fig. 3h). Hydroxyphenylpyruvate and hydroxyphenyllactate had relative

https://doi.org/10.1038/s42255-026-01550-8

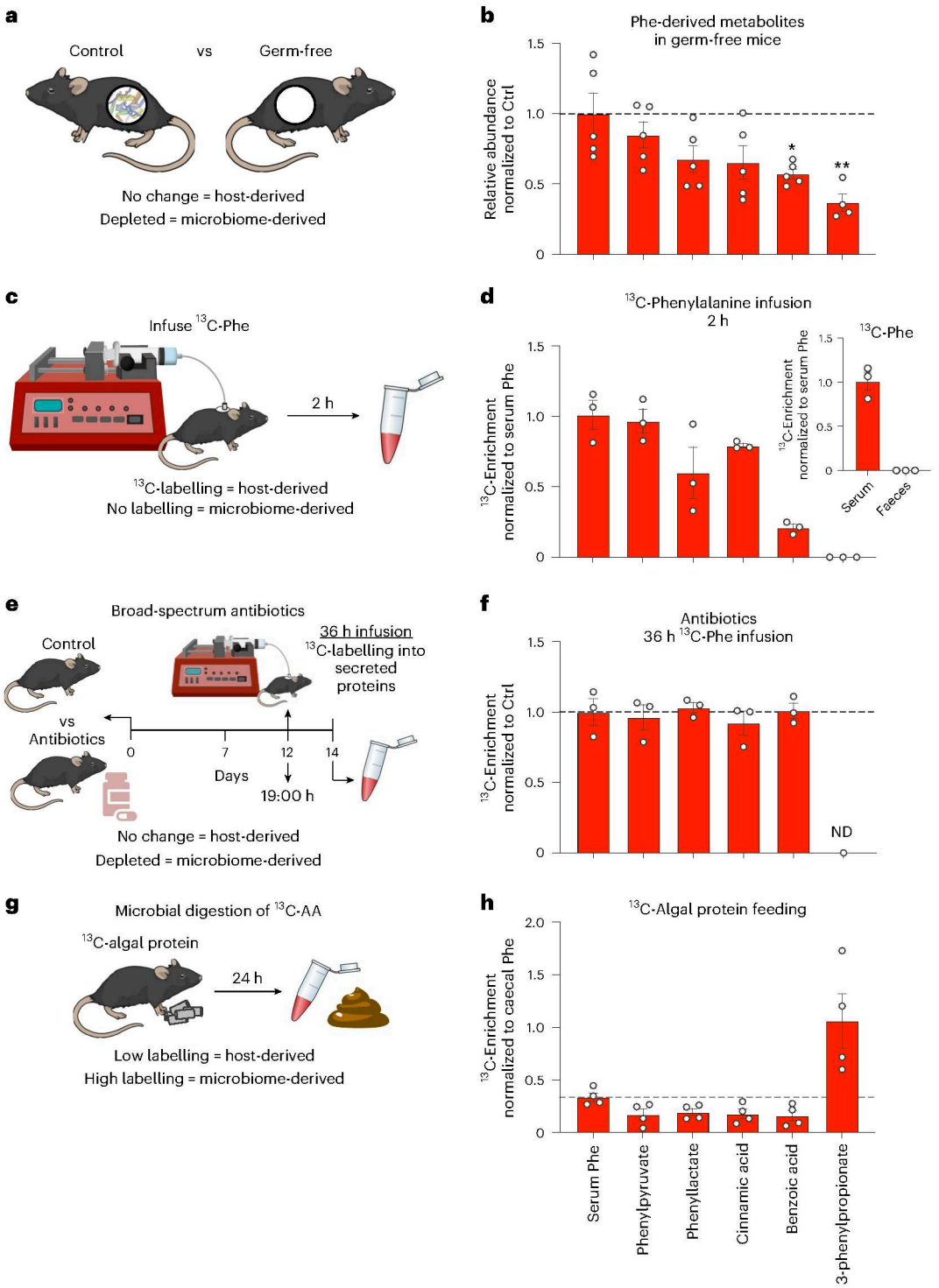


Fig. 2 | Most products of phenylalanine can be produced by the host. **a**,**b**, Scheme (**a**) and measurement (**b**) of phenylalanine-derived phenol metabolites in serum of germ-free mice ($n = 5$ female mice) compared to control mice ($n = 4$ female mice). **c**,**d**, Scheme describing the infusion of ^{13}C-phenylalanine for 2 h (**c**) and ^{13}C-enrichment of phenylalanine-derived metabolites in serum normalized to serum ^{13}C-phenylalanine enrichment ($n = 3$ male mice) (**d**). Inset shows ^{13}C-phenylalanine in serum and faeces at 2 h. **e**,**f**, Scheme showing a 14 day antibiotic treatment with ^{13}C-phenylalanine infused for the last 36 h (**e**) and ^{13}C-phenol enrichment in antibiotic-treated male mice ($n = 3$) compared to Ctrl male mice ($n = 3$) (**f**). **g**,**h**, Scheme for tracing circulating phenylalanine-derived phenols coming from microbial digestion of dietary protein (**g**) and ^{13}C-phenol enrichment in portal vein from male mice fed a ^{13}C-algal protein-containing diet for 24 h ($n = 4$) (**h**). All data are graphed as mean values; error bars, s.e.m. P values are notated as $^{*}P < 0.05$ and $^{**}P < 0.01$ as calculated by a two-tailed t-test. ND, not detected.

enrichment matching or below circulating ^{13}C-tyrosine enrichment (26%), further supporting phenol sulfate and p-cresol sulfate, but not hydroxyphenylpyruvate and hydroxyphenyllactate, originating from microbial metabolism.

Indole and phenol metabolites are also made by host metabolism in rats

To assess the generality of the above findings, we carried out short (2 h) ^{13}C-tryptophan and ^{13}C-phenylalanine tracer studies in

https://doi.org/10.1038/s42255-026-01550-8

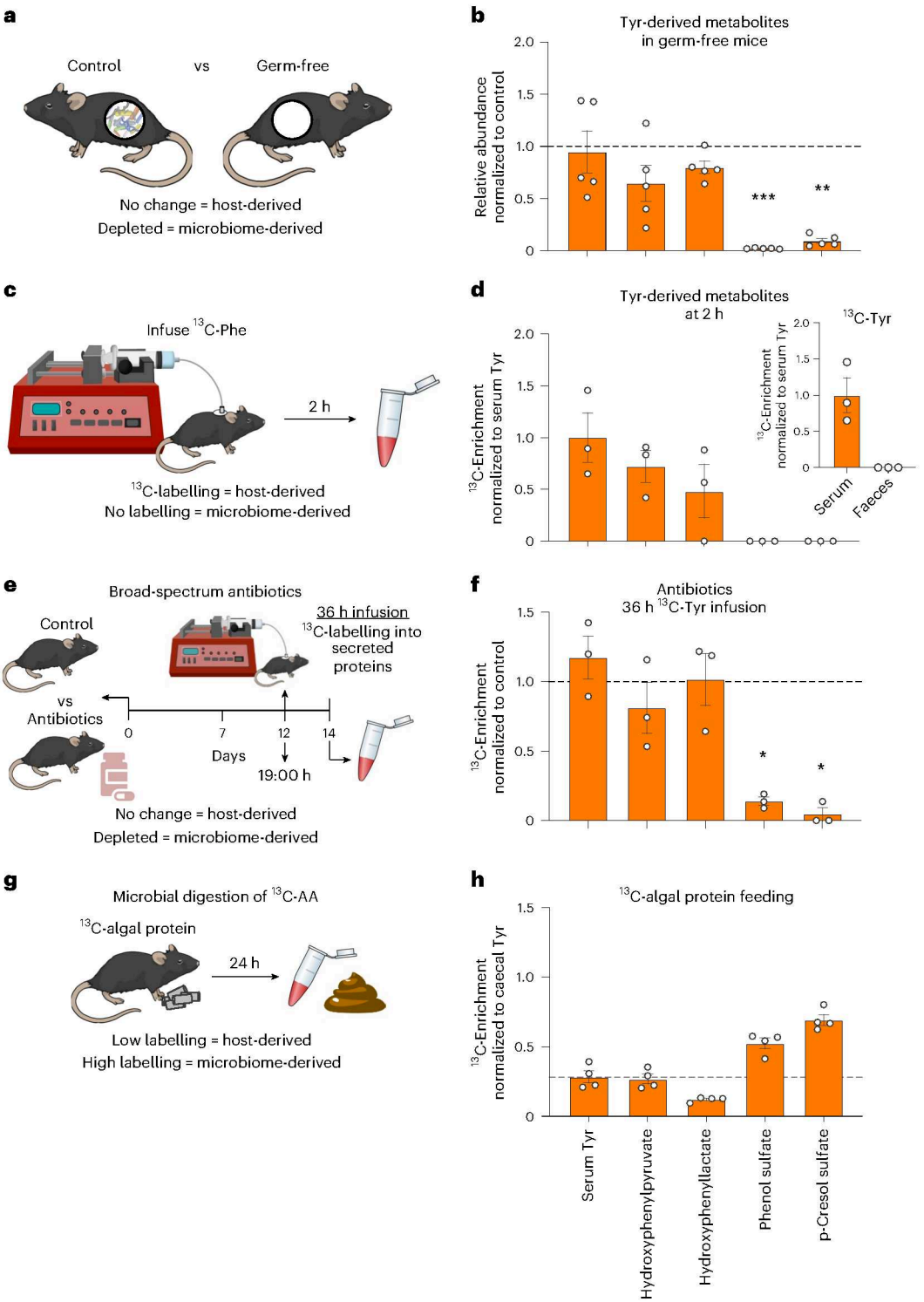


Fig. 3 | The transaminase products of tyrosine are host-derived. a,**b**, Scheme (**a**) and measurement (**b**) of phenol metabolites in serum of germ-free mice ($n = 5$ female mice) compared to Ctrl mice ($n = 4$ female mice). **c**,**d**, Scheme describing the infusion of ^{13}C-phenylalanine for 2 h (**c**) and ^{13}C-enrichment of tyrosine-derived metabolites in serum normalized to serum ^{13}C-tyrosine enrichment from the ^{13}C-phenylalanine infusion ($n = 3$ male mice) (**d**). Inset shows ^{13}C-tyrosine in serum and faeces at 2 h. **e**,**f**, Scheme showing a 14 day antibiotic treatment with ^{13}C-tyrosine infused for the last 36 h (**e**) and ^{13}C-phenol enrichment in antibiotic-treated male mice ($n = 3$) compared to enrichment of ^{13}C-phenol metabolites in Ctrl male mice ($n = 3$) (**f**). **g**,**h**, Scheme for tracing circulating tyrosine-derived phenols coming from microbial digestion of dietary protein (**g**) and ^{13}C-phenol enrichment in portal vein from male mice fed a ^{13}C-algal protein-containing diet for 24 h ($n = 4$) (**h**). All data are graphed as mean values; error bars, s.e.m. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$ as calculated by a two-tailed t-test.

jugular-vein-catheterized Sprague Dawley rats (Extended Data Fig. 4a). As in mice, these tracers appear in the circulation but not faecal contents over this timescale (Extended Data Fig. 4b–d). All of the metabolites that were assigned as host-derived in mice (indole-3-pyruvate, indole-3-lactate, indole-3-acetate, indole-3-carboxylic acid, phenylpyruvate, phenyllactate, cinnamic acid, benzoic acid, hydroxyphenylpyruvate and hydroxyphenyllactate) became substantially labelled, whereas those that were assigned as microbiome-dependent

(indole-3-propionate, indoxyl sulfate, 3-phenylpropionate, phenol sulfate and p-cresol sulfate) did not. These data suggest that mice and rats share conserved metabolic pathways for the production of indole and phenol metabolites.

Overall concentrations of indole and phenol metabolites are similar in faeces and circulation

As the gut microbiome has been thought to be required for the production of indole and phenol metabolites, many studies measure faecal concentrations of these metabolites[7,28–30]. We measured, from mice on a purified diet, absolute indole and phenol metabolite concentrations from blood, faeces and urine. The most abundant metabolites overall were p-cresol sulfate, phenylpyruvate, hydroxyphenyllactate and indole-3-pyruvate (Extended Data Fig. 5). For a majority of metabolites, urinary concentrations significantly exceed serum ones, consistent with these metabolites being filtered but not reabsorbed. The opposite pattern was observed for p-cresol sulfate, consistent with its high plasma protein binding[31,32]. Serum and faecal concentrations differed significantly for only five metabolites: two indoles were serum-enriched, and three phenols were faeces-enriched. The faeces-enriched phenols included hydroxyphenyllactate, whose circulating levels are substantially host-derived. This example highlights the potential for extensive local gut microbiome production of metabolites also made by the host.

Human cell lines can produce indole and phenol metabolites

The above data support many indole and phenol metabolites being made by the host in vivo. We next explored whether cultured mammalian cells can autonomously produce these metabolites. To this end, we cultured colorectal cancer cells (HCT116) and traced the metabolism of tryptophan, phenylalanine and tyrosine (Fig. 4a). In cells cultured with media containing ^{13}C-tryptophan, indole-3-pyruvate and indole-3-lactate were labelled by the added ^{13}C-tryptophan tracer (Fig. 4b; indole-3-lactate labelling is substantial but not quantitative owing to dilution by its presence in the media). The other host-derived indole metabolites (indole-3-acetate and indole-3-carboxylic acid) were not produced by HCT116 cells.

Similarly, we cultured HCT116 cells in media containing either ^{13}C-phenylalanine or ^{13}C-tyrosine. The cells produced phenylpyruvate, phenyllactate and cinnamic acid from the added ^{13}C-phenylalanine, further supporting these being host-derived metabolites. The microbiome metabolite 3-phenylpropionate was, as expected, not made from the added tracer (Fig. 4c). Benzoic acid, which comes at least partially from host metabolism in mice, was not produced by HCT116 cells. From ^{13}C-tyrosine, HCT116 cells produced labelled hydroxyphenylpyruvate and hydroxyphenyllactate but not phenol or p-cresol sulfate (Fig. 4d). The HCT116 cells were confirmed to be negative for *Mycoplasma* contamination (Extended Data Fig. 6a,b). Production of these metabolites was also observed in various other common cell lines (Extended Data Fig. 6c–q). Overall, these cultured mammalian cells exhibited a subset of the in vivo transformations observed in rodents, with cells in culture converting aryl amino acids to their respective aryl-pyruvates, aryl-lactates and cinnamic acid.

The metabolic pathways from aryl amino acids to the indole and phenol metabolites could logically proceed initially through transamination to the respective pyruvate analogues. To assess whether transaminase inhibition could block production of the downstream products, we treated HCT116 cells with the pan-transaminase inhibitor aminooxyacetic acid and traced production of indoles and phenols through the addition of ^{13}C-aryl amino acids (Fig. 4e). Aminooxyacetic acid was cytostatic at the given dose (500 μM). Upon treatment with aminooxyacetic acid, production of indole and phenol metabolites by HCT116 cells decreased (Fig. 4f–h). Therefore, isolated mammalian cells can autonomously produce many indole and phenol metabolites through transamination followed by subsequent metabolic reactions.

Characterization of the reactions in cell and tissue lysates

To further characterize the mammalian synthetic pathway to indoles and phenols, we performed biochemical assays in crude lysates from HCT116 cells and from mouse liver. The conversion of the amino acid to aryl-pyruvate is probably through a transaminase, which typically requires the cofactor pyridoxal 5′-phosphate (PLP) as well as a nitrogen acceptor such as alpha-ketoglutarate (α-KG). With the addition of PLP and α-KG, both types of lysates had robust transaminase activity, converting tryptophan to indole-3-pyruvate and phenylalanine to phenylpyruvate (Fig. 4i and Extended Data Fig. 7a–c). Both lysates also reduced the aryl-pyruvates to aryl-lactates with the addition of NADH (Fig. 4j and Extended Data Fig. 7d,e) and converted phenylpyruvate to cinnamic acid (Fig. 4l and Extended Data Fig. 7f). Interestingly, when phenyllactate was added, no cinnamic acid was formed, suggesting that cinnamic acid production proceeds directly from phenylpyruvate and not through phenyllactate (Extended Data Fig. 7f). Furthermore, when indole-3-pyruvate was added to the lysates (without any further cofactor requirement), the liver lysates, but not cell lysates, produced indole-3-acetate and indole-3-carboxylic acid. Similarly, liver lysates converted phenypyruvate into benzoic acid (Fig. 4k). Therefore, liver lysates can make all the host-derived indole and phenol metabolites.

