DISSERTATION

MULTI-OMICS INVESTIGATION OF INTERACTIONS BETWEEN PERSISTENT BACTERIA AND *SALMONELLA* IN THE INFLAMED GUT

Submitted by

Ikaia Leleiwi

Graduate Degree Program in Cell and Molecular Biology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2023

Doctoral Committee:

Advisor: Kelly C. Wrighton

Jessica Prenni
Erika Szymanski
Tiffany Weir

ABSTRACT

MULTI-OMICS INVESTIGATION OF INTERACTIONS OF PERSISTENT BACTERIA AND *SALMONELLA* IN THE INFLAMED GUT

Salmonella is a globally relevant enteric pathogen responsible for numerous outbreaks and debilitating illness yearly. Expansive tropism allows *Salmonella* to find bastion in zoonotic reservoirs including prominent food animals. Continued prophylactic antibiotic use in livestock and therapeutic antibiotic use in humans has increased selection for multi-drug resistant *Salmonella* varieties. Most of the current research on *Salmonella* enteric disease is performed absent complete native gut microbiota. Further, common murine models that could facilitate study of *Salmonella* in a robust community setting lack model-specific microbiome resources to accomplish the feat. Presented in this dissertation is a comprehensive catalogue of CBA/J mouse gut microbial genomes created as a resource for the research community. The genome database was used to recruit various omics data types to expand the current knowledge of *Salmonella* infection in a complex community setting, identifying community members robust to inflammation and with potential to further explore as probiotics.

In Chapter 1, I review the current state of *Salmonella* pathogenesis in the context of the gut microbiome. The focus here is to survey the literature for prominent *Salmonella* mechanisms of infection and how they relate to both host and commensal microbes. I explore host responses to *Salmonella* and microbial metabolites capable of affecting *Salmonella* pathogenesis. This microbiome-centric take on *Salmonella* infection implies a need for comprehensive methods to examine microbes and their processes *in vivo*, including queries of genes and gene products. A special emphasis on multi-omics approaches is mentioned in this section as powerful tools to holistically study the complete *Salmonella*-included gut microbiome and to address deficiencies in prior work, ultimately providing more translatable results impacting human health.

Chapter 2 outlines the creation of the CBAJ-DB – a first of its kind bacteria and virus genome collection produced from the gut communities of *Salmonella* infected and uninfected CBA/J mice. Relevance of this work to *Salmonella* research is explained, emphasizing the CBA/J model advantages to study enteric infection in unperturbed gut communities. Robust genome recovery from deep sequencing yielded over 2,000 bacterial metagenome-assembled genomes including novel bacteria strains and taxa with implications for other mouse breeds and human microbiomes. Viral genomes reconstructed from metagenomic sequencing were linked to bacteria hosts and mined for genes germane to bacteria function. The complete functional potential of the CBA/J gut community in infected and uninfected mice was also explored, detailing a decrease in immune-modulatory functional potential following *Salmonella* infection, and implying a potentially important role of *Alistipes* sp. in butyrate production. Importantly, work from this chapter provides the infrastructure for genome-resolved multi-omics investigations detailed in Chapter 3 that are critical to determine functional links between *Salmonella* and the commensal microbiota.

In Chapter 3 additional metagenomic sequencing is combined with the CBAJ-DB and used to recruit metatranscriptomic and metabolomic data from infected and uninfected CBA/J mice. We reveal expression and metabolites that implicate numerous commensal bacteria with the flow of sulfur in the inflamed intestine, making it available for host oxidation to tetrathionate in support of *Salmonella* anaerobic respiration. Current dogma surrounding *Salmonella* lactate utilization from the host is also confronted by our data, which implies potential cross feeding on microbially derived D-lactate by *Salmonella* during peak infection. These expression data are supported by random forest and logistic regression modeling which determined genes for D-lactate production or utilization are important to *Salmonella*-association of other bacteria in the inflamed gut. Relatively abundant bacteria observed in Chapter 2 were confirmed to be active in infected communities and to be expressing genes relevant to *Salmonella* processes like chitinase, lactate dehydrogenase, and sulfatase. Not only does this chapter illustrate the utility of the CBAJ-DB but it highlights how multi-omics investigation in complete

ecosystems can unveil results that may be different than claims made based on *in vitro* or reduced community *in vivo* studies.

The final chapter presented here summarizes the key findings from Chapters 2 and 3 and offers avenues for future research including specific strain isolation from infected communities and subsequent *Salmonella* competition experiments to determine probiotic therapeutic potential. This dissertation aims to (1) Examine the diversity of the CBA/J mouse gut and provide a genomic resource to the microbiome community, (2) using various omics techniques, discover interactions between *Salmonella* and commensal bacteria that could impact pathogenesis, and (3) identify members of the inflamed community with probiotic potential that are indifferent to *Salmonella* or that display niche overlap for substrate competition with *Salmonella*. Ultimately, this dissertation provides a comprehensive examination of *Salmonella* infection amidst a whole and robust microbiome identifying important membership in the inflamed community and linking autochthonous processes with pathogenic ones to better understand *Salmonella* enteric disease.

ACKNOWLEDGEMENTS

My dissertation work presented in this document would not have been possible without the assiduous effort of members from the Wrighton, Wilkins, Ahmer, and Wysocki labs. Beyond the professional and technical roles instrumental to this science, my support network of family and friends kept me mentally strong during the years of study to gain my Ph.D. Reflecting on my life between 2019 and now (2023), I shudder to think where I'd be without the catharsis these people provided. In acknowledgment of the times, here are some events that occurred during my Ph.D. – a global pandemic, I got married, the birth of my first child, multiple deaths in the family, life changing medical diagnoses, record breaking forest fires, an insurrection in Washington D.C., William Shatner went to space, ChatGPT passed the bar exam, England crowned a king, and Putin started a war.

Dr. Kelly Wrighton, I am eternally grateful for the opportunity you gave me by inviting me to join your lab. The mentorship you provided has forever changed my life for the better. How can I ever thank you for seeing my potential all those years ago? The interactions I've had with the fine humans you chose to be in your lab have shaped me as a scientist and an individual. You're a wonderful leader and I'm so lucky to have met you. You taught me to believe in myself, Kelly. Thank you.

Dr. Carol Wilusz thank you for giving me a chance. You pushed me past my comfort zone and expected me to do more than the minimum, always. Your strength and leadership are the keystones of the CMB program, and you hold us all up. Thank you, Carol, for your advocacy, your drive, and your hustle.

To my wonderful committee, I'm so grateful I chose each one of you, and that you let me be part of your busy schedules these past years. **Dr. Jessica Prenni**, **Dr. Erika Szymanski**, and **Dr. Tiffany Weir** thank you for your time and your support during my PhD.

I'd like to thank the members of the **Wrighton and Wilkins labs.** These people are just the best. **Dr. Josue Rodriguez-Ramos** thank you for immediately being my friend and for always checking in on me when my life became stressful. **Dr. Kayla Borton** thank you for your mentorship and your example. You set a high bar in the lab and our endeavor to meet it is part of what makes this lab so great. **Dr.**

Michael Shaffer you are an inspiration and a talented bioinformatician and I'm lucky to have learned from you, thank you. **Katherine Kokkinias** thank you for your friendship and for always making yourself available to help me or to brainstorm with me about *Salmonella*. **Dr. Brian Ahmer** and **Dr. Vicki Wysocki** thank you both for your time and your student's time helping to make the *Salmonella* project so great.

Clint and **Mary Leleiwi**, thanks for being my parents. Thanks for watching Ronan and for always believing in me. **Lisa Leleiwi** thanks for being my best friend and my wife and for making Ronan. **Ronan Leleiwi** thanks for being goofy and most of all thank you for reminding me of my ambition.

DEDICATION

This dissertation is dedicated

To Ronan Kaluaiko'olau Clinton Leleiwi who makes me look forward to everyday

TABLE OF CONTENTS

Chapter 1: Introduction

1.1 Salmonella Pathogenesis

Non-typhoidal *Salmonella* (NTS), including *Salmonella enterica* serovar Typhimurium (hereon referred to as *Salmonella*), are a significant source of diarrheal illness globally, causing more than 50,000 deaths and 95 million cases of enterocolitis in 2017[1]. NTS tropism spans a wide range of animal hosts including reptiles, birds, and various livestock animals[2–4]. Preemptive antibiotic administration to food animals continues to be widely used to boost yield despite its contribution to the prevalence of antibiotic resistant *Salmonella* serovars in recent years[5,6]. Human infections are treated with fluoroquinolones, cephalosporins, and macrolides while also contributing to *Salmonella* antimicrobial resistance[7]. Aggressive antibiotic use in humans and common zoonotic reservoirs of *Salmonella* are instrumental to increased rise in multi-drug-resistant serovars — highlighting a need for alternative methods of treatment for bacterial pathogens including *Salmonella*[8]. Various biotic interventions to control *Salmonella* have been proposed that reduce the reliance on antibiotics for treatment. These include probiotics, prebiotics, synbiotics, phytobiotics, postbiotics, vaccines, and phage therapy[9,10]. Efficacy of many of these approaches depends on the therapy remaining effective during *Salmonella*-induced enteric inflammation and amidst a consortium of resident microbiota. It is for these reasons it is crucial to investigate *Salmonella* pathogenesis within the context of a robust and complete microbiome.

Salmonella infection in humans most often occurs after consuming contaminated food products like meat or dairy[11]. Once consumed, *Salmonella* persists through the upper gastrointestinal tract (GI) escaping stomach acidity by shielding in food or via the activation of an acid tolerance response[12]. *Salmonella* next travels to the distal ileum and initiates infection of host epithelia with a Type Three Secretion System I (T3SS-I), genes for which reside on *Salmonella* Pathogenicity Island 1 (SPI-1)[12]. Attachment to host cells is facilitated by various *Salmonella* fimbriae and the induction of SPI-1 genes is triggered by specific environment triggers in the distal ileum and cecum[12,13]. T3SS-I facilitates S*almonella* invasion of gut epithelium and initial suppression of macrophage proinflammatory

responses[14]. Many environmental modulators of SPI-1 expression are microbial metabolites like short chain fatty acids (SCFA), a subclass of fatty acids produced from microbial polysaccharide fermentation all which contain fewer than 6 carbons[12,15]. SCFAs are key microbial metabolites in the gut, nourishing gut epithelia and contributing to mucosal immunity and agonism of host signaling pathways involved in central nervous system function, endocrine function, and cellular metabolism[16,17]. SCFAs propionate and butyrate and long chain fatty acids either directly or indirectly affect araC-like transcriptional regulator HilD which controls the expression of *hilA* SPI-1 transcriptional regulator[18]. The SCFAs acetate and formate have been shown to activate *hilD* translation, potentiating the regulatory cascade for SPI-1 expression, and it is thought the relative concentrations of these and other microbially produced SCFAs are important signals dictating where *Salmonella* initiates infection in the GI[18]. Beyond SCFAs, researchers observed reduction in *hilA* expression and subsequent downregulation of SPI-1 following indole treatment as well[19]. Initial suppression of macrophage proinflammatory cytokine expression is necessary for *Salmonella* intracellular life stages, but induction of inflammation in non-phagocytic cells via T3SS-I is conversely critical to luminal expansion and planktonic *Salmonella* proliferation[14,20]. Clearly, *Salmonella* is not agnostic to the influence of microbial metabolites, and in fact it may depend on them as important signals in the GI to initiate and sustain infection. Because of this, there is a need to examine the inflamed metabolome during infection to truly understand how cross feeding and microbial metabolites can impact *Salmonella* enteric disease. In this dissertation, I used metabolomics in concert with other omics types in this dissertation to interrogate microbial interactions in the *Salmonella*-included GI and to investigate some of the metabolites mentioned above that are known to affect SPI-1 regulation.

Salmonella is a facultative intracellular pathogen capable of T3SS-I-mediated invasion of non-phagocytic cells but also is well adapted to an intracellular lifestyle within phagosomes or *Salmonella*-containing vacuoles (SCVs)[12]. Type 3 Secretion System II (T3SS-II) genes encoded on *Salmonella* Pathogenicity Island 2 (SPI-2) are activated during intracellular *Salmonella* infection and produce effectors to control host immune responses and SCV maintenance and movement within the cell[21]. The T3SS-II modification of SCVs allow for *Salmonella* intracellular replication[22]. Furthermore, T3SS-II

effectors can disrupt host cell metabolism, shifting it from oxidative phosphorylation to aerobic glycolysis and inducing an accumulation of citric acid cycle constituents including succinate[22]. Consequently, T3SS-II is further upregulated via SPI-2 activation by host succinate uptake[22]. *Salmonella* systemic infection results from successful intracellular proliferation of *Salmonella*, particularly within Th2 macrophages that can spread the infection to mesenteric lymph nodes and the spleen[12,23]. Crucial to *Salmonella* replication within SCVs is iron, and in mouse models systemic *Salmonella* infection can be avoided by Nramp1/Slc11a1 competency which increases lipocalin-2 expression and limits *Salmonella* iron acquisition intracellularly[12,21,24]. Importantly, lipocalin-2 is used as a measure of inflammation severity in Chapter 2 of this dissertation, to correlate *Salmonella* relative abundance with GI inflammation. Gut microbiome research in the context of *Salmonella* infection has embraced Nramp1$^{+/+}$ mouse models like CBA, sv129S6, and C3H/HeN to avoid systemic infections and observe prolonged gut-localized *Salmonella* challenge in whole microbial communities[12,25,26]. In this dissertation, I develop a first of its kind microbial genomic catalog of the gut microbiome in CBA mice and use it to investigate many of the metabolite mediators of *Salmonella* T3SS-I and T3SS-II previously discussed. The resultant gut microbial database was published as a resource for other microbiome researchers using the CBA/J mouse model and its utility to recruit other omics types is illustrated in Chapter 3.

1.2 Salmonella and the microbiome

A healthy GI is host to a diverse consortium of bacteria unique to each individual[27]. The human GI is colonized by bacteria which exist in a synergy between host and microbiota that persists a lifetime[28]. Host diet, immunity, and physiology impact the composition of the gut microbiota, and the microbial community structure of the GI can be altered by changes to these host factors[20,27,29]. The gut microbiota contribute to host nutrition via fermentation of fiber to SCFAs[16,28]. Anaerobic bacteria in the gut make up the majority of the community, and while *Clostridia* are often considered the dominant butyrate producing bacteria, other prominent classes responsible for SCFA production include *Bacteroidia* and

Bacilli[30,31]. Mucin degrading microbiota, like *Akkermansia muciniphila*, live in close proximity with the gut epithelia and can stimulate mucin production and increase gut barrier integrity of the host[32].

The barrier between the luminal space and gut epithelium is critical for the maintenance of an anaerobic environment in the lumen and for immune homeostasis[28]. By physically separating the epithelia from immune triggers like pathogen and microbial associated molecular patterns, the mucosal barrier limits innate immune stimulation via NOD-like receptor (NLR) and toll-like receptor (TLR) binding[28,33]. Microbial derived SCFAs, particularly butyrate, also contribute to gut barrier maintenance by increasing tight junction integrity[27,34]. Permeability of the gut is consequently associated with various gut pathologies including inflammatory bowel disease[16,34]. *Salmonella* is known to affect gut barrier integrity via SpvB virulence factor redistribution of apical junctional complex proteins between epithelial cells, further antagonizing the host inflammatory response during enteric infection[35].

Commensal bacteria in the gut also play an important role in colonization resistance against *Salmonella* and the ability of *Salmonella* to alter the luminal environment and potentiate proliferation[20,28,29,36]. A key factor to pathogen success during *Salmonella* enteritis is an induction of the host inflammation response, underpinning the importance of commensal mechanisms of host immune modulation either via microbially derived metabolites or direct priming of the immune response as seen with the administration of various *Lactobacillus* probiotic species[31,37]. These mechanisms may become instrumental in future probiotic control of *Salmonella* enteric infection. Experimental designs including whole community microbiomes to discover potential bacteria for probiotic therapeutic interventions help to ensure the viability of any new probiotic strains amidst full microbial consortia. Careful consideration of potentially probiotic bacteria identified in the CBA/J inflamed gut communities is presented in this dissertation. The work done here identified multiple *Bacilli* bacteria, including *Lactobacillus*, that persist during inflammation and may participate in cross feeding with *Salmonella,* showing how even bacteria generally recognized as safe harbor potential to indirectly exacerbate certain enteric infections. Bacteria with implications to human health like *Akkermansia* and *Enterococcus* are identified as prominent

members of the CBA/J post-infection community as well, and their role as both potential probiotics and potential pathobionts are considered[38,39].

Consequences of *Salmonella* infection on the gut microbiome include an increased redox potential in the lumen and subsequent niche alteration that drastically impacts the community structure of the microbiota[40]. Interactions between host pattern recognition receptors and *Salmonella* initiate the innate immune response[41]. *Salmonella* lipopolysaccharide and flagellin can agonize TLRs and NLRs triggering interleukin (IL) IL-23/IL-17 neutrophil recruitment and IL-23/IL-22 T cell response[41]. Induction of nitric oxide synthase (iNOS) by IL-22 and chemokine recruitment of neutrophils results in reactive nitrogen species (RNS) and reactive oxygen species (ROS) in the infected environment respectively[41–43]. T3SS-I effectors can further promote the inflammatory response as in the case of SopE-mediated increased expression of iNOS and caspase-1 inflammasome activation [41,44]. This induction of host inflammation provides multiple terminal electron acceptors that favor *Salmonella* anaerobic respiration allowing *Salmonella* to proliferate in the lumen and outcompete resident obligate fermenting members of the gut community[20,44]. Hydrogen sulfide mitochondrial oxidation to thiosulfate provides inorganic sulfur for further oxidation to tetrathionate by ROS during inflammation[45]. *Salmonella* use of tetrathionate as a terminal electron acceptor for anaerobic respiration is an important mechanism during early infection for *Salmonella* success overcoming colonization resistance[45]. Later, nitrate formed by isomerization of peroxynitrite originating from NOS and ROS interaction makes available the preferred electron acceptor (NO_3^-) for *Salmonella* respiration . *Salmonella* also uses molecular oxygen as a terminal electron acceptor when respiring hydrogen, a microbial fermentation product[47–49]. *Salmonella* use of abundant commensal bacteria metabolites like hydrogen, but also 1-2 propanediol, propionate, and succinate provide immediate niche occupation critical to success in the lumen when *Salmonella* initially infects[48,50–52].

Broadly, prior to the start of my Ph.D., it was recognized that *Salmonella* expansion in the gut and subsequent host inflammation alters gut community composition characterized by a reduction in *Clostridia* and persistence of *Enterobacteriaceae* and some *Bacilli*[40,53]. It was also widely accepted that host inflammation created electron acceptors that *Salmonella* can use to overcome commensal

colonization resistance. The studies mentioned above briefly review these processes, however none adequately consider confounding factors inherent with *Salmonella* infection within a robust microbiome. I address the knowledge gap left by these studies by using multi-omic approaches to test *Salmonella* infection with a CBA/J model including an unperturbed microbiome. Furthermore, I explore the involvement of persistent microbiota during inflammation in *Salmonella* anaerobic respiration of tetrathionate and the wider importance of sulfur reduction to gut community resilience to inflammatory perturbation.

1.3 Multi-omics to study Salmonella

The majority of prior work done to elucidate many of the aforementioned mechanisms of *Salmonella* pathogenesis and metabolism were performed *in vitro* or in reduced community mice or mice first treated with antibiotics, lacking a robust and complete microbiome[54–57]. Studies that did consider complete microbiomes used CBA mice and did not comprehensively examine the role of the microbiota in processes like *Salmonella* nitrate and tetrathionate respiration or succinate, propionate, and 1,2-propanodiol utilization[50–52,58,59]. A common mouse model used in *Salmonella* research is C57BL/6, requiring antibiotic perturbation to foster the level of cecal inflammation necessary to potentiate *Salmonella* luminal expansion and which lacks NRAMP1 competency[50]. Swiss Webster mice and C57BL/6 mice were used in experiments outlining *Salmonella* lactate utilization[60]. However, despite an extensive series of experiments with *Salmonella* lactate dehydrogenase knockouts in whole community Swiss Webster mice, gnotobiotic Swiss Webster mice, and antibiotic treated C57BL/6 mice, the investigators failed to account for the role of the microbiota in *Salmonella* lactate utilization where *Salmonella* lactate dehydrogenase expression and lactate enantiomer preference could vary depending on environmental conditions[60]. Unfortunately, examples like this are rife in the literature, where incomplete consideration of commensal activity leaves the role of the gut microbiota and commensal interactions up to question in many of the mechanisms touted critical to *Salmonella* enteric infection and proliferation.

Salmonella resistant mice, namely NRAMP1$^{+/+}$ breeds like CBA and 129/SvJ, can provide models to observe prolonged, gut-localized, *Salmonella* infection in the context of a complete microbiome[53,61]. A study conducted by our lab showed CBA/J fostering *Salmonella* enteric infection for up to 16 days[53]. Another study performed in 129/SvJ mice used multi-omics to observe *Salmonella* gut-localized infection for up to 28 days, showing *Salmonella* alteration of the microbiota, enrichment of *Enterococcus*, and potential fucose utilization[61]. One group studied *Salmonella* in NRAMP1$^{+/+}$ C57BL/6 mice and 129X1/SvJ mice, each with complete communities[62]. Their results concluded propionate from *Bacteroides* mediates colonization resistance against *Salmonella* based on differences in *Salmonella* shedding and microbiota community composition between mice breeds and *in vitro* evaluation of *Bacteroides* inhibition of *Salmonella*[62]. While compelling, these results cannot conclude *Bacteroides* alone is responsible for the varied level of *Salmonella* infection severity observed *in vivo*. Results in a separate paper show *Salmonella* utilizing propionate to respire nitrate, though unperturbed CBA mice were used for some of the experiments, no conclusions about microbiota interactions affecting the propionate nitrate respiratory process could be made because of the limitations in their methods (optical density growth assays, qRT-PCR, and LC-MS/MS propionate abundance quantified to a standard)[51]. *Salmonella* 1,2-propanediol mediated fitness advantage was observed in CBA mice with full communities, but further investigations of the involvement of specific microbiota members was performed in germ-free Swiss Webster mice monocolonized by individual bacteria strains[50]. Prior work either does not include a robust and complete microbiome or it does not evaluate whole community *Salmonella* infection in a way that considers the possibility of multiple microbiota members affecting the mechanism under investigation. Here I use multi-omics evaluation of the complete microbiome found in feces of CBA mice. This approach can be used to holistically interrogate prior mechanistic assertions done in reduced complexity models, as well as generate new hypothesis. In comparison to prior studies which targeted specific metabolites or metabolisms, multi-omics is an untargeted approach which enables the tracking of many metabolisms simultaneously.

To holistically evaluate *Salmonella* impact on the gut microbiome I integrated metagenomic, metatranscriptomic, and metabolomic evaluation of inflamed *Salmonella*-infected CBA/J mice with 16S rRNA amplicon sequencing and compared the results to multi-omics from uninfected mice. With metagenomics, whole community DNA is extracted from feces and sequenced. We then bioinformatically reconstructed the bacterial community and observed taxonomy representation in communities with and without *Salmonella*. Reconstructed genomes also provided information about the functional potential of the communities in each of our treatments. Paired metatranscriptomics delivered a comprehensive collection of community transcripts that may then be mapped to reconstructed genomes ascertaining individual bacteria expression during *Salmonella* infection. Metabolites extracted from feces can be examined and linked to bacteria expression. Enzyme products in our metabolomics data from abundant and highly expressed pathways can increase our confidence that a certain process is occurring in the gut. Additionally, by observing the presence or absence of specific microbial or host modifications in the metabolite data, further inference can be made about the microbiota activity and interactions between pathogen, host, and commensal bacteria. By observing genome-resolved transcription and metabolomics data we can also infer individual bacteria substrate utilization and the oxidation state of the gut, allowing for new connections between commensal bacteria and *Salmonella* and *Salmonella* and the host to be discovered. By mining reconstructed genomes for 16S sequences and linking them to 16S rRNA amplicon sequencing timeseries data, we can expand inferences made about expression and functional potential to timepoints during the experiment that lack multi-omics sampling.

Objectives of this work were as follows:

1. Create a first-of-its-kind comprehensive bacterial and viral genome database from CBA/J mice as a critical resource for further recruitment of other omics types in a genome-resolved multi-omics study of *Salmonella* infection (Chapter 2).
2. Leverage the CBA/J microbiota database and whole community transcripts and metabolites to determine persistent and important commensal membership both relatively abundant and

functionally involved in mechanisms and metabolisms critical to *Salmonella* successful infection (Chapter 2 / Chapter 3).

3. Discover important gene content of persistent commensal bacteria during *Salmonella* infection and mine microbiome data from *Salmonella* infected mice for potentially novel pathogen/commensal interactions to enlighten probiotic potential of strains for *Salmonella* therapeutic administration (Chapter 2 / Chapter 3).

Collectively, these works provide a valuable genome database resource to the microbiome community and to researchers who utilize the CBA/J mouse model. More central to *Salmonella* impact on the microbiome, the studies presented here identify coenriched and inflammation resistant commensal bacteria with probiotic potential that are prominent in the *Salmonella*-included gut. Simultaneously, this work identfies novel commensal involvement in substrate availability for *Salmonella* anaerobic respiration. Beyond promising probiotic targets for future exploration, this work also confronts current paradigms of *Salmonella* substrate utilization and offers hypotheses seeding further studies that are ultimately more translatable owing to the holistic and robust multi-omics approach taken herein.

Chapter 1 References

1. Stanaway JD, Parisi A, Sarkar K, Blacker BF, Reiner RC, Hay SI, Nixon MR, Dolecek C, James SL, Mokdad AH, et al. The global burden of non-typhoidal salmonella invasive disease: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet Infectious Diseases 2019; 19:1312–24.

2. Cheng RA, Eade CR, Wiedmann M. Embracing Diversity: Differences in Virulence Mechanisms, Disease Severity, and Host Adaptations Contribute to the Success of Nontyphoidal Salmonella as a Foodborne Pathogen. Frontiers in Microbiology [Internet] 2019 [cited 2023 May 1]; 10. Available from: https://www.frontiersin.org/articles/10.3389/fmicb.2019.01368

3. Rabsch W, Andrews HL, Kingsley RA, Prager R, Tschäpe H, Adams LG, Bäumler AJ. Salmonella enterica Serotype Typhimurium and Its Host-Adapted Variants. Infection and Immunity 2002; 70:2249–55.

4. Zając M, Skarżyńska M, Lalak A, Kwit R, Śmiałowska-Węglińska A, Pasim P, Szulowski K, Wasyl D. Salmonella in Captive Reptiles and Their Environment—Can We Tame the Dragon? Microorganisms 2021; 9:1012.

5. Davidson KE, Byrne BA, Pires AFA, Magdesian KG, Pereira RV. Antimicrobial resistance trends in fecal Salmonella isolates from northern California dairy cattle admitted to a veterinary teaching hospital, 2002-2016. PLoS ONE [Internet] 2018 [cited 2023 May 1]; 13. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6023112/

6. Martin MJ, Thottathil SE, Newman TB. Antibiotics Overuse in Animal Agriculture: A Call to Action for Health Care Providers. Am J Public Health 2015; 105:2409–10.

7. Kariuki S, Gordon MA, Feasey N, Parry CM. Antimicrobial resistance and management of invasive Salmonella disease. Vaccine 2015; 33:C21–9.

8. Plumb ID, Brown AC, Stokes EK, Chen JC, Carleton H, Tolar B, Sundararaman P, Saupe A, Payne DC, Shah HJ, et al. Increased Multidrug-Resistant Salmonella enterica I Serotype 4,[5],12:i:- Infections Associated with Pork, United States, 2009–2018 - Volume 29, Number 2—February 2023 - Emerging Infectious Diseases journal - CDC. [cited 2023 Apr 3]; Available from: https://wwwnc.cdc.gov/eid/article/29/2/22-0950_article

9. Ruvalcaba-Gómez JM, Villagrán Z, Valdez-Alarcón JJ, Martínez-Núñez M, Gomez-Godínez LJ, Ruesga-Gutiérrez E, Anaya-Esparza LM, Arteaga-Garibay RI, Villarruel-López A. Non-Antibiotics Strategies to Control Salmonella Infection in Poultry. Animals (Basel) 2022; 12:102.

10. Nale JY, Vinner GK, Lopez VC, Thanki AM, Phothaworn P, Thiennimitr P, Garcia A, AbuOun M, Anjum MF, Korbsrisate S, et al. An Optimized Bacteriophage Cocktail Can Effectively Control Salmonella in vitro and in Galleria mellonella. Frontiers in Microbiology [Internet] 2021 [cited 2023 May 1]; 11. Available from: https://www.frontiersin.org/articles/10.3389/fmicb.2020.609955

11. Ehuwa O, Jaiswal AK, Jaiswal S. Salmonella, Food Safety and Food Handling Practices. Foods 2021; 10:907.

12. Palmer AD, Slauch JM. Mechanisms of Salmonella pathogenesis in animal models. Hum Ecol Risk Assess 2017; 23:1877–92.

13. Darwin KH, Miller VL. Molecular Basis of the Interaction of Salmonella with the Intestinal Mucosa. Clin Microbiol Rev 1999; 12:405–28.

14. Lou L, Zhang P, Piao R, Wang Y. Salmonella Pathogenicity Island 1 (SPI-1) and Its Complex Regulatory Network. Frontiers in Cellular and Infection Microbiology [Internet] 2019 [cited 2023 Apr 4]; 9. Available from: https://www.frontiersin.org/articles/10.3389/fcimb.2019.00270

15. Tan J, McKenzie C, Potamitis M, Thorburn AN, Mackay CR, Macia L. Chapter Three - The Role of Short-Chain Fatty Acids in Health and Disease [Internet]. In: Alt FW, editor. Advances in Immunology. Academic Press; 2014 [cited 2023 May 7]. page 91–119.Available from: https://www.sciencedirect.com/science/article/pii/B9780128001004000039

16. Silva YP, Bernardi A, Frozza RL. The Role of Short-Chain Fatty Acids From Gut Microbiota in Gut-Brain Communication. Frontiers in Endocrinology [Internet] 2020 [cited 2023 May 2]; 11. Available from: https://www.frontiersin.org/articles/10.3389/fendo.2020.00025

17. Kelly CJ, Zheng L, Campbell EL, Saeedi B, Scholz CC, Bayless AJ, Wilson KE, Glove LE, Kominsky DJ, Magnuson A, et al. Crosstalk between Microbiota-Derived Short-Chain Fatty Acids and Intestinal Epithelial HIF Augments Tissue Barrier Function. Cell Host Microbe 2015; 17:662–71.

18. Golubeva YA, Ellermeier JR, Cott Chubiz JE, Slauch JM. Intestinal Long-Chain Fatty Acids Act as a Direct Signal To Modulate Expression of the Salmonella Pathogenicity Island 1 Type III Secretion System. mBio 2016; 7:e02170-15.

19. Kohli N, Crisp Z, Riordan R, Li M, Alaniz RC, Jayaraman A. The microbiota metabolite indole inhibits Salmonella virulence: Involvement of the PhoPQ two-component system. PLoS ONE 2018; 13:e0190613.

20. Rogers AWL, Tsolis RM, Bäumler AJ. Salmonella versus the Microbiome. Microbiol Mol Biol Rev 2021; 85.

21. Ibarra JA, Steele-Mortimer O. Salmonella – the ultimate insider. Salmonella virulence factors that modulate intracellular survival. Cellular Microbiology 2009; 11:1579.

22. Rosenberg G, Yehezkel D, Hoffman D, Mattioli CC, Fremder M, Ben-Arosh H, Vainman L, Nissani N, Hen-Avivi S, Brenner S, et al. Host succinate is an activation signal for Salmonella virulence during intracellular infection. Science 2021; 371:400–5.

23. Monack DM, Bouley DM, Falkow S. Salmonella typhimurium Persists within Macrophages in the Mesenteric Lymph Nodes of Chronically Infected Nramp1 + / + Mice and Can Be Reactivated by IFNγ Neutralization. J Exp Med 2004; 199:231–41.

24. Fritsche G, Nairz M, Libby SJ, Fang FC, Weiss G. Slc11a1 (Nramp1) impairs growth of Salmonella enterica serovar typhimurium in macrophages via stimulation of lipocalin-2 expression. J Leukoc Biol 2012; 92:353–9.

25. Brown DE, Libby SJ, Moreland SM, McCoy MW, Brabb T, Stepanek A, Fang FC, Detweiler CS. Salmonella enterica Causes More Severe Inflammatory Disease in C57/BL6 Nramp1G169Mice Than Sv129S6 Mice. Vet Pathol 2013; 50:867–76.

26. Sabag-Daigle A, Blunk HM, Gonzalez JF, Steidley BL, Boyaka PN, Ahmer BMM. Use of Attenuated but Metabolically Competent Salmonella as a Probiotic To Prevent or Treat Salmonella Infection. Infection and Immunity 2016; 84:2131–40.

27. Rinninella E, Raoul P, Cintoni M, Franceschi F, Miggiano GAD, Gasbarrini A, Mele MC. What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases. Microorganisms 2019; 7:14.

28. Litvak Y, Bäumler AJ. Microbiota-Nourishing Immunity: A Guide to Understanding Our Microbial Self. Immunity 2019; 51:214–24.

29. Aljahdali NH, Sanad YM, Han J, Foley SL. Current knowledge and perspectives of potential impacts of Salmonella enterica on the profile of the gut microbiota. BMC Microbiology 2020; 20:353.

30. Sweeney TE, Morton JM. The Human Gut Microbiome. JAMA Surg 2013; 148:563–9.

31. Singh V, Lee G, Son H, Koh H, Kim ES, Unno T, Shin J-H. Butyrate producers, "The Sentinel of Gut": Their intestinal significance with and beyond butyrate, and prospective use as microbial therapeutics. Frontiers in Microbiology [Internet] 2023 [cited 2023 May 2]; 13. Available from: https://www.frontiersin.org/articles/10.3389/fmicb.2022.1103836

32. Kim S, Shin Y-C, Kim T-Y, Kim Y, Lee Y-S, Lee S-H, Kim M-N, O E, Kim KS, Kweon M-N. Mucin degrader Akkermansia muciniphila accelerates intestinal stem cell-mediated epithelial development. Gut Microbes 13:1892441.

33. Schaefer AK, Melnyk JE, He Z, Del Rosario F, Grimes CL. Chapter 14 - Pathogen- and Microbial- Associated Molecular Patterns (PAMPs/MAMPs) and the Innate Immune Response in Crohn's Disease [Internet]. In: Chatterjee S, Jungraithmayr W, Bagchi D, editors. Immunity and Inflammation in Health and Disease. Academic Press; 2018 [cited 2023 May 2]. page 175–87.Available from: https://www.sciencedirect.com/science/article/pii/B9780128054178000147

34. Bischoff SC, Barbara G, Buurman W, Ockhuizen T, Schulzke J-D, Serino M, Tilg H, Watson A, Wells JM. Intestinal permeability – a new target for disease prevention and therapy. BMC Gastroenterology 2014; 14:189.

35. Sun L, Yang S, Deng Q, Dong K, Li Y, Wu S, Huang R. Salmonella Effector SpvB Disrupts Intestinal Epithelial Barrier Integrity for Bacterial Translocation. Front Cell Infect Microbiol 2020; 10:606541.

36. Galán JE. Salmonella Typhimurium and inflammation: a pathogen-centric affair. Nat Rev Microbiol 2021; 19:716–25.

37. Hakansson A, Molin G. Gut Microbiota and Inflammation. Nutrients 2011; 3:637–82.

38. Ganesh BP, Klopfleisch R, Loh G, Blaut M. Commensal Akkermansia muciniphila Exacerbates Gut Inflammation in Salmonella Typhimurium-Infected Gnotobiotic Mice. PLoS One 2013; 8:e74963.

39. Repoila F, Le Bohec F, Guérin C, Lacoux C, Tiwari S, Jaiswal AK, Santana MP, Kennedy SP, Quinquis B, Rainteau D, et al. Adaptation of the gut pathobiont Enterococcus faecalis to deoxycholate and taurocholate bile acids. Sci Rep 2022; 12:8485.

40. Rivera-Chávez F, Zhang LF, Faber F, Lopez CA, Byndloss MX, Olsan EE, Xu G, Velazquez EM, Lebrilla CB, Winter SE, et al. Depletion of butyrate-producing Clostridia from the gut microbiota drives an aerobic luminal expansion of Salmonella. Cell Host Microbe 2016; 19:443–54.

