



The CIAMIB: a Large and Metabolically Diverse Collection of Inflammation-Associated Bacteria from the Murine Gut

Erin Oi-Yan Wong,^a Emma Joan Elizabeth Brownlie,^a Katharine Michelle Ng,^{b,c} Sornnujah Kathirgamanathan,^a  Feiqiao Brian Yu,^d  Bryan D. Merrill,^e  Kerwyn Casey Huang,^{d,e,f} Alberto Martin,^g  Carolina Tropini,^{b,c,h}  William Wiley Navarre^a

^aDepartment of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada

^bSchool of Biomedical Engineering, University of British Columbia, Vancouver, Canada

^cDepartment of Microbiology and Immunology, University of British Columbia, Vancouver, Canada

^dChan Zuckerberg Biohub, San Francisco, California, USA

^eDepartment of Microbiology and Immunology, Stanford University School of Medicine, Stanford, California, USA

^fDepartment of Bioengineering, Stanford University, Stanford, California, USA

^gDepartment of Immunology, University of Toronto, Toronto, Ontario, Canada

^hHumans and the Microbiome Program, Canadian Institute for Advanced Research, Toronto, Canada

ABSTRACT Gut inflammation directly impacts the growth and stability of commensal gut microbes and can lead to long-lasting changes in microbiota composition that can prolong or exacerbate disease states. While mouse models are used extensively to investigate the interplay between microbes and the inflamed state, the paucity of cultured mouse gut microbes has hindered efforts to determine causal relationships. To address this issue, we are assembling the Collection of Inflammation-Associated Mouse Intestinal Bacteria (CIAMIB). The initial release of this collection comprises 41 isolates of 39 unique bacterial species, covering 4 phyla and containing 10 previously uncultivated isolates, including 1 novel family and 7 novel genera. The collection significantly expands the number of available Muribaculaceae, Lachnospiraceae, and Coriobacteriaceae isolates and includes microbes from genera associated with inflammation, such as *Prevotella* and *Klebsiella*. We characterized the growth of CIAMIB isolates across a diverse range of nutritional conditions and predicted their metabolic potential and anaerobic fermentation capacity based on the genomes of these isolates. We also provide the first metabolic analysis of species within the genus *Adlercreutzia*, revealing these representatives to be nitrate-reducing and severely restricted in their ability to grow on carbohydrates. CIAMIB isolates are fully sequenced and available to the scientific community as a powerful tool to study host-microbiota interactions.

IMPORTANCE Attempts to explore the role of the microbiota in animal physiology have resulted in large-scale efforts to cultivate the thousands of microbes that are associated with humans. In contrast, relatively few lab mouse-associated bacteria have been isolated, despite the fact that the overwhelming number of studies on the microbiota use laboratory mice that are colonized with microbes that are quite distinct from those in humans. Here, we report the results of a large-scale isolation of bacteria from the intestines of laboratory mice either prone to or suffering from gut inflammation. This collection comprises dozens of novel isolates, many of which represent the only cultured representatives of their genus or species. We report their basic growth characteristics and genomes and are making them widely available to the greater research community.

KEYWORDS *Adlercreutzia*, *Atopobiaceae*, Lachnospiraceae, Muribaculaceae, *Prevotella*, gut inflammation, microbial growth, microbiota

While the vast majority of studies aiming to experimentally dissect interactions between microbes and their mammalian hosts rely on mouse models, the large majority of cultivated gut microbes are derived from humans (1–3). Large-scale culture

Invited Editor Lora V. Hooper, University of Hawaii at Manoa

Editor Margaret J. McFall-Ngai, University of Hawaii at Manoa

Copyright © 2022 Wong et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to William Wiley Navarre, william.navarre@utoronto.ca.

The authors declare no conflict of interest.

Received 1 October 2021

Accepted 9 February 2022

Published 10 March 2022

collections of human gut microbes have advanced our understanding of microbial growth requirements and facilitated investigations of the ways specific microbes influence host physiology (4, 5). While human and lab mice microbiotas are both dominated by species from the phyla Bacteroidota and Bacillota, they differ considerably at lower taxonomic levels and vary widely in the relative abundance of many species (6–8). Hence, the degree to which the mouse and human microbiota are functionally similar is largely unknown. The few studies that have directly addressed this question found that many isolates of human origin do not stably colonize the mouse gut, are rapidly outcompeted by mouse-derived isolates, and differ in their effects on intestinal immune development (9, 10). These differences confound efforts to align findings from lab mice with conventional (undefined) microbiotas to those from gnotobiotic mice colonized with human fecal samples or synthetic communities. Furthermore, widespread but poorly understood and uncontrolled variability across the microbiotas of lab mice from different facilities has been a major source of irreproducibility in mouse models (11, 12). A comprehensive collection of murine-derived microbial isolates will greatly improve the quality and reproducibility of gut microbiota studies conducted in mice, enable a more complete understanding of how human and mouse microbes are and are not comparable, and also allow the field to make sense of decades of prior research.

16S rRNA gene and shotgun metagenomic sequencing efforts have played a key role in characterizing the gut microbiota of lab mice, but the paucity of cultured species limits our understanding of their functions. A substantial proportion of short 16S rRNA gene sequences and metagenome-assembled genomes (MAGs) cannot be taxonomically assigned at the species level using currently available pipelines and databases, complicating both identification and subsequent targeted infection experiments (13–16). Even within a given species, individual strains may not fill the same ecological niche or behave in a functionally equivalent manner. For example, different strains of *Bacteroides ovatus* promote the production of various levels of immunoglobulin A (17).

To date, there have been two large-scale cultivations of microbes from laboratory mice: the Mouse Intestinal Bacterial Collection (miBC) and the Mouse Gut Microbial Biobank (mGMB). The miBC comprises 76 species isolated from mice from various genetic backgrounds, providers, and laboratories (18). The mGMB comprises 126 bacterial species obtained from *Lep^{ob/ob}* mice, a model for understanding metabolic diseases (19). Together, the miBC and the mGMB account for only ~44% of total reads identified in the amplicon data set generated by the miBC study at the species level, indicating that considerable work remains to be done to more completely capture the set of microbes that inhabit the murine gut (19).

A major focus of study relates to the roles various gut microbes play in initiating and prolonging the course of intestinal inflammation (20). Inflammation dramatically changes the intestinal environment via host-produced antimicrobial peptides, infiltration of the intestinal lumen by immune cells, iron limitation, and an increase in oxidative stress through immune-mediated production of peroxides, superoxide, and nitric oxide (21–24). Gut microbiota composition during inflammation often shifts away from strict anaerobes and toward oxygen-tolerant microbes like Pseudomonadota that tolerate or even thrive in the inflamed intestinal milieu (25–28). Some gut microbes can trigger inflammation (pathogens), exacerbate or prolong inflammation (pathobionts), persist and ultimately promote recovery and repair of host tissues, or merely act as resilient but neutral bystanders that presumably evolved to withstand the transient but frequent storms of inflammation occurring during the life of their animal hosts.

Here, we report the results of our efforts to cultivate novel bacterial species and strains that thrive or survive in the inflamed mouse intestine, the Collection of Inflammation-Associated Mouse Intestinal Bacteria (CIAMIB). We exploited a previous observation that different cages in our animal facility of interleukin-10-deficient (*IL10^{-/-}*) mice and mice deficient in the adenomatous polyposis coli (*Apc*) and MutS homolog 2 (*Msh2*) genes (*Apc^{Min/+} Msh2^{-/-}*) displayed different degrees of pathology dependent on

the composition of their gut microbiota (29–31). These mouse models are commonly used to investigate the relationship between the gut microbiota and intestinal inflammation, a precursor to diseases such as inflammatory bowel disease (IBD) and colorectal cancer (CRC) (32–34). Cohousing experiments demonstrated that these differences were due almost entirely to gut microbiota composition, as the transfer of microbes from affected mice promoted the development of colitis or polyps in cages of previously healthy isogenic mice (29, 30).

The first release from the CIAMIB reported here comprises 39 species and 41 strains, 10 of which represent novel bacterial species that have never previously been isolated or cultured to our knowledge. We sequenced the genomes of all 41 strains and characterized the growth of each species under a wide range of conditions, leveraging genomic information to examine trends in metabolic profiles across bacterial families and for particular species. The CIAMIB will provide a powerful resource to experimentally dissect the role that specific microbes play in health and disease, particularly inflammation.

RESULTS

Establishment of a collection of bacteria from the inflamed mouse gut. The primary goal of this study was to generate a collection of inflammation-associated bacteria from the mouse gut as a resource for the scientific community. At this stage, we did not seek to isolate specific causal microbes or define whether our cultivated microbes are pathogens, pathobionts, or inflammation-resistant bystanders. To this end, cecal and colonic contents were collected from seven mice that either had developed spontaneous colitis (*IL10*^{-/-} mice; four mice total) or polyps (*Apc*^{Min/+} *Msh2*^{-/-}, two mice total) or were *Apc*^{Min/+} *Msh2*^{+/-} mice cohoused with *Apc*^{Min/+} *Msh2*^{-/-} mice with polyps (one mouse total). Colonic and cecal contents were plated on 18 types of agar media and incubated under anaerobic, microaerobic (5% O₂), and aerobic conditions for 3 to 7 days. Both selective and nonselective media were used to try to recover slow-growing, rare, and fastidious bacteria (Fig. S1). Each isolate was restreaked onto solid media twice, with single colonies picked each time, to increase the likelihood that the isolates were pure and clonal.

The identities of 608 isolates were determined by 16S rRNA gene sequencing. Using a sequence identity cutoff of 97% to define species, we identified 39 distinct bacterial species from 18 families, spanning the four major phyla found in human and the mouse gut microbiotas: Bacillota (formerly Firmicutes) (13 species), Bacteroidota (formerly Bacteroidetes) (11 species), Actinomycetota (formerly Actinobacteria) (9 species), and Pseudomonadota (formerly Proteobacteria) (6 species) (Fig. 1). The genomes of the unique isolates were sequenced, and individual draft genomes were generated. Detailed information about each isolate, including isolation conditions, 16S rRNA gene sequences, accession numbers, and assembly statistics, is provided in Table S1.

The 16S rRNA of 10 species from the CIAMIB had <97% identity to any species cultivated to date. *In silico* DNA-DNA hybridization (DDH) and average nucleotide identity (ANI) comparisons were carried out between the whole-genome sequences of novel taxa and their closest neighboring taxa identified from 16S rRNA gene sequences using the Genome-to-Genome Distance Calculator 2.1 and OrthoANLu calculator (35, 36). The genomes of these 10 isolates were below the species-level DDH threshold of 70% and below the 95% ANI species cutoff, establishing them as novel bacterial species belonging to seven novel genera and one novel family (Table S2) (37). To determine novelty at different taxonomic ranks, we performed a classification analysis using the TypeMat tool available through the Microbial Genomes Atlas (MiGA) web server. This tool estimates the probability of novelty at each rank based on ANI and average amino acid identity (AAI) (Table S2) (38). These novel species have been deposited at the DSMZ to ensure long-term availability (Table S2) and are currently identified by the lowest taxonomic groups to which they can confidently be assigned (e.g., *Bacteroides* [new species]).