GOT1 produces indole-3-pyruvate and phenylpyruvate

To understand what enzymes may be responsible for the conversion of aryl amino acids to their downstream products, we started by assessing IL4I1 activity, an enzyme initially discovered as being induced by IL-4 in B cells, but since found to be expressed and excreted predominantly by monocyte lineage cells. IL4I1 reportedly converts phenylalanine to phenylpyruvate and tryptophan to indole-3-pyruvate in vitro[19–21]. When comparing plasma concentrations from *Il4i1*$^{-/-}$ and wild-type mice, there were no significant changes in any of the host-derived indole or phenol metabolites, suggesting that IL4I1 is not required for these host pathways (Extended Data Fig. 8a–d).

To find additional host enzymes capable of tryptophan or phenylalanine transamination, we tracked these activities in mouse liver lysates as they were fractionated sequentially by size exclusion (SEC), anion exchange and cation exchange chromatography. The resulting fractions were assayed for indole-3-pyruvate and phenylpyruvate production, with the most active fraction from each step selected for further fractionation. Both activities co-eluted consistently through SEC, anion exchange and cation exchange (Fig. 5a–f). To find candidate enzymes responsible for the activity observed in these lysates, we performed proteomics on fractions exhibiting a wide range of activities from the final cation exchange separation and correlated the abundance of all measured proteins with phenylalanine transaminase enzymatic activity. Out of the top 20 best-correlated proteins (all Spearman correlation > 0.96), keratin 79 (KRT79) had the highest abundance in the most active fraction assayed (Fig. 5g). Glutamate oxaloacetate transaminase 1 (GOT1), one of the most abundant mammalian transaminases, was the next highest abundance. Given that KRT79 does not have known enzymatic activity and GOT1 is a well-characterized amino acid transaminase, we selected GOT1 for further biochemical characterization. With the addition of PLP and α-KG, recombinant purified GOT1 efficiently converted both tryptophan and phenylalanine to their respective aryl-pyruvates (Fig. 5h). Therefore, GOT1 can initiate the host pathways to indole and phenol metabolites.

Alignment with clinical data involving antibiotic therapy

The above data differentiate indole and phenol metabolites into host-derived and microbiome-dependent classes. We next asked whether this categorization would align with metabolic responses in humans to antibiotics. To this end, we analysed the abundance of indole and phenol metabolites in a publicly available metabolomics dataset of serum from patients in general admissions at Boston Children's Hospital[33,34]. The original purpose of this dataset was to compare

https://doi.org/10.1038/s42255-026-01550-8

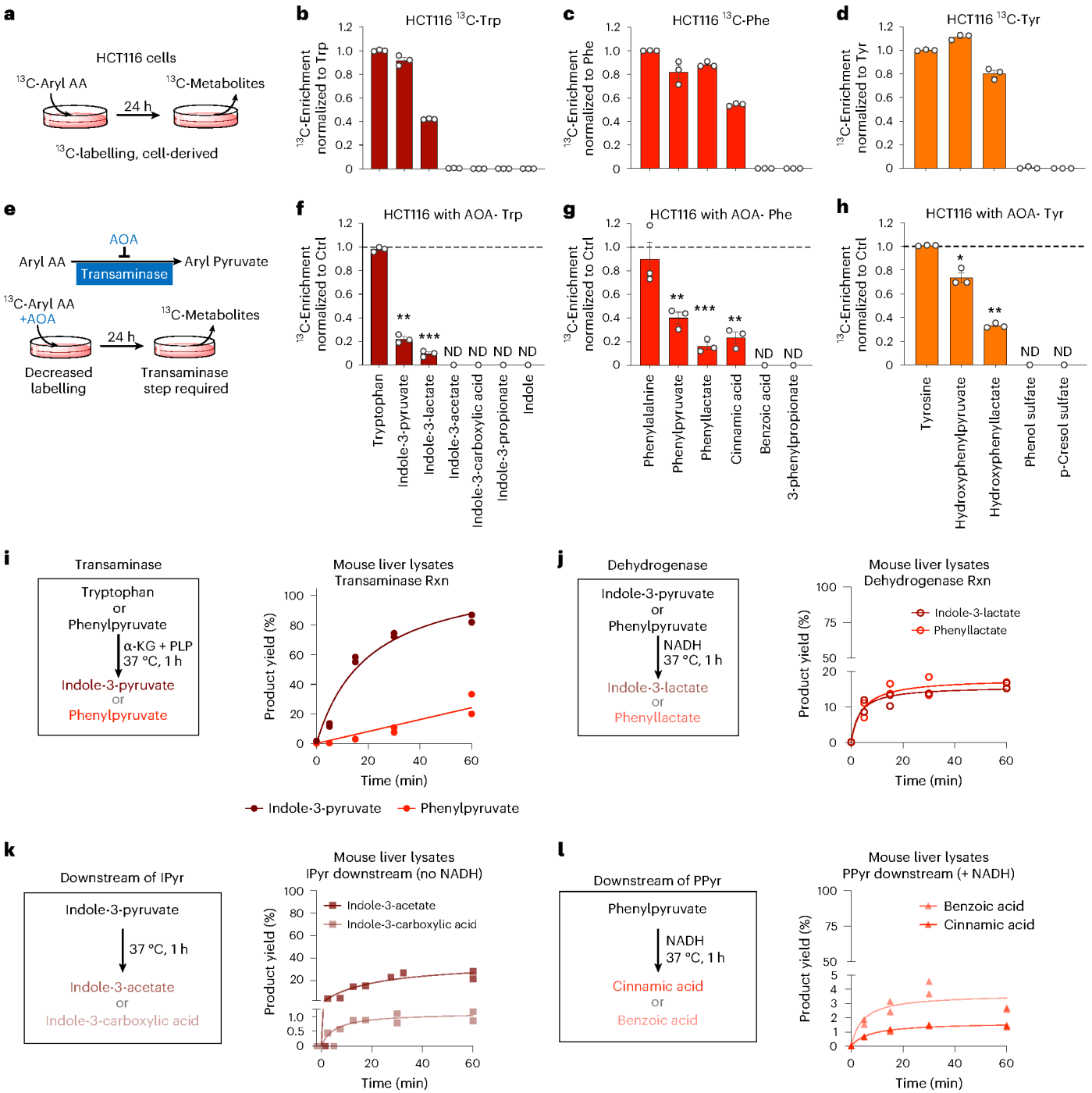


Fig. 4 | Many indoles and phenols are made in mammalian cells and tissue lysates. a, Scheme for tracing metabolism of ^{13}C-amino acids into indoles and phenols in HCT116 cells. **b**, ^{13}C-enrichment of indole metabolites in HCT116 cells 24 h after addition of 16 mg ml^{-1} ^{13}C-tryptophan to the media ($n = 3$). **c**,**d**, ^{13}C-enrichment of phenol metabolites from 66 mg ml^{-1} ^{13}C-phenyalalnine ($n = 3$) (**c**) or 80 mg ml^{-1} ^{13}C-tyrosine ($n = 3$) (**d**). Enrichment was normalized to the respective ^{13}C-amino acids in HCT116 cells. **e**, Schematic showing HCT116 cells being treated with 500 µM aminooxyacetic acid (AOA) to inhibit transaminases, the presumptive first step for indole and phenol synthesis. **f**–**h**, Relative ^{13}C-enrichment of tryptophan-derived (**f**), phenylalanine-derived (**g**) or tyrosine-derived (**h**) metabolites 24 h following the addition of AOA ($n = 3$ for each amino acid) compared to cells treated with vehicle ($n = 3$). **i**–**l**, Transamination (**i**), dehydrogenation (**j**) and further reduction or decarboxylation reactions in mouse liver lysates (**k**,**l**). All reactions were run in duplicates. All data points graphed are mean values; error bars, s.e.m. ND, not detected in either vehicle or AOA-treated cells. *$P < 0.01$; **$P < 0.001$; ***$P < 0.0001$ as calculated by a two-tailed *t*-test. IPyr, indole-3-pyruvate; PPyr, phenylpyruvate.

metabolic profiles of generally healthy individuals (ages 2 months to 55 years) against a cohort of paediatric patients with a variety of rare metabolic or genetic diseases. Although the intention was not to study the effect of antibiotics, many patients had detectable levels of antibiotics in their serum. Out of the 169 patients in the dataset, 13 had detectable circulating levels of one of the following antibiotics: moxifloxin, clindamycin, erythromycin, ciprofloxin (or its sulfate metabolite sulfociprofloxacin), doxycycline, metronidazole or sulfamethoxazole, and three patients had multiple of these antibiotics detected in circulation (Fig. 6a).

https://doi.org/10.1038/s42255-026-01550-8

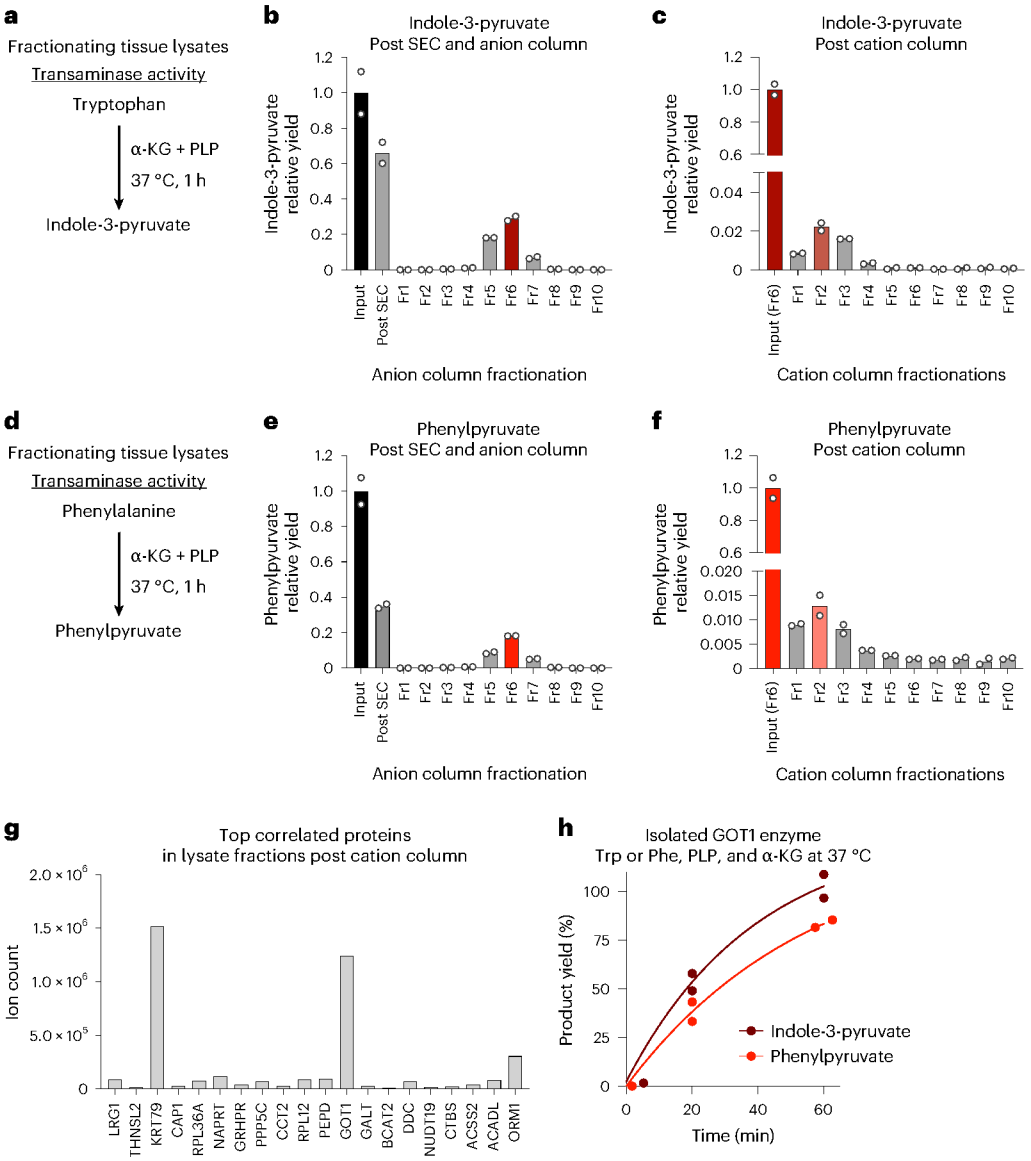


Fig. 5 | GOT1 produces indole-3-pyruvate and phenylpyruvate. a, Scheme for the tryptophan transaminase reaction in the fractions of mouse liver lysates. **b**, Indole-3-pyruvate production in the initial (input) mouse liver lysate, the most active SEC fraction (Fr) and further fractionation using an anion column. **c**, Further fractionation of fraction 6 from the anion column using a cation column. **d**–**f**, Same for phenylalanine transaminase. All reactions were run in duplicates, and activity was assayed after 1 h at 37 °C. Data are graphed as mean values; error bars, s.e.m. **g**, Raw ion counts of the proteins in fraction 1 after cation column fractionation. Data shown are the 20 proteins whose abundances best correlated with enzyme activity across the assayed fraction. **h**, Production of indole-3-pyruvate and phenylpyruvate by a commercial GOT1 enzyme. Reactions were run in duplicate. Data are shown as mean values; error bars, s.e.m.

When comparing serum from patients who had detectable levels of antibiotics (single or multiple) to those without, there was no change in any of the host-derived metabolites (indole-3-lactate, indole-3-acetate, indole-3-carboxylic acid, phenylpyruvate, phenyllactate and hydroxyphenyllactate; Fig. 6b–d). Indole-3-propionate and phenylpropionate were both significantly decreased in patients with one circulating antibiotic, and all microbiome-dependent metabolites (indole-3-propionate, indoxyl sulfate, 3-phenylpropionate, phenol sulfate and p-cresol sulfate) were depleted in patients with multiple circulating antibiotics. This dataset did not include levels of the other metabolites of interest.

To complement these data, we sought a second dataset with heavy antibiotic exposure, serum metabolite measurements and paired stool samples to enable full microbiome sequencing. We identified three patients with acute myeloid leukaemia (AML) who had faecal and serum samples taken before and after receiving broad-spectrum antibiotics (Fig. 6e and Table 1). In each patient, gut microbial diversity and relative abundance of obligate anaerobic taxa were dramatically lower in stool samples during the broad-spectrum antibiotic treatment. One patient demonstrated loss of diversity (Fig. 6f), one showed recovery of diversity after stopping antibiotics (Fig. 6g) and one exhibited recovery followed by loss upon re-initiation of antibiotics (Fig. 6h). Serum concentrations of the host-derived indole and phenol metabolites (indole-3-lactate, indole-3-acetate, indole-3-carboxylic acid, phenylpyruvate, phenyllactate and hydroxyphenyllactate) were not affected by microbiome compositional changes (Fig. 6i). By contrast, the microbiome-dependent metabolites indole-3-propionate and p-cresol sulfate were decreased in low-diversity, antibiotic-exposed samples. Other indoles and phenols were not detected in these patient samples. Therefore, from mouse to human, the circulating levels of many indole and phenol metabolites are robust to drastic changes in gut microbiota composition.

https://doi.org/10.1038/s42255-026-01550-8

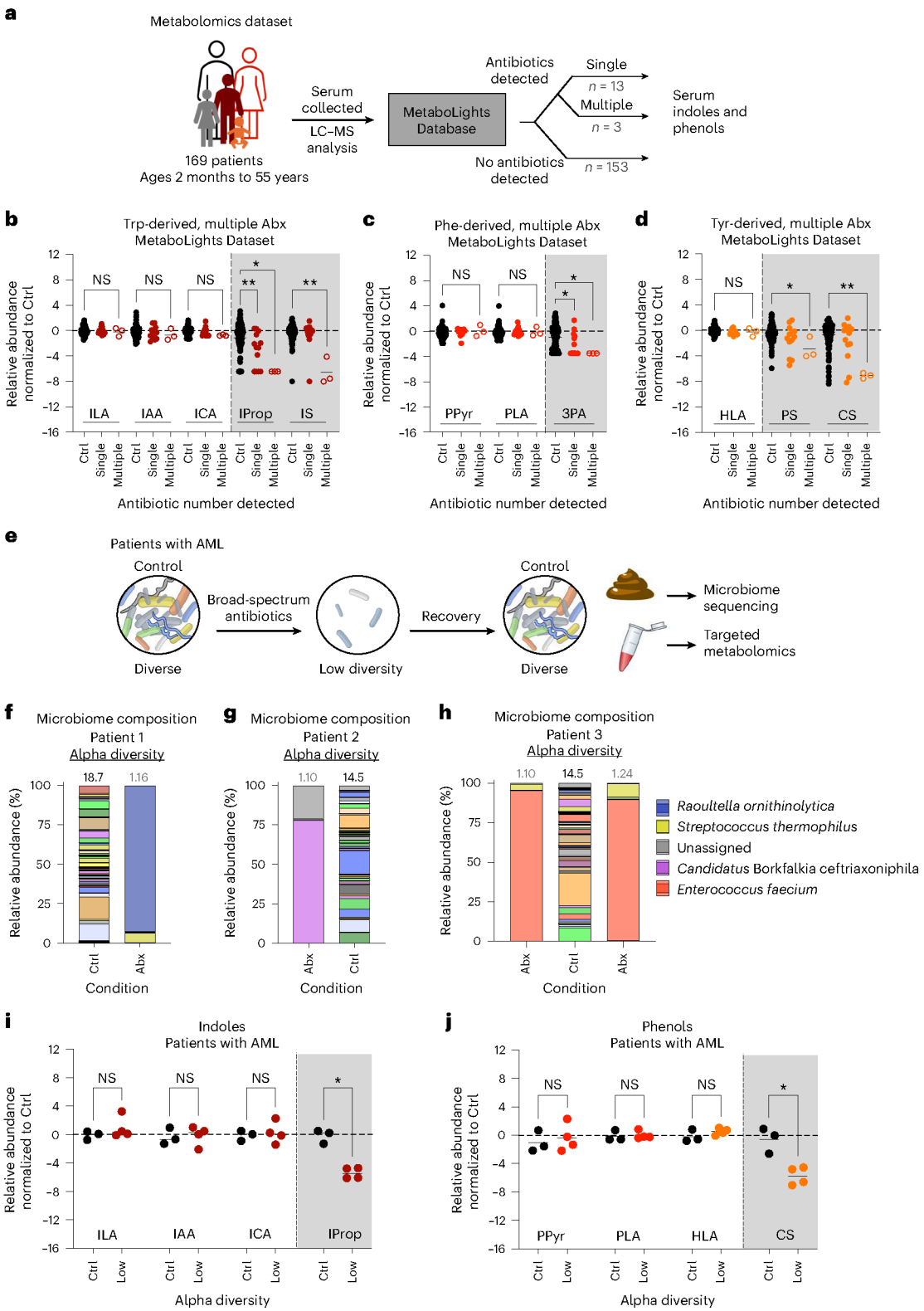


Fig. 6 | Metabolites produced by host metabolism are insensitive to antibiotics in clinical samples. a, Scheme depicting the analysis of serum indole and phenol metabolites in patients with or without antibiotics detected in serum by LC–MS. **b**–**d**, Relative (log_2-transformed fold change) serum tryptophan-derived (**b**), phenylalanine-derived (**c**) or tyrosine-derived (**d**) metabolites in patients without detectable levels of antibiotics ($n = 153$), patients with a single antibiotic detected ($n = 13$) or patients with multiple antibiotics detected in circulation ($n = 3$). $^*P < 0.05$; $^{***}P < 0.001$; $^{****}P < 0.0001$ by Kruskal–Wallis test (non-parametric ANOVA). NS, not significant. **e**, Scheme depicting the analysis of serum indole and phenol metabolites in patients with AML with or without the treatment of broad-spectrum antibiotics. Control denotes samples collected while patients were not receiving broad spectrum antibiotics. **f**–**h**, Microbial composition, as measured by Shotgun sequencing, and alpha diversity (inverse Simpson iIndex), for the three patients with AML. **i**,**j**, Relative indole (**i**) and phenol (**j**) metabolites for patients on days with standard or low alpha diversity. $^*P < 0.05$ and $^{**}P < 0.01$ as calculated by a two-sided Welch t-test. Abx, broad-spectrum antibiotics; Ctrl, control; ILA, indole-3-lactate; IAA, indole-3-acetate; ICA, indole-3-carboxylic acid; IProp, indole-3-propionate; IS, indoxyl sulfate; PLA, phenyllactate; 3PA, 3-phenylpropionate; HLA, hydroxyphenyllactate; PS, phenol sulfate; CS, p-cresol sulfate.

https://doi.org/10.1038/s42255-026-01550-8

Table 1 | Clinical details for patients with AML receiving antibiotics

Clinical details				Antibiotic treatment		
	Sex	Age	Race/ ethnicity	First sample	Second sample	Third sample
Patient 1	F	59	White	Day 0 6 days of levofloxacin	Day +19 12 days of cefepime 9 days of vancomycin 9 days of metronidazole 6 days of piperacillin/tazobactam	Not applicable
Patient 2	F	19	Hispanic	Day 0 20 days of piperacillin/tazobactam 5 days of doxycycline 3 days of vancomycin	Day +74 Cefidinir prophylaxis	Not applicable
Patient 3	M	63	Other	Day 0 24 days of piperacillin/tazobactam 10 days of daptomycin 6 days of cefdinir 3 days of vancomycin 1 day of metronidazole	Day +120 Cefdinir prophylaxis	Day +135 3 days of metronidazole 3 days of cefepime 2 days of fidaxomicin *Clostridium difficile* diagnosis on day of sample

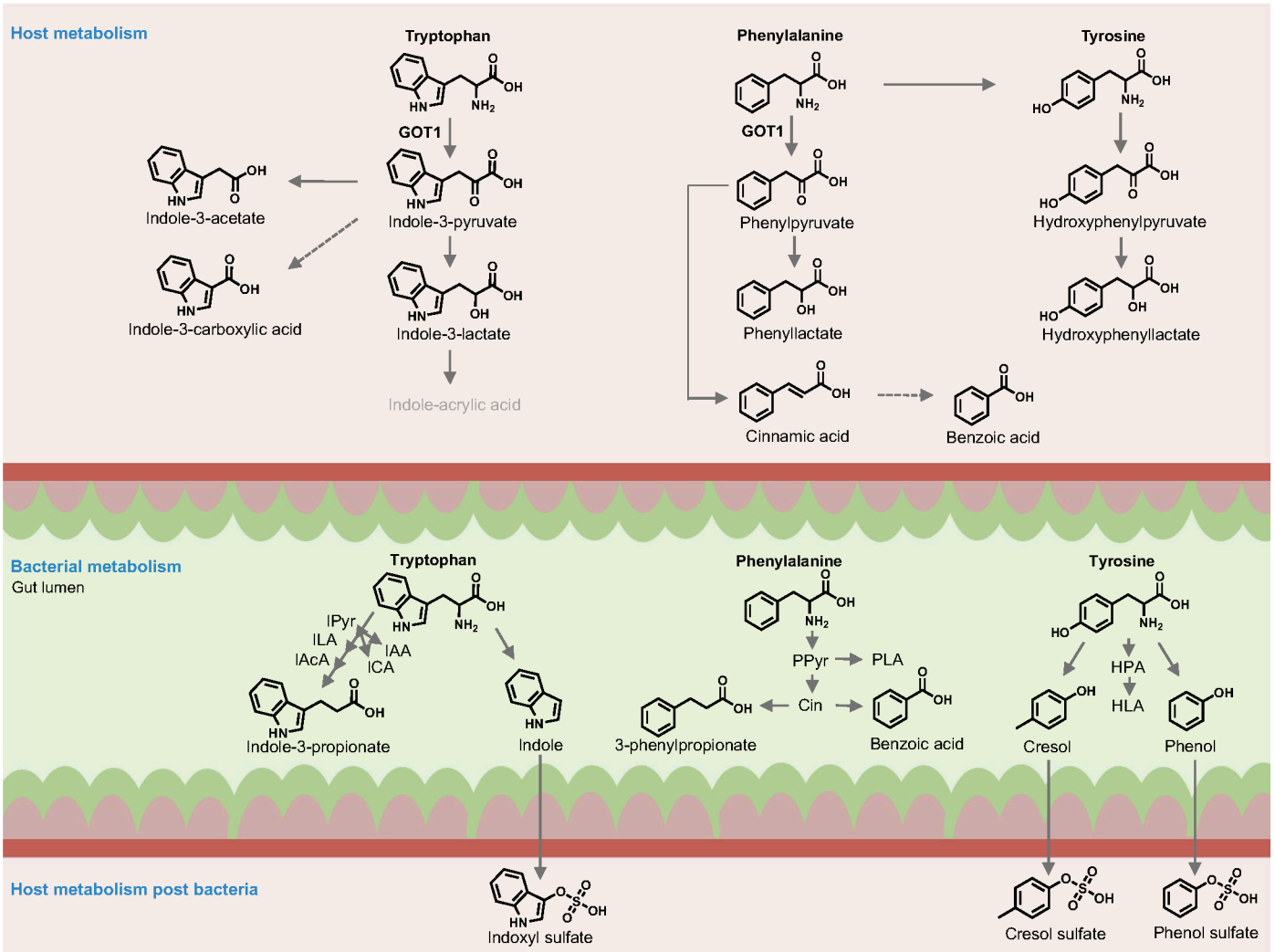


Fig. 7 | Metabolism of aryl amino acids to indole and phenol metabolites. Mammals encode the metabolic capacity to transaminate tryptophan, phenylalanine and tyrosine to their respective aryl-pyruvates and reduce to aryl-lactates. Additionally, indole-3-acetate, indole-3-carboxylic acid and cinnamic acid (Cin) are produced by host metabolism. Benzoic acid can also be made through host metabolism but is probably primarily a microbially derived metabolite, as isotope tracing of host metabolic pathways could only account for about 30% of benzoic acid production. Bacterial metabolism is required to make aryl-propionates, indole, phenol and p-cresol. Solid arrows represent reactions that we have directly shown in the tissue lysates or that are known in the literature. Dashed arrows represent proposed pathways. HPA, hydroxyphenylpyruvate; IAcA, indole-acrylic acid.

Discussion

We demonstrate, by isotope tracing, that many physiologically important circulating indole and phenol metabolites are derived from endogenous metabolism in mice, rats and cultured human cells. We conclude that there is a conserved metabolic pathway for all three aryl amino acids in mammalian cells that can convert tryptophan, phenylalanine and tyrosine to their respective aryl-pyruvates and aryl-lactates (Fig. 7).

Additionally, there is a mammalian metabolic pathway that can convert phenylalanine to cinnamic acid. Cinnamic acid appears to be made from phenylpyruvate directly, as there was no production of cinnamic acid from added phenyllactate. Although benzoic acid was not made in cell culture, it was produced in the mouse liver lysates and labelled from ^{13}C-phenylalanine over 2 h in vivo, which suggests that benzoic acid is at least partially derived from host metabolism.

Host metabolism is also capable of making indole-3-acetate and indole-3-carboxylic acid from tryptophan. Although the responsible mammalian enzymes are not known, indole-3-acetate was produced in mouse liver lysates from indole-3-pyruvate, the precursor of indole-3-acetate in bacteria and plants[35,36]. The pathway to indole-3-carboxylic acid remains an open question in both mammals and bacteria[37].

Furthermore, there is probably a similar metabolic pathway in humans, given that the indole and phenol metabolites that are host-derived in rodents are not affected by antibiotic-mediated depletion of the gut microbiome in humans. This conclusion is based on two studies. One is a retrospective analysis of metabolomics data from a cohort of 169 patients, of whom only 13 had antibiotics detected in circulation. The other involves three patients with cancer who went on or off broad-spectrum antibiotics. Although the findings of these studies are consistent, studies with larger patient populations are merited to validate the hypothesis that host metabolism is sufficient to maintain the circulating levels of these metabolites.

Not all indole and phenol metabolites can be made through host metabolic pathways. Bacterial metabolism is required for fully freeing the aryl group into indole, phenol or p-cresol. Additionally, the gut microbiome is necessary for reducing the amino acid all the way to indole-3-propionate or 3-phenylpropionate.

Indole-3-lactate and indole-3-acetate are frequently studied metabolites as they mediate important anti-inflammatory effects, with supplementation, in mouse models, ameliorating diseases ranging from inflammatory bowel disease to extreme nearsightedness[2,7,38,39]. Changes in the levels of these indoles have been correlated with microbiome composition changes in patient populations or mouse models. These correlations are classically interpreted as evidence that the microbiome metabolism is responsible for the production of the metabolites. Our data argue for caution in this interpretation. Although it is clear that the microbiome has the capacity to produce all the indole and phenol metabolites[12,28,40], and introducing bacteria with the ability to produce metabolites such as indole-3-lactate has been shown to increase circulating levels[28,41], we show that the contribution of host metabolism also needs to be considered in these correlations. For example, correlations may arise owing to diet or disease state affecting both microbial composition and host indole metabolite production or through crosstalk between host and microbiome metabolism.

Overall, an understanding of the source of indole and phenol metabolites is the fundamental step both to developing tools to control their circulating concentrations and to understanding the physiological factors controlling their levels and functions. Our data argue for host metabolism being a major source for physiologically important circulating metabolites such as indole-3-lactate, indole-3-acetate and cinnamic acid.

Methods

Mouse and rat studies

All mouse and rat studies were completed in compliance with protocols approved by the Princeton University Animal Care and Use Committee. Unless otherwise indicated, all mice were C57BL/6NCrl mice (Charles River Laboratories, strain 027). Sprague Dawley jugular-vein-catheterized rats were used (Charles River Laboratories, strain 001). Both mice and rats were 8–12 weeks old and fed a standard purified diet (Research Diets, D11112201). Mice were group-housed on a normal light–dark cycle (08:00–20:00 h), and rats were individually housed on a reverse light–dark cycle (20:30–08:30 h). The rats were acclimated to human handling for 2 weeks before experiments began. All rodents had ad libitum access to food and water.

Germ-free mouse model

C57BL/6N wild-type female germ-free or Murine Pathogen Free mice (Taconic, strain B6-F GF or MPF) were housed in individually vented cages (Allentown Sentry Sealed Positive Pressure system) under a 12 h light and 12 h dark cycle. All mice were 9–10 weeks old, group-housed and had free access to water and doubly irradiated standard purified diet (D11112201-1.5Vii). Contamination quality checks were performed on mouse faecal pellets, cage samples and research diet pellets using aerobic and anaerobic cultures. Samples were collected 10 days after initiation of the purified diet.

Human subject details

Patients with a diagnosis of AML who were receiving treatment at the University of Chicago Medicine were recruited to the protocol (IRB23-1365). Written informed consent was obtained, and there was no compensation for participation. Serial collection of serum and stool was performed while hospitalized. Patient samples with examples of microbiome diversity loss and gain were selected to determine whether key metabolites correlate with microbiome composition. All serum–faecal sample pairs were taken within 24 h of each other. Information requests regarding the samples can be directed to Christopher J. Lehmann.

In vivo sample collection and metabolite extraction

Serum samples were collected by tail bleed. Blood was placed on ice immediately following collection, centrifuged at 4 °C, 14,000*g* for 10 min and moved to another tube to be stored at −80 °C. Faeces were collected fresh, immediately placed on dry ice and stored at −80 °C until processing.

Serum samples were extracted with methanol (40× dilution) and incubated for 10 min on ice. Samples were then centrifuged at 15,000*g* for 30 min at 4 °C, and the resulting supernatant was transferred to a mass spectrometry vial. Faecal samples were collected fresh and placed on dry ice. Faecal and dietary samples were ground at liquid nitrogen temperature with a cryomill (Restch). The resulting powder was extracted with 40:40:20 methanol:acetonitrile:water (40× dilution) for 10 min on ice and centrifuged at 15,000*g* for 30 min. The resulting extract was moved to LC–MS vials (Thermo Fisher Scientific, 200046; caps, Thermo Fisher Scientific, 501313).

Metabolite measurement

An Orbitrap Exploris 240 or Q Exactive Plus hybrid quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific) coupled with hydrophilic interaction chromatography (HILIC) was used. For LC, an XBridge BEH Amide column (150 mm × 2.1 mm, 2.5 µM particle size; Waters) was used with gradient solvent A (95%: 5% H_2O: acetonitrile with 20 mM ammonium acetate, 20 mM ammonium hydroxide, pH 9.4) and solvent B (100% acetonitrile). The gradient ran as follows: 0 min, 90% B; 2 min, 90% B; 3 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; 14 min, 25% B; 16 min, 0% B, 20.5 min, 0% B; 21 min, 90% B; and 25 min, 90% B with a flow rate of 150 µl min^{-1}. Injection volume was 10 µl and column temperature was 25 °C. A negative-ion mode was used for the MS scans with a resolution of 140,000 at m/z 200. Automatic gain control target was 5×10^6, and scan range was m/z 75–1,000. Xcalibur (v.4.3; Thermo Fisher Scientific) was used to collect raw data.

All LC–MS data were analysed using El-Maven (v.0.12.1)[42]. All ^{13}C-labelling data were corrected for natural ^{13}C abundance using the Accucor package[43] (https://github.com/XiaoyangSu/AccuCor). All reported measurements are based on relative ion counts.

Metabolite assignment by LC–MS/MS

LC–MS/MS spectra were collected for sample metabolites and standards using the targeted MS2 scan function on an Orbitrap Exploris 240 at higher-energy collisional dissociation energy of 30 eV. Mass resolution of 15,000, *m/z* isolation window of 2, automated gain control of 5×10^5 and a maximum injection time of 200 ms were used. Unlabelled standards were from Sigma-Aldrich: tryptophan, phenylalanine, tyrosine, indole-3-pyruvate, indole-3-lactate, indole-3-carboxylic acid, indole-3-propionate, indoxyl sulfate, phenylpyruvate, phenyllactate, cinnamic acid, 3-phenylpropionate, hydroxyphenylpyruvate, hydroxyphenyllactate and phenol sulfate. D5-Benzoic acid (DLM-122), D7-Indole-3-acetic acid (DLM-8040) and D7-p-cresol sulfate (DLM-9786) were from Cambridge Isotope Laboratories. LC–MS/MS spectral data for compounds and standards can be found in Supplementary Table 1.

^{13}C-amino acid compounds

All amino acids were from Cambridge Isotope Laboratories ($^{13}C_9$-phenylalanine, CLM-2250-H; $^{13}C_9$-tyrosine, CNLM-439-H; and $^{13}C_{11}$-tryptophan, CLM-4290-H). Upon suspicion of contamination of the ^{13}C-tryptophan with ^{13}C-indole-3-acetate, the ^{13}C-tryptophan was further purified by Lotus Separations using a Phenomenex Luna NH_2 column (100 Å, 5 µm, 150 mm (length) × 21 mm (internal diameter, i.d.)). Isocratic conditions with 60% A (20 mM ammonium acetate, 20 mM ammonium hydroxide, 95:5 water:acetonitrile, pH 9.6) and 40% B (acetonitrile) at 45 °C and a flow rate of 15 ml min^{-1} were used.

Jugular vein and carotid artery catheterization

For jugular-vein-catheterized mice, using aseptic surgical techniques, a vascular access button (Instech, VABM1B/25, 25 gauge, one-channel button) was implanted under the back skin and connected to a catheter (Instech, C20PV-MJV1301, 2 French, 10 cm) placed in the right jugular vein. Jugular-vein-catheterized rats were purchased from Charles River Laboratories, with externalized catheters secured by a wound clip. Mice and rats were allowed at least 5 days to recover from surgery before any tracer infusion. Catheters were flushed with sterile saline and refilled with sterile heparin glycerol locking solution (SAI, HGS) at least once per week.

2 h ^{13}C-amino acid infusions

Mice were fed ad libitum for all infusions. At 18:00 h, mice were attached to the infusion line with a swivel and tether (Instech, swivel, SMCLA; line, KVABM1T/25) and acclimated for 1 h. With an infusion pump (SyringePump, NE-1000), the infusate was advanced at a rate of 0.1 µl min^{-1} per gram mouse weight starting at 19:00 h. Infusates were 5 mM ^{13}C-tryptophan, 2 mM ^{13}C-tyrosine or 15 mM ^{13}C-phenylalanine in sterile saline. For rat infusions, infusion lines were connected by a blunt needle at 07:00 h (during light phase). The lines included a swivel and tether (Instech, swivel, CM375KRP; line, KVABR1T/22 modified by removal of the vascular access button connector). Infusions were started at 08:00 h at a rate of 0.5 µl min^{-1} per gram weight using the SyringePump pump. Infusates were 5 mM ^{13}C-tryptophan and 30 mM ^{13}C-phenylalanine. After 2 h of infusions, serum was collected by tail snip and faeces were collected fresh. All collected samples were stored at −80 °C.

36 h infusion with and without antibiotic treatment

The antibiotic cocktail (1 g l^{-1} ampicillin, 1 g l^{-1} neomycin, 1 g l^{-1} metronidazole, 0.5 g l^{-1} vancomycin and 0.5% aspartame) was administered through drinking water. Infusion of the ^{13}C-tracers (5 mM ^{13}C-tryptophan, 2 mM ^{13}C-tyrosine, 15 mM ^{13}C-phenylalanine) was initiated on day 12 with or without antibiotics at 19:00 h, and mice were maintained on antibiotics throughout the infusion. Serum and faeces were collected before infusion started and 2, 12 and 36 h after initiation of the infusion. Mice were fed ad libitum throughout the infusion.

Alpha diversity measurements in antibiotic-treated mice

Bacterial DNA was extracted from faecal samples with the Power Soil DNA Isolation kit (QIAGEN). The V4 region of the 16S rRNA gene was amplified using the primers 27F (AGRGTTYGATYMTGGCTCAG) and 1492R (CGGYTACCTTGTTACGACTT) and LongAmp Taq 2X Master Mix (New England Biolabs, M0287L). Thermocycler conditions were 94 °C for 3 min, 30 cycles of 94 °C for 30 s, 55 °C for 30 s, 65 °C for 90 s, followed by 10 min hold at 65 °C. The amplified V4 region DNA was isolated by AMPure XP beads. Correct amplicon size was confirmed by gel electrophoresis. Amplicons were quantified using the Qubit dsDNA HS assay kit (Life Technologies, Q32854) and concentrations were standardized to 15 ng µl^{-1}. Library preparations were prepared using the Native Barcoding Kit 96 V14 (SQK-NBD114.96). Sequencing was completed using full-length 16S rRNA gene amplicon sequencing using an Oxford Nanopore MinION sequencer. Species-level taxonomic annotation was assigned to reads with Emu (v.3.4.5) using default settings and database[44–46]. Phyloseq[47] was used for alpha diversity (Shannon index) analysis.

Bacterial load quantification

Bacterial density was estimated using universal 16S qPCR with primers 515F (GGACTACNVGGGTWTCTAAT) and 806R (GTGYCAGCMGCCGCGGTAA). Reaction mixture included 1 µl of template DNA, 12.5 µl PerfeCTa SYBR Green SuperMix Reaction Mix (QuantaBio), 6 µl nuclease-free H_2O and 2.25 µl of each primer. Thermocycler conditions were 95 °C for 10 min, 40 cycles of 95 °C for 15 s, 60 °C for 40 s and 72 °C for 30 s. Reactions were performed in triplicate. Bacterial DNA concentrations were normalized to whole-genome equivalents based on a mean genome size of 4.50 Mbp, then multiplied by the total template DNA volume (50 µl). DNA concentrations were normalized by the grams of faeces used for DNA isolation.

Gavage with ^{13}C-amino acids

Mice were fasted for 7 h before a gavage (10 ml kg^{-1}) with 15 mM $^{13}C_9$-phenylalanine, 5 mM ^{13}C11-tryptophan and 2 mM $^{13}C_9$-^{15}N-tyrosine. Serum and caecal content were collected at 15, 45, 90 or 180 min following the gavage.

Feeding with ^{13}C-digestion-resistant algal protein

Mice were fed a diet with 20% overall protein (10% ^{13}C-algal protein (Sigma-Aldrich, 642878) and 10% rice protein isolate (NAKED Supplements)). Rice protein was used as both rice and algal protein are digestion-resistant. Blood from the portal vein and caecal contents were collected 24 h after diet initiation (08:00 h). For portal vein collection, mice were anaesthetized with isoflurane, and the portal vein was exposed by carefully displacing the intestines. The portal vein was nicked with dissection scissors, and pooled blood was collected from the peritoneal cavity by pipette and transferred to an Eppendorf tube on ice.

Cell culture

All cell lines were obtained from ATCC. Stock cells were maintained in either DMEM (HCT116, CCL-247; A549, CCL-185; HepG2, HB-8065; and HEK293, CRL-1573) or RPMI (THP-1, TIB-202; Jurkat, CRL-2899) with 10% complete FBS. All cell lines were cultured at 37 °C and 5% CO_2. Cells were grown to approximately 70–80% confluence. Comparable cell density was confirmed visually and by cell count with a Countess II Hemocytometer. Cells were routinely tested for *Mycoplasma* contamination using the PlasmoTest *Mycoplasma* Detection Kit (Invivogen).

^{13}C-tracing in cell culture

All cell lines were plated into six-well plates. Then, 24 h after plating, the media was aspirated and switched to media containing 10% dialyzed FBS and one of the ^{13}C-aryl amino acids added (DMEM: 16 mg ml^{-1} ^{13}C-tryptophan, 66 mg ml^{-1} ^{13}C-phenylalaine or 80 mg ml^{-1} ^{13}C-tyrosine; RPMI: 5 mg ml^{-1} ^{13}C-tryptophan, 15 mg ml^{-1} ^{13}C-phenylalaine or 29 mg ml^{-1} ^{13}C-tyrosine). Cells were collected by scraping in 500 µl ice-cold 80% methanol after media aspiration. To inhibit transaminases, HCT116 cells were treated with 500 µM aminooxyacetic acid (Sigma-Aldrich, C13408) or sterile water in DMEM with dialyzed media for 24 h.

Il4i1$^{-/-}$ plasma samples

Plasma samples from three female (22–23 weeks old) and two male (16 weeks old) whole body *Il4i1*-knockout mice and three female (22–23 weeks old) and two male (16 weeks old) wild-type mice were provided by C. Opitz. All mice were fed ad libitum with irradiated extruded chow pellets (KLIBA, 3437), and at 14 days of age, they also began receiving autoclaved oat flakes. Plasma samples were stored at −80 °C and prepared for MS analysis in the same way as serum samples.

Tissue lysates

Freshly collected mouse tissues were homogenized using a KIMBLE Dounce tissue grinder set (Sigma-Aldrich, DWK885300-0000) in a 1:4 ratio of tissue to assay buffer (100 mM Tris NaCl, 1 mM dithiothreitol (DTT) and 1× Halt Protease Inhibitor Cocktail, EDTA-Free (Thermo Scientific)). After homogenization, the sample was centrifuged at 4,000*g* for 5 min. The supernatant was diluted further 1:1 with assay buffer containing the relevant substrates and cofactors, and a baseline sample was taken immediately (and diluted by 500× in ice-cold methanol).

Cell lysates

HCT116 cells at 50–70% confluence were trypsinized, pelleted and washed twice in ice-cold PBS before resuspending in lysis buffer (150 mM NaCl, 20 mM Tris-HCl at pH 8, 1 mM beta-mercaptoethanol and 1× complete mini protease inhibitor cocktail with ethylene-diaminetetraacetic acid (Thermo Fisher, 78429)). Pelleted cells were lysed with Emulsiflex (Avestin). Lysed cells were spun for 60 min at 4 °C at 10,000 rpm and allocated into 100 µl aliquots with 20% glycerol, then snap-frozen with liquid nitrogen.

In vitro activity assays

For all in vitro transaminase reactions, including isolated GOT1 (from porcine heart; Sigma-Aldrich 9000-97-9), activity was assessed by the addition of either phenylalanine or tryptophan (2 mM) with PLP (200 µM) and α-KG (20 mM). For dehydrogenase and downstream reaction conditions, 2 mM indole-3-pyruvate or phenylpyruvate was added to lysates, and, except for indole-3-acetate and indole-3-carboxylic acid producing reactions, 20 mM NADH was added to the reaction mixtures. All lysates were placed at 37 °C and protected from light. At the designated times, samples from the lysates were taken and diluted 500× in ice-cold methanol. Reactions were run in duplicate, and metabolite concentrations were quantified by LC–MS using internal spike-in standards.

Protein fractionation

Protein fractionation was completed on an ÄKTA pure chromatography system at 4 °C at a flow rate of 1 ml min^{-1}. For SEC, a HiPrep 16/60 Sephacryl S 200 HR column with a column volume (CV) of 120 ml was equilibrated with 2 CVs of running buffer before the application of 1 ml of sample. Proteins were eluted isocratically over 1.5 CVs of running buffer, and 5 ml fractions were collected. The running buffer composition was 150 mM NaCl, 25 mM Tris (pH 8.0), 1 mM TCEP and 10% glycerol.

For anion exchange chromatography, a HiTrap Capto Q ImpRes column (CV of 1 ml) was equilibrated with 5 CVs of 2.5% buffer B before application of 1 ml of sample. The gradient conditions were as follows: 2.5% buffer B, 5 CVs; ramp to 100% buffer B over 20 CVs; 100% buffer B, 5 CVs; 2.5% buffer B, 5 CVs. Next, 1 ml fractions were collected starting at the ramp step. Buffer A was 25 mM Tris-HCl (pH 8.0), 1 mM TCEP, 10% glycerol, and buffer B was 1 M NaCl, 25 mM Tris (pH 8.0), 1 mM TCEP, 10% glycerol.

For cation exchange chromatography, a HiTrap Capto SP ImpRes column (CV of 1 ml) was used under the same chromatography conditions as for anion exchange, except for the buffer composition. Buffer A was 25 mM Bis-Tris (pH 6.0), 1 mM TCEP, 10% glycerol and buffer B was 1 M NaCl, 25 mM Bis-Tris (pH 6.0), 1 mM TCEP, 10% glycerol.

For sequential fractionation, the fraction of interest from the first chromatographic run was exchanged into the initial buffer composition for the second run using a HiTrap Desalting column (CV of 5 ml). The column was equilibrated with 1.5 CVs of the target buffer before application of 1 ml of the sample. Protein was eluted isocratically in 1.5 CVs of the target buffer, and 0.5 ml fractions were collected. Fractions containing >90% of the protein according to the absorbance readings at 280 nm were pooled for the second round of fractionation. SEC fractions were first concentrated from 5 ml to 1 ml using Amicon Ultra-15 spin filter tubes with a 10 kDa cutoff.

Proteomics

Fraction 6 from the anion column and fractions 1, 3, 5 and 7 following the cation column were used for proteomic analysis. For each tested fraction, a 10 µg sample was diluted with RIPA buffer (125 µl). DTT was added to a final concentration of 10 mM, and lysate was heated at 95 °C for 5 min. Then, iodoacetamide was added to a final concentration of 15 mM and tubes were incubated in the dark with rotation. After 30 min, an additional 10 mM DTT was added (final DTT concentration of 20 mM) and incubated at room temperature with rotation. After 15 min, 2 µl of a prepared 50 µg µl^{-1} Sera-Mag bead stock was added, followed by 450 µl ethanol, and the samples were incubated at room temperature for 5 min with shaking at 1,000 rpm. Beads were then pelleted on a magnetic rack and supernatant was discarded. The beads were washed three times with 200 µl 80% ethanol before being resuspended in 50 mM ammonium bicarbonate. 0.5 µg trypsin was added to each tube and sonicated for 30 s in a water bath, followed by incubation for 18 h at 37 °C with rotation. Then, 1 µl Optima LC–MS formic acid was added, and supernatants were passed through a 0.22 µm spin filter column (CoStar, 98231-UT-1). The resulting sample was transferred to an LC–MS vial.

Label-free, data-independent acquisition (DIA) proteomics was performed on Bruker TimsTOF Pro 2 connected to a nanoElute LC. A 1 µl aliquot of each sample was injected onto a trap column (C18 Pepmap, 5 µM particle size, 5 mm length, 300 µM i.d.), followed by separation by an analytical column (C18 ReproSil AQ, 1.9 µM particle size, 100 mm length, 75 µM i.d.). Peptides were eluted by an acetonitrile/water gradient at a column temperature of 40 °C (buffer A, 0.1% formic acid/water; buffer B, 0.1% formic acid/acetonitrile; flow rate of 0.5 µl min^{-1}; gradient: start at 2% B, increase to 35% B over 20 min, then increase to 95% B over 0.5 min, hold at 95% for 2.25 min). Scans were performed in positive ion, diaPASEF mode over a m/z range of 100–1,700 and ramp time of 100 ms, accumulation time of 100 ms and duty cycle of 100%, ramp rate of 9.43 Hz, MS averaging set to 1. Absolute thresholds were set to 5,000 for mobility peaks and 10 for MS peaks; data were collected with Bruker Compass HyStar v.6.2. Resulting raw data (.d files) were processed using DIA-NN v.1.8.1 with the following parameters: trypsin/P digestion, three missed cleavages, three max. variable modifications, N-term M excision, Ox(M), Ac(N-term) and C-carbamidomethylation, peptide length range of 7–30, precursor charge range 1–4, m/z range 300–1,800, fragment ion range 200–1,800, mass accuracy and MS accuracy both set to ten, precursor false discovery rate set to 1%. Within the DIA-NN algorithm, the following settings were applied: 'Use isotopologues', 'MBR' (match between runs), 'No shared spectra' and 'Heuristic protein

inference'. A spectral library generated in DIA-NN from all known *Mus musculus* proteins was used (in silico spectral library, generated in DIA-NN using FASTA of Uniprot *M. musculus* proteome UP000000589, options selected were 'FASTA digest for library-free search/library generation' and 'Deep learning-based spectra, RTs and IMs prediction'; other parameters same as described above). After processing, matrix.pg files were analysed in Perseus (v.2.0.7.0)[48]; intensities are inputted as 'main' while the other descriptors are listed as 'categorical'. Data are presented as raw ion counts in each chromatography fraction obtained during purification.

Boston Children's patient dataset

Patient sample data were obtained from a study completed at Boston Children's Hospital[33]. The metabolomics dataset from this study was deposited in MetaboLights[33,34]. A missing value imputation was performed, allowing for samples in which the metabolite was below detectable limits to be filled with the minimally detected ion count for that metabolite. Out of this dataset, patients with antibiotics were selected based on antibiotics detected in serum.

Metabolomics and microbiome sequencing of samples from patients with AML

For metabolic analysis details, methods can be accessed under MassIVE MSV000099027. In brief, for extraction, four volumes of extraction solvent (100% methanol spiked with 500 nM (acetyl-)D3-melatonin (Cambridge Isotope Laboratories, DLM-7101-0 as internal standard)) were added to one volume (200 µl) of plasma. Samples were incubated at −80 °C for at least 1 h to overnight. Indole metabolites were analysed on an Agilent 1290 Infinity II liquid chromatography system coupled to an Agilent 6470 triple quadrupole mass spectrometer, operating in positive mode, with an Agilent Jet Stream Electrospray Ionization source. Data analysis was performed using MassHunter Quant software (v.B.10, Agilent Technologies) and confirmed by comparison with authentic standards (Supplementary Table 2). For phenol metabolites, plasma samples and standards were analysed on a Thermo Fisher liquid chromatography system coupled to an Orbitrap IQ-X mass spectrometer, operating in positive mode. Raw data files were analysed in Thermo Freestyle software. Peak integration using the ICIS algorithm in Freestyle was done on extracted ion chromatograms for the m/z of each compound within 3 ppm of the m/z in each standard compound injection (Supplementary Table 3). A missing value imputation was performed, allowing for samples in which the metabolite was below detectable limits to be filled with the minimally detected ion count for that metabolite.

For sequencing, shotgun DNA sequencing was used for the faecal samples from patients with AML. A bead beater (BioSpec Product) was used for mechanical disruption, followed by sample purification with QIAamp mini spin columns (Qiagen). Purified DNA was quantified by a Qubit 2.0 fluorometer, and an Illumina HiSeq platform was used for sequencing. A total of 7–8 million paired-end DNA sequencing reads (read length of 150 bp) were produced per sample. Adaptors were trimmed from the raw reads, and quality was assessed using Trimmomatic (v.0.39)[49]. Human genome reads were removed by kneaddata (https://github.com/biobakery/kneaddata, v.0.7.10), and taxonomy of the resultant high-quality reads was profiled with metaplan4 (ref. 50). MEGAHITS (v.1.2.9)[51] was used to assemble microbial reads with genes called by prodigal (https://github.com/hyattpd/Prodigal, v.2.6.3).

Statistical analysis and figure production

A P value of <0.05 was used to determine statistical significance using a two-tailed, unpaired Student's t-test. No statistical methods were used to pre-determine sample sizes. Data collection and analysis were not performed blind to the conditions of the experiments, and there was no randomization. Unless otherwise stated, data distribution was assumed to be normal, but this was not formally tested. No animals were excluded unless the infusions did not reach their desired endpoint, either through technical error or death of the mouse. Figures were made with GraphPad Prism (v.10.2.3), ChemDraw (v.23.1.1) and NIH BioArt (https://bioart.niaid.nih.gov).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data generated or analysed during this study are included in this article and its Supplementary Information files. The datasets supporting the findings are available from the corresponding author upon reasonable request. MS metabolomics data for the Boston Children's study are available from MetaboLights (MTBLS948). Microbiome sequencing and metabolomic data from the patients with AML have been deposited on Zenodo (https://doi.org/10.5281/zenodo.18829320)[52]. All MS proteomics data have been deposited to the MassIVE database (accession no.: MSV000100980), including an in silico spectral library generated in DIA-NN using FASTA of Uniprot mouse proteome UP000000589. Source data are provided with this paper.

References

1. Wu, Y. et al. Cinnamic acid ameliorates nonalcoholic fatty liver disease by suppressing hepatic lipogenesis and promoting fatty acid oxidation. *Evid. Based Complement. Alternat. Med.* **2021**, 9561613 (2021).
2. Lavelle, A. & Sokol, H. Gut microbiota-derived metabolites as key actors in inflammatory bowel disease. *Nat. Rev. Gastroenterol. Hepatol.* **17**, 223–237 (2020).
3. Agus, A., Planchais, J. & Sokol, H. Gut microbiota regulation of tryptophan metabolism in health and disease. *Cell Host Microbe* **23**, 716–724 (2018).
4. Alexeev, E. E. et al. Microbiota-derived indole metabolites promote human and murine intestinal homeostasis through regulation of interleukin-10 receptor. *Am. J. Pathol.* **188**, 1183–1194 (2018).
5. Kikuchi, K. et al. Gut microbiome-derived phenyl sulfate contributes to albuminuria in diabetic kidney disease. *Nat. Commun.* **10**, 1835 (2019).
6. Vanholder, R., De Smet, R. & Lesaffer, G. P-Cresol: a toxin revealing many neglected but relevant aspects of uraemic toxicity. *Nephrol. Dial. Transplant.* **14**, 2813–2815 (1999).
7. Li, M. et al. Gut microbiota metabolite indole-3-acetic acid maintains intestinal epithelial homeostasis through mucin sulfation. *Gut Microbes* **16**, 2377576 (2024).
8. Dimopoulou, C. et al. Potential of using an engineered indole lactic acid producing *Escherichia coli* Nissle 1917 in a murine model of colitis. *Sci. Rep.* **14**, 17542 (2024).
9. Zhao, Z. H. et al. Indole-3-propionic acid inhibits gut dysbiosis and endotoxin leakage to attenuate steatohepatitis in rats. *Exp. Mol. Med.* **51**, 1–14 (2019).
10. Mnafgui, K. et al. Anti-obesity and cardioprotective effects of cinnamic acid in high fat diet-induced obese rats. *J. Food Sci. Technol.* **52**, 4369–4377 (2015).
11. Saito, Y., Sato, T., Nomoto, K. & Tsuji, H. Identification of phenol- and p-cresol-producing intestinal bacteria by using media supplemented with tyrosine and its metabolites. *FEMS Microbiol. Ecol.* **94**, fiy125 (2018).
12. Dodd, D. et al. A gut bacterial pathway metabolizes aromatic amino acids into nine circulating metabolites. *Nature* **551**, 648–652 (2017).
13. Zelante, T. et al. Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22. *Immunity* **39**, 372–385 (2013).

14. Lamas, B. et al. CARD9 impacts colitis by altering gut microbiota metabolism of tryptophan into aryl hydrocarbon receptor ligands. *Nat. Med.* **22**, 598–605 (2016).
15. Zheng, X. et al. The footprints of gut microbial-mammalian co-metabolism. *J. Proteome Res.* **10**, 5512–5522 (2011).
16. Asnicar, F. et al. Microbiome connections with host metabolism and habitual diet from 1,098 deeply phenotyped individuals. *Nat. Med.* **27**, 321–332 (2021).
17. Wu, J. et al. Microbiota-induced alteration of kynurenine metabolism in macrophages drives formation of creeping fat in Crohn's disease. *Cell Host Microbe* **32**, 1927–1943.e9 (2024).
18. Krisko, T. I. et al. Dissociation of adaptive thermogenesis from glucose homeostasis in microbiome-deficient mice. *Cell Metab.* **31**, 592–604.e9 (2020).
19. Sadik, A. et al. IL4I1 is a metabolic immune checkpoint that activates the AHR and promotes tumor progression. *Cell* **182**, 1252–1270.e34 (2020).
20. Zhang, X. et al. Endogenous indole pyruvate pathway for tryptophan metabolism mediated by IL4I1. *J. Agric. Food Chem.* **68**, 10678–10684 (2020).
21. Boulland, M. L. et al. Human IL4I1 is a secreted L-phenylalanine oxidase expressed by mature dendritic cells that inhibits T-lymphocyte proliferation. *Blood* **110**, 220–227 (2007).
22. Morgan, E. W. et al. Contribution of circulating host and microbial tryptophan metabolites toward Ah receptor activation. *Int. J. Tryptophan Res.* **16**, 11786469231182510 (2023).
23. Jang, C., Chen, L. & Rabinowitz, J. D. Metabolomics and isotope tracing. *Cell* **173**, 822–837 (2018).
24. Zeng, X. et al. Gut bacterial nutrient preferences quantified in vivo. *Cell* **185**, 3441–3456.e19 (2022).
25. Demarco, M. et al. Digestibility, bioaccessibility and bioactivity of compounds from algae. *Trends Food Sci. Technol.* **121**, 114–128 (2022).
26. Rong, Y. & Kiang, T. K. L. Characterization of human sulfotransferases catalyzing the formation of p-cresol sulfate and identification of mefenamic acid as a potent metabolism inhibitor and potential therapeutic agent for detoxification. *Toxicol. Appl. Pharmacol.* **425**, 115553 (2021).
27. Tsiara, I. et al. Characterizing the sulfated and glucuronidated (poly)phenol metabolome for dietary biomarker discovery. *J. Agric. Food Chem.* **73**, 6702–6710 (2025).
28. Qian, X. et al. Bifidobacteria with indole-3-lactic acid-producing capacity exhibit psychobiotic potential via reducing neuroinflammation. *Cell Rep. Med.* **5**, 101798 (2024).
29. Lloyd-Price, J. et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* **569**, 655–662 (2019).
30. Mallaret, G. et al. The autism-linked gut microbial metabolite p-cresol inhibits host catecholamine biosynthesizing enzymes to elicit social deficits. *Commun. Biol.* **8**, 1800 (2025).
31. Pham, N. M., Recht, N. S., Hostetter, T. H. & Meyer, T. W. Removal of the protein-bound solutes indican and p-cresol sulfate by peritoneal dialysis. *Clin. J. Am. Soc. Nephrol.* **3**, 85–90 (2008).
32. Poesen, R., de Loor, H., Kuypers, D., Augustijns, P. & Meijers, B. Metabolism, protein binding, and renal clearance of microbiota-derived p-cresol in patients with CKD. *Clin. J. Am. Soc. Nephrol.* **11**, 1136–1144 (2016).
33. Kong, S. & Hernandez-Ferrer, C. Assessment of plasma metabolomic biomarkers in various pediatric diseases. *Environ. Epidemiol.* **3**, 213 (2019).
34. Yurekten, O. et al. MetaboLights: open data repository for metabolomics. *Nucleic Acids Res.* **52**, D640–D646 (2024).
35. Cao, X. et al. The roles of auxin biosynthesis YUCCA gene family in plants. *Int. J. Mol. Sci.* **20**, 6343 (2019).
36. Spaepen, S., Vanderleyden, J. & Remans, R. Indole-3-acetic acid in microbial and microorganism–plant signaling. *FEMS Microbiol. Rev.* **31**, 425–448 (2007).
37. Jia, D., Kuang, Z. & Wang, L. The role of microbial indole metabolites in tumor. *Gut Microbes* **16**, 2409209 (2024).
38. Li, H. et al. Gut microbiota-derived indole-3-acetic acid suppresses high myopia progression by promoting type I collagen synthesis. *Cell Discov.* **10**, 89 (2024).
39. Ye, X. et al. Dual role of indoles derived from intestinal microbiota on human health. *Front. Immunol.* **13**, 903526 (2022).
40. Sinha, A. K. et al. Dietary fibre directs microbial tryptophan metabolism via metabolic interactions in the gut microbiota. *Nat. Microbiol.* **9**, 1964–1978 (2024).
41. Zhu, P. et al. Exploring the fecal metabolome in infants with cow's milk allergy: the distinct impacts of cow's milk protein tolerance acquisition and of synbiotic supplementation. *Mol. Nutr. Food Res.* **69**, e202400583 (2025).
42. D'Alessandro, A. (ed.) *High-Throughput Metabolomics: Methods and Protocols* (Humana Press, 2019).
43. Su, X., Lu, W. & Rabinowitz, J. D. Metabolite spectral accuracy on orbitraps. *Anal. Chem.* **89**, 5940–5948 (2017).
44. Stoddard, S. F., Smith, B. J., Hein, R., Roller, B. R. K. & Schmidt, T. M. rrnDB: Improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Res.* **43**, D593–D598 (2015).
45. O'Leary, N. A. et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **44**, D733–D745 (2016).
46. Schoch, C. L. et al. NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database* **2020**, baaa062 (2020).
47. McMurdie, P. J. & Holmes, S. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217 (2013).
48. Tyanova, S. et al. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740 (2016).
49. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
50. Blanco-Míguez, A. et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. *Nat. Biotechnol.* **41**, 1633–1644 (2023).
51. Li, D., Liu, C. M., Luo, R., Sadakane, K. & Lam, T. W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).
52. sulakhe-uc. uchicago-dfi-informatics/Metabolite-and-gut-microbial-composition-in-AML-Patients: March 1, 2026. *Zenodo* https://doi.org/10.5281/zenodo.18829320 (2026).

Acknowledgements

This research is the result of funding in whole or in part provided by the National Institutes of Health (NIH). It is subject to the NIH Public Access Policy. Through acceptance of this federal funding, NIH has been given the right to make this paper publicly available in PubMed Central upon the Official Date of Publication, as defined by NIH. Thank you to J. Pratas at Lotus Separations for purifying the ^{13}C-tryptophan tracer. J.E.A. is funded by F30DK139739 and M.S.D. by R01AI172144, both from the NIH. J.D.R. was supported by Ludwig Institute for Cancer Research, Princeton Branch, DP1DK113643 from the National Institute of Diabetes, Digestion and Kidney Disease and an internal grant from the Princeton Alliance for Collaborative Research and Innovation, Princeton University. We are thankful to all members of the Rabinowitz lab for their scientific discussion and insightful input.

Author contributions

J.E.A. and J.D.R. conceived and designed the study. J.E.A. performed the experiments, completed data analysis and made the figures. K.O. and S.Y. assisted with in vitro assays and germ-free mouse experiments, respectively. S.M. and C.J.H. performed rat infusions. C.J.H. performed the catheterization surgeries. J.A.B. and S.D.K. completed the proteomics analysis under the supervision of D.W.C.M. J.S. and A.S. generated the UChicago LC–MS data. L.S. completed MS/MS data analysis. M.R.M. performed data analysis on the patient database. R.P.R performed the *Mycoplasma* testing. A.H. and C.A.O. provided plasma samples from *Il4i1*$^{-/-}$ mice. M.S.D. provided crucial insights for project conceptualization. S.I. and C.J.L. designed the studies at the University of Chicago, obtained ethical approval, enrolled patients, developed sample collection and handled protocols. E.P. and O.O. oversaw microbial sequencing for the patient samples. J.E.A. and J.D.R. wrote the manuscript. All authors reviewed the manuscript.

Competing interests

J.D.R. is a member of the Rutgers Cancer Institute of New Jersey (RCINJ) and the University of Pennsylvania Diabetes Research Center (U Penn DRC); a director of the U Penn DRC–Princeton inter-institutional metabolomics core and RCINJ metabolomics core; an advisor and stockholder in Colorado Research Partners, Bantam Pharmaceuticals and Rafael Pharmaceuticals; a founder, director and stockholder of Farber Partners, Raze Therapeutics and Sofro Pharmaceuticals; a founder, advisor and stockholder in Marea Therapeutics; an advisor and stockholder in Empress Therapeutics (which investigates microbiome metabolites); and inventor of patents held by Princeton University. The remaining authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s42255-026-01550-8.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42255-026-01550-8.

Correspondence and requests for materials should be addressed to Joshua D. Rabinowitz.

Peer review information *Nature Metabolism* thanks Henrik Roager, Alpaslan Tasdogan, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Yunling Xie and Alfredo Gimenez-Cassina, in collaboration with the *Nature Metabolism* team.

Reprints and permissions information is available at www.nature.com/reprints.

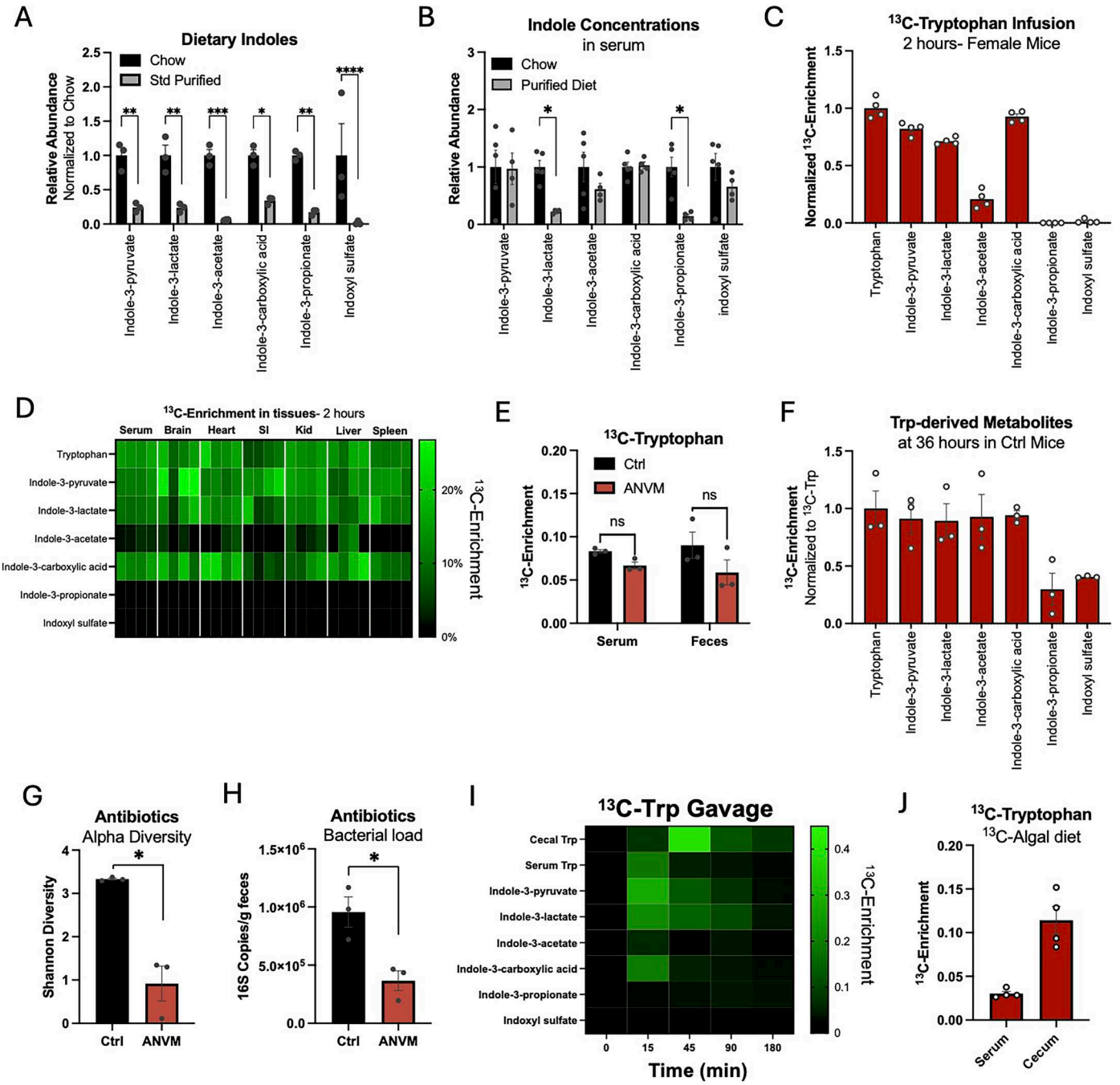
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2026

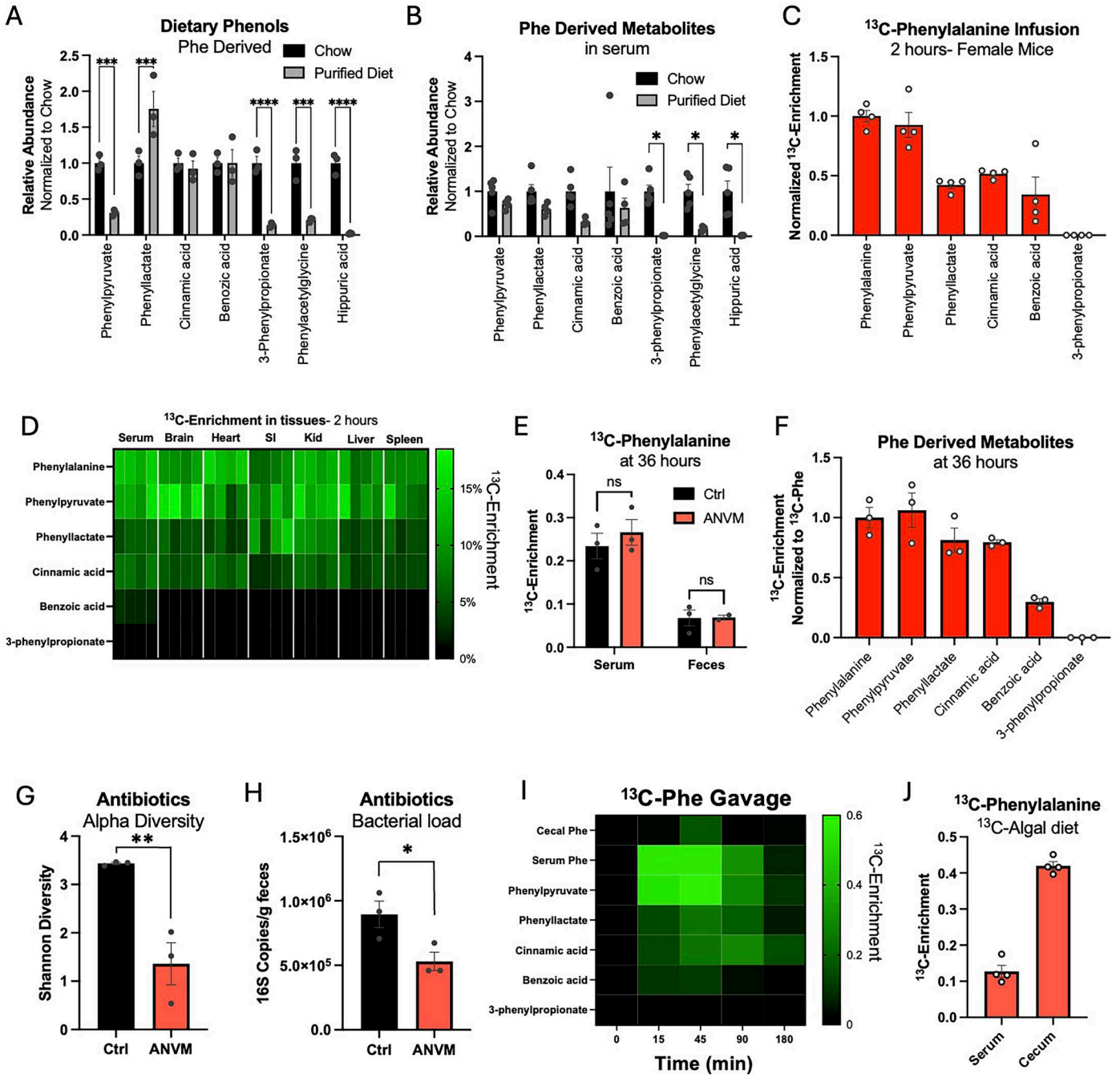
[1]Department of Molecular Biology, Princeton University, Princeton, NJ, USA. [2]Ludwig Institute for Cancer Research, Princeton University, Princeton, NJ, USA. [3]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA. [4]Duchossois Family Institute, University of Chicago, Chicago, IL, USA. [5]Department of Chemistry, Princeton University, Princeton, NJ, USA. [6]German Cancer Research Center (DKFZ), Heidelberg, Division of Metabolic Crosstalk in Cancer and the German Cancer Consortium (DKTK), DKFZ Core Center Heidelberg, Heidelberg, Germany. [7]Department of Medicine, Section of Infectious Diseases & Global Health, University of Chicago Medicine, Chicago, IL, USA. [8]Division of Infectious Diseases, Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA. [9]Department of Medicine, Section of Hematology and Oncology, University of Chicago, Chicago, IL, USA. ✉e-mail: joshr@princeton.edu

https://doi.org/10.1038/s42255-026-01550-8



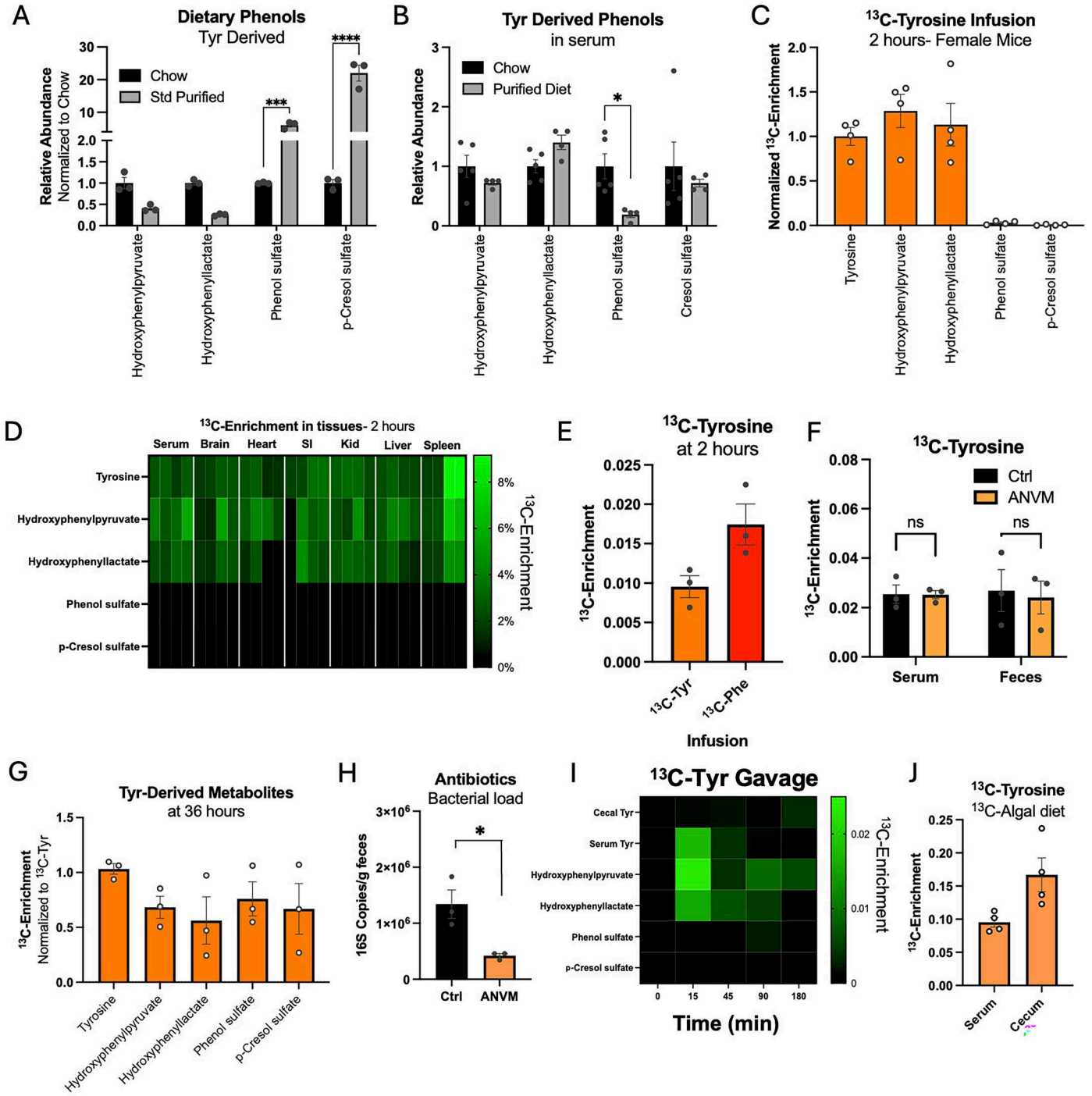
Extended Data Fig. 1 | Assessing host contribution to indole metabolites. (**a**) Indole metabolites in PicoLab Chow (n = 3) vs a standard purified diet (n = 3). (**b**) Indole metabolites in mice fed PicoLab Chow (n = 5) vs a standard purified diet (n = 4). (**c-d**) ^{13}C-enrichment of indole metabolites in serum (**c**) and tissues (**d**) following a 2-hour infusion of ^{13}C-tryptophan in female mice (n = 3). (e) ^{13}C-enrichment of tryptophan in serum and feces after a 36-hour infusion in antibiotic treated or control male mice (both n = 3 serum and feces). (**f**) ^{13}C-enrichment of indole metabolites in serum at the end of a 36-hour infusion of ^{13}C-tryptophan, which allows for microbially derived metabolites to be enriched. Enrichment was normalized to serum ^{13}C-tryptophan enrichment and data is from control male mice (n = 3). This data is used as the comparison to antibiotics-treated (ANVM) animals as seen in Fig. 1f. (**g**) Alpha diversity as calculated by Shannon index (n = 3) and (**h**) fecal bacterial load for the control and ANVM treated mice (both n = 3). (**i**) Enrichment of tryptophan and indole metabolites in portal vein serum following a gavage with ^{13}C-trytpophan (n = 3-4 per time point). (**j**) Cecal tryptophan enrichment following 24 hours of feeding with ^{13}C-algal protein (n = 4). All bar graphs represent mean ± SEM. ns= not significant, $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$ by a two-tailed *t*-test.

https://doi.org/10.1038/s42255-026-01550-8



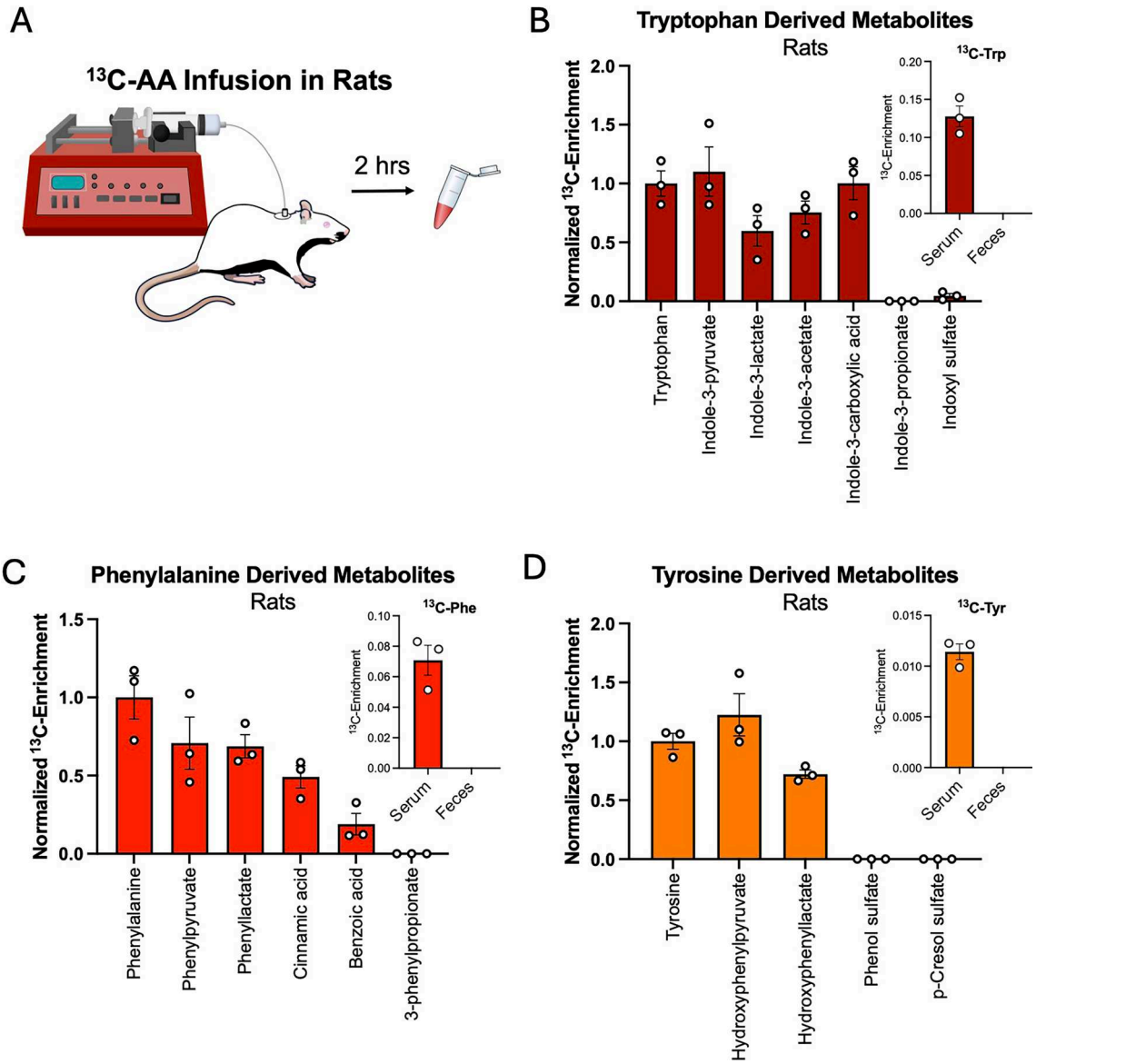
Extended Data Fig. 2 | Assessing host contribution to phenylalanine-derived metabolites. (**a**) Phenylalanine-derived metabolites in PicoLab Chow (n = 3) vs a standard purified diet (n = 3). (**b**) Phenylalanine-derived metabolites in mice fed PicoLab Chow (n = 5) vs a standard purified diet (n = 4). (**c-d**) ^{13}C-enrichment of phenol metabolites in serum (**c**) and tissues (**d**) following a 2-hour infusion of ^{13}C-phenylalanine in female mice (n = 3). (**e**) ^{13}C-enrichment of phenylalanine in serum and feces after a 36-hour infusion in antibiotic treated (n = 3) or control male mice (n = 3). (**f**) ^{13}C-enrichment of phenol metabolites in serum at the end of a 36-hour infusion of ^{13}C-phenylalanine, which allows for microbially derived metabolites to be enriched. Enrichment was normalized to serum ^{13}C-phenylalanine enrichment and data is from control male mice (n = 3). This data is used as the comparison to ANVM treated animals as seen in Fig. 2f. (**g**) Alpha diversity as calculated by Shannon index (n = 3) and (**h**) fecal bacterial load for the control and ANVM treated mice (both n = 3). (**i**) Enrichment of phenylalanine and phenol metabolites in portal vein serum following a gavage with ^{13}C-phenylalanine (n = 3). (**j**) Cecal phenylalanine enrichment following 24 hours of feeding with ^{13}C-algal protein (n = 4). All bar graphs represent mean ± SEM. ns= not significant, $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$ as calculated by a two-tailed *t*-test.

https://doi.org/10.1038/s42255-026-01550-8



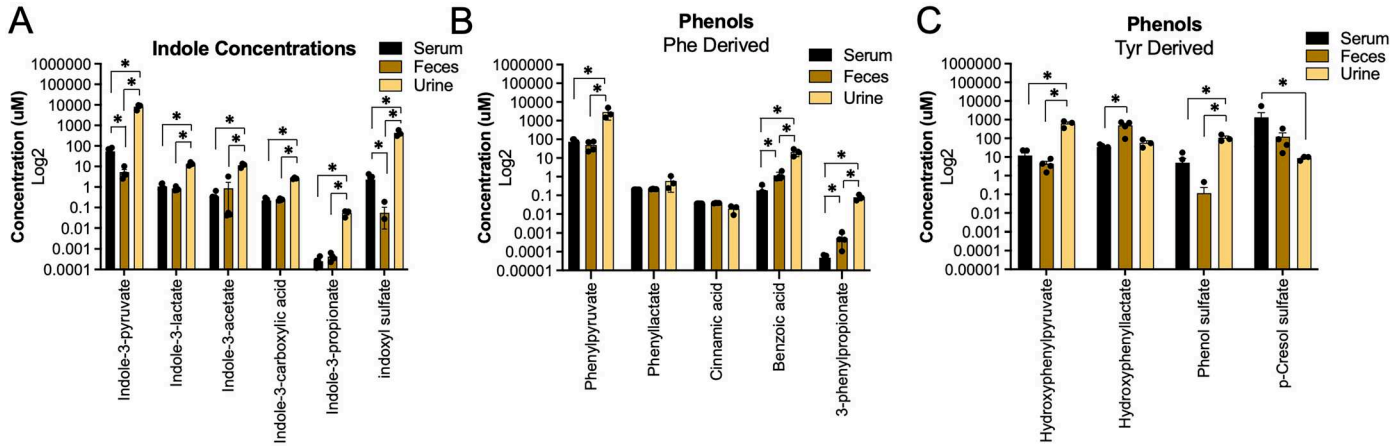
Extended Data Fig. 3 | Assessing host contribution to tyrosine-derived metabolites. (**a**) Tyrosine-derived metabolites in in PicoLab Chow (n = 3) vs a standard purified diet (n = 3). (**b**) Tyrosine-derived metabolites in mice fed PicoLab Chow (n = 5) vs a standard purified diet (n = 4). (**c-d**) ^{13}C-enrichment of phenol metabolites in serum (**c**) and tissues (**d**) following a 2-hour infusion of ^{13}C-phenylalanine in female mice (n = 3). (**e**) ^{13}C-enrichment of tyrosine in serum of male mice infused either ^{13}C-tyrosine (n = 3) or ^{13}C-phenylalanine (n = 3) at 2 hours. (**f**) ^{13}C-enrichment of tyrosine in serum and feces after a 36-hour infusion in antibiotic treated (n = 3) or control male mice (n = 3). (**g**) ^{13}C-enrichment of phenols in serum at the end of a 36-hour infusion of ^{13}C-tyrosine in control male mice (n = 3). Enrichment was normalized to serum ^{13}C-tyrosine enrichment. This data is used as the comparison to antibiotics-treated (ANVM) animals as seen in Fig. 3f. (**h**) Fecal bacterial load for the control and ANVM treated mice (n = 3). (**i**) Enrichment of tyrosine and phenol metabolites in portal vein serum following a gavage with ^{13}C-tyrosine (n = 3-4 per time point). (**j**) Cecal tyrosine enrichment following 24 hours of feeding with ^{13}C-algal protein (n = 4). All bar graphs represent mean ± SEM. ns= not significant, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, and $^{****}p < 0.0001$ as calculated by a two-tailed *t*-test.

https://doi.org/10.1038/s42255-026-01550-8



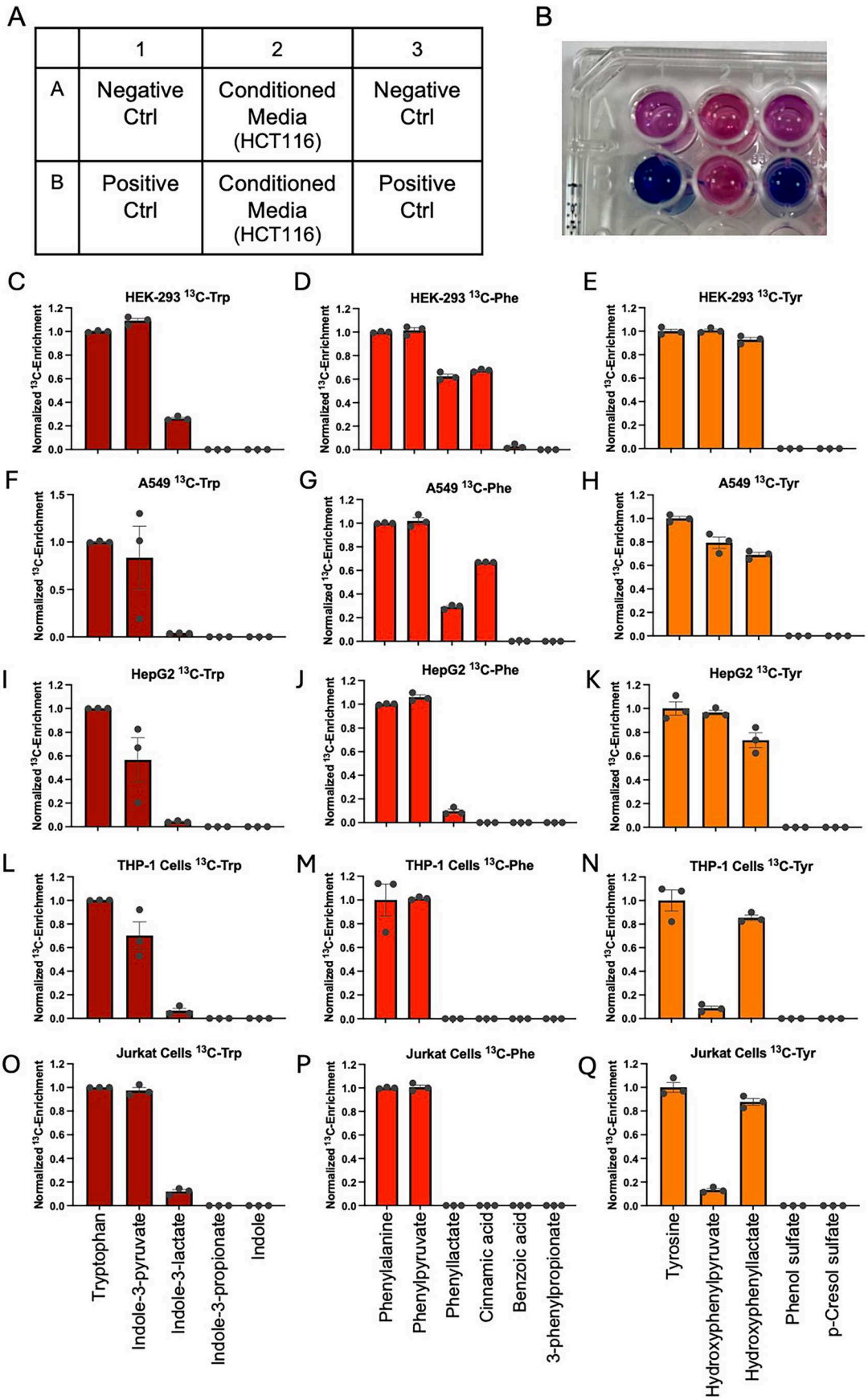
Extended Data Fig. 4 | Rats also produce many indole and phenol metabolites. (**a**) Scheme describing the infusion of ^{13}C-amino acids for two hours in rats. (**b**) ^{13}C-enrichment of indole metabolites in serum normalized to serum ^{13}C-tryptophan enrichment (n = 3 male rats). Inset showing ^{13}C-tyrptophan in serum and feces at 2 hours. (**c**) ^{13}C-enrichment of phenylalanine-derived metabolites in serum normalized to serum ^{13}C-phenylalanine enrichment (n = 3 male rats). Inset showing ^{13}C-phenylalanine in serum and feces at 2 hours. (**d**) ^{13}C-enrichment of tyrosine-derived metabolites in serum normalized to serum ^{13}C-tyrosine enrichment. Inset showing ^{13}C-tyrosine in serum and feces at 2 hours (n = 3 male rats). All bar graphs represent mean ± SEM.

https://doi.org/10.1038/s42255-026-01550-8



Extended Data Fig. 5 | Absolute serum, fecal, and urine concentrations of indole and phenol metabolites. (**a**) Indole metabolite concentrations in serum, feces, and urine of male mice fed a control purified diet. (**b**) Same for phenylalanine-derived and (**c**) Tyrosine-derived phenols. For all metabolites, n = 5 serum samples, n = 4 fecal samples, n = 3 urine samples. Statistical analysis was completed by two-way ANOVA with $^{*}p < 0.01$.

https://doi.org/10.1038/s42255-026-01550-8

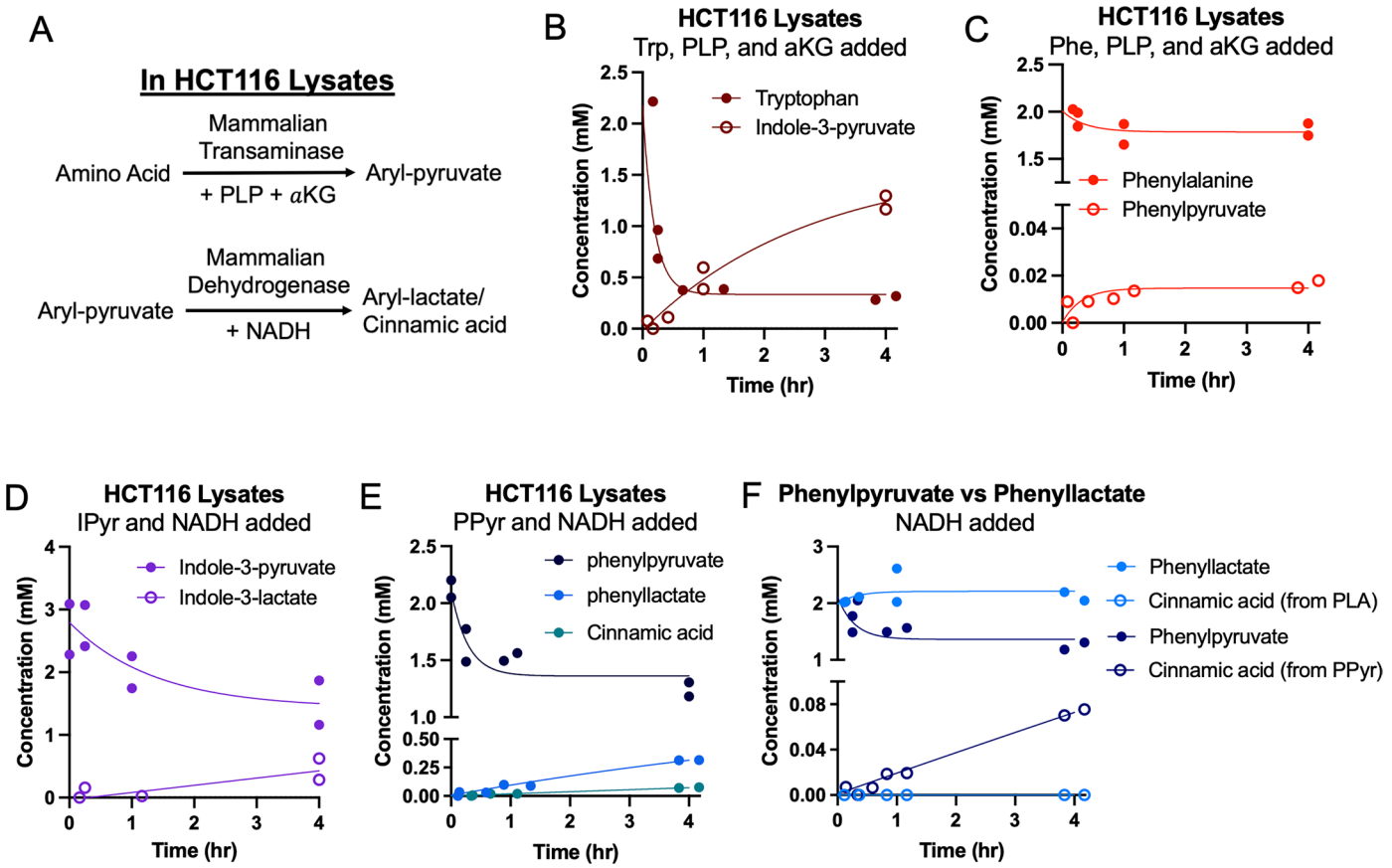


Extended Data Fig. 6 | See next page for caption.

https://doi.org/10.1038/s42255-026-01550-8

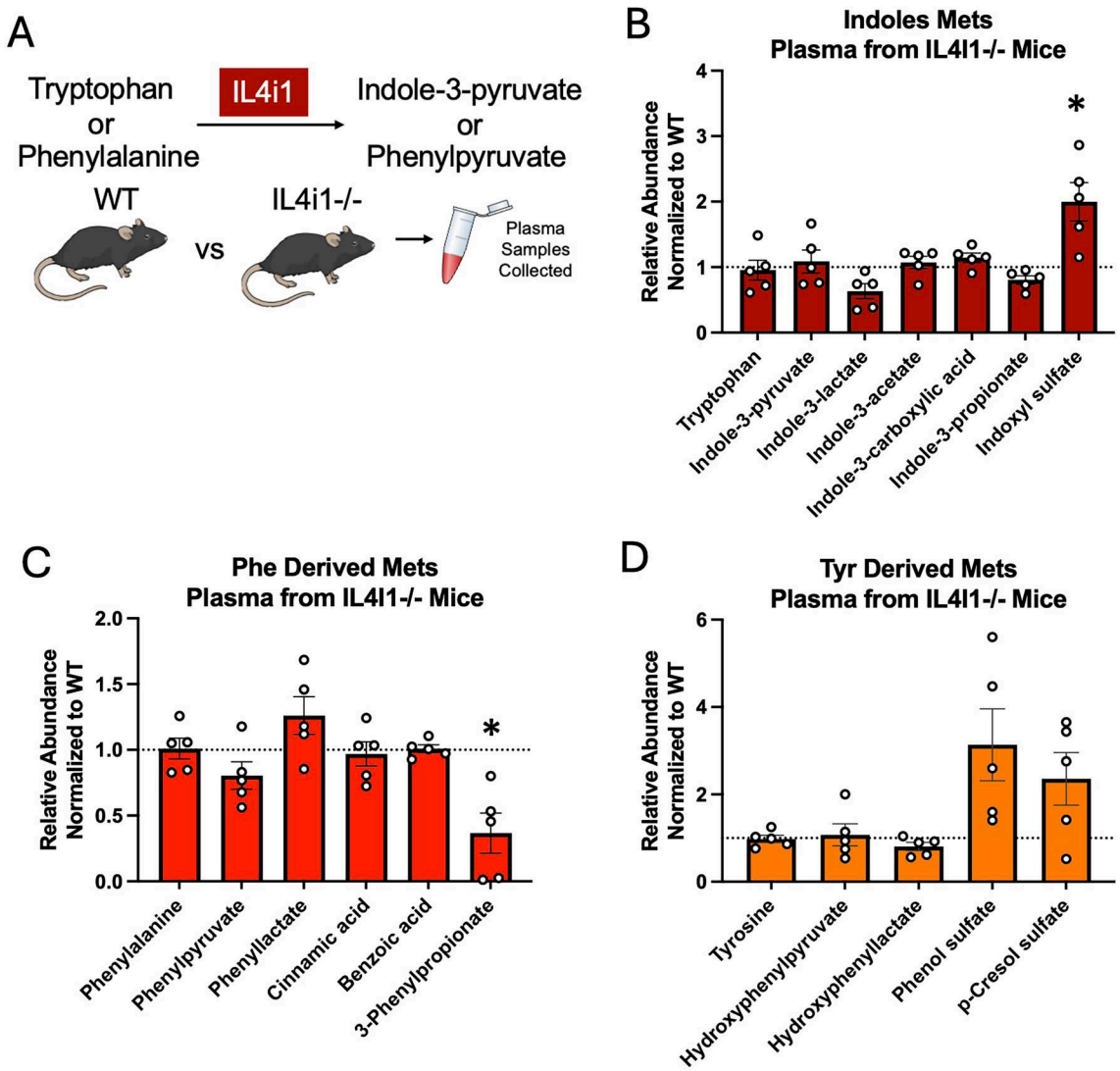
Extended Data Fig. 6 | Synthesis of indoles and phenols is conserved in many different cell lines. (**a**) Plate layout for the PlasmoTest *Mycoplasma* Detection Kit. (**b**) Results showing no *Mycoplasma* contamination of the HCT116 cells. (**c-q**) Enrichment of indole, phenylalanine-derived, and tyrosine-derived metabolites in cells treated with ^{13}C-tryptophan, ^{13}C-phenylalanine, or ^{13}C-tyrosine for 20 hours in (**c-e**) A549 lung cancer cells, (**f-h**) HepG2 liver cancer cells, (**i-k**) HEK-293 immortalized human embryonic kidney cells, (**l-n**) THP-1 leukemia cells, and (**o-q**) Jurkat T cell lymphoma cells. All metabolite enrichment was normalized to their respective ^{13}C-amino acid enrichment. All measurements are represented as the mean ± SEM with an n = 3 biological replicates.

https://doi.org/10.1038/s42255-026-01550-8



Extended Data Fig. 7 | Aryl-amino acid transaminase and dehydrogenase activities are present in mammalian cell lysates. (**a**) Scheme describing the reactions observed and cofactors needed in HCT116 lysates. (**b**-**c**) The conversion of tryptophan to indole-3-pyruvate (**b**) and phenylalanine to phenylpyruvate (**c**) occurs with the addition of pyridoxal phosphate (PLP) and alpha-ketoglutarate (*a*KG). No reaction occurred in lysates without the addition of PLP and *a*KG. (**d**-**e**) The production of indole-3-lactate (**d**) from the addition of indole-3-pyruvate (IPyr) and phenyllactate, and cinnamic acid (E) from phenylpyruvate (PPyr) with NADH. No reaction occurred without the addition of NADH. (**f**) Cinnamic acid production from phenylpyruvate or phenyllactate and NADH in the HCT116 lysates. NADH was required for conversion. All measurements were performed with an n = 2 independent lysate reactions. Mean ± SEM is graphed for all data points.

https://doi.org/10.1038/s42255-026-01550-8



Extended Data Fig. 8 | IL4I1 is not the main contributor to indole or phenol circulating metabolites. (**a**) Schematic showing the metabolic reactions of IL4I1 and the comparison of plasma from IL4I1−/− vs wild type (WT) mice. (**b-d**)Plasma metabolite abundances in whole body IL4I1−/− mice (n = 3 female and 2 male mice) compared to WT mice (n = 3 female and 2 male mice). Mean ± SEM is graphed for all data points. *$p < 0.05$ as calculated by a two-tailed *t*-test.

nature portfolio

Corresponding author(s): Joshua D. Rabinowitz

Last updated by author(s): Mar 2, 2026

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study.
For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

LC-MS for all rodent and in vitro data: Orbitrap Exploris 240 or Q Exactive Plus hybrid quadrupole-orbitrap mass spectrometer (Thermo Fisher Scientific).
LC-MS indoles measurement for human data: Agilent 1290 infinity II liquid chromatography system coupled to an Agilent 6470 triple quadrupole mass spectrometer
LC-MS phenols measurement for human data: Thermo Fisher liquid chromatography system coupled to an Orbitrap IQ-X mass spectrometer
Protein fractionation: ÄKTA pure chromatography system
Proteomics: Bruker TimsTOF Pro 2 connected to a nanoElute LC
For Shannon Diversity measures: Oxford Nanopore MinION™ sequencer.
Human study microbiome composition: Illumina HiSeq platform

Data analysis	For analysis of microbiome composition for the AML patients: Human genome reads were removed by kneaddata (https://github.com/bFiiobakery/kneaddata, v 0.7.10.) taxonomy of the resultant high-quality reads was profiled with metaplan448. MEGAHITS (v 1.2.9) was used to assemble microbial reads with genes called by prodigal (https://github.com/hyattpd/Prodigal, v. 2.6.3). For Shannon Diversity: Species level taxonomic annotation was assigned to reads with Emu (v. 3.4.5) and Phyloseq45 was used for shannon index. All LC-MS data for rodent studies was analyzed using El-Maven (v.0.12.1). All 13C-labeling data was corrected for natural 13C abundance using the Accucor package (https://github.com/XiaoyangSu/AccuCor). LC-MS data analysis for human samples: MassHunter Quant software (version B.10, Agilent Technologies) Proteomics analysis: Perseus (v 2.0.7.0) Figures were created using GraphPad Prism (v.10.2.3)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data generated or analyzed during this study are included in this article and its Supplementary Information files. The datasets supporting the findings are available from the corresponding author upon reasonable request.

All MS proteomics data have been deposited to the MassIVE database (accession: MSV000100980; ftp://MSV000100980@massive.ucsd.edu) including an in silico spectral library generated in DIA-NN via FASTA of Uniprot mouse proteome UP000000589.

Microbiome sequencing and metabolomic data for the AML patients has been deposited on Zendo (https://doi.org/10.5281/zenodo.18829320).

Mass spectrometry data from the Boston Children's study is available from MetaboLights dataset MTBLS948.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Biological sex or gender were not considered when designing or consenting participants for this study. Biological sex is reported in the demographics table (Extended Data table 1).
Reporting on race, ethnicity, or other socially relevant groupings	Race and ethnicity were not considered for consenting participants for this study. Patients with a diagnosis of acute myeloid lymphoma (AML) receiving treatment at the University of Chicago Medicine were enrolled without regard for social grouping. Data on race and ethnicity are reported in the demographics table (Extended Data table 1).
Population characteristics	Sex, age, race or ethnicity, and clinical course are all reported in Extended Data table 1.
Recruitment	Human subjects were recruited to this study based on pre-specified criteria: having new diagnosis AML, admitted to the hospital, greater than 17 years of age, not pregnant, and having provided informed consent. Patients did not self-identify for inclusion. All eligible patients were approached for enrollment during the study period. Patient characteristics could be influenced by the population served by University of Chicago, and are described in Extended Table 1
Ethics oversight	The University of Chicago Institutional Review Board approved this study (IRB23-1365).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were done to predetermine sample sizes. For all rodent experiments, we aimed for 3-5 animals. For cell culture, we also completed triplicates. For in vitro analysis, we completed independent duplicates for each measurement.

Data exclusions	**Human studies: For the study completed at the University of Chicago, participants were excluded if they did not have serum and fecal samples** taken within 24 hours of each other. **Rodent studies: For infusion studies, samples from rodents were excluded if the infusion was stopped prematurely, either due to a technical** issue, or the mouse or rat not surviving for the duration of the infusion.
Replication	**Rodent studies: The number of mice used for each study is indicated in the figure legends in the manuscript. While optimizing the conditions** for the long term infusions, the short term infusions were replicated multiple times, with all attempts successfully replicating the representative data presented in the figures. **For in vitro work: Each assay was repeated at least in duplicate or greater and all attempts for replication were successful.**
Randomization	**Mice were grouped randomly between control and treatment groups.**
Blinding	**Human studies: The study including patients with AML was observational so blinding was not required, but microbiome sequencing and** metabolomics analysis were completed by core facility staff who were not aware of the clinical data (antibiotic exposure). **Rodent studies: There was no formal blinding completed, but sample preparation and work up were done in a blinded manner.**

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ **Eukaryotic cell lines**
☒	☐ **Palaeontology and archaeology**
☐	☒ Animals and other organisms
☐	☒ Clinical data
☒	☐ **Dual use research of concern**
☒	☐ **Plants**

Methods

n/a	Involved in the study
☒	☐ **ChIP-seq**
☒	☐ **Flow** cytometry
☒	☐ **MRI-based neuroimaging**

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	**All cell lines were obtained from ATCC.**
Authentication	**All cell lines were either recently purchased from ATCC or authenticated with STR fingerprinting.**
Mycoplasma contamination	**Cells were routinely tested for Mycoplasma contamination using the PlasmoTest Mycoplasma Detection Kit (Invivogen).**
Commonly misidentified lines (See ICLAC register)	**No commonly misidentified lines were used in this study.**

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	**Mouse models included C57BL/6NCrl male and female mice from Charles River (Charles River Laboratories, strain 027) and C57BL/6N wild type female germ-free or murine pathogen free mice (Taconic strain B6-F GF or MPF). Rat models include Sprague Dawley jugular vein catheterized male rats (Charles River Laboratories, strain 001).**
Wild animals	**No wild animals were used for this study.**
Reporting on sex	**Both females and males were included for both mouse studies. Sex of mice was detailed in the manuscript and the data is represented in a disaggregated fashion with the exception of IL4i1 -/- mice as there was not enough sample size (n=2 male for each** group) to disaggregate the sexes.
Field-collected samples	**No field-collected samples were used for this study.**

Ethics oversight	**All mouse and rat studies were completed in compliance with protocols approved by the Princeton University Animal Care and Use Committee.**

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	**There was not a clinical trial completed for this study.**
Study protocol	**The clinical data from the University of Chicago was completed under the protocol IRB23-1365.**
Data collection	**Patients with a diagnosis of acute myeloid lymphoma (AML) receiving treatment at the University of Chicago Medicine. Serial collection of serum and stool was performed while hospitalized. All serum/fecal sample pairs were taken within 24 hours of each other.**
Outcomes	**Microbiome alpha diversity and indole and phenol metabolite concentrations were the outcomes for the clinical study completed at the University of Chicago Medicine.**

Plants

Seed stocks	**No plants were used for this study**
Novel plant genotypes	**No plants were used for this study**
Authentication	**No plants were used for this study**

This checklist template is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/