41. Santos RL, Raffatellu M, Bevins CL, Adams LG, Tükel Ç, Tsolis RM, Bäumler AJ. Life in the inflamed intestine, Salmonella style. Trends in Microbiology 2009; 17:498–506.

42. Nguyen GT, Green ER, Mecsas J. Neutrophils to the ROScue: Mechanisms of NADPH Oxidase Activation and Bacterial Resistance. Frontiers in Cellular and Infection Microbiology [Internet] 2017 [cited 2023 May 2]; 7. Available from: https://www.frontiersin.org/articles/10.3389/fcimb.2017.00373

43. Rhen M. Salmonella and Reactive Oxygen Species: A Love-Hate Relationship. JIN 2019; 11:216–26.

44. Bliska JB, van der Velden AWM. Salmonella “Sops” Up a Preferred Electron Receptor in the Inflamed Intestine. mBio 2012; 3:e00226-12.

45. Winter SE, Thiennimitr P, Winter MG, Butler BP, Huseby DL, Crawford RW, Russell JM, Bevins CL, Adams LG, Tsolis RM, et al. Gut inflammation provides a respiratory electron acceptor for Salmonella. Nature 2010; 467:426–9.

46. Shelton CD, Yoo W, Shealy NG, Torres TP, Zieba JK, Calcutt MW, Foegeding NJ, Kim D, Kim J, Ryu S, et al. Salmonella Typhimurium uses anaerobic respiration to overcome propionate-mediated colonization resistance. bioRxiv 2021; :2021.05.25.445690.

47. Litvak Y, Mon KKZ, Nguyen H, Chanthavixay G, Liou M, Velazquez EM, Kutter L, Alcantara MA, Byndloss MX, Tiffany CR, et al. Commensal Enterobacteriaceae Protect against Salmonella Colonization through Oxygen Competition. Cell Host & Microbe 2019; 25:128-139.e5.

48. Maier RJ, Olczak A, Maier S, Soni S, Gunn J. Respiratory Hydrogen Use by Salmonella enterica Serovar Typhimurium Is Essential for Virulence. Infect Immun 2004; 72:6294–9.

49. Maier L, Vyas R, Cordova CD, Lindsay H, Schmidt TSB, Brugiroux S, Periaswamy B, Bauer R, Sturm A, Schreiber F, et al. Microbiota-Derived Hydrogen Fuels Salmonella Typhimurium Invasion of the Gut Ecosystem. Cell Host & Microbe 2013; 14:641–51.

50. Faber F, Thiennimitr P, Spiga L, Byndloss MX, Litvak Y, Lawhon S, Andrews-Polymenis HL, Winter SE, Bäumler AJ. Respiration of Microbiota-Derived 1,2-propanediol Drives Salmonella Expansion during Colitis. PLOS Pathogens 2017; 13:e1006129.

51. Shelton CD, Yoo W, Shealy NG, Torres TP, Zieba JK, Calcutt MW, Foegeding NJ, Kim D, Kim J, Ryu S, et al. Salmonella enterica serovar Typhimurium uses anaerobic respiration to overcome propionate-mediated colonization resistance. Cell Reports [Internet] 2022 [cited 2022 Jul 14]; 38. Available from: https://www.cell.com/cell-reports/abstract/S2211-1247(21)01680-6

52. Spiga L, Winter MG, de Carvalho TF, Zhu W, Hughes ER, Gillis CC, Behrendt CL, Kim J, Chessa D, Andrews-Polymenis HL, et al. An oxidative central metabolism enables Salmonella to utilize microbiota-derived succinate. Cell Host Microbe 2017; 22:291-301.e6.

53. Borton MA, Sabag-Daigle A, Wu J, Solden LM, O'Banion BS, Daly RA, Wolfe RA, Gonzalez JF, Wysocki VH, Ahmer BMM, et al. Chemical and pathogen-induced inflammation disrupt the murine intestinal microbiome. Microbiome [Internet] 2017 [cited 2019 Aug 26]; 5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5408407/

54. Hapfelmeier S, Hardt W-D. A mouse model for S. typhimurium-induced enterocolitis. Trends in Microbiology 2005; 13:497–503.

55. Barthel M, Hapfelmeier S, Quintanilla-Martínez L, Kremer M, Rohde M, Hogardt M, Pfeffer K, Rüssmann H, Hardt W-D. Pretreatment of Mice with Streptomycin Provides a Salmonella enterica Serovar Typhimurium Colitis Model That Allows Analysis of Both Pathogen and Host. Infection and Immunity 2003; 71:2839–58.

56. Coburn B, Li Y, Owen D, Vallance BA, Finlay BB. Salmonella enterica Serovar Typhimurium Pathogenicity Island 2 Is Necessary for Complete Virulence in a Mouse Model of Infectious Enterocolitis. Infect Immun 2005; 73:3219–27.

57. Stecher B, Macpherson AJ, Hapfelmeier S, Kremer M, Stallmach T, Hardt W-D. Comparison of Salmonella enterica Serovar Typhimurium Colitis in Germfree Mice and Mice Pretreated with Streptomycin. Infect Immun 2005; 73:3228–41.

58. Lopez CA, Winter SE, Rivera-Chávez F, Xavier MN, Poon V, Nuccio S-P, Tsolis RM, Bäumler AJ. Phage-Mediated Acquisition of a Type III Secreted Effector Protein Boosts Growth of Salmonella by Nitrate Respiration. mBio 2012; 3:e00143-12.

59. Rivera-Chávez F, Winter SE, Lopez CA, Xavier MN, Winter MG, Nuccio S-P, Russell JM, Laughlin RC, Lawhon SD, Sterzenbach T, et al. Salmonella Uses Energy Taxis to Benefit from Intestinal Inflammation. PLoS Pathog [Internet] 2013 [cited 2020 Feb 6]; 9. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3630101/

60. Gillis CC, Hughes ER, Spiga L, Winter MG, Zhu W, Furtado de Carvalho T, Chanin RB, Behrendt CL, Hooper LV, Santos RL, et al. Dysbiosis-Associated Change in Host Metabolism Generates Lactate to Support Salmonella Growth. Cell Host & Microbe 2018; 23:54-64.e6.

61. Kaiser BLD, Li J, Sanford JA, Kim Y-M, Kronewitter SR, Jones MB, Peterson CT, Peterson SN, Frank BC, Purvine SO, et al. A Multi-Omic View of Host-Pathogen-Commensal Interplay in Salmonella-Mediated Intestinal Infection. PLOS ONE 2013; 8:e67155.

62. Jacobson A, Lam L, Rajendram M, Tamburini F, Honeycutt J, Pham T, Treuren WV, Pruss K, Stabler SR, Lugo K, et al. A Gut Commensal-Produced Metabolite Mediates Colonization Resistance to Salmonella Infection. Cell Host & Microbe 2018; 24:296-307.e7.

Chapter 2 – Exposing New Taxonomic Variation with Inflammation – A Murine Model-Specific Genome Database for Gut Microbiome Researchers

2.1 Summary

The murine CBA/J mouse model widely supports immunology and enteric pathogen research. This model has illuminated *Salmonella* interactions with the gut microbiome since pathogen proliferation does not require disruptive pretreatment of the native microbiota, nor does it become systemic, thereby representing an analog to gastroenteritis disease progression in humans. Despite the value to broad research communities, microbiota in CBA/J mice are not represented in current murine microbiome genome catalogs.

Here we present the first microbial and viral genomic catalog of the CBA/J murine gut microbiome. Using fecal microbial communities from untreated and *Salmonella*-infected, highly inflamed mice, we performed genomic reconstruction to determine the impacts on gut microbiome membership and functional potential. From high depth whole community sequencing (~42.4 Gbps/sample), we reconstructed 2,281 bacterial and 4,516 viral draft genomes. *Salmonella* challenge significantly altered gut membership in CBA/J mice, revealing 30 genera and 98 species that were conditionally rare and unsampled in non-inflamed mice. Additionally, inflamed communities were depleted in microbial genes that modulate host anti-inflammatory pathways and enriched in genes for respiratory energy generation. Our findings suggest decreases in butyrate concentrations during *Salmonella* infection corresponded to reductions in the relative abundance in members of the *Alistipes.* Strain-level comparison of CBA/J microbial genomes to prominent murine gut microbiome databases identified newly sampled lineages in this resource, while comparisons to human gut microbiomes extended the host relevance of dominant CBA/J inflammation resistant strains.

This CBA/J microbiome database provides the first genomic sampling of relevant, uncultivated microorganisms within the gut from this widely used laboratory model. Using this resource, we curated a functional, strain-resolved view on how *Salmonella* remodels intact murine gut communities, advancing pathobiome understanding beyond inferences from prior amplicon-based approaches. *Salmonella*-induced inflammation suppressed *Alistipes* and other dominant members, while rarer commensals like *Lactobacillus* and *Enterococcus* endure. The rare and novel species sampled across this inflammation gradient advance the utility of this microbiome resource to benefit the broad research needs of the CBA/J scientific community, and those using murine models for understanding the impact of inflammation on the gut microbiome more generally.

2.2 Introduction

Non-typhoidal *Salmonella* (NTS) is one of the leading causes of gastroenteritis and associated mortality worldwide, resulting in nearly 1 million cases and more than 50 thousand deaths in 2017 [1–3]. *Salmonella enterica* serovar Typhimurium (referred to hereon as *Salmonella*) is a NTS model enteric pathogen that exploits inflammation to increase its pathogenicity and fitness relative to other bacteria [4–6]. Prior research in murine models showed gut microbiota are remodeled during *Salmonella*-induced inflammation because of innate immunity, diminished resources, and altered chemical environment [6,7]. Similar inflammation-associated changes in commensal gut ecology are observed in patients with Crohn's disease, irritable bowel disease, and metabolic syndrome [4,8]. The *Salmonella* disease model may represent a human-relevant system for investigating host-pathogen-microbiota interactions in the inflamed GI tract germane to microbiome changes during other human chronic inflammatory conditions.

Earlier work showed that *Salmonella* induced inflammation created reactive oxygen and nitrogen species in the luminal environment, increasing the availability of oxygen and potentiating formation of tetrathionate and nitrate [9]. Increased concentrations of these terminal electron acceptors allowed *Salmonella* to respire and out-compete obligate fermentative commensals like members of the *Clostridia*

[6,10]. Additionally, in the remodeled gut ecosystem respiring *Salmonella* benefited from unique access to non-competitive carbon sources like propionate and ethanolamine [11,12]. Prior competition studies often focused on decreases in the *Clostridia* [7,13], ignoring implications for the other members of the community, and those that can withstand *Salmonella* infection. Here we provide a holistic, strain resolved view of the changes in microbiome membership and function during *Salmonella* infection.

Previous investigations of the impacts of *Salmonella* on the gut microbiome relied on the use of pre-treated reduced diversity communities [14–16]. Mice in past studies (e.g., BALB/c or C57BL/6) required antibiotic conditioning prior to pathogen introduction, preventing investigation of *Salmonella* physiology in response to an intact microbial community [14,15,17]. As an alternative, CBA/J mice are gaining appreciation as a model to interrogate *Salmonella* pathogenicity as they support longer non-systemic *Salmonella* infection without prior antibiotic perturbation, similar to enteric disease manifestation in humans [17–20]. While *in vitro* and *in vivo* studies with reduced microbiota consortia provide an important theoretical framework, additional research is needed in the presence of native microbial communities to understand how specific inflammation-induced changes to microbiota membership and function impact *Salmonella* physiology and pathogenicity.

Currently the capability to study intact resident gut microbiota during *Salmonella*-induced inflammation is hindered because CBA/J mice are missing from available murine microbiome genomic catalogs [21–23]. Furthermore, existing murine gut genomic databases exclude inflamed mice, and mice colonized by enteric pathogens (e.g. like *Salmonella, Klebsiella*, and *Citrobacter)*, thus limiting the extension of these existing microbiome resources to pathobiome models. Despite being recognized as important contributors to human health [24], these existing curated murine gut microbial genome catalogs also lack virome sampling [25,26]. Accurate interrogation of complete microbial community functions during *Salmonella* infection requires comprehensive model-specific knowledge of gene content and community membership both in healthy and inflamed guts.

To evaluate the functional potential of microbial communities during *Salmonella*-induced inflammation, and to explore if the CBA/J inflammation model harbors unique and previously understudied microorganisms, we constructed a metagenome assembled genome catalog from healthy and *Salmonella* infected CBA/J mice. We employed high depth metagenomic sequencing and used several assembly strategies to increase the *de novo* reconstruction of viral and bacterial genomes. These efforts resulted in a comprehensive culture-independent genome resource that (i) revealed novel taxonomy unique to CBA/J and inflamed mice, (ii) included taxa with relevance to human systems and with anti-inflammatory effector potential, and (iii) showed how enteric inflammation remodels the functional profile of the gut by selecting for bacteria that encode mechanisms to withstand oxidative stress. Ultimately our findings reframe existing responses of the microbiota during *Salmonella* infection and provide new insights into specific bacteria that can withstand inflammation to maintain critical gut functions, perhaps revealing promising future probiotic targets.

2.3 Results

2.3.1 Pathogen perturbation extends the genomic sampling of the CBA/J microbiome #

To examine the microbial community response to *Salmonella* colonization, 14 CBA/J mice were infected with 10^9 CFU *Salmonella enterica* serovar Typhimurium strain 14028 with results compared to 16 uninoculated control mice sampled at the same time points (n=30 mice, Fig. 1A). Feces were collected prior to infection (Day -1) and in late stages of infection (Days 10 and 11), with 16S rRNA microbial community analyses performed at early and late time points (n=60) and lipocalin-2, an indicator of enteric inflammation, measured on late time points from select mice from each treatment group (n=12). The 60 fecal samples yielded 2,047,287 paired end 16S rRNA reads, which identified 23,022 unique amplicon sequencing variants (ASVs) from both inflamed and control treatments. To confirm infection, we established that inoculated mice had *Salmonella* relative abundance greater than 25% on Day 11 and had significantly higher lipocalin-2 concentrations than control mice on Day 10. From these mice we selected

feces at Day 11 from 3 *Salmonella* infected mice and 3 uninfected mice for deep metagenomic sequencing.

The 16S rRNA gene findings confirmed *Salmonella* infection resulted in statistically discernable microbial communities by Day 11 following infection (Fig. 1B). A *Salmonella enterica* relative abundance increase (≥25%) was concomitant with increased inflammation evidenced by a 2.5 log-fold rise in lipocalin-2 compared to levels in uninfected mice (Fig. 1B). The microbial community of inflamed mice statistically differed from control uninfected mice at the same time point, and pre-pathogen treated mice from both treatments (Fig. 1B, Fig. 1C, Fig. 7A). Pre-infection (Day-1) mice that later became *Salmonella* inoculated, and uninoculated mice had fecal microbial communities that were not discernable from each other, indicating observed community differences by Day 11 were due to *Salmonella* infection (Fig. 7B). As others have reported [13,27], *Salmonella*-induced inflammation significantly changed gut microbial diversity, it reduced ASV richness by more than half (76.2%) and decreased Shannon's diversity by 2.6-fold. These findings demonstrate that pathogen perturbation changes microbial membership and structure, offering a strategy for differential genomic sampling of the CBA/J gut microbiome.

To extend the relevance of our findings we compared uninfected amplicon sequenced communities to communities from CBA mice in two other studies and showed a strong overlap in taxonomy between the three (Fig. 8) [27,28]. Dominate taxa in the CBAJ-DB were shared across studies despite myriad differences including mouse breeding facility, chow type, experimental facility, and experimental methods (Fig. 8). These findings highlight the relevance of this first CBA mouse genomic resource to microbiome research in this model more broadly by expounding notable community consistency between pathogen-free CBA mice despite other confounding factors.

These 16S rRNA analyses revealed inflamed communities were enriched in members of the *Gammaproteobacteria* and *Bacilli,* while gut communities of uninoculated mice included higher relative abundances of *Bacteroidia*, *Mollicutes*, and *Clostridia* (Fig. 1C). From the mice also sampled for metagenomic analysis (n=6), *Alistipes* sp. was the most dominant commensal and the most reduced during inflammation (from 37.2%), but notably still detectable (7.68%). *Salmonella enterica* Typhimurium dominated the *Gammaproteobacteria* in inflamed communities, contributing up to a mean relative abundance of 94% in infected samples. Certain low abundant members of the CBA/J microbiome significantly increased in relative abundance following pathogen treatment, including some members of *Lactobacillus*, *Enterococcus*, and *Lachnospiraceae* (Fig. 1D). Control mouse communities are consistent with findings from prior work showing uninfected CBA/J mouse gut community membership dominated by *Bacteroidetes* and *Firmicutes*, especially *Clostridia* of various *Lachnospiraceae* and *Ruminococcaceae* families [27,28]. These 16S rRNA gene analyses revealed abundant members in both healthy and inflamed CBA/J gut microbiomes that represented microorganismal genome 'targets' for our database.

2.3.2 Microbial genomic reconstruction from CBA/J mice recovers relevant members sampled in amplicon surveys

To thoroughly catalog the CBA/J gut microbiota high sequencing depth was required to sequence through *Salmonella* dominance (25.8 – 94.2% by amplicon analyses) and recover some of the first genomes from rare, but persistent co-occurring members of the pathogen inflamed gut. We obtained 254.2 Gbps of metagenomic sequencing data from 6 representative mice (inflamed n=3, uninfected n=3, Fig. 1A), 7-fold more sequencing/sample than is commonly done in murine catalogs (Fig. 9A)[21,22,29]. Additionally, we used iterative, targeted assembly approaches (single, co-assembly, subtractive assembly) as well as two different assemblers to attempt to enhance genome quality and recovery, especially from less dominant members (Fig. 2A, Fig. 9B). Subtractive and co-assembly methods derived 259 additional metagenomically assembled genomes (MAGs) beyond those from single sample assemblies, with the distribution of MAGs from each assembler reported (Fig. 2A). In total, we recovered 2,281 MAGs. Quality assessment revealed 504 MAGs to be either medium or high-quality (MQHQ) with

sufficiently low contamination to be included in further analyses. Of the genome quality tools used, CheckM [30] provided the most conservative MAG set (n=504 MQHQ MAGs), compared to CheckM2 [31] (n=531 MQHQ MAGs) and GUNC [32] (n=790, MQHQ MAGs) as assessed by contamination and completeness. There existed no significant differences between the quality metrics of MAGs from either treatment (uninfected, infected) or between assembly methods (Fig. 10A). These quality genomes contained 156,921 uniquely called predicted genes [33,34] (Fig. 2).

Dereplication of our metagenome assembled genomes (99% identity) resulted in 113 medium and high-quality MAGs (dMQHQ) from both treatments. These MAGs were assigned to 7 Phyla – *Actinobacteriota* (1), *Bacteroidota* (4) *Firmicutes* (7), *Firmicutes*_A (98), *Firmicutes*_B (1), *Proteobacteria* (1), and *Verrucomicrobiota* (1) (Fig. 2D). Nearly a third (30 of the 113) of the dereplicated MAGs were assigned to 30 genera and 98 to species that were only recognized by alphanumeric numbering in GTDB-Tk, hinting that novelty sampled here may be undescribed not only in murine but larger MAG collections. Reflecting the richness of these samples, the majority of MAGs originated from uninfected mice (59%) and their co-assemblies (35%) while 13% came from inflamed mice. Specifically, *Enterococcus_D gallinarum, Erysipelatoclostridium cocleatum, Kineothrix sp000403275, and Lactobacillus_B animalis* MAGs were uniquely recovered from inflamed mice, consistent with their 16S rRNA membership (Fig. 1D). This finding indicates how perturbation can aid in the sampling of genomes from conditionally rare members, however we also acknowledge the risk of facility specificity and other factors influencing these microbiota.

Expanding this resource beyond solely bacterial genomes, we also reconstructed viral genomes from our CBA/J assemblies, recovering 4,516 viral metagenome assembled genomes (vMAGs). Of these, 2,351 vMAGs were ≥10kb which were then dereplicated into 609 viral genomes (Fig. 2C, Fig. 4D).

We first sought to verify if this microbial dereplicated MAG set represented the key members identified in our amplicon sequenced CBA/J communities from both inflamed (n=14) and uninfected (n=16) individuals (Fig. 1B). The relative abundance of dMQHQ MAGs closely mirrored the full community 16S rRNA amplicon at the class level from both uninfected Day 11 (rho=0.68) and inflamed Day 11 (rho=0.86) mice, indicating the dMQHQ database is representative of CBA/J untreated and inflamed communities (Fig. 3A, Fig. 1C). More specifically, a linear discriminate analysis of MAG relative abundance indicated similar dynamics between our genome and amplicon data sets. For example, *Salmonella* and *Enterococcus_D* were the most significant genomes in determining infected communities, while genomes from *Alistipes, Duncaniella and Lachnospiraceae* COE1 were most significant in determining uninfected communities (Fig. 3B). Additional to these genera, relative abundance of other key taxa is consistent with amplicon sequencing, including *Akkermansia,* and *Muribaculaceae* prominence in uninfected mice and persistence in infected mice. *Lactobacillus* genome and ASV relative abundance also similarly increased during infection (Fig. 3C).

To link these reconstructed genomes more precisely to the amplicon data, we identified 96 MAGs that contained a partial to full 16S rRNA gene sequence. A relatively low proportion of MAGs containing 16S rRNA sequences may be attributed to the difficulty *de novo* assembly algorithms have with the conserved regions and tetranucleotide variation associated with this gene [35]. A pairwise comparison of MAG-derived 16S rRNA sequences and the V4 region sequences from our ASVs identified 33 unique genomes containing sequences matching ASVs in our 16S rRNA dataset. Many MAGs with 16S rRNA matches were among the most enriched taxa including *Lactobacillus johnsonii*, *Alistipes* sp002428825, and *Clostridia* in order 4C28d-15 (Fig. 10B). Together these findings indicate significant membership congruence in our MAG database and our amplicon data, demonstrating that inferences made with the CBAJ-DB have relevance to the more broadly sampled amplicon sequenced gut communities from inflamed and uninfected mice.

2.3.3 This CBA/J microbial genomic resource includes mouse and human relevant lineages

One goal of developing a genome-resolved CBA specific microbiome resource is to advance future multi-omics studies in this mouse model. Metaproteomes and metatranscriptomes are mapped using high stringency to illuminate changes in gene expression, often at a genome-resolved level, typically the strain level (>99% average nucleotide identity) [36–38]. To assess the unique members captured in this resource, we compared strain level identity of our sampled MAGs to similar quality MAGs from two prevalent mouse gut genome catalogs: (i) Integrated Mouse Gut Metagenomic Catalog (iMGMC) [21] and (ii) The Mouse Gastrointestinal Bacteria Catalogue (MGBC) [22]. Notably, many of these CBA derived genomes represented unique strains from the classes *Bacilli* (n=3), *Bacteroidia* (n=2), *Clostridia* (n=24), *Coriobacteriia* (n=1), and *Dehalobacteriia* (n=1) not represented in iMGMC, and MAGs from *Bacilli* (n=1), *Dehalobacteriia* (n=1), and *Clostridia* (n=30) not represented in MGBC (Fig. 4A). Additionally, of the strains that were sampled in our dataset and prior curated catalogs, 33 (30 *Clostridia*, 3 *Bacilli*) received a higher quality score indicating the value of these recovered MAGs to advance knowledge of cultivated and uncultivated genomes in murine models more broadly. At a coarser taxonomic level (e.g. 95% genome sequence identity) we still detected novel taxa in the CBAJ-DB that were absent in these existing genomic collections. For example, clustered at 95% identity the CBAJ-DB harbored a novel *Faecalicatena* sp. and *Provencibacterium* sp. not found in the iMGMC.

We also examined CBAJ-DB MAGs against genomes derived from human hosts. To analyze shared genera and species, our dMQHQ database was dereplicated with isolate genomes from the Human Microbiome Project (HMP) [39] (n= 813) and MQHQ MAGs (n=2,560) from a human cohort (PRJNA725020) [40,6/21/2023 11:14:00 PM39]. *Akkermansia muciniphila* (CBAJDB_482) and *Enterococcus_D gallinarum* (CBAJDB_497), two defining members of the commensal and inflamed gut respectively, clustered with species previously recovered from human hosts. Recovery of *Enterococcus_D gallinarum* from the uninfected CBA/J gut demonstrates the applicability of perturbation techniques to uncover conditionally rare members. As has been reported by others, there was more similarity at higher

taxonomic levels (e.g. genus) between our murine and human gut microbial members [41,42], with 27 MAGs from *Bacilli, Bacteroidia, Clostridia, Corriobacteriia, Gammaproteobacteria, and Verrucomicrobiae* sharing similarity (Fig. 4A).

We were particularly interested if the microbial members recovered from our pathogen-inflamed CBA/J had relevance to inflammation in humans. To test this, sequencing reads from the Lloyd-Price et al cohort [43] containing 972 inflamed and 365 healthy gut metagenomes were stringently mapped to the CBAJ-DB MQHQ MAGs[43]. Consistent with their distribution across our treatments, sequencing reads from healthy and inflamed humans mapped to 11 of our *Akkermansia muciniphila* MAGs, while 3 *Enterococcus_D gallinarum* MAGs derived only in our inflamed treatments recruited sequences from inflamed human subjects (Fig. 4B, C). While it can be challenging to extend specific microorganismal findings from murine to human conditions [22,41,42], inferences from critical lineages (e.g. *A. muciniphila* or *E. gallinarum)* in our database may have more direct human relevance as *A. muciniphila* is recognized to promote gut barrier integrity and *E. gallinarum* has multiple documented cases as a pathobiont [44,45]. Beyond inflammation, we also dereplicated the CBAJ-DB with the Unified Human Gastrointestinal Genome (UHGG) catalog [46]. Species from the *Muribaculaceae* (including genus CAG-485) *and Oscillospiraceae* family (including *Lawsonibacter* spp.*, Oscillibacter* spp.*,* genus UBA9475, and an undescribed genus) were well represented in the CBAJ-DB and human fecal samples. Together these findings show that the CBAJ-DB recovered new species, but also many that have genomic coherence with members in the human gut.

2.3.4 Salmonella infection and inflammation restructures the metabolic potential of the murine gut microbiome

Given this is one of the first genome-resolved analyses of a pathogen-impacted microbiota, and the first for *Salmonella*, it offered a new opportunity to assess functional potential remodeling during infection. Prior reports indicated that pathogen induced inflammation created oxidative conditions that generated terminal electron acceptors like oxygen, tetrathionate, nitrate, and sulfate [9,19]. As such, we

wanted to evaluate the respiratory capacity of inflamed communities and compare it to uninflamed communities. Interestingly, individual respiration functions were not significantly different between treatments, however when analysis was expanded to consider the entire respiration category, we found MAGs with respiration capability to be significantly enriched in infected samples (ANOVA $p=1.23e^{-4}$), likely due to the increased n afforded at the category level of analysis (Fig. 5A). In infected communities, *Salmonella* has the highest mean genome relative abundance, and encoded gene sets for respiring oxygen (both high and low affinity oxidases), fumarate, tetrathionate, and trimethylamine N-oxide (TMAO) (Fig. 5A). Outside *Salmonella,* no other organisms had the capability for respiring with low affinity oxidases, but we infer *Enterococcus and Lactobacillus* have the capability to reduce low levels of oxygen for detoxification (due to the absence of complex I in electron transport chain) while *Akkermansia muciniphila* and *Muribaculaceae* likely respire low levels of oxygen using high affinity oxygenase. Similarly, we observed genes for detoxifying reactive oxidative damage (SOD, catalase, thioredoxin reductase) were more enriched in the inflamed community than the uninfected community. Together these findings demonstrated organisms co-existing with *Salmonella* in the inflamed gut encode the metabolic abilities to withstand or leverage the oxidative redox conditions caused by inflammation (Fig. 5C). Markedly, there were members in the uninflamed gut with respiratory metabolic potential that were not maintained in the inflamed gut (*Duncaniella sp, Hungatella_A sp*), demonstrating there are other selective forces besides the ability to respire that dictate persistence in response to pathogen colonization (Fig. 5C).

Prior reports by our team and others demonstrated that butyrate, a key gut short chain fatty acid (SCFA), decreased by 15-fold in the *Salmonella* inflamed gut, most likely due to inflammation induced redox changes with detrimental impacts on members of the class *Clostridia* [13,27]. Here we sought to better understand the relationship between taxonomy and SCFA production potential. In uninflamed communities the most prevalent butyrate producing bacteria were members of the *Alistipes* and *Lachonospiraceae,* members of classes *Bacteroidia* and *Clostridia* respectively. Interestingly, while the most dominant *Clostridia* did decrease in relative abundance with inflammation, replacement *Clostridia*

members (*Lachnospiracea*, *Dorea*, *Faecalicatena*) were enriched which encoded overlapping butyrate production potential. For example, a MAG belonging to the genus *Dorea* within the *Clostridia* was enriched 16-fold and likely most contributed to butyrate production stability, while the dominant *Alistipes* MAG (a member of the *Bacteroidia*) reduced in abundance by a third was not replaced by taxonomically similar members. Together, these data suggest the notion that decreased butyrate concentrations observed in the CBA/J mouse model during *Salmonella* infection [27] may be attributed to *Bacteroidia* reduction and less so to *Clostridia*, a hypothesis needing further validation using gene expression to track butyrate production and consumption activities in the inflamed gut.

Salmonella-induced inflammation alters carbon usage patterns with more favorable redox conditions enabling the use of less energetically favorable substrates like 1,2-propanediol and ethanolamine [47,48]. While *Salmonella* encodes this metabolic capacity, we were interested if any of the other persisting microorganisms could compete for use of these substrates. *Enterococcus_D* and multiple *Oscillospirales* genomes contain genes from the *eut* gene cluster for ethanolamine utilization and *pdu* genes for 1,2-propanediol utilization. These genera increase in relative abundance with inflammation, particularly *Enterococcus_D*, which is one of the next most abundant members after *Salmonella* (expanding to 2.6% of the inflamed community). Additionally, we showed that the polymer utilization profile was also impacted with inflammation, as infected communities can utilize more alpha-galactan and chitin (Fig. 5). In a similar fashion, the community utilization potential of sugars fructose, fucose, and mannose increased with inflammation. Collectively, these data can inform probiotic approaches for controlling *Salmonella* abundance through competitive exclusion targeting select substrate use patterns using inflammation resistant strains.

Next, we quantified genes commonly reported in humans to impact inflammation and examined if they were depleted in this inflamed mouse model. Consistent with literature reporting healthy individuals have a greater potential for tryptophan degradation [49–51], we observed the potential for tryptophanase

mediated conversion of tryptophan to indole by members of *Bacteroidia, Clostridia,* and *Verrucomicrobiae* in both inflamed and uninfected mice. However, the proportion of *Bacteroidia* with this gene was much lower in inflamed guts (Fig. 6B). Tryptophan Indole/AhR pathway representation in infected mice is concurrent with lower proportions of *Verrucomicrobiae* and *Bacteroidia* spp. (Fig. 6). Also, like human microbiomes, we observed microbial genes responsible for cleaving taurine or glycine from primary bile acids and metabolizing secondary bile acid products (*bsh*, *baiN*, *baiA*, and *hdhA*) were significantly lower in relative abundance in mice infected with *Salmonella* (Fig. 6). These data provide promising insights that the functional gene profiles for modulating inflammation are present in the CBA/J model, and may suggest its relevance for study of similar inflammation-associated mechanisms in humans.

2.3.5 Viral AMGs contribute to the bacterial community functional potential in CBA/J mice via Firmicutes

In the creation of the first murine gut viral database, we sought to compare viral genomic content cataloged here to other mammalian gut systems. Of the 609 dereplicated vMAGs that were recovered from both treatments, less than 1% had taxonomic assignments (Additional file 3: Data S4). These three vMAGs were assigned to the Caudovirales in the families Siphoviridae (n=2), and Myoviridae (n=1). To perform biogeographic analyses, we collated phage genomes previously reported from mammalian guts [24,38,43,52,53] and clustered these with our mouse recovered vMAGs. We found that 322 of the CBA/J derived vMAGs (53%) had similar representatives in other phage gut metagenome studies, meaning over half of our vMAGs clustered with viruses from at least one additional study (Fig. 4D). This suggests a potentially cosmopolitan phage seedbank that may be conserved across a wide variety of animals, geographies and, in the case of humans, ethnicities and health statuses. Ultimately, viral content in the CBAJ-DB can have relevance to other mouse models and human guts.

To explore if viral communities could potentially influence the structure and function of the CBAJ-DB uninflamed and inflamed microbial communities, we verified that microbial and viral genome-

based ordinations were coordinated (Fig. 11). With informatics we conservatively determined that of the 609 vMAGs, 11.5% were putatively linked to 43 MAGs that encompassed 27 unique taxonomies (Fig. 11). All putative hosts corresponded to members of the *Firmicutes*, and included members of the *Lachnospiraceae*, *Ruminococcaceae*, *Oscillospiraceae*, *Anaerotignaceae*, *Borkfalkiaceae* and *Acutalibacteraceae* families. Among the vMAGs that putatively infected hosts, we identified 36 auxiliary metabolic genes (AMGs) with functionalities including regulation of the TCA cycle (citrate synthase), glycolysis (orthophosphate dikinase), phosphate metabolism (PhoH), and oxidative stress response (rubrerythrin). These phage genomes also encoded AMGs for the induction of germination (Peptidase A25), spore formation (M50B), the cleavage of amorphous cellulose (GH2), and low pH resistance (ornithine carbamoyltransferase). Among the putative viral hosts were members within the *Clostridia* class, exhibiting some of the largest MAG relative abundance differences between inflammation states. For example, *Dorea* and *Faecalicatena* enriched in inflamed mice, and *Lachnospiraceae* COE1 enriched in uninfected mice. Together these findings indicate phages may be underappreciated top down (predation) and bottom up (resource) controllers of microbiota functionality in the murine gut.

2.4 Discussion

2.4.1 Perturbation expanded the microbial and viral genomic cataloging of the CBA/J gut microbiome

Genome resolved catalogs like CBAJ-DB are valuable resources for interrogating metabolic potential of the microbiome, yet these databases are biased by the by environment, organism, or disease state they were generated from; and host associated microbiomes can vary drastically between different species, model organisms, and even within an individual [54–56]. Previous murine gut bacterial databases lacked membership from inflamed individuals and CBA/J mice, and none have curated viral content [21,22,57], underscoring a clear value of this resource to the community. Our findings highlight the power of using biological perturbation, in this case *Salmonella*-induced inflammation, to genomically sample taxa that are conditionally rare and obscured by their low abundance only to be critical contributors to ecosystem functionality under altered states.

While assemblies and binning are prone to missing key lineages or under-sampling diversity, we used paired 16S rRNA analyses to affirm the representation of critical community members in the inflamed and uninfected gut. Our paired amplicon sequencing indicates the CBAJ-DB contains membership similar to previously reported *Salmonella*-inflamed CBA/J communities and proportional representation of similar bacteria to those found in other mouse breeds [13,27,58]. It is our intent to create a resource for other researchers conducting microbiome analyses using CBA/J mice, such that this genome content can be accessed by taxonomic naming or linkages to the 16S rRNA gene. Likewise, this genome library can be used by others for read recruitment of future metagenome and metatranscriptome sequencing, or to substantiate metabolomic insights from the CBA/J microbiome.

The vMAG database also provides interesting context for researchers in other mammalian gut habitats, but especially mouse gut which has been historically under-sampled in this regard. While collections of human gut viruses are available, the mouse virome is understudied [24,59]. This collection of over 4,000 vMAGs contains a significant number of cosmopolitan genera also found in other mammalian guts including humans. The existence of a core mammalian gut virome is an exciting proposition that alludes to an intricacy in gut community function and begs further exploration. The gut microbiome is a complex system involving the interplay of host, microbes, and abiotic elements [60–62]. Beyond functional characterization of gut communities during *Salmonella* infection, the CBAJ-DB offers a bacterial and viral resource for holistic microbiome study in a common mouse model.

2.4.2 A genome resolved inventory of functional potential changes in the pathogen inflamed gut

Previous studies of gut microbiomes highlight the role microbial metabolites play in gastrointestinal inflammation as signaling effectors in host immune regulation [63,64].
The CBAJ-DB showcases the juxtapose of pro-inflammatory and anti-inflammatory microbial membership and gene content in CBA/J mice and during *Salmonella* infection. *Salmonella*-induced

inflammation shifted the functional potential of infected communities favoring respiring organisms, marked by a reduction of butyrate and acetate producers and an increase in bacteria with anaerobic respiration capability. Convention indicates butyrate agonism of PPAR-γ and SCFA engagement with G-protein coupled receptors respectively help to maintain luminal anaerobiosis and promote colonic regulatory T cell development [6,65,66]. We showed specific bacteria reduction coincides with an inflamed state and diminished SCFA production potential in the gut.

Our findings indicate *Alistipes* reduction following inflammation as a possible cause of butyrate production potential loss, contrasting with current dogma linking butyrate production in the gut chiefly to *Clostridia* abundance [11,67–69]. Furthermore, *Salmonella* infection enriched certain *Clostridia* including *Dorea, Faecalicatena*, and a novel bacteria described only at the class level, highlighting the functional redundancy that may be provided within this class. Interestingly, the mouse with the lowest *Salmonella* relative abundance had the greatest diversity of *Clostridia*. It is interesting to speculate that this lineage and the diversity within it, may be important for microbiota recolonization and return to homeostasis following gastric infection, a notion supported by previous research [70].

Salmonella-induced inflammation also caused a reduction in bacteria with the capability to mediate anti-inflammatory microbial metabolites. Bacteria with genes for secondary bile acid production and tryptophan catabolism were decreased in inflamed metagenomes, a response previously shown to increase host susceptibility to infection [71]. Specifically, reduced bile acid can limit ligands like pregnane X (PXR) and farnesoid X (FXR) which are important regulators of the host anti-inflammatory response, thus reduced production of these genes could have further feedback on inflammation [64,65]. Similarly, indole and indole derivatives like indole acrylic acid, indole-3-probionate, and indole-3-lactate are AhR and PXR agonists and thus anti-inflammatory [49,72]. Here we demonstrate how the reduction of bacteria capable of producing these important host pathway modulators can further promote inflammation and *Salmonella* expansion, evidenced by high lipocalin-2 levels concomitant with high *Salmonella* relative

abundance and lower functional potential for bile acid deconjugation, secondary bile acid production, and tryptophan catabolism in inflamed mice. Future research directly measuring transcription and metabolite concentrations can be used in concert with the CBAJ-DB to determine the anti-inflammatory impact of individual bacteria on the microbiome.

2.4.3 Commensal bacteria that can withstand inflammation may represent future biological therapeutic opportunities

A recent rise of antibiotic resistant *Salmonella* strains globally underpins the need for alternative treatments and prevention measures against foodborne pathogens. One avenue may be the use of probiotic bacteria to reduce the intensity or duration of infection [49,73,74]. *Alistipes* sp002428825 and *Akkermansia muciniphila* clustered closely with genomes from human microbial communities and we explore here their potential as anti-inflammatory probiotics.

Akkermansia muciniphila is a well-known commensal gut bacterium in mammals that lives in the lumen mucosal layer and contributes to epithelial gut barrier integrity [75,76]. Nevertheless, one study showed *Akkermansia muciniphila* exacerbated inflammation and increased *Salmonella typhimurium* relative abundance [77]. These findings are inconsistent with our data however, where *Akkermansia muciniphila* is relatively abundant and consistently present in uninfected and *Salmonella*-infected mice. Genome analysis reveals the capacity of *Akkermansia muciniphila* to produce indole, potentially an important anti-inflammatory mechanism and *Salmonella* deterrent. Other studies have shown the effectiveness of indole from *E. coli* increasing tight junctions of the gut epithelial and decreasing *Salmonella* pathogenicity, though more work is needed to confirm similar action from *Akkermansia* in buffering gut inflammation [78].

Alistipes sp002428825 was also consistently detected in both treatments. Analysis of this genome suggests it can respire oxygen and directly compete with *Salmonella* for arabinan, arabinose, and pectin, whilst maintaining critical gut homeostasis functionalities like butyrate production. *Alistipes* spp. are

often associated with healthy human microbiomes [79,80] and have even been shown to facilitate microbiota recolonization following perturbation [81]. We have shown *Alistipes* has saccharolytic genes to metabolize rhamnose and fucose and it is possible that Gram-positive bacteria eliminated by inflammation and fucosylated proteins from host epithelial erosion could support *Alistipes* during *Salmonella* infection as sources of these sugars [82,83]. Given this genus, like *Akkermansia muciniphila,* remained abundant despite host inflammation, and both bacteria contain genes for indole derivative production, it also may be directly antagonistic to *Salmonella* through anti-inflammatory effector potential.

It also bears mentioning the presence of lactic acid bacteria *E. gallinarum* and *L. johnsonii* in the CBAJ-DB in relatively high abundance during *Salmonella* infection. In fact, these members were previously illuminated in prior studies from mice as well as clinical patients, demonstrating co-occurrence may exist beyond this single study or mouse model [84]. Beyond resisting or maybe even responding to conditions created by *Salmonella* infection, we provide genomic evidence supporting *L. johnsonii* nutrient competition for arabinose, mixed-linkage glucans, and amorphous cellulose, while *E. gallinarum* may compete for chitin, pectin, and arabinan. Additionally, closely related species to these have been shown to produce anti-*Salmonella* agents like bacteriocin and organic acids putatively decreasing *Salmonella* abundance over time [85–87]. Given these species have many members already approved as probiotics and our data indicate CBA/J mice harbor at least one species of *Enterococcus* common to human guts and *Lactobacillus* spp. resistant to host inflammation, a probiotic lactic acid bacteria strain resistant to *Salmonella* may reside in the CBAJ-DB [88]. This genome resolved research identifies future targets with promising potential for exploration as probiotics robust to *Salmonella* infection. Yet, we recognize that commensal bacteria like these can be pathobionts provided the right setting [54,77,89,90], such that future research in this mouse model would first include challenge experiments with isolates or even targeted consortia that provide multiple avenues of overlapping pathogen colonization resistance.

2.5 Conclusion

The CBAJ-DB uniquely captures gut community variation in CBA/J mice. MAG reconstruction from metagenomic sequencing enabled us to profile the functional potential of the murine gut microbiome during acute *Salmonella* inflammation, contrasting community membership and gene content with uninfected mice. Persisting taxa in the inflamed gut encoded the capability to withstand or utilize changing redox conditions, while bacteria producing SCFA and producing host anti-inflammatory effectors decreased in mice with high *Salmonella* burden. Further, our phage analyses leave open the possibility that phage infection could alter *Firmicutes* energy regulation and spore formation.

Together, these findings validate physiological investigations performed with reduced complexity synthetic or modified gut microbiota. We also provide new perspectives that advance the understanding of *Salmonella* effect on an intact microbiome and provide model-specificity to CBA/J gut consortia. Our efforts show novel bacteria unique to the CBA/J mouse model and enriched by *Salmonella* infection. An exploration of potential probiotic targets in the CBAJ-DB revealed multiple lactic acid bacteria capable of withstanding the host immune response to *Salmonella* and that may be indifferent to *Salmonella* competition. Additionally, genomes were recovered of *A. muciniphila* and *E. gallinarum* with species similarity to bacteria in human guts. The CBAJ-DB is the first culture-independent murine genome catalog to include sampling from CBA/J mice and inflamed individuals, providing a resource with application to multi-omic microbiome investigation, gut inflammation research, and studies involving the CBA/J mouse model broadly.

2.6 Methods

2.6.1 Strains and media

S. enterica serovar Typhimurium strain 14028 (*S. typhimurium* 14028) was cultured overnight in Luria-Bertani (LB) broth at 37 °C with constant agitation. Overnight culture was washed and resuspended

in water. *S. typhimurium* 14028 ASV was determined manually from an identical sequence match to the 16S region of NCBI Reference Sequence NZ_CP034230.1 mapped with Geneious Prime® 2020.1.2.

2.6.2 Animals and experimental design

Female CBA/J mice were purchased from The Jackson Laboratory (Bar Harbor, ME) and housed 5 per cage in conventional enclosures maintained in a temperature controlled 12 hour light/dark cycle. To mitigate microbiome differences caused by variables other than *Salmonella*, we housed all mice in the same room, and mice were chosen at random when populating cages and assigning treatment. Irradiated mouse chow (Teklad, 7912) was made available *ad libitum* to mice in the control group (n=16) and to infected mice (n= 14) housed separately. Individuals in this study were chosen based on fecal sample availability at Day 11 and only infected mice with ≥25% *S. typhimurium* 14028 at the sampling time were used. Mice in the infected group received 10^9 CFU *S. typhimurium* 14028 oral gavage on Day 0 with no subsequent treatment, and control group mice were left without treatment. Animal experiment protocol was approved by The Ohio State University Institutional Animal Care and Use Committee (IACUC; OSU 2009A0035).

2.6.3 Sample collection

Fecal pellets were collected from each mouse one day prior to treatment and 10 and 11 days after treatment initiation on autoclaved aluminum foil. Pellets were immediately placed in labeled microcentrifuge tubes and flash frozen with EtOH/dry-ice prior to storage at -80 °C until further processing.

2.6.4 Lipocalin-2 quantification

Vortex homogenization of fecal sample in PBS containing 0.1% Tween 20 (100 mg/ml) for 20 min was performed prior to centrifugation of the resultant suspension at 12,000 rpm for 10 min at 4 °C. The resulting supernatant was used to measure levels of inflammation marker Lipocalin-2 using the Duoset murine Lcn-2 ELISA kit (R&D Systems, Minneapolis, MN).

2.6.5 DNA extraction and sequencing

Total nucleic acids were extracted using the *Quick*-DNA Fecal/Soil Microbe Microprep Kit (Zymo Research) and stored at −20 °C until amplicon or metagenomic sequencing. DNA was submitted for amplicon sequencing at Argonne National Lab at the Next Generation Sequencing facility using Illumina MiSeq with 2 × 251 bp paired end reads following established HMP protocols [91]. Briefly, universal primers 515F and 806R were used for PCR amplification of the V4 hypervariable region of 16S rRNA gene using 30 cycles. The 515F primer contained a unique sequence tag to barcode each sample. Both primers contained sequencer adapter regions. DNA for metagenomes was submitted to the Genomics Shared Resource facility at Ohio State University and was prepared for sequencing with a Nextera XT library system followed by solid-phase reversible immobilization size selection. Libraries were quantified and then sequenced using an Illumina HiSeq platform.

2.6.6 16S rRNA preprocessing

Amplicon sequencing fastq data were processed in a QIIME2 2019.10.0 environment [92]. Reads were demultiplexed and then denoised with DADA2 [93]. For all sequencing runs (n=4), forward reads were truncated at 246 bps and reverse reads were truncated at 167 bps. Feature tables from each sequencing run were combined and ASVs were assigned taxonomy with the silva-132-99-515-806-nb-classifier [94]. Before further analysis the ASV table was filtered with R version 4.0.2 to (i) remove samples with no ASVs, (ii) to remove samples with a combined ASV count <1000 across all samples, (iii) to remove ASVs with 0 abundance in every sample, and (iv) to remove ASVs designated as mitochondria and chloroplast. The resulting filtered feature table contained 23,022 ASVs. Raw reads are deposited on NCBI (PRJNA348350) and the final ASV table is published in the supplementary materials (Additional file1: Data S1). Reads for external study comparison were obtained from the NCBI Sequence Read Archive under accession number SRP057511 (O'loughlin [28]) and under bioproject PRJNA348350 (Borton [27]). All 16S rRNA amplicon data for comparison analyses were treated as described above. ASV relative

abundances were then summed within genera or the lowest annotated taxonomy level when no classification was provided.

2.6.7 Genome reconstruction from metagenomes

All bioinformatic tools were run with default parameters unless otherwise specified. Quality scores of raw metagenomic reads were evaluated using FastQC (v0.11.9, [95]). Reads were trimmed, adapters removed, and mouse reads removed using BBDuk (ktrim=r, k=23, mink=11, hdist=1, qtrim=rl, trimq=20, minlen=75, maq=10) from BBTools (v38.89, https://jgi.doe.gov/data-and-tools/bbtools). Trimmed reads from each individual sample were assembled with Megahit (v1.1.1, [96]) and with IDBA_UD (v1.1.3, [97]). Each assembler was also used to perform co-assembly of all the samples (n=6) at once.

Subsequently, each single-sample and co-sample assembly was binned separately. To obtain bins, reads were mapped to assembly contig or scaffold set filtered to ≥2500 bps using BBMap (https://jgi.doe.gov/data-and-tools/bbtools), sorted using SAMtools (v1.9, [98]). and then binned with Metabat2 (v2.12.1, [99]). The resulting MAGs were checked for quality and contamination with CheckM (v1.1.2, [30]), CheckM2 (v0.1.3, [31]), and GUNC (v1.0.5, [32]). Using the resulting medium quality (completeness ≥ 50%, contamination < 10%) and high-quality (completeness ≥ 90%, contamination < 5%) MAGs from the initial single-sample and co-sample assemblies, trimmed reads were mapped using BBMap in perfect mode. Reads that did not map were assembled individually by-sample and co-assembled by treatment with IDBA_UD (v1.1.3, [97]) to create subtractive assemblies. The resulting subtractive contigs or scaffolds were then subject to previously described filtering (≥2500 bps), processing, binning, and quality check.

To construct the final CBAJ-DB, all MQHQ MAGs from all assemblies were assigned taxonomy with GTDB-Tk (v1.3.0, r95, [100], classify_wf) and dereplicated with dRep (v2.5.4, 99% ANI, [101]).

Conventional assembly and binning of reads rarified with BBTools for MAG recovery sensitivity to read depth was performed with the Snakemake [102] workflow included here (https://github.com/ileleiwi/metaG_pipeline/blob/main/Snakefile_5.88_onesamp).

2.6.8 16S rRNA linked to MAGs

Mining of MAGs for 16S rRNA genes was performed with Wrighton Lab software (https://github.com/WrightonLabCSU/join_asvbins) MMseqs2 (v13.45111, [103]) and Barrnap (v0.9, [104]) and SILVA reference database (SILVA_138_SSURef_NR99_tax_silva, [94]) followed by pairwise comparison to the V4 region sequences from our amplicon sequencing.

2.6.9 Database mapping and comparison to other MAG resources

To calculate relative abundance of individual dMQHQ MAGs, BBmap (https://jgi.doe.gov/data-and-tools/bbtools) was used to randomly map trimmed reads to the dMQHQ with minid = .95. Then CoverM (v0.6.0, https://github.com/wwood/CoverM) was used to estimate read counts per scaffold and per bin. Values are the mean number of aligned reads calculated after removing positions with the most and the least coverage as determined by default values (methods:–proper-pairs-only -m trimmed_mean –min-read-percent-identity-pair 0.95, --proper-pairs-only -m mean –min-read-percent-identity-pair 0.95 –min-covered-fraction 0.75, --proper-pairs-only -m reads_per_base –min-read-percent-identity-pair 0.95 –min-covered-fraction 0). Relative abundances of mapped reads were then GeTMM normalized (Additional file 2: Data S2) [105] using the edgeR package (v3.36.0, [106]). To place MAGs in either treatment (control, infected) or both treatments, CoverM was used and mapping was only considered if the subject covered fraction exceeded 75% with 95% sequence identity and a minimum of 3 reads per base depth.

To estimate the relative abundance of each vMAG, the metagenomic reads were mapped using Bowtie2 [107]. Reads were mapped using BBMap with minid = 0.95. Afterwards, CoverM was run using the -mean option to consider only those vMAGs that have >75% of their fraction covered. Relative abundances for each vMAG were calculated as their coverage proportion from the sum of the whole

coverage of all bins for each set of metagenomic reads prior to GeTMM normalization and are reported in Additional file 3: Data S4.

Isolate MAGs were obtained from the Human Microbiome Project [39] and medium and high-quality MAGs were obtained from https://doi.org/10.5281/zenodo.4977712. Metagenomic reads from the Human cohort and Lloyd-Price cohort [43] were obtained from PRJNA725020 and the SRA Database Commons respectively. Human cohort rarefication to 2Gbps was performed with BBTools reformat. Human metagenomic reads were mapped to MQHQ bins with fastANI (v1.32, [108]). Genome matches with ANI ≥ 94% and alignment fraction (AF) ≥ 33% were deemed of the same species and genome matches with ANI ≥ 73% and AF ≥ 33% of the same Genus [109].

Comparison to mouse database MAGs was performed on MGBC [22] non-redundant high-quality genomes (Additional file 4: Data S5, MGBC-hqnr_26640) and iMGMC [21] dereplicated medium and high-quality MAGs (Additional file 4: Data S5, iMGMC-mMAGs-dereplicated_genomes). iMGMC MAGs were re-classified with GTDB-Tk version 1.3.0 r95 to align with CBAJ-DB and MGCB taxonomy, and dereplication with each MAG set and CBAJ-DB was performed with dRep (v2.5.4, 99% ANI).

2.6.10 MAG function analysis

CBAJ-DB MAGs were annotated using DRAM (v1.1.1) [34] and dbCAN (v3.0.2, dbCAN-HMMdb-V10, [110]). CAZy genes called with DRAM were updated with the latest dbCAN database and parsed with HMMer (v3.3, [111]) to include only significant hits (e-value < 10^{-18}) with > 35% coverage. Wrighton lab software rule_adjectives (https://github.com/rmFlynn/rule_adjectives) and functions_pa (https://github.com/ileleiwi/functions_pa) were used to parse KEGG ids, Enzyme Commission numbers, and dbCAN ids from gene annotations referencing function rule sheets (Additional file 6: Data S7) to determine function presence or absence in each MAG. Function relative abundance significance was determined with ANOVA performed using the R stat package (v4.1.3).

2.6.11 Viral host-linkage and MAGs

Metagenomic assemblies and subassemblies (n=16) were screened for DNA viral sequences using VirSorter2 (v2.2.2, [112]) using the published protocol in Protocols.io [113]. Briefly, VirSorter2 was run using parameters "–include-groups dsDNAphage, ssDNA", "--min-length 10000", and "–min-score 0.5". The resulting VirSorter2 output was then run through CheckV [114] to ensure quality viral sequences using the "end_to_end" function. The trimmed viral sequences output by CheckV were then once again run through VirSorter2 using options above with additional flags "--seqname-suffix-off" "--viral-gene-enrich-off", and "--prep-for-dramv". The final output was then manually curated using the CheckV output as described in Protocols.io. Briefly, 1) viral-like scaffolds that had more than 1 viral gene were kept and deemed viral 2) viral-like scaffolds with no viral genes, host genes equal to zero, or score ≥0.95, or that had 1 host gene with a length of ≥10kb were separated and further inspected. Scaffolds not meeting the above criteria or manually inspected to be non-viral were discarded. After generation of a curated, quality-controlled viral vMAGs, they were clustered at 95% identity across 85% of the shortest contig representing viral populations using ClusterGenomes [113].

To determine taxonomic affiliation, vMAGs were clustered to viruses belonging to viral reference taxonomy databases NCBI Bacterial and Archaeal Viral RefSeq V85 with the International Committee on Taxonomy of Viruses (ICTV) and NCBI Taxonomy using the network-based protein classification software vConTACT2 (v0.9.8, [115]) using default methods [116]. To determine if the viruses present in this study represented relevant communities across mammalian gut ecosystems, we included viruses mined from 5 publicly available datasets in our vConTACT2 analyses from: 1) human guts [24,43,52] , moose rumen [38], and bird / other mammalian guts [53] . The viral sequences that were identified from these systems and the genes used for vConTACT2 are deposited on Zenodo with doi https://doi.org/10.5281/zenodo.7144664 with more information of downloaded datasets found in Additional file 3 Data S4.

Viral contigs were annotated with DRAM-v [34]. Auxiliary scores were assigned by DRAM, based on the following ranking system: A gene is given an auxiliary score of 1 if there is at least one hallmark gene on both the left and right flanks, indicating the gene is likely viral. An auxiliary score of 2 is assigned when the gene has a viral hallmark gene on one flank and a viral-like gene on the other flank. An auxiliary score of 3 is assigned to genes that have a viral-like gene on both flanks. All vMAG annotations are reported in Additional file 3: Data S4. To identify likely vMAG hosts, we used two strategies which included 1) Linking viral spacers found in CRISPR systems assembled using CRASS [117] and 2) Oligonucleotide frequencies between virus and hosts using VirHostMatcher and a threshold of d2* measurements of <0.25 [118]. The lowest d2* value for each viral contig <0.2 was used, and only vMAGs for which the top 3 hits had taxonomic consensus at the genera level were considered "good" hits [118].

2.6.12 Spearman Correlation of Metagenomic and Amplicon Communities

Relative abundance of mapped reads to dMQHQ MAGs were averaged within treatment and the total relative abundance of each MAG in *Bacteriodia, Clostridia, Verrucomicrobiae, Gammaproteobacteria, Bacilli,* and *Coriobacteriia* Classes was summed respectively. Relative abundance of MAGs from all other classes were summed together into an additional category. Similarly, 16S amplicon sequence ASV relative abundances from high responder mice and control mice were summed within each treatment and by taxonomy according to the classes previously mentioned, combining all other ASV abundances into an additional category. Spearman correlation was then performed with the R stats package comparing Class abundances between metagenomic communities and ASV communities within the same treatment.

2.6.13 Statistical analysis

Alpha (Shannon's diversity) and Beta diversity (Bray-Curtis dissimilarity) were calculated with ASV or MAG relative abundances from the filtered data using the Vegan package 2.5.7 in R [119]. The NMDS Beta diversity visualization was produced using ggplot in R [120]. To determine significant grouping

of samples by treatment, we performed an analysis of similarities and mrpp [121], with a stress test determining goodness of fit [121]. Lipocalin-2 (ng/g) and *S. typhimurium* 14028 relative abundance significance between treatments was determined using a Wilcoxon Rank Sum Test, the same test used to compare class relative abundance between different treatments and timepoints. Linear discriminate analysis on MAGs and ASVs was done with LEfSe [122]. MAG and vMAG ordination coordination was determined by Procrustes analysis [119]. MAG quality quartile significance between groups was calculated with a Chi Squared test in

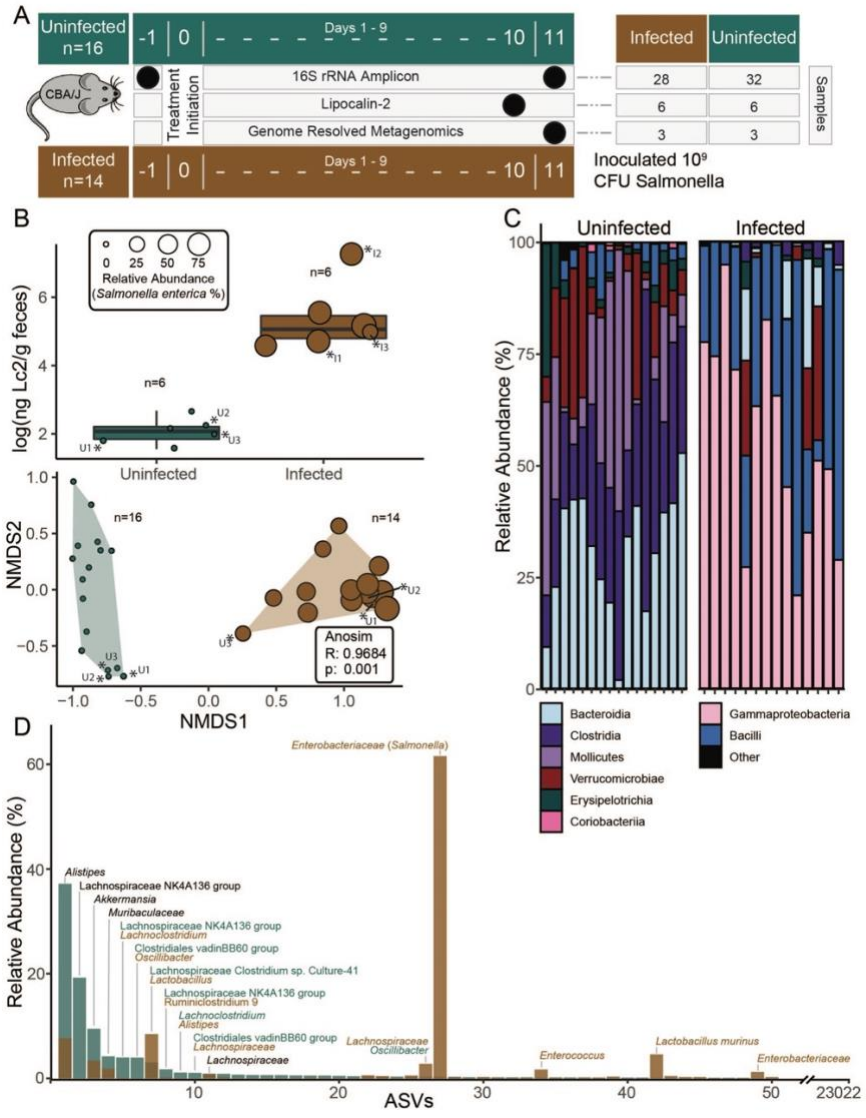


Fig.1 Amplicon sequencing of the CBA/J gut reveals shifts in microbiome composition with inflammation. A) Experimental design shows the number of mice in healthy (green) and infected (brown, infected with 10^9 CFU *Salmonella*) treatments with fecal sampling times and corresponding analysis indicated by black circles. B) Boxplots show lipocalin-2 (Lc2) levels of mice in each treatment during peak infection (top), with *Salmonella* relative abundance indicated by circle size. Non-metric multidimensional scaling of Amplicon Sequence Variant (ASV) Bray-Curtis distances showing significantly distinct communities between treatments at time of sampling with points scaled to *Salmonella* relative abundance (bottom). Asterisks indicate mice used to create the CBAJ-DB. C) ASV Class distribution is depicted by stacked bar charts of healthy and infected communities, with each bar representing a single mouse at the Day 11 timepoint. D) Rank abundance curve of mean ASV relative abundance by treatment in mice sampled to create the CBAJ-DB. Bars represent a single ASV and are colored by treatment, with ASVs ranked separately for each treatment to show the changes in rank and abundance with inflammation. Bars are labeled with taxonomic identity in text if greater than 3% mean relative abundance in either treatment, with text color denoting treatment. Black text indicates both treatments have the same ASV taxonomy within that rank.

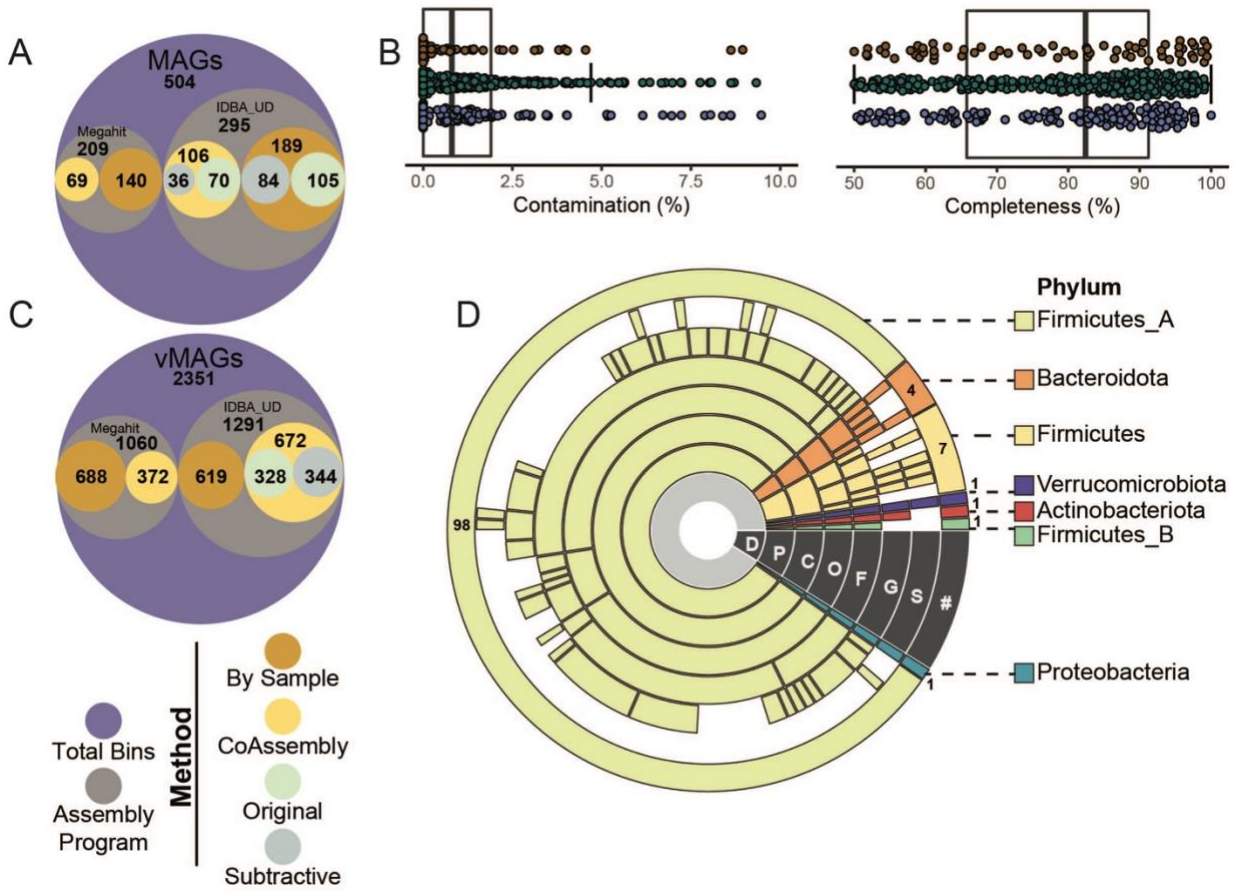


Fig. 2 CBA/J mouse database (CBAJ-DB) genomic methods and composition. A) Circle plot shows the number of medium and high quality (MQHQ) metagenome assembled genomes (MAGs) reconstructed from CBA/J mouse gut metagenomes and corresponding assembly method and assembly software. B) Completeness and contamination of all MQHQ MAGs colored by sample treatment origin (brown = infected; green = uninfected; purple = co-assembly) are shown by box plots, with bold horizontal lines indicating median across all MQHQ MAGs. C) Circle plot shows the number of viral metagenome assembled genomes (vMAGs) reconstructed from CBA/J mouse gut metagenomes and corresponding assembly method and assembly software. D) Dereplicated MQHQ (dMQHQ) MAG phylum distribution is shown by sequential colored rings listed from least specific (Domain, D) to most specific (Species, S) moving outwards from the plot center. Gaps at each taxonomic level represent MAGs that are previously undescribed. The outer ring is labeled with the number of MAGs from the dMQHQ CBAJ-DB database within each Phylum.

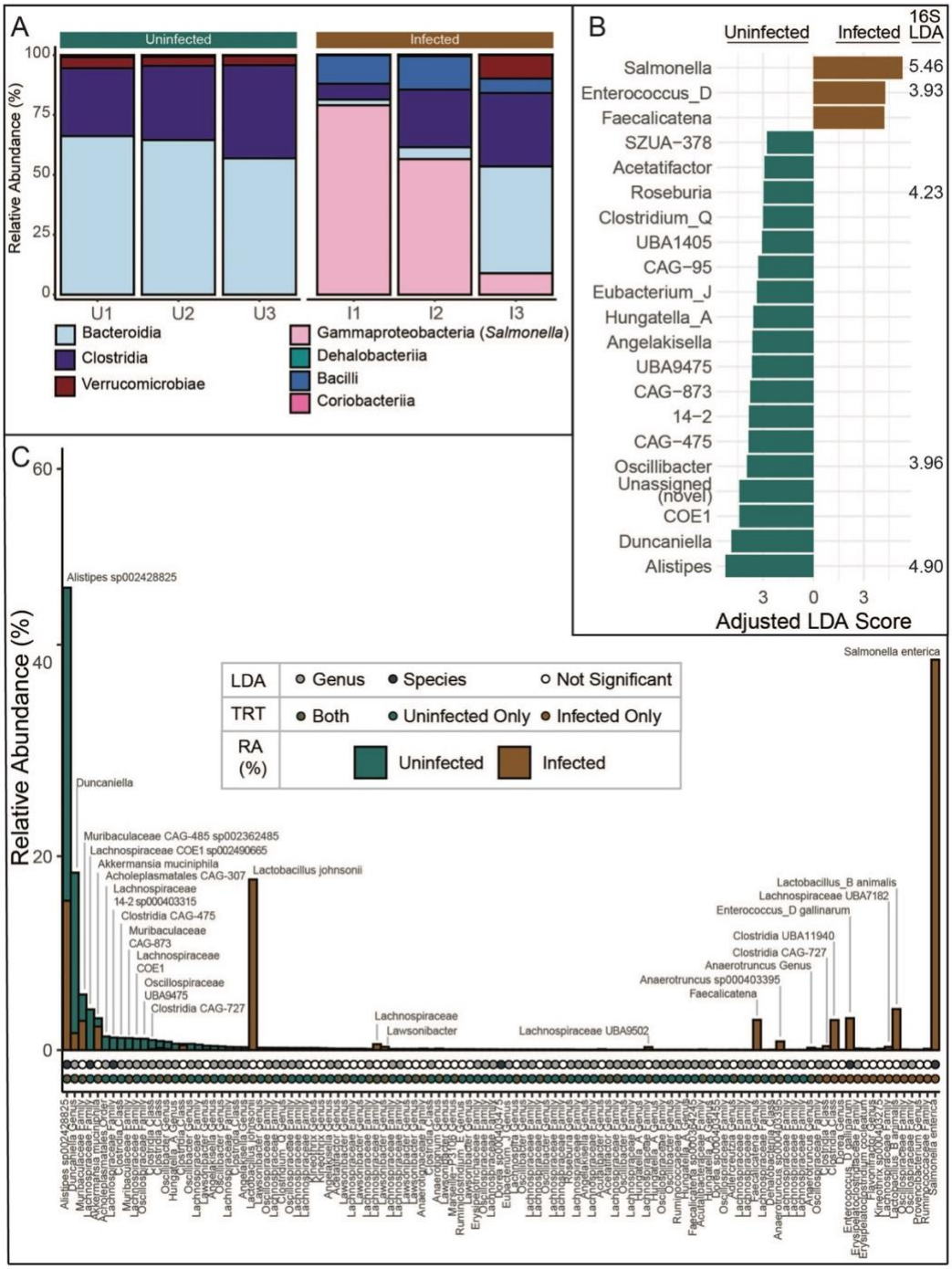


Fig. 3 The CBA/J database (CBAJ-DB) genomes are representative key members in amplicon microbiome data. A) Class distribution of dereplicated MQHQ (dMQHQ) metagenome assembled genome (MAG) relative abundance across individual metagenomes (U# = uninfected, I# = infected) labeled by treatment. B) Linear discriminant analysis Effect Size (LEfSe) scores of most important genera in either treatment in metagenomes, including genera LEfSe linear discriminant analysis (LDA) scores for amplicon sequencing variants (ASVs) with matching taxonomy. C) Rank abundance of dMQHQ MAGs (n=113) showing mean relative abundance (RA) in uninfected (green) and infected (brown) treatments (TRT). Circles below bars highlight LDA significant species (black) and genera (gray) (top row) and treatment origin of each MAG (bottom row). MAGs are labeled with the most resolved GTDB-Tk taxonomy.

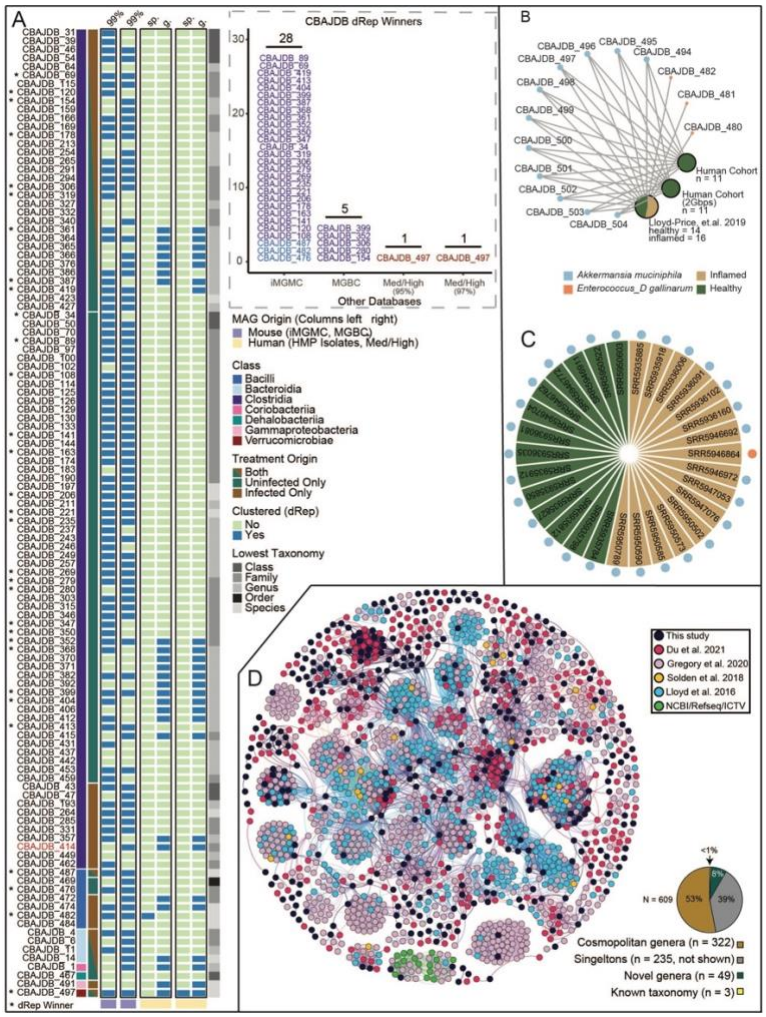


Fig. 4 CBA/J database (CBAJ-DB) genomes link to other murine and human studies. A) Medium and high quality (MQHQ) metagenome assembled genomes (MAGs) that clustered with genomes from either murine databases (light purple columns) or the HMP/Human cohort (yellow columns). CBAJ-DB MAGs that were the highest quality genome in each cluster are marked with an asterisk and are displayed in the stacked bar chart grouped by database. MAGs are grouped by class (first bar annotation) and treatment origin (second bar annotation), with the lowest assigned taxonomy indicated in gray scale (right bar). For each database in black outline, blue cells indicate CBAJ-DB MAGs that clustered, while green cells show no clustering. Accompanying bar chart shows the number of CBAJ-DB MAGs with higher quality scores corresponding genomes in other databases. The MAG indicated with red font in the heatmap was determined to contain possible chimeric by GUNC but not CheckM. B) CBAJ-DB MAGs that recruited human reads are shown by blue (*Akkermansia muciniphila*) and orange nodes (*Enterococcus_D gallinarum*), with size indicating number of CBAJ-DB MAGs. MAG nodes are linked to databases shown as pie charts, where green (healthy human) and tan (inflamed human) indicate sequencing origin of mapped reads and node size indicates sample number. C) SRA accession ID's of samples from the Lloyd-Price cohort that mapped to CBAJ-DB MAGs *Akkermansia muciniphila* (blue) or *Enterococcus_D gallinarum* (orange). D) vContact2 network that shows 609 clustered viral metagenome assembled genomes (vMAG) populations present in CBAJ-DB. The network colors represent the vMAG study origin. Pie chart shows proportion of CBAJ-DB vMAGs that clustered to other studies that are cosmopolitan genera (brown), singletons (gray), novel genera (clusters of >1 vMAGs from this study only, green), or known taxonomy (yellow).

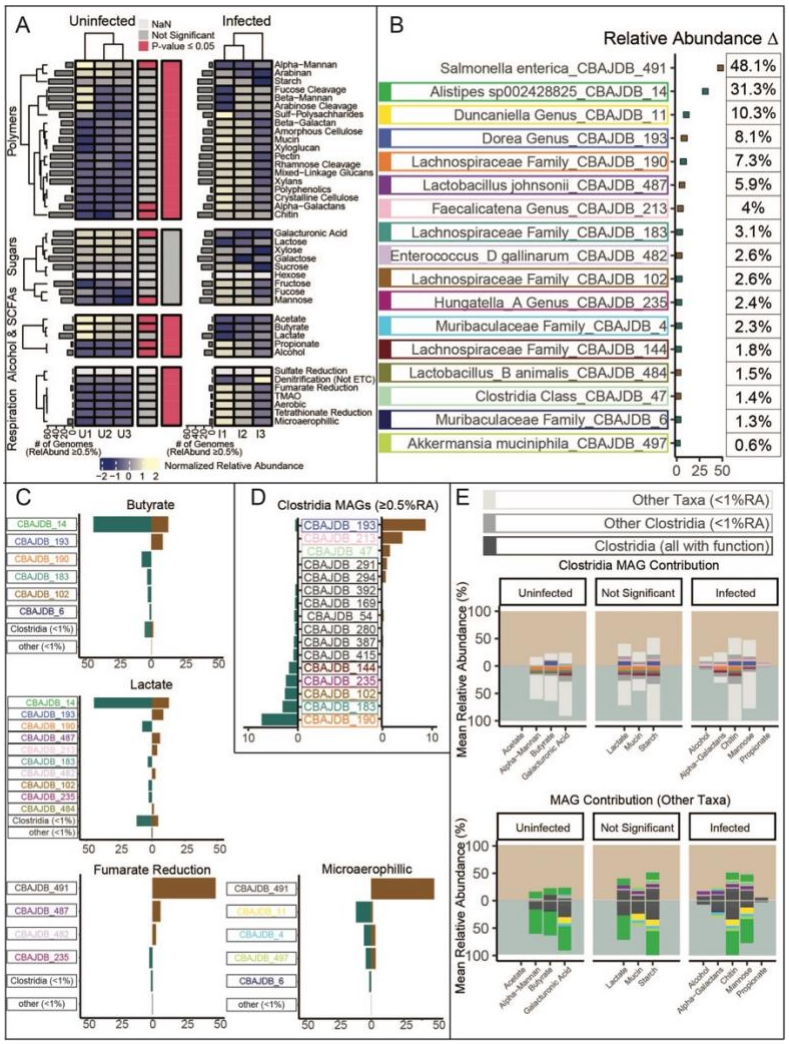


Fig. 5 The CBA/J database (CBAJ-DB) highlights differential metabolisms encoded in uninfected and inflamed mice. A) Normalized relative abundance is shown by a heatmap. Values are mean GeTMM relative abundance of all MAGs with each function center scaled across rows. Functions that are significantly different between treatments as determined by analysis of variance (ANOVA) ($p \leq 0.05$) are indicated by horizontal bars between heatmaps with red highlighting significance at the function (first bar) or functional group level (second bar). Gray bars on each heatmap indicate the number dereplicated medium and high quality (dMQHQ) MAGs that comprise at least 0.5% of the community with a given function. B) Percent change in relative abundance (RA) between treatments the 17 most divergent MAGs. Points are colored by the treatment (uninfected, green; infected, brown) with higher RA. Individual MAGs are uniquely colored by surrounding boxes, acting as color legend for subsequent figure sections. C) Individual MAG contribution to specific functions is shown, with bar magnitude denoting mean RA in either treatment. D) *Clostridia* MAGs with mean RA greater than 0.5% in both treatments are shown. E) *Clostridia* contribution to significant functions (top) and contribution of other taxa (bottom). Plot fields are colored by treatment and bar magnitude indicates mean RA within a treatment.

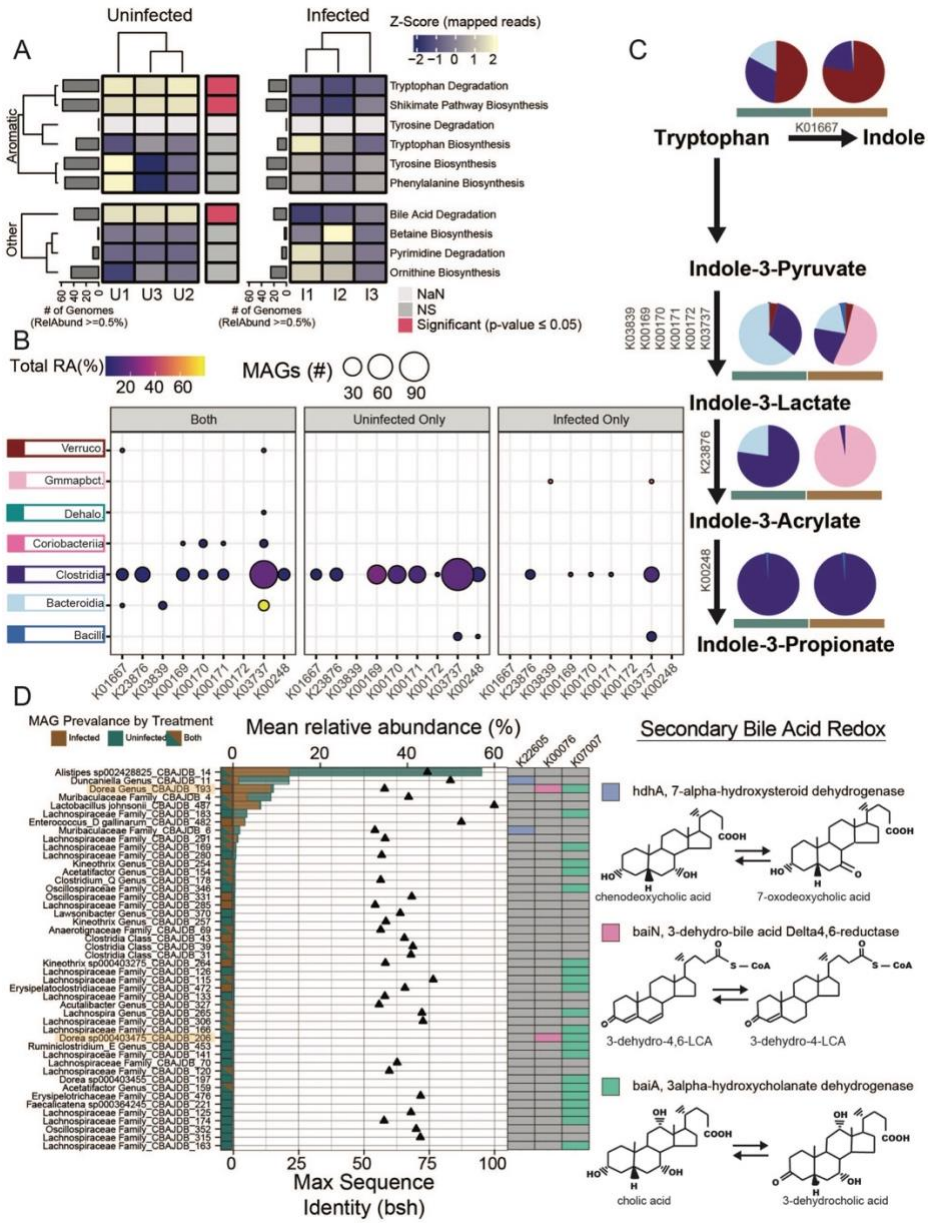


Fig. 6 Tryptophan and bile acid metabolism in inflamed and uninfected gut microbiomes. A) Mean relative abundance summed for each function (rows). Functions that are significantly different between treatments as determined by analysis of variance (ANOVA) ($p \leq 0.05$) are indicated by horizontal bar between heatmaps with red highlighting significance at the function level. Gray bars indicate the number medium and high quality (MQHQ) metagenome assembled genomes (MAGs) that comprise at least 0.5% of the community with a given function. B) Relative abundance (point color) and number of MAGs (size) in each class with each gene for tryptophan degradation separated by MAG presence in each treatment, where both indicates MAGs that recruited strictly mapped reads from both treatments. C) Tryptophan degradation to indole and indole derivatives pathway with pie charts colored by proportion of MAGs in each class (coloring from B) for each treatment. D) Relative abundance (bars) of MAGs in each treatment encoding bile salt hydrolase (bsh, points show sequence similarity to bsh. or *hdhA* (K22605), *baiN* (K00076), or *baiA* (K07007) involved in secondary bile acid metabolism. *Dorea* are highlighted as MAGs with more than one gene for metabolizing secondary bile acid products.

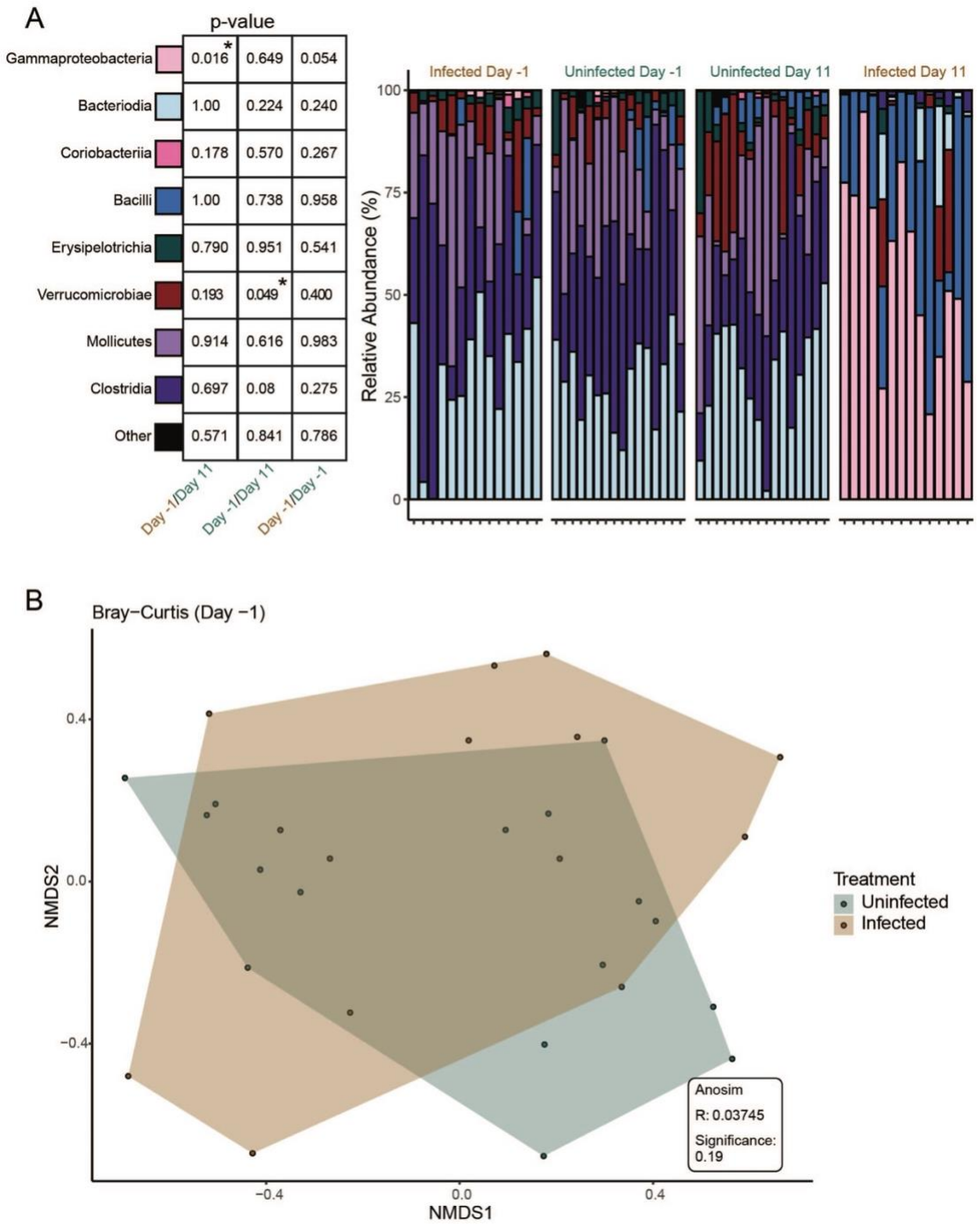


Fig. 7. Relative abundance of classes in pre-infection communities are not statistically different than uninfected communities, indicating a shared starting microbiome prior to infection.
A) Wilcoxon Rank Sum significance revealing no difference between Infected Day -1 communities and uninfected communities from either Day -1 or Day 11. Summed abundance within each class was examined comparing different timepoints and treatments (uninfected = green, infected = brown). B) NMDS of Day -1 communities colored by treatment. No significant difference exists between communities of mice from either treatment at Day -1 as determined by Anosim.

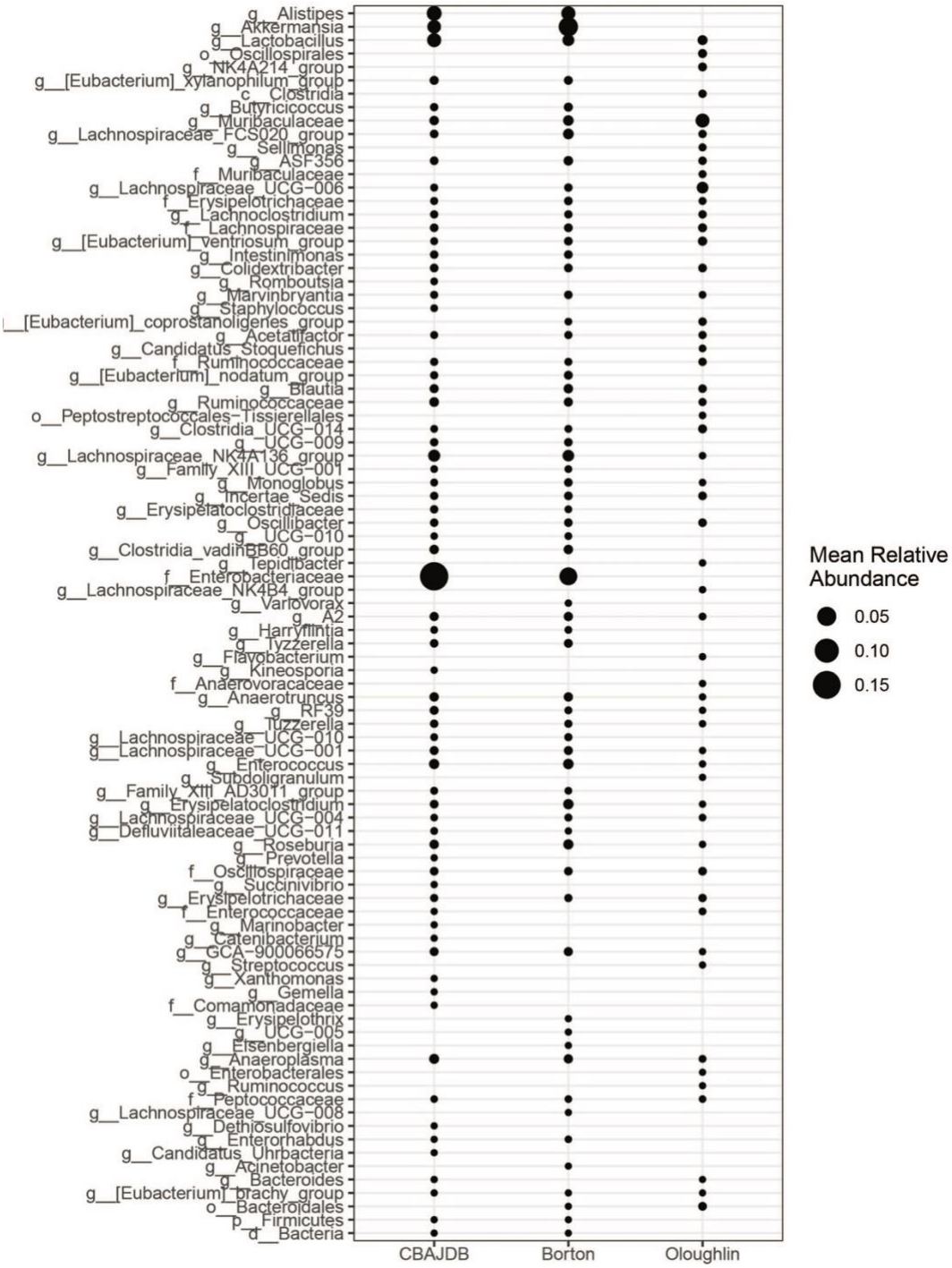


Fig. 8. CBAJ-DB uninfected (no *Salmonella*) amplicon sequenced communities show considerable taxonomic overlap with communities from other CBA studies. Points indicate lowest meaningful taxonomy presence in CBA mice from the CBAJ-DB, Borton et.al. 2017[27], or O'loughlin et.al. 2015[28]. Points are sized by average relative abundance of ASVs contributing to each taxonomy.

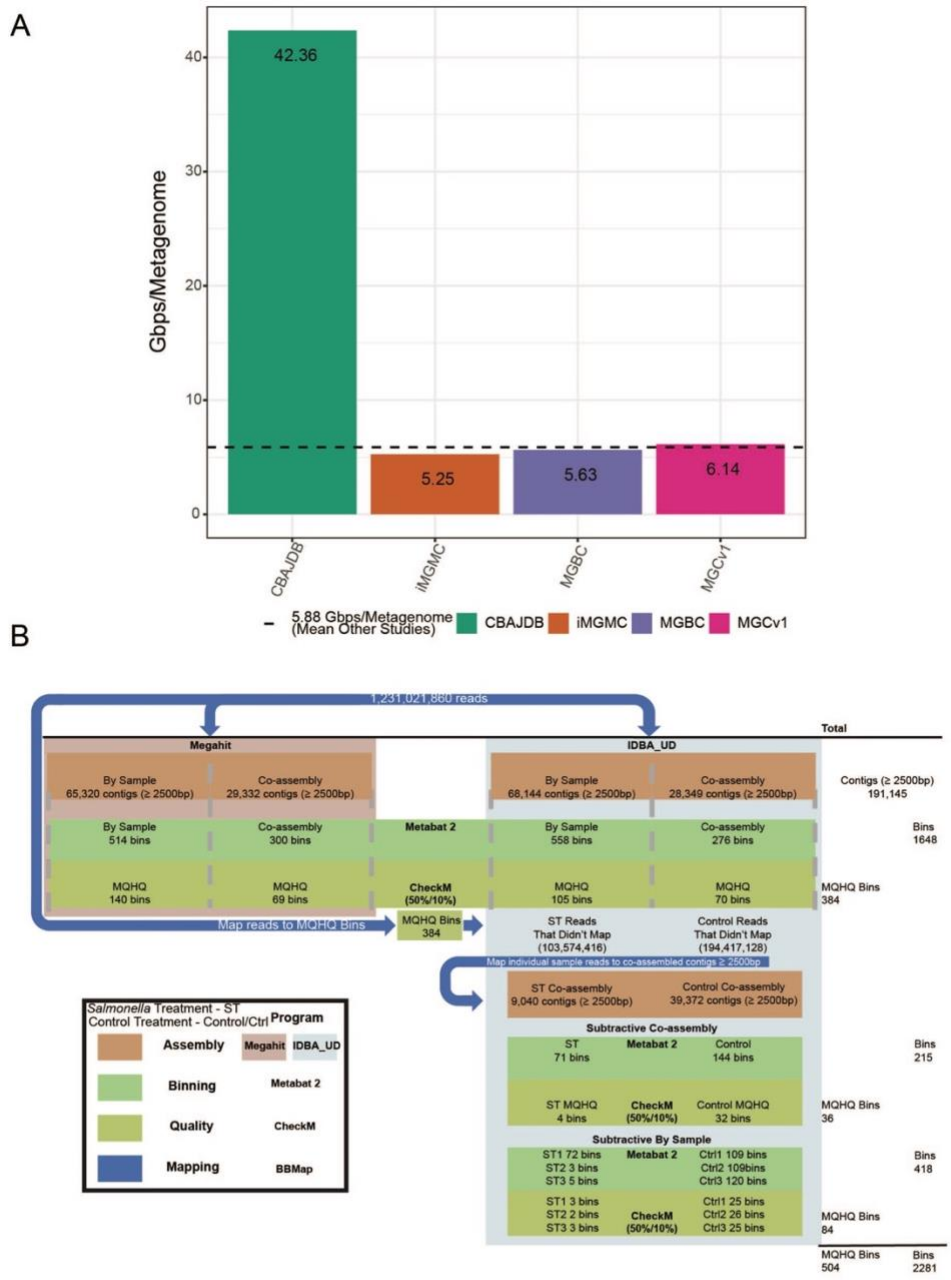
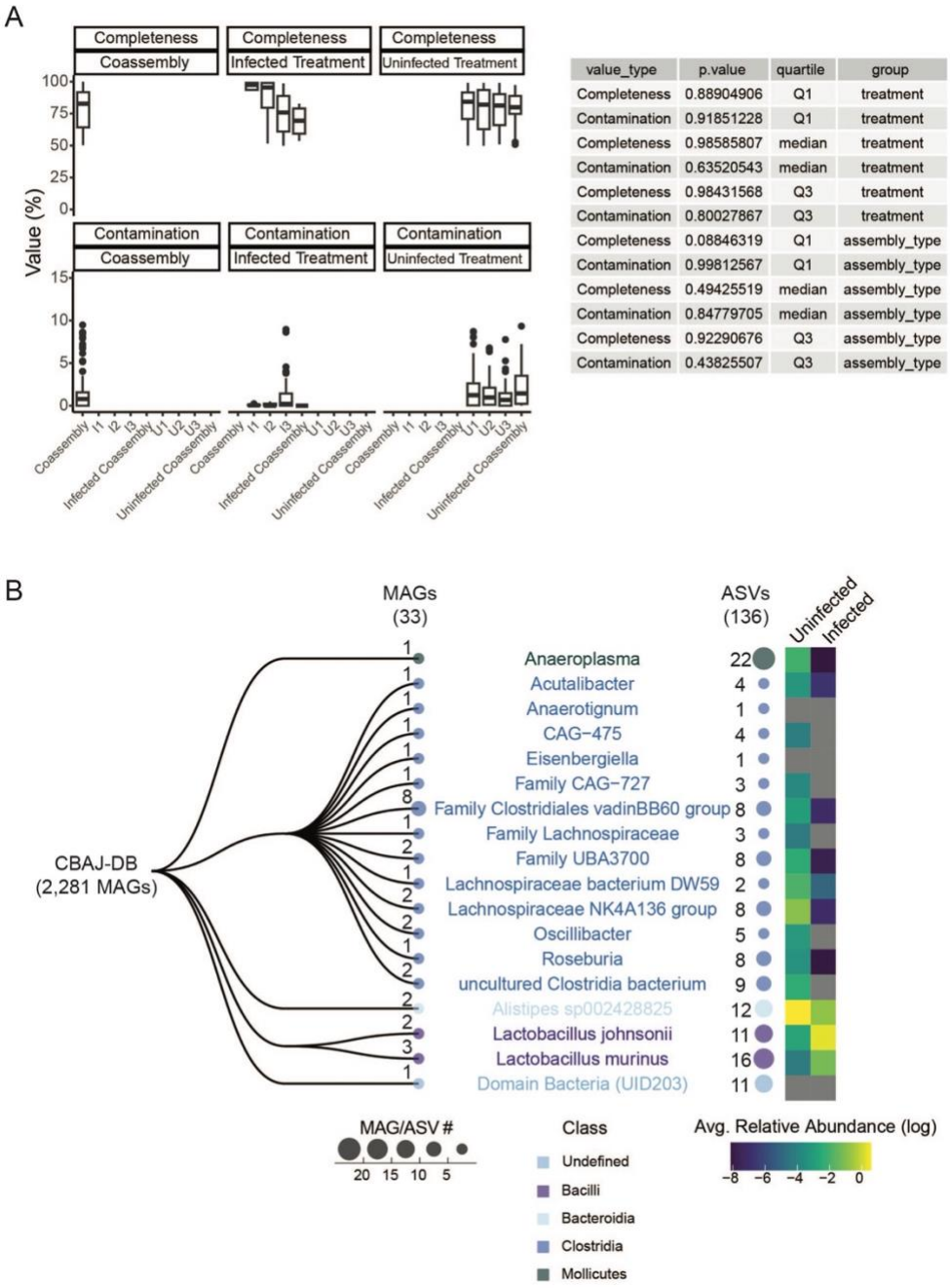


Fig. 9. A) Gigabase pairs (Gbps) per sample of prevalent murine genome databases. integrated Mouse Gut Metagenomic Catalog (iMGMC), The Mouse Gastrointestinal Bacterial Catalogue (MGBC), and Xiao et.al. 2015 (https://doi.org/10.1038/nbt.3353) (MGCv1) compared to the per sample sequencing effort of the CBAJ-DB. **B) Assembly and binning workflow.** Total and medium and high quality (MQHQ) bins derived from Megahit assembler (left) and IDBA_UD assembler (right) shown, blue arrows indicate read mapping for subtractive assembly.



value_type	p.value	quartile	group
Completeness	0.88904906	Q1	treatment
Contamination	0.91851228	Q1	treatment
Completeness	0.98585807	median	treatment
Contamination	0.63520543	median	treatment
Completeness	0.98431568	Q3	treatment
Contamination	0.80027867	Q3	treatment
Completeness	0.08846319	Q1	assembly_type
Contamination	0.99812567	Q1	assembly_type
Completeness	0.49425519	median	assembly_type
Contamination	0.84779705	median	assembly_type
Completeness	0.92290676	Q3	assembly_type
Contamination	0.43825507	Q3	assembly_type

Fig. 10. Contamination and completion statistics and the most resolved taxonomy groups for MAGs containing amplicon sequencing variants (ASVs). A) Distribution of contamination and completion of medium and high quality MAGs by treatment and assembly method. Table shows Chi Squared test p-values of quartiles between groups. B) MAG groups with matching ASVs from 16S rRNA sequencing as determined by Mmseqs2 and Barrnap (see methods). Colored text indicates the lowest resolved taxonomy. Mean ASV relative abundance within each taxonomy group in each treatment is represented in Log relative abundance (right).

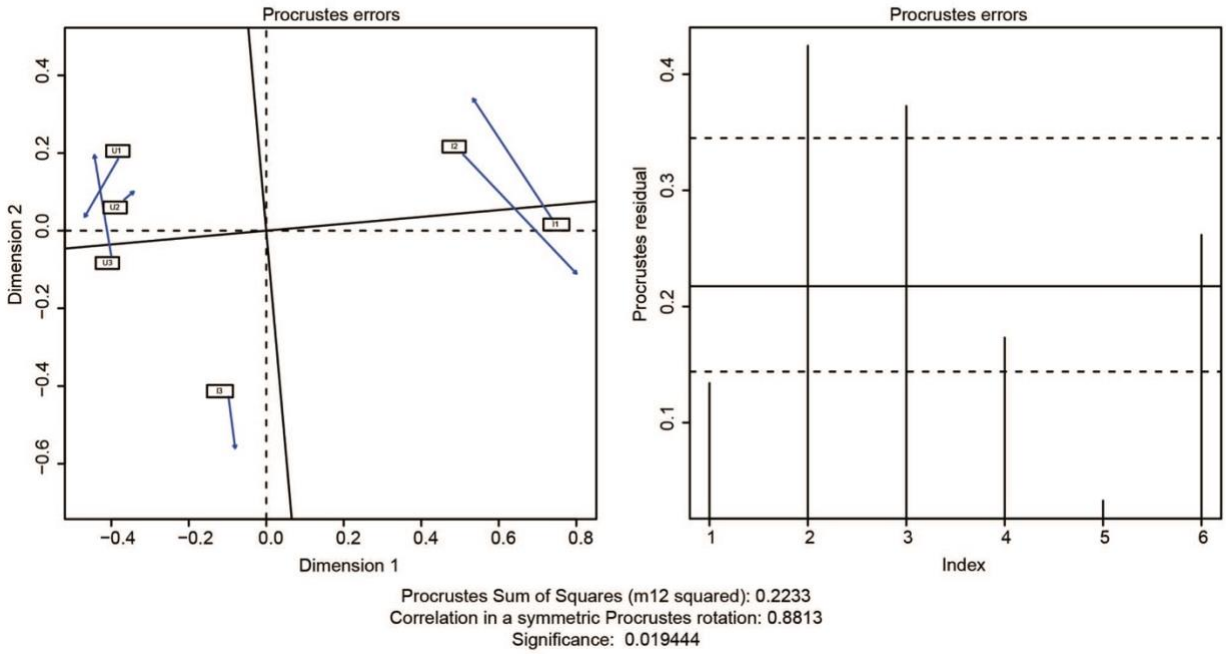


Fig. 11. Procrustes analysis of dereplicated medium and high quality (dMQHQ) metagenome assembled genomes (MAGs) and viral metagenome assembled genomes (vMAGs). dMQHQ and vMAG relative abundance non-metric multidimensional scaling (NMDS) and viral genome-database NMDS showing high significant similarity of ordinations.

Chapter 2 References

1. Stanaway JD, Parisi A, Sarkar K, Blacker BF, Reiner RC, Hay SI, Nixon MR, Dolecek C, James SL, Mokdad AH, et al. The global burden of non-typhoidal salmonella invasive disease: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Infect Dis 2019; 19:1312–24.

2. Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, Jones TF, Fazil A, Hoekstra RM, for the International Collaboration on Enteric Disease "Burden of Illness" Studies. The Global Burden of Nontyphoidal Salmonella Gastroenteritis. Clin Infect Dis 2010; 50:882–9.

3. Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson M-A, Roy SL, Jones JL, Griffin PM. Foodborne Illness Acquired in the United States—Major Pathogens. Emerg Infect Dis 2011; 17:7–15.

4. Stecher B, Robbiani R, Walker AW, Westendorf AM, Barthel M, Kremer M, Chaffron S, Macpherson AJ, Buer J, Parkhill J, et al. Salmonella enterica Serovar Typhimurium Exploits Inflammation to Compete with the Intestinal Microbiota. PLOS Biol 2007; 5:e244.

5. Barman M, Unold D, Shifley K, Amir E, Hung K, Bos N, Salzman N. Enteric Salmonellosis Disrupts the Microbial Ecology of the Murine Gastrointestinal Tract. Infect Immun 2008; 76:907–15.

6. Rogers AWL, Tsolis RM, Bäumler AJ. Salmonella versus the Microbiome. Microbiol Mol Biol Rev 2020; 85:e00027-19.

7. Stecher B. Establishing causality in Salmonella-microbiota-host interaction: The use of gnotobiotic mouse models and synthetic microbial communities. Int J Med Microbiol 2021; 311:151484.

8. Vos WM de, Tilg H, Hul MV, Cani PD. Gut microbiome and health: mechanistic insights. Gut 2022; 71:1020–32.

9. Winter SE, Thiennimitr P, Winter MG, Butler BP, Huseby DL, Crawford RW, Russell JM, Bevins CL, Adams LG, Tsolis RM, et al. Gut inflammation provides a respiratory electron acceptor for Salmonella. Nature 2010; 467:426–9.

10. Shelton CD, Yoo W, Shealy NG, Torres TP, Zieba JK, Calcutt MW, Foegeding NJ, Kim D, Kim J, Ryu S, et al. Salmonella Typhimurium uses anaerobic respiration to overcome propionate-mediated colonization resistance. bioRxiv 2021; :2021.05.25.445690.

11. Rivera-Chávez F, Bäumler AJ. The Pyromaniac Inside You: Salmonella Metabolism in the Host Gut. Annu Rev Microbiol 2015; 69:31–48.

12. Walker GT, Raffatellu M. Salmonella respiration turns the tables on propionate. Trends Microbiol 2022; 30:206–8.

13. Rivera-Chávez F, Zhang LF, Faber F, Lopez CA, Byndloss MX, Olsan EE, Xu G, Velazquez EM, Lebrilla CB, Winter SE, et al. Depletion of butyrate-producing Clostridia from the gut microbiota drives an aerobic luminal expansion of Salmonella. Cell Host Microbe 2016; 19:443–54.

14. Sekirov I, Tam NM, Jogova M, Robertson ML, Li Y, Lupp C, Finlay BB. Antibiotic-Induced Perturbations of the Intestinal Microbiota Alter Host Susceptibility to Enteric Infection. Infect Immun 2008; 76:4726–36.

15. Woo H, Okamoto S, Guiney D, Gunn JS, Fierer J. A Model of Salmonella Colitis with Features of Diarrhea in SLC11A1 Wild-Type Mice. PLOS ONE 2008; 3:e1603.

16. Ferreira RBR, Gill N, Willing BP, Antunes LCM, Russell SL, Croxen MA, Finlay BB. The Intestinal Microbiota Plays a Role in Salmonella-Induced Colitis Independent of Pathogen Colonization. PLoS ONE 2011; 6:e20338.

17. Ahmer BM, Gunn JS. Interaction of Salmonella spp. with the Intestinal Microbiota. Front Microbiol [Internet] 2011 [cited 2020 Apr 16]; 2. Available from: https://www.frontiersin.org/articles/10.3389/fmicb.2011.00101/full

18. Karlinsey JE, Maguire ME, Becker LA, Crouch M-LV, Fang FC. The Phage Shock Protein PspA Facilitates Divalent Metal Transport and is Required for Virulence of Salmonella enterica sv. Typhimurium. Mol Microbiol 2010; 78:669–85.

19. Shelton CD, Yoo W, Shealy NG, Torres TP, Zieba JK, Calcutt MW, Foegeding NJ, Kim D, Kim J, Ryu S, et al. Salmonella enterica serovar Typhimurium uses anaerobic respiration to overcome propionate-mediated colonization resistance. Cell Rep [Internet] 2022 [cited 2022 Jul 14]; 38. Available from: https://www.cell.com/cell-reports/abstract/S2211-1247(21)01680-6

20. Spiga L, Winter MG, de Carvalho TF, Zhu W, Hughes ER, Gillis CC, Behrendt CL, Kim J, Chessa D, Andrews-Polymenis HL, et al. An oxidative central metabolism enables Salmonella to utilize microbiota-derived succinate. Cell Host Microbe 2017; 22:291-301.e6.

21. Lesker TR, Durairaj AC, Gálvez EJC, Lagkouvardos I, Baines JF, Clavel T, Sczyrba A, McHardy AC, Strowig T. An Integrated Metagenome Catalog Reveals New Insights into the Murine Gut Microbiome. Cell Rep 2020; 30:2909-2922.e6.

22. Beresford-Jones BS, Forster SC, Stares MD, Notley G, Viciani E, Browne HP, Boehmler DJ, Soderholm AT, Kumar N, Vervier K, et al. The Mouse Gastrointestinal Bacteria Catalogue enables translation between the mouse and human gut microbiotas via functional mapping. Cell Host Microbe 2022; 30:124-138.e8.

23. Wong EO-Y, Brownlie EJE, Ng KM, Kathirgamanathan S, Yu FB, Merrill BD, Huang KC, Martin A, Tropini C, Navarre WW. The CIAMIB: a Large and Metabolically Diverse Collection of Inflammation-Associated Bacteria from the Murine Gut. mBio 2022; 13:e02949-21.

24. Gregory AC, Zablocki O, Zayed AA, Howell A, Bolduc B, Sullivan MB. The Gut Virome Database Reveals Age-Dependent Patterns of Virome Diversity in the Human Gut. Cell Host Microbe 2020; 28:724-740.e8.

25. Adiliaghdam F, Jeffrey KL. Illuminating the human virome in health and disease. Genome Med 2020; 12:66.

26. Cao Z, Sugimura N, Burgermeister E, Ebert MP, Zuo T, Lan P. The gut virome: A new microbiome component in health and disease. eBioMedicine [Internet] 2022 [cited 2022 Jul 11];

81. Available from: https://www.thelancet.com/journals/ebiom/article/PIIS2352-3964(22)00294-8/fulltext

27. Borton MA, Sabag-Daigle A, Wu J, Solden LM, O'Banion BS, Daly RA, Wolfe RA, Gonzalez JF, Wysocki VH, Ahmer BMM, et al. Chemical and pathogen-induced inflammation disrupt the murine intestinal microbiome. Microbiome [Internet] 2017 [cited 2019 Aug 26]; 5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5408407/

28. O'Loughlin JL, Samuelson DR, Braundmeier-Fleming AG, White BA, Haldorson GJ, Stone JB, Lessmann JJ, Eucker TP, Konkel ME. The Intestinal Microbiota Influences Campylobacter jejuni Colonization and Extraintestinal Dissemination in Mice. Appl Environ Microbiol 2015; 81:4642–50.

29. Xiao L, Feng Q, Liang S, Sonne SB, Xia Z, Qiu X, Li X, Long H, Zhang J, Zhang D, et al. A catalog of the mouse gut metagenome. Nat Biotechnol 2015; 33:1103–8.

30. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes [Internet]. PeerJ Inc.; 2015 [cited 2019 Dec 3]. Available from: https://peerj.com/preprints/554

31. Chklovski A, Parks DH, Woodcroft BJ, Tyson GW. CheckM2: a rapid, scalable and accurate tool for assessing microbial genome quality using machine learning [Internet]. 2022 [cited 2023 Feb 5]; :2022.07.11.499243. Available from: https://www.biorxiv.org/content/10.1101/2022.07.11.499243v1

32. Orakov A, Fullam A, Coelho LP, Khedkar S, Szklarczyk D, Mende DR, Schmidt TSB, Bork P. GUNC: detection of chimerism and contamination in prokaryotic genomes. Genome Biol 2021; 22:178.

33. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloe-Fadrosh EA, et al. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat Biotechnol 2017; 35:725–31.

34. Shaffer M, Borton MA, McGivern BB, Zayed AA, La Rosa SL, Solden LM, Liu P, Narrowe AB, Rodríguez-Ramos J, Bolduc B, et al. DRAM for distilling microbial metabolism to automate the curation of microbiome function. Nucleic Acids Res 2020; 48:8883–900.

35. Mise K, Iwasaki W. Unexpected absence of ribosomal protein genes from metagenome-assembled genomes. ISME Commun 2022; 2:1–9.

36. Microbial Genome-Resolved Metaproteomic Analyses Frame Intertwined Carbon and Nitrogen Cycles in River Hyporheic Sediments [Internet]. 2021 [cited 2023 Feb 27]; Available from: https://www.researchsquare.com

37. McGivern BB, Tfaily MM, Borton MA, Kosina SM, Daly RA, Nicora CD, Purvine SO, Wong AR, Lipton MS, Hoyt DW, et al. Decrypting bacterial polyphenol metabolism in an anoxic wetland soil. Nat Commun 2021; 12:2466.

38. Solden LM, Naas AE, Roux S, Daly RA, Collins WB, Nicora CD, Purvine SO, Hoyt DW, Schückel J, Jørgensen B, et al. Interspecies cross-feeding orchestrates carbon degradation in the rumen ecosystem. Nat Microbiol 2018; 3:1274–84.

39. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The Human Microbiome Project. Nature 2007; 449:804–10.

40. Borton MA, Shaffer M, Hoyt DW, Jiang R, Ellenbogen J, Purvine S, Nicora CD, Eder EK, Wong AR, Smulian AG, et al. Targeted curation of the gut microbial gene content modulating human cardiovascular disease [Internet]. 2022 [cited 2022 Aug 12]; :2022.06.20.496735. Available from: https://www.biorxiv.org/content/10.1101/2022.06.20.496735v1

41. Nguyen TLA, Vieira-Silva S, Liston A, Raes J. How informative is the mouse for human gut microbiota research? Dis Model Mech 2015; 8:1–16.

42. Park JC, Im S-H. Of men in mice: the development and application of a humanized gnotobiotic mouse model for microbiome therapeutics. Exp Mol Med 2020; 52:1383–96.

43. Lloyd-Price J, Arze C, Ananthakrishnan AN, Schirmer M, Avila-Pacheco J, Poon TW, Andrews E, Ajami NJ, Bonham KS, Brislawn CJ, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. Nature 2019; 569:655–62.

44. Ghaffari S, Abbasi A, Somi MH, Moaddab SY, Nikniaz L, Kafil HS, Ebrahimzadeh Leylabadlo H. Akkermansia muciniphila: from its critical role in human health to strategies for promoting its abundance in human gut microbiome. Crit Rev Food Sci Nutr 2022; 0:1–21.

45. Choi S-H, Lee S-O, Kim TH, Chung J-W, Choo EJ, Kwak YG, Kim M-N, Kim YS, Woo JH, Ryu J, et al. Clinical Features and Outcomes of Bacteremia Caused by Enterococcus casseliflavus and Enterococcus gallinarum: Analysis of 56 Cases. Clin Infect Dis 2004; 38:53–61.

46. Almeida A, Nayfach S, Boland M, Strozzi F, Beracochea M, Shi ZJ, Pollard KS, Sakharova E, Parks DH, Hugenholtz P, et al. A unified catalog of 204,938 reference genomes from the human gut microbiome. Nat Biotechnol 2021; 39:105–14.

47. Faber F, Thiennimitr P, Spiga L, Byndloss MX, Litvak Y, Lawhon S, Andrews-Polymenis HL, Winter SE, Bäumler AJ. Respiration of Microbiota-Derived 1,2-propanediol Drives Salmonella Expansion during Colitis. PLOS Pathog 2017; 13:e1006129.

48. Thiennimitr P, Winter SE, Winter MG, Xavier MN, Tolstikov V, Huseby DL, Sterzenbach T, Tsolis RM, Roth JR, Bäumler AJ. Intestinal inflammation allows Salmonella to use ethanolamine to compete with the microbiota. Proc Natl Acad Sci 2011; 108:17480–5.

49. Hyland NP, Cavanaugh CR, Hornby PJ. Emerging effects of tryptophan pathway metabolites and intestinal microbiota on metabolism and intestinal function. Amino Acids 2022; 54:57–70.

50. Agus A, Planchais J, Sokol H. Gut Microbiota Regulation of Tryptophan Metabolism in Health and Disease. Cell Host Microbe 2018; 23:716–24.

51. Cussotto S, Delgado I, Anesi A, Dexpert S, Aubert A, Beau C, Forestier D, Ledaguenel P, Magne E, Mattivi F, et al. Tryptophan Metabolic Pathways Are Altered in Obesity and Are Associated

With Systemic Inflammation. Front Immunol [Internet] 2020 [cited 2022 Jul 5]; 11. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2020.00557

52. Du J, Zayed AA, Kigerl KA, Zane K, Sullivan MB, Popovich PG. Spinal Cord Injury Changes the Structure and Functional Potential of Gut Bacterial and Viral Communities. mSystems 2021; 6:e01356-20.

53. Wang H, Ling Y, Shan T, Yang S, Xu H, Deng X, Delwart E, Zhang W. Gut virome of mammals and birds reveals high genetic diversity of the family Microviridae. Virus Evol 2019; 5:vez013.

54. Yang Y, Nguyen M, Khetrapal V, Sonnert ND, Martin AL, Chen H, Kriegel MA, Palm NW. Within-host evolution of a gut pathobiont facilitates liver translocation. Nature 2022; :1–8.

55. Hugenholtz F, de Vos WM. Mouse models for human intestinal microbiota research: a critical evaluation. Cell Mol Life Sci 2018; 75:149–60.

56. Ursell LK, Metcalf JL, Parfrey LW, Knight R. Defining the Human Microbiome. Nutr Rev 2012; 70:S38–44.

57. Kieser S, Zdobnov EM, Trajkovski M. Comprehensive mouse microbiota genome catalog reveals major difference to its human counterpart. PLOS Comput Biol 2022; 18:e1009947.

58. Velazquez EM, Nguyen H, Heasley KT, Saechao CH, Gil LM, Rogers AWL, Miller BM, Rolston MR, Lopez CA, Litvak Y, et al. Endogenous Enterobacteriaceae underlie variation in susceptibility to Salmonella infection. Nat Microbiol 2019; 4:1057–64.

59. Nayfach S, Páez-Espino D, Call L, Low SJ, Sberro H, Ivanova NN, Proal AD, Fischbach MA, Bhatt AS, Hugenholtz P, et al. Metagenomic compendium of 189,680 DNA viruses from the human gut microbiome. Nat Microbiol 2021; 6:960–70.

60. Mousa WK, Chehadeh F, Husband S. Recent Advances in Understanding the Structure and Function of the Human Microbiome. Front Microbiol 2022; 13:825338.

61. Yadav M, Verma MK, Chauhan NS. A review of metabolic potential of human gut microbiome in human nutrition. Arch Microbiol 2018; 200:203–17.

62. Berg G, Rybakova D, Fischer D, Cernava T, Vergès M-CC, Charles T, Chen X, Cocolin L, Eversole K, Corral GH, et al. Microbiome definition re-visited: old concepts and new challenges. Microbiome 2020; 8:103.

63. Zheng D, Liwinski T, Elinav E. Interaction between microbiota and immunity in health and disease. Cell Res 2020; 30:492–506.

64. Belkaid Y, Hand TW. Role of the Microbiota in Immunity and Inflammation. Cell 2014; 157:121–41.

65. Zeng MY, Inohara N, Nuñez G. Mechanisms of inflammation-driven bacterial dysbiosis in the gut. Mucosal Immunol 2017; 10:18–26.

66. Ang Z, Ding JL. GPR41 and GPR43 in Obesity and Inflammation – Protective or Causative? Front Immunol [Internet] 2016 [cited 2022 Jul 11]; 7. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2016.00028

67. Louis P, Flint HJ. Formation of propionate and butyrate by the human colonic microbiota. Environ Microbiol 2017; 19:29–41.

68. Parada Venegas D, De la Fuente MK, Landskron G, González MJ, Quera R, Dijkstra G, Harmsen HJM, Faber KN, Hermoso MA. Short Chain Fatty Acids (SCFAs)-Mediated Gut Epithelial and Immune Regulation and Its Relevance for Inflammatory Bowel Diseases. Front Immunol [Internet] 2019 [cited 2022 Jul 11]; 10. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2019.00277

69. Louis P, Flint HJ. Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. FEMS Microbiol Lett 2009; 294:1–8.

70. Kim Y-G, Sakamoto K, Seo S-U, Pickard JM, Gillilland MG, Pudlo NA, Hoostal M, Li X, Wang TD, Feehley T, et al. Neonatal acquisition of Clostridia species protects against colonization by bacterial pathogens. Science 2017; 356:315–9.

71. Sinha SR, Haileselassie Y, Nguyen LP, Tropini C, Wang M, Becker LS, Sim D, Jarr K, Spear ET, Singh G, et al. Dysbiosis-induced Secondary Bile Acid Deficiency Promotes Intestinal Inflammation. Cell Host Microbe 2020; 27:659-670.e5.

72. Guzior DV, Quinn RA. Review: microbial transformations of human bile acids. Microbiome 2021; 9:140.

73. Qin X, Yang M, Cai H, Liu Y, Gorris L, Aslam MZ, Jia K, Sun T, Wang X, Dong Q. Antibiotic Resistance of Salmonella Typhimurium Monophasic Variant 1,4,[5],12:i:-in China: A Systematic Review and Meta-Analysis. Antibiotics 2022; 11:532.

74. Sabag-Daigle A, Blunk HM, Gonzalez JF, Steidley BL, Boyaka PN, Ahmer BMM. Use of Attenuated but Metabolically Competent Salmonella as a Probiotic To Prevent or Treat Salmonella Infection. Infect Immun 2016; 84:2131–40.

75. Ouyang J, Lin J, Isnard S, Fombuena B, Peng X, Marette A, Routy B, Messaoudene M, Chen Y, Routy J-P. The Bacterium Akkermansia muciniphila: A Sentinel for Gut Permeability and Its Relevance to HIV-Related Inflammation. Front Immunol 2020; 11:645.

76. Zhang T, Li Q, Cheng L, Buch H, Zhang F. Akkermansia muciniphila is a promising probiotic. Microb Biotechnol 2019; 12:1109–25.

77. Ganesh BP, Klopfleisch R, Loh G, Blaut M. Commensal Akkermansia muciniphila Exacerbates Gut Inflammation in Salmonella Typhimurium-Infected Gnotobiotic Mice. PLoS ONE 2013; 8:e74963.

78. Bansal T, Alaniz RC, Wood TK, Jayaraman A. The bacterial signal indole increases epithelial-cell tight-junction resistance and attenuates indicators of inflammation. Proc Natl Acad Sci U S A 2010; 107:228–33.

79. Rautio M, Eerola E, Väisänen-Tunkelrott M-L, Molitoris D, Lawson P, Collins MD, Jousimies-Somer H. Reclassification of Bacteroides putredinis (Weinberg et al., 1937) in a New Genus Alistipes gen. nov., as Alistipes putredinis comb. nov., and Description of Alistipes finegoldii sp. nov., from Human Sources. Syst Appl Microbiol 2003; 26:182–8.

80. Parker BJ, Wearsch PA, Veloo ACM, Rodriguez-Palacios A. The Genus Alistipes: Gut Bacteria With Emerging Implications to Inflammation, Cancer, and Mental Health. Front Immunol 2020; 11:906.

81. Chng KR, Ghosh TS, Tan YH, Nandi T, Lee IR, Ng AHQ, Li C, Ravikrishnan A, Lim KM, Lye D, et al. Metagenome-wide association analysis identifies microbial determinants of post-antibiotic ecological recovery in the gut. Nat Ecol Evol 2020; 4:1256–67.

82. Pickard JM, Chervonsky AV. Intestinal fucose as a mediator of host-microbe symbiosis. J Immunol Baltim Md 1950 2015; 194:5588–93.

83. Mistou M-Y, Sutcliffe IC, van Sorge NM. Bacterial glycobiology: rhamnose-containing cell wall polysaccharides in Gram-positive bacteria. FEMS Microbiol Rev 2016; 40:464–79.

84. Singh P, Teal TK, Marsh TL, Tiedje JM, Mosci R, Jernigan K, Zell A, Newton DW, Salimnia H, Lephart P, et al. Intestinal microbial communities associated with acute enteric infections and disease recovery. Microbiome 2015; 3:45.

85. De Keersmaecker SCJ, Verhoeven TLA, Desair J, Marchal K, Vanderleyden J, Nagy I. Strong antimicrobial activity of Lactobacillus rhamnosus GG against Salmonella typhimurium is due to accumulation of lactic acid. FEMS Microbiol Lett 2006; 259:89–96.

86. Jiang H, Li P, Gu Q. Heterologous expression and purification of plantaricin NC8, a two-peptide bacteriocin against Salmonella spp. from Lactobacillus plantarum ZJ316. Protein Expr Purif 2016; 127:28–34.

87. Perumal V, Venkatesan A. Antimicrobial, cytotoxic effect and purification of bacteriocin from vancomycin susceptible Enterococcus faecalis and its safety evaluation for probiotization. LWT 2017; 78:303–10.

88. Fijan S. Microorganisms with Claimed Probiotic Properties: An Overview of Recent Literature. Int J Environ Res Public Health 2014; 11:4745–67.

89. Hanchi H, Mottawea W, Sebei K, Hammami R. The Genus Enterococcus: Between Probiotic Potential and Safety Concerns—An Update. Front Microbiol [Internet] 2018 [cited 2022 Aug 15]; 9. Available from: https://www.frontiersin.org/articles/10.3389/fmicb.2018.01791

90. Antoun M, Hattab Y, Akhrass F-A, Hamilton LD. Uncommon Pathogen, Lactobacillus, Causing Infective Endocarditis: Case Report and Review. Case Rep Infect Dis 2020; 2020:e8833948.

91. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci 2011; 108:4516–22.

92. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 2019; 37:852–7.

93. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High resolution sample inference from Illumina amplicon data. Nat Methods 2016; 13:581–3.

94. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 2013; 41:D590–6.

95. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data [Internet]. [cited 2021 May 10]; Available from: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

96. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 2015; 31:1674–6.

97. Peng Y, Leung HCM, Yiu SM, Chin FYL. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics 2012; 28:1420–8.

98. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. Bioinformatics 2009; 25:2078–9.

99. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ [Internet] 2019 [cited 2019 Nov 18]; 7. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6662567/

100. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics 2019; :btz848.

101. Olm MR, Brown CT, Brooks B, Banfield JF. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. ISME J 2017; 11:2864–8.

102. Köster J, Rahmann S. Snakemake—a scalable bioinformatics workflow engine. Bioinformatics 2012; 28:2520–2.

103. Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat Biotechnol 2017; 35:1026–8.

104. Seemann T. Barrnap [Internet]. 2022 [cited 2022 Aug 16]; Available from: https://github.com/tseemann/barrnap

105. Smid M, Coebergh van den Braak RRJ, van de Werken HJG, van Riet J, van Galen A, de Weerd V, van der Vlugt-Daane M, Bril SI, Lalmahomed ZS, Kloosterman WP, et al. Gene length corrected trimmed mean of M-values (GeTMM) processing of RNA-seq data performs similarly in

intersample analyses while improving intrasample comparisons. BMC Bioinformatics 2018; 19:236.

106. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010; 26:139–40.

107. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012; 9:357–9.

108. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 2018; 9:5114.

109. Barco RA, Garrity GM, Scott JJ, Amend JP, Nealson KH, Emerson D. A Genus Definition for Bacteria and Archaea Based on a Standard Genome Relatedness Index. mBio 2020; 11:e02475-19.

110. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res 2012; 40:W445–51.

111. Eddy SR. Accelerated Profile HMM Searches. PLOS Comput Biol 2011; 7:e1002195.

112. Guo J, Bolduc B, Zayed AA, Varsani A, Dominguez-Huerta G, Delmont TO, Pratama AA, Gazitúa MC, Vik D, Sullivan MB, et al. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. Microbiome 2021; 9:37.

113. Bolduc B, Roux S. Clustering Viral Genomes in iVirus [Internet]. protocols.io2017 [cited 2022 Aug 15]; Available from: https://www.protocols.io/view/clustering-viral-genomes-in-ivirus-gwebxbe

114. Nayfach S, Camargo AP, Schulz F, Eloe-Fadrosh E, Roux S, Kyrpides NC. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. Nat Biotechnol 2020; :1–8.

115. Bin Jang H, Bolduc B, Zablocki O, Kuhn JH, Roux S, Adriaenssens EM, Brister JR, Kropinski AM, Krupovic M, Lavigne R, et al. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. Nat Biotechnol 2019; 37:632–9.

116. Merchant N, Lyons E, Goff S, Vaughn M, Ware D, Micklos D, Antin P. The iPlant Collaborative: Cyberinfrastructure for Enabling Data to Discovery for the Life Sciences. PLOS Biol 2016; 14:e1002342.

117. Skennerton CT, Imelfort M, Tyson GW. Crass: identification and reconstruction of CRISPR from unassembled metagenomic data. Nucleic Acids Res 2013; 41:e105.

118. Ahlgren NA, Ren J, Lu YY, Fuhrman JA, Sun F. Alignment-free \documentclass[12pt]{minimal} \usepackage{amsmath} \usepackage{wasysym} \usepackage{amsfonts} \usepackage{amssymb} \usepackage{amsbsy} \usepackage{upgreek} \usepackage{mathrsfs} \setlength{\oddsidemargin}{-69pt} \begin{document} }{}d_2^*\end{document} oligonucleotide frequency dissimilarity measure improves prediction of hosts from metagenomically-derived viral sequences. Nucleic Acids Res 2017; 45:39–53.

119. Dixon P. VEGAN, a package of R functions for community ecology. J Veg Sci 2003; 14:927–30.

120. Create Elegant Data Visualisations Using the Grammar of Graphics [Internet]. [cited 2022 Aug 12]; Available from: https://ggplot2.tidyverse.org/

121. Anderson MJ. Analysis of Ecological Communities: Bruce McCune and James B. Grace, MjM Software Design, Gleneden Beach, USA, 2002, ISBN 0 9721290 0 6, US$ 35 (Pbk). J Exp Mar Biol Ecol 2003; 289:303–5.

122. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. Metagenomic biomarker discovery and explanation. Genome Biol 2011; 12:R60.

Chapter 3 – Multi-omics Illuminates the *Salmonella*-included Microbiome: Discovering Roles of Persistent Bacteria in the Inflamed Gut

3.1 Summary

Salmonella-induced inflammation results in oxidation of the lumen. Widely regarded as critical to successful *Salmonella* enteric infection, host inflammation makes available terminal electron sinks potentiating *Salmonella* anaerobic respiration. Energetically favorable respiratory metabolism of both host derived and microbially derived substrate allows *Salmonella* to outcompete obligate fermenting bacteria and shapes the microbial community structure following infection.

Here we perform genome-resolved metatranscriptomic and metabolomic evaluation of the *Salmonella*-altered microbiome in CBA/J mice to determine the resident microbial community role in *Salmonella* pathogenesis, importantly addressing a lack of *in vivo* studies with complete microbiomes. Machine learning modeling of metagenomically derived bacteria genomes shows functional potential and gene content of commensal bacteria can predict *Salmonella*-association in the inflamed gut. Transcript and metabolite data indicate a potentially critical role of host derived organic sulfur to *Salmonella* tetrathionate respiration and reveals expression from persistent taxa to increase sulfur availability in support of tetrathionate resultant from host inflammation.

Furthermore, our multi-omics investigation of *Salmonella* infection in the context of a robust and complete microbiome belies current dogma of *Salmonella* lactate utilization. We present evidence of potential *Salmonella*/commensal cross feeding and show persistence of certain bacteria during infection, including *Akkermansia*, irrespective of host diet or genetics. The comprehensive evaluation of the *Salmonella*-included microbiome presented here offers multiple bacteria for further investigation as probiotics among the *Bacilli* class, including *Lactobacillus* and *Enterococcus*, that are either enriched

during *Salmonella* infection or which persist in the community and express genes indicating potential substrate competition with *Salmonella*.

3.2 Introduction

Non-typhoidal *Salmonella*, like *Salmonella enterica* serovar *Typhimurium*, causes more than 1.3 million infections and is responsible for 420 deaths annually in the United States[1,2]. The global burden of non-typhoidal *Salmonella* is magnitudes larger and the World Health Organization reports 550 million infections globally per annum, nearly half (220 million) occurring in children[3]. These infections are on the rise in recent decades[1] due to multidrug resistant strains with one report indicating 77% of the *Salmonella* genomes sequenced between 2015 and 2018 having genetic markers for antibiotic resistance[1]. *Salmonella* is often transmitted by contaminated food products and is harbored in zoonotic reservoirs of common food animals like chickens and pigs[4]. To better understand *Salmonella* impacts on the gastrointestinal (GI) tract, our team[5,6] and others[7,8] use *Salmonella* resistant murine models like the CBA/J mouse which are *Slc11A1* competent, a trait that effectively prevents *Salmonella* systemic infection[9]. Infection in these mice mimics *Salmonella enterica* serovar *Typhimurium* GI infection in humans where a strong GI inflammatory response occurs, which remodels the gut ecosystem[5,6,10,11]. Like humans, *Salmonella* resistant mice have variable infection rates[5,6], requiring the induction of inflammation for reproducible *Salmonella* GI infection.

Salmonella enterica serovar *Typhimurium* inflammation-adapted metabolic range allows it to proliferate using products abundant in the inflamed gut like succinate, lactate, 1,2 propanediol, and ethanolamine[8,12–14]. Reactive oxygen species (ROS) and reactive nitrogen species produced from oxidative bursts and neutrophil degranulation make terminal electron acceptors nitrate and tetrathionate available for *Salmonella* anaerobic respiration during infection[7,15,16]. Hydrogen sulfide produced by sulfur reducing bacteria (SRB) is detoxified in colonic epithelia to sulfate and thiosulfate that is released into the lumen providing substrate for inflammatory oxidation to tetrathionate[11,17]. Despite hydrogen sulfide

toxicity and tetrathionate mediated *Salmonella* anaerobic respiration, sulfur is an important component of a homeostatic gut microbiome[18]. Members of *Desulfovibrio* are the predominant SRB in the human gut, able to use fermentation products like molecular hydrogen and lactate or acetate to reduce sulfate[18,19]. Lactic acid produced by bacteria including *Lactobacillus* spp. can be used by SRBs as electron donors for sulfur reduction; and a balance of SRB and lactic acid bacteria (LAB) as well as their associated metabolites is important for gut homeostasis[18]. Conversely, an accumulation of hydrogen sulfide has been associated with chronic inflammatory conditions like ulcerative colitis and inflammatory bowel disease while an overabundance of lactate affects pH homeostasis and is associated with conditions like lactic acidosis[18,20]. Notably, *Salmonella* impacts on the microbiome include increases in host-derived lactate and sulfur oxidation by ROS, two processes that can shape the microbiota affecting SRB and LAB levels in the community[13,21]. Groundbreaking papers displaying *Salmonella* utilization of host-derived lactate and tetrathionate in the inflamed gut neglect any potential for broader community contributions to these processes beyond a presence/absence paradigm acknowledging a reduction in butyrate producing bacteria[16,22]. More comprehensive interrogation of the complete inflamed microbiome is necessary to determine if any specific and persistent commensal bacteria are involved in *Salmonella* lactate and tetrathionate metabolism and if they affect key substrate availability for these processes.

The effect of *Salmonella* GI colonization on the resident microbiota is commonly characterized by a reduction of *Clostridia* spp. and lower overall richness of the community[23,24]. These community impacts are attributed to successful *Salmonella* competition dependent on inflammation[9]. Most bacteria in the healthy gut are classified as obligate anaerobes and they gain energy from fermenting fiber and sugars from the host diet into short chain fatty acids (SCFAs) that in turn nourish the intestinal epithelium[25]. Members of the microbiota play an important role in colonization resistance to enteric pathogens like *Salmonella*[26]. For instance, prominent microbiota members like *Akkermansia muciniphila* can increase gut barrier integrity by stimulating host mucin production[27]. Fermentation products like propionate can increase colonization resistance to *Salmonella*, and many LAB like *Lactobacillus* and *Enterococcus*

produce bacteriocins that can suppress *Salmonella* growth[28–31]. Furthermore, bacteria expressing bile salt hydrolase can deconjugate bile acids from their conjugate amino acids (glycine and taurine), making them more cytotoxic[32]. Deconjugated bile acids are more potent inhibitors of *Salmonella* pathogenicity island-1 (SPI-1) transcription than their conjugated counterparts[33–35]. Other common metabolites in the gut, either from the diet or microbial products, can affect SPI-1 translation and effector potential including the flavonoid quercetin, tryptophan and indole derivatives, and the SCFA propionate[36,37]. SPI-1 and *Salmonella* pathogenicity island-2 are critical for *Salmonella* successful infection and their products are important triggers of host inflammation and thus colonization of the gut[9,36]. Though microbiota mediated colonization resistance to enteric pathogens is well documented, there is comparatively limited work on the post-colonization community interaction with *Salmonella*, after peak infection and inflammation. Additionally, many studies evaluating the *Salmonella*-included metabolome do so in either germfree models or mice previously treated with antibiotics or with only a limited infection period (as little as 4 days)[7,12,16,22]. To understand the community dynamics during a *Salmonella* GI infection, microbiome chemistry and individual microbiota activity need to be interrogated in studies that are ultimately more translatable to human disease.

Here we use a comprehensive meta-omics approach to evaluate the *Salmonella*-included microbiome up to 15 days post-infection. Our holistic evaluation of key metabolites known to impact *Salmonella* colonization and infection provides a necessary link between mechanistic study of *Salmonella* metabolism and pathogenicity and therapeutic use of probiotics or biological control or prevention of *Salmonella* enteritis that may be vital to the global effort against antibiotic resistance. Both diet and mouse strain are considered during our community analysis, identifying microbes prevalent during *Salmonella* infection despite either variable. We identify *Salmonella*-associated bacteria and important mechanisms for survival in the *Salmonella*-included gut using machine learning and metagenomics, and we leverage metatranscriptomics and metabolomics to both confirm and confront current dogma surrounding *Salmonella* lactate utilization, sulfur reduction, and the gut microbiota role during infection.

Specifically, individual bacteria interactions with host metabolites are identified as potential crucial links between the inflamed host-provided substrates and *Salmonella* ability to use them.

3.3 Results

3.3.1 Salmonella colonization alters microbiota community structure concordant with pathogen burden and infection duration

To holistically evaluate *Salmonella enterica* serovar Typhimurium strain 14028 (*Salmonella*) effect on the murine gut microbiome we introduced 10^9 colony forming units of *Salmonella* to CBA/J mice (n=44) and an uninfected cohort (n=23) were left without *Salmonella.* Mice from both infection treatment groups were either fed chow (n=42) or a high fat diet (HFD, n=25) to interrogate the impact of diet on a *Salmonella*-included gut community. Total mice in each treatment cohort are as follows: 24 (chow + infected), 18 (chow + uninfected), 20 (HFD + infected), 5 (HFD + uninfected). Fecal samples from all mice were collected and analyzed up to 15 days post infection to determine the impacts of *Salmonella* colonization on the gut microbiome as infection progressed (Fig. 12a, 12b). Fecal samples were also obtained up to 6 days (day -6, -3, -2, -1) before infection (day 0) from mice (n=65) in both uninfected and infected groups (Fig. 12a, 12b). Amplicon sequencing was performed on all sampled feces (n=712) to establish the level of *Salmonella* colonization and the relative abundance of community members in each treatment cohort. These sequencing efforts resulted in 35,749,706 paired end 16S rRNA reads and a final filtered feature table with 6,308 unique ASVs (Fig. 17a).

Median *Salmonella* relative abundance in high-responder mice (mice with 25% *Salmonella* relative abundance in at least one timepoint) steadily rose following *Salmonella* introduction and peaked at day 8 (94.4% *Salmonella* relative abundance, Fig. 12a). We performed non-metric multidimensional scaling on the Bray-Curtis distances for all 16S rRNA amplicon sequenced communities in the study (n=712 post-filtering) to broadly characterize *Salmonella* impact on community structure over time. CBA/J gut communities were similar in uninfected mice regardless of the sampling day. Infected

communities clustered most strongly when *Salmonella* levels were highest during timepoints near day 8 (Fig. 12d). Samples were divided into three roughly equal groups for statistical analysis (Mann Whitney U) – Early (days -6 – 2, n=293), Middle (days 3 – 8, n=241), and Late (days 9 – 15, n=178). Communities from high-responder mice in the Early group had significantly lower *Salmonella* relative abundance, alpha diversity (Shannon's D), and ratio of *Firmicutes* to *Bacteriodia* compared to both Middle and Late communities from the same mice (Fig. 12d, S6b). Microbiota richness in high-responder mice was significantly changed between all groups (Early, Middle, and Late) and between Late and the other groups for communities from uninfected and low-responder mice (*Salmonella* relative abundance never above 25%) (Fig. 12d, S6b). These results support the notion that *Salmonella* is altering the microbiota commensurate with its relative abundance in the community and the duration of infection by reducing bacteria richness and diversity and selecting for specific taxa (e.g. *Bacilli*).

Next, we analyzed the class distribution of ASVs from mice with paired metatranscriptomic and metabolomic data and observed significant ($p < 0.05$, Wilcoxson Rank-Sum test) community changes following *Salmonella* infection (Fig. 12c, S6d). These results indicate similarity in early timepoint community structure across mouse treatment groups and implicate community shifts with *Salmonella* presence. To determine specific impacts of *Salmonella* on whole community structure, we performed LEfSe linear discriminant analysis (LDA) on late timepoint (days 10–12) communities from mice with paired metatranscriptomic and metabolomic data[38]. *Clostridia* genera dominate the significant taxa in uninfected mice, and the total mean relative abundance of all significant *Clostridia* is 8.33x lower in infected communities compared to uninfected (Fig. 12f). *Bacilli* genera were the only taxa significantly enriched in infected mice other than *Salmonella*. *Enterococcus* genera were 129.8x enriched in the infected mice and *Lactobacillus* were 20.5x enriched during infection (Fig. 12f).

Co-occurrence of genera in the amplicon sequenced communities with *Salmonella* was determined via Spearman's correlation. The ASV table of high-responder samples from day 12 or later

(n=32) was collapsed to the genus level and we determined that *Lactobacillus*, *Enterococcus*, and *Lachnospiraceae* NK4A136_group were significantly ($p<0.05$) positively correlated with *Salmonella* during infection (Fig. 12e). Genera with significant positive correlations ($\boldsymbol{\rho} \geq 0.5$) to the three *Salmonella*-correlated genera were *Akkermansia*, *Alistipes*, and *Clostridia* vandinBB60 group (Fig. 12e). These results paired with the LDA results indicate a strong association of *Lactobacillus* and *Enterococcus* with *Salmonella* during later stages of infection.

3.3.2 Diet significantly impacts amplicon community structure regardless of infection state

ASV community structure significantly differed between chow fed mice and HFD fed mice as early as day -3 (chow diet initiation) and day -6 (HFD diet initiation) as determined by multi-response permutation procedures and analysis of similarities (Fig. 18a). The magnitude of HFD impact on beta-diversity (Bray-Curtis) increased from diet initiation (A=0.05, R=0.28) to one day post infection (A=0.22, R=0.86, Fig. 18a). Beta-diversity compared at the latest time point where amplicon data was collected from HFD mice (day 8) indicated significant community differences between mice in all diet and treatment groups (chow + infected, chow + uninfected, HFD + infected, HFD + uninfected) (A=0.36, R=0.73, Fig. 18a). Class distributions of communities from each diet/treatment group indicate an increased proportion of *Verrucomicrobiae* associated with HFD and a more immediate and pronounced bloom of *Salmonella* following treatment initiation compared to chow fed mice (Fig. 18b). All mice with metatranscriptomic sequencing and metabolite data were fed chow so as not to confound mechanistic conclusions drawn from expression and chemistry differences observed between infected and uninfected treatments. Nonetheless, these results indicate a strong diet effect on community structure, underpinning diet importance in future work examining the inflamed microbiome.

3.3.3 Metagenomic sequencing of Salmonella-infected mice indicates particular taxa persistence regardless of mouse diet or breed

Deep metagenomic sequencing was performed on 15 samples from both infected (n=9) and uninfected (n=6) treatments fed either chow (n=6) or HFD (n=9) resulting in 858.16 Gbps of sequencing

per sample and 6,221,387,714 total paired end reads (Fig. 17c). Total number of mice in each treatment cohort with metagenomic sequencing are: 3 (chow + infected), 3 (chow + uninfected), 4 (HFD + infected), 1 (HFD + uninfected). The three chow fed mice from each infection treatment were used previously to produce the CBAJ-DB and were sampled on day 11 post infection[5]. The remaining metagenomic sequencing originated from infected and uninfected mice fed HFD from either day 1 (n=2), day 5 (n=2), or day 8 (n=5). Metagenomic sequencing from HFD mice was assembled and binned (see methods) and resulted in 1,386 metagenomically assembled genomes (MAGs) that were dereplicated with the CBAJ-DB (n=2,281) to create a non-redundant medium and high quality (contamination < 10% and completeness ≥ 50%) database of 160 MAGs.

Trimmed metagenomic reads from the infected treatment samples of CBA/J mice in this study (n=7) and *Salmonella* infected C57BL/6J mice with a humanized microbiome originating from an external study (n=8, 179.56 Mbps of sequencing) were mapped to the MAG database to determine the effect of diet and mouse breed on individual bacteria prevalence during infection[39]. Reads from CBA/J chow, CBA/J HFD, and C57BL/6J mice mapped to MAGs from *Akkermansia muciniphila* and *Muribaculaceae* CAG-485. Other than *Salmonella,* these genomes were among the top four most highly mapped MAGs from each cohort (Fig. 12g). Additional prevalent genomes recruited reads only from CBA/J mice and included *Lactobacillus johnsonii*, *Enterococcus D gallinarum*, *Alistipes* sp002428825, and *Luxibacter massiliensis* (Fig. 12g). The lower read mapping from C57BL/6J mice was expected since no MAGs in the database originated from that sequencing and only 64% of the 255,206,015 reads from these samples mapped to the database. None the less, it is notable that the two genera found in all three metagenomic cohorts were in the highest relative abundance in C57BL/6J mice – *A. muciniphila* was 48% of the C57BL/6J community and CAG-485 was 7% of the community from that breed (Fig. 12g). These values are compared to the same genera prevalence in either CBA/J chow or CBA/J HFD mice where *A. muciniphila* accounted for 4% and 8% of the community respectively, and *M.* CAG-485 3% and 5% of

the community respectively (Fig. 12g). These results suggest at least some gut bacteria persist during infection due to factors beyond host genetics and diet.

3.3.4 Genes for succinate and D-lactate dehydrogenase cytochromes are predictive of bacteria co-occurrence with Salmonella

The MAG database was next evaluated to establish a subset of the community associated with *Salmonella*. These bacteria were determined by a combination of MAG abundance, association rule mining, and taxonomy linkages between *Salmonella*-correlated ASVs and MAGs (see methods). Two machine learning models (LASSO logistic regression and random forest) were trained using refined DRAM annotations from each MAG as predictor variables and *Salmonella*-association (associated MAGs n=51, not associated n=108) as the target variable. A 10-fold cross validated random forest model was trained on 75% of the data and tested on the remaining 25%, then variable importance was calculated to determine MAG functions with the most predictive power to classify a bacterium as *Salmonella*-associated (Fig. 13a). Next, we trained a logistic regression model with LASSO penalty on the same data. The regression model training was repeated 400 times with different random training and test splits (training 75%, test 25%) and the variable coefficients were examined (Fig. 13b).

Both machine learning approaches indicated that the presence of genes for D-lactate dehydrogenase was important in determining *Salmonella*-association (Fig. 13a, 13b). Additionally, both models indicate *Salmonella*-association is best predicted by the presence of cytochrome complexes in a bacterial genome, specifically complex III cytochrome bd ubiquinol oxidase and complex II succinate dehydrogenase (Fig. 13a, 13b). Functions in the energy category were also determined some of the most predictive variables by both models (Fig. 13a, 13b). Succinate oxidation to fumarate is potentiated by succinate dehydrogenase while lactate oxidation is dependent on cytochrome bd oxidase[40,41]. These results may indicate the importance of specific substrate utilization linked with the ability to oxidize D-lactate and succinate for bacteria to persist during *Salmonella* infection and inflammation.

3.3.5 Substrate utilization by relatively abundant and active bacteria favors microaerophilic and fumarate reduction processes during Salmonella infection

To evaluate community activity during *Salmonella* infection we extracted whole community RNA and performed metatranscriptomic sequencing on the corresponding cDNA from 10 chow fed CBA/J mice (5 infected treatment, 5 uninfected treatment) resulting in 135.4 Gbp and 1.073 billion paired end reads of transcript data (Fig. 12b, S6c). All metatranscriptomic data were collected from day 11 fecal samples, and all infected samples had a *Salmonella* ASV relative abundance of >30% except the sample from mouse I7 (Fig 7c). Because of the low *Salmonella* relative abundance in this sample, it was not included in the differential expression analysis. Infected sample GeTMM normalized counts of mapped expression from CBA/J chow mice (n=5) to the MAG database were compared to GeTMM normalized mapped metagenomic reads from CBA/J chow infected samples (n=3). *L. johnsonii* received the most mapping from the metatranscriptomic reads (98.8 mean GeTMM) and it was the third most abundant genome by normalized metagenomic read mapping (10% average relative abundance, Fig. 13c). The next most active commensal bacterium during infection was *L. massiliensis* (48.7 mean GeTMM) though *L. massiliensis* mean relative abundance in infected samples was only 4.32% (Fig. 13c). *Salmonella* was the most abundant genome (50.3% average relative abundance) in the infected samples, but *Salmonella* expression (62.4 mean GeTMM) was only 63.2% that of *L. johnsonii* (Fig. 13c). *Alistipes* sp002428825 was the next most abundant genome (13.6% average relative abundance) yet its average expression level was only 1.53 GeTMM (Fig. 13c). *A. muciniphila* was another taxon with disparity between MAG relative abundance and expression level. It was the fifth most relatively abundant MAG (3.59% average relative abundance) in infected CBA/J mice, yet its expression level was the tenth highest (2.43 mean GeTMM, Fig. 13c). These results suggest that certain bacteria become more active with *Salmonella* infection (e.g. *L. johnsonii, L. massiliensis*) and others remain relatively abundant but reduce their activity in response to *Salmonella* (e.g. *Alistipes* sp002428825, *A. muciniphila*). Active or relatively abundant bacteria in this instance may be benefitting from host inflammation products or synergy with *Salmonella*.

Substrate utilization expression profiles within treatments showed that *E. gallinarum*, *L. johnsonii*, and *Enterococcus faecalis* fumarate reduction increases during infection (Fig. 13d). Both *A. muciniphila* and *M.* CAG-485 sp002362485 reduce expression of microaerophilic genes during infection. Surprisingly, *A. muciniphila* down regulated expression of cytochrome bd ubiquinol oxidase (cydB, cydA) by more than 1000x during infection and *M.* CAG-485 sp002362485 reduced expression of NADH-quinone oxidoreductase (nuoCD, nuoL) more than 115x. *Salmonella* displays the most varied substrate utilization, expressing sufficient genes for us to confidently assert it is utilizing tetrathionate, reducing sulfur and fumarate and nitrate, as well as performing aerobic respiration and microaerophilic processes (Fig. 13d). Of note, any *Salmonella* expression observed in the uninfected treatment can be attributed to nonspecific mapping to other *Enterobacteraceae* and should be considered as such. Together, these results corroborate our observations of *A. muciniphila* and *M.* CAG-485 sp002362485 general reduction of gene expression during *Salmonella* infection, and *Bacilli* spp. increase in activity commensurate with high levels of *Salmonella* (Fig. 13c). Additionally, these data suggest fumarate is an important compound linked to *Bacilli* enrichment, and other persistent classes (*A. muciniphila, M.* CAG-485 sp00236485) shift away from oxygen respiration with high affinity oxidases, perhaps because of competition for molecular oxygen with *Salmonella* and sulfide autoxidation.

3.3.6 Metatranscriptomic differential gene expression reveals both substrate exclusivity and competition between Salmonella and commensal bacteria

Differential gene expression analysis indicated a significant shift in lactate dehydrogenase gene expression in the community following *Salmonella* infection (Fig. 14a). *L. johnsonii* and *E. gallinarum* significantly upregulated ldh (L-lactate dehydrogenase, K00016) 4.85 and 5.68 (mean logfc) respectively in the infected treatment (Fig. 14a). *Coprocola* sp910575055 and *A. muciniphila* both significantly downregulated ldhA (D-lactate dehydrogenase, K03778) following *Salmonella* infection, reducing expression 5.18 and 4.82 (mean logfc) respectively (Fig. 14a). The stereospecific lactate dehydrogenase (ldhA) was significantly upregulated by *L. johnsonii*, displaying a 3.48 log fold-change increase in the infected treatment compared to the uninfected (Fig. 14a). *Salmonella* lactate dehydrogenase expression

encompassed two genes (dld K03777, ldhA K03778) specific to the D-lactate enantiomer and one gene specific to the L-lactate enantiomer (lldD K00101)[13,22]. *Salmonella* expression of ldhA averaged 63.5 counts (GeTMM), and *Salmonella* was the only bacteria expressing dld, averaging 31 counts (GeTMM). These counts are far greater than the mean *Salmonella* lldD expression of 0.82 (GeTMM). Combined, these lactate expression profiles indicate a strong preference for D-lactate by *Salmonella* and *Bacilli* spp. preference for L-lactate.

We observed significant overlap in expression of genes associated with chitin backbone cleavage between members of the *Bacilli* class and *Salmonella* (Fig. 14a). Gene expression of chitinase (GH18, EC.3.2.1.14) by *E. gallinarum* increased 7.67 logs during *Salmonella* infection (Fig. 14a). *Salmonella* also expressed genes for chitinase (GH19, EC.3.2.1.4), averaging 4.69 read counts (GeTMM) across infected samples (Fig. 14a). Interestingly, *Salmonella* was actively expressing G-type lysozyme (GH23, EC.3.2.1.17) and *L. johnsonii, E. gallinarum,* and *E. faecalis* all displayed significant increases in lysozyme (GH73, EC.3.2.1.17) expression during infection (4.84, 5.86, and 4.07 logfc increase respectively). Each of these glycoside hydrolases may be involved in chitin backbone cleavage, suggesting potential competition for chitin as a metabolic substrate or it may indicate more direct competition of *Salmonella* and *Bacilli* members via antimicrobial enzymatic activity.

3.3.7 Sulfur reduction genes are expressed by abundant bacteria to persist in the inflamed gut

Examination of gene expression of prominent families during *Salmonella* infection revealed many genes associated with oxidative stress response and sulfur reduction were upregulated (Fig. 14c). Among the families *Muribaculaceae, Akkermansiaceae, Enterobacteriaceae, Lactobacillaceae, Lachnospiraceae,* and *Enterococcaceae* thioredoxin (trxA, trxB, trxC, tpx) genes were being actively expressed during infection (Fig. 14c). The average expression of thioredoxin genes in *Lactobacillaceae* was increased 15x with infection and increased 197.7x in *Enterococcaceae*. Glutathione reductase (gor) expression was only present during infection in *Enterococcaceae, Lactobacillaceae,* and

Enterobacteriaceae and increased during infection in all three families (Fig. 14c). *Akkermansiaceae* decreased expression of sulfite reductase (cysJ), thioredoxin genes (trxA, trxB,), and peptide methionine sulfoxide reductases (mrsB, mrsAB) during infection (Figuer 3c). *Muribaculaceae* expression of thioredoxin genes (trxA, trxB, tpx) was also decreased during infection (Fig. 14c). These results coincide with generally lower expression profiles of predominant members in *Akkermansiaceae* (*A. muciniphila* 86.2x lower expression in infected treatment) and *Muribaculaceae* (*M.* CAG-485 sp002362485 12.3x lower expression in infected treatment) following *Salmonella* infection (Fig. 14a, 14c). Thioredoxin peroxidase (tpx) was expressed by all the families previously listed bar *Akkermansiaceae* (Fig. 14c).

Genes involved in the dissimilatory sulfate reduction pathway were also transcribed in the inflamed system. Sulfate adenylyltransferase (cysNC/cysN/cysD) catalyzing the conversion of sulfate to adenoside 5'-phosphosulfate (APS) was transcribed by *Salmonella* and *Clostridia* in genus *Dysosmobacter* and *Anaerotruncus colihominis_A* during infection (Fig. 14c, 16). Sulfite reductase (cysI/cysJ/asrA/asrB/asrC) responsible for converting bisulfite to persulfurate and then hydrogen sulfide was expressed by *A. muciniphila, Lachnospiraceae* spp. (n=4), and *Salmonella* during infection (Fig. 14c, 16)[42,43]. Additionally, evidence of oxidative stress responses to reduce methionine sulfoxide (msrA/msrB/msrC/msrAB) and to protect DNA with sulfur modification (dndE) were displayed during infection by members of *Enterococcaceae, Lactobacillaceae, Akkermansiaceae,* and *Salmonella* (Fig. 14c, 16). Together these results suggest a key role for sulfur reduction in response to oxidative stress during *Salmonella* infection and show some of the most successful bacteria in the inflamed gut employ multiple sulfur antioxidant mechanisms.

3.3.8 Sulfatase activity and prevalent sources of sulfur implicate the microbiota with Salmonella tetrathionate respiration

It is generally accepted that *Salmonella* can respire tetrathionate to outcompete commensal bacteria during infection[15,16,44]. Our expression data corroborate this showing *Salmonella* expressing tetrathionate reductase genes ttrA, ttrB, and ttrC (Fig. 14c, 16). *Salmonella* is also expressing three thiosulfate reductase genes (phsA, phsB, and phsC) capable of reducing thiosulfate produced by mitochondrial oxidation of hydrogen sulfide in colonocytes[45,46] (Fig. 14c, 16). Furthermore, evidence of sulfatase expression by relatively abundant commensal bacteria *Akkermansia, Enterococcus, Muribaculaceae* CAG-485 sp002362485, and *Alistipes* sp002428825 was indicated in our metatranscriptomics data (Fig. 14b). Annotated genes were mined for cystine and serine sulfatase signatures (C/S)XPXR and CXAXR[47]. Genes with the sulfatase motif were then cross-referenced with Pfam annotations and only Sulfatase [PF00884.26] and Sulfatase-modifying factor enzyme [PF03781.19] annotated genes were considered. *Enterococcus* and *Akkermansia* both expressed Sulfatase-modifying factor enzyme capable of activating CXAXR motifs to FGlyXAXR active sites of arylsufatases[47] (Fig. 14b, 16). These results indicate commensal sulfatase expression during *Salmonella* infection, and the potential for sulfate liberated by the microbiota to oxidize in the inflamed lumen to then act as a terminal electron acceptor for *Salmonella* respiration.

To examine any changes in the chemistry of *Salmonella* infected microbiomes, we performed LC-MS/MS untargeted metabolomics on 7 uninfected samples and 6 infected samples from day 12 chow fed CBA/J mice, including 3 technical replicates for each (Fig. 12a, 12b). The resulting annotated list of compounds included 342 uniquely identified metabolites in both treatments. We performed a principal component analysis (PCA) (see methods) to determine overall difference in the Pareto scaled metabolite chemistry of each treatment (Fig. 15a). A PERMANOVA test on Bray-Curtis distances calculated from the metabolite abundance table indicated sample clusters from each treatment were significantly different ($p=0.001$), and *Salmonella* treatment group accounted for 72% of the variation in the PCA. Most of the variance in the PCA across principal component 1 (63%) was driven by flavonoid, amino acid, and some bile acid compounds (Fig. 15a).

Metabolomic evidence of abundant sulfated bile acids and sulfated flavonoids in the infected treatment offer potential sources of sulfate from commensal and *Salmonella* sulfatase activity (Fig. 15b, 15c). These sulfate groups, added to bile acids and flavonoids in the liver, are ultimately host derived sources of sulfur that can be oxidized by ROS to tetrathionate following commensal and *Salmonella* sulfatase liberation. Evidence of commensal and *Salmonella* sulfate reduction to hydrogen sulfide is present in our data (Fig. 14b, 14c, 16). Hydrogen sulfide is a toxic product of microbial sulfur reduction and colonic epithelial cells oxidize it to sulfate and thiosulfate[11]. In an oxidizing environment like the inflamed gut, these compounds can become further oxidized to tetrathionate or hydrogen sulfide can be directly oxidized by ROS[16]. The high prevalence of methionine sulfoxide in infected metabolomes and the low prevalence of hydrogen sulfide suggest an oxidative environment affecting inorganic sulfur and thiol redox during *Salmonella* infection (Fig. 15b, 15c). From our metatranscriptomic and metabolomic data we infer commensal bacteria and *Salmonella* sulfatase activity during infection is releasing host-derived sulfate from sulfonated bile acids and sulfonated flavonoids that can then be respired by *Salmonella*, though further work is needed to confirm this (Fig. 14b, 11).

The total flavonoid content averaged within treatments is 8.6x higher in infected than in uninfected guts. This ratio is far higher when considering only the sulfated flavonoid species in our metabolomics data. Apigenin 7-sulfate is 362.4x more abundant in infected guts while quercetine 3-rhamnoside-3'-sulfate is absent completely in the uninfected treatment and has an average abundance of 15,435.4 in the infected treatment (Fig. 15c).

Free sulfate in the lumen may be internalized and reduced by active bacteria in the gut, suggested by certain sulfur reduction pathways expressed by *Salmonella* and commensal bacteria during infection. Sulfate freed by microbial sulfatase activity may be internalized prior to ROS oxidation to tetrathionate by sulfate/thiosulfate ABC transport systems (cysP, cysW, cysA, sbp) expressed during infection by

Salmonella and members of *Clostridia* (Fig. 16)[48]. Once inside the cell, reduction of sulfate to hydrogen sulfide is possible by gene products (cysD, cysN, cysC, cysH, cysI, cysJ, cysG, asrA, asrB, and asrC) expressed by *Salmonella* in our transcript data (Fig. 16). Additionally, members of *Clostridia* were expressing anaerobic sulfite reductase genes (asrA, asrB, and asrC) and some were expressing cysN (ATP sulfurlyase) for the conversion of sulfate to APS[48]. *L. massiliensis* expressed cysC (APS kinase) which converts APS to 3'-phosphoadenosine 5'-phosphosulfate (PAPS) and *A. muciniphila* expressed cysJ (sulfite reductase) to produce sulfide from sulfite (Fig. 16)[48]. These results indicate sulfur reduction occurring across prominent taxa groups to produce hydrogen sulfide in the inflamed intestine.

3.3.9 Bile acid pool composition is significantly altered in the inflamed gut

Our metabolomic data show 10 bile acid compounds in either infected or uninfected mice and the overall bile acid pool was largest in uninfected samples (Fig. 15b, 16). In both treatments, there was a higher ratio of secondary bile acids to primary bile acids, on average 3.7:1 and 12.3:1 in uninfected and infected respectively. There was also a disproportionate amount of sulfated bile acid in the infected treatment compared to the uninfected, an average ratio of 40:1 (sulfated : unsulfated) in infected and a ratio of 3.7:1 (unsulfated : sulfated) in uninfected. Finally, the ratio of amino acid conjugated bile acids to unconjugated bile acids was dramatically different between treatments. In the infected treatment the average ratio of conjugated to unconjugated bile acids was 32.5:1 and in uninfected treatment there was 62.4x more unconjugated bile acids than conjugated ones. Implications of these data are that microbial modification of the bile acid pool is diminished during infection, but also that new primary bile acids are not being added to the pool, and that host-derived sulfur is far more prominent in the inflamed gut in the form of bile acid sulfates.

Metatranscriptomic data indicates the only gene expression in infected guts for conversion of primary bile acid to secondary bile acid is baiA by *L. massiliensis* (9x more expression in infected treatment, Fig. 16). In the uninfected treatment, genes (baiA, baiB, baiCD, and baiN) for the 7α-

dehydroxylation pathway are expressed by *L. massiliensis* and 23 other *Clostridia*. The presence of secondary bile acids derivative of lithocholic acid, deoxycholic acid, and ursodeoxycholic acid indicates further microbial activity to oxidize and epimerize bile acids remains uncaptured by our metatranscriptomic sequencing in inflamed guts[32]. A prerequisite to 7α-dehydroxylation of primary bile acids is the removal of conjugated amino acids taurine and glycine by microbial bile salt hydrolase (bsh)[32]. In uninfected mice bsh expression is almost entirely from *Coriobacteriia* and *Clostridia*. However, in infected mice *L. johnsonii* and *E. gallinarum* express nearly all the bsh. Together these data indicate a shift in the bacteria responsible for microbial bile acid modification from *Clostridia* to *Bacilli* during infection, and the metabolomic data would suggest an overall reduction in bsh activity and increased secondary bile acid sulfation by the host in *Salmonella*-included guts (Fig. 15c).

3.3.10 Commensal activity under oxidative stress provides free sulfur from amino acids

Metabolomic evidence of sulfur containing amino acids and amino acid derivatives depict significantly higher levels of free taurine, histidinyl-methionine, hypotaurine, and L-homocysteine sufinic acid in uninfected samples compared to infected (Fig. 15c). We speculate less free taurine and hypotaurine is resultant from reduced *Clostridia* bsh activity on conjugated bile acids evidenced by the higher proportion of bile acid conjugates in infected samples (Fig. 15c). However, infected samples contain higher levels of oxidized methionine (methionine sulfoxide) and ample evidence of bacterial utilization of methionine and related amino acids (Fig. 16).

As an important antioxidant mechanism in the oxidative stress response to luminal inflammation during *Salmonella* infection, bacteria maintain the reduced form of methionine by expressing methionine sulfoxide reductases (msrA,B,C)[43,49]. Our data show these genes are expressed in the inflamed gut by members of the *Bacilli* including *L. johnsonii*, E. *gallinarum*, and *E faecalis*. *Salmonella* and *A. muciniphila* are also actively reducing methionine sulfoxide during infection (Fig. 16). Further amino acid associated sulfur reduction is occurring in the inflamed system potentially producing hydrogen sulfide.

Cystathionine-γ-lyase (mccB) converts methionine to homocysteine and mccB expression during infection is detected in *E. gallinarum, E. faecalis, L. johnsonii, L. masilliensis,* and *Salmonella* (Fig. 16)[43]. Hydrogen sulfide production from homocysteine is catalyzed by 5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase (metE) expressed by *E. gallinarum* in the inflamed system[43]. Homocysteine can also be produced from cysteine by conversion to cystathionine by cystathionine-γ-synthase (metB) or cystathione-γ-lyase (mccB) and finally homocysteine by cysteine-S-conjugate beta-lyase (malY, metC)[43]. Homocysteine conversion to methionine is facilitated by 5-methyltetrahydrofolate-homocysteine methyltransferase (metH)[43]. *Clostridia* (n=10) including *L. massiliensis* as well as *Salmonella, A. muciniphila,* and *Alistipes* sp002428825 are expressing metH during infection to produce methionine from homocysteine (Fig. 16) and many bacteria including prominent *Bacilli* and *M.* CAG-485 along with *Salmonella* are expressing the genes malY, metC, and mccB converting cysteine to homocysteine (Fig. 16)[43]. Furthermore, direct hydrogen sulfide production from cysteine catalyzed by cysteine synthase (cysK), S-sulfo-L-cysteine synthase (cysM), and cysteine desulfurase (iscS) is expressed by *Bacilli* including *L. johnsonii, E. gallinarum,* and *E. faecalis* as well as members of the *Clostridia* including *L. massiliensis. A. muciniphilia* expresses cysK and cysM during infection and *Salmonella* is expressing cysK, cysM, iscS, as well as dcyD (D-cysteine desulfhydrase) (Fig. 16).

The expression and chemical evidence for commensal production of hydrogen sulfide from amino acid sources during *Salmonella* infection is strong, driven primarily by the dominate *Bacilli* spp. in the inflamed gut and supported by members of *Clostridia* and *Bacteroidia*. *Salmonella* also participates in all the mentioned sulfur amino acid reductions (Fig. 16). It can be speculated that free hydrogen sulfide is either immediately oxidized by ROS in a highly inflamed environment or is sufficiently proximal to colonocytes to be absorbed by the host, mechanisms that may explain the significantly lower hydrogen sulfide levels detected in inflamed metabolomic samples (Fig. 15c).

3.4 Discussion

3.4.1 Multi-omics reveals persistent bacteria may benefit from inflammation products and participate in Salmonella cross feeding during infection

The diverse genetic repertoire of *Salmonella* to both induce and capitalize on host inflammation is well documented[8,9,11–14,16,23]. Few studies, however, approach *Salmonella* remodeling of the gut ecosystem with multi-omic synthesis as was presented here – considering the microbiome autochthonous community, chemistry, and microbial genetic potential and expression. Not only were we able to show *Salmonella* obliteration of the normal commensal community using 16S rRNA amplicon sequencing and metagenomics, but we were able to identify select members of the community that were seemingly immune to both *Salmonella* competition and non-specific immunity of the host regardless of host diet or genetics. Additionally, we were able to show the impact of this community shift on microbial function and interaction with paired genome resolved metatranscriptomics and comprehensive untargeted metabolomics. Notably, this multi-omics approach facilitated the discovery of mechanistic connections between host provided sulfur and lactate aiding in *Salmonella* proliferation via critical commensal activity and individual expression of genes for sulfatase, sulfur reduction, and lactate utilization and production. With the revelation that *Salmonella* may benefit from certain synergistic relationships with resident bacteria, we leveraged association rule mining and correlation analysis to create a cohort of bacteria considered to be *Salmonella*-associated during inflammation. Two machine learning models trained on the gene content of MAGs reconstructed from our model ecosystem corroborated our multi-omics data, indicating both lactate and succinate cytochrome importance to predict *Salmonella*/commensal association. Our modeling connected these well-known substrates used by *Salmonella* with expressed pathways from prominent *Salmonella*-associated bacteria involved with succinate and lactate accumulation during infection.

3.4.2 Salmonella lactate dehydrogenase expression indicates both D-lactate production and oxidation

Convention shows a reduction of *Clostridia* and butyrate starvation of intestinal epithelia as hallmarks of *Salmonella* impact on the microbiome[11,24]. Inflammation promotes colonocytes to release L-

lactate into the lumen as they perform less oxidative phosphorylation and more glycolysis due to diminished peroxisome proliferator-activated receptor (PPARγ) signaling[11,13,22,50]. Our data substantiates this, indicating a reduction of *Clostridia* in the microbiome during *Salmonella* infection and an increase of lactate in the lumen. Prior work involving *Salmonella* lldD (L-lactate dehydrogenase) and dld (D-lactate dehydrogenase) knockouts indicated *Salmonella* utilization of specifically host-derived L-lactate[22]. Contrarily, our data shows *Salmonella* D-lactate dehydrogenases (dld, ldhA) expressed 43.4x more than L-lactate dehydrogenase (lldD). Considering the absence of lactate racemase expression, these results suggest *Salmonella* is primarily using lactate from a source other than the host[51]. Furthermore, *Clostridia* species *Coprocola* sp910575055, *A. muciniphila*, *L. johnsonii*, and *Salmonella* highly expressed (mean 44 GeTMM) ldhA, a fermentive lactate dehydrogenase known to produce D-lactate from pyruvate[52]. D-lactate is also potentially produced from racemic products of *Bacilli* ldh conversion of pyruvate to lactate, indicating the potential for *Salmonella* cross feeding on microbially derived D-lactate[53]. Bacteria expressing ldh may also be converting host derived L-lactate to pyruvate suggesting some commensal bacteria may metabolically benefit from the host immune response. The potential for a more intricate role of lactate in *Salmonella* physiology also exists, considering *Salmonella* dld expression. Researchers investigating D-lactate dehydrogenase in *Shewanella oneidensis* MR-1 (*Gammaproteobacteria*) showed ldhA and dld used to cycle NAD+ and NADH in N-acetylglucosamine metabolism[52]. Considering, N-acetylglucosamine is a primary component of chitin, and we observed transcription evidence of competition for chitin between *Salmonella* and *L. johnsonii, E. gallinarum,* and *E. faecalis*, these results are particularly intriguing and may imply a similar D-lactate mechanism occurring in *Salmonella*.

Alternative fates for lactate in the inflamed intestine are microbial transformation to succinate or propionate via the mixed acid pathway (MAP) or the acrylate pathway (AP)[54,55]. Succinate levels in our metabolomics data were 2.6x (mean abundance) higher in infected samples than in uninfected, and propionate levels in uninfected samples were 23x (mean abundance) higher than in infected (Fig. 16). Lactate derived pyruvate and subsequent phosphoenolpyruvate can enter MAP to produce succinate[55,56].

Genes involved in MAP are expressed by *Salmonella* and multiple members of the commensal microbiota during infection, including *Alistipes* sp., multiple prominent *Bacilli* spp., *M.* CAG-485 sp002362485, *Clostridia* spp., and *A. muciniphila* (Fig. 16). Succinate can be further catabolized to propionate via the succinate pathway actively expressed in our data during infection by members of the *Clostridia*, *Dehalobacteriia*, *A. muciniphila*, *M.* CAG-485 sp002362485, *Alstipes* sp., *L. massiliensis*, and *Salmonella* (Fig. 16)[56]. Evidence that lactate is being converted to propionate via AP is clear in the infected gene expression of *L. massiliensis*, 7 other *Clostridia* spp., *E. gallinarum*, *E. faecalis*, and *Salmonella* (Fig. 16)[54]. In this way various community members are producing propionate during active *Salmonella* infection. Prior reports show *Salmonella* catabolism of propionate to support growth and our data indicate *Salmonella* expression of the complete *prpBCDE* operon necessary for this function (Fig. 16)[7]. Furthermore, other groups have shown *Salmonella* utilization of succinate during infection, a notion supported by our data which shows *Salmonella* expression of the complete oxidative TCA cycle necessary for succinate utilization as a carbon source (sdh, sucABCD)(Fig. 16)[12].

Seven of the top ten most prominent bacteria (Fig. 13c) in the infected intestine express L-lactate dehydrogenase. Of these, four are *Bacilli* including *L. johnsonii* and *E. gallinarum* both of which are also expressing L-lactate oxidase (lctO) during infection. Other groups have shown *Lactobacillus plantarum* to utilize lactate in the presence of oxygen to produce acetate, potentially catalyzed by lactate oxidase[57]. It is possible that *Lactobacillus* and *Enterococcus* are utilizing host L-lactate as an energy source to persist during *Salmonella* infection. Prior work has shown *Clostridia* are some of the main lactate utilizing bacteria in the healthy gut, converting lactate to butyrate or propionate[58,59]. Certain SRBs can also use lactate to produce acetate[58]. High lactate and acetate levels can lower the pH in the gut, further affecting the microbiota through loss of acid sensitive taxa and exacerbating a dysbiotic SCFA profile by decreasing butyrate and propionate production[60]. Our metabolomic data indicate low levels of butyrate and propionate in the infected gut and high levels of lactate and acetate. Moreover, the high levels of succinate in the infected treatment provide metabolite evidence of dysbiosis. The accumulation of

succinate in the large intestine is associated with multiple inflammatory pathologies including ulcerative colitis and inflammatory bowel disease, and is associated with microbial community imbalance and augmented inflammatory response via succinate receptor 1 activation[61]. Succinate levels in the *Salmonella*-included mice from our study are elevated compared to levels in uninfected mice (Fig. 16). We propose that the altered gut community during inflammation is capitalizing on host-derived L-lactate to produce acetate, and high succinate and lactate levels contribute to the community dysbiosis and exacerbate the host inflammatory response via acidification and succinate receptor 1 agonism respectively.

3.4.3 Commensal bacteria and Salmonella benefit from multiple sources of sulfur during inflammation including bile constituents and amino acids

Prior studies showed *Salmonella* using tetrathionate as a terminal electron acceptor for anaerobic respiration to outcompete resident microbiota in the gut, and a key source of tetrathionate during *Salmonella* infection is thiosulfate oxidized by ROS[15,16,44]. It is well known that colonic epithelial cells produce thiosulfate from hydrogen sulfide as a mechanism to detoxify the lumen[45,46]. Our work identified multiple pathways for hydrogen sulfide production being expressed by inflammation resistant members of the microbiota as well as *Salmonella*. Sulfatase expression by some of the most relatively abundant and active bacteria (*Salmonella*, *Enterococcus*, *Akkermansia, Alistipes, M.* CAG-485) was detected, as were high abundances of sulfated bile acids and sulfated flavonoids during infection.

Flavonoids and bile acids are conjugated to sulfonate in the liver and released back to the GI via enterohepatic circulation[62,63]. Sulfation of these molecules changes their biochemical impact in the microbiome[64,65]. Sulfated bile acids have lower absorption rates in the intestine and amino acid conjugated bile acids undergo less passive absorption and are less toxic than unconjugated bile acids[62]. Both bile acid classes (sulfated and conjugated) were increased during infection in our data. Sulfated flavonoids have reduced anti-inflammatory radical scavenging potential because one or more hydroxyl groups of the B ring are replaced by sulfate[66,67]. Quercetin 3-rhamnoside-3'-sulfate was one of the

abundant sulfated flavonoids in our infected metabolite data (Fig. 15c). Quercetin rhamnose conjugates are common in food and are poorly absorbed without prior glycoside hydrolase separation of the flavone from the sugar, and 3'sulfated quercetin has diminished antioxidant potential that may be restored by bacterial sulfatase activity[66,68]. Speculating, the diminished diversity of the infected community may leave diet derived compounds like flavonoids underutilized by the microbiota, contributing to their relatively high levels in the infected treatment mice. Consequently, our data show the bile composition during *Salmonella* infection becomes a rich source of sulfur and has less antioxidant and antimicrobial potential. Interestingly, *Clostridia* (*Anaerotruncus* sp000403395) was among the few bacteria to express bsh during infection, a process that makes bile acids available for further microbial modification like 7α-dehydroxylation[69]. Our data indicates the potential for microbial sulfate liberation from abundant bile constituents and a reduction of other microbial bile acid transformations, both processes which can impact the community composition of the gut microbiome. These expression and metabolite data suggest a potential mechanism that links commensal microbiota to *Salmonella* tetrathionate respiration during inflammation via host derived organic sulfur bile constituents.

Salmonella and commensal expression of tetrathionate reductase, thiosulfate reductase, sulfite reductase, and various genes involved in assimilatory sulfate reduction indicate a continuous reduction of sulfur species during inflammation regardless of the sulfur oxidation state. Once sulfur is reduced to sulfate, *Clostridia* continue the reduction to sulfite and then hydrogen sulfide or incorporate it in cysteine synthesis (Fig. 16). *Clostridia* and relatively abundant members of *Bacilli* and *Bacteroidia* are expressing genes (metE, iscS, cysK, cysM, dcyD) to produce hydrogen sulfide from cysteine and homocysteine during infection. Indeed, our transcription data indicate prominent and active members of the microbiota are interacting with compounds that are copious during infection, like methionine sulfoxide, to reduce it and produce homocysteine and cysteine (Fig. 16). Oxygen tolerance in microaerophilic lactic acid bacteria involves oxidase (pyruvate oxidase, lactate oxidase, and NADH oxidase) conversion of molecular oxygen to water or hydrogen peroxide, necessitating further mechanisms to both limit Fenton

reaction and protect DNA and protein[70]. Mechanisms to withstand intracellular oxidative stress, including thioredoxin and glutathione reductase genes (Fig. 14c), are expressed by prominent members of the inflamed community and take advantage of sulfur multivalent antioxidant potential. The reduction of ROS by tpx is perpetuated by tpx disulfide bond reduction by trxA, and glutathione reductase is important in lactic acid bacteria for maintaining reduced glutathione amidst oxidant stress from host immunity[71,72] . Glutathione-based oxidative stress response may be more important than thioredoxin-based response for persistence in the inflamed gut as only the most active and enriched bacteria (*Salmonella, L. johnsonii,* and *Enterococcus* spp.) expressed gor during infection, and the most successful *Enterococcus* (*E. gallinarum*) expressed gor 5x more than *E. faecalis*. Maintenance of a reduced intracellular space underpins the importance of sulfur to bacteria persistence in the inflamed gut, linking sulfur reduction potential to community robustness during inflammatory events like *Salmonella* infection.

3.5 Conclusion

The work presented here has identified a set of commensal bacteria that persist during *Salmonella* infection and actively participate in the utilization and production of key metabolites in the inflamed gut ecosystem. Co-occurrence with *Salmonella* in the inflamed gut was predicted by individual bacteria gene content, and expression profiles from prominent persistent commensal bacteria indicated both synergistic and competitive potential with *Salmonella* during infection. These results offer multiple exciting avenues for probiotic strain identification of bacteria robust to enteric inflammation and pathogen perturbation that may be pivotal to reestablishing normal flora following infection. We confirmed *Salmonella* expression of genes to utilize succinate, propionate, tetrathionate, and thiosulfate *in vivo* amidst a complete microbial community. Our data also challenges current paradigms of *Salmonella* lactate utilization during infection and proposes a role for commensal bacteria in *Salmonella* tetrathionate respiration. Further work is necessary to confirm the mechanistic assertions made here, but the inferences drawn underpin the value in holistic examination of the microbiome afforded by a multi-omics approach.

3.6 Methods

3.6.1 Strains and media

Salmonella enterica serovar Typhimurium strain 14028 (*S. typhimurium* 14028) cultures were washed and resuspended in water after overnight incubation in Luria-Bertani (LB) broth at 37 °C with constant agitation.

3.6.2 Animals and experimental design

Female CBA/J mice were procured from The Jackson Laboratory (Bar Harbor, ME). Mice were randomly selected to populate cages and treatment groups (infected n=44, uninfected n=23, chow = 42, HFD = 25), and were kept 5 per cage in conventional enclosures in a temperature controlled 12 hour light/dark cycle. Irradiated mouse chow (Teklad, 7912) was made available *ad libitum* to the chow group mice and BioServ Product #S3282 to the HFD group mice. Mice in the infected group were presented with 10^9 CFU *S. typhimurium* 14028 via oral gavage on Day 0 with no subsequent treatment. Uninfected mice were left with no treatment. Animal experiment protocol was approved by The Ohio State University Institutional Animal Care and Use Committee (IACUC; OSU 2009A0035).

3.6.3 Sample collection

Mouse fecal pellets were collected from each mouse before and after treatment initiation (on days -6, -3 – 15) on autoclaved aluminum foil. Fecal pellets were immediately placed in labeled microcentrifuge tubes and flash frozen with EtOH/dry-ice prior to storage at -80 °C until further processing.

3.6.4 DNA/RNA Extraction and Sequencing

All nucleic acid extraction for 16S rRNA amplicon sequencing was performed using ZymoBIOMICS DNA/RNA Miniprep Kit (Zymo Research) and stored at −20 °C until further processed. PCR amplification of the V4 hypervariable region of 16S rRNA gene was performed using 30 cycles and

unique sequence barcodes in each primer were used to identify multiplexed samples. Both primers contained sequencer adapter regions. Total nucleic acid extraction and sequencing for metagenomics was performed as described in Leleiwi et.al[5]. Metatranscriptomic RNA extraction and isolation was performed on feces from 10 mice sampled 11 days post infection with Zymo-Seq Ribo Free Total RNA Library Kit Cat No. R3000 and 2 x 151 bp paired end reads were produced from corresponding cDNA at UC Denver Sequencing Facility using an Illumina HiSeq 2500 Sequencing System.

3.6.5 Metabolite Sample Preparation and Analysis

1 mL of a solution composed of three different solvents (water/methanol/dichloromethane, 1/2/3, v/v/v) was used for fecal metabolite extraction, followed by physical disruption with a sonicator (Bioruptor®, Diagenode, Belgium). The disrupted fecal suspension was vortexed and incubated at room temperature. The top aqueous layer was transferred, dried under vacuum, and reconstituted for metabolomics analysis.

The Agilent 6545 QTOF mass spectrometer equipped with Agilent 1290 UHPLC systems liquid chromatography was used for an untargeted metabolomics study on mouse fecal pellet extractions (n=13) from day 12. To improve the metabolite coverage, the prepared fecal samples were subjected to reversed-phase liquid chromatography with C18 column (ACQUITY UPLC® HSS T3 1.8 µm, 2.1 x 100 mm, Waters Corporation, MA, USA) and hydrophilic interaction liquid chromatography (HILIC, ACQUITY UPLC® BEH HILIC 1.7 µm 2.1 X 150 mm, Waters Corporation, MA, USA). Water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B) were used for reversed-phase separation. The flow rate was set at 0.3 mL/min with the gradient as follows: 2% B for 0-2 min, from 2% B to 30% B for 4 min, to 50% B for 8 min, and 98% B for 1.5 min and held at 98% B for 1min, then returning into initial gradient for equilibrium for 1.5 min. For HILIC separation, Water/acetonitrile (95/5) with 0.1% formic acid and 10 mM ammonium formate (A) and water/acetonitrile (5/95) with 0.1% formic acid and 10 mM ammonium formate (B) were prepared. For gradient elution, 99% B was held for 2 min, gradually

reduced to 75% B for 7 min and reduced again to 45% B for 5 min. And the gradient was held at 45% B for 2 min, and returned to 99% B. The flow rate was set at 0.3 mL/min. The quality control (QC) sample was prepared by mixing an equal volume of each sample. The QC sample was analyzed after every 6 samples.

The collected mass spectra were converted into mzML format with MSconvert in Proteowizard[73]. The data analysis for untargeted, global profiling experiments was performed with Progenesis QI (Waters Corporation/Nonlinear Dynamics). The Compound MS/MS library used in the study for metabolite annotation was the embedded Progenesis database with METLIN/Waters collaboration for fragmentation patterns, human metabolome database (HMDB), E. coli metabolome database (ECMDB), and Lipid Maps. All annotated metabolites were manually inspected to be reported and metabolite significance between treatments was determined with ANOVA.

3.6.6 Principal Component Analysis of Metabolites

Any missing values (0 abundance) in the metabolite abundances were replaced with half the lowest value for each specific compound. Values were log transformed and Pareto scaled. Principal components were calculated from the data with the prcomp function in the stats (v4.1.3) R package with arguments center and scale set to TRUE and FALSE respectively. Significant loadings were determined with principal component thresholds set to > 0.09 and < -0.09. Significant compounds (ANOVA) contribution to the PCA variance was calculated in Euclidian distance from the centroid. Bray-Curtis distances were calculated for the metabolite table with vegdist (R vegan package v2.6-2) and PERMANOVA analysis of treatment clusters was calculated with adonis2 (R vegan package v2.6-2).

3.6.7 16S rRNA Analysis

Amplicon sequencing fastq data generated for this experiment were processed in a QIIME2 2019.10.0 environment and reads were demultiplexed and then denoised with DADA2[74,75]. For all sequencing runs (n=6), forward reads were truncated at 246 bps and reverse reads were truncated at 167

bps. Feature tables from each sequencing run were combined and ASVs were later assigned taxonomy with the silva-138-99-515-806-nb-classifier in a QIIME2 2022.8.0 environment[74,76]. The resulting ASV table was filtered with R version 4.1.3 to remove any samples with zero counts for every ASV. Next, ASVs were removed with zero counts in every remaining sample, and subsequently samples with fewer than 1000 counts across all ASVs remaining ASVs were dropped. Finally, any ASVs designated as mitochondria, chloroplast, unassigned at the domain level, or assigned Eukaryota was removed. The resulting feature table contained 6308 unique ASVs. Median and mean *Salmonella* ASV abundances were calculated in R and the 16S V4 region from *Salmonella enterica* Typhimurium ATCC 14028 reference genome was manually aligned with Geneious Prime® 2020.1.2 to the *Salmonella* ASV with 100% sequence identity to determine confirm taxonomy. The designation "High responder" was applied to any mouse with *Salmonella* ASV relative abundance ≥ 25% in at least one sample.

3.6.8 ASV Community Metrics and Class Significance

ASV relative abundance, richness, and Shannon's Diversity were calculated in R version 4.1.3 and mapped to NMDS ordinations produced with the vegan package (v2.6-2). All filtered ASV data was divided into categories based on sampling day: Early (Day -6 – Day 2, n = 496), Middle (Day 3 – Day 8, n = 503), and Late (Day 9 – Day 23, n = 579) and metric significance was generated by the Mann-Whitney U test comparing each pair of groups independently. ASV communities from mice with metabolomics data and metatranscriptomics data (n=14) were categorized by sampling day: Early (Day -2 – Day 0) and Late (Day 10 – Day 12). HFD and chow NMDS ordination significance was determined with the mrpp and anosim functions from the vegan R package (v2.6-2). Average ASV abundance within each class was calculated and the 8 most abundant classes were retained and any other ASVs were classified as "Other". Significant differences between classes and treatments were calculated with the Mann-Whitney U test in R.

3.6.9 16S rRNA Linear Discriminant Analysis

The filtered ASV table was limited to samples (n=41) from mice that also had metabolomics or metatranscriptomic data including only samples from days 10, 11, and 12. ASVs with ambiguous genus designations were removed and then relative abundance of each ASV was calculated within samples. The table was collapsed to the genus level and linear discriminant analysis was performed with LEfSe to determine important taxa from each treatment[38].

3.6.10 ASV High Responder Correlation Network

First the filtered ASV feature table was collapsed to the genus level and features were removed with 5000 or fewer counts and ambiguously named genera were removed ("uncultured", "uncultured bacterium", "unidentified"). The genus table was further curated to include only data from high responder mice collected from Day 12 or later in the experiment. Any genus with 1000 or fewer counts from the subsequent table was also removed from further analysis. Spearman correlation between genera was performed with the Hmisc package (v.4.7-1) in R version 4.1.3. Significant ($p<0.05$) positive interactions from the resulting correlation matrix were retained. Any genera with a positive correlation coefficient with *Salmonella* were further considered along with any genera with $\boldsymbol{\rho} \geq 0.5$ correlation to the *Salmonella* correlated genera. The resulting correlation network was plotted with igraph (v.1.3.4) and ggraph (v.2.0.6) R packages.

3.6.11 Metagenome Assembled Genome Database

All metagenomic sequencing was quality checked, assembled, and binned as described in Leleiwi et. al.[5].The complete MAG set from the CBAJ-DB (https://doi.org/10.1101/2022.10.24.513540) and MAGs produced for this project were dereplicated with dRep (v2.6.2) at a 99% average nucleotide identity including only MAGs with at least 50% completeness and less than 10% contamination[77]. The dereplicated MAG set (n=160) was used for mapping metagenomic and metatranscriptomic reads. MAG taxonomy was assigned using GTDB-Tk (v.2.1.1) and gene annotations for each MAG in the database were produced with DRAM (v1.4.0). MAG quality was confirmed with CheckM (v.1.1.2).

3.6.12 Metagenomic Mapping

Raw paired-end metagenomic sequencing reads from *Salmonella*-infected mice used to create MAGs in the mapping database (n=7) were included with processed paired-end metagenomic reads from *Salmonella*-infected mice from an external study (n=8) (PRJNA491522) and mapped to the MAG database using Bowtie2 (v.2.4.5)[39]. Preprocessing of raw reads prior to mapping included quality trimming with Sickle (v.1.33) to remove sequences with phred quality scores below 20. Additionally, reads originating from CBA/J mice were also filtered with RQCFilter2 to remove adapter sequences, RNA artifacts, phiX kmers, rRNA sequences, and any sequences that map to mouse, cat, dog, human, or common microbial contaminate genomes. BAM files produced during mapping were sorted by sequence name and filtered to include only mappings with 95% identity with Samtools (v.1.9) and reformat.sh (https://github.com/BioInfoTools/BBMap/blob/master/sh/reformat.sh) respectively. Gene counts were produced for each MAG in the database with Featurecounts (v.1.5.3) and the resulting counts table was GeTMM normalized[78].

3.6.13 Sulfatase Identification

DRAM annotated amino acid sequences from the MAG database were mined for cysteine and serine sulfatase activation signatures - (C/S)*XPX*R and CXAXR[47]. The resulting gene set was further curated to include only genes with a sulfatase motif and Pfam Sulfatase annotation, either sulfatase PF00884.26, or the sulfatase modifying factor enzyme PF03781.19.

3.6.13 Association Rule Mining and Salmonella-associated Classification

Association rule mining (ARM) analysis was performed using custom functions available at https://github.com/ileleiwi/ARM_analysis. Calculations to determine feature associations from the GeTMM normalized MAG read counts are as follows:

$$Support(Taxa1 \rightarrow Taxa2) = \frac{Freq(Taxa1, Taxa2)}{N}$$

$$Confidence(Taxa1 \rightarrow Taxa2) = \frac{Freq(Taxa1, Taxa2)}{Freq(Taxa1)}$$

$$ExpectedConfidence(Taxa1 \rightarrow Taxa2) = \frac{Freq(Taxa2)}{N}$$

$$Lift(Taxa1 \rightarrow Taxa2) = \frac{Confidence}{ExpectedConfidence}$$

The GeTMM table was first converted to presence absence and then ARM metrics were calculated for association rules including 2 and 3 MAGs (n=153,648 individual rules) with each MAG in the place of consequent. MAGs that were consequent in rules where lift ≥ 1.5 and *Salmonella* was one of the antecedents were retained as *Salmonella*-associated (n=54). Next, all MAG (n=160) GTDB-Tk taxonomy was examined and only MAGs with red_value ≥ 0.99 or fastani_ani ≥ 0.99 were included in ASV taxonomy matching. Similarly, only ASV SILVA taxonomies with confidence ≥ 0.99 were considered for taxonomy matching to MAGs[76]. High confidence ASVs included in the genus-level correlation network were collapsed to the genus level (n=356) and high confidence MAG taxonomy strings were also aggregated at the genus level, then any MAGs with matching genus-level taxonomy to correlated amplicon taxonomy strings were retained as *Salmonella*-associated (n=42). Finally, MAG relative abundance was averaged across each mapped sample and the top 10 most relatively abundant MAGs were added to the *Salmonella*-associated group, in this way 6 additional bins were declared *Salmonella*-associated. The resulting *Salmonella*-associated MAG set from ARM analysis, correlation network taxonomy matching, and MAG abundance included 51 MAGs (including *Salmonella* MAG).

3.6.14 Logistic Regression LASSO model

Input for the LASSO logistic regression model included 56 predictors from the DRAM product columns produced from annotation of the dereplicated MAG database (n=160) after removing any variables with 0 or FALSE values for every observation and any Boolean variables with only one TRUE instance. The *Salmonella* MAG was also omitted from the analysis and the target variable was amended by changing the classification of the 10 least relatively abundant *Salmonella*-associated MAGs to not-

Salmonella-associated (negative class). Data were partitioned into training (75%) and test (25%) sets and a L1 penalty was employed. A series of 400 models with different random partitions were trained and their results compared. The minimum lambda value (lambda.min) each round was determined by 11-fold cross validation using cv.glmnet from the glmnet R package (v.4.1-6). The final best model was trained with the glmnet function from the same package. All cross validations and the final model were standardized to a mean of 0 and a standard deviation of 1 with the glmnet function parameter 'standardization=TRUE'. Classification threshold was set to 0.2515723. All model accuracy statistics and each model coefficients can be found in Supplemental data 2.

3.6.15 Random Forest Classification Model

Input for the random forest model was the same as that for the LASSO logistic regression model. Data were split into training (75%) and test (25%) sets and were center scaled with the caret R package (v.6.0-93). The model was optimized with a 10-fold cross validation tuning model parameters mtry, splitrule, and min.node.size and the best model was selected based on Cohen's Kappa. Cross validation and final model training were performed via ranger (v.0.14.1) and caret R packages. The best performing model was trained with a random sampling of 6 variables at each split using the extratrees random splitting function with a minimum node size of 5[79]. Final variable importance measures were calculated with the varImp function from the caret package where each predictor was permuted and the difference between model accuracies was averaged and normalized by the standard error for each tree.

3.6.16 Metatranscriptomic Sequencing Analysis and Mapping

Raw metatranscriptomic paired fastq data collected on day 11 from 5 CBA/J mice in each treatment were quality trimmed to remove reads shorter than 75 bps and of lower quality than phred 20 with bbduk.sh (https://github.com/BioInfoTools/BBMap/blob/master/sh/bbduk.sh). PhiX and rRNA sequences were removed with bbmap.sh (https://github.com/BioInfoTools/BBMap/blob/master/sh/bbmap.sh) at minid=0.90. The resulting reads

were then mapped to the MAG database using Bowtie2 (v.2.4.5) using flags -D 10 -R 2 -N 0 -L 22 -i S,0,2.50. Binary alignment files were filtered to only include mappings with high sequence identity (≥97%) with reformat.sh (https://github.com/BioInfoTools/BBMap/blob/master/sh/reformat.sh) and were sorted by sequence name and with Samtools (v.1.9). Individual gene counts were produced with Featurecounts (v.1.5.3) and the resulting counts table was GeTMM normalized[78]. The count table was further filtered before differential gene expression was analyzed, to first remove any genes with 0 counts in every sample, then to include only genes with at least 15 counts in 3 or more samples. Differential gene expression and significance of GeTMM normalized filtered transcript counts was produced with limma (v.3.50.3) and edgeR (v.3.36.0) R packages. DRAM gene annotation refinement allowed transcripts to be grouped into functional categories. Rule_adjectives (https://github.com/rmFlynn/rule_adjectives) was used to query a set of expertly curated genes (https://github.com/rmFlynn/rule_adjectives/blob/main/dram2/rule_adjectives/rules.tsv) for MAG substrate utilization and individual gene expression in MAGs of interest was totaled within each substrate category (MetaT_carbon_substrates_fig.R). To consider all possible gene expression, unfiltered GeTMM transformed transcript counts were used to produce Figures 8c, 9b, 9c, and 11.

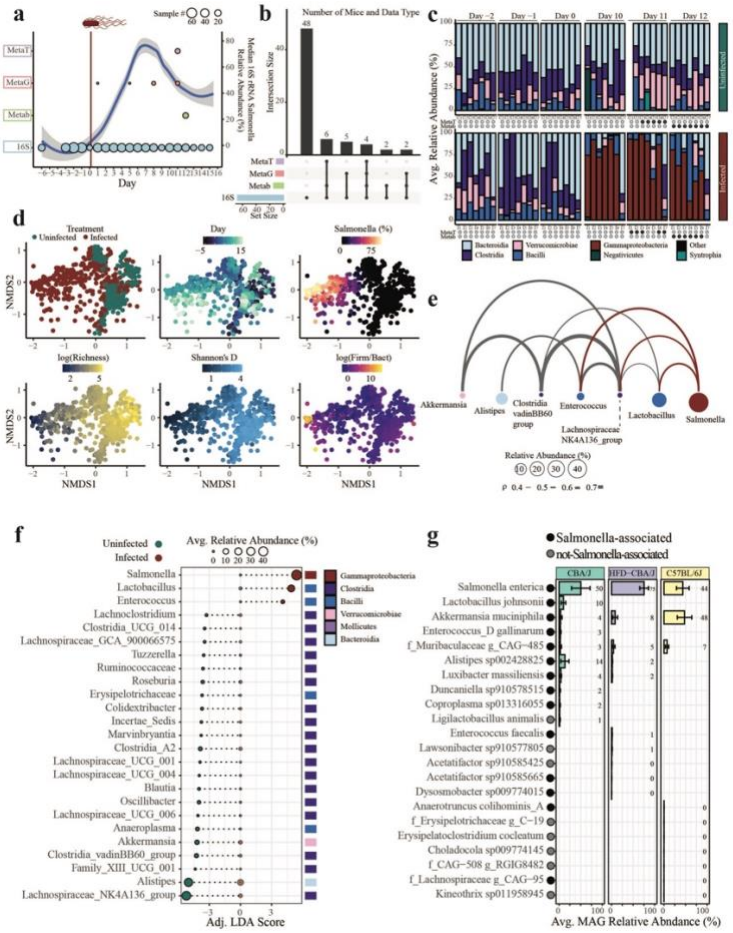


Fig. 12. *Salmonella* infection decreases gut diversity and enriches for *Bacilli* in amplicon and whole genome sequenced communities regardless of diet or mouse breed. (a) Experimental design showing high responder mice *Salmonella* amplicon sequencing variant (ASV) relative abundance (blue curve) and sample number (point size) at each timepoint colored by analysis type. *Salmonella* treatment start is indicated by the red vertical line on day 0. (b) Upset plot showing the number of samples in each analysis. (c) Stacked bars showing the ASV class distribution of mice that have metatranscriptomic and metabolomic data. Mouse U1 Day 11 sample was omitted via ASV table filtering (see methods) (d) Non-metric multidimensional scaling of Bray-Curtis distances calculated from ASV relative abundances showing *Salmonella* treatment, timepoint (Day), *Salmonella* relative abundance, community richness (log), alpha diversity (Shannon's index), and *Firmicutes*/*Bacteriodia* ratio (log). (e) Arc diagram depicting positive significant Spearman correlations between community genera and *Salmonella*. (f) Linear discriminant analysis of late timepoint samples (day 10-12) from mice with metatranscriptomic and metabolomic data. Points are sized by relative abundance of each genus within a treatment (infected or uninfected) and are colored by treatment. Points aligning with x-axis value 0 are the relative abundance of each genus in the non-significant treatment. Genera classes are listed by color on the right of the plot. (g) Boxplots depicting the mean relative abundance of mapped metagenomic reads from either CBA/J chow fed mice (green), CBA/J high fat diet fed mice (purple), or C57BL/6 mice to the metagenome assembled genome (MAG) database. Relative abundances are summed at the lowest meaningful taxonomy designation. Colored circles on the y-axis indicate whether each taxonomy is *Salmonella*-associated (black) or not (gray) as determined by Spearman correlation and association rule mining or MAG relative abundance (see methods).

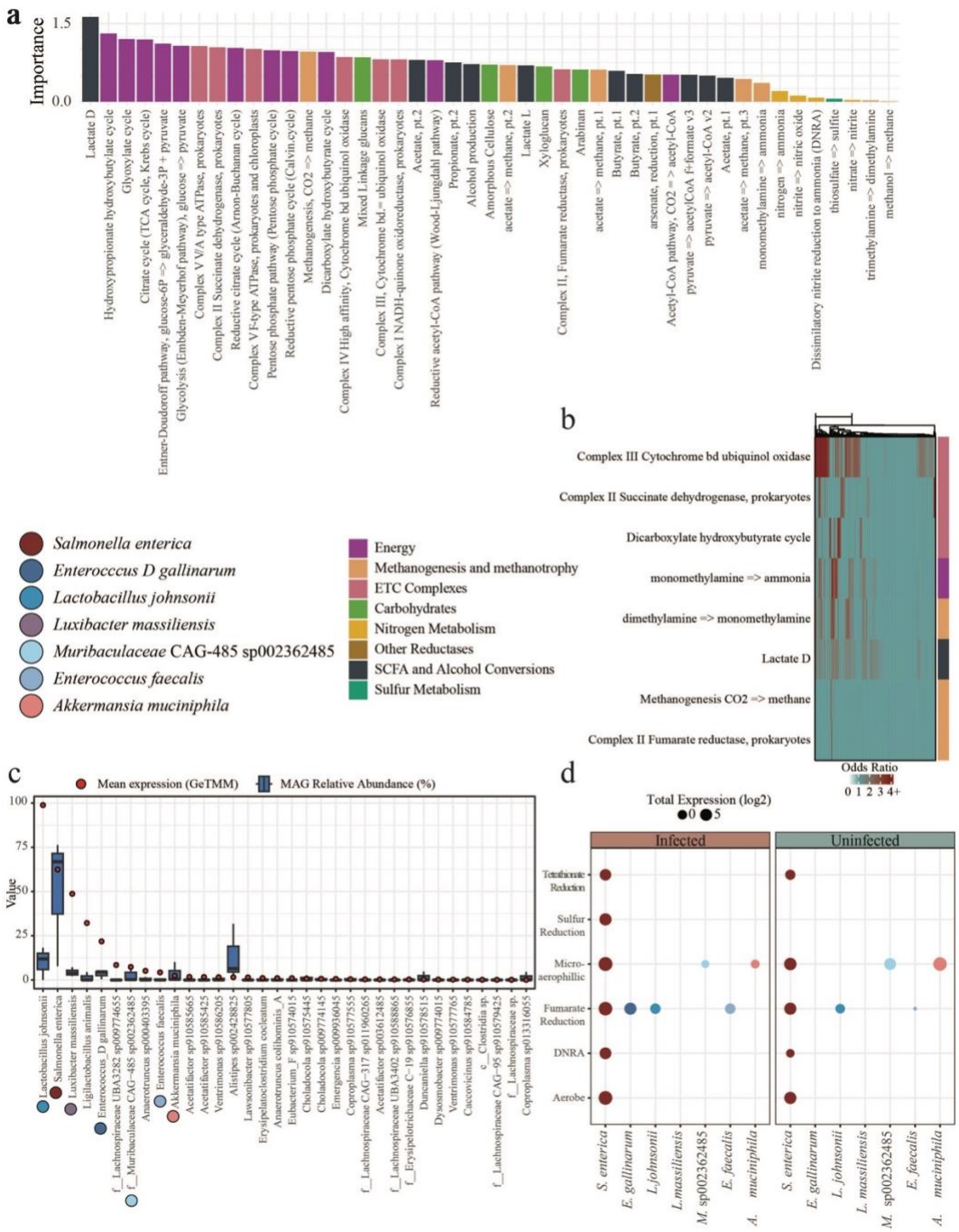


Fig. 13. Predictive gene categories of *Salmonella* co-occurrence and gene expression of prominent bacteria during infection reveal persistent taxa and genetic potential in the infected gut. (a) Important predictive variables for *Salmonella*-associated bacteria as determined by random forest modeling. (b) Most important exponentiated coefficient values from 400 logistic regression models to predict *Salmonella*-association from community gene content. (c) Box plots (blue) showing the MAG relative abundance (GeTMM normalized mapped reads from CBA/J chow mice, n=3) of the 35 bacteria with the highest average expression (red points) (CBA/J chow mice, n=5) in *Salmonella*-infected mice. (d) Point size indicates the total mean expression (log2) of genes contributing to a given substrate utilization in either infected or uninfected treatments.

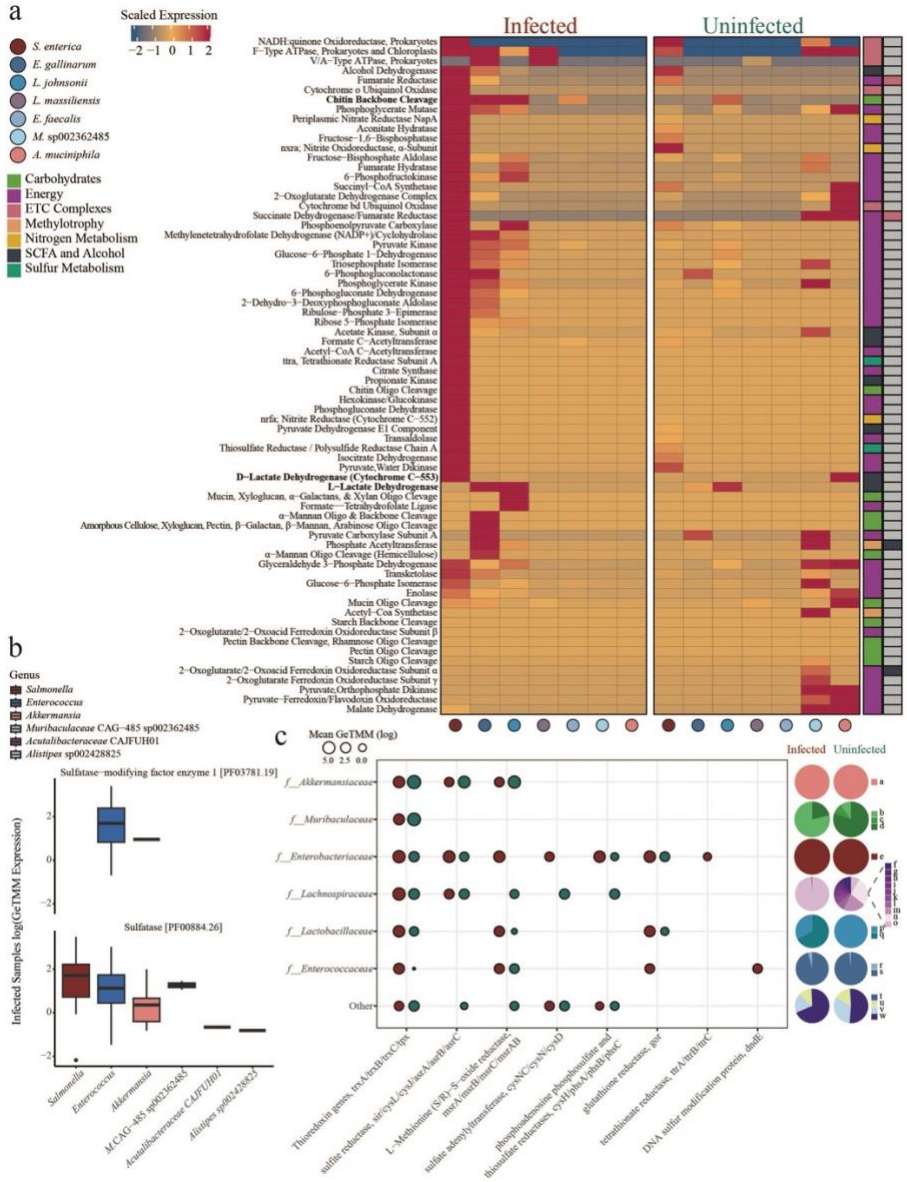


Fig. 14. Differential gene expression provides evidence for function overlap with *Salmonella* and commensal niche divergence. (a) Heatmap of actively expressed genes during infection that are also significantly differentially expressed between treatments. Cell values are GeTMM scaled totals of all genes linked with a particular gene description (y-axis) averaged across samples within a treatment and expressed by an individual taxon (x-axis). Colored boxes on the right indicated gene class. Gene ID's included in each gene description may fall into multiple classes as indicated by the second row of boxes. (b) Boxplots depicting important taxa expression of sulfatase (bottom) and sulfatase activating enzyme (top). (c) Dot plot depicting average log expression (GeTMM) of genes involved in sulfur reduction and oxidative stress response. Families of important taxa are listed on the y-axis and pie charts indicate the proportion of expression attributed to individual genera in each family. Specific genera indicated by lowercase letter for each family and specific classes for the Other category are as follows: a. *Akkermansia*, b. CAG-485, c. CAG-873, d. *Duncaniella*, e. *Salmonella*, f. Other, g. UBA34021, h. *Caccovicinus*, i. MD308 COE1, j. MGBC164599, k. UMGS1370, l. *Ventrimonas*, m. 14–2, n. *Sporofaciens*, o. *Luxibacter*, p. *Lactobacillus*, q. *Ligilactobacillus*, r. *Enterococcus*, s. *Enterococcus D*, t. *Bacilli*, u. *Coriobacteriia*, v. *Bacteroidia*, w. *Clostridia*.

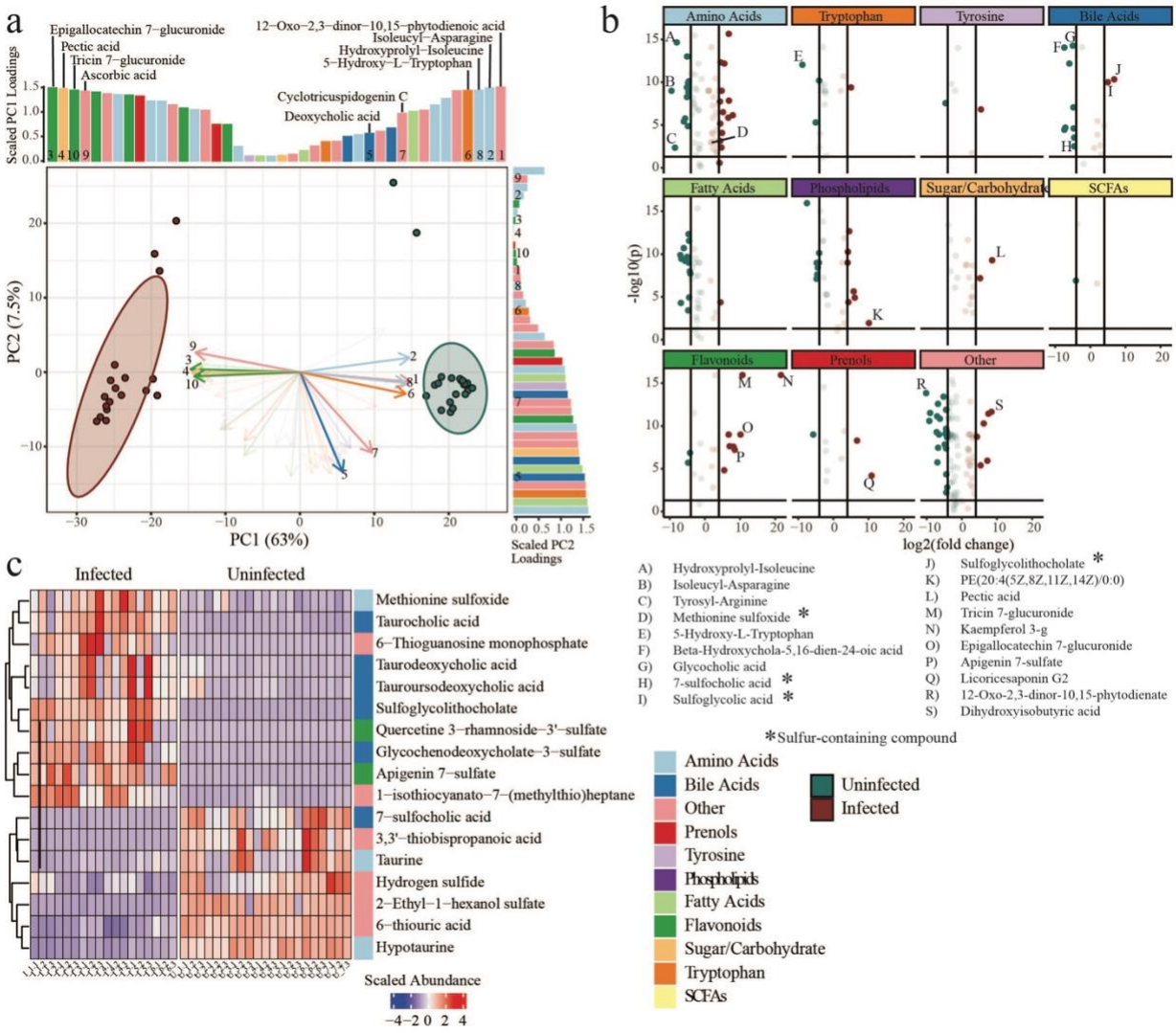


Fig. 15. Uninfected metabolomes differ from inflamed microbiomes where sulfonated bile acids and flavonoids are prevalent. (a) Principal component analysis (PCA) of metabolites from infected and uninfected mice colored by treatment (red = infected, green = uninfected). Arrows and bars are significant (ANOVA $p < 0.05$) compounds. Dark arrows are the top 10 compounds that explain the PCA ordination variance (Euclidian distance from tip to centroid). Bars are arranged by absolute value of compound loading for each component, and they are colored by compound group. (b) Volcano plots of significant metabolites separated by compound group. Points are colored by treatment (red = infected, green = uninfected) and dark points indicate significant (ANOVA $p < 0.05$) compounds with at least 4x log2 fold

change between treatments. Select compounds (most changed or system relevant) are labeled below the plot and sulfur containing compounds are denoted with an asterisk. (c) Heatmap showing the center scaled abundance of significantly different sulfur containing compounds in each treatment.

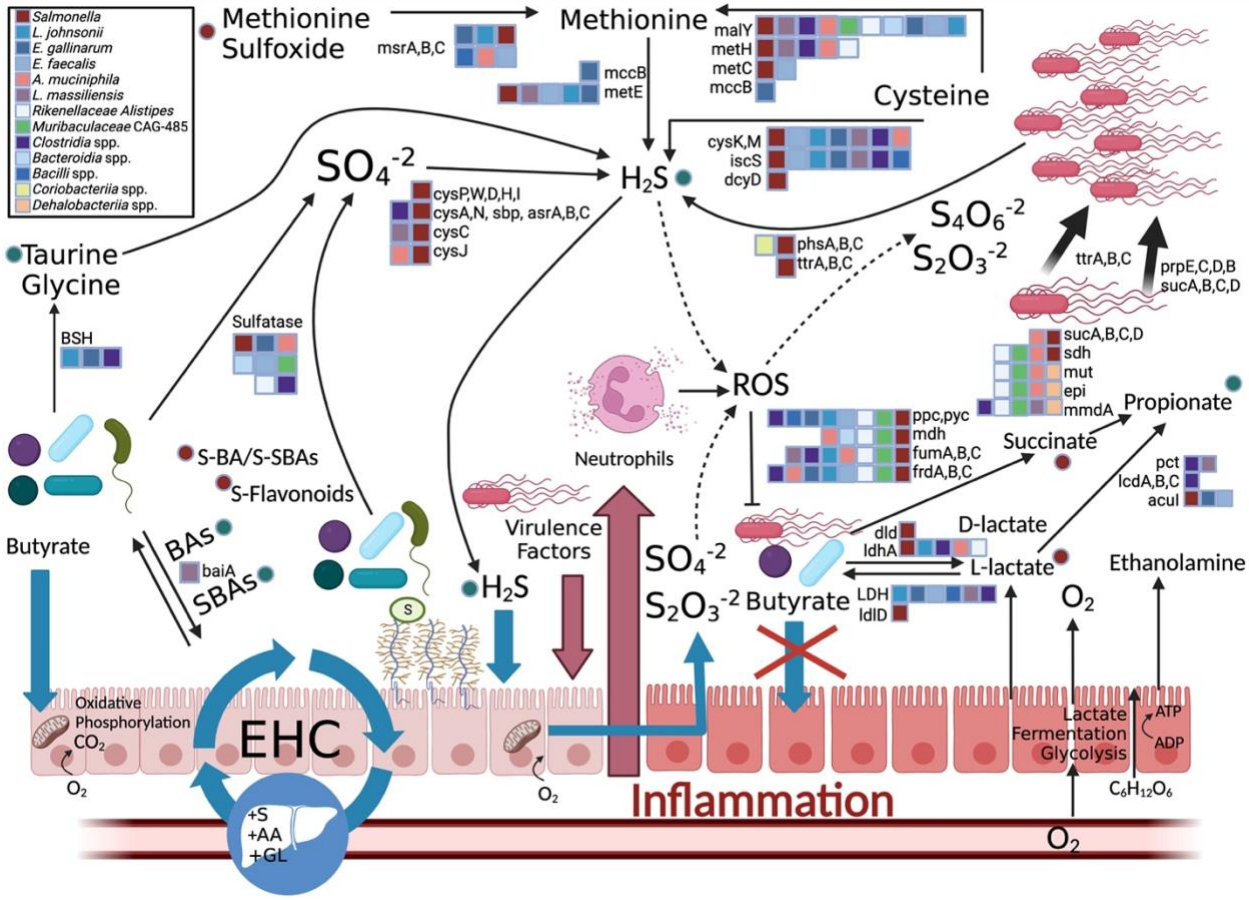


Fig. 16 Inflammation and microbial activity make sulfur and carbon from the host available to *Salmonella*. Colored circles indicate metabolite abundance was higher in either uninfected (green) or infected (red) samples. Colored squares indicate individual taxa active gene expression during infection. Dotted arrows indicate potential sulfur oxidation. Compound abbreviations are as follows: bile acids (BAs), secondary bile acids (SBAs), sulfonated bile acids (S-BA), sulfonated secondary bile acids (S-SBAs), sulfonated flavonoids (S-Flavonoids), reactive oxygen species (ROS), sulfonate (S), amino acids (AA), gluconate (GL).

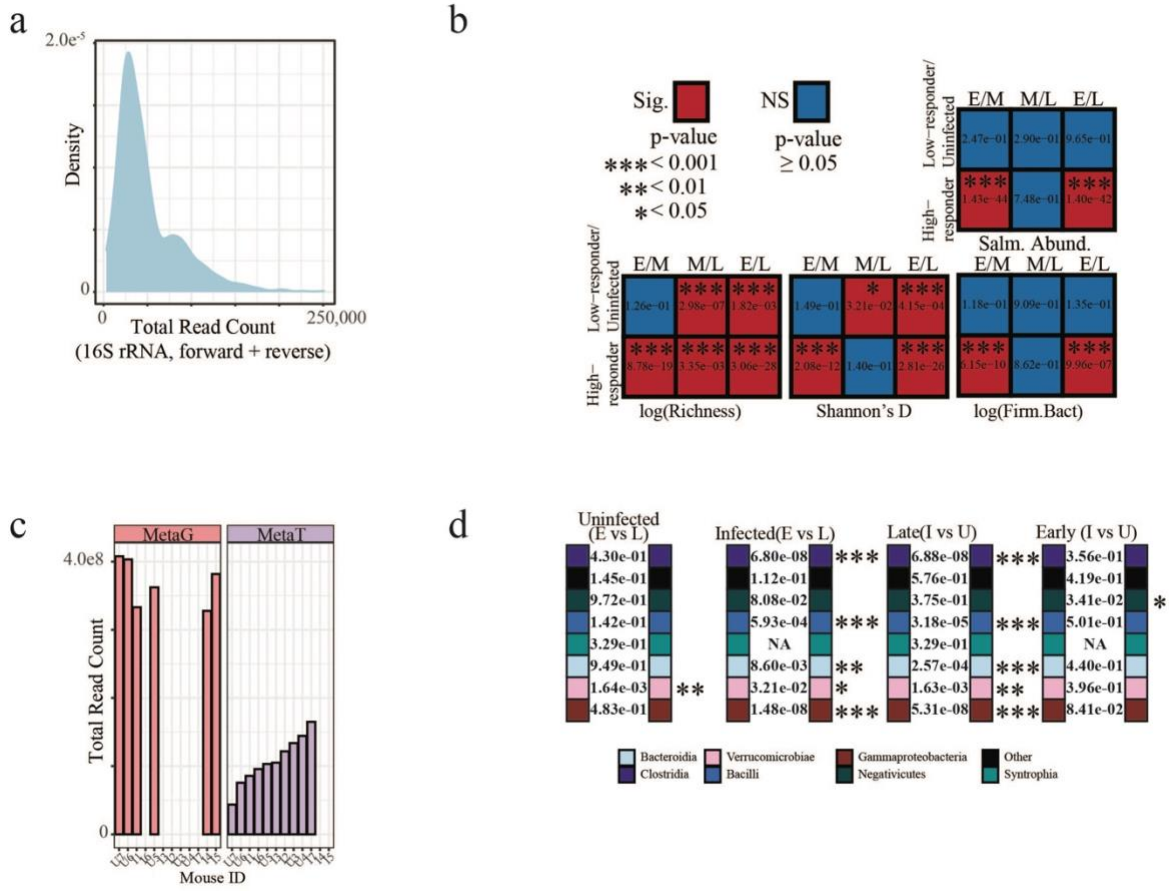


Fig. 17. Sequencing effort and statistical analysis of microbial communities. (a) Distribution of total 16S rRNA reads from all samples. (b) Mann Whitney U significance between Early (E, days -6 – 2, n=293), Middle (M, days 3 – 8, n=241), and Late (L, days 9 – 15, n=178) for the different metrics, either *Salmonella* relative abundance (Salm. Abund., top right), ASV richness (bottom left), beta diversity (Shannon's D, bottom middle), or *Firmicutes/Bacteriodia* ratio (F/B, bottom right). (c) Metagenomic and metatranscriptomic reads counts. (d) Significant differences (Wilcoxon Rank Sum) between classes of either Early timepoints (E, days -2–0), late time points (L, days 10–12), infected samples, or uninfected samples.

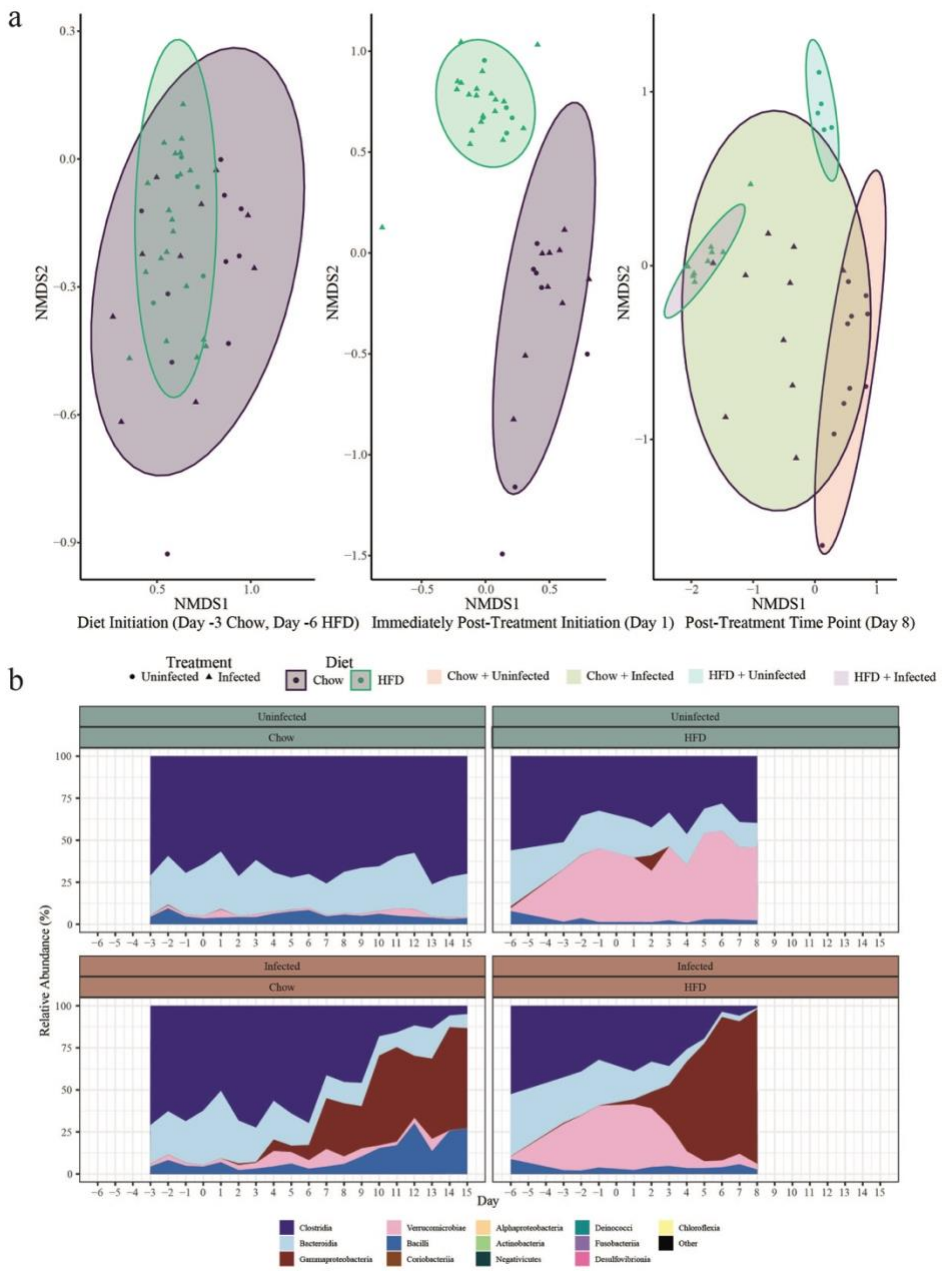


Fig. 18. Diet effect on amplicon sequenced community structure. (a) NMDS ordination of Bray-Curtis distances from 16S rRNA amplicon sequenced communities in mice fed chow (purple) or HFD (green) and sampled on either day -3 and day -6 (left), day 1 (middle), or day 8 (right). Triangle points designate communities from mice in the infected treatment and circle points are mice in the uninfected treatment. (b) Density plots depicting the class distribution of bacteria in mice fed either chow (left) or HFD (right) from infected treatment mice (bottom) or uninfected mice (top).

Chapter 3 References

1. Plumb ID, Brown AC, Stokes EK, Chen JC, Carleton H, Tolar B, Sundararaman P, Saupe A, Payne DC, Shah HJ, et al. Increased Multidrug-Resistant Salmonella enterica I Serotype 4,[5],12:i:- Infections Associated with Pork, United States, 2009–2018 - Volume 29, Number 2—February 2023 - Emerging Infectious Diseases journal - CDC. [cited 2023 Apr 3]; Available from: https://wwwnc.cdc.gov/eid/article/29/2/22-0950_article

2. Salmonella By the Numbers | Food Safety and Inspection Service [Internet]. [cited 2023 Apr 3]; Available from: http://www.fsis.usda.gov/inspection/inspection-programs/inspection-poultry-products/reducing-salmonella-poultry/salmonella

3. Salmonella (non-typhoidal) [Internet]. [cited 2021 Feb 3]; Available from: https://www.who.int/news-room/fact-sheets/detail/salmonella-(non-typhoidal)

4. Jajere SM. A review of Salmonella enterica with particular focus on the pathogenicity and virulence factors, host specificity and antimicrobial resistance including multidrug resistance. Vet World 2019; 12:504–21.

5. Leleiwi I, Rodriguez-Ramos J, Shaffer M, Sabag-Daigle A, Kokkinias K, Flynn RM, Daly RA, Kop LF, Solden LM, Ahmer BMM, et al. Exposing New Taxonomic Variation with Inflammation – A Murine Model-Specific Genome Database for Gut Microbiome Researchers [Internet]. 2022 [cited 2023 Feb 7]; :2022.10.24.513540. Available from: https://www.biorxiv.org/content/10.1101/2022.10.24.513540v1

6. Borton MA, Sabag-Daigle A, Wu J, Solden LM, O'Banion BS, Daly RA, Wolfe RA, Gonzalez JF, Wysocki VH, Ahmer BMM, et al. Chemical and pathogen-induced inflammation disrupt the murine intestinal microbiome. Microbiome [Internet] 2017 [cited 2019 Aug 26]; 5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5408407/

7. Shelton CD, Yoo W, Shealy NG, Torres TP, Zieba JK, Calcutt MW, Foegeding NJ, Kim D, Kim J, Ryu S, et al. Salmonella enterica serovar Typhimurium uses anaerobic respiration to overcome propionate-mediated colonization resistance. Cell Reports [Internet] 2022 [cited 2022 Jul 14]; 38. Available from: https://www.cell.com/cell-reports/abstract/S2211-1247(21)01680-6

8. Faber F, Thiennimitr P, Spiga L, Byndloss MX, Litvak Y, Lawhon S, Andrews-Polymenis HL, Winter SE, Bäumler AJ. Respiration of Microbiota-Derived 1,2-propanediol Drives Salmonella Expansion during Colitis. PLOS Pathogens 2017; 13:e1006129.

9. Ahmer BM, Gunn JS. Interaction of Salmonella spp. with the Intestinal Microbiota. Front Microbiol [Internet] 2011 [cited 2020 Apr 16]; 2. Available from: https://www.frontiersin.org/articles/10.3389/fmicb.2011.00101/full

10. Coburn B, Li Y, Owen D, Vallance BA, Finlay BB. Salmonella enterica Serovar Typhimurium Pathogenicity Island 2 Is Necessary for Complete Virulence in a Mouse Model of Infectious Enterocolitis. Infect Immun 2005; 73:3219–27.

11. Rogers AWL, Tsolis RM, Bäumler AJ. Salmonella versus the Microbiome. Microbiol Mol Biol Rev 2021; 85.

12. Spiga L, Winter MG, de Carvalho TF, Zhu W, Hughes ER, Gillis CC, Behrendt CL, Kim J, Chessa D, Andrews-Polymenis HL, et al. An oxidative central metabolism enables Salmonella to utilize microbiota-derived succinate. Cell Host Microbe 2017; 22:291-301.e6.

13. Stecher B, Jung K. LACTATEing Salmonella: A Host-Derived Fermentation Product Fuels Pathogen Growth. Cell Host & Microbe 2018; 23:3–4.

14. Roof DM, Roth JR. Ethanolamine utilization in Salmonella typhimurium. J Bacteriol 1988; 170:3855–63.

15. Hensel M, Hinsley AP, Nikolaus T, Sawers G, Berks BC. The genetic basis of tetrathionate respiration in Salmonella typhimurium. Molecular Microbiology 1999; 32:275–87.

16. Winter SE, Thiennimitr P, Winter MG, Butler BP, Huseby DL, Crawford RW, Russell JM, Bevins CL, Adams LG, Tsolis RM, et al. Gut inflammation provides a respiratory electron acceptor for Salmonella. Nature 2010; 467:426–9.

17. Stoffels L, Krehenbrink M, Berks BC, Unden G. Thiosulfate Reduction in Salmonella enterica Is Driven by the Proton Motive Force. J Bacteriol 2012; 194:475–85.

18. Dordević D, Jančíková S, Vítězová M, Kushkevych I. Hydrogen sulfide toxicity in the gut environment: Meta-analysis of sulfate-reducing and lactic acid bacteria in inflammatory processes. Journal of Advanced Research 2021; 27:55–69.

19. Gibson GR, Cummings JH, Macfarlane GT. Growth and activities of sulphate-reducing bacteria in gut contents of healthy subjects and patients with ulcerative colitis. FEMS Microbiology Letters 1991; 86:103–12.

20. Vitetta L, Coulson S, Thomsen M, Nguyen T, Hall S. Probiotics, D–Lactic acidosis, oxidative stress and strain specificity. Gut Microbes 2017; 8:311–22.

21. Rhen M. Salmonella and Reactive Oxygen Species: A Love-Hate Relationship. JIN 2019; 11:216–26.

22. Gillis CC, Hughes ER, Spiga L, Winter MG, Zhu W, Furtado de Carvalho T, Chanin RB, Behrendt CL, Hooper LV, Santos RL, et al. Dysbiosis-Associated Change in Host Metabolism Generates Lactate to Support Salmonella Growth. Cell Host & Microbe 2018; 23:54-64.e6.

23. Aljahdali NH, Sanad YM, Han J, Foley SL. Current knowledge and perspectives of potential impacts of Salmonella enterica on the profile of the gut microbiota. BMC Microbiology 2020; 20:353.

24. Rivera-Chávez F, Zhang LF, Faber F, Lopez CA, Byndloss MX, Olsan EE, Xu G, Velazquez EM, Lebrilla CB, Winter SE, et al. Depletion of butyrate-producing Clostridia from the gut microbiota drives an aerobic luminal expansion of Salmonella. Cell Host Microbe 2016; 19:443–54.

25. Wong JMW, de Souza R, Kendall CWC, Emam A, Jenkins DJA. Colonic Health: Fermentation and Short Chain Fatty Acids: Journal of Clinical Gastroenterology 2006; 40:235–43.

26. Ducarmon QR, Zwittink RD, Hornung BVH, van Schaik W, Young VB, Kuijper EJ. Gut Microbiota and Colonization Resistance against Bacterial Enteric Infection. Microbiology and Molecular Biology Reviews 2019; 83:e00007-19.

27. Ouyang J, Lin J, Isnard S, Fombuena B, Peng X, Marette A, Routy B, Messaoudene M, Chen Y, Routy J-P. The Bacterium Akkermansia muciniphila: A Sentinel for Gut Permeability and Its Relevance to HIV-Related Inflammation. Front Immunol 2020; 11:645.

28. Jacobson A, Lam L, Rajendram M, Tamburini F, Honeycutt J, Pham T, Treuren WV, Pruss K, Stabler SR, Lugo K, et al. A Gut Commensal-Produced Metabolite Mediates Colonization Resistance to Salmonella Infection. Cell Host & Microbe 2018; 24:296-307.e7.

29. Kim JY, Young JA, Gunther IV NW, Lee J-L. Inhibition of Salmonella by Bacteriocin-Producing Lactic Acid Bacteria Derived from U.S. Kimchi and Broiler Chicken. Journal of Food Safety 2015; 35:1–12.

30. Zamfir M, Callewaert R, Cornea PC, Savu L, Vatafu I, De Vuyst L. Purification and characterization of a bacteriocin produced by Lactobacillus acidophilus IBB 801. Journal of Applied Microbiology 1999; 87:923–31.

31. Piazentin ACM, Mendonça CMN, Vallejo M, Mussatto SI, de Souza Oliveira RP. Bacteriocin-like inhibitory substances production by Enterococcus faecium 135 in co-culture with Ligilactobacillus salivarius and Limosilactobacillus reuteri. Braz J Microbiol 2022; 53:131–41.

32. Ridlon JM, Harris SC, Bhowmik S, Kang D-J, Hylemon PB. Consequences of bile salt biotransformations by intestinal bacteria. Gut Microbes 2016; 7:22–39.

33. Yang X, Stein KR, Hang HC. Anti-infective bile acids bind and inactivate a Salmonella virulence regulator. Nat Chem Biol 2023; 19:91–100.

34. Prouty AM, Gunn JS. Salmonella enterica serovar typhimurium invasion is repressed in the presence of bile. Infect Immun 2000; 68:6763–9.

35. Eade CR, Hung C-C, Bullard B, Gonzalez-Escobedo G, Gunn JS, Altier C. Bile Acids Function Synergistically To Repress Invasion Gene Expression in Salmonella by Destabilizing the Invasion Regulator HilD. Infection and Immunity 2016; 84:2198–208.

36. Lou L, Zhang P, Piao R, Wang Y. Salmonella Pathogenicity Island 1 (SPI-1) and Its Complex Regulatory Network. Frontiers in Cellular and Infection Microbiology [Internet] 2019 [cited 2023 Apr 4]; 9. Available from: https://www.frontiersin.org/articles/10.3389/fcimb.2019.00270

37. Hung C-C, Garner CD, Slauch JM, Dwyer ZW, Lawhon SD, Frye JG, McClelland M, Ahmer BMM, Altier C. The Intestinal Fatty Acid Propionate Inhibits Salmonella Invasion through the Post-translational Control of HilD. Mol Microbiol 2013; 87:1045–60.

38. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. Metagenomic biomarker discovery and explanation. Genome Biol 2011; 12:R60.

39. Bratburd JR, Keller C, Vivas E, Gemperline E, Li L, Rey FE, Currie CR. Gut Microbial and Metabolic Responses to Salmonella enterica Serovar Typhimurium and Candida albicans. mBio 2018; 9:e02032-18.

40. Gillis CC, Winter MG, Chanin RB, Zhu W, Spiga L, Winter SE. Host-Derived Metabolites Modulate Transcription of Salmonella Genes Involved in l-Lactate Utilization during Gut Colonization. Infect Immun 2019; 87:e00773-18.

41. Pecsi I, Hards K, Ekanayaka N, Berney M, Hartman T, Jacobs WR, Cook GM. Essentiality of Succinate Dehydrogenase in Mycobacterium smegmatis and Its Role in the Generation of the Membrane Potential Under Hypoxia. mBio 2014; 5:e01093-14.

42. Kushkevych I, Cejnar J, Treml J, Dordević D, Kollar P, Vítězová M. Recent Advances in Metabolic Pathways of Sulfate Reduction in Intestinal Bacteria. Cells 2020; 9:698.

43. Wolf PG, Cowley ES, Breister A, Matatov S, Lucio L, Polak P, Ridlon JM, Gaskins HR, Anantharaman K. Diversity and distribution of sulfur metabolic genes in the human gut microbiome and their association with colorectal cancer. Microbiome 2022; 10:64.

44. Hinsley AP, Berks BC. Specificity of respiratory pathways involved in the reduction of sulfur compounds by Salmonella enterica. Microbiology 2002; 148:3631–8.

45. Furne J, Springfield J, Koenig T, DeMaster E, Levitt MD. Oxidation of hydrogen sulfide and methanethiol to thiosulfate by rat tissues: a specialized function of the colonic mucosa. Biochemical Pharmacology 2001; 62:255–9.

46. Hildebrandt TM, Grieshaber MK. Three enzymatic activities catalyze the oxidation of sulfide to thiosulfate in mammalian and invertebrate mitochondria. The FEBS Journal 2008; 275:3352–61.

47. Berteau O, Guillot A, Benjdia A, Rabot S. A New Type of Bacterial Sulfatase Reveals a Novel Maturation Pathway in Prokaryotes *. Journal of Biological Chemistry 2006; 281:22464–70.

48. van der Ploeg J, Eichhorn E, Leisinger T. Sulfonate-sulfur metabolism and its regulation in Escherichia coli. Archives of Microbiology 2001; 176:1–8.

49. Singh VK, Singh K, Baum K. The Role of Methionine Sulfoxide Reductases in Oxidative Stress Tolerance and Virulence of Staphylococcus aureus and Other Bacteria. Antioxidants (Basel) 2018; 7:128.

50. Byndloss MX, Olsan EE, Rivera-Chávez F, Tiffany CR, Cevallos SA, Lokken KL, Torres TP, Byndloss AJ, Faber F, Gao Y, et al. Microbiota-activated PPAR-γ-signaling inhibits dysbiotic Enterobacteriaceae expansion. Science 2017; 357:570–5.

51. Sheridan PO, Louis P, Tsompanidou E, Shaw S, Harmsen HJ, Duncan SH, Flint HJ, Walker AW. Distribution, organization and expression of genes concerned with anaerobic lactate utilization in human intestinal bacteria. Microbial Genomics 2022; 8:000739.

52. Kasai T, Suzuki Y, Kouzuma A, Watanabe K. Roles of d-Lactate Dehydrogenases in the Anaerobic Growth of Shewanella oneidensis MR-1 on Sugars. Appl Environ Microbiol 2019; 85:e02668-18.

53. Pohanka M. D-Lactic Acid as a Metabolite: Toxicology, Diagnosis, and Detection. Biomed Res Int 2020; 2020:3419034.

54. Reichardt N, Duncan SH, Young P, Belenguer A, McWilliam Leitch C, Scott KP, Flint HJ, Louis P. Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. ISME J 2014; 8:1323–35.

55. Zhang X, Jantama K, Moore JC, Jarboe LR, Shanmugam KT, Ingram LO. Metabolic evolution of energy-conserving pathways for succinate production in Escherichia coli. Proceedings of the National Academy of Sciences 2009; 106:20180–5.

56. Yang T, Mbadinga SM, Zhou L, Yang S-Z, Liu J-F, Gu J-D, Mu B-Z. Propionate metabolism and diversity of relevant functional genes by in silico analysis and detection in subsurface petroleum reservoirs. World J Microbiol Biotechnol 2017; 33:182.

57. Murphy MG, O'Connor L, Walsh D, Condon S. Oxygen dependent lactate utilization by Lactobacillus plantarum. Arch Microbiol 1985; 141:75–9.

58. Belenguer A, Duncan SH, Holtrop G, Flint HJ, Lobley GE. Quantitative Analysis of Microbial Metabolism in the Human Large Intestine. Current Nutrition & Food Science 4:109–26.

59. Llibre A, Grudzinska FS, O'Shea MK, Duffy D, Thickett DR, Mauro C, Scott A. Lactate cross-talk in host–pathogen interactions. Biochemical Journal 2021; 478:3157–78.

60. Wang SP, Rubio LA, Duncan SH, Donachie GE, Holtrop G, Lo G, Farquharson FM, Wagner J, Parkhill J, Louis P, et al. Pivotal Roles for pH, Lactate, and Lactate-Utilizing Bacteria in the Stability of a Human Colonic Microbial Ecosystem. mSystems 2020; 5:e00645-20.

61. Connors J, Dawe N, Van Limbergen J. The Role of Succinate in the Regulation of Intestinal Inflammation. Nutrients 2018; 11:25.

62. Alnouti Y. Bile Acid Sulfation: A Pathway of Bile Acid Elimination and Detoxification. Toxicological Sciences 2009; 108:225–46.

63. Chen L, Cao H, Huang Q, Xiao J, Teng H. Absorption, metabolism and bioavailability of flavonoids: a review. Critical Reviews in Food Science and Nutrition 2022; 62:7730–42.

64. Huijghebaert S, Parmentier G, Eyssen H. Specificity of bile salt sulfatase activity in man, mouse and rat intestinal microflora. Journal of Steroid Biochemistry 1984; 20:907–12.

65. Hervert-Hernández D, Goñi I. Dietary Polyphenols and Human Gut Microbiota: a Review. Food Reviews International 2011; 27:154–69.

66. Kumar S, Pandey AK. Chemistry and Biological Activities of Flavonoids: An Overview. ScientificWorldJournal 2013; 2013:162750.

67. Kashyap P, Shikha D, Thakur M, Aneja A. Functionality of apigenin as a potent antioxidant with emphasis on bioavailability, metabolism, action mechanism and in vitro and in vivo studies: A review. Journal of Food Biochemistry 2022; 46:e13950.

68. Human Gut Microbiome and Quercetin [Internet]. [cited 2023 Mar 27]; Available from: https://encyclopedia.pub/entry/11806

69. Funabashi M, Grove TL, Wang M, Varma Y, McFadden ME, Brown LC, Guo C, Higginbottom S, Almo SC, Fischbach MA. A metabolic pathway for bile acid dehydroxylation by the gut microbiome. Nature 2020; 582:566–70.

70. Feng T, Wang J. Oxidative stress tolerance and antioxidant capacity of lactic acid bacteria as probiotic: a systematic review. Gut Microbes 12:1801944.

71. Gabrielsen M, Beckham KSH, Feher VA, Zetterström CE, Wang D, Müller S, Elofsson M, Amaro RE, Byron O, Roe AJ. Structural Characterisation of Tpx from Yersinia pseudotuberculosis Reveals Insights into the Binding of Salicylidene Acylhydrazide Compounds. PLOS ONE 2012; 7:e32217.

72. Pophaly SD, Singh R, Pophaly SD, Kaushik JK, Tomar SK. Current status and emerging role of glutathione in food grade lactic acid bacteria. Microbial Cell Factories 2012; 11:114.

73. Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, Gatto L, Fischer B, Pratt B, Egertson J, et al. A cross-platform toolkit for mass spectrometry and proteomics. Nat Biotechnol 2012; 30:918–20.

74. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 2019; 37:852–7.

75. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High resolution sample inference from Illumina amplicon data. Nat Methods 2016; 13:581–3.

76. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research 2013; 41:D590–6.

77. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloe-Fadrosh EA, et al. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat Biotechnol 2017; 35:725–31.

78. Smid M, Coebergh van den Braak RRJ, van de Werken HJG, van Riet J, van Galen A, de Weerd V, van der Vlugt-Daane M, Bril SI, Lalmahomed ZS, Kloosterman WP, et al. Gene length corrected trimmed mean of M-values (GeTMM) processing of RNA-seq data performs similarly in intersample analyses while improving intrasample comparisons. BMC Bioinformatics 2018; 19:236.

79. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. Mach Learn 2006; 63:3–42.

Chapter 4 Conclusion

4.1 Summary

My dissertation work employed various omics techniques to explore *Salmonella* infection in a robust microbiome and to consider the impact of persistent and relatively abundant members of the gut community on the *Salmonella* pathogenic process. Using metagenomics and amplicon sequencing techniques I reconstructed a comprehensive sampling of CBA/J mouse gut bacteria and viruses to show *Salmonella* restructuring of the gut community while also providing a genome resource for microbiome researchers (Chapter 2). I displayed how *Salmonella* alters the functional potential of the gut microbiome, limiting its genetic capacity to modulate host anti-inflammatory processes and opening the possibility of phage regulation of certain mechanisms in *Firmicutes* during *Salmonella* infection and inflammation (Chapter 2). The CBAJ-DB identifies bacteria with implications to human health, like *Akkermansia muciniphila* and *Enterococcus D gallinarum*, in the inflamed CBA/J gut, ultimately expanding the relevance of this work and the work presented in Chapter 3.

Next, I demonstrated the utility of the CBAJ-DB to recruit metatranscriptomics and metabolomics data, providing a genome-resolved view of *Salmonella* infection and illuminating interactions between *Salmonella* and prominent inflammation-resistant bacteria in the gut community (Chapter 3). I was able to predict *Salmonella* co-occurrence based on annotated genome content from the gut community and identify important functional capacity to bacteria persistence with *Salmonella* during the later stages of infection (Chapter 3). Highly expressive and relatively abundant lactic acid bacteria including *Lactobacillus johnsonii* and *Enterococcus D gallinarum* play roles in substrate availability for *Salmonella*, particularly regarding bile acids and lactate (Chapter 3). The current paradigm of *Salmonella* lactate enantiomer preference was challenged in our data, implying a more important role for other bacteria in *Salmonella* lactate utilization than was previously realized (Chapter 3). Additionally, sulfur metabolism in the inflamed gut was explored to suggest multiple avenues for host-derived and diet-derived sulfur to be liberated and oxidized in support of *Salmonella* anaerobic respiration of tetrathionate,

suggesting a more intricate involvement of the microbiota and host processes in *Salmonella* sulfur metabolism than implied by prior reports (Chapter 3). The multi-omics study of *Salmonella* infection in CBA/J mice indicated a continual reduction of sulfur regardless of oxidation state, underpinning the importance of sulfur redox regulation to host-associated microbiome resiliency in the face of immune activation and pathogen perturbation (Chapter 3).

4.2 Future Research Directions

Salmonella poses a continuous global health threat, particularly as multi-drug resistant strains increase in prevelance[1,2]. With the need for alternative treatment for salmonellosis clear, probiotic bacteria offer a compelling avenue for further research[3]. Various lactic acid bacteria including *Lactobacillus* spp. have been used probiotically and therapeutically to stave *Salmonella* infection and modulate host immunity or directly compete with the pathogen[3]. Several persistent community members identified in this dissertation work offer compelling candidates for future probiotic exploration. Isolation of the *Lactobacillus johnsonii* strain identified in the CBAJ-DB and then co-culturing with *Salmonella* to determine competitive potential is one possibility for future work. *Akkermansia muciniphila* is another potential candidate for probiotic trials identified here. *Akkermansia muciniphila* is consistently persistent in *Salmonella*-included microbiomes, and it appeared ubiquitous in all the human databases and mouse strains tested in the entirety of this work. *Akkermansia muciniphila* is known to add to host gut barrier integrity and to modulate the adaptive immune response[4–6]. Isolation and subsequent competition experiments with *Salmonella* are some of the next steps in identifying its probiotic potential prior to the, perhaps prohibitive, process of approval for therapeutic administration to humans.

Often CBA/J mice infected with *Salmonella* would clear the infection and did not foster the levels of inflammation necessary for *Salmonella* enteric expansion. In this dissertation I briefly explored data from CBA/J mice both infected with *Salmonella* and fed a high fat diet (HFD). While HFD did increase the proportion of high-responder (25% *Salmonella* relative abundance in at least one measurement) mice, the ability to predict which chow fed mice would become high responders based on pre-infection data

would be useful. Members of our lab and myself have tried to predict CBA/J high-responders from 16S rRNA amplicon community data with little success. Pre-infection multi-omics measurements with high enough replication to employ machine learning to this classification problem have thus far been cost prohibitive. It is my hope that with the publication of the CBAJ-DB, more CBA/J microbiome datasets will become publicly available that we can then use to profile the microbiome of high-responding CBA/J mice to better inform future experiments.

In summary, comprehensive whole microbiome multi-omics discovery offers a unique glimpse at the invisible communities that are so impactful to our health. Broad discovery followed by targeted and controlled validation presents a pathway to harness biological processes and players for the benefit of the human condition. The exciting frontier of biotic therapies offers an endless expanse to explore strange new ecosystems, to seek out new life and new mechanisms, and to boldly go where no scientist has gone before.

Chapter 4 References

1. Popa GL, Papa MI. Salmonella spp. infection - a continuous threat worldwide. Germs 2021; 11:88–96.

2. Plumb ID, Brown AC, Stokes EK, Chen JC, Carleton H, Tolar B, Sundararaman P, Saupe A, Payne DC, Shah HJ, et al. Increased Multidrug-Resistant Salmonella enterica I Serotype 4,[5],12:i:- Infections Associated with Pork, United States, 2009–2018 - Volume 29, Number 2—February 2023 - Emerging Infectious Diseases journal - CDC. [cited 2023 Apr 3]; Available from: https://wwwnc.cdc.gov/eid/article/29/2/22-0950_article

3. Castillo NA, de Moreno de LeBlanc A, Galdeano CM, Perdigón G. Probiotics: An alternative strategy for combating salmonellosis: Immune mechanisms involved. Food Res Int 2012; 45:831–41.

4. Ansaldo E, Slayden LC, Ching KL, Koch MA, Wolf NK, Plichta DR, Brown EM, Graham DB, Xavier RJ, Moon JJ, et al. Akkermansia muciniphila induces intestinal adaptive immune responses during homeostasis. Science 2019; 364:1179–84.

5. Zhang T, Li Q, Cheng L, Buch H, Zhang F. Akkermansia muciniphila is a promising probiotic. Microb Biotechnol 2019; 12:1109–25.

6. Ghaffari S, Abbasi A, Somi MH, Moaddab SY, Nikniaz L, Kafil HS, Ebrahimzadeh Leylabadlo H. Akkermansia muciniphila: from its critical role in human health to strategies for promoting its abundance in human gut microbiome. Crit Rev Food Sci Nutr 2022; 0:1–21.