Development of Analytical Tools and Animal Models for Studies of Small-Intestine Dysbiosis

Thesis by

Said R. Bogatyrev

In Partial Fulfillment of the Requirements for

the degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2020

(Defended September 20, 2019)

© 2019

Said R. Bogatyrev
ORCID: 0000-0003-0486-9451

DEDICATION

I dedicate this work, my biggest accomplishment in life, to my family, particularly to my parents and my brother, who were and are the highest professionals in their art yet always remained the people with great humility; and to all my family members for all the sacrifices the generations of them made when they perished in Gulag and while fighting in WWII, or during the disorderly years of the Post-Soviet Russia, in order for me to be here now and do the work I love.

I dedicate this work to all the teachers, friends, and colleagues who always saw and treated me as a person and not as a dispensable tool on their path to success. To all those who appreciated my loyalty, trusted me, and remained loyal.

You are the giants on whose shoulders I am standing right now. I aspire to live up to your expectations.

ACKNOWLEDGEMENTS

I would like to thank Caltech for giving me the opportunity to pursue the doctoral degree within its walls. I would like to thank Chin-Lin Guo, Richard Murray, Rob Phillips, and Paul Patterson for evaluating me highly during the PhD admissions interviews and welcoming me to this university.

I would like to thank my thesis committee members Michael Elowitz, Sarkis Mazmanian, and Paul Sternberg, as well as Graduate Dean Doug Rees, for their encouragement and advice.

BE program coordinator Linda Scott and administrative staff of the BBE and CCE departments, the Registrar's Office, and the Dean's Office are acknowledged for their help with keeping my academic affairs in good order.

Dr. Karen Lencioni and Dr. Janet Baer are kindly acknowledged for their support of our efforts in the development of novel animal models. All OLAR staff and veterinary technicians are acknowledged for their dedication in keeping our research subjects healthy and happy.

All members of the Ismagilov Lab who taught me something new, performed their group jobs diligently to keep the lab running, and provided good partnership are greatly acknowledged.

This dissertation in its entirety or in parts would not be presentable to the reader without the enormous efforts and time Natasha Shelby has put into helping me organize the results of the research into comprehensible writing.

The research described here would not be possible without the intellectual and material resources provided by my adviser, Prof. Rustem F. Ismagilov, for which he is kindly acknowledged.

ABSTRACT

Our appreciation of the role of human-associated microbial communities in the context of human health and disease has grown dramatically in the past two decades, with modern research tools enabling deeper insights into the mechanisms of host-microbial interactions. The elusive notion of dysbiosis, a state of microbial imbalance related to a disease, has achieved widespread distribution across popular, scientific, and medical literature (on September 16, 2019 PubMed search yielded 6,064 records of scientific and medical publications containing this keyword). The conventional wisdom further narrows down the definition and understanding of dysbiosis towards a compositional "imbalance" of the microbiota (a community of microorganisms inhabiting the human body). There exists an additional and frequently overlooked aspect of microbial imbalance in the context of the human gastrointestinal system, something that we can define as a "spatial imbalance": a state of the microbial community in the host gastrointestinal system where even a "healthy" and "balanced" microbiota may be associated with or causative of a disease by being present in sections of the gastrointestinal tract where it is not "supposed" to be, with the most prominent example being small intestinal bacterial overgrowth (SIBO). This thesis describes the progress in the development of analytical tools (quantitative microbiome profiling described in Chapter I) and refinement of animal mouse models (non-coprophagic mouse model described in Chapter II) for exploring the normal function of small-intestine microbiota in health and for dissecting the mechanisms of emergence and the persistence of the small-intestine dysbiosis (SIBO) in the future.

PUBLISHED CONTENT AND CONTRIBUTIONS

1. **Said R. Bogatyrev** and Rustem F. Ismagilov. Quantitative microbiome profiling in lumenal and tissue samples with broad coverage and dynamic range via a single-step 16S rRNA gene DNA copy quantification and amplicon barcoding. *In preparation*.
    - **S.R.B. contributions: Conception, assay optimization for different types of samples, animal study execution, animal study sample processing for quantitative 16S rRNA gene amplicon sequencing, quantitative 16S rRNA gene amplicon sequencing and data analysis, manuscript preparation.**

2. **Said R. Bogatyrev**, Justin C. Rolando, and Rustem F. Ismagilov. Self-reinoculation with fecal flora changes microbiota density and composition leading to an altered bile-acid profile in the mouse small intestine. *Pre-submitted*.
    - **S.R.B. contributions: Conception, mouse tail cup development, animal study execution, animal study sample processing for quantitative 16S rRNA gene amplicon sequencing, quantitative 16S rRNA gene amplicon sequencing and data analysis, animal study sample processing for metabolomic analysis, bile acid metabolomics data analysis, manuscript preparation.**
    - J.C.R.: Metabolomics method development and validation, animal study sample processing for metabolomic analysis, UPLC-MS instrument setup and sample analysis, chromatographic and mass spectra data analysis.

3. Jacob T. Barlow, **Said R. Bogatyrev**, and Rustem F. Ismagilov. A quantitative sequencing framework for absolute abundance measurements of mucosal and lumenal microbial communities. *In review*.
    - **S.R.B. contributions and authorship order are in revision.**

4. Tahmineh Khazaei, Rory L. Williams, **Said R. Bogatyrev**, John C. Doyle, Christopher S. Henry, Rustem F. Ismagilov. Metabolic bi-stability and hysteresis in a model microbiome community. *In review*.
    - **S.R.B. contributions: Hypothesis ideation with T.K. Designed and performed preliminary experiments with T.K.: evaluating Kp-Bt community growth in batch culture as a function of substrate concentration, selectivity, and redox potential in the system.**
    - T.K.: Hypothesis ideation with SