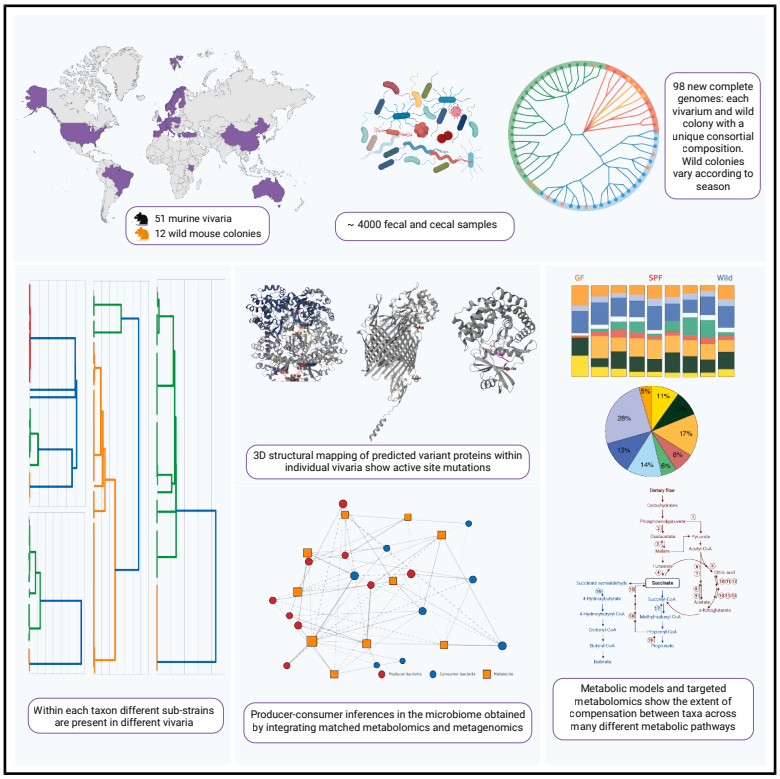
Resource

Cell Host & Microbe

A global survey of taxa-metabolic associations across mouse microbiome communities

Graphical abstract



Highlights

- Highly variable murine microbiota compositions occur in both lab and wild colonies
- Functional outputs converge despite taxonomic and habitat differences
- Within each taxon, there are multiple variants, usually unique to that vivarium
- Some species variants map to active sites in enzymes linked to core metabolic pathways

Authors

Bahtiyar Yilmaz, Isabel Baertschi, Karin H.U. Meier, ..., Uwe Sauer, Jörg Stelling, Andrew J. Macpherson

Correspondence

bahtiyar.yilmaz@unibe.ch (B.Y.), andrew.macpherson@unibe.ch (A.J.M.)

In brief

Yilmaz et al. provide a global atlas of murine microbiota and their species variants across many vivaria and wild colonies. By integrating strain-resolved metagenomics, metabolomics, and structural annotation, they show that conserved metabolic outputs occur independently of taxonomic diversity, linking species variants to host-relevant metabolic phenotypes through a searchable resource.

Yilmaz et al., 2025, Cell Host & Microbe *33*, 1960–1976
November 12, 2025
https://doi.org/10.1016/j.chom.2025.10.010

CellPress

CellPress
OPEN ACCESS

Cell Host & Microbe

Resource

A global survey of taxa-metabolic associations across mouse microbiome communities

Bahtiyar Yilmaz,[1,2,*] Isabel Baertschi,[1,2] Karin H.U. Meier,[3] Constance Le Gac,[4,5] Sebastian B.U. Jordi,[1,2] Caitlin Black,[6] Jiaqi Li,[1,2] Anna K. Lindholm,[6] International Mouse Microbiota Investigators, Barbara König,[6] Uwe Sauer,[3,7] Jörg Stelling,[4,5,7] and Andrew J. Macpherson[1,2,7,8,*]

[1]Department of Visceral Surgery and Medicine, Bern University Hospital, University of Bern, 3010 Bern, Switzerland
[2]Maurice Müller Laboratories, Department for Biomedical Research, University of Bern, 3008 Bern, Switzerland
[3]Institute of Molecular Systems Biology, Swiss Federal Institute of Technology (ETH) Zurich, 8093 Zurich, Switzerland
[4]Department of Biosystems Science and Engineering, ETH Zurich, 4056, Basel, Switzerland
[5]SIB Swiss Institute of Bioinformatics, 1015, Lausanne, Switzerland
[6]Department of Evolutionary Biology and Environmental Studies, University of Zurich, 8057, Zurich, Switzerland
[7]Senior author
[8]Lead contact
*Correspondence: bahtiyar.yilmaz@unibe.ch (B.Y.), andrew.macpherson@unibe.ch (A.J.M.)
https://doi.org/10.1016/j.chom.2025.10.010

SUMMARY

Host-microbiota mutualism is rooted in the exchange of dietary and metabolic molecules. Microbial diversity broadens the metabolite pool, with each taxon contributing distinct compounds in varying proportions. In the human microbiome, high variability in consortial composition is largely compensated by similar metabolic functions across different taxa. However, the extent of compensation in lower diversity mouse models, and whether vivaria are metabolically equivalent, is unknown. We provide a searchable resource of microbiome composition variability across 51 murine vivaria and 12 wild mouse colonies worldwide, with vivarium-specific variants mapped according to predicted 3D structures for each microbial species. Our matched metabolomics data show that realized metabolic potential has relatively low variability, providing functional evidence for metabolic compensation. Additionally, variability is related to taxonomic composition rather than vivarium, revealing taxa-metabolite associations that are potentially relevant to phenotypic differences between vivaria. Collectively, this resource offers tools to strengthen microbiome studies and collaborative science.

INTRODUCTION

The mouse is a key model for host-microbial studies since its genetic background, microbiota composition, environmental conditions, and diet can be experimentally manipulated to address the underlying mechanisms of mutualism.[1,2] The profound impact of the microbiota on host cells and metabolism has been shown by comparing germ-free and colonized animal models of the same inbred strain.[2–5] The effects of different microbiotas on animal phenotypes, and some correlations with individual metabolites, are well accepted. However, the extent to which the realized metabolite diaspora depends on the microbiota consortial composition remains unclear. Recent studies have highlighted the translational limitations of conventional specific pathogen-free (SPF) mouse models and advocated for the use of natural microbiota-based systems to better recapitulate host physiology and immune complexity.[6,7]

Although most members of the microbiota are well contained in the intestinal lumen or on other body surfaces,[5] serum and urine metabolomes of colonized animals contain many metabolites derived from microbial metabolism.[8] These metabolites include those released by effete microbial cells or the products of microbial metabolism of foods or other ingested xenobiotics, such as drugs. The molecules reaching central body tissues, therefore, depend on (1) the overall metabolic capacity of the resident microbiota; (2) the extent to which missing biochemical pathways in one taxon can be complemented in another; and (3) that the pathway concerned is functional within the microbial consortium and its metabolites are generated.

There are many different microbial consortia containing taxa with interacting metabolism in nature.[9] The mammalian intestine contains a relatively open system with highly nested taxa networks, which are very heterogeneous in composition between hosts, with different diets or under varying environmental conditions.[10] Despite this heterogeneity, landmark human microbiome sequencing studies showed that there is a considerable functional overlap between different possible microbiota members, so that the aggregate of pathways encoded by the overall consortium broadly showed comparatively little variation despite the variations in taxa composition.[11–13] Measurements

of intestinal luminal metabolites using capsule systems in 15 healthy volunteers, with pH-dependent deployment at different positions along the intestine, have shown some dietary and microbiota correlations. There were strong interparticipant differences and clear differences between intestinal and stool metabolomes.[14] Therefore, genomic compensation within the microbiota is unlikely to fully compensate for the realized human intraluminal metabolite portfolio.

Vivarium mice exist under more standardized dietary and environmental conditions than humans. If complementation of missing metabolic pathways in one taxon is actually realized by compensation with others in different vivarium mouse microbiotas, differences in the resulting metabolite diaspora reaching host tissues may be negligible, even with different taxa compositions.[13] Provided such complementation is sufficient, different (non-pathogenic) microbiota compositions may have similar effects on experimental *in vivo* mouse phenotypes. However, the well-known phenotypic variations (including immune function, behavior, susceptibility to infection, gastrointestinal physiology, and epigenetic regulation) observed among genetically identical (isogenic) mice housed in different vivaria or following microbiota rewilding challenge this assumption.[15] There are also known circumstances—such as on certain body sites[11] or for specialist bile salt metabolic pathways[16]—where encoded pathways vary across time or between experimental animals. It is therefore important to interpret experiments not only in terms of microbiota taxonomic composition but also in terms of the extent of metagenomic pathway variation, genomic variants in key metabolic transport proteins and enzymes, and the consequent metabolomic output in the mouse microbiotas under study.

We have analyzed 51 different vivaria and 12 wild mouse colonies to find how far the overall metagenome-encoded metabolic pathways of the microbiomes vary in relation to the differences in microbiota composition and how far this affects the diversity of the resultant intraluminal intestinal metabolome that potentially penetrates the host. Within this metagenomic variability, we also report the extent of synonymous and non-synonymous variation within individual taxa from different vivaria and wild mouse colonies and identify non-synonymous variation at active sites and quaternary interfaces of key metabolic enzymes and transport proteins of individual taxa. We present a range of metabolites that vary between mouse colonies according to microbiota composition and diet. We also provide a searchable resource of the relationships between taxonomic, metagenomic, gene variant, and realized metabolomic variations across a wide range of experimental and natural contexts to guide researchers in the interpretation of mouse experimental results.

RESULTS

Mapping the global diversity of mouse gut microbiota across laboratory and wild populations

As the starting point for addressing the depth of diversity and its functional consequences across experimental mouse vivaria, we established a resource for the heterogeneity of mouse microbiota compositions by analyzing a global reach of cecal and fecal samples of 51 laboratory colonies from 48 different institutions. Given the importance of extending the available model traits and phenotypes in laboratory mice that have been rewilded with free-living microbiotas, we also included samples from 12 different colonies of wild mice (Figure S1; Table S1). As expected, genus-level analysis using 16S gene or species-level analysis via metagenomic sequencing revealed that each vivarium had a distinct microbiota composition (Figures 1 and S2). While different papers have reported either a higher or lower proportion of Bacillota/Bacteroidota (formerly Firmicutes/Bacteroidetes (F/B) ratio) as the main phyla in wild mice compared with those in vivaria, likely depending on different wild mouse colonies studied and the different vendors of lab mice,[17,18] we found that within each vivarium and over the different wild mice sampled, there was a very wide range of the ratios between these phyla (Figure 1A). The microbiotas of vivarium mice were generally different from the microbiota composition of wild mice, with selected vivaria having Verrucomicrobiota, Mycoplasmatota (formerly Tenericutes), and Actinomycetota (Actinobacteria) that were only present in very low proportions in all the wild mouse colonies studied (Figure 1). The picture emerging from this broad geographical dataset confirms the immense taxonomic heterogeneity of the different microbiotas being used in experiments worldwide. This detailed dataset can be accessed through two interactive Shiny apps, where all taxonomic and metabolic results (https://yilmazlab.shinyapps.io/vivaria_app/) and the structural variants in gut microbial strains (https://yilmazlab.shinyapps.io/vivaria_app2/) reported in this paper can be searched and explored according to vivarium or wild colony.

Genomic diversity defines mouse gut microbiota across vivaria and wild populations

Despite extensive sequencing efforts, both humans and mice have high proportions of undefined taxa and sub-strain variants of sequenced taxa. The breadth of our sampling from worldwide vivaria and wild mouse colonies allowed us to ask how incomplete our sequence information is, even from SPF vivaria routinely being used to model health and disease. We prioritized samples from 16 vivaria and 12 different locations of wild mice across four countries for metagenomic sequencing (Figure 2). Metagenome-assembled genomes (MAGs) using three different binning tools (i.e., MetaBAT,[19] MaxBin,[20] and CONCOCT[21]) were integrated to eliminate duplicates and enhance the accuracy of the genome assemblies.[22] A total of 20,345 MAGs from this pipeline exceeded medium quality (≥50% completeness and <10% contamination) based on the minimum information about a MAG (MIMAG) standard.[23] Species-level phylogenies and classified species-level genome bins (SGBs) were then reconstructed from these MAGs using the *PhyloPhlAn* pipeline.[24] SGB genomes were divided into distinct subgroups to show their respective completeness and contamination percentages (Figures 2A–2C): those with completeness ranging from 50% to 90% were designated as medium quality; 90%–99.9% as high quality, and 100% as completed. Low or high contamination levels were 0.001%–5% or greater than 5%, respectively. The near-complete genomes (Figure 2C) displayed lower strain heterogeneity compared with the medium-quality genomes (Table S2). Overall, 49% of bins were of high quality, and 282 were fully closed, of which 98 were without contamination (Figure 2C; Table S2). Additional genome statistics (including contig number and N50, average mash distance) supported

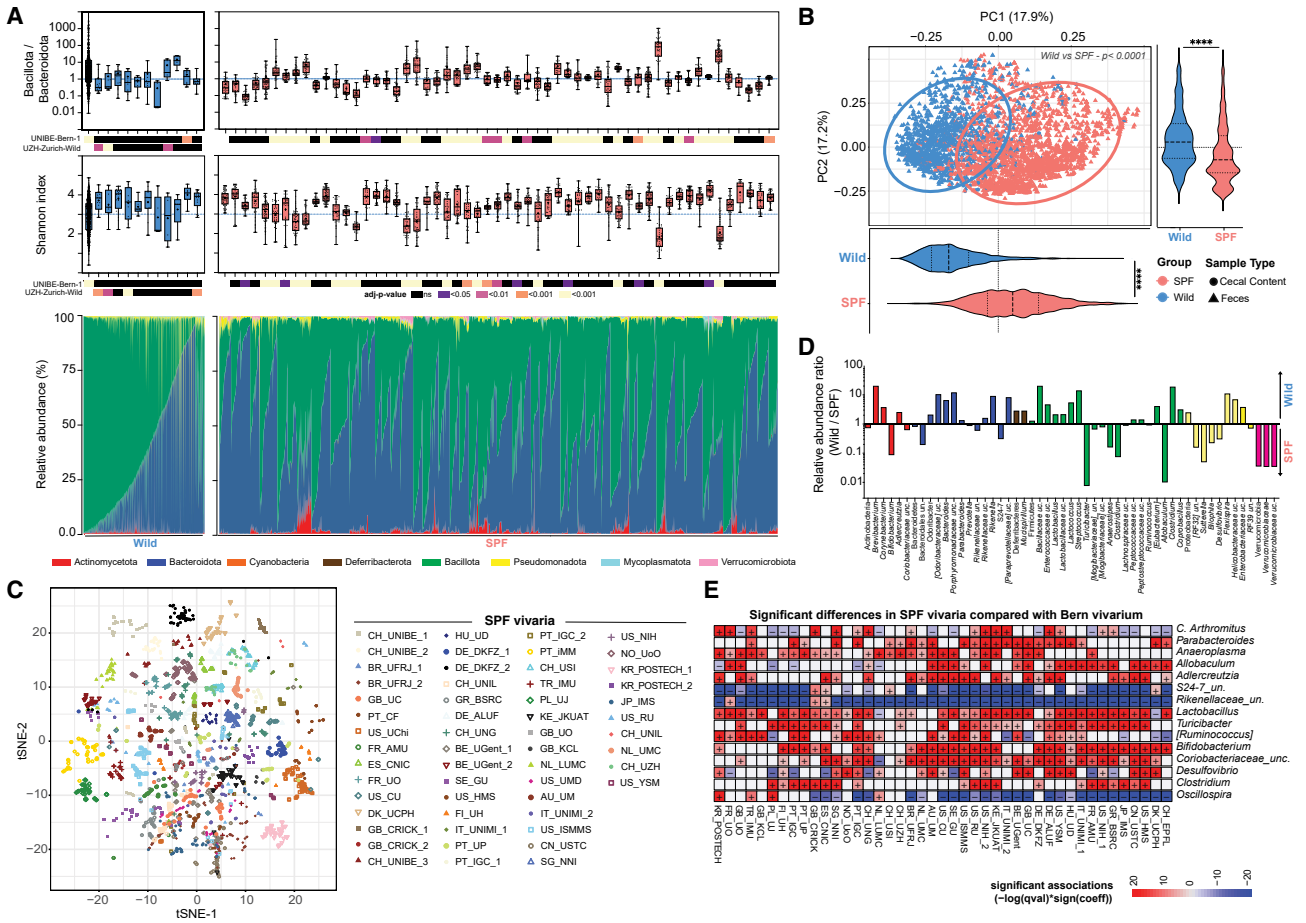


Figure 1. Overall gut microbiota profile in SPF and wild mice around the world

(A) Bacillota:Bacteroidota ratio (upper), Shannon index (middle), and relative abundance (%) (lower) calculated from 16S rRNA gene sequencing dataset are shown for wild mice colonies and different SPF mice vivaria. Each box plot represents one sample location. Each vivaria or wild mice colony was compared with CH_UNIBE or UZH_Zurich_Wild. Heatmap shows adj-*p* values.

(B) PCoA of Bray-Curtis distances depicting microbial community differences between wild and SPF mice. Ellipsoids show 95% confidence intervals. Group differences were tested by Adonis on the PCoA plot ($p < 0.0001$) and by the Mann-Whitney U test for PC1 and PC2 ranks (±SD), with ****$p < 0.0001$.

(C) Projection of all microbiota samples from different vivaria (origin) onto two-dimensional space using t-distributed stochastic neighbor embedding (t-SNE).

(D) Significant bacterial abundance differences (adj-$p < 0.05$) between all wild mice colonies and all SPF mice vivaria samples are shown.

(E) Significant taxa differences (adj-$p < 0.05$) in microbiome composition across SPF vivaria. General linear mixed models were run relative to the University of Bern colony (Bern 1).

(D and E) Differences between groups were determined using MaAsLin2.

See also https://yilmazlab.shinyapps.io/vivaria_app/.

the consistent high quality of near-complete MAGs compared with medium-quality MAGs (Table S2).

As expected, human samples had a relatively higher proportion of genomes that could be assigned to known SGBs (kSGBs) compared with either vivarium or wild mice, where most were assigned to previously unidentified SGBs (uSGBs) (Figures 2D and 2E). In two locations of wild mouse colonies, more than 90% of binned taxa were unassigned (Figure 2D), and there was clear heterogeneity of k- and uSGB composition in SPF vivaria (Figure 2E). The previously uSGBs in SPF mice were mainly within Bacteroidota, whereas wild mice also had uSGBs within Pseudomonadota (formerly Proteobacteria) and Bacillota (Figure 2E). Despite the overlap between the composition of vivarium and natural wild mice microbiota, many of the known and previously uSGBs in this analysis were shared either between different colonies of wild mice or between different colonies of SPF vivarium mice, and thus largely differentiated wild and SPF mice in unsupervised clustering (Figure 2E). Therefore, even in mouse facilities with strictly controlled experimental conditions, there is still a high proportion of assigned but previously unknown taxa, some of which are only found in individual vivaria.

Since incomplete taxonomy and annotation are well-known issues in human microbiotas, we compared our mouse data with human samples to assess potential overlap (Figures S2A–S2C). This analysis showed that both known and previously unknown SGBs obtained from human samples were almost completely distinct (>99%) from the mouse samples

CellPress
OPEN ACCESS

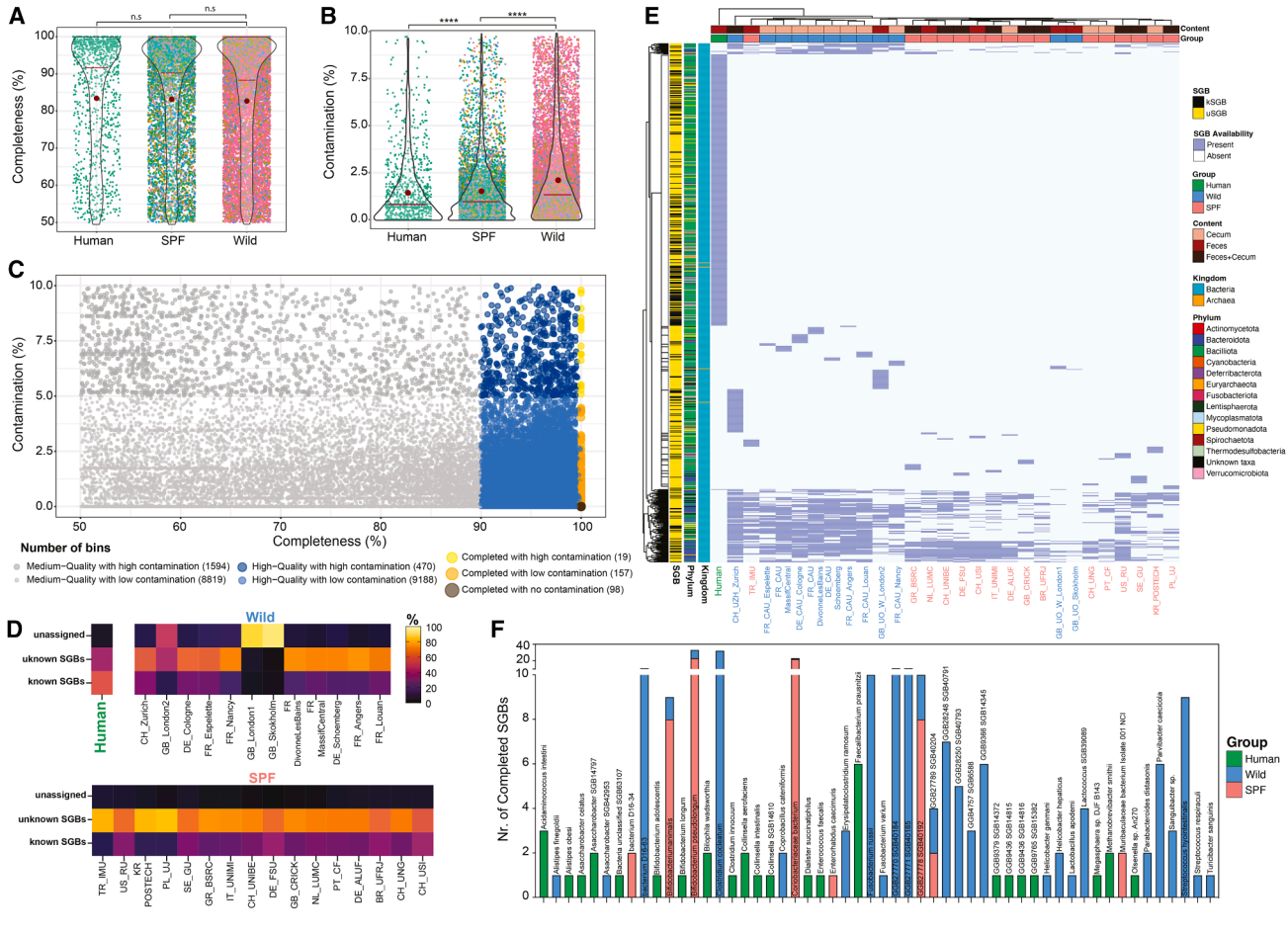


Figure 2. Metagenomic assemblies of strains from worldwide experimental mouse vivaria and wild colonies

(A and B) MAGs were generated from 16 SPF vivaria and 12 wild mouse colonies across four countries, with >50% completeness (A) and <10% contamination (B).

(C) MAG distributions are shown, with completeness (*x* axis) and contamination (*y* axis); each dot represents a MAG, colored by colony (colors random, not informative).

(D) SGBs of MAGs using the *PhyloPhlAn* pipeline showing proportions of known (kSGB) taxa, unknown (uSGB) and unassigned SGBs in humans and different colonies of wild mice (upper) and different SPF vivaria (lower).

(E) Sub-strain variants were shown according to kSGB or uSGB stratified by human, wild mouse (Zürich-NAA) or SPF (Greece-Jagiellonian) mouse source.

(F) The origin and identity of newly completed SGBs without contamination. Bar chart displaying the origin of samples as human (green), wild mice (blue), and SPF mice (red), with corresponding identities shown.

(Figures 2E, 2F, and S2D; Table S2), arguing against accidental contamination of clean mouse facilities from human personnel.

Since mouse rewilding has so far only been carried out on a small number of starting microbiota, there likely remains considerable potential to increase the breadth of traits and phenotypes by using different diverse wild consortia (Figures S3A and S3B). We confirmed seasonal variations of the wild mouse microbiota[25] (Figures S3C–S3E). These included significant summer decreases in the Bacillota:Bacteroidota phylum ratios. Most taxa driving these cyclical shifts between summer and non-summer months were found within 5 specific bacterial orders, including substantial depletions in the *Lactobacillales* and *Clostridiales* orders of the Bacillota phylum as well as significant blooms in *Bacteroidales* and *Campylobacterales* orders of Bacteroidota. The facultative anaerobic *Enterobacteriales* order also showed significant summer expansions. These findings indicate that the gut microbiome of wild mice is not only immensely heterogeneous, depending on the location of the colony sampled, but also undergoes substantial seasonal remodeling, with reductions in obligate anaerobes of the Bacillota, the aerotolerant opportunistic Pseudomonadota phylum, and increases in the Bacteroidota during the warmer, more humid summer (Figure S3E). This compositional plasticity is likely driven by consumer resources (microbe: carbon source) and seasonal dietary differences.[18,26–30]

Strain-level genetic divergence in distinct microbiota across humans and mice

Genetic polymorphism will exist in each microbial constituent species of the microbiota. We aimed to develop a resource that demonstrates these effects across different microbial species and mouse colonies and their potential to affect the combined metabolic output of the microbial consortium concerned. We had previously shown genetic evolution through drift and

positive selection, with the emergence of sub-strains within individual bacterial species over time, in isobiotic mice created from inoculating C57BL/6 germ-free recipients with 12 different taxa, each starting from a single clone, and maintained with strictly aseptic protocols for over 6 years.[31] Taking a population average nucleotide identity (popANI) > 99.9999% as indicative of strain identity, we carried out variant analysis using the *inStrain* pipeline[32] of co-occurring genome populations to ask how far sub-strains (1) result in different polymorphic individual microbial species within each colony or between different colonies and (2) how genetic variants within individual species compare between experimental or wild mice.

The taxonomic overlap and genetic heterogeneity of 200 different bacterial species were analyzed from 115 wild mice and 115 lab mice. Out of these, 48 taxa between different individuals and/or mice showed genetic similarities of over 95% marking a species boundary,[33,34] with a coverage overlap ranging from 70% to 99.9%. As a control, we included healthy human samples in the analysis, which showed that very few taxa were shared between humans and mice or were uniquely present in either laboratory or wild mice (Figure 3A). Highly identical strains within a species (defined as 99.99999% popANI, where, in the case of two alleles, only the predominant variant is considered) were normally derived just from a single vivarium (Figures 3B–3E) or within a single colony of wild mice (Figures 3D and 3E). Almost no identical strains within the same bacterial taxon were found when comparing between laboratory and wild mice or between humans and either wild or laboratory mice.

Twelve different taxa were commonly shared between different wild or between different SPF mice colonies (average popANI ranging between 97% and 99.5%), but strong genetic identities (popANI > 99.999%) were only found for taxa within each individual vivarium or specific wild mouse location (Figure 3E). Within the vivarium, microdiversity (popANI > 97%) was widely seen within species of the Bacteroidota and Bacillota phyla, as well as the *Enterobacteriaceae* family. However, there were only a few examples where closely related strains within a species (ANI > 97%) were found between different vivaria (*Akkermansia muciniphila*, *Turicimonas muris*, and *Muribaculum* sp002492595; Figure 3C) or between different wild colonies (*Helicobacter cinaedi* and *Muribaculum* sp001701195; Figure 3C). A representative phylogenetic tree further highlights the genetic similarity of a strain shared between wild and SPF mice (Figure 3D). As a comparison, microdiversity (97%–99.5%) between different hosts was verified within different human fecal samples[35] (Figure 3F) although only *B. uniformis* showed genetic resemblance between human and SPF mouse samples, and only six human strains had popANI overlaps >97% with wild mice colonies (Figure 3G). These results show that the genetic identity of individual taxa within a microbiota is highly restricted to each individual experimental mouse vivarium or specific wild mouse community.

Adaptive metabolic pathway evolution shapes functional divergence in diverse microbiota

Since we had found that microbial strain variants are largely specific to individual vivaria, and we have previously reported evidence for evolution in microbiota taxa in an interbreeding mouse colony across time,[31] we questioned whether individual taxa in different vivaria had developed genetic variability likely to be relevant for metabolic compensation in enzymes or transport proteins independently of the levels of that taxon in the consortium. Since an excess of non-synonymous variants can provide indirect evidence for positive selection at that locus,[31] we reasoned that we would be able to detect non-synonymous variants predicted to be important for the function or protein structure of enzymes at committal points for microbial metabolism. We specifically focused on microbial species displaying microdiversity between vivaria, characterized by popANI values ranging from 97% to 99%. Following gene annotation within individual microbial species, we performed a systematic screening to identify variants exhibiting a dN/dS ratio greater than one across at least three distinct vivaria. We then took four examples (Figure 4A) for detailed structural analysis in combination with *Alphafold3 3D* reconstruction (Figures 4B–4J; https://yilmazlab.shinyapps.io/vivaria_app2/). First, by examining TonB-dependent receptor P3,[36] we observed that non-synonymous variants were structurally constrained, predominantly occurring in the predicted external polysaccharide utilization domain. Synonymous variants were also concentrated in this domain, likely through translational constraints,[37] with only one variant located in the transmembrane beta-sheets forming the protein pore (Figures 4I and 4J). Secondly, non-synonymous variants also occur at the predicted active site of the committal step of threonine synthesis in homoserine kinase[38] (Figures 4D and 4F) and at the initial step of *de novo* purine synthesis in amidophosphoribosyltransferase[39,40] (Figures 4G and 4H). Non-synonymous variants were also clustered at the predicted subunit interfaces of the homoserine kinase (Figures 4D and 4E) and GTP cyclohydrolase 1 type 2[41,42] (Figures 4B and 4C) quaternary protein structure. Many of these non-synonymous variants were vivarium specific and absent from the same genes in other vivaria, even allowing for the limitations of sequence depth obtained. Whereas it is possible that variants shared between different experimental vivaria could be the result of transfer of animals and their microbiotas between breeders and different experimental institutions, we also found some overlaps in non-synonymous variants between vivarium and wild mice supporting their independent derivation through positive selection (e.g., amidophosphoribosyltransferase—the committal step in *de novo* purine synthesis; and GTP cyclohydrolase—a key enzyme in the folate and biopterin biosynthesis pathways) (Figures 4B and 4C). Although these variants have not yet been tested biochemically, a sufficient depth of a forward genetic approach can potentially contribute to an understanding of the function of per-taxon-specific metabolic pathways within individual microbiota consortia.

Metabolic flexibility and resource-driven specialization differentiate wild and SPF mouse microbiota

Given the immense diversity of species composition and genomic variants in the microbiome among individual animals, we next assessed whether different metabolic functions within the microbiota theoretically combine in different combinations to achieve metabolic homeostasis. To examine the metabolic potential of the different microbial communities, we first analyzed cecal and fecal metagenomic shotgun sequencing data from different vivaria and wild mouse colonies (Figures S1B and S2) using the *HUMAnN3* pipeline for microbial

CellPress
OPEN ACCESS

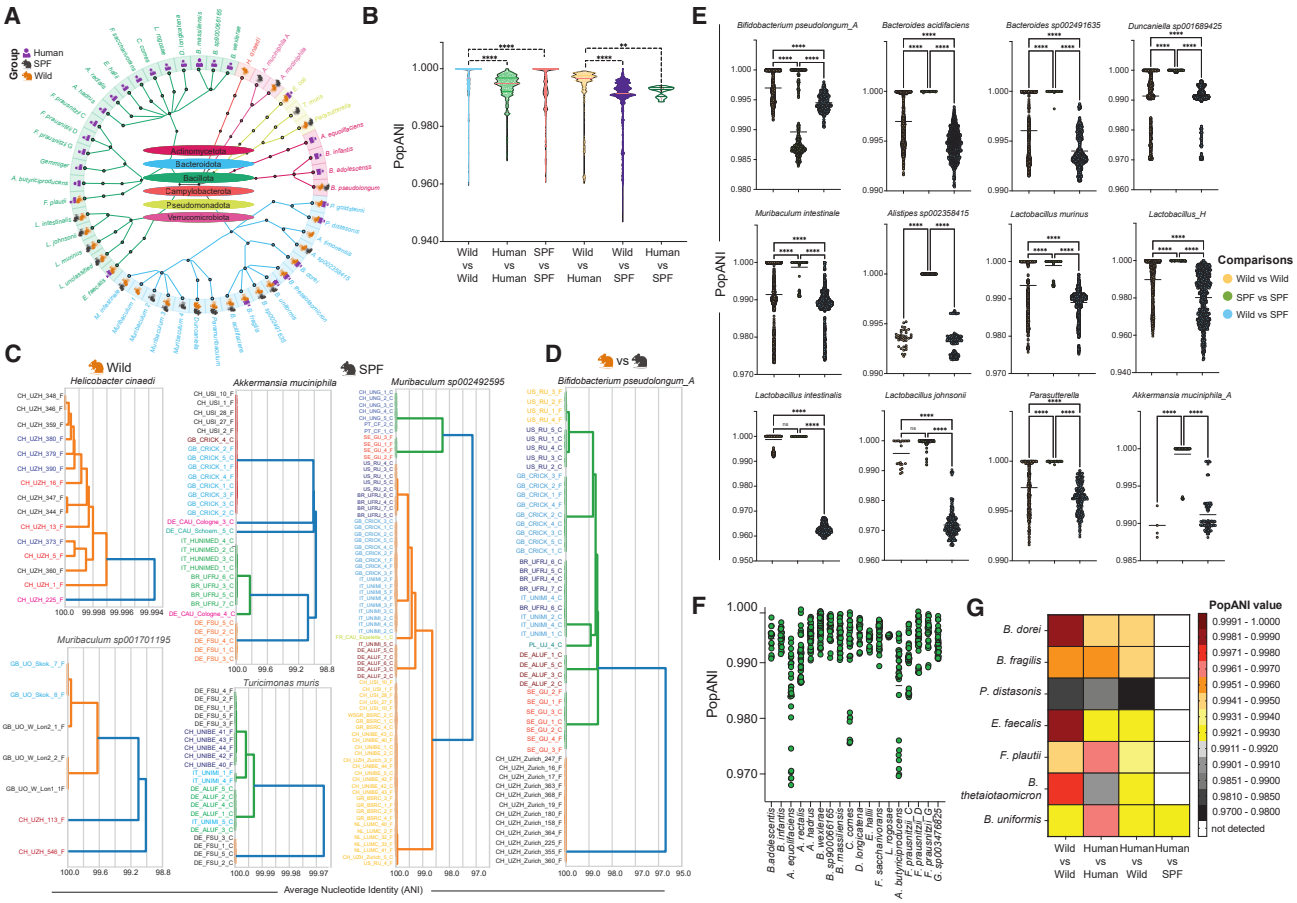


Figure 3. Comparative analysis of genetic homogeneity in gut microbiota across mouse and human hosts

(A) Microbial strains with identified genetic similarities between humans, wild mice, and SPF mice are shown on the tree. The color code used on the tree indicates the phylum level of the identified taxa.

(B) The overall popANI per group is shown on the violin plot. The data are reported at both interspecies and intraspecies levels.

(C) Genetically congruent microbial taxa observed solely within colonies of wild mice or SPF vivaria are delineated.

(D) Representative depiction of a phylogenetic tree illustrating genetic similarities in a particular strain between wild and SPF mice.

(E) Dot plots of popANI values, color coded by comparisons, show the comparative analysis of bacterial species between colonies of wild mice, SPF mice, and between wild and SPF mice.

(F) Dot plots illustrate the genetic differences among strains that have been found only in humans.

(G) Heatmap shows strain-level genetic similarities between mouse and human samples. Significance: * adj-$p < 0.05$, **adj-$p < 0.01$, ***adj-$p < 0.001$, ****adj-$p < 0.0001$; n.s., not significant.

metabolic pathways.[43] This showed increased pathway richness in both fecal and cecal samples from wild mice compared with SPF mice, as well as in control humans compared with mice, albeit with overlapping ranges (Figure 5A). Principal coordinates analysis (PCoA) based on the Bray-Curtis distance confirmed that 83 annotated metabolic pathways present in SPF mice overlapped with those in wild mice and annotated pathways in fecal human samples clustered adjacent to the wide beta-diversity spread from cecal and fecal samples in wild mice (Figure 5B). These results confirmed that despite the huge diversity of the microbiota, cumulative metabolic potentials overlap.[11,12,16]

Statistical analysis with the MaAsLin[44] revealed broadly shared pathway representation among wild mice, SPF mice, and humans, with no higher-level clusters uniquely distinguishing any of the groups (Figure 5C). Cecal samples from 6 of 18 SPF vivaria clustered within a group that was shared mainly with control healthy human fecal samples, although the microbial communities from humans specifically contained the pathway of glycogen degradation (Figure 5C host cluster 2, pathway PA233). Clusters mainly containing wild mice (Figure 5C, host clusters 1 and 3) featured genes annotated for increased representation of *de novo* nucleoside synthesis pathways and synthesis of other oligosaccharides, likely due to greater carbon source limitation in wild mice compared with *ad libitum* feeding of humans and vivarium mice (Figure 5C; Table S3). The requirement for *de novo* nucleoside synthesis by the microbiota depends on nucleobase bioavailability for energy-efficient salvage pathways derived from cell turnover or dietary sources. Although we were unable to assess the diets or food preferences of the sampled wild mice, plant material is a poor dietary source of nucleotides, and we confirmed the presence of dietary nucleotides in the vivarium diets (Figure S4; Table S4).

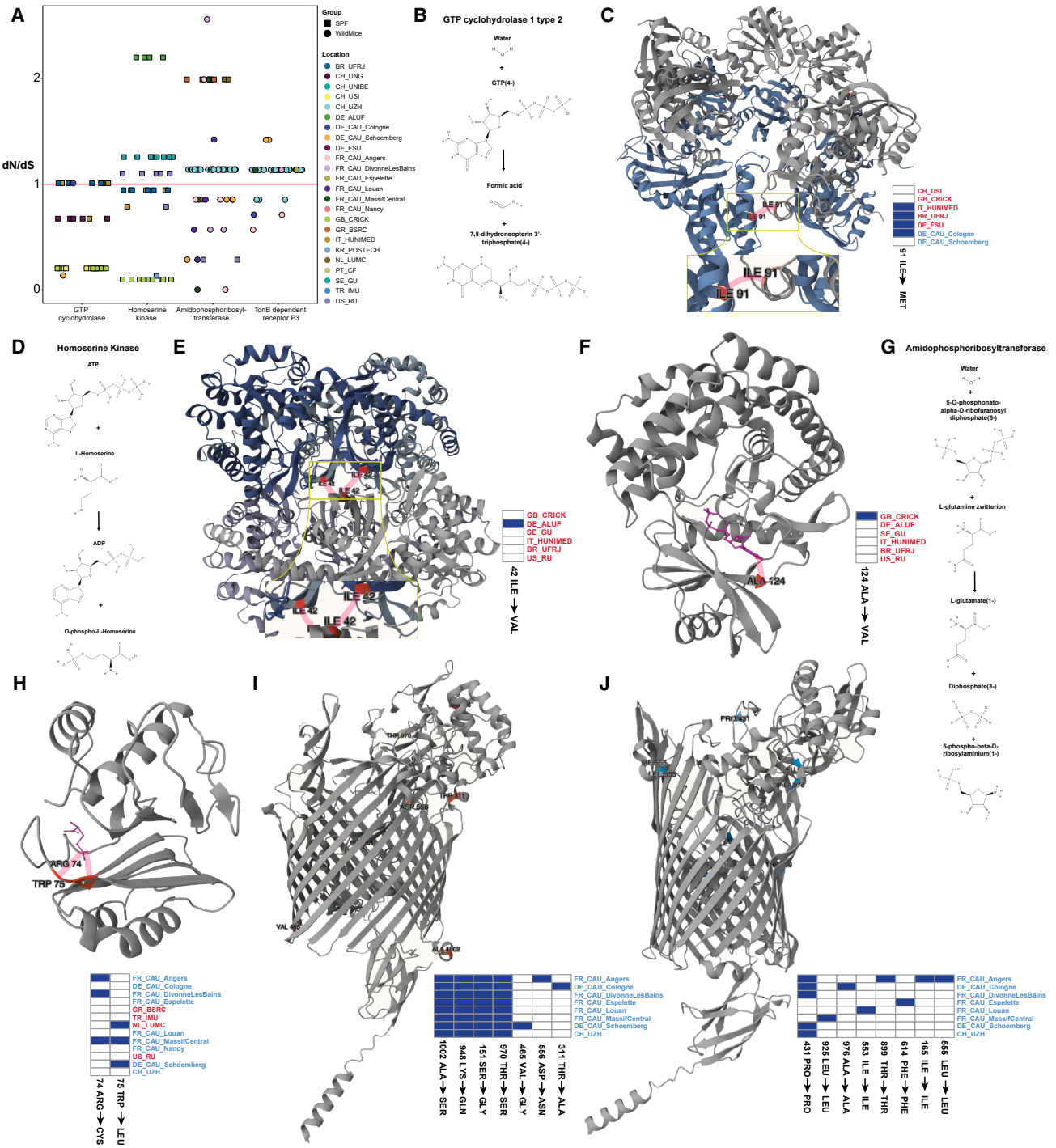


Figure 4. Structural and evolutionary insights into vivarium-specific microbial strain variants reveal evidence of positive selection in selected key functional genes

(A) Distribution of dN/dS ratios across different vivaria of wild and SPF mice for four key genes. Each dot represents an individual mouse, highlighting the variations within and between vivaria. The four genes are of different microbial origin: GTP cyclohydrolase (GTPCH) of *A. muciniphila* A, homoserine kinase (HSK) of *B. pseudolongum*, amidophosphoribosyltransferase (ATase) of *Bacteroides sp002491635*, and TonB-dependent receptor P3 of *B. acidifaciens*.

(B) Schematic representation of the reaction catalyzed by GTPCH, a key enzyme in the folate and biopterin biosynthesis pathways.

(C) Toroid structure of GTPCH generated with AlphaFold3. The non-synonymous substitutions sites potentially involved in multimerization, according to the protein-ligand interaction profiler (PLIP), are highlighted in red. Vivaria samples presenting these substitutions are indicated in dark blue in the accompanying heatmap. Vivaria written in red represent SPF mice and vivaria written in blue wild mice.

(D) Schematic illustration of the reaction catalyzed by HSK, the pacemaker enzyme in the threonine biosynthesis pathway.

(legend continued on next page)

Therefore, although the microbiome composition taxonomy was distinct for each vivarium (Figures 1C–1E) and the wild colony (Figure S3B), unsupervised analysis of the variability in the metabolic potential of the annotated gene content does not cluster according to vivarium or for wild vs. laboratory mice (Figure 5C). Most features of the inferred cumulative metabolic potential did not even segregate across the host groups. This reflects substantial functional compensation among microbial species across environments, enabling diverse taxonomic assemblages of consumers to exploit resource availability under specific dietary conditions.

Although we had prioritized intestinal samples from 18 vivaria for metagenomic sequencing, we also had 16S gene sequencing data from 51 vivaria, 12 colonies of wild mice, and 107 human samples (see also https://yilmazlab.shinyapps.io/vivaria_app). We therefore tested whether we could find evidence for the different metabolic potentials of these three groups in the wider sample collection. Some samples were excluded from this analysis due to missing annotations necessary for model fitting (see STAR Methods). The analyzed samples covered 21 vivaria, 1 wild mice colony, and 105 human samples. We mapped amplicon sequence variants (ASVs) to metabolic reactions via a previously published collection of genome-scale metabolic models (GSMMs)[45] and used this mapping to quantify reaction abundance in each sample (Figures S5E–S5I). This mapping yielded 7,604 unique metabolic reactions, of which 3,146 could be mapped to one of 134 different metabolic subsystems in Kyoto Encyclopedia of Genes and Genomes (KEGG)[46–48] and one ''unknown'' subsystem for reactions that could not be mapped to an identified subsystem. Differences in reaction abundance between all vivaria (Figures S5E–S5I), wild colonies, and human samples were then related to groups of interest by fitting reaction-specific linear mixed-effect models. We estimated the mean differences in reaction abundance between the groups of interest and tested for significance using the Holm-Bonferroni multiple testing correction, (adj-$p \leq 0.05$). For each group of interest, the enrichment of significant reactions in each metabolic subsystem was then tested using Fisher's exact test ($p \leq 0.05$). Reactions mapped to the unknown subsystem were considered in the enrichment test, but the unknown subsystem itself was not tested. As expected, there were clear differences when comparing pups that were still being nursed with milk (containing lactose and milk glycoproteins) to adult mice. Those differences were visible both at the microbial level (Figures S5A–S5D) and the metabolic level (Figure 5D). Subsystems such as galactose metabolism, glycolysis, peptidoglycan biosynthesis, and glycerophospholipid metabolism were enriched in reactions with significantly higher mean abundance in pups than in adults (Figure 5D). However, consistent with the metagenomic analysis of cumulative metabolic potential across different microbiotas, there were significant differences between the metabolism of wild and SPF mice, notably in higher tryptophan metabolism, indole alkaloid biosynthesis, and glucosinolate biosynthesis in wild mice (Figure S5G). The mean reaction abundance of a few metabolic pathways was comparatively higher in control humans than in mice, including pathways involved in valine, leucine, and isoleucine biosynthesis, lysine biosynthesis, fatty acid biosynthesis, and elongation (Figures S5E–S5I). Notwithstanding general metabolic compensation between different mouse microbiota, differences in subsystem representation occur at different life stages, likely due to consumer-resource effects of milk vs. a weaned diet, and potentially also between wild and vivarium mice.

Compensatory stability and metabolite sensitivity reveal functional adaptation in diverse microbiota

Generally, the interpretation of functional metabolic compensation of the microbiota from metagenomic analysis or metabolic reaction mapping of ASVs is subject to limitations because these methods only provide an estimate of the metabolic potential, which can result in different metabolomes, depending on the realized fluxes. Rather than relying on indirect estimates of metabolic potential, the crucial information for experimentalists is to know the *realized* portfolio of metabolites, where changes in the microbiota consortium composition potentially cause significant deviations in their intraluminal concentrations. To assess the realized portfolio, we therefore determined steady-state metabolome compositions in the luminal contents of the cecum across vivaria (62 adult SPF mice, 10–18 weeks old), wild mice (9), and germ-free mice (5). Targeted analysis of 112 typical intestinal metabolites showed similar representation across all broad compound classes in the cecal contents of animals in the different vivaria except for germ-free mice, which, as expected,[49] had reduced metabolite diversity (Figures 6A, 6B, S6A, and S6B; Table S5). Ordination analysis of this data revealed low ($\leq$12%) variance in the principal components, with only germ-free mice differentiated significantly (Figure 6C). Mice did not segregate by vivarium in an unsupervised analysis, either (Figure 6D; Table S5).

To assess how microbiome composition as well as diet relate to the realized portfolio of metabolites, we focused on 47 mice (42 SPF, 5 wild mice, from 10 vivaria) with matched metagenomics, luminal metabolomics, and (vivaria-specific) diet metabolomics

(E) Tetrameric structure of HSK modeled with AlphaFold3. Predicted multimerization residues (predicted with PLIP) are shown in red. The heatmap highlights non-synonymous substitutions in vivaria samples (dark blue), with SPF vivaria in red and wild vivaria in blue.
(F) HSK monomer bound to ATP (pink). ATP-binding residues (predicted by COACH, C-score = 0.19) carrying non-synonymous substitutions in at least one vivarium are marked in red. The heatmap shows these substitutions (dark blue), with SPF vivaria in red and wild vivaria in blue.
(G) Schematic representation of the reaction catalyzed by ATase. This reaction is the committing step in *de novo* purine synthesis.
(H) Monomeric structure of ATase with bound glutamic acid, colored in pink. Two ligand binding residues (predicted by COACH, C-score = 1) displaying non-synonymous substitutions are highlighted in red. The accompanying heatmap shows in dark blue which vivaria present the respective substitution. Vivaria in red are SPF mice and vivaria in blue are wild mice.
(I) Distribution of non-synonymous substitutions in the TonB-dependent receptor P3. The heatmap (dark blue) indicates vivaria with substitutions; red = SPF, blue = wild.
(J) Structure of the TonB-dependent receptor P3 with synonymous mutations in light blue. The heatmap indicates their presence across vivaria (dark blue); red indicates SPF, while blue indicates wild.
See also https://yilmazlab.shinyapps.io/vivaria_app2.

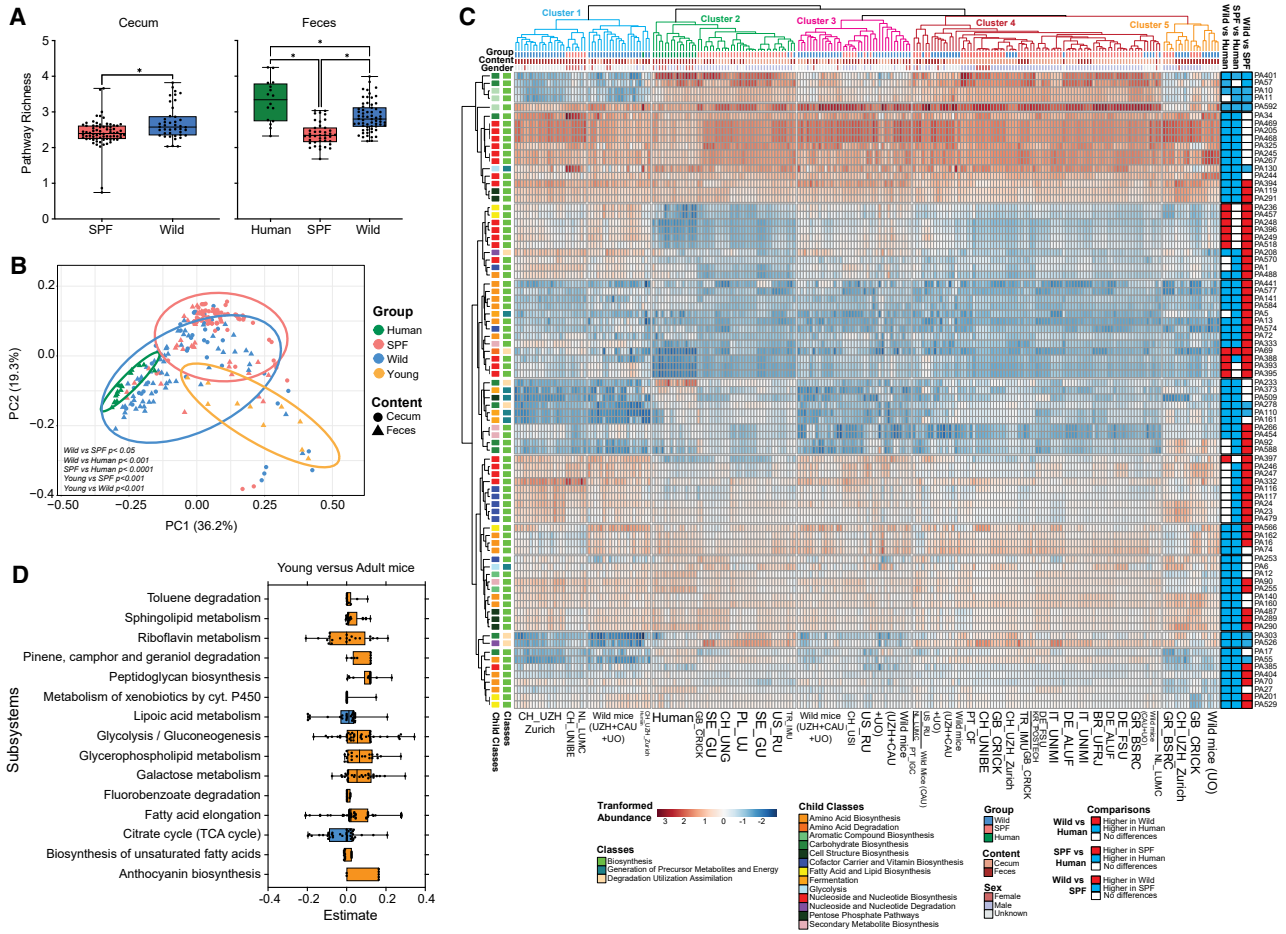


Figure 5. The metabolic potential of microbial communities in mice and humans

(A) Box-and-whisker plots identified pathway (analyzed using the HUMAnN3 pipeline) richness differences among wild mice, SPF mice and human samples (quartiles, ranges, and standard deviations). *$p < 0.05$.

(B) PCoA based on Bray-Curtis distance showing inferred microbial metabolic pathways between wild mice, SPF mice and human samples. The positions of each group are represented by ellipsoids (95% confidence intervals). Non-parametric Mann-Whitney U test and Adonis test were used for statistical differences between groups of alpha diversity (A) or beta diversity (B), respectively.

(C) Relative abundance of microbial metabolic pathways that were not present in at least 30% of our samples exhibiting significant associations with different groups were identified using MaAsLin2. The heatmap displays inferred microbial pathways varying significantly in relative abundance between sample types (adj-$p < 0.05$). Pathway relative abundance is indicated on a color scale from low (blue) to high (red).

(D) Estimated significant differences in mean reaction abundance imputable to mouse age differences (adult/pup) for significantly enriched subsystems ($p \leq 0.05$) based on GSMM/mixed-effect model analysis. Each point represents an estimate, only significant estimates (adj-$p \leq 0.05$) are shown. Positive values show subsystem enrichment in pups.

See also https://yilmazlab.shinyapps.io/vivaria_app/.

data. We first analyzed the variability in the three data types individually (Figure S6C). Variability of luminal metabolite concentrations across mice, as characterized by the coefficient of variation (CV) for individual metabolites, had a median CV of 179% (interquartile range [IQR] [138%, 258%]), which is substantially higher than the typical technical variability of ~10%. Because the variability of diet metabolites across vivaria was of the same magnitude (median CV: 130%, IQR [96%, 186%]), we tested for significant differences in variability between luminal and diet metabolite abundances across vivaria (one-sided F tests, α = 0.05).[49] The results in Figure 6E do not support systematically higher variability of luminal metabolites, indicating functional compensation also in the realized metabolic portfolio.

We then used linear mixed-effects models to quantify relations between taxonomic (71 taxa identified at the species level), diet (124 metabolites), and metabolome composition. Specifically, we selected per-metabolite models that best explained luminal metabolite variability in terms of diet and species composition (see STAR Methods for details). Models including only diet predictors explained significantly less variability than those with species predictors, and they did not lead to optimal per-metabolite models, which all showed high explanatory power ($R^2 \geq 0.9$; Figure S6D). A few of the significant predictors were related to diet, and these had small estimated effects (Figure S6E). Vivaria or microbiome models were not identified as significant (Table S6), confirming that

CellPress
OPEN ACCESS

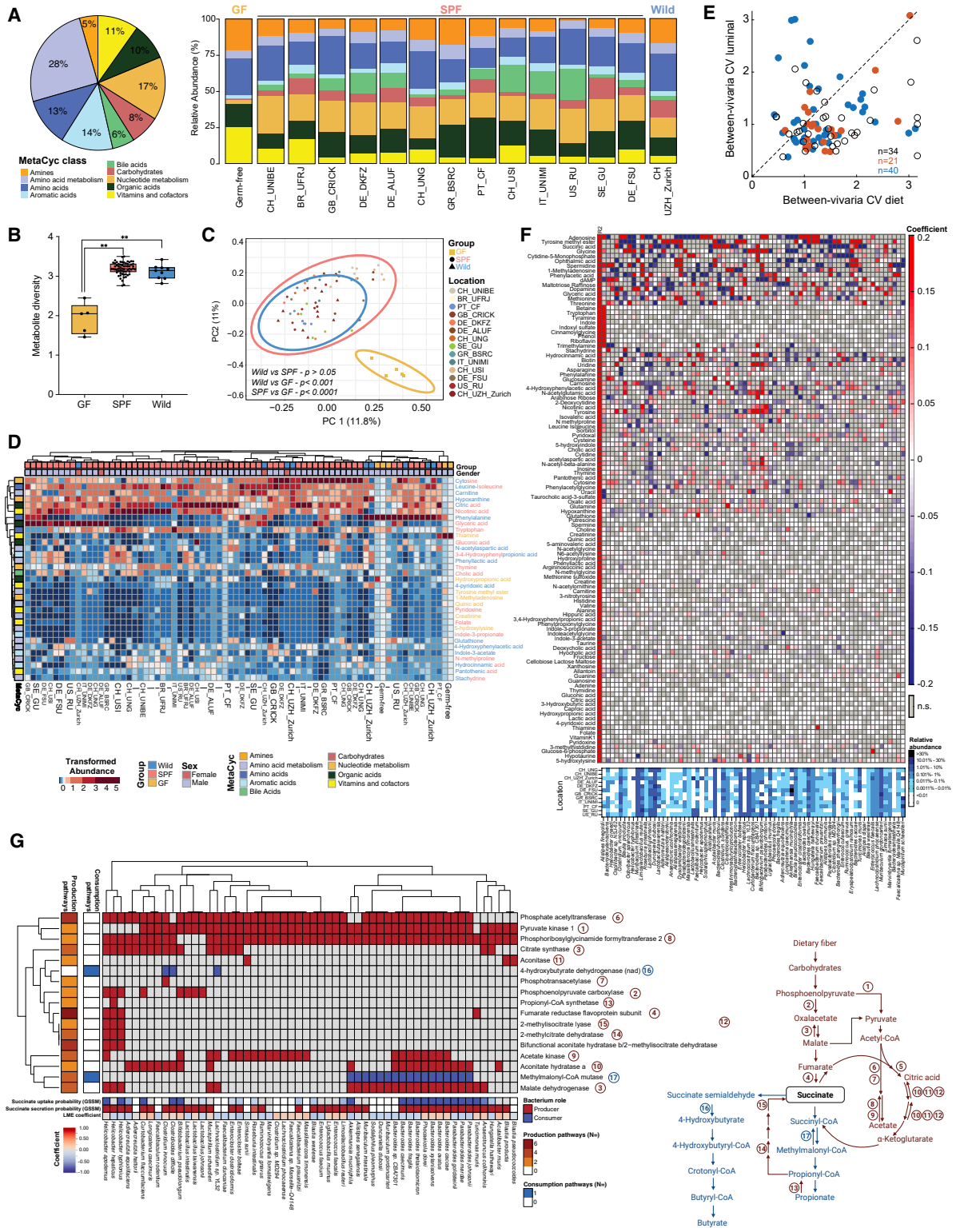


Figure 6. Gut metabolic profile of GF, wild, and SPF mice

(A) Categorization of abundant metabolites in 62 cecal samples from GF, wild, and 13 SPF vivaria, based on MetaCyc annotations. Bar graphs show mean relative abundance per class.

(B) Boxplots of metabolite richness in GF, wild, and SPF mice (quartiles, ranges, and standard deviations). $^{**}p < 0.01$.

(C) PCoA of Bray-Curtis distances showing metabolite differences among GF, wild, and SPF mice.

(legend continued on next page)

species composition primarily shapes the luminal metabolome. Figure 6F shows the estimated, significant consumer and producer relationships across metabolites; here, color indicates estimated metabolite production or consumption, and color intensity reflects the change in model R^2 upon species exclusion. Many models include contributions by many (producer and consumer groups of) species. This indicates that the gut metabolic environment, at least for these metabolites, may be a result of widespread cross-feeding and collaboration, rather than being determined by a few, specialized ''keystone'' species. A lack of correlation between species abundance and number or strength of estimated effects on the luminal metabolome (Figure S6E) supports this notion. Finally, for many models, the analysis highlights limited sets of putative contributor species with high estimated effects; this may be due to actual metabolic specialization, or it may be related to the limited share of taxa identified at the species level, or to other factors such as un-modeled species-species interactions. We concluded that the realized metabolic potential has a variability that is consistent with functional compensations and that it can be explained primarily by taxonomic composition and extensive metabolic interactions between taxa.

For a more in-depth analysis, we focused on succinate, which is recognized as a product of microbiota complex carbohydrate fermentation and is a carbon source for downstream short-chain fatty acid (SCFA) production. In view of the general importance of SCFAs and succinate in host physiology and immunity,[50] we analyzed the annotated pathways for the production and consumption of succinate in these taxa, as well as their predicted capabilities for succinate production and consumption based on GSMMs (see STAR Methods). These two characterizations of metabolic potential were highly consistent with the inferred contributions from linear mixed-effect models (Figure 6G), suggesting that inferences from matched metabolomics and metagenomics data generate valid hypotheses on producer-consumer relations in the microbiome.

Metabolite concentrations reflect a steady-state shaped by diet, host metabolite shedding, and microbial metabolism. Despite pronounced taxonomic variation between vivaria and wild populations, metabolomic profiles aligned with metagenomic and genome-scale modeling, indicating that cumulative metabolic potential broadly compensates for individual differences. The degree of compensation, however, depends on compound, class, diet, and environmental context. Notably, taxa differ in their contribution to functional metabolic outcomes, which cannot be inferred solely from encoded potential.

DISCUSSION

Beyond the observed taxonomic and species variant diversity, our analyses across 51 different colonies of healthy wild-type lab mice and 12 different colonies of wild mice reveal extensive and previously underappreciated sub-species diversification within individual taxa. This deep strain-level heterogeneity differs markedly between individual animals, vivaria, and habitats, substantiating each microbiota as its own unique ecosystem. Metagenomic data have indicated evidence for the potential of extensive metabolic compensation between different microbiota taxa within each consortium, but this is quite different from whether this potential is actually realized. In this study, we found that many metabolite concentrations show low variability across wild and SPF mice and that for many metabolites, this variability can be explained by differences in microbiota composition. Although our data support the overall conclusions of microbiome sequencing studies, which reveal considerable functional overlap in encoded metabolic pathways, we also demonstrate that the extent of compensation in healthy wild-type mice at steady state is variable according to metabolite class. As shown for the heterogeneity of bile salt metabolism,[16] it is certainly incorrect to assume general adequacy of metabolic compensation to eliminate phenotypic differences between intestinal microbial compositions of healthy mice, as some of these differences, such as succinate and other SCFA precursors,[51] likely cause important differences in experimental phenotypes.

There are also differences between wild and vivarium mice, as well as between mice in different vivaria, which can be identified by community-level metagenomic-encoded metabolic pathways or in genome-scale reaction-specific linear mixed-effect models. We found community-level metabolism differences in the early-life microbiota during lactation consistent with ecological consumer-resource models.[52,53] Another example was *de novo* nucleotide synthesis, found mainly in wild mice but also in mice from some vivaria, likely driven by the reduced availability of nucleotides for salvage pathways from effete cells or the diet. These findings are likely also consistent with the ecological consumer-resource model[28] and with the concept that nucleotides are semi-essential nutritional components since the intestinal mucosa preferentially uses preformed pyrimidine and purine

(D) Heatmap showing the significant metabolite differences between GF, wild mice, and SPF mice. Color coding on the right-hand *y* axis of the heatmap indicates metabolites with a significantly elevated concentration in wild mice (blue), SPF mice (red), or germ-free (yellow).
(E) Between-vivaria CVs for 95 metabolites in diet (vivaria-specific) vs. luminal content (mean across animals) from 47 wild and SPF mice across 10 vivaria. Blue/red indicate significantly higher/lower variability in luminal samples (single-tailed F test, $\alpha = 0.05$); black, not significant. Numbers denote metabolite counts.
(F) Heatmap illustrating associations of species abundances (per animal) and diet metabolite concentrations (per vivarium) with quantified luminal metabolites across samples in (E). Estimated contributions by species included in optimal models with $R^2 > 0.9$ are shown by color (blue: consumers; red: producers; gray: not significant) and color intensity (change in model R^2>, ΔR^2, when the respective species is excluded, capturing its marginal contribution to explained variance). Column R^2 gives the full model's value scaled by 0.2. Only covariates with at least one large effect ($\Delta R^2 > 0.1$) were included.
(G) Overview of succinate-producing (red arrows) and succinate-consuming (blue arrows) pathways in the gut microbiota. The metabolic map illustrates the conversions from carbohydrates to succinate and SCFAs, such as butyrate and propionate, with each reaction step annotated by a number corresponding to the enzyme function. The accompanying heatmap displays the presence of predicted enzymes, predicted productions/consumption capabilities, and inferred consumer-producer relationships across individual bacterial species. In the rows corresponding to metabolic enzymes, colored cells indicate the presence of the enzymatic reaction (red: succinate production; blue: consumption) in individual bacterial species according to MetaCyc pathways. ''Succinate uptake/secretion probability'' indicates species-level probabilities from AGORA2 GSMMs, averaged across mapped models (blank = no mapping). ''LME coefficient'' refers to estimates shown in (F).

bases for replication.[54–56] Unlike fortified rodent chows, wild diets are probably low in nucleotide content.[57] This implies that the generation of wider phenotype ranges through rewilding experiments should also take into account the nutritional content. In parallel, recent studies have proposed that transplantable, co-evolved natural microbiota offer a promising route to enhance model robustness and reproducibility across institutions, supporting the broader adoption of natural microbiota-based mouse models in biomedical research.[6,7]

The reproducibility crisis in preclinical mouse research has been increasingly attributed to hidden variables in microbiota composition across animal facilities. Despite standardization of genetics, diet, and husbandry, microbiota-driven phenotypic differences continue to limit generalizability. Our findings directly highlight this problem by demonstrating that taxonomic variation between vivaria and wild populations underlies significant divergence in metabolic outputs. This underscores recent calls to shift toward natural microbiota-based models that retain ecological complexity and evolutionary co-adaptation, which may improve reproducibility and translational relevance in biomedical research.[6,7,17,18]

The complexity of the microbiota and the technical challenges of following *in vivo* metabolite fluxes within individual taxa of a consortium have so far limited the analysis to steady-state concentrations of metabolites in intestinal fluid. We have shown examples of genetic variations consistent with positive selection that occur at active or critical structural sites within individual taxa at committal points for nucleotide, amino acid, and vitamin microbial metabolism. A future-forward genetic approach may be useful to complement flux studies between taxa in the disambiguation of overall consortial metabolism and how this is compensated for with different taxon compositions.

Unlike a stable ecological environment, the intestine needs to cope with changes in feeding, fasting, circadian rhythms, and differences in dietary intake. We suggest that maintaining sufficient microbiota species diversity and species variants (in some cases, variants that can potentially bloom from a low background allele frequency) is required for the flexibility to adapt blooms and contractions within different metabolic niches, thereby allowing community metabolism to respond to changing environmental and dietary conditions.

Beyond metabolic flexibility, an additional possibility is that microbiota taxonomic and strain-level variability may serve as a source of antigenic diversity, influencing host immune development and function. Rather than being purely stochastic, these microbiota differences, particularly across vivaria and environmental contexts, may reflect adaptive ecological tuning, offering a dynamic repertoire of microbial-associated molecular patterns (MAMPs) that shape immune education.[58–60] This concept aligns with the broader idea that immune system maturation in early life is not only influenced by microbial density or load but also by the diversity and novelty of antigens encountered. Such a model would suggest that microbiota variation contributes to host adaptability by preparing the immune system for a wider range of environmental exposures and pathogens.

Notably, despite the clear taxonomic divergence between wild mice, SPF mice, and even human microbiota, we observed a marked convergence in the encoded metabolic potential and realized intestinal metabolite profiles. This raises the possibility that host-derived selective pressures constrain microbiota function toward conserved metabolic outputs, regardless of taxonomic identity. In this view, host-microbiota interactions may reflect a form of functional phylo-symbiosis or co-evolution, where different microbial taxa are filtered and maintained to meet shared physiological demands across mammalian hosts.[61,62] This observation is consistent with prior studies showing that gut microbial communities from non-native hosts (e.g., human microbiota introduced into germ-free mice) fail to engraft robustly or elicit appropriate host responses, highlighting the strong selective pressures exerted by the host environment on microbiota assembly.[63–65] While speculative, this framework would be consistent with recent findings suggesting that microbial functions, particularly those involved in nutrient processing, immune modulation, and barrier maintenance, may be prioritized over taxonomic composition, and that host environments play an active role in shaping this outcome.[66]

Limitations of the study

While our study provides a comprehensive global survey of mouse microbiota diversity and its metabolic consequences across 51 vivaria and 12 wild mouse colonies, several limitations should be noted. First, although the dataset includes samples from institutions across six continents, geographic representation remains skewed toward European facilities, and broader institutional inclusion from underrepresented regions would further enhance the global scope. This resource is intended to remain extensible, and future efforts may incorporate additional institutions to expand taxonomic and ecological diversity. Second, while the study is anchored in a large 16S rRNA gene sequencing dataset spanning over 4,000 samples, only a subset of animals with matched cecal and fecal material were selected for deep metagenomic and metabolomic profiling. This design enabled direct multi-omic integration but limited the number of samples in which functional resolution could be achieved. Third, current bioinformatic tools and reference databases remain biased toward human-associated taxa, potentially underestimating the functional potential of wild or understudied microbial species. Lastly, while we observed robust patterns of metabolic compensation at steady state, future longitudinal studies with controlled perturbations will be necessary to evaluate dynamic responses and causal mechanisms.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Andrew Macpherson (andrew.macpherson@unibe.ch).

Materials availability

This study did not generate new and unique reagents.

Data and code availability

- Sequencing data: raw data are deposited in NCBI BioProject PRJNA1049830 (SRA: SUB14023600).
- 16S rRNA data: processed files and sample details on Figshare (DOI: 10.6084/m9.figshare.24523324).
- Metabolomics data: deposited at MassIVE (MSV000095413, https://massive.ucsd.edu/ProteoSAFe/static/massive.jsp).

- Code and metadata: no new code was developed in this study. All representative scripts and metadata are available at https://github.com/yilmazlab/Vivaria-Microbiota.
- Reaction modeling scripts: GSMM mapping, normalized reaction abundance, and subsystem annotation in MATLAB R2022b; model fitting and data processing in R 4.2.1; downstream analyses, *p* value correction, subsystem identification, and figure generation in Python 3.10.9. Relevant code at https://gitlab.com/csb.ethz/mouse-vivaria-metabolomics.
- Metagenome-diet-metabolome integration: command lines and metadata at https://gitlab.com/csb.ethz/mouse-vivaria-reactions.
- Interactive exploration: R Shiny apps available at https://yilmazlab.shinyapps.io/vivaria_app/ (OMICS data) and https://yilmazlab.shinyapps.io/vivaria_app2/ (vivaria variants).
- Any additional information required to reanalyze the data reported in this paper is available from the corresponding authors (bahtiyar.yilmaz@unibe.ch [B.Y.] and andrew.macpherson@unibe.ch [A.J.M.]).

CONSORTIA

The members of International Mouse Microbiota Investigators are Pinar Ciftci, Julian Evans, Bruce Boatman, Sarah C.L. Knowles, Eveliina Hanski, John F. Baines, Peter J. Cowan, Fabio Grassi, Sara Maffei, Marcelo T. Bozza, Vinicius M. Vidal, Arthur Kaser, Lorraine M. Holland, Raffaella Gozzelino, Silvia Cardoso, Pilar Martín, Cecile Fremond, Patricia Lopes, Rui M. Costa, Ana Mafalda Vicente, Yuuki Obata, Vassilis Pachnis, Ali Altıntaş, Lewin Small, Ali Adnan, Mathias Heikenwalder, Sandra Offner, Didier Trono, Sinem Karaman, Myrto Denaxa, Omar Mossad, Marco Prinz, Thomas Blank, Mirko Trajkovski, Dorothée Rigo, Andy Wullaert, Lars Vereecke, Wendy S. Garrett, Sena Bae, Viktoria Jeney, Sofia Lamas, Bruno Silva-Santos, Natacha Gonçalves-Sousa, Ilker Karacan, Sebastian Weis, Daniel Kariuki, Ewoud N. Speksnijder, Iqbal Hamza, Maria Rescigno, Michela Lizier, Andrea Cerutti, Emilie K. Grasset, Fiona Powrie, Claire Pearson, Lilian Lam, Jerzy Kotlinowski, Sin-Hyeog Im, Hiroshi Ohno, Takashi Kanaya, George M. H. Birchenough, Stephanie Claudinot, Pascalle L. P. Van Loo, Marco Cassano, Zuri Sullivan, Sven Pettersson, Kwan Soon Kim, Aikaterini Nanou, and Hai Li.

ACKNOWLEDGMENTS

We are grateful to Daniel Mucida and Gregory Donaldson (Rockefeller University), Ruslan Medzhitov (Yale University School of Medicine), Miguel P. Soares (Instituto Gulbenkian de Ciência), Jo Spencer and William Guesdon (King's College London), Anita S.F. Chong and Qiang Wang (University of Chicago), Mathilde J.H. Girard-Madoux and Elisabeth Wieduwild (Aix-Marseille Université), Chris J. Janse (Leiden University Medical Center), Mario Bonalli (University of Zurich), and Henrik Rasmussen (Oslo University Hospital) for generously providing samples. Shotgun metagenomic sequencing was conducted on UBELIX, the HPC cluster at the University of Bern, with sequencing performed by the University of Bern's NGS Platform. We further thank Matthew R. Olm (University of Colorado Boulder) and Alexander Crits-Christoph (University of California) for guidance on inStrain pipeline and Francesco Asnicar and Nicola Segata (Università di Trento) for support with HUMAnN pipeline. This research was funded by the Swiss National Science Foundation (SNF Starting grant TMSGI3_211300 to B.Y.; SNF Sinergia CRSII3_154414 and CRSII5_177164 to A.J.M., U.S., and J.S.). Additionally, A.J.M. received support from the European Research Council under the HHMM-Neonates Project (grant number 742195).

AUTHOR CONTRIBUTIONS

Conceptualization, B.Y. and A.J.M.; study supervision, B.Y. and A.J.M.; investigation, B.Y., A.J.M., I.B., C.L.G., K.H.U.M., S.B.U.J., B.K., A.K.L., C.B., J.S., and U.S.; sample collection, B.Y., C.B., A.K.L., J.L., and I.M.M.I.; data analysis, B.Y., A.J.M., I.B., S.B.U.J., C.L.G., K.H.U.M., J.S., and U.S.; resources, A.J.M. and B.Y.; writing – original draft, B.Y. and A.J.M.; writing – review and editing, A.J.M., U.S., J.S., and B.Y.; and funding acquisition, B.Y., A.J.M., U.S., and J.S.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

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

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Gut microbiota sampling from lab (SPF) mice and wild mice
  - Human subjects
  - Ethics statement
- METHOD DETAILS
  - Sample preparation for microbiota profiling in metagenomics and 16S rRNA gene sequencing
  - Amplicon sequencing on IonTorrent PGM™ platform
  - Library preparation, sequencing for metagenomics and data analysis
  - The inStrain pipeline
  - Prediction of vivarium-specific 3D protein structures
  - Prediction of protein-ligand binding and multimerization
  - Metabolite extraction from cecum samples
  - Metabolite extraction from diet pellets
  - Chemicals
  - Metabolite profiling using LC-TOF-MS
  - Preparation of standards for calibration and quantification
  - Data processing, analysis, and visualization of metabolomic data
  - GSMM mapping
  - Normalized reaction abundance and statistical model
  - Subsystem annotation and identification of significantly different subsystems
  - Statistical model for metabolome analysis
  - Succinate metabolism – analysis of metabolic potential in selected bacterial species
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.chom.2025.10.010.

Received: June 6, 2025
Revised: September 18, 2025
Accepted: October 9, 2025
Published: November 3, 2025

REFERENCES

1. Stappenbeck, T.S., and Virgin, H.W. (2016). Accounting for reciprocal host-microbiome interactions in experimental science. Nature *534*, 191–199. https://doi.org/10.1038/nature18285.
2. McCafferty, J., Mühlbauer, M., Gharaibeh, R.Z., Arthur, J.C., Perez-Chanona, E., Sha, W., Jobin, C., and Fodor, A.A. (2013). Stochastic changes over time and not founder effects drive cage effects in microbial community assembly in a mouse model. ISME J. *7*, 2116–2125. https://doi.org/10.1038/ismej.2013.106.
3. Macpherson, A.J., and McCoy, K.D. (2015). Standardised animal models of host microbial mutualism. Mucosal Immunol. *8*, 476–486. https://doi.org/10.1038/mi.2014.113.

4. Eberl, C., Ring, D., Münch, P.C., Beutler, M., Basic, M., Slack, E.C., Schwarzer, M., Srutkova, D., Lange, A., Frick, J.S., et al. (2019). Reproducible Colonization of Germ-Free Mice With the Oligo-Mouse-Microbiota in Different Animal Facilities. Front. Microbiol. *10*, 2999. https://doi.org/10.3389/fmicb.2019.02999.

5. Hooper, L.V., Littman, D.R., and Macpherson, A.J. (2012). Interactions between the microbiota and the immune system. Science *336*, 1268–1273. https://doi.org/10.1126/science.1223490.

6. Runge, S., von Zedtwitz, S., Maucher, A.M., Bruno, P., Osbelt, L., Zhao, B., Gernand, A.M., Lesker, T.R., Gräwe, K., Rogg, M., et al. (2025). Laboratory mice engrafted with natural gut microbiota possess a wild-ling-like phenotype. Nat. Commun. *16*, 5301. https://doi.org/10.1038/s41467-025-60554-2.

7. Bruno, P., Schüler, T., and Rosshart, S.P. (2025). Born to be wild: utilizing natural microbiota for reliable biomedical research. Trends Immunol. *46*, 17–28. https://doi.org/10.1016/j.it.2024.11.013.

8. Holmes, E., Li, J.V., Athanasiou, T., Ashrafian, H., and Nicholson, J.K. (2011). Understanding the role of gut microbiome-host metabolic signal disruption in health and disease. Trends Microbiol. *19*, 349–359. https://doi.org/10.1016/j.tim.2011.05.006.

9. Thompson, L.R., Sanders, J.G., McDonald, D., Amir, A., Ladau, J., Locey, K.J., Prill, R.J., Tripathi, A., Gibbons, S.M., Ackermann, G., et al. (2017). A communal catalogue reveals Earth's multiscale microbial diversity. Nature *551*, 457–463. https://doi.org/10.1038/nature24621.

10. Wang, J., Pan, Z., Yu, J., Zhang, Z., and Li, Y.Z. (2023). Global assembly of microbial communities. mSystems *8*, e0128922. https://doi.org/10.1128/msystems.01289-22.

11. Human; Microbiome; Project Consortium (2012). Structure, function and diversity of the healthy human microbiome. Nature *486*, 207–214. https://doi.org/10.1038/nature11234.

12. Louca, S., Polz, M.F., Mazel, F., Albright, M.B.N., Huber, J.A., O'Connor, M.I., Ackermann, M., Hahn, A.S., Srivastava, D.S., Crowe, S.A., et al. (2018). Function and functional redundancy in microbial systems. Nat. Ecol. Evol. *2*, 936–943. https://doi.org/10.1038/s41559-018-0519-1.

13. Tian, L., Wang, X.W., Wu, A.K., Fan, Y., Friedman, J., Dahlin, A., Waldor, M.K., Weinstock, G.M., Weiss, S.T., and Liu, Y.Y. (2020). Deciphering functional redundancy in the human microbiome. Nat. Commun. *11*, 6217. https://doi.org/10.1038/s41467-020-19940-1.

14. Folz, J., Culver, R.N., Morales, J.M., Grembi, J., Triadafilopoulos, G., Relman, D.A., Huang, K.C., Shalon, D., and Fiehn, O. (2023). Human metabolome variation along the upper intestinal tract. Nat. Metab. *5*, 777–788. https://doi.org/10.1038/s42255-023-00777-z.

15. Rehermann, B., Graham, A.L., Masopust, D., and Hamilton, S.E. (2025). Integrating natural commensals and pathogens into preclinical mouse models. Nat. Rev. Immunol. *25*, 385–397. https://doi.org/10.1038/s41577-024-01108-3.

16. Wang, M., Osborn, L.J., Jain, S., Meng, X., Weakley, A., Yan, J., Massey, W.J., Varadharajan, V., Horak, A., Banerjee, R., et al. (2023). Strain drop-outs reveal interactions that govern the metabolic output of the gut microbiome. Cell *186*, 2839–2852.e21. https://doi.org/10.1016/j.cell.2023.05.037.

17. Rosshart, S.P., Vassallo, B.G., Angeletti, D., Hutchinson, D.S., Morgan, A.P., Takeda, K., Hickman, H.D., McCulloch, J.A., Badger, J.H., Ajami, N.J., et al. (2017). Wild Mouse Gut Microbiota Promotes Host Fitness and Improves Disease Resistance. Cell *171*, 1015–1028.e13. https://doi.org/10.1016/j.cell.2017.09.016.

18. Rosshart, S.P., Herz, J., Vassallo, B.G., Hunter, A., Wall, M.K., Badger, J. H., McCulloch, J.A., Anastasakis, D.G., Sarshad, A.A., Leonardi, I., et al. (2019). Laboratory mice born to wild mice have natural microbiota and model human immune responses. Science *365*, eaaw4361. https://doi.org/10.1126/science.aaw4361.

19. Kang, D.D., Froula, J., Egan, R., and Wang, Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. PeerJ *3*, e1165. https://doi.org/10.7717/peerj.1165.

20. Wu, Y.W., Simmons, B.A., and Singer, S.W. (2016). MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics *32*, 605–607. https://doi.org/10.1093/bioinformatics/btv638.

21. Alneberg, J., Bjarnason, B.S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., Lahti, L., Loman, N.J., Andersson, A.F., and Quince, C. (2014). Binning metagenomic contigs by coverage and composition. Nat. Methods *11*, 1144–1146. https://doi.org/10.1038/nmeth.3103.

22. Mallawaarachchi, V., and Lin, Y. (2022). Accurate Binning of Metagenomic Contigs Using Composition, Coverage, and Assembly Graphs. J. Comput. Biol. *29*, 1357–1376. https://doi.org/10.1089/cmb.2022.0262.

23. Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T.B.K., Schulz, F., Jarett, J., Rivers, A.R., Eloe-Fadrosh, E.A., et al. (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat. Biotechnol. *35*, 725–731. https://doi.org/10.1038/nbt.3893.

24. Asnicar, F., Thomas, A.M., Beghini, F., Mengoni, C., Manara, S., Manghi, P., Zhu, Q., Bolzan, M., Cumbo, F., May, U., et al. (2020). Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. Nat. Commun. *11*, 2500. https://doi.org/10.1038/s41467-020-16366-7.

25. Maurice, C.F., Knowles, S.C.L., Ladau, J., Pollard, K.S., Fenton, A., Pedersen, A.B., and Turnbaugh, P.J. (2015). Marked seasonal variation in the wild mouse gut microbiota. ISME J. *9*, 2423–2434. https://doi.org/10.1038/ismej.2015.53.

26. Linnenbrink, M., Wang, J., Hardouin, E.A., Künzel, S., Metzler, D., and Baines, J.F. (2013). The role of biogeography in shaping diversity of the intestinal microbiota in house mice. Mol. Ecol. *22*, 1904–1916. https://doi.org/10.1111/mec.12206.

27. Graham, A.L. (2021). Naturalizing mouse models for immunology. Nat. Immunol. *22*, 111–117. https://doi.org/10.1038/s41590-020-00857-2.

28. MacArthur, R. (1970). Species packing and competitive equilibrium for many species. Theor. Popul. Biol. *1*, 1–11. https://doi.org/10.1016/0040-5809(70)90039-0.

29. Goldford, J.E., Lu, N., Bajić, D., Estrela, S., Tikhonov, M., Sanchez-Gorostiaga, A., Segrè, D., Mehta, P., and Sanchez, A. (2018). Emergent simplicity in microbial community assembly. Science *361*, 469–474. https://doi.org/10.1126/science.aat1168.

30. Goyal, A., Dubinkina, V., and Maslov, S. (2018). Multiple stable states in microbial communities explained by the stable marriage problem. ISME J. *12*, 2823–2834. https://doi.org/10.1038/s41396-018-0222-x.

31. Yilmaz, B., Mooser, C., Keller, I., Li, H., Zimmermann, J., Bosshard, L., Fuhrer, T., Gomez de Agüero, M., Trigo, N.F., Tschanz-Lischer, H., et al. (2021). Long-term evolution and short-term adaptation of microbiota strains and sub-strains in mice. Cell Host Microbe *29*, 650–663. e9. https://doi.org/10.1016/j.chom.2021.02.001.

32. Olm, M.R., Crits-Christoph, A., Bouma-Gregson, K., Firek, B.A., Morowitz, M.J., and Banfield, J.F. (2021). inStrain profiles population microdiversity from metagenomic data and sensitively detects shared microbial strains. Nat. Biotechnol. *39*, 727–736. https://doi.org/10.1038/s41587-020-00797-0.

33. Viver, T., Conrad, R.E., Rodriguez-R, L.M., Ramírez, A.S., Venter, S.N., Rocha-Cárdenas, J., Llabrés, M., Amann, R., Konstantinidis, K.T., and Rossello-Mora, R. (2024). Towards estimating the number of strains that make up a natural bacterial population. Nat. Commun. *15*, 544. https://doi.org/10.1038/s41467-023-44622-z.

34. Rodriguez-R, L.M., Conrad, R.E., Viver, T., Feistel, D.J., Lindner, B.G., Venter, S.N., Orellana, L.H., Amann, R., Rossello-Mora, R., and Konstantinidis, K.T. (2024). An ANI gap within bacterial species that advances the definitions of intra-species units. mBio *15*, e0269623. https://doi.org/10.1128/mbio.02696-23.

35. Carter, M.M., Olm, M.R., Merrill, B.D., Dahan, D., Tripathi, S., Spencer, S. P., Yu, F.B., Jain, S., Neff, N., Jha, A.R., et al. (2023). Ultra-deep

sequencing of Hadza hunter-gatherers recovers vanishing gut microbes. Cell *186*, 3111–3124.e13. https://doi.org/10.1016/j.cell.2023.05.046.

36. Silale, A., and van den Berg, B. (2023). TonB-Dependent Transport Across the Bacterial Outer Membrane. Annu. Rev. Microbiol. *77*, 67–88. https://doi.org/10.1146/annurev-micro-032421-111116.
37. Gudkov, M., Thibaut, L., and Giannoulatou, E. (2024). Quantifying negative selection on synonymous variants. HGG Adv. *5*, 100262. https://doi.org/10.1016/j.xhgg.2024.100262.
38. Zhou, T., Daugherty, M., Grishin, N.V., Osterman, A.L., and Zhang, H. (2000). Structure and mechanism of homoserine kinase: prototype for the GHMP kinase superfamily. Structure *8*, 1247–1257. https://doi.org/10.1016/s0969-2126(00)00533-5.
39. Smith, J.L. (1998). Glutamine PRPP amidotransferase: snapshots of an enzyme in action. Curr. Opin. Struct. Biol. *8*, 686–694. https://doi.org/10.1016/s0959-440x(98)80087-0.
40. Hove-Jensen, B., Andersen, K.R., Kilstrup, M., Martinussen, J., Switzer, R.L., and Willemoës, M. (2017). Phosphoribosyl Diphosphate (PRPP): Biosynthesis, Enzymology, Utilization, and Metabolic Significance. Microbiol. Mol. Biol. Rev. *81*, e00040–e00016. https://doi.org/10.1128/MMBR.00040-16.
41. Gräwert, T., Fischer, M., and Bacher, A. (2013). Structures and reaction mechanisms of GTP cyclohydrolases. IUBMB Life *65*, 310–322. https://doi.org/10.1002/iub.1153.
42. Ladner, J.E., Obmolova, G., Teplyakov, A., Howard, A.J., Khil, P.P., Camerini-Otero, R.D., and Gilliland, G.L. (2003). Crystal structure of Escherichia coli protein ybgI, a toroidal structure with a dinuclear metal site. BMC Struct. Biol. *3*, 7. https://doi.org/10.1186/1472-6807-3-7.
43. Beghini, F., McIver, L.J., Blanco-Míguez, A., Dubois, L., Asnicar, F., Maharjan, S., Mailyan, A., Manghi, P., Scholz, M., Thomas, A.M., et al. (2021). Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. eLife *10*, e65088. https://doi.org/10.7554/eLife.65088.
44. Mallick, H., Rahnavard, A., McIver, L.J., Ma, S., Zhang, Y., Nguyen, L.H., Tickle, T.L., Weingart, G., Ren, B., Schwager, E.H., et al. (2021). Multivariable association discovery in population-scale meta-omics studies. PLoS Comput. Biol. *17*, e1009442. https://doi.org/10.1371/journal.pcbi.1009442.
45. Heinken, A., Hertel, J., Acharya, G., Ravcheev, D.A., Nyga, M., Okpala, O. E., Hogan, M., Magnúsdóttir, S., Martinelli, F., Nap, B., et al. (2023). Genome-scale metabolic reconstruction of 7,302 human microorganisms for personalized medicine. Nat. Biotechnol. *41*, 1320–1331. https://doi.org/10.1038/s41587-022-01628-0.
46. Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M., and Ishiguro-Watanabe, M. (2023). KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Res. *51*, D587–D592. https://doi.org/10.1093/nar/gkac963.
47. Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. *28*, 27–30. https://doi.org/10.1093/nar/28.1.27.
48. Kanehisa, M. (2019). Toward understanding the origin and evolution of cellular organisms. Protein Sci. *28*, 1947–1951. https://doi.org/10.1002/pro.3715.
49. Meier, K.H.U., Trouillon, J., Li, H., Lang, M., Fuhrer, T., Zamboni, N., Sunagawa, S., Macpherson, A.J., and Sauer, U. (2023). Metabolic landscape of the male mouse gut identifies different niches determined by microbial activities. Nat. Metab. *5*, 968–980. https://doi.org/10.1038/s42255-023-00802-1.
50. Wei, Y.H., Ma, X., Zhao, J.C., Wang, X.Q., and Gao, C.Q. (2023). Succinate metabolism and its regulation of host-microbe interactions. Gut Microbes *15*, 2190300. https://doi.org/10.1080/19490976.2023.2190300.
51. Hosomi, K., Saito, M., Park, J., Murakami, H., Shibata, N., Ando, M., Nagatake, T., Konishi, K., Ohno, H., Tanisawa, K., et al. (2022). Oral administration of Blautia wexlerae ameliorates obesity and type 2 diabetes via metabolic remodeling of the gut microbiota. Nat. Commun. *13*, 4477. https://doi.org/10.1038/s41467-022-32015-7.
52. Orlando, P.A., Brown, J.S., and Wise, D.H. (2012). Coexistence and community structure in a consumer resource model with implicit stoichiometry. Theor. Popul. Biol. *82*, 77–84. https://doi.org/10.1016/j.tpb.2012.05.005.
53. Silverstein, M., Bhatnagar, J.M., and Segrè, D. (2023). Metabolic complexity drives divergence in microbial communities. Preprint at bioRxiv. https://doi.org/10.1101/2023.08.03.551516.
54. Sánchez-Pozo, A., and Gil, A. (2002). Nucleotides as semiessential nutritional components. Br. J. Nutr. *87*, S135–S137. https://doi.org/10.1079/bjn2001467.
55. Sanderson, I.R., and He, Y. (1994). Nucleotide uptake and metabolism by intestinal epithelial cells. J. Nutr. *124*, 131S–137S. https://doi.org/10.1093/jn/124.suppl_1.131S.
56. Bronk, J.R., and Hastewell, J.G. (1988). The transport and metabolism of naturally occurring pyrimidine nucleosides by isolated rat jejunum. J. Physiol. *395*, 349–361. https://doi.org/10.1113/jphysiol.1988.sp016923.
57. Sauer, N., Mosenthin, R., and Bauer, E. (2011). The role of dietary nucleotides in single-stomached animals. Nutr. Res. Rev. *24*, 46–59. https://doi.org/10.1017/S0954422410000326.
58. Belkaid, Y., and Hand, T.W. (2014). Role of the microbiota in immunity and inflammation. Cell *157*, 121–141. https://doi.org/10.1016/j.cell.2014.03.011.
59. Thaiss, C.A., Zmora, N., Levy, M., and Elinav, E. (2016). The microbiome and innate immunity. Nature *535*, 65–74. https://doi.org/10.1038/nature18847.
60. Ramanan, D., Bowcutt, R., Lee, S.C., Tang, M.S., Kurtz, Z.D., Ding, Y., Honda, K., Gause, W.C., Blaser, M.J., Bonneau, R.A., et al. (2016). Helminth infection promotes colonization resistance via type 2 immunity. Science *352*, 608–612. https://doi.org/10.1126/science.aaf3229.
61. Groussin, M., Mazel, F., Sanders, J.G., Smillie, C.S., Lavergne, S., Thuiller, W., and Alm, E.J. (2017). Unraveling the processes shaping mammalian gut microbiomes over evolutionary time. Nat. Commun. *8*, 14319. https://doi.org/10.1038/ncomms14319.
62. Moeller, A.H., Caro-Quintero, A., Mjungu, D., Georgiev, A.V., Lonsdorf, E. V., Muller, M.N., Pusey, A.E., Peeters, M., Hahn, B.H., and Ochman, H. (2016). Cospeciation of gut microbiota with hominids. Science *353*, 380–382. https://doi.org/10.1126/science.aaf3951.
63. Chung, H., Pamp, S.J., Hill, J.A., Surana, N.K., Edelman, S.M., Troy, E.B., Reading, N.C., Villablanca, E.J., Wang, S., Mora, J.R., et al. (2012). Gut immune maturation depends on colonization with a host-specific microbiota. Cell *149*, 1578–1593. https://doi.org/10.1016/j.cell.2012.04.037.
64. Moeller, A.H., Suzuki, T.A., Phifer-Rixey, M., and Nachman, M.W. (2018). Transmission modes of the mammalian gut microbiota. Science *362*, 453–457. https://doi.org/10.1126/science.aat7164.
65. Muegge, B.D., Kuczynski, J., Knights, D., Clemente, J.C., González, A., Fontana, L., Henrissat, B., Knight, R., and Gordon, J.I. (2011). Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science *332*, 970–974. https://doi.org/10.1126/science.1198719.
66. Ushey, K., and Wickham, H. (2023). Renv: Project Environments. GitHub. https://github.com/rstudio/renv.
67. Almeida, A., Nayfach, S., Boland, M., Strozzi, F., Beracochea, M., Shi, Z. J., Pollard, K.S., Sakharova, E., Parks, D.H., Hugenholtz, P., et al. (2021). A unified catalog of 204,938 reference genomes from the human gut microbiome. Nat. Biotechnol. *39*, 105–114. https://doi.org/10.1038/s41587-020-0603-3.
68. Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G.A., Sonnhammer, E.L.L., Tosatto, S.C.E., Paladin, L., Raj, S., Richardson, L.J., et al. (2021). Pfam: The protein families database in 2021. Nucleic Acids Res. *49*, D412–D419. https://doi.org/10.1093/nar/gkaa913.
69. Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., and Ogata, H. (2020). KofamKOALA: KEGG Ortholog

assignment based on profile HMM and adaptive score threshold. Bioinformatics *36*, 2251–2252. https://doi.org/10.1093/bioinformatics/btz859.

70. Drula, E., Garron, M.L., Dogan, S., Lombard, V., Henrissat, B., and Terrapon, N. (2022). The carbohydrate-active enzyme database: functions and literature. Nucleic Acids Res. *50*, D571–D577. https://doi.org/10.1093/nar/gkab1045.
71. Alcock, B.P., Huynh, W., Chalil, R., Smith, K.W., Raphenya, A.R., Wlodarski, M.A., Edalatmand, A., Petkau, A., Syed, S.A., Tsang, K.K., et al. (2023). CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. Nucleic Acids Res *51*, D690–D699. https://doi.org/10.1093/nar/gkac920.
72. Yilmaz, B., Juillerat, P., Øyås, O., Ramon, C., Bravo, F.D., Franc, Y., Fournier, N., Michetti, P., Mueller, C., Geuking, M., et al. (2019). Microbial network disturbances in relapsing refractory Crohn's disease. Nat. Med. *25*, 323–336. https://doi.org/10.1038/s41591-018-0308-z.
73. Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. *37*, 852–857. https://doi.org/10.1038/s41587-019-0209-9.
74. Morgan, X.C., Tickle, T.L., Sokol, H., Gevers, D., Devaney, K.L., Ward, D. V., Reyes, J.A., Shah, S.A., LeLeiko, N., Snapper, S.B., et al. (2012). Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. Genome Biol. *13*, R79. https://doi.org/10.1186/gb-2012-13-9-r79.
75. Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods *13*, 581–583. https://doi.org/10.1038/nmeth.3869.
76. Uritskiy, G.V., DiRuggiero, J., and Taylor, J. (2018). MetaWRAP-a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome *6*, 158. https://doi.org/10.1186/s40168-018-0541-1.
77. McMurdie, P.J., and Holmes, S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS One *8*, e61217. https://doi.org/10.1371/journal.pone.0061217.
78. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A.J., Bambrick, J., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature *630*, 493–500. https://doi.org/10.1038/s41586-024-07487-w.
79. Sehnal, D., Bittrich, S., Deshpande, M., Svobodová, R., Berka, K., Bazgier, V., Velankar, S., Burley, S.K., Koča, J., and Rose, A.S. (2021). Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. Nucleic Acids Res. *49*, W431–W437. https://doi.org/10.1093/nar/gkab314.
80. Yang, J., Roy, A., and Zhang, Y. (2013). Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. Bioinformatics *29*, 2588–2595. https://doi.org/10.1093/bioinformatics/btt447.
81. Adasme, M.F., Linnemann, K.L., Bolz, S.N., Kaiser, F., Salentin, S., Haupt, V.J., and Schroeder, M. (2021). PLIP 2021: expanding the scope of the protein-ligand interaction profiler to DNA and RNA. Nucleic Acids Res. *49*, W530–W534. https://doi.org/10.1093/nar/gkab294.
82. Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. Bioinformatics *30*, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
83. Wood, D.E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. Genome Biol. *20*, 257. https://doi.org/10.1186/s13059-019-1891-0.
84. Lu, J., Breitwieser, F.P., Thielen, P., and Salzberg, S.L. (2017). Bracken: estimating species abundance in metagenomics data. PeerJ Comput. Sci. *3*, e104. https://doi.org/10.7717/peerj-cs.104.
85. Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., and Glöckner, F.O. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. *35*, 7188–7196. https://doi.org/10.1093/nar/gkm864.
86. Li, D., Liu, C.M., Luo, R., Sadakane, K., and Lam, T.W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics *31*, 1674–1676. https://doi.org/10.1093/bioinformatics/btv033.
87. Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P.A. (2017). metaSPAdes: a new versatile metagenomic assembler. Genome Res. *27*, 824–834. https://doi.org/10.1101/gr.213959.116.
88. Wang, J., Linnenbrink, M., Künzel, S., Fernandes, R., Nadeau, M.J., Rosenstiel, P., and Baines, J.F. (2014). Dietary history contributes to enterotype-like clustering and functional metagenomic content in the intestinal microbiome of wild mice. Proc. Natl. Acad. Sci. USA *111*, E2703–E2710. https://doi.org/10.1073/pnas.1402342111.
89. König, B., Lindholm, A.K., Lopes, P.C., Dobay, A., Steinert, S., and Buschmann, F.J.-U. (2015). A system for automatic recording of social behavior in a free-living wild house mouse population. Anim. Biotelem. *3*, 39. https://doi.org/10.1186/s40317-015-0069-0.
90. Sundquist, A., Bigdeli, S., Jalili, R., Druzin, M.L., Waller, S., Pullen, K.M., El-Sayed, Y.Y., Taslimi, M.M., Batzoglou, S., and Ronaghi, M. (2007). Bacterial flora-typing with targeted, chip-based Pyrosequencing. BMC Microbiol. *7*, 108. https://doi.org/10.1186/1471-2180-7-108.
91. Schreiner, P., Yilmaz, B., Rossel, J.B., Franc, Y., Misselwitz, B., Scharl, M., Zeitz, J., Frei, P., Greuter, T., Vavricka, S.R., et al. (2019). Vegetarian or gluten-free diets in patients with inflammatory bowel disease are associated with lower psychological well-being and a different gut microbiota, but no beneficial effects on the course of the disease. U. Eur. Gastroenterol. J. *7*, 767–781. https://doi.org/10.1177/2050640619841249.
92. Callahan, B.J., Sankaran, K., Fukuyama, J.A., McMurdie, P.J., and Holmes, S.P. (2016). Bioconductor Workflow for Microbiome Data Analysis: from raw reads to community analyses. F1000Res *5*, 1492. https://doi.org/10.12688/f1000research.8986.2.
93. Anderson, M.J. (2006). Distance-based tests for homogeneity of multivariate dispersions. Biometrics *62*, 245–253. https://doi.org/10.1111/j.1541-0420.2005.00440.x.
94. McMurdie, P.J., and Holmes, S. (2014). Waste not, want not: why rarefying microbiome data is inadmissible. PLoS Comput. Biol. *10*, e1003531. https://doi.org/10.1371/journal.pcbi.1003531.
95. Wood, D.E., and Salzberg, S.L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. *15*, R46. https://doi.org/10.1186/gb-2014-15-3-r46.
96. Alcock, B.P., Huynh, W., Chalil, R., Smith, K.W., Raphenya, A.R., Wlodarski, M.A., Edalatmand, A., Petkau, A., Syed, S.A., Tsang, K.K., et al. (2023). CARD 2023: expanded curation, support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance Database. Nucleic Acids Res. *51*, D690–D699. https://doi.org/10.1093/nar/gkac920.
97. Zhao, S., Lieberman, T.D., Poyet, M., Kauffman, K.M., Gibbons, S.M., Groussin, M., Xavier, R.J., and Alm, E.J. (2019). Adaptive Evolution within Gut Microbiomes of Healthy People. Cell Host Microbe *25*, 656–667.e8. https://doi.org/10.1016/j.chom.2019.03.007.
98. Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). circlize Implements and enhances circular visualization in R. Bioinformatics *30*, 2811–2812. https://doi.org/10.1093/bioinformatics/btu393.
99. Yang, J., Roy, A., and Zhang, Y. (2013). BioLiP: a semi-manually curated database for biologically relevant ligand-protein interactions. Nucleic Acids Res. *41*, D1096–D1103. https://doi.org/10.1093/nar/gks966.
100. Hai, L., Limenitakis, J.P., Fuhrer, T., Geuking, M.B., Lawson, M.B., Wyss, M., Brugiroux, S., Keller, I., Macpherson, J.A., Rupp, S., et al. (2015). The outer mucus layer hosts a distinct intestinal microbial niche. Nat. Commun. *6*, 8292.

101. Heinken, A., and Thiele, I. (2015). Systematic prediction of health-relevant human-microbial co-metabolism through a computational framework. Gut Microbes *6*, 120–130. https://doi.org/10.1080/19490976.2015.1023494.

102. Kuznetsova, A., Brockhoff, P.B., and Christensen, R.H.B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. J. Stat. Softw. *82*, 1–26. https://doi.org/10.18637/jss.v082.i13.

103. Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Software *67*, 1–48. https://doi.org/10.18637/jss.v067.i01.

104. Halekoh, U., and Højsgaard, S. (2014). A Kenward-Roger Approximation and Parametric Bootstrap Methods for Tests in Linear Mixed Models – The R Package pbkrtest. J. Stat. Softw. *59*, 1–32. https://doi.org/10.18637/jss.v059.i09.

105. Lenth, R.V. (2023). emmeans: Estimated Marginal Means, aka Least-Squares Means. CRAN.R-Project. https://cran.r-project.org/web/packages/emmeans/emmeans.pdf.

106. Caspi, R., Billington, R., Ferrer, L., Foerster, H., Fulcher, C.A., Keseler, I. M., Kothari, A., Krummenacker, M., Latendresse, M., Mueller, L.A., et al. (2016). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res. *44*, D471–D480. https://doi.org/10.1093/nar/gkv1164.

107. Lewis, N.E., Hixson, K.K., Conrad, T.M., Lerman, J.A., Charusanti, P., Polpitiya, A.D., Adkins, J.N., Schramm, G., Purvine, S.O., Lopez-Ferrer, D., et al. (2010). Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models. Mol. Syst. Biol. *6*, 390. https://doi.org/10.1038/msb.2010.47.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
Methanol	Sigma	Cat#: 34860
Isopropanol	Sigma	Cat#: 34863
Acetic acid	Sigma	Cat#: 45754
9-anthracene carboxylic acid	Sigma	Cat#: A89405
Biological samples		
Human stool samples	This paper	N/A
Experimental models: Organisms/strains		
Mice: SwissAlbino	Jomo Kenyatta University of Agriculture and Technology	N/A
Mice: BALB/C	Jomo Kenyatta University of Agriculture and Technology	N/A
Mice: C57BL/6J	Jomo Kenyatta University of Agriculture and Technology	N/A
Mice: C57BL/6J	University of Science and Technology of China	N/A
Mice: C57BL/6N	RIKEN Center for Integrative Medical Sciences	N/A
Mice: C57BL/6J	RIKEN Center for Integrative Medical Sciences	N/A
Mice: C57BL/6J	National Neuroscience Institute	N/A
Mice: C57BL/6J	Pohang University of Science and Technology	N/A
Mice: C57BL/6J	University of Melbourne	N/A
Mice: C57BL/6J	Ghent University	N/A
Mice: C57BL/6J	University of Copenhagen	N/A
Mice: C57BL/6J	The Francis Crick Institute	N/A
Mice: C57BL/6J	King's College London	N/A
Mice: C57BL/6J	University of Oxford	N/A
Mice: C57BL/6JRccHsd	University of Helsinki	N/A
Mice: C57BL/6JOlaHsd	University of Helsinki	N/A
Mice: C57BL/6J	Aix-Marseille Université	N/A
Mice: C57BL/6J	University of Orleans	N/A
Mice: C57BL/6J	German Cancer Research Center Heidelberg	N/A
Mice: C57BL/6N	University of Freiburg	N/A
Mice: C57BL/6J	University of Freiburg	N/A
Mice: C57BL/6J	Biomedical Sciences Research Center "Alexander Fleming"	N/A
Mice: C57BL/6JLumc	Leiden University Medical Center	N/A
Mice: C57BL/6JOlaHsd	Utrecht University Medical Center	N/A
Mice: C57BL/6J	University of Debrecen	N/A
Mice: C57BL/6J	Milano University	N/A
Mice: C57BL/6JOlaHsd	Milano University	N/A
Mice: C57Bl/6N	University of Oslo	N/A
Mice: C57BL/6J	Jagiellonian University	N/A
Mice: C57BL/6J	Champalimaud Foundation	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mice: C57BL/6J	Instituto de Medicina Molecular	N/A
Mice: C57BL/6J	Instituto Gulbenkian de Ciência	N/A
Mice: C57BL/6J	Universidade do Porto	N/A
Mice: C57BL/6J	Centro Nacional de Investigaciones Cardiovasculares	N/A
Mice: C57BL/6J	University of Gothenburg	N/A
Mice: C57BL/6J	University of Bern	N/A
Mice: C57BL/6J	University of Zurich	N/A
Mice: C57BL/6J	Università della Svizzera Italiana	N/A
Mice: C57BL/6J	University of Geneva	N/A
Mice: B6.CBA-Tg(Pou5f1-EGFP)2Mnn/J.Kap1	École polytechnique fédérale de Lausanne	N/A
Mice: B6D2F1-Tg(Dux-KO)	École polytechnique fédérale de Lausanne	N/A
Mice: C57BL/6JOlaHsd	Université de Lausanne	N/A
Mice: C57BL/6J	Université de Lausanne	N/A
Mice: BALB/c	Istanbul Medeniyet University	N/A
Mice: C57BL/6J	Yale University School of Medicine	N/A
Mice: CD45.1.CD45.2.FoxP3GFP	University of Chicago	N/A
Mice: CD45.1.Igha	University of Chicago	N/A
Mice: CD45.1.CD45.2.Igha	University of Chicago	N/A
Mice: CD45.1	University of Chicago	N/A
Mice: CD45.1.FoxP3GFP	University of Chicago	N/A
Mice: C57BL/6J	University of Maryland	N/A
Mice: 3/4_C57BL/6J:_1/4_SVJ129	National Institutes of Health	N/A
Mice: C57BL/6J	National Institutes of Health	N/A
Mice: C57BL/6J	Harvard Medical School	N/A
Mice: C57BL/6J	Columbia University	N/A
Mice: C57BL/6J	Icahn School of Medicine at Mount Sinai	N/A
Mice: C57BL/6J	Rockefeller University	N/A
Mice: C57BL/6J	Universidade Federal do Rio de Janeiro	N/A
Mice: C57BL/6J	Friedrich-Schiller University	N/A
Mice: *Mus musculus domesticus*	University of Zurich	N/A
Mice: *Mus musculus domesticus*	Kiel University and Max Planck Institute for Evolutionary Biology	N/A
Mice: *Mus musculus domesticus*	University of Oxford	N/A
Mice: *Apodemus sylvaticus*	University of Oxford	N/A
Mice: *Apodemus flavicollis*	University of Oxford	N/A
Critical Commercial Assays		
PowerFecal Pro Kit	Qiagen	Cat#: 51804
QIAquick Gel Extraction Kit	Qiagen	Cat#: 28704
Ion PGM HiQ View Sequencing 400 kit	Thermo Fisher	Cat#: A30044
Qubit™ 1X dsDNA High Sensitivity (HS) Assay Kit	Thermo Fisher	Cat#: Q33231
Fragment Analyzer DNA Kit	Agilent	Cat#:5191-6570
NovaSeq 6000 S1 Reagent Kit v1.5 (300 cycles)	Illumina	Cat#:20028317
KAPA HiFi HotStart ReadyMix	Roche	Cat#: 07958935001
Ion PGM™ Sequencing 400 Kit	Thermo Fisher	Cat#: 4482002
Ion Chip 316™ V2	Thermo Fisher	Cat#: 4488149

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
Metabolomics Raw Data	This paper	MSV000095413
Metagenomic Raw Sequencing Data	This paper	PRJNA1049650
16S rRNA Amplicon Sequencing Data	This paper	10.6084/m9.figshare.24523324
UHGGv1	Almeida et al.[67]	http://ftp.ebi.ac.uk/pub/databases/metagenomics/mgnify_genomes/human-gut/v1.0
Pfam	Mistry et al.[68]	http://pfam.xfam.org/
Kofam	Aramaki et al.[69]	ftp://ftp.genome.jp/pub/db/kofam
CAZyDB	Drula et al.[70]	https://www.cazy.org/
CARD/broadstreet-v3.2.5	Alcock et al.[71]	https://card.mcmaster.ca/download
Oligonucleotides		
5′-CCTCTCTATGGGCAGTCGGTGATACGAGCTGACGACARCCATG-3′	Yilmaz et al.[72]	N/A
5′-CCATCTCATCCCTGCGTGTCTCCGACTCAG-BARCODE-ATTAGATACCCYGGTAGTCC-3	Yilmaz et al.[72]	N/A
Software and Algorithms		
QIIME v2	Bolyen et al.[73]	https://qiime2.org
R (v3.6.2 and v4.2.1)	R Core Team	https://www.r-project.org/
MaAsLin2	Morgan et al.[74]	https://huttenhower.sph.harvard.edu/maaslin/
DADA2	Callahan et al.[75]	https://benjjneb.github.io/dada2/
MetaWRAP	Uritskiy et al.[76]	https://github.com/bxlab/metaWRAP
phyloseq	McMurdie and Holmes[77]	https://joey711.github.io/phyloseq/
Graphpad Prism v10.0	www.graphpad.com	https://ggplot2.tidyverse.org/index.html
inStrain	Olm et al.[32]	https://instrain.readthedocs.io/en/latest
AlphaFold3	Abramson et al.[78]	https://alphafoldserver.com/about
Mol*viewer	Sehnal et al.[79]	https://molstar.org/viewer/
COACH	Yang et al.[80]	https://zhanggroup.org/COACH/
PLIP	Adasme et al.[81]	https://plip-tool.biotec.tu-dresden.de/plip-web/plip/index
Prokka	Seemann[82]	https://github.com/tseemann/prokka
kofam_scan-1.3.0	Aramaki et al.[69]	https://github.com/takaram/kofam_scan
Kraken2	Wood et al.[83]	https://ccb.jhu.edu/software/kraken2/
Bracken	Lu et al.[84]	https://ccb.jhu.edu/software/bracken/
SILVA 16S rRNA sequence database	Pruesse et al.[85]	https://www.arb-silva.de/
HUMAnN3	Beghini et al.[43]	https://github.com/biobakery/humann
MegaHit	Li et al.[86]	https://github.com/voutcn/megahit
metaSPAdes	Nurk et al.[87]	N/A
MetaBAT	Kang et al.[19]	https://bitbucket.org/berkeleylab/metabat/src/master/
MaxBin	Wu et al.[20]	https://sourceforge.net/projects/maxbin/
CONCOCT	Alneberg et al.[21]	https://github.com/BinPro/CONCOCT
MassHunter Quant Software	Agilent Technologies	https://www.agilent.com/en-us/support/software-informatics/mhquant-g3336aa-ssb
MATLAB R2021b	MathWorks	https://ch.mathworks.com/
R Shiny app for all dataset	This paper	https://yilmazlab.shinyapps.io/vivaria_app/

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Gut microbiota sampling from lab (SPF) mice and wild mice

This global analysis encompassed 2,388 murine stool samples predominantly derived from the C57BL/6 strain, collected from 51 research vivaria spanning 24 countries. Each participating institution contributed an average of 43 ± 13 stool and cecal samples, all collected under local specific pathogen–free (SPF) conditions as defined by the respective facilities. Additional strains included Swiss Albino (Kenya) and C57BL/6J:SVJ129 (3:1 hybrid) (Maryland). Animals were typically 8–18 weeks old and balanced for sex and age across groups (*see* Table S1 *for detailed metadata*). Mice were maintained in cages of up to five adults at 20–22.5 °C, 30–75 % humidity, and a 12 h light/12 h dark cycle, with unrestricted access to autoclaved chow and water.

Sample handling followed standardized procedures across institutions: stool or cecal content was collected into sterile 2 ml tubes, immediately frozen at -20°C, and transported on dry ice to Bern, where samples were stored at −80°C. Stool samples from the United Kingdom were preserved in DNA/RNA Shield before freezing and shipment, while DNA extracts from France and Germany were sent directly for integration into the study.

Wild mouse populations (n = 1,489) were sampled from a total of 12 designated locations in four different countries (Switzerland: *Mus musculus domesticus*; United Kingdom: *Mus musculus domesticus*, *Apodemus sylvaticus*, and *Apodemus flavicollis*; Germany[26,88]; *Mus musculus domesticus* and France[26,88]; *Mus musculus domesticus*), yielding an average of 11 ± 5 samples per location. Stool samples collected in the UK, wild mice were trapped with live traps (Small Sherman traps) provisioned with peanuts, non-absorbent cotton wool for bedding, and with a spray of sesame oil outside the trap used as a lure overnight. A notable contribution to this collection was made by the barn near Zürich, providing a total of 1376 samples.[89] Sample handling followed the same preservation and shipping procedures as for SPF mice. Colony codes and sampling locations are listed in Table S1. The annotation indicating the presence of wild mice was utilized in certain figure panels: Angers (AN), Louan (LO), MassifCentral (MC), Espelette (ES), Divonne les bains (DB), Nancy (NAA), Schoemberg (SL), Cologne (CB), Skokholm (SKO), London1 (UK1) and London2 (UK2).

All animal work adhered to national and institutional animal welfare regulations. Specific ethical approvals included: Veterinary Office Zurich (ZH 091/16, ZH 098/19); UK Home Office Project License PPL PB0178858 with additional permits from the Islands Conservation Advisory Committee and Natural Resources Wales; the Royal Veterinary College's Animal Welfare and Ethical Review Body (UK, *Apodemus spp.*); and Veterinäramt Kreis Plön (Germany, Permit 1401–144/PLÖ-004697). Equivalent permissions were obtained for France and Kenya.

Both male and female mice were included in all analyses, and sex ratios were balanced across age groups and experimental sources (Table S1). No significant sex-dependent differences were observed in microbiota alpha or beta diversity or in metabolomic profiles. Therefore, sex was not considered a major covariate in subsequent analyses.

All authors provided data in Excel sheets, including locally recorded data. This metadata was then imported and processed in R (http://www.r-project.org) as a *dataframe* using the *xlsx* package.

Human subjects

To provide a human perspective to the study, 106 stool samples were obtained from healthy Swiss participants (with an average age of 43 for males and 50 for females) attending the Department of Visceral Surgery and Medicine at Inselspital University Hospital for routine check-ups. These individuals showed no signs of gastrointestinal symptoms or suspected functional intestinal problems. No treatment or diagnostic details were gathered, as the participants were healthy individuals. Stool samples from humans and mice were assigned with unique ID numbers, and hence, until sequencing biopsy samples for microbial profiling, there was no metadata information on each sample. Besides, blinding is not relevant to this study, since no experiments involved allocation of samples to a test group.

Ethics statement

The Bern Human Intestinal Community project was approved by the Bern Cantonal Ethics Commission (Ref: KEK-BE: 251/14 and 336/14), and signed informed consent was obtained from all participants. Bern cohort data were anonymized and collected in the EDC (electronic data capture) system (REDCap). The EDC system and associated database are maintained and hosted by the Clinical Trial Unit (CTU) of the University of Bern to ensure high-level data protection and integrity.

METHOD DETAILS

Sample preparation for microbiota profiling in metagenomics and 16S rRNA gene sequencing

3939 samples of frozen fecal or intestinal matter were mixed with glass beads and homogenized. Genomic DNA was extracted using the PowerFecal Pro Kit from Qiagen, following the manufacturer's instructions using sterile water free of DNase and RNase. The quality and quantity of the extracted DNA were checked using a NanoDrop 2000 and a Qubit 3.0 fluorometer (ThermoFisher Scientific). Samples were then stored at -20°C until amplicon PCR or shotgun metagenomic sequencing.[31] Notably, Bahtiyar Yilmaz was responsible for processing and preparing 90% of the samples for sequencing. Pinar Ciftci and Jiaqi Li managed the preparation of the remaining samples.

Amplicon sequencing on IonTorrent PGM™ platform

The amplification of the V5/V6 hypervariable region of the 16S rRNA gene was conducted using the KAPA HiFi HotStart ReadyMix (Roche) DNA polymerase using samples ranging from 200–1000 ng with an expected targeted amplicon length of approximately 350 base pairs. For the polymerase chain reaction, specific primer pairs designed for bacteria were employed. The forward primer included a barcode and the following sequence: 5' CCATCTCATCCCTGCGTGTCTCCGACTCAGC barcode ATTAGATAC CCYGGTAGTCC 3' and the reverse primer contained the sequence: 5' CCTCTCTATGGGCAGTCGGTGATACGAGCTGACGAC ARCCATG-3′.[90]

To conduct PCR, the following conditions were used: i) an initial denaturation step of 5 minutes at 94°C, followed by ii) 35 cycles of denaturation for 1 minute at 94°C, annealing for 20 seconds at 46°C, and extension for 30 seconds at 72°C; iii) the final extension step lasted for 7 minutes at 72°C. PCR products were run on a 1% gel for 2 hours, and amplicons were purified using a Gel Extraction Kit from Qiagen. The presence (positive control) and absence (negative control) of PCR products were verified via gel electrophoresis, but these samples were not submitted for sequencing. The concentration of the amplicons was determined using a Qubit 3.0 Fluorometer (Thermo Fisher Scientific). A library tube was then prepared by pooling 26pM of each sample, and sequencing was carried out using the Ion PGM™ Sequencing 400 Kit and Ion Chip 316™ V2 within the IonPGM™ System (Thermo Fisher Scientific).[72]

In this segment of the study, 32 individual libraries were created for 16S rRNA sequencing. Each batch includes samples from different locations of SPF mice or from different time intervals for samples of wild mice.

Raw sequences were first loaded into the QIIME2 pipeline[73] on the UBELIX Linux cluster of the University of Bern, as described.[91] The trimming and filtering process utilizes Q-scores, employing a strategy where a random subset of reads at each base position is selected for generating a boxplot reflecting the associated Q-scores. The nucleotide quality scores within these raw reads were scrutinized via the *demux summarize* command. We obtained 156,550,955 sequences and 21876 features from 4205 samples with more than 3000 reads per sample (37148 reads on average per sample). Within the DADA2 framework, sequences underwent the removal of primers and chimeras, followed by an evaluation of their quality scores. All the sequences with good quality (Q-score > 30) were used for downstream analysis. Our methodology did not explicitly integrate an external mechanism for the identification and elimination of chimeras within either pipeline. However, by default, DADA2, as embedded in QIIME2, addresses chimeric sequences through a 'consensus' approach.[75] After trimming, quality control, denoising, sequence correction, chimeric sequence filtration, paired-end overlap sequence filtration, and amplicon sequence variant (ASV) generation with 100% identity using DADA2 amplicon sequencing with QIIME2, we generated the table.qza, an ASV count data-related table and rep-seqs.qza, representative sequences of ASVs. Subsequently, the taxonomy was assigned utilizing the pre-trained classifier generated using SILVA 132 database[85] through the application of the *feature-classifier classify-sklearn* command. To establish the composition of taxa and occupancy of each taxon in a sample, as well as the percentage of reads at each taxonomic rank, the taxonomy.qza file was turned into a bar graph using the *taxa barplot* command and was visualized using QIIME 2 view (https://view.qiime2.org) The occupancy of each taxon at level 6 (genus) in each individual sample was examined using the taxonomy.qza and table.qza or filtered-table.qza files using the *taxa collapse* command to generate collapsed-table.qza. Species richness was assessed using alpha diversity analysis, which included the Shannon and Simpson indices. Differences in community composition (β-diversity) were calculated and plotted as Bray-Curtis genus-level community dissimilarity distances. For a Bray-Curtis dissimilarity measure in beta diversity analysis, relative abundance data of each taxon in every sample were used. Differences in beta-diversity between groups were assessed (PERMANOVA) with pairwise comparison (including Benjamini-Hochberg false discovery rate correction) using *pairwiseAdonis* R package.[77,92] To determine if the dispersion of any group was significantly different from the others, we calculated the average distance of the groups using the *betadisper* command from the vegan R package. We implemented M. Anderson's PERMDISP2 for this analysis.[93]

We utilized MaAsLin2 with implemented multiple testing (Benjamini-Hochberg false discovery rate correction; a false discovery rate [FDR], q-value) to identify correlations between metadata and microbial community abundance.[44,74] ASVs that were not present in at least 10% of our samples or with a low abundance (<0.0001% of the total counts) were filtered out from the downstream analysis. The differences between groups were tested at phylum, family, and genus levels *with analysis method = LM* after converting the feature table into relative abundance. Following the general consensus not to rarefy the data, this was omitted from the downstream analysis.[94]

The following commands were the representative command line for MaAsLin2 analysis when the taxonomy data was relative abundance: fit_data <- Maaslin2 (bacteria, metadata, 'AnalysisName, transform = "AST", min_abundance=0.000001, normalization="none", method="LM", max_significance=0.2, min_prevalence=0.1, fixed_effects = c('Strain','BMI', 'Cage', 'Age', 'Sex'), random_effects = c('SampleID', 'Origin'), standardize = FALSE)

After adjusting for the false discovery rate, significance was determined with an adj-p value <0.05. In certain statistical analyses, differences with a *p*-value <0.05 were considered significant. Significant differences are denoted with asterisks: * for adj-p < 0.05, ** for adj-p < 0.01, *** for adj-p < 0.001, and **** for adj-p < 0.0001. Non-significant data is indicated as "n.s." The relevant codes for running MaAsLin2[44] can be found on the webpage of the Group of Curtis Huttenhower (https://huttenhower.sph.harvard.edu). Plots were generated using *ggplot2* with either a *phyloseq* object or GraphPad Prism v10.

Library preparation, sequencing for metagenomics and data analysis

6 individual libraries containing 246 samples were prepared for shotgun metagenomic sequencing. Each library includes samples from different locations of SPF mice or different sample types from different groups of mice. Bacterial DNA libraries were prepared according to the TruSeq DNA PCR-Free (for fecal samples) and Nextera DNA Library Prep (for some cecal samples due to the low

DNA amount) and sequenced on NovaSeq 6000 (Illumina) in paired-end mode to produce 2 x 150bp reads. The total number of reads generated across all 246 metagenomic samples was approximately 9.18 billion, reflecting deep sequencing coverage. On average, each sample yielded 37.3 million ± 22.9 million reads, with individual sample read counts ranging from 9.1 million to 109.1 million reads. This sequencing depth provides a robust basis for comprehensive taxonomic and functional profiling, ensuring adequate coverage even in samples with lower biomass or reduced microbial complexity.

The *metaWRAP-Read_qc* module was employed to filter out human genome-contaminated reads by aligning them to the hg38 (UCSC) human reference genome.[76] Additionally, this step removed adaptor sequences and low-quality reads, generating quality reports for each sequenced sample before proceeding with microbial abundance estimation.[76]

Taxonomic evaluations were conducted using the Kraken2 pipeline, employing a customized RefSeq database in accordance with the developer's instructions.[83,95] To obtain the necessary taxonomic information for our analysis, we employed the *kraken2-build tool with the "–download-taxonomy"* option. This allowed us to download and store the following components in the CustomDB taxonomy folder: i) Accession number to taxon maps, which establish the relationship between accession numbers and corresponding taxa; ii) Taxonomic names, providing the specific names associated with each taxon; iii) Tree information, sourced from NCBI, encompassing complete genomes of bacteria. Subsequently, we utilized the kraken2-build command to construct a custom database. To enhance the accuracy of read counts at lower taxonomic levels and perform re-estimation at the species, genus, and family levels, we employed Bracken (Bayesian Re-estimation of Abundance with Kraken).[84] This was achieved by utilizing the output from Kraken2. The default command line for Bracken was implemented as follows: https://genomics.sschmeier.com/ngs-taxonomic-investigation/index.html#bracken. Across all samples, an average of approximately 12.7 million ± 14.6 million reads per sample were assigned to bacteria, which constituted the majority of classified reads. To ensure sufficient taxonomic resolution, we applied a post-classification threshold of 1.5 million bacterial reads per sample, below which samples were excluded from downstream analyses. Additional reads were assigned to Eukaryota (mean: 429,080 ± 504,868 reads), Archaea (mean: 4,335 ± 56,299 reads), and viruses (mean: 18,462 ± 80,487 reads). These thresholds and filtering steps ensured that only high-quality, taxonomically informative samples were retained for downstream analyses.

Sequences associated with the functional genes were assigned to taxonomic groups using HUMAnN v3.0 (http://huttenhower.sph.harvard.edu/humann) to determine the functional profiles in samples. This pipeline was employed with default parameters using the MetaCyc database, which provides pathway definitions by a gene family.[43] Output data shows that there is a negligible difference in the read assignment rates in pathway coverage between SPF and wild mouse samples. Specifically, the mean percentage of reads aligned for SPF mice was 50.46% (±11.42%), and for wild mice, it was 50.65% (±10.90%). This similarity underscores the consistent performance of HUMAnN in processing mouse samples irrespective of their microbiological status. However, human samples demonstrated a markedly better mean alignment rate of 23.50% (±4.49%). This divergence underscores the assertion that HUMAnN exhibits enhanced precision in assigning reads for human samples as opposed to mouse samples.

High-quality genome assemblies of complex microbial populations from shotgun metagenomic data were generated using metaWRAP, a modular Unix-based pipeline (https://github.com/bxlab/metaWRAP).[76] Briefly, the read_qc module was used to pre-process raw Illumina sequencing reads in preparation for assembly and alignment. Then, the assembly module enabled assembly of metagenomic reads using either metaSPAdes[87] or MegaHit (default).[86] The binning module is a convenient wrapper around three metagenomic binning software: MetaBAT,[19] MaxBin,[20] and CONCOCT[21] and then integrated and refined to eliminate duplicates and enhance the accuracy of the assembled genomes.[22] Subsequently, the *bin_refinement* module utilizing a hybrid approach to produce a consolidated, improved bin set, the *reassemble_bins* module used to improve the set of bins. Lastly, the species-level phylogenies and classified species-level genome bins (SGBs) of metagenome-assembled genomes (MAGs) were constructed using the *PhyloPhlAn* pipeline.[24]

The inStrain pipeline

This pipeline was used to compare the genomes of different populations and analyze gene coverage and breadth. It allowed the identification of SNPs and their location within genes, as well as the distinction between synonymous and nonsynonymous mutations.

We ran the inStrain pipeline with the publicly available Unified Human Gastrointestinal Genome (UHGG) collection as the reference database.[67] The UHGG catalog version 1.0 was obtained from the European Bioinformatics Institute (EBI) (http://ftp.ebi.ac.uk/pub/databases/metagenomics/mgnify_genomes/human-gut/v1.0/). Reference genomes were annotated using Prokka[82] and annotated genes were further classified into protein families (PFAM),[68] CAZymes,[70] and KEGG orthologs (KO)[69] according to the user manual in the inStrain documentation (https://instrain.readthedocs.io/en/latest/user_manual.html). Antibiotic resistance genes were identified using the Comprehensive Antibiotic Resistance Database (CARD).[96] Additionally, we calculated nucleotide diversity and the dN/dS ratio.[32] The genome annotations were merged with the inStrain output using a custom R script.

We calculated the base-pair coverage and number of covered genome bases for each sample's genome using sequencing data obtained from the *metaWRAP-Read_qc* module. To obtain the SNPs and dN/dS profile for each sample, we executed the following commands: https://instrain.readthedocs.io).[32] We calculated the base-pair coverage and number of covered genome bases for each sample's genome using sequencing data obtained from the *metaWRAP-Read_qc* module. To obtain the SNPs and dN/dS profile for each sample, we executed the following commands:

- **inStrain profile** SampleX.sam UHGGv1/UHGG_reps.fasta -p 24 -g UHGGv1/UHGG_reps.genes.fna -o Vivarium/Instrain/SampleX.IS -p 24 -g UHGGv1/UHGG_reps.genes.fna -s UHGGv1/UHGG_reps.stb –database_mode.

- **inStrain** compare -i Instrain/SampleX.IS -s UHGGv1/UHGG_reps.stb -p 24 -o Vivarium/Instrain/IS.COMPARE –database_mode

For further analysis, genomes were required to have a coverage depth of at least 5 times the cumulative average across all samples. In addition, a species was required to have a coverage breadth of at least 50% of the genome in one sample to be considered present in the cohort. We only included SNPs above zero in more than 80% of the samples. To calculate the nucleotide diversity (π) of each position, the following formula was used: $\pi = 1 - [$(number of 'A' bases/total bases)2 + (number of 'C' bases/total bases)2 + (number of 'T' bases/total bases)2 + (number of 'G' bases/total bases)2]. This calculation was implemented in the inStrain pipeline, and the resulting values were used to determine the average nucleotide diversity for each genome.

For the strain comparisons between samples, a distance matrix was then created for each subspecies based on population average nucleotide diversity (popANI) values: this matrix was used to cluster strains into a number of individual strains using '*average*' hierarchical clustering with a threshold of 99.999% ANI with the *scipy* cluster package. Strains with PopANI > 99.999% were called identical '*shared*' strains. Additionally, the 'pairwise approach' identified strains that shared >99.999% popANI across consecutive sampling. We chose the cutoff of 99.999% popANI for persistence following the calculation performed before describing the expected rate of *in situ* bacterial evolution in human gut (~0.9 single-nucleotide polymorphisms (SNPs)/genome/year).[97]

Single nucleotide substitutions and variants were visualized using circular plots generated with the *circlize* R package.[98] We considered only single nucleotide polymorphisms (SNPs) with a position coverage exceeding 10 and a higher frequency in the consensus base compared to the reference base. Additionally, only genes with more than 20 SNPs observed across all samples were included in the visualizations.

Heatmaps of dN/dS and pN/pS ratios were produced with pheatmap package in R to visualize genes with potential positive selection. For visualization of dN/dS and pN/pS only SNPs passing a coverage threshold of at least 5 and a breadth of more than 80% were selected. Moreover, as for SNP visualizations, only genes with more than 20 SNPs observed across all samples were included in the dN/dS and pN/pS visualizations.

Prediction of vivarium-specific 3D protein structures

For selected genes of interest, the reference amino acid sequences were retrieved from the UHGG catalog version 1.0, available from the European Bioinformatics Institute (EBI) (http://ftp.ebi.ac.uk/pub/databases/metagenomics/mgnify_genomes/human-gut/v1.0/). Vivarium-specific amino acid sequences were generated using a custom R script, which replaced amino acids based on detected nonsynonymous substitutions. If two substitutions occurred within the same codon, the resulting new amino acid was calculated and incorporated into the vivarium-specific sequence accordingly. The amino acid sequences were submitted to the AlphaFold Server, powered by AlphaFold3[78] (https://alphafoldserver.com) to generate predicted 3D protein structures. For further visualization of the 3D protein structures, the web-based open-source tool Mol*viewer[79] was used (https://molstar.org/viewer/).

Prediction of protein-ligand binding and multimerization

Predicted monomer 3D protein structures were submitted to the COACH tool available under (http://zhanglab.ccmb.med.umich.edu/COACH/) for prediction of potential ligand binding sites.[80,99] COACH predicts ligand binding sites in proteins using a consensus approach that combines the outputs of five established programs, integrating protein substructure analysis with sequence profile alignments of known ligand binding sites.

To predict interactions between the chains of a protein multimer, the multimer structures were analyzed using the PLIP tool with the advanced protein-protein interactions option (https://plip-tool.biotec.tu-dresden.de/plip-web/plip/index).[81] By indicating the monomer chains in the multimer, PLIP identifies several types of inter-chain interactions, including hydrophobic interactions, hydrogen bonds, salt bridges, and more. We have applied the default thresholds for interaction identification, as recommended by PLIP for optimal accuracy.

Metabolite extraction from cecum samples

Conventionally colonized SPF mice on a C57BL/6J background from 13 vivaria, three to six samples per vivarium, 10-18 weeks of age, male, and confirmed to be pathogen-free, were used. All mice received a standard lab chow diet and water *ad libitum*. Mice were sacrificed by cervical dislocation. Samples from the luminal contents of the cecum were collected as described previously.[100] Samples were snap-frozen in liquid nitrogen and stored at -80 °C until extraction. Metabolites were extracted from the samples in 1 ml of 80°C Millipore water. Samples were vortexed for 10 sec, incubated for 3 min at 80°C in an Eppendorf Thermomixer at maximal speed (1500 rpm), vortexed for 10 sec, and centrifuged for 3 min at 20000*g* at room temperature. From the supernatant, 150 µl were transferred to 96 well microtiter plates and stored at -80°C until analysis. For mass-spectrometry analysis, the samples were diluted 1:1 in the mobile phase (49.6:49.6:0.8 v/v/v methanol:water:acetic acid) containing a retention time reference compound (9-anthracene carboxylic acid, Sigma Aldrich) and stored at -20°C until analysis.

Metabolite extraction from diet pellets

Per vivarium, one mouse food pellet was weighed, dissolved in 20 ml of Millipore water and vortexed for 10 seconds. After incubation overnight, an additional 10 ml were added and the samples vortexed for 10 seconds. From this, 1 ml of food pellet suspension was heated to 80°C and metabolites were extracted analogous to the cecum sample metabolite extraction (3 min incubation at 80°C in an

Eppendorf Thermomixer at maximal speed, vortexed for 10 sec, centrifuged for 3 min at 20000*g* at room temperature). From the supernatant, 150 µl were transferred to 96 well microtiter plates and stored at -80°C until analysis. Prior to measurements, samples were 1:20 diluted in MilliQ water.

Chemicals

HPLC-grade methanol and isopropanol, all chemicals, buffer additives for online mass referencing and sample preparation chemicals were purchased from Sigma-Aldrich, Agilent Technologies and Cayman Chemicals. HPLC grade water was obtained from an IQ7000 MilliQ water purification system equipped with an LC-Pak (Merck).

Metabolite profiling using LC-TOF-MS

The cecum and diet samples were analyzed on an Agilent 1290 Infinity LC system coupled to an Agilent 6550 accurate-mass Quadrupole-TOF Mass-Spectrometer using Dual Agilent Jet Stream Electrospray Ionization (Agilent Technologies, Santa Clara, USA). The injection volume was 4 µl and chromatographic separation was achieved using a Zorbax SB-Aq 1.8 µM 2.1 x 50 mm column with a Zorbax-SB-C8 guard column with rapid resolution cartridge (2.1 x 30 mm 3.5 um). Elution was achieved using a linear gradient at a flow rate of 0.6 ml min^{-1} starting with 2% mobile phase B (0.2% (v/v) acetic acid in methanol) over 13 minutes, gradually changing to 98% eluent B. This was followed by 1.5 min isocratic flow of 2% mobile phase A (0.2% (v/v) acetic acid in water) and 1-minute equilibration at 2% eluent B. Data acquisition was carried out using electrospray ionization in positive and negative mode using full scan analysis over a 50-1700 m/z mass range in 2 GHz extended dynamic range acquisition mode. For online mass axis correction, purine and hexakis (1H, 1H, 3H-tetrafluoropropoxy)phosphazine (HP-0921, Agilent Technologies) were added to the mobile phase. Electrospray settings were as follows: ion spray voltage, 3.5 kV negative mode and 4 kV positive mode; capillary temperature, 325°C; drying gas flow, 10 L/min, and 45 psi nebulizer pressure.

Preparation of standards for calibration and quantification

Stock solutions of 138 purified single compounds obtained from Sigma-Aldrich (Merck KGaA) were prepared by dissolving them separately in either MilliQ water, ethanol- or methanol-water mixtures, or DMSO and subsequently mixed to 600 µM. The mixture was aliquoted and dried at 0.12 mbar to complete dryness in a SpeedVac setup (Christ, Osterode am Harz, Germany) and stored at -80 °C. Working solutions were prepared in MilliQ water with 0.4 % acetic acid (v/v) and filtered before use. Calibration curves were obtained from 24-point serial dilution series spanning seven orders of magnitude (from 17.9 pM to 150 µM) in MilliQ water and, to assess matrix effects, in pooled study sample background. The diluted matrix was spiked with the serially diluted standard mixture. Standards were measured at the beginning and the end of the measurement in positive and negative ionization mode.

Linear regression of log-transformed ion counts versus log-transformed concentrations was performed on measurements from standard dilution series, allowing for the iterative removal of up to six dilution steps from the upper concentration limit and 12 data points from the lower dilutions, with the goal of maximizing the R^2 value. Lower limit of quantification and upper limit of linearity were determined as the lowest and highest accepted dilution steps, respectively. Metabolite concentrations were then calculated based on the derived slopes and intercepts. Only values within the linear range were considered for further analyses. Of 138 metabolites, 111 could be reliably detected and quantified based on these criteria. For each unique metabolite, we measured up to 5 different modifications (deprotonated anion, protonated cation, sodium adduct cation, dimer cation and dimer sodium adduct cation) and manually selected the modification based on linear fit (highest R^2 value), linear range and the reliable coverage within the data set (maximal number of detected samples). Metabolite concentrations are reported as $\frac{nmol}{mg\ sample}$.

Data processing, analysis, and visualization of metabolomic data

For LC-TOF-MS data, pre-processing, peak picking, and annotation were performed using the MassHunter Quant software (Agilent Technologies, Santa Clara, USA). Metabolites were classified according to MetaCyc.[60] Further data analysis, statistics and visualization were performed after raw data export in MATLAB R2021b (MathWorks, MA, USA) using functions embedded in the Bioinformatics and Statistics toolboxes. Graphs, illustrations, and plots were finalized using Illustrator (Adobe Inc.). Metabolic diversity was calculated using normalized counts on Shannon diversity index.

GSMM mapping

ASVs identified from the *Phyloseq* pipeline were used as input for the mapping. In order to be able to fit the models described in the next section, we had to select samples for which all desired annotations were present. We kept 3804 out of 3917 available samples, spanning 51 vivaria but only one wild mice colony, and 105 human samples. We selected only ASVs that were annotated at the family and/or genus and/or species level and considered ASVs with identical annotations to be identical. 457 different ASVs were recovered in this manner. Using the taxonomy annotations, we mapped the ASVs to a recent collection of reconstructed GSMMs from the human microbiome[101] and normalized them into relative abundances. The total mapped abundance (sum of relative abundances of ASVs that could be mapped to a model) was satisfactory, with a median larger than 75% in all models. Note that some individual samples have very low total mapped abundance but were still included in the analysis. ASVs that could be mapped to several models (e.g., ASVs that were identified at the family or genus level only) were mapped to all the corresponding models (see Normalized reaction abundance).

Normalized reaction abundance and statistical model

Using the GSMM mapping, we transformed the relative ASV abundances into normalized reaction abundances as described in a previous study.[72] The normalized reaction abundance of reaction *i* in a microbiome sample was described as:

$$a_r(i) = \sum_{k=1}^{N} a_{asv}(k) \cdot E(i,k)$$

with $a_{asv}(k)$ the relative abundance of ASV k in the microbiome sample, N the total number of ASVs mapping to at least one GSMM, and $E(i,k)$ the expected probability of reaction i in ASV k, which is computed from the different GSMMs mapping to ASV k. From the GSMM mapping, we gathered 7604 different metabolic reactions (excluding exchange and biomass reactions). The mean reaction abundance over all samples and reactions was $1.04 \cdot 10^{-1}$ and the median reaction abundance was $5.73 \cdot 10^{-3}$.

For each of the identified metabolic reactions, we fitted a reaction-specific linear mixed effect model or linear model to express the reaction abundance as a function of the species (mouse or human), the microbiome model (Wild, SPF, Human), the sex (male/female), the age group (adult/pup, for mice only), the origin of the sample, the experimental chip, and the SPF mouse strain. We excluded samples with unknown sex from the analysis, which resulted in all used samples being of the same sample type (feces). The model syntax was expressed as:

$$a_r(i) = 1 + (1|Species:Origin) + 1(Species:Strain) + (1|Chip) + Species + Species:Model + Species:Sex + Species:Age$$

where $(1|X)$ denotes that X is a random effect, and $A:B$ denotes that effect B is nested into effect A. $1+$ denotes the presence of an intercept.

The code for fitting the reaction-specific models was written in R. For each reaction-specific model, the significance of each random effect was tested. Random effects were only included in the model if deemed significant by a likelihood ratio test (p-value ≤ 0.05) using the *step* function from the lmerTest[102] package. Linear mixed effect models were fit using the lme4[103] packages, and fixed effect models using the *lm* function from the stats package.[103] P-values for the fixed effect estimates of the mixed effect models were obtained using the lmerTest[102] package and the Kenward-Roger approximation of the degrees of freedom as implemented in the pbkrTest package.[104] To compare the effects of the microbiome model on the reaction abundances, we used the emmeans[105] package to compute the estimated marginal means of the effect.

For the downstream analysis, we gathered for each reaction-specific model the estimates for the difference in mean reaction abundance due to sex (species-specific), the difference in mean reaction abundance due to age (species-specific), and the estimated marginal means for the microbiome models, as well as the associated p-values. We then corrected the p-values for multiple testing using the Holm-Bonferroni correction. Each difference in mean with adj-p-value ≤ 0.05 was considered significant. We used all mapped reactions for this study, independent of how low the total abundance of this reaction across all samples could be. However, as reaction models are fitted separately, it is possible to only select reactions with the desired minimal abundance for the downstream analysis.

Subsystem annotation and identification of significantly different subsystems

We developed a pipeline to map each reaction to metabolic subsystems from the KEGG database using the different available reaction identifiers.[46–48] We used the third level of the KEGG subsystem annotation and counted 134 different metabolic subsystems. Of the 7604 reactions analyzed with mixed models, 3146 could be mapped to an identified subsystem. Reactions that could not be mapped to an identified subsystem were designated as an 'Unknown' subsystem and were still considered in the enrichment analysis.

Once the reactions were annotated with metabolic subsystems, we used Fisher's exact test to test for each effect (sex, age, microbiome model) and each subsystem (except the 'Unknown' subsystem) whether this subsystem was enriched in significant estimates for this effect. The subsystems with a p-value ≤ 0.05 were considered significantly different between the two groups tested. No p-value correction was applied at this step. Plots shown in the results section of this article show the estimated value of the significant estimates in each subsystem identified as significantly different for the considered effect.

Statistical model for metabolome analysis

For each of the 112 profiled metabolites, we estimated a metabolite-specific linear mixed effect model to express the abundance of metabolite i in a microbiome sample, $m(i)$, as a function of the abundances of microbial species j, $s(j)$, the concentrations of metabolites k in the diet (identical per vivarium), $d(k)$, the microbiome model (Wild, SPF), the origin, and the sex. The model was defined in the same syntax as the reaction-specific linear mixed-effects models as:

$$m(i) = 1 + Model + (1|Sex) + \sum_{j \in S} s(j) + \sum_{k \in D} d(k),$$

where S (D) denotes the index set of microbial species (diet metabolites) included in a given model. We iteratively selected the best model per metabolite in terms of included covariates according to the Bayesian information criterion (BIC). Each iteration involved the following sequence of possible model modifications until convergence to a (local) minimum: (i) augmenting S or D by the covariate that most improved the BIC after fitting models with each of the potential covariates not yet in S/D included individually; (ii) when none

of the augmented models from (i) improved the BIC, removing the covariate with the highest p-value for its coefficient whose removal individually improved the BIC; and (iii) when none of the modifications (ii) improved the BIC, removing the covariate that most improved the BIC upon removal, irrespective of its coefficient's significance. To assess the relative contributions of microbiome and diet, we estimated three types of models: (i) only including diet covariates; (ii) only including microbiome covariates; and (iii) including both types of covariates. Because the number of species (71) and diet metabolites (124) exceeded the number of animals with available luminal and diet metabolite data (47 mice in 10 vivaria), per model class, we initialized estimations in three different ways: (i) one estimation with the null model (only intercept); (ii) 100 estimations with randomly selected covariates in proportion to available diet / microbiome covariates, where the total number of covariates in the initial model was uniformly sampled between zero and the maximum possible number (42); (iii) 100 estimations with a full model of 42 randomly selected covariates. For the random strategies, the number of covariates was reduced when no full-rank fixed effects matrices could be sampled. Per metabolite, we selected the model with the lowest BIC among the 603 models as the final model. After model selection, we additionally assessed the significance of covariate contributions to luminal metabolite concentrations by the unadjusted p-values of the covariates' estimated coefficients ($P < 0.05$). To estimate corresponding effect sizes, for each covariate j, covariate k, and metabolite i, we fitted a reduced model without this covariate and computed the difference in R^2 between full and reduced model. All computations were performed in Matlab R2024a, with a total run time of ~4000 CPU hours on a cluster with 256 cores used.

Succinate metabolism – analysis of metabolic potential in selected bacterial species

We investigated in MetaCyc for pathways involved in succinate production and consumption and created a *MetaCyc SmartTable* for all associated enzymes and their corresponding gene names.[106] The *SmartTable* was exported and compared to the reference genomes of interest. Specifically, gene names and associated products from the reference genomes were matched to those in the *SmartTable* using string similarity thresholds of 91% for gene product names and 85% for gene names to ensure high detection sensitivity. Ambiguous but highly similar matches were manually excluded. We then used custom R Markdown scripts to generate heatmaps comparing gene presence across bacterial species, as well as HTML outputs listing succinate-related genes and their corresponding gene products by species. Last, a subset of genes with known gene products present in at least one bacterial species of interest, was used to generate a pathway collage in MetaCyc, illustrating the potential interplay between succinate-related pathways.

To investigate capabilities using genome-scale metabolic models, we mapped bacteria to models of the AGORA2 collection at the species level.[45] For each model, we blocked oxygen uptake to simulate an anaerobic environment and constrained the biomass reaction to be larger than an arbitrary threshold (0.1/h) to ensure growth. All other exchanges were left unconstrained. We then ran parsimonious FBA,[107] either maximizing (to check for succinate secretion) or minimizing (to check for succinate uptake) the succinate exchange flux. Uptake or secretion was validated when the absolute value of the succinate exchange flux was above 0.01 au. In both cases, we constrained the succinate exchange reaction to operate in the chosen direction only. We then averaged the number of hits over all models mapped to a given species to obtain probabilities of succinate uptake and secretion.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical methods were not employed to predetermine the sample size. Given the impracticality of blinding in the experiment, samples were not randomized, and investigators were not blinded to the conditions. However, human and mouse stool samples were anonymized using unique IDs, and the associated metadata remained concealed until after microbial profiling through sequencing. While the data distribution was assumed to follow a normal distribution, this assumption was not formally tested.

Unless otherwise stated in the individual method sections above, all statistical analyses were performed using R version 3.6.1 (2019-07-05) or Prism 10 (GraphPad Software, San Diego, CA). The performed tests were introduced along the manuscript. Briefly, the differences between the two groups were evaluated using Mann–Whitney U test (non-parametric) and more than 3 groups were tested using one-way ANOVA (parametric), followed by Tukey's honest significant difference test or two-stage step-up method of Benjamini, Krieger, and Yekutieli, as a post hoc test. Microbial changes were tested using multivariate analysis by linear models (MaAsLin2) R package.[44,74] The beta diversity with pairwise comparison (Benjamini-Hochberg false discovery rate correction) using *pairwiseAdonis* R package to confirm the strength and statistical significance of groups in the same distance metrics with the *phyloseq* pipeline in R.[77,92] Statistically significant differences are shown with asterisks as follows: *; adj-p < 0.05, **; adj-p < 0.01, ***; adj-p < 0.001 and ****; adj-p < 0.0001. Notable near-significant differences (0.05<p and 0.05< adj-p-value <0.2) were also highlighted in some of the figures and following asterisks were also used: *; p < 0.05, **; p < 0.01, and ***; adj-p < 0.001.