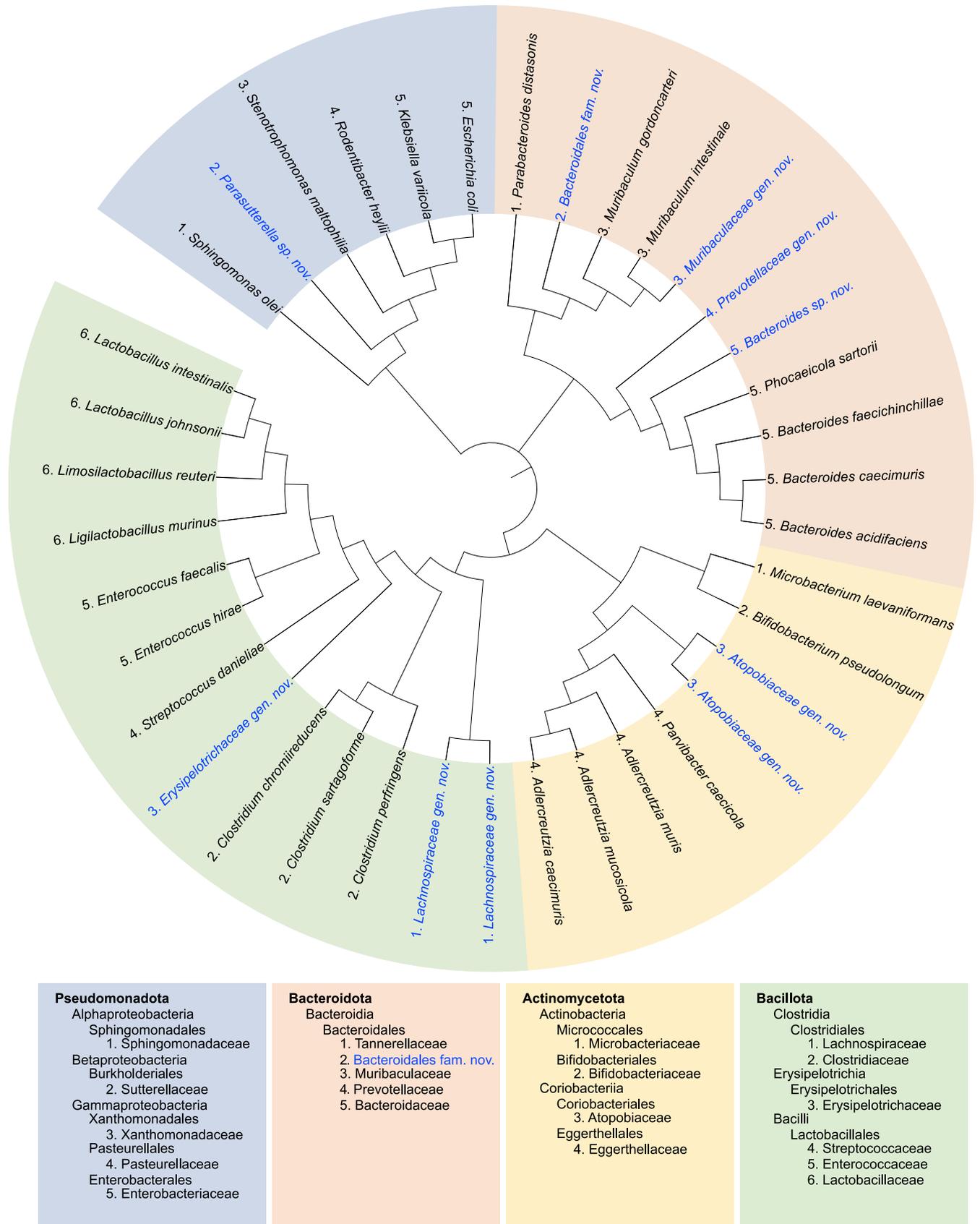


FIG 1 The CIAMIB is phylogenetically diverse. The cladogram illustrates the taxonomic distribution of all 39 species in the CIAMIB based on 16S rRNA gene sequences. Color coding is by phylum, and taxonomy is indicated numerically at the family level. Novel taxa from this study with their candidate names are in blue text.

Determining novelty of previously uncharacterized taxa in the CIAMIB.

(i) Novel Bacteroidota. Microbes within the family Muribaculaceae (previously known as S24-7) have been consistently shown to be highly prevalent in the mouse gut in 16S rRNA gene and metagenomic surveys (39, 40). However, only five isolates from this family have been cultivated to date (10, 18, 40–42). The CIAMIB contains two novel taxa most closely related to *Muribaculum intestinale* YL27, the first member of the family Muribaculaceae that was cultivated. MiGA classification analysis of isolate NM86_A22 places it in the family Muribaculaceae, potentially within the genus *Muribaculum*. However, the 16S rRNA sequence of NM86_A22 has 90.3% sequence similarity to *M. intestinale* YL27, suggesting that it likely belongs to a novel genus. MiGA classification analysis of the isolates NM04_E33, NM74_B14, and NM76_A32 indicated that these isolates belong to one species in a novel family, corroborated by a 16S rRNA gene sequence similarity of 87.7% to *M. intestinale* YL27.

MiGA classification analysis identified isolate NM73_A23 as a member of the family Prevotellaceae, potentially in a novel genus. By 16S rRNA identity, this isolate is most similar to *Prevotella loescheii* at 90.2%. Isolate NM69_E16B represents a novel *Bacteroides* species that is most closely related to *Bacteroides uniformis* (16S rRNA identity of 96.7%).

(ii) Novel Bacillota. The CIAMIB contains novel isolates from two families within the phylum Bacillota. Isolates NM01_1-7b and NM72_1-8 were identified by MiGA classification analysis as members of potentially two novel genera within the family Lachnospiraceae, with their 16S rRNA gene sequences most closely matching *Ruminococcus gnavus* at 91.0% and *Murimonas intestini* at 92.9% identity, respectively. The Lachnospiraceae are diverse and abundant in the gut but remain poorly represented by available isolates and reference genomes (43, 44). MiGA classification analysis identified isolate NM09_H32 as a member of the family Erysipelotrichaceae, possibly in a novel genus. The 16S rRNA gene sequence of this isolate most closely matched *Dubosiella newyorkensis* at 90.5% identity (45).

(iii) Novel Actinomycetota. Bacteria in the phylum Actinomycetota are ubiquitous in human and mouse gut microbiotas, albeit at a lower abundance than those of the phyla Bacillota and Bacteroidota. Within Actinomycetota, the CIAMIB contains two novel species in the family Atopobiaceae: isolate NM07_P-09 and isolate NM08_P-01. MiGA classification analysis identifies these isolates as likely each belonging to a novel genus. The 16S rRNA gene sequence of NM07_P-09 is most similar to *Olsenella umbonata* at 92.8% identity, and NM08_P-01 is most similar to *Olsenella uli* at 94.1%.

(iv) Novel Pseudomonadota. Among the Betaproteobacteria class, members of the genus *Parasutterella* are highly prevalent in the human and mouse gut, but only two type strains (*Parasutterella excrementihominis* and *P. secunda*) have been cultivated (46, 47). MiGA classification analysis indicates that isolate NM82_D38, whose 16S rRNA gene sequence most closely matches *P. excrementihominis* at 96.0% identity, is a novel species in the genus *Parasutterella*.

Strain variation within isolates of the same species. Although 16S rRNA sequencing is frequently used to differentiate species, it is of limited utility in distinguishing between strains within a species. Given that strains can be functionally distinct, we sought to ascertain whether multiple strains coexisted within the same mouse. Of the 608 isolates analyzed, 44% were from either the genus *Bacteroides* or the family Lactobacillaceae. Many isolates were identified as *Limosilactobacillus reuteri* (formerly *Lactobacillus reuteri*), *Ligilactobacillus murinus* (formerly *Lactobacillus murinus*), or *Phocaeicola sartorii*.

To determine whether there were substantial genomic differences between these isolates, we employed random amplification of polymorphic DNA (RAPD). RAPD is a molecular technique used to differentiate strains based on PCR amplification of genomic DNA with a random primer, with the resulting amplicons acting as a unique fingerprint for individual strains (48). RAPD was used to examine eight isolates of *L. reuteri*, nine isolates of *L. murinus*, and five isolates of *P. sartorii* (Fig. S2). Two distinct strains were identified among the eight *L. reuteri* isolates (CIAMIB strains NM11_1-41 and NM12_1-47) and among the nine *L. murinus* isolates (CIAMIB strains NM26_J9 and

NM28_3M-8). The five isolates of *P. sartorii* appeared to be clonal. After verifying the distinct strains by whole-genome sequencing, the CIAMIB was expanded to encompass 41 unique isolates.

A comparison of *L. reuteri* protein-coding genes (as annotated by NCBI) using a 95% amino acid sequence identity threshold revealed that 28.6% of the 2,111 genes in strain NM11_1-41 were not found in NM12_1-47. Conversely, of the 1,924 protein-coding genes in strain NM12_1-47, 21.7% were not found in NM11_1-41. The *L. murinus* strain genomes were similarly distinct, with 24.4% of the 2,305 protein-coding genes in NM26_J9 not found in NM28_3M-8 and 23.6% of the 2,281 genes in NM28_3M-8 not found in NM26_J9 (Table S3). Examples of strain-specific genes include several that encode surface proteins, glycosyltransferases, and flippases, suggesting that the cell surface of each strain is distinct (Table S3).

We also examined the predicted metabolic functions of each strain using the KofamKOALA analysis pipeline provided by the Kyoto Encyclopedia of Genes and Genomes (KEGG) (49). While core metabolisms were largely similar between strains, there were several differences that could be functionally important in the context of the murine gastrointestinal tract. Among the *L. reuteri* strains, NM11_1-41 uniquely carries genes necessary to synthesize threonine *de novo* from homoserine, while NM12_1-47 uniquely encodes citrate lyase activity (Table S3). NM12_1-47 also encodes acetyl coenzyme A (acetyl-CoA) carboxylase enzymes (*accC* and *accD*) to initiate fatty acid synthesis, a gene cluster that appears to be undergoing loss via pseudogenization and deletion in NM11_1-41. We found that this particular pseudogenization event appears to be common among a large subset of the sequenced *L. reuteri* isolates available on NCBI, suggesting that subclades of *L. reuteri* have formed with respect to fatty acid metabolism. Each *L. reuteri* strain harbors two prophage-like elements that are not found in the other strain, as well as several small genetic islands that are sporadically distributed across the 50 currently sequenced *L. reuteri* genomes. Differences in core metabolism between the two *L. murinus* strains may be minor and related to the ability to metabolize lactose and rhamnose; many of the genes that differ between NM26_J9 and NM28_3M-8 are currently annotated as hypothetical. NM26_J9 appears to have two prophage-like elements that are absent in NM28_3M-8. Likewise, NM28_3M-8 has a prophage-like genetic element that is absent from NM26_J9.

Defining the conditions that permit growth of strains in the CIAMIB. Gut microbiotas contain many fastidious anaerobic organisms that have complex nutritional and environmental requirements, which makes culturing these species in the laboratory a challenging task. We investigated the ability of each strain from the CIAMIB to grow anaerobically in isolation in five commonly used liquid media and on nine solid-agar media (Fig. 2). Pseudomonadota strains that could be grown aerobically were also assayed for growth under aerobic conditions on the nine solid-agar medium formulations.

Agars that yielded consistent growth in anaerobic or aerobic conditions were Media that Grows Everything - Anaerobic enrichment (MTGE) agar, brain heart infusion (BHI) agar, BHI agar supplemented with horse blood and taurocholate, *Brucella* agar supplemented with sheep blood, and chocolate agar. Interestingly, many species that could be propagated on a particular solid medium were unable to grow in the same formulation of liquid medium. For example, species belonging to the families Eggerthellaceae and Lachnospiraceae grew poorly in all liquid media tested. This observation may help to explain the poor representation of the family Lachnospiraceae in culture collections, despite its abundance in human and mouse guts (44).

The CIAMIB contains 4 new isolates from the family Muribaculaceae and 3 isolates from a novel, but closely related, family (NM04_E33, NM74_B14, and NM76_A32). Given their prevalence in the guts of many mammals, it is remarkable that so few species in either family have been successfully cultivated. Furthermore, it has been reported that several Muribaculaceae isolates failed to revive after cryogenic storage (40). We hypothesized that these species were more difficult to propagate in liquid media because they require a high inoculum density to resume growth. To evaluate this hypothesis, we tested

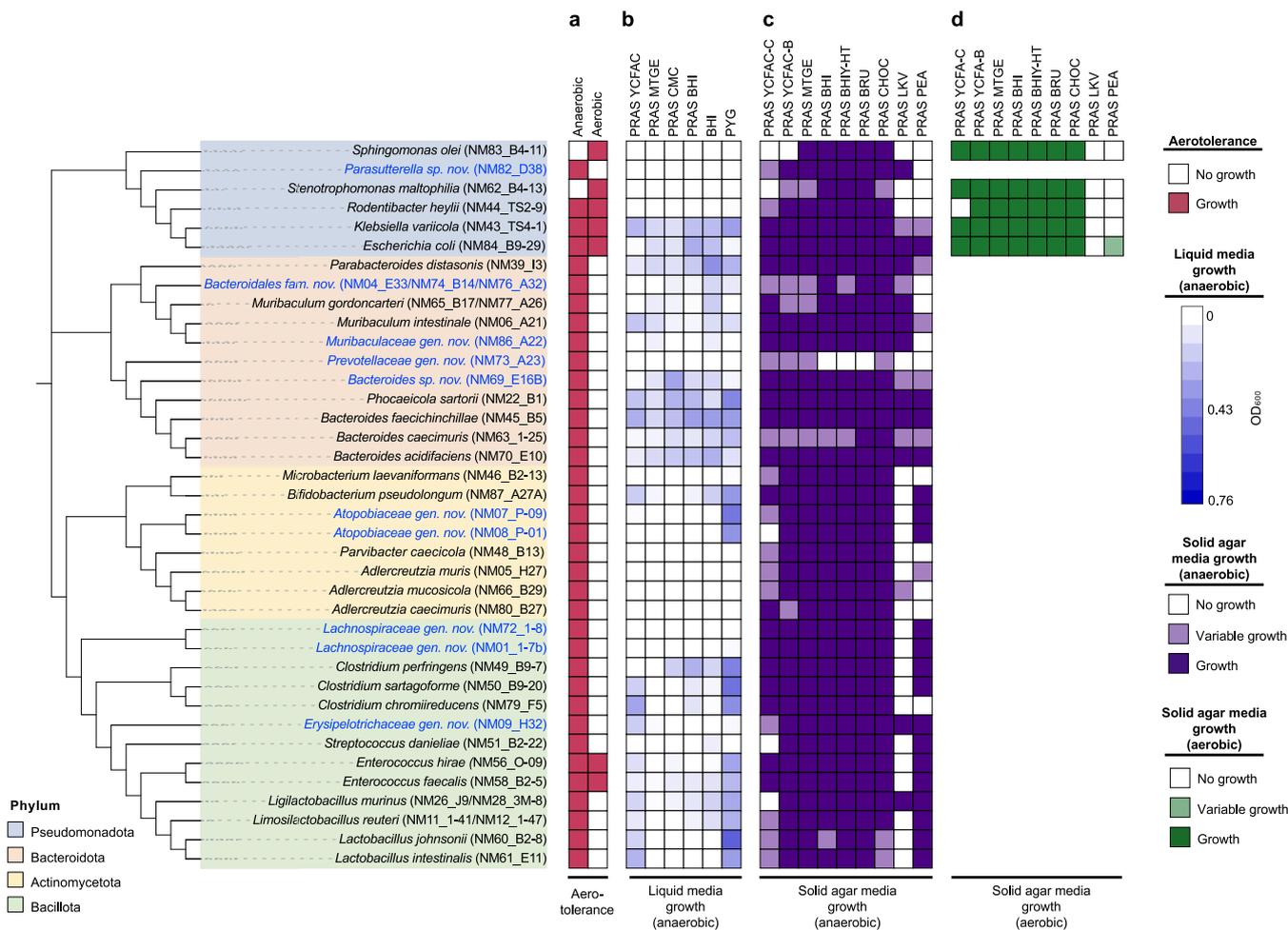


FIG 2 CIAMIB species display diverse growth capacities. The cladogram on the left was generated using 16S rRNA gene sequences and includes all 39 species in the CIAMIB. It is color coded by phylum. Novel taxa with their candidate names are in blue text. From left to right, the growth of each species is shown in different oxygen availabilities (a), liquid media under anaerobic conditions (b), agar media under anaerobic conditions (c), and agar media under aerobic conditions (d). A species was considered to grow in a given agar medium if growth was observed in all three biological replicates. Variable growth was recorded if growth was observed in at least one biological replicate. Liquid media tested include PRAS (pre-reduced anaerobically sterilized) BHI (brain heart infusion broth), PRAS YCFAC (yeast Casitone fatty acids broth with carbohydrates), PRAS MTGE (anaerobic enrichment broth), PRAS CMC (chopped-meat medium), PYG (peptone yeast glucose broth), and BHI. All solid agar media were PRAS, including LKV (laked *Brucella* blood agar with kanamycin and vancomycin), PEA (phenylethyl alcohol blood agar), MTGE (anaerobic enrichment agar), BHIIY-HT (BHI agar with horse blood and taurocholate), BHI (BHI agar), YCFAC-B (yeast Casitone fatty acids agar with carbohydrates and sheep blood), YCFAC-C (yeast Casitone fatty acids agar with carbohydrates), and CHOC (chocolate agar).

dilution series for all CIAMIB species within the order Bacteroidales to determine the minimal inoculum optical density necessary for reliable growth initiation. Next, we determined the number of cells required for growth of the two Bacteroidales species that required the highest inoculum (*M. intestinale* isolate NM75_1-48a and isolate NM76_A32) and the lowest inoculum (*Bacteroides faecichinchillae* isolate NM45_B5 and *Parabacteroides distasonis* isolate NM39_I3) (Fig. S3A and B). *M. intestinale* and *Bacteroidales* (new family) NM76_A32 required 10-fold and 10,000-fold more cells, respectively, than *B. faecichinchillae* and *P. distasonis* to reliably resume growth after subculturing. Thus, strategies to cultivate individual bacteria from complex mixtures via limiting dilution may be strongly biased against species that grow poorly in liquid media, or those such as NM76_A32 that grow significantly better only at high cell densities, which may include many Muribaculaceae members.

Predicting metabolic profiles of CIAMIB isolates. We analyzed the draft genome of each CIAMIB isolate to gain insight into its functional metabolic capacity and the considerable differences in growth observed across taxa in the collection. Putative functions were assigned to each predicted protein-coding gene using the KEGG

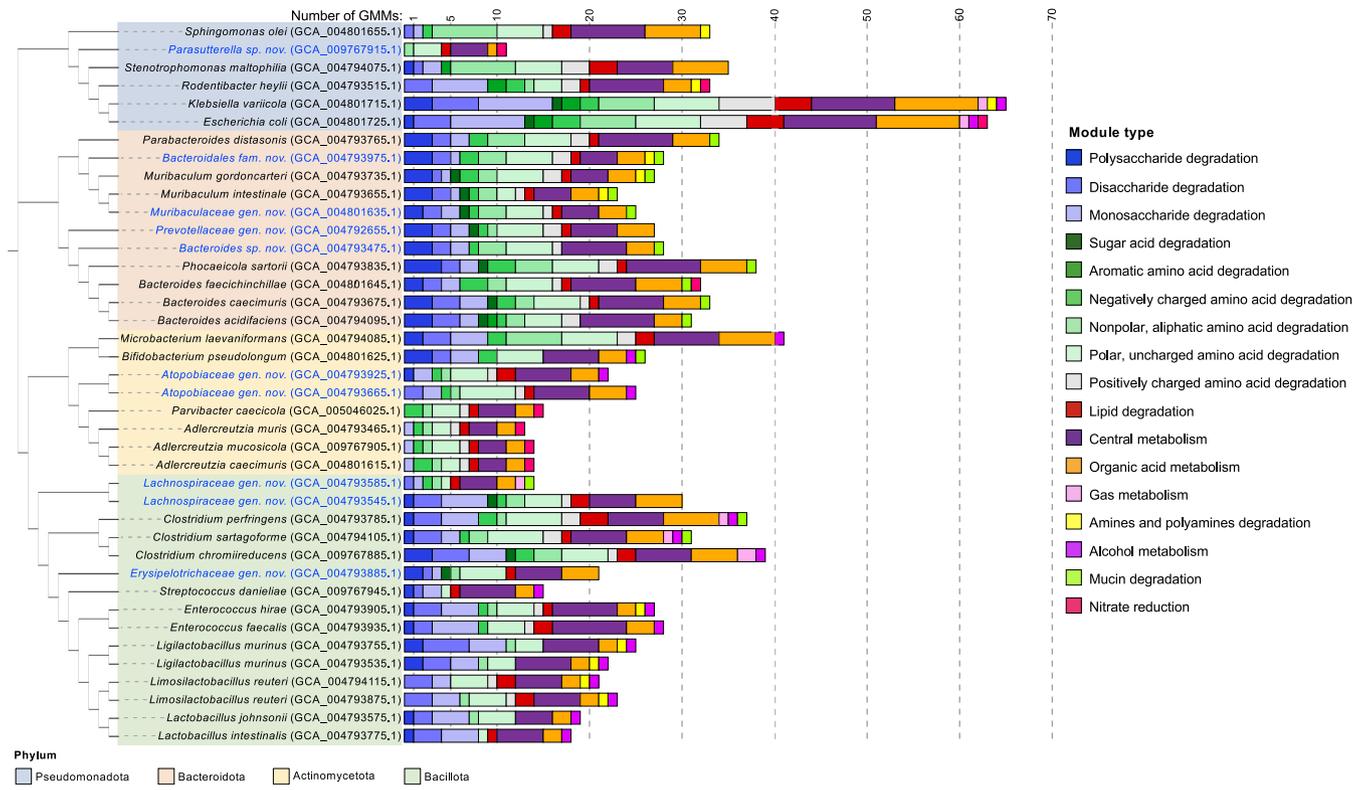


FIG 3 CIAMIB strains display a wide range of metabolic potential. The cladogram on the left shows the taxonomic distribution of CIAMIB strains based on 16S rRNA gene sequences and is color coded by phylum. Gut metabolic modules (GMMs) representing sets of closely related enzymatic functions were identified for all 41 CIAMIB strains.

KofamScan pipeline. Strains were clustered on the basis of their functional gene content (Table S4), which revealed that the predicted functional similarity of these isolates approximately, but not perfectly, aligns with their taxonomic relationships. Particularly notable was the clustering of all *Bacteroides* isolates and their separation from Bacillota, Actinomycetota, and Pseudomonadota isolates.

The genomes of CIAMIB strains were also bioinformatically assessed using a targeted gut metabolic analysis framework based on a set of manually curated reference modules (50). This analysis searches genomes for the presence/absence of 103 gut metabolic modules (GMMs), each representing a metabolic pathway predefined by a set of orthologous groups involved in the pathway. These GMMs encompass several functions, including catabolism of carbohydrates, amino acids, and lipids; cross-feeding interactions; production of fermentation end products; methanogenesis; and mucus degradation (50).

Bacteroidota species all contained GMMs for polysaccharide degradation, whereas Bacillota were enriched in GMMs for monosaccharide and disaccharide degradation (Fig. 3). These observations align with previous reports on species in the genus *Bacteroides* and the family Lactobacillaceae (51, 52). Many Bacteroidota species contained GMMs for sugar acid degradation, a trait that appeared in only a few species from other phyla, such as the GMM-rich Pseudomonadota members *Klebsiella variicola* and *Escherichia coli*. Most species in the CIAMIB contained GMMs to degrade polar and uncharged amino acids, whereas the ability to degrade negatively charged and aromatic amino acids was primarily present in Bacteroidota species, *K. variicola*, and *Escherichia fergusonii*. Specific GMMs for each strain are detailed in Table S5.

Unsurprisingly, isolates differed widely in their predicted metabolic capacities, largely depending on their phylogeny. Isolate NM82_D38 (the novel *Parasutterella* species) had the fewest putative metabolic modules (11 GMMs), while *K. variicola* isolate NM43_TS4-1 contained the most (65 GMMs). The mean across all species was 28

GMMs. However, fastidious obligate anaerobic organisms—defined here as obligate anaerobes that failed to reach an optical density at 600 nm (OD_{600}) of 0.15 in any liquid medium—had a mean of only 17 GMMs per genome. These findings suggest that fastidious anaerobic organisms fail to grow robustly in liquid culture at least in part due to an inability to synthesize one or more essential metabolites *de novo* or to utilize or access potential sources of essential metabolites in the media.

GMM analysis of species in the family Eggerthellaceae was particularly enlightening, as few complete genomes were previously available and no prior metabolic analysis had been reported. The only carbohydrate degradation module contained in the *Adlercreutzia* genomes (*Adlercreutzia muris*, *A. caecimuris*, and *A. mucosicola*) was for rhamnose. Even more surprising was that *Parvibacter caecicola*, an Eggerthellaceae species neighboring members of the genus *Adlercreutzia*, contained no carbohydrate degradation GMMs. Analysis of Eggerthellaceae genomes in the CIAMIB for carbohydrate active enzymes (CAZymes) revealed that these isolates contain few glycoside hydrolases, which mediate the hydrolysis of glycosidic bonds, compared to other Actinomycetota from the CIAMIB (Fig. S4) (53, 54). Furthermore, all four Eggerthellaceae species possessed the GMM for nitrate reduction, a feature present in only four other species in the CIAMIB (*B. faecichinchillae* and the Pseudomonadota members *E. coli*, *Rodentibacter heyllii*, and the novel *Parasutterella* sp.). Our understanding of nitrate-reducing bacteria in the gut microbiota is primarily limited to lactic acid bacteria and enterobacterial pathogens like *E. coli* and *Salmonella* (55–57). Nitrate is found at high levels in processed foods and is also a stable end product generated from nitric oxide by activated phagocytes. Phagocyte-derived nitrate can be utilized by gut pathogens and pathobionts as a terminal electron acceptor for anaerobic respiration during inflammation (58). These data suggest that these members of the family Eggerthellaceae represent a novel group of nitrate-reducing members of the gut microbiota that is severely restricted in its ability to grow on carbohydrates.

Genomes of the CIAMIB strains overlap variably with metagenomic assemblies.

Over the past few years, attempts to assemble complete microbial genomes from shotgun metagenomic sequence data of gut microbiota samples have generated numerous metagenome-assembled genomes (MAGs). A recently published effort, the Integrated Mouse Gut Metagenome Catalog (iMGMC), used data from several independently sequenced and biologically distinct murine gut samples to improve confidence in assigning particular genes to a putative genome (59). The iMGMC effort generated thousands of contigs that were subsequently placed into >1,300 “bins,” where each bin is intended to represent a MAG or the genomic contigs from an individual species or strain. Notably, a significant proportion of the contigs assembled from metagenomic data by the iMGMC were not assignable to a particular MAG and hence were left unbinned.

The genome sequences of our isolates allowed us to retrospectively determine how well MAGs assembled from short-read metagenomic data performed in terms of correct and complete assignments to known strains. Nucleotide sequences of open reading frames (ORFs) identified by the NCBI Prokaryotic Genome Annotation Pipeline from the whole-genome sequences of CIAMIB isolates were matched by BLAST (95% sequence identity) to metagenome-derived contigs generated by the iMGMC. We then determined the percentage of CIAMIB ORFs in the iMGMC that were assigned to a bin with the same taxonomic classification as identified by the 16S rRNA gene of the CIAMIB isolate (Table S6). Among the CIAMIB strains with MAGs represented in the iMGMC data, the proportion of binned sequences with the same taxonomic classification as determined by the 16S rRNA gene varied considerably (Fig. 4). For example, 85% of the ORFs in *Clostridium perfringens* strain NM49_B9-7 were binned by the iMGMC into a single MAG of same predicted taxonomy. Conversely, the genomic sequences from *E. coli* and the Bacteroidales species within the CIAMIB were less successfully predicted by the iMGMC, with approximately 50% of the protein-coding genes of most isolates found in contigs not assigned to a MAG by the iMGMC. Many ORFs from CIAMIB isolates were assigned to a bin with a phylogeny inconsistent with the one determined by the 16S rRNA gene, or in some cases were simply not

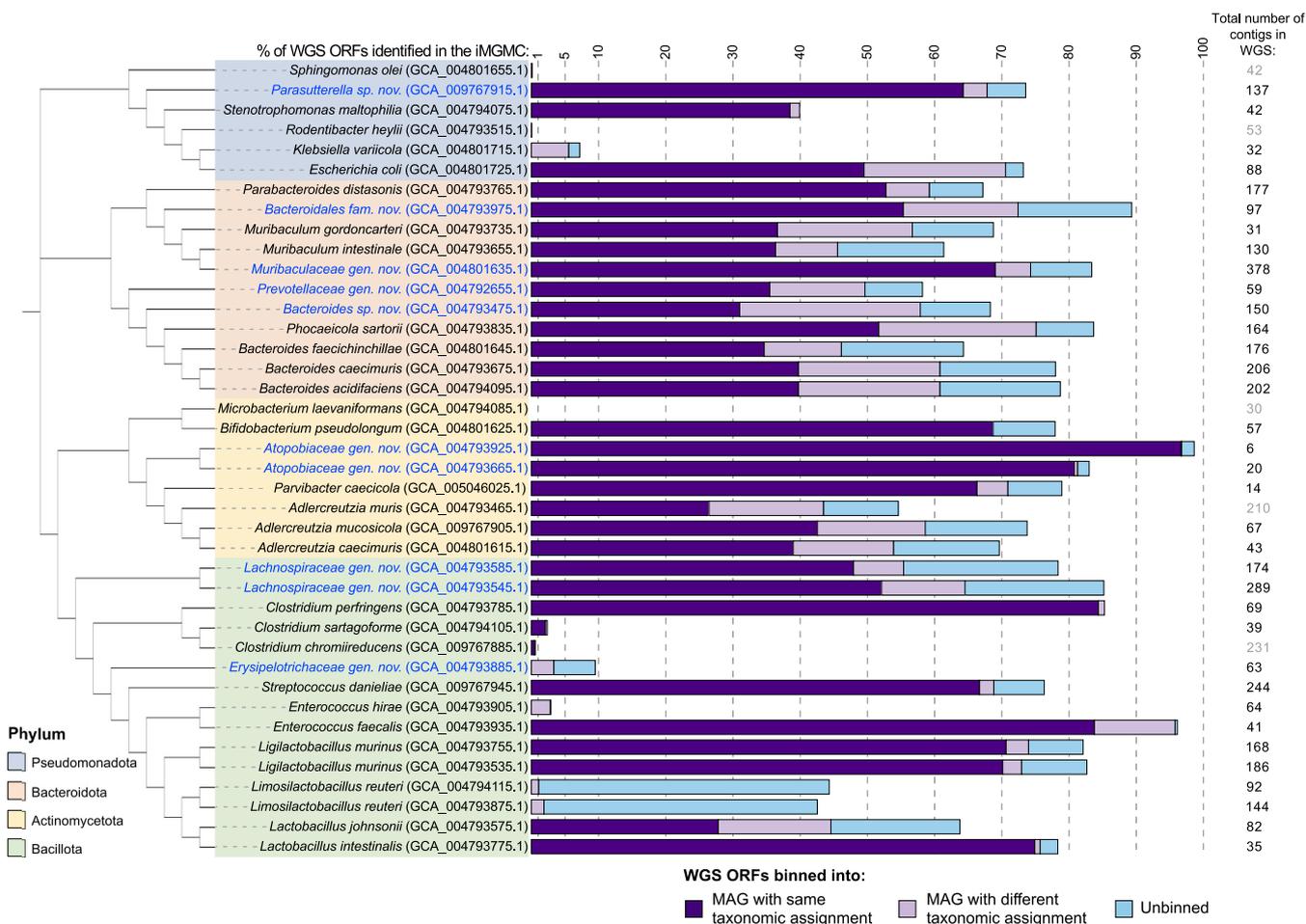


FIG 4 Comparison of CIAMIB ORFs to MAG ORFs. Nucleotide sequences of ORFs identified by the NCBI Prokaryotic Annotation Pipeline from the whole-genome sequences (WGS) of CIAMIB isolates were matched by BLAST (95% sequence identity) to metagenome-derived contigs generated by the Integrated Mouse Gut Metagenome catalog (iMGMC). Shown for each strain are the percentage of total ORFs that were identified in a MAG assigned to the same taxonomic classification as determined by the 16S rRNA gene sequence of CIAMIB isolates (dark purple), a MAG of a different taxonomic assignment (light purple), and to contigs within the iMGMC that were not assigned to a MAG (light blue). The remaining ORFs were not identified among the contigs assembled in the iMGMC. The total number of contigs in each assembly is indicated on the far right. Contig numbers in gray represent assemblies in which fewer than 1% of ORFs were identified in the iMGMC.

represented in the iMGMC metagenomic assemblies at all. Notably, the genomes of 11 CIAMIB isolates were not represented or were very poorly represented in the iMGMC data, suggesting that species similar to these particular isolates were not present in the microbiota of the mice used to generate the data for the iMGMC. These findings confirm the idea that, while useful, most MAGs represent only incomplete and partly accurate genomes of individual strains and that some taxa, like Bacteroidales, have genomes that present a greater challenge to assemble into MAGs than other taxa, like *C. perfringens*.

DISCUSSION

Here, we report the first iteration of the CIAMIB, a collection of bacteria from the colons and ceca of mice with colonic inflammation. The CIAMIB contributes to the growing but still limited number of mouse gut isolates available to the research community (18, 19, 60). Future microbiota research will be greatly accelerated by expanding the number of available isolates associated with a variety of diseases, including obesity, diabetes, nonalcoholic fatty liver disease, and neurologic and psychiatric disorders. Our functional characterizations of the CIAMIB isolates, including optimal growth conditions, will enable the construction of defined microbial communities that can be tailored to recapitulate the microbiota during various physiological states. Gnotobiotic

mice colonized with a defined microbiota constructed entirely from well-characterized isolates will be an important tool to experimentally define the role that any individual species or strain plays during the host inflammatory response.

Several isolates within the CIAMIB are particularly notable with respect to the role they may play during inflammation. *Enterococcus faecalis*, Pasteurellaceae species (*R. heyliae*), and *K. variicola* have been found at a higher abundance in IBD and CRC patients than healthy cohorts (61–63). We isolated novel species from the genus *Parasutterella* and the family Prevotellaceae, both of whose increased abundance has been associated with several diseases, including IBD (64–66). Our isolate from the family Erysipelotrichaceae is of interest, as there have been numerous reports of the potential of this family to modulate host physiology and/or disease (67–71), including high immunogenicity, increased abundance after treatment with broad-spectrum antibiotics, and enrichment in CRC (45, 67, 68, 72, 73).

Other taxa of note in the CIAMIB include novel Lachnospiraceae and Muribaculaceae species. Both of these families are deficient in taxonomic resolution and cultured isolates despite being diverse and abundant members of the mouse gut microbiota (10, 18, 39, 41). While higher Lachnospiraceae abundance has primarily been associated with a healthy intestinal state, 16S rRNA metagenomic sequencing of our *IL10^{-/-}* mice showed that Lachnospiraceae remained abundant even under inflammation conditions, an observation supported by other studies (30, 62, 74, 75). The limited literature discussing the family Muribaculaceae has demonstrated both increases and decreases during intestinal inflammation and both blooms and depletions following treatment with broad-spectrum antibiotics (28, 76, 77). These conflicting associations remain common in gut microbiota research and are not surprising given the bacterial diversity encompassed at high taxonomic levels.

Metagenomic studies have become increasingly important for understanding the functional potential of a given microbiota and for *in silico* reconstruction of uncultivated bacterial genomes. Our results suggest caution regarding the increasing reliance on correlating specific functions to particular microbiota members based on short-read 16S sequencing surveys, as two species with nearly identical 16S sequences can harbor very different metabolic or pathogenic potential. We also observed that a large fraction of the putative genomes computationally assembled from very large and diverse short-read shotgun metagenomic experiments were neither accurate nor complete. We therefore urge caution against assuming that MAGs represent accurate genome reconstructions of individual strains or species.

Future work will aim to identify and cultivate species that are underrepresented or not represented in current culture collections. One potential bottleneck we identified in isolating individual microbes from a complex sample are cases, such as Muribaculaceae species, in which the microbe will not reliably resume growth when diluted to a single cell. Additionally, in trying to elucidate growth conditions for CIAMIB species, we found that several that could be propagated on a particular agar medium could not be cultured in the liquid medium of the same formulation. It is possible that certain species will be refractory to isolation using dilution-based protocols and might grow only in the presence of other microbes or with the addition of spent culture media (18, 19, 78). Despite the caveats mentioned above, MAGs can be a powerful tool for identifying the metabolic bottlenecks of uncultivated species and can aid in the design of targeted strategies and growth media for their cultivation.

In conclusion, our work provides a foundational collection of inflammation-associated microbes from the mouse gut. The CIAMIB can be used independently or in combination with other collections to define the roles that each microbe plays in the gut ecosystem, in the absence and presence of inflammation. We provide useful information on culturing each species in the collection and report on the growth patterns within related taxa. The collection will facilitate future functional studies and help to refine strategies for targeted isolation as the field works to expand the availability of mouse gut bacterial isolates.

MATERIALS AND METHODS

Mouse background and colon sample collection. All mice used were of the C57BL/6J background. *Apc*^{Min/+} mice were purchased from Jackson Laboratories and carry a nonsense mutation in exon 15 of the *Apc* gene. *IL10*^{-/-} mice and *Msh2*^{+/-} mice were provided by Ken Croitoru and Tak Wah Mak, respectively, of the University of Toronto. *IL10*^{-/-} and *Apc*^{Min/+} *Msh2*^{-/-} mice developed colitis spontaneously without any treatment. *Apc*^{Min/+} *Msh2*^{+/-} mice did not spontaneously develop colitis but harbored a microbiota similar to that of *Apc*^{Min/+} *Msh2*^{-/-} mice based on 16S rRNA gene sequencing. *IL10*^{-/-} mice were raised under conventional conditions and fed a Teklad Global 18% protein rodent chow (Harlan, WI, USA).

Seven- to 9-week-old mice were sacrificed by cervical dislocation and immediately transferred to an anaerobic chamber (Anaerobe Systems AS-580) containing 10% CO₂, 10% H₂, and 80% N₂ (Praxair NI CD10H1C-K). The working area, dissection tools, and abdomen of each mouse were soaked with 70% (vol/vol) ethanol before dissection to avoid contamination. Cecal, mucosal, and colonic contents were collected by scraping the intestinal wall. For each mouse, these contents were weighed and resuspended in pre-reduced phosphate-buffered saline (PBS) with 0.1% cysteine to a final concentration of 15 mL of PBS per gram of contents. Suspensions were vortexed for 5 min and allowed to settle for 5 min, after which the supernatant was transferred to a new tube and used as starting material. Aliquots of supernatant and supernatant with equal volumes of pre-reduced 50% glycerol were frozen for subsequent genomic extractions and isolations, respectively. The remaining supernatant was used in isolations.

Culture media. All quantities are per 1 L of medium unless otherwise indicated. Postautoclaving reagents were added after the medium cooled to 50°C. Media containing sheep blood were poured immediately, as sheep blood addition cools down agar substantially.

(i) Stock solutions. (a) *Vitamin K (1 mg/mL)*. One hundred microliters of vitamin K1 (Sigma-Aldrich V3501) were diluted into 10 mL of 95% ethanol.

(b) *Hemin (5 mg/mL)*. Fifty milligrams of hemin (Sigma-Aldrich H9039) were dissolved in 1 mL of 1 M NaOH. The volume was adjusted to 10 mL using distilled water, then filter-sterilized.

(c) *L-Cysteine (500 mg/mL)*. Five grams of L-cysteine hydrochloride monohydrate (Sigma-Aldrich C6852) were dissolved in 10 mL distilled water and then filter-sterilized.

(ii) Types of media used for isolations. Descriptions of medium are adapted from the work of Lau et al. [44].

(a) *AIA*. We followed the manufacturer's instructions for Actinomycete isolation agar (AIA; Himedia M490).

(b) *Beef agar*. Beef agar with L-cysteine (0.5 g/L), hemin (10 mg/L), and vitamin K (1 mg/L) was prepared as follows. We followed the manufacturer's instructions for cooked meat medium (Himedia M1495) with several modifications. Before autoclaving, we added 15 g agar (BD 214010). After autoclaving, we added 1 mL L-cysteine stock solution, 2 mL hemin stock solution, and 1 mL vitamin K stock solution.

(c) *BHI*. For brain heart infusion (BHI) agar with L-cysteine (0.5 g/L), hemin (10 mg/L), and vitamin K (1 mg/L), we followed the manufacturer's instructions for BHI broth (Oxoid CM1135) with 15 g agar added and several modifications. After autoclaving, we added 1 mL of L-cysteine stock solution, 2 mL of hemin stock solution, and 1 mL vitamin K stock solution.

(d) *Brucella agars*. For *Brucella* agar with sheep blood (5%), hemin (5 mg/L), and vitamin K (1 mg/L), we followed the manufacturer's instructions for *Brucella* broth (BD 211088) with several modifications. Before autoclaving, we added 15 g agar (BD 214010). After autoclaving, we added 1 mL hemin stock solution, 1 mL vitamin K stock solution, and 50 mL defibrinated sheep blood (Cedarlane CL2581-100D). For *Brucella* agar with sheep blood (5%), trimethoprim (5 mg/L), vancomycin (10 mg/L), and polymyxin B (2500 IU/L), we followed the manufacturer's instructions for *Brucella* broth (BD 211088) with several modifications. Before autoclaving, we added 15 g agar (BD 214010). After autoclaving, we added 50 mL defibrinated sheep blood (Cedarlane CL2581), 1 mL trimethoprim (Sigma-Aldrich 92131) dissolved to 5 mg/mL in dimethyl sulfoxide (DMSO), 2,500 IU polymyxin B sulfate (Sigma-Aldrich P0972) dissolved in distilled water and filter-sterilized, and 1 mL vancomycin hydrochloride (Sigma-Aldrich V2002) dissolved to 10 mg/mL in distilled water and filter-sterilized.

(e) *Columbia agars*. For Columbia agar with sheep blood (5%), we followed the manufacturer's instructions for Columbia agar (BD 211124) with one modification. After autoclaving, we added 50 mL defibrinated sheep blood (Cedarlane CL2581). For Columbia agar with colistin and nalidixic acid (CNA), we followed the manufacturer's instructions for CNA (BD 212104) with one modification. After autoclaving, we added 50 mL defibrinated sheep blood (Cedarlane CL2581).

(f) *Laked blood agar*. For laked blood agar with kanamycin (100 mg/L) and vancomycin (7.5 mg/L), we followed the manufacturer's instructions for tryptic soy broth (BD 211825) with several modifications. Before autoclaving, we added 5 g yeast extract (BD 212750) and 15 g agar (BD 214010). After autoclaving, we added 10 mL vitamin K stock solution, 1 mL hemin stock solution, 0.8 mL L-cysteine stock solution, 2 mL kanamycin monosulfate (BioShop KAN201) dissolved to 50 mg/mL in distilled water and filter-sterilized, 75 μL vancomycin hydrochloride (Sigma-Aldrich V2002) dissolved to 100 mg/mL in distilled water, and 50 mL laked sheep blood (Cedarlane CL2581-100L).

(g) *LB agar*. We followed the manufacturer's instructions for Luria-Bertani (LB) agar (BD 244520).

(h) *MacConkey agar*. We followed the manufacturer's instructions for MacConkey agar (BD 281810).

(i) *Mannitol salt agar*. We followed the manufacturer's instructions for mannitol salt agar (BD 211407).

(j) *McKay agar*. McKay agar was 13.3 g nutrient broth (BD 234000), 5 g dextrose (BioShop GLU501), 15 g agar (BD 214010), 10 g yeast extract (BD 212750), 5 g tryptone (BD 211699), 2 g K₂HPO₄, 40 mL salt solution (10 g/L NaHCO₃, 2 g/L NaCl, 1 g/L K₂HPO₄, 1 g/L KH₂PO₄, 0.5 g/L MgSO₄·7H₂O, 0.25 g/L CaCl₂·2H₂O), 1 mL Tween 80, 1 mg crystal violet, 60 mg bromocresol purple, 10 μg vitamin K, 0.05 g hemin, and 15 g Bacto agar in 1 L of distilled water. After autoclaving, we added 20 mL L-arginine (2.5% [wt/vol]),

sulfadiazine (Sigma-Aldrich S8626), colistin sulfate, and oxolinic acid to final concentrations of 500 mg/L, 10 mg/L, and 5 mg/L, respectively (79).

(k) *Phenylethyl alcohol agar*. For phenylethyl alcohol agar, we followed the manufacturer's instructions (BD 211539) with one modification. After autoclaving medium, we added 50 mL defibrinated sheep blood (Cedarlane CL2581).

(l) *Reinforced clostridial agar*. We followed the manufacturer's instructions for reinforced clostridial agar (Oxoid CM0151).

(m) *Rifampin blood agar*. For rifampin blood agar, we followed the manufacturer's instructions for *Brucella* broth (BD 211088) with several modifications. Before autoclaving, we added 15 g agar (BD 214010). After autoclaving, we added 50 μ L of 1,000 μ g/mL rifampin (Sigma R3501) (10 mg dissolved in 2 mL absolute ethyl alcohol, with 8 mL distilled water added), and 50 mL defibrinated sheep blood (80).

(n) *Sulfite polymyxin sulfadiazine agar*. Sulfite polymyxin sulfadiazine agar was 15 g pancreatic digest of casein, 15 g agar (BD 214010), 10 g yeast extract (BD 212750), 0.5 g ferric citrate, 0.5 g Na₂SO₃, 0.12 g sulfadiazine (Sigma-Aldrich S8626), and 0.01 g polymyxin sulfate in 1 L of distilled water.

(o) *THB agar*. For Todd-Hewitt broth (THB) agar with yeast extract, we followed the manufacturer's instructions for THB (BD 249240) with two modifications. Before autoclaving, we added 10 g yeast extract (BD 212750) and 15 g agar (BD 214010).

(p) *Tryptic soy agars*. For tryptic soy agar with sheep blood (5%), we followed the manufacturer's instructions for tryptic soy broth (BD 211825) with two modifications. Before autoclaving, we added 15 g agar (BD 214010). After autoclaving, we added 50 mL defibrinated sheep blood (Cedarlane CL2581). For tryptic soy yeast agar with sheep blood (5%), L-cysteine (0.5 g/L), hemin (10 mg/L), and vitamin K (1 mg/L), we followed the manufacturer's instructions for tryptic soy broth (BD 211825) with several modifications. Before autoclaving, we added 3 g yeast extract (BD 212750) and 15 g agar (BD 214010). After autoclaving, we added 50 mL defibrinated sheep blood (Cedarlane CL2581), 1 mL L-cysteine stock solution, 2 mL hemin stock solution, and 1 mL vitamin K stock solution.

(iii) Media used to propagate CIAMIB isolates after isolation. Once isolated, most strains were found to grow on *Brucella* agar supplemented with sheep blood, hemin, and vitamin K (Hardy Diagnostics A30) and were therefore routinely propagated on this medium. The two exceptions were *Streptococcus danie-liae* and Prevotellaceae (new genus) isolate NM73_A23, which were propagated on THB agar with yeast extract and in BHI broth, respectively.

Isolation and identification of bacterial strains. Gut sample resuspensions were diluted 10-fold, and 100 μ L were plated on 17 types of agar medium (see compositions in "Culture media"). All media were placed in an anaerobic chamber at least 48 h prior to use to remove oxygen. Plates were then incubated at 37°C for 2 to 14 days under the following conditions: aerobic; anaerobic with a mixture of 10% hydrogen, 10% carbon dioxide, and 80% nitrogen; or hypoxic with a mixture of 5% oxygen, 10% carbon dioxide, and 85% nitrogen. Each day, plates were checked for growth, and when single colonies appeared, they were restreaked onto fresh plates. This procedure was always performed at least twice to obtain a single-colony morphology.

Once colonies appeared morphologically pure, several were picked and resuspended in 50 μ L sterile distilled water. Samples were boiled for 15 min and then centrifuged for 5 min at 16,000 \times g. The supernatant was transferred to a new tube, and 5 μ L were used as the template DNA for PCR-based amplification of the 16S rRNA gene. PCRs were carried out using the 16S rRNA gene universal primers 8f (5'-AGAGTTGATCCTGGCTCAG-3') and 1492r (5'-CGGTTACCTTGTACGACTT-3') (81). Amplification was performed using the ProFlex 3 by 32-well PCR system, in 50- μ L reaction mixtures with 5 μ L of 10 \times buffer, 1.5 μ L of 50 mM MgCl₂, 1 μ L of 10 mM deoxynucleoside triphosphates (dNTPs), 5 μ L of 10 mM forward primer, 5 μ L of 10 mM reverse primer, 5 μ L of DNA template, 27.75 μ L of nuclease-free water and 0.25 μ L of *Taq* DNA polymerase. The cycler was programmed for 2 min at 94°C for initial denaturation, followed by 32 cycles of 1 min at 94°C for denaturation, 1 min at 56°C for annealing, 2.5 min at 72°C for extension, and finally 10 min at 72°C for the final extension. To check for amplification, PCR products were run at 100 V for 30 min on a 1% (wt/vol) agarose gel containing SYBR Safe DNA gel stain (Invitrogen S33102) diluted to 1/20,000 in 1 \times Tris-acetate-EDTA (TAE) buffer. Successfully amplified samples were purified using the GeneJET PCR purification kit (Thermo Scientific K0701). Next, Sanger sequencing of the 16S rRNA gene was carried out at The Centre for Applied Genomics (Toronto, Canada) using the 8f primer. Sequences were identified by performing a BLAST search of the 16S rRNA gene sequences against the NCBI reference RNA sequences (refseq_rna) database. A cutoff of 97% 16S rRNA gene identity was used to delineate novel taxa (82).

Cryo-stocks were prepared by resuspending cells grown for 3 to 5 days (depending on colony morphology) from plates in *Brucella* broth with glycerol (25% vol/vol), and stored at -80°C.

Bacterial growth assays. Strains were initially grown anaerobically from frozen stocks on media as described above for propagation at 37°C for 72 h. Strains were then expanded twice in order to obtain enough bacterial cells for assay. For the assay, bacterial cells were resuspended in peptone yeast glucose (PYG) medium (following the instructions for DSMZ medium 104). To test for growth on solid medium, resuspensions were plated separately on various agar media divided into 6 sections. Two microliters of the suspension were deposited into each sector and spread using 1- μ L inoculation loops. Afterward, plates under anaerobic conditions were incubated for 72 h at 37°C and plates under aerobic conditions were incubated for 48 h at 37°C. After incubation, plates were observed visually, photographed, and scored as growth (presence of colonies) or no growth (absence of colonies). Assays were performed in biological triplicate. A species was considered to grow in a given medium if growth was observed in all three replicates. Variable growth was recorded if growth was observed in at least one biological replicate. Growth under the various conditions was plotted using iTOL (83).

Solid growth was evaluated on nine types of prerduced, anaerobically sterilized (PRAS) media provided by Anaerobe Systems: laked *Brucella* blood agar with kanamycin and vancomycin (LKV), phenethyl alcohol blood agar (PEA), MTGE anaerobic enrichment agar (MTGE), *Brucella* blood agar (BRU), brain heart infusion agar with horse blood and taurocholate (BHIY-HT), brain heart infusion agar (BHI), yeast casitone fatty acids agar with carbohydrates and sheep blood (YCFAC-B), yeast Casitone fatty acids agar with carbohydrates (YCFAC), and chocolate agar (CHOC).

To test for growth in liquid media, resuspensions were diluted to an OD₆₀₀ of 1 and inoculated into 200 μ L of each of six types of liquid medium to a starting OD₆₀₀ of 0.1 in a 96-well plate. Liquid cultures were incubated anaerobically for 72 h at 37°C. Wells were mixed every 24 h by pipetting. To measure the final OD₆₀₀, 100 μ L of culture from each well were transferred to a new 96-well plate, and the OD was recorded using a Tecan 96-well plate reader (Infinite 200 Pro). Assays were performed in technical duplicate and biological triplicate. Raw OD₆₀₀ values were first averaged between technical replicates, normalized by subtracting the medium blank to correct for medium turbidity, and averaged between three biological replicates. Growth under various conditions was plotted using iTOL (83).

Liquid growth was evaluated in six types of broth (four PRAS media provided by Anaerobes Systems and two media made in house and reduced before use): BHI, YCFA, MTGE, chopped-meat medium with carbohydrates, BHI with 10 mg/L hemin and 1 mg/L vitamin K (Oxoid), and PYG.

To determine the minimum number of cells needed to promote growth in BHI for four of the *Bacteroidales* species, strains were initially grown anaerobically from frozen stocks on *Brucella* blood agar with hemin and vitamin K (Hardy Diagnostics A30) at 37°C for 48 h. Strains were then expanded by streaking twice more on *Brucella* blood agar to obtain enough bacterial cells for assay. For the assay, cells were resuspended to a standardized OD₆₀₀ of 0.02 and serially diluted 20-fold in a 96-well plate. Five microliters from each dilution were also spotted onto *Brucella* agar supplemented with sheep blood, hemin, and vitamin K and incubated anaerobically at 37°C for 72 h to determine the concentration of the initial resuspension in CFU per milliliter. This concentration was used to calculate the starting CFU in each well. Ninety-six-well plates were incubated anaerobically at 37°C for 48 h and mixed by pipetting at 24 h. To measure the final OD₆₀₀, 100 μ L of each dilution were transferred to a new 96-well plate, and the OD was recorded using a Tecan 96-well plate reader (Infinite 200 Pro). The assay was performed in technical duplicate and biological triplicate. Raw OD₆₀₀ values were first averaged between technical replicates and then normalized by subtracting the medium blank to correct for medium turbidity. A minimum OD₆₀₀ of 0.1 was used to determine successful growth. The number of CFU inoculated into the last well with successful growth in BHI was determined to be the minimum number of cells needed to promote growth. The minimum number of cells needed to promote growth was averaged across three biological replicates.

Whole-genome sequencing. In this study, 46 draft genomes were constructed, representing each strain in the CIAMIB and all Muribaculaceae isolates. DNA sequencing was carried out by the Genomics Platform of the Chan Zuckerberg Biohub or performed in house (detailed in NCBI BioSample information).

Genomic DNA for most isolates was extracted using the DNeasy PowerSoil kit (Qiagen 12888) following manufacturer's instructions. *S. danieliae* genomic DNA was extracted using the PureLink genomic DNA minikit (Invitrogen K182001) with several modifications. Briefly, *S. danieliae* was streaked onto two full THB agar plates and incubated anaerobically at 37°C for 72 h. Cells were collected and resuspended in tubes containing 200 μ L TE (10 mM Tris, 1 mM EDTA; pH 8), 500 U/mL mutanolysin, and 15 mg/mL lysozyme and then incubated with gentle shaking for 4 h at 37°C. Next, 25 μ L of 28.8 mg/mL proteinase K were added, and cells were incubated for 1 h at 56°C. Finally, 265 μ L of PureLink genomic lysis/binding buffer were added, and tubes were briefly vortexed before incubation at 55°C for 30 min. The remaining steps followed the manufacturer's instructions for binding, washing, and eluting DNA. Genomic DNA was eluted with 50 μ L UltraPure DNase/RNase-free distilled water.

All DNA libraries were prepared for sequencing using the Nextera XT DNA library preparation kit (Illumina FC-131) following the manufacturer's instructions. In-house sequencing was carried out using an Illumina MiniSeq according to manufacturer's instructions. At the Chan Zuckerberg Biohub, Nextera XT libraries with final volumes of 4 μ L were constructed using a miniaturized and high-throughput process leveraging low-volume liquid handlers, including a Mantis (Formulatrix), a Mosquito X1, and a Mosquito HTS (SPT Labtech). Libraries were pooled and sequenced on an Illumina NovaSeq S4 flow cell.

Reads were assembled using the comprehensive genome analysis service provided by the Pathosystems Resource Integration Center (PATRIC) through their Bacterial Bioinformatics Resource Centre (84). This service assembled reads using SPAdes v. 3.10.0 with the activated BayesHammer tool and MismatchCorrector (85, 86). Assemblies were evaluated using QUAST v. 2.3 (87). Predictions and annotations were performed with the NCBI Prokaryotic Genome Annotation Pipeline (88). Genome sequences were submitted to NCBI under the accession numbers provided in Table S1. Assemblies were also evaluated using CheckM (v1.0.18) through the KBase UI (version 2.5.2) platform with the default parameters to determine genome completion and contamination (89, 90).

Genome-based analyses. Multiple alignment of the full-length (>1,400-nucleotide [nt]) 16S rRNA gene sequence of each species (using a 97% cutoff value) was performed using Muscle in MEGA7 (91). This alignment was used to construct a phylogenetic tree based on the maximum-likelihood algorithm, using the Tamura-Nei model in MEGA7 (91). The tree was then annotated using iTOL (83).

Taxonomic classification based on average nucleotide identity (ANI) and average amino acid identity (AAI) was performed using the TypeMat tool from the Microbial Genomes Atlas (MiGA) (r2021-04 update). MiGA is a web server that allows the classification of genomes against all taxonomically assigned taxa with available genome sequences in the NCBI RefSeq and NCBI Genome Database, available at <http://microbial-genomes.org> (38).

Digital DNA-DNA hybridization (dDDH) to ascertain novel taxa was performed through the Genome-to-Genome Distance Calculator (GGDC) 2.1 using formula 2, a web service provided by the DSMZ (<http://ggdc.dsmz.de>) (35, 36). The genome of each species belonging to a novel taxon was compared to the reference genomes of the three phylogenetically closest species based on 16S rRNA sequence identity (37). The ANI between novel taxa and their closest neighboring species based on 16S rRNA sequence identity was calculated using the OrthoANI algorithm, available at <https://www.ezbiocloud.net/tools/ani> (92).

To obtain gut metabolic modules (GMMs), the set of functional genes from each strain was first defined using BLAST, filtering out non-coding sequence regions. DNA sequences were then converted into amino acid sequences and annotated using the Kyoto Encyclopedia of Genes and Genomes (KEGG) KofamScan pipeline using their database of prokaryotic hidden Markov models (HMMs) (93). Through this pipeline, KEGG Orthology (KO) numbers were assigned to the predicted coding genes of each bacterial strain on the basis of their similarity to the HMMs. To assess functional relatedness, CIAMIB strains were clustered by the presence/absence of the various KO functions identified using the clustermap function (method = "ward") of the Seaborn data visualization library based on the Python matplotlib library (94, 95) in a custom Python script.

Separately, the KO lists generated for each strain were used as input into Omixer-RPM using the human GMM database to map KO-annotated genes to predicted metabolic pathways (50, 96), using a module coverage cutoff 66% as indicated in the original publication (50). Total numbers of GMMs identified in each functional group were plotted using iTOL (83).

Carbohydrate-active enzymes were identified using the dbCAN2 meta server, a webserver for automated carbohydrate-active enzyme annotation, using their HMMER (hits: E value < 1e-15, coverage > 0.35), DIAMOND (hits: E value < 1e-102, hits per query [-k] = 1), and Hotpep (hits: frequency > 2.6, hits > 6) tools. Carbohydrate hydrolases identified by two or more tools were quantified and plotted.

For each strain in the CIAMIB, the protein-coding sequences were used to search assembled iMGMC contigs using standalone BLAST (BLASTn) from NCBI with a sequence identity threshold of 90% identity over 90% of the gene sequence. The protein-coding genes in each strain were assigned to specific contigs and, subsequently, MAGs and predicted taxonomic groups based on tables provided by the iMGMC (97). Taxonomic classifications assigned to MAGs by the iMGMC were manually compared to classifications of the 16S rRNA gene sequence of CIAMIB isolates assigned by BLAST (refseq_rna) to determine percentage of ORFs with the same classification. For novel taxa, the comparison was made to the highest taxonomic level of confidence (>94.5% for genus level, >86.5% for family level).

RAPD. Six sets of primers were used in RAPD (random amplification of polymorphic DNA) analysis of *L. reuteri*, *L. murinus*, and *P. sartorii*: GS10 (5'-TGACACGACG-3'), GS15 (5'-CCCACATCCC-3'), GS16 (5'-CCCAACCACC-3'), GS17 (5'-GGGAAGGAGG-3'), GS19 (5'-GGCACTGACG-3'), and GS20 (5'-TCCACCGACC-3'). Genomic DNA was extracted using the GenElute bacterial genomic DNA kit (Sigma-Aldrich) following the manufacturer's instructions and diluted to a final concentration of 10 ng/ μ L. PCR amplification was carried out using a Bio-Rad PCR machine, in a 50- μ L reaction mixture with 5 μ L of 10 \times buffer, 1.5 μ L of 50 mM MgCl₂, 1 μ L of 10 mM dNTPs, 10 μ L of 10 mM primer, 5 μ L of DNA template, 27.25 μ L of nuclease-free water, and 0.25 μ L of *Taq* DNA polymerase (Thermo Fisher). The cyclor was programmed for 20 s at 94°C for initial denaturation, followed by 35 cycles of 30 s at 95°C for denaturation, 30 s at 35°C for annealing, 1.5 min at 72°C for extension, and 10 min at 72°C for the final extension. To visualize, PCR products were run at 100 V for 30 min in a 1% (wt/vol) agarose gel containing SYBR Safe DNA gel stain (Invitrogen S33102) diluted 1/20,000 in 1 \times TAE buffer.

Data availability. All genome assemblies are publicly available in NCBI under BioProject accession number [PRJNA474907](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA474907). Genome sequences were submitted to NCBI under the accession numbers provided in Table S1. All novel taxa are available as individual cultures from the German Collection of Microorganisms and Cell Cultures (DSMZ) or upon request.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, EPS file, 0.3 MB.

FIG S2, EPS file, 0.3 MB.

FIG S3, EPS file, 0.3 MB.

FIG S4, EPS file, 0.2 MB.

TABLE S1, XLSX file, 0.04 MB.

TABLE S2, XLSX file, 0.04 MB.

TABLE S3, XLSX file, 0.3 MB.

TABLE S4, XLSX file, 0.9 MB.

TABLE S5, XLSX file, 0.02 MB.

TABLE S6, XLSX file, 0.02 MB.

ACKNOWLEDGMENTS

We thank Thomas Clavel (University Hospital of RWTH Aachen, Germany) and Emma Allen-Vercoe (University of Guelph, Canada) for their expertise and discussion regarding culturing microbes. We thank Thomas Reidel (Leibniz Institute, DSMZ—German

Collection of Microorganisms and Cell Cultures, Germany) for his help in the deposition of novel taxa to the DSMZ. We thank Dana Philpott (University of Toronto, Canada) for sharing their anaerobic chamber, funded by the Canadian Foundation for Innovation Leading Edge fund. We thank Steve Cox from Anaerobe Systems for providing the growth media used to catalog growth conditions. We thank Patrick Lavoie, Sudipta Saha, and Zhuyi Wang for their technical help in identifying a subset of the isolates. We thank Danica Chaharlangi for technical and language proofreading of the manuscript.

This work was supported by a team grant from the Canadian Institutes of Health Research (number 144628, to W.W.N. and A.M.), a Weston Family Microbiome Initiative Grant from the W. Garfield Weston Foundation (to W.W.N. and C.T.), a Grant from Crohn's and Colitis Foundation (625155 to C.T.), and National Institutes of Health grants RM1 GM135102 and R01 AI147023 (to K.C.H.). K.C.H. is a Chan Zuckerberg Biohub Investigator.

REFERENCES

- Hugon P, Dufour J-C, Colson P, Fournier P-E, Sallah K, Raoult D. 2015. A comprehensive repertoire of prokaryotic species identified in human beings. *Lancet Infect Dis* 15:1211–1219. [https://doi.org/10.1016/S1473-3099\(15\)00293-5](https://doi.org/10.1016/S1473-3099(15)00293-5).
- Browne HP, Forster SC, Anonye BO, Kumar N, Neville BA, Stares MD, Goulding D, Lawley TD. 2016. Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature* 533:543–546. <https://doi.org/10.1038/nature17645>.
- Forster SC, Kumar N, Anonye BO, Almeida A, Viciani E, Stares MD, Dunn M, Mkandawire TT, Zhu A, Shao Y, Pike LJ, Louie T, Browne HP, Mitchell AL, Neville BA, Finn RD, Lawley TD. 2019. A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nat Biotechnol* 37:186–192. <https://doi.org/10.1038/s41587-018-0009-7>.
- Tramontano M, Andrejev S, Pruteanu M, Klünemann M, Kuhn M, Galardini M, Jouhten P, Zelezniak A, Zeller G, Bork P, Typas A, Patil KR. 2018. Nutritional preferences of human gut bacteria reveal their metabolic idiosyncrasies. *Nat Microbiol* 3:514–522. <https://doi.org/10.1038/s41564-018-0123-9>.
- Geva-Zatorsky N, Sefik E, Kua L, Pasman L, Tan TG, Ortiz-Lopez A, Yanortsang TB, Yang L, Jupp R, Mathis D, Benoist C, Kasper DL. 2017. Mining the human gut microbiota for immunomodulatory organisms. *Cell* 168:928–943.E11. <https://doi.org/10.1016/j.cell.2017.01.022>.
- Nguyen TLA, Vieira-Silva S, Liston A, Raes J. 2015. How informative is the mouse for human gut microbiota research? *Dis Model Mech* 8:1–16. <https://doi.org/10.1242/dmm.017400>.
- Krych L, Hansen CHF, Hansen AK, van den Berg FWJ, Nielsen DS. 2013. Quantitatively different, yet qualitatively alike: a meta-analysis of the mouse core gut microbiome with a view towards the human gut microbiome. *PLoS One* 8:e62578. <https://doi.org/10.1371/journal.pone.0062578>.
- Hugenholtz F, de Vos WM. 2018. Mouse models for human intestinal microbiota research: a critical evaluation. *Cell Mol Life Sci* 75:149–160. <https://doi.org/10.1007/s00018-017-2693-8>.
- Chung H, Pamp SJ, Hill JA, Surana NK, Edelman SM, Troy EB, Reading NC, Villablanca EJ, Wang S, Mora JR, Umesaki Y, Mathis D, Benoist C, Relman DA, Kasper DL. 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>.
- Seedorf H, Griffin NW, Ridaura VK, Reyes A, Cheng J, Rey FE, Smith MI, Simon GM, Scheffrahn RH, Woebken D, Spormann AM, Van Treuren W, Ursell LK, Pirrung M, Robbins-Pianka A, Cantarel BL, Lombard V, Henrissat B, Knight R, Gordon JL. 2014. Bacteria from diverse habitats colonize and compete in the mouse gut. *Cell* 159:253–266. <https://doi.org/10.1016/j.cell.2014.09.008>.
- Franklin CL, Ericsson AC. 2017. Microbiota and reproducibility of rodent models. *Lab Anim (NY)* 46:114–122. <https://doi.org/10.1038/labani.1222>.
- Laukens D, Brinkman BM, Raes J, De Vos M, Vandenabeele P. 2016. Heterogeneity of the gut microbiome in mice: guidelines for optimizing experimental design. *FEMS Microbiol Rev* 40:117–132. <https://doi.org/10.1093/femsre/fuv036>.
- Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto J-M, Hansen T, Le Paslier D, Linneberg A, Nielsen HB, Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Li S, Jian M, Zhou Y, Li Y, Zhang X, Li S, Qin N, Yang H, Wang J, Brunak S, Doré J, Guarner F, Kristiansen K, Pedersen O, Parkhill J, Weissenbach J, MetaHIT Consortium, et al. 2010. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464:59–65. <https://doi.org/10.1038/nature08821>.
- Daniel H, Gholami AM, Berry D, Desmarchelier C, Hahne H, Loh G, Mondot S, Lepage P, Rothballer M, Walker A, Böhm C, Wenning M, Wagner M, Blaut M, Schmitt-Kopplin P, Kuster B, Haller D, Clavel T. 2014. High-fat diet alters gut microbiota physiology in mice. *ISME J* 8:295–308. <https://doi.org/10.1038/ismej.2013.155>.
- Parks DH, Rinke C, Chuvochina M, Chaumeil P-A, Woodcroft BJ, Evans PN, Hugenholtz P, Tyson GW. 2017. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat Microbiol* 2:1533–1542. <https://doi.org/10.1038/s41564-017-0012-7>.
- Pasolli E, Asnicar F, Manara S, Zolfo M, Karcher N, Armanini F, Beghini F, Manghi P, Tett A, Ghensi P, Collado MC, Rice BL, DuLong C, Morgan XC, Golden CD, Quince C, Huttenhower C, Segata N. 2019. Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. *Cell* 176:649–662.E20. <https://doi.org/10.1016/j.cell.2019.01.001>.
- Yang C, Mogno I, Contijoch EJ, Borgerding JN, Aggarwala V, Li Z, Siu S, Grasset EK, Helmus DS, Dubinsky MC, Mehandru S, Cerutti A, Faith JJ. 2020. Fecal IgA levels are determined by strain-level differences in *Bacteroides ovatus* and are modifiable by gut microbiota manipulation. *Cell Host Microbe* 27:467–475.E6. <https://doi.org/10.1016/j.chom.2020.01.016>.
- Lagkouvardos I, Pukall R, Abt B, Foesele BU, Meier-Kolthoff JP, Kumar N, Bresciani A, Martinez I, Just S, Ziegler C, Brugiroux S, Garzetti D, Wenning M, Bui TPN, Wang J, Hugenholtz F, Plugge CM, Peterson DA, Hornef MW, Baines JF, Smidt H, Walter J, Kristiansen K, Nielsen HB, Haller D, Overmann J, Stecher B, Clavel T. 2016. The Mouse Intestinal Bacterial Collection (miBC) provides host-specific insight into cultured diversity and functional potential of the gut microbiota. *Nat Microbiol* 1:16131. <https://doi.org/10.1038/nmicrobiol.2016.131>.
- Liu C, Zhou N, Du MX, Sun YT, Wang K, Wang YJ, Li DH, Yu HY, Song Y, Bai BB, Xin Y, Wu L, Jiang CY, Feng J, Xiang H, Zhou Y, Ma J, Wang J, Liu HW, Liu SJ. 2020. The Mouse Gut Microbial Biobank expands the coverage of cultured bacteria. *Nat Commun* 11:79. <https://doi.org/10.1038/s41467-019-13836-5>.
- Belkaid Y, Hand TW. 2014. Role of the microbiota in immunity and inflammation. *Cell* 157:121–141. <https://doi.org/10.1016/j.cell.2014.03.011>.
- Muniz LR, Knosp C, Yeretsian G. 2012. Intestinal antimicrobial peptides during homeostasis, infection, and disease. *Front Immunol* 3:310. <https://doi.org/10.3389/fimmu.2012.00310>.
- Soufli I, Toumi R, Rafa H, Touil-Boukoffa C. 2016. Overview of cytokines and nitric oxide involvement in immuno-pathogenesis of inflammatory bowel diseases. *World J Gastrointest Pharmacol Ther* 7:353–360. <https://doi.org/10.4292/wjgpt.v7.i3.353>.
- Wessling-Resnick M. 2010. Iron homeostasis and the inflammatory response. *Annu Rev Nutr* 30:105–122. <https://doi.org/10.1146/annurev.nutr.012809.104804>.
- Lugrin J, Rosenblatt-Velin N, Parapanov R, Liaudet L. 2014. The role of oxidative stress during inflammatory processes. *Biol Chem* 395:203–230. <https://doi.org/10.1515/hsz-2013-0241>.
- Arthur JC, Perez-Chanona E, Mühlbauer M, Tomkovich S, Uronis JM, Fan T-J, Campbell BJ, Abujamel T, Dogan B, Rogers AB, Rhodes JM, Stintzi A, Simpson KW, Hansen JJ, Keku TO, Fodor AA, Jobin C. 2012. Intestinal

- inflammation targets cancer-inducing activity of the microbiota. *Science* 338:120–123. <https://doi.org/10.1126/science.1224820>.
26. Lupp C, Robertson ML, Wickham ME, Sekirov I, Champion OL, Gaynor EC, Finlay BB. 2007. Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of Enterobacteriaceae. *Cell Host Microbe* 2:119–129. <https://doi.org/10.1016/j.chom.2007.06.010>.
 27. Zeng MY, Inohara N, Núñez G. 2017. Mechanisms of inflammation-driven bacterial dysbiosis in the gut. *Mucosal Immunol* 10:18–26. <https://doi.org/10.1038/mi.2016.75>.
 28. Borton MA, Sabag-Daigle A, Wu J, Solden LM, O'Banion BS, Daly RA, Wolfe RA, Gonzalez JF, Wysocki VH, Ahmer BMM, Wrighton KC. 2017. Chemical and pathogen-induced inflammation disrupt the murine intestinal microbiome. *Microbiome* 5:47. <https://doi.org/10.1186/s40168-017-0264-8>.
 29. Belcheva A, Irrazabal T, Robertson SJ, Streutker C, Maughan H, Rubino S, Moriyama EH, Copeland JK, Surendra A, Kumar S, Green B, Geddes K, Pezo RC, Navarre WW, Milosevic M, Wilson BC, Girardin SE, Wolever TMS, Edelmann W, Guttman DS, Philpott DJ, Martin A. 2014. Gut microbial metabolism drives transformation of *msh2*-deficient colon epithelial cells. *Cell* 158:288–299. <https://doi.org/10.1016/j.cell.2014.04.051>.
 30. Irrazabal T, Thakur BK, Kang M, Malaise Y, Streutker C, Wong EOY, Copeland J, Gryfe R, Guttman DS, Navarre WW, Martin A. 2020. Limiting oxidative DNA damage reduces microbe-induced colitis-associated colorectal cancer. *Nat Commun* 11:1802. <https://doi.org/10.1038/s41467-020-15549-6>.
 31. Su LK, Kinzler KW, Vogelstein B, Preisinger AC, Moser AR, Luongo C, Gould KA, Dove WF. 1992. Multiple intestinal neoplasia caused by a mutation in the murine homolog of the APC gene. *Science* 256:668–670. <https://doi.org/10.1126/science.1350108>.
 32. Keubler LM, Buettner M, Häger C, Bleich A. 2015. A multihit model: colitis lessons from the interleukin-10-deficient mouse. *Inflamm Bowel Dis* 21:1967–1975. <https://doi.org/10.1097/MIB.0000000000000468>.
 33. Reitmair AH, Cai JC, Bjerknes M, Redston M, Cheng H, Pind MT, Hay K, Mitri A, Bapat BV, Mak TW, Gallinger S. 1996. MSH2 deficiency contributes to accelerated APC-mediated intestinal tumorigenesis. *Cancer Res* 56:2922–2926.
 34. Sohn K-J, Choi M, Song J, Chan S, Medline A, Gallinger S, Kim Y-I. 2003. Msh2 deficiency enhances somatic Apc and p53 mutations in Apc+/- Msh2-/- mice. *Carcinogenesis* 24:217–224. <https://doi.org/10.1093/carcin/24.2.217>.
 35. Meier-Kolthoff JP, Klenk HP, Göker M. 2014. Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. *Int J Syst Evol Microbiol* 64:352–356. <https://doi.org/10.1099/ijs.0.056994-0>.
 36. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14:60. <https://doi.org/10.1186/1471-2105-14-60>.
 37. Auch AF, von Jan M, Klenk HP, Göker M. 2010. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci* 2:117–134. <https://doi.org/10.4056/signs.531120>.
 38. Rodríguez-R LM, Gunturu S, Harvey WT, Rosselló-Mora R, Tiedje JM, Cole JR, Constantinidis KT. 2018. The Microbial Genomes Atlas (MiGA) web-server: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. *Nucleic Acids Res* 46:W282–W288. <https://doi.org/10.1093/nar/gky467>.
 39. Ormerod KL, Wood DLA, Lachner N, Gellatly SL, Daly JN, Parsons JD, Dal'Molin CGO, Palfreyman RW, Nielsen LK, Cooper MA, Morrison M, Hansbro PM, Hugenholtz P. 2016. Genomic characterization of the uncultured Bacteroidales family S24-7 inhabiting the guts of homeothermic animals. *Microbiome* 4:36. <https://doi.org/10.1186/s40168-016-0181-2>.
 40. Lagkouvardos I, Lesker TR, Hitch TCA, Gálvez EJC, Smit N, Neuhaus K, Wang J, Baines JF, Abt B, Stecher B, Overmann J, Strowig T, Clavel T. 2019. Sequence and cultivation study of *Muribaculaceae* reveals novel species, host preference, and functional potential of this yet undescribed family. *Microbiome* 7:28. <https://doi.org/10.1186/s40168-019-0637-2>.
 41. Salzman NH, de Jong H, Paterson Y, Harmsen HJM, Welling GW, Bos NA. 2002. Analysis of 16S libraries of mouse gastrointestinal microflora reveals a large new group of mouse intestinal bacteria. *Microbiology (Reading)* 148:3651–3660. <https://doi.org/10.1099/00221287-148-11-3651>.
 42. Miyake S, Ding Y, Soh M, Low A, Seedorf H. 2020. *Muribaculum gordoncarteri* sp. nov., an anaerobic bacterium from the faeces of C57BL/6J mice. *Int J Syst Evol Microbiol* 70:4725–4729. <https://doi.org/10.1099/ijsem.0.004338>.
 43. Meehan CJ, Beiko RG. 2014. A phylogenomic view of ecological specialization in the Lachnospiraceae, a family of digestive tract-associated bacteria. *Genome Biol Evol* 6:703–713. <https://doi.org/10.1093/gbe/evu050>.
 44. Lau JT, Whelan FJ, Herath I, Lee CH, Collins SM, Bercik P, Surette MG. 2016. Capturing the diversity of the human gut microbiota through culture-enriched molecular profiling. *Genome Med* 8:1–10. <https://doi.org/10.1186/s13073-016-0327-7>.
 45. Kaakoush NO. 2015. Insights into the role of Erysipelotrichaceae in the human host. *Front Cell Infect Microbiol* 5:84–84. <https://doi.org/10.3389/fcimb.2015.00084>.
 46. Morotomi M. 2014. The family Sutterellaceae, p 1005–1012. In Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (ed), *The prokaryotes: Alphaproteobacteria and Betaproteobacteria*. Springer, Berlin, Germany.
 47. Ju T, Kong JY, Stothard P, Willing BP. 2019. Defining the role of *Parasutterella*, a previously uncharacterized member of the core gut microbiota. *ISME J* 13:1520–1534. <https://doi.org/10.1038/s41396-019-0364-5>.
 48. Micheli MR, Bova R, Pascale E, D'Ambrosio E. 1994. Reproducible DNA fingerprinting with the random amplified polymorphic DNA (RAPD) method. *Nucleic Acids Res* 22:1921–1922. <https://doi.org/10.1093/nar/22.10.1921>.
 49. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, 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>.
 50. Vieira-Silva S. 2016. Species-function relationships shape ecological properties of the human gut microbiome. *Nat Microbiol* 1:16088. <https://doi.org/10.1038/nmicrobiol.2016.88>.
 51. Salyers AA, West SEH, Vercellotti JR, Wilkins TD. 1977. Fermentation of mucins and plant polysaccharides by anaerobic bacteria from the human colon. *Appl Environ Microbiol* 34:529–533. <https://doi.org/10.1128/aem.34.5.529-533.1977>.
 52. Gänzle MG, Follador R. 2012. Metabolism of oligosaccharides and starch in Lactobacilli: a review. *Front Microbiol* 3:340. <https://doi.org/10.3389/fmicb.2012.00340>.
 53. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZY) in 2013. *Nucleic Acids Res* 42:D490–D495. <https://doi.org/10.1093/nar/gkt1178>.
 54. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y, Yin Y. 2018. DbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res* 46:W95–W101. <https://doi.org/10.1093/nar/gky418>.
 55. Tiso M, Schechter AN. 2015. Nitrate reduction to nitrite, nitric oxide and ammonia by gut bacteria under physiological conditions. *PLoS One* 10:e0127490. <https://doi.org/10.1371/journal.pone.0127490>.
 56. Sparacino-Watkins C, Stolz JF, Basu P. 2014. Nitrate and periplasmic nitrate reductases. *Chem Soc Rev* 43:676–706. <https://doi.org/10.1039/c3cs60249d>.
 57. Rocha BS, Laranjinha J. 2020. Nitrate from diet might fuel gut microbiota metabolism: minding the gap between redox signaling and inter-kingdom communication. *Free Radic Biol Med* 149:37–43. <https://doi.org/10.1016/j.freeradbiomed.2020.02.001>.
 58. Winter SE, Winter MG, Xavier MN, Thiennimitr P, Poon V, Keestra AM, Laughlin RC, Gomez G, Wu J, Lawhon SD, Popova IE, Parikh SJ, Adams LG, Tsolis RM, Stewart VJ, Bäuml AJ. 2013. Host-derived nitrate boosts growth of *E. coli* in the inflamed gut. *Science* 339:708–711. <https://doi.org/10.1126/science.1232467>.
 59. Lesker TR, Durairaj AC, Gálvez EJC, Lagkouvardos I, Baines JF, Clavel T, Sczyrba A, McHardy AC, Strowig T. 2020. An integrated metagenome catalog reveals new insights into the murine gut microbiome. *Cell Rep* 30:2909–2922.E6. <https://doi.org/10.1016/j.celrep.2020.02.036>.
 60. Eberl C, Ring D, Münch PC, Beutler M, Basic M, Slack EC, Schwarzer M, Srutkova D, Lange A, Frick JS, Bleich A, Stecher B. 2019. Reproducible colonization of germ-free mice with the oligo-mouse-microbiota in different animal facilities. *Front Microbiol* 10:2999–2915. <https://doi.org/10.3389/fmicb.2019.02999>.
 61. Zhou Y, Chen H, He H, Du Y, Hu J, Li Y, Li Y, Zhou Y, Wang H, Chen Y, Nie Y. 2016. Increased *Enterococcus faecalis* infection is associated with clinically active Crohn disease. *Medicine (Baltimore)* 95:e5019. <https://doi.org/10.1097/MD.0000000000005019>.
 62. Gevers D, Kugathasan S, Denson LA, Vázquez-Baeza Y, Van Treuren W, Ren B, Schwager E, Knights D, Song SJ, Yassour M, Morgan XC, Kostic AD, Luo C, González A, McDonald D, Haberman Y, Walters T, Baker S, Rosh J, Stephens M, Heyman M, Markowitz J, Baldassano R, Griffiths A, Sylvester F, Mack D, Kim S, Crandall W, Hyams J, Huttenhower C, Knight R, Xavier RJ. 2014. The treatment-naïve microbiome in new-onset Crohn's disease. *Cell Host Microbe* 15:382–392. <https://doi.org/10.1016/j.chom.2014.02.005>.

63. Lee IA, Kim DH. 2011. *Klebsiella pneumoniae* increases the risk of inflammation and colitis in a murine model of intestinal bowel disease. *Scand J Gastroenterol* 46:684–693. <https://doi.org/10.3109/00365521.2011.560678>.
64. Chen Y-J, Wu H, Wu S-D, Lu N, Wang Y-T, Liu H-N, Dong L, Liu T-T, Shen X-Z. 2018. *Parasutterella*, in association with irritable bowel syndrome and intestinal chronic inflammation. *J Gastroenterol Hepatol* 33:1844–1852. <https://doi.org/10.1111/jgh.14281>.
65. Elinav E, Strowig T, Kau AL, Henao-Mejia J, Thaiss CA, Booth CJ, Peaper DR, Bertin J, Eisenbarth SC, Gordon JI, Flavell RA. 2011. NLRP6 inflammasome regulates colonic microbial ecology and risk for colitis. *Cell* 145:745–757. <https://doi.org/10.1016/j.cell.2011.04.022>.
66. Hu B, Elinav E, Huber S, Strowig T, Hao L, Hafemann A, Jin C, Wunderlich C, Wunderlich T, Eisenbarth SC, Flavell RA. 2013. Microbiota-induced activation of epithelial IL-6 signaling links inflammasome-driven inflammation with transmissible cancer. *Proc Natl Acad Sci U S A*. 110:9862–9867. <https://doi.org/10.1073/pnas.1307571110>.
67. Palm NW, de Zoete MR, Cullen TW, Barry NA, Stefanowski J, Hao L, Degnan PH, Hu J, Peter I, Zhang W, Ruggiero E, Cho JH, Goodman AL, Flavell RA. 2014. Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. *Cell* 158:1000–1010. <https://doi.org/10.1016/j.cell.2014.08.006>.
68. Chen W, Liu F, Ling Z, Tong X, Xiang C. 2012. Human intestinal lumen and mucosa-associated microbiota in patients with colorectal cancer. *PLoS One* 7:e39743. <https://doi.org/10.1371/journal.pone.0039743>.
69. Craven M, Egan CE, Dowd SE, McDonough SP, Dogan B, Denkers EY, Bowman D, Scherl EJ, Simpson KW. 2012. Inflammation drives dysbiosis and bacterial invasion in murine models of ileal Crohn's Disease. *PLoS One* 7:e41594. <https://doi.org/10.1371/journal.pone.0041594>.
70. Schaubeck M, Clavel T, Calasan J, Lagkouvardos I, Haange SB, Jehmlich N, Basic M, Dupont A, Hornef M, von Bergen M, Bleich A, Haller D. 2016. Dysbiotic gut microbiota causes transmissible Crohn's disease-like ileitis independent of failure in antimicrobial defence. *Gut* 65:225–237. <https://doi.org/10.1136/gutjnl-2015-309333>.
71. Dinh DM, Volpe GE, Duffalo C, Bhalchandra S, Tai AK, Kane AV, Wanke CA, Ward HD. 2015. Intestinal microbiota, microbial translocation, and systemic inflammation in chronic HIV infection. *J Infect Dis* 211:19–27. <https://doi.org/10.1093/infdis/jiu409>.
72. Zhao Y, Wu J, Li JV, Zhou N-Y, Tang H, Wang Y. 2013. Gut microbiota composition modifies fecal metabolic profiles in mice. *J Proteome Res* 12:2987–2999. <https://doi.org/10.1021/pr400263n>.
73. Zhu Q, Jin Z, Wu W, Gao R, Guo B, Gao Z, Yang Y, Qin H. 2014. Analysis of the intestinal lumen microbiota in an animal model of colorectal cancer. *PLoS One* 9:e90849. <https://doi.org/10.1371/journal.pone.0090849>.
74. Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR. 2007. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci U S A* 104:13780–13785. <https://doi.org/10.1073/pnas.0706625104>.
75. Baumgart M, Dogan B, Rishniw M, Weitzman G, Bosworth B, Yantiss R, Orsi RH, Wiedmann M, McDonough P, Kim SG, Berg D, Schukken Y, Scherl E, Simpson KW. 2007. Culture independent analysis of ileal mucosa reveals a selective increase in invasive *Escherichia coli* of novel phylogeny relative to depletion of Clostridiales in Crohn's disease involving the ileum. *ISME J* 1:403–418. <https://doi.org/10.1038/ismej.2007.52>.
76. Yao J, Carter RA, Vuagniaux G, Barbier M, Rosch JW, Rock CO. 2016. A pathogen-selective antibiotic minimizes disturbance to the microbiome. *Antimicrob Agents Chemother* 60:4264–4273. <https://doi.org/10.1128/AAC.00535-16>.
77. Choo JM, Kanno T, Zain NMM, Leong LEX, Abell GCJ, Keeble JE, Bruce KD, Mason AJ, Rogers GB. 2017. Divergent relationships between fecal microbiota and metabolome following distinct antibiotic-induced disruptions. *mSphere* 2:e00005-17. <https://doi.org/10.1128/mSphere.00005-17>.
78. Goodman AL, Kallstrom G, Faith JJ, Reyes A, Moore A, Dantas G, Gordon JI. 2011. Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. *Proc Natl Acad Sci U S A* 108:6252–6257. <https://doi.org/10.1073/pnas.1102938108>.
79. Sibley CD, Grinwis ME, Field TR, Parkins MD, Norgaard JC, Gregson DB, Rabin HR, Surette MG. 2010. McKay agar enables routine quantification of the '*Streptococcus milleri*' group in cystic fibrosis patients. *J Med Microbiol* 59:534–540. <https://doi.org/10.1099/jmm.0.016592-0>.
80. Sutter VL, Sugihara PT, Finegold SM. 1971. Rifampin-blood-agar as a selective medium for the isolation of certain anaerobic bacteria. *Appl Microbiol* 22:777–780. <https://doi.org/10.1128/am.22.5.777-780.1971>.
81. Muyzer G, Teske A, Wirsén CO, Jannasch HW. 1995. Phylogenetic relationships of *Thiomicrospira* species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S rDNA fragments. *Arch Microbiol* 164:165–172. <https://doi.org/10.1007/BF02529967>.
82. Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y, Seo H, Chun J. 2017. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 67:1613–1617. <https://doi.org/10.1099/ijsem.0.001755>.
83. Letunic I, Bork P. 2019. Interactive Tree of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 47:W256–W259. <https://doi.org/10.1093/nar/gkz239>.
84. Wattam AR, Davis JJ, Assaf R, Boisvert S, Brettin T, Bun C, Conrad N, Dietrich EM, Disz T, Gabbard JL, Gerdes S, Henry CS, Kenyon RW, Machi D, Mao C, Nordberg EK, Olsen GJ, Murphy-Olson DE, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Vonstein V, Warren A, Xia F, Yoo H, Stevens RL. 2017. Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. *Nucleic Acids Res* 45:D535–D542. <https://doi.org/10.1093/nar/gkw1017>.
85. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshtik AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
86. Nikolenko SI, Korobeynikov AI, Alekseyev MA. 2013. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genomics* 14:S7. <https://doi.org/10.1186/1471-2164-14-S1-S7>.
87. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUASt: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>.
88. Tatusova T, DiCuccio M, Badretdin A, Chetverin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
89. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
90. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia J-M, Chia J-M, Colasanti R, Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Jain R, et al. 2018. KBase: the United States Department of Energy systems biology knowledgebase. *Nat Biotechnol* 36:566–569. <https://doi.org/10.1038/nbt.4163>.
91. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol E* 33:1870–1874. <https://doi.org/10.1093/molbev/msw054>.
92. Yoon SH, Min Ha S, Lim J, Kwon S, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 110:1281–1286. <https://doi.org/10.1007/s10482-017-0844-4>.
93. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2016. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* 44:D457–D462. <https://doi.org/10.1093/nar/gkv1070>.
94. Waskom M. 2021. seaborn: statistical data visualization (version JOSS paper). <https://doi.org/10.21105/joss.03021>.
95. Hunter JD. 2007. Matplotlib: a 2D graphics environment. *Comput Sci Eng* 9:90–95. <https://doi.org/10.1109/MCSE.2007.55>.
96. Darzi Y, Falony G, Vieira-Silva S, Raes J. 2016. Towards biome-specific analysis of meta-omics data. *ISME J* 10:1025–1028. <https://doi.org/10.1038/ismej.2015.188>.
97. Lesker TR, Chakravarthy A, Galvez JC, Lagkouvardos I, Baines JF, Clavel T, Sczyrba A, McHardy AC, Strowig T. 2019. An integrated metagenome catalog reveals novel insights into the murine gut microbiome. <https://doi.org/10.1101/528737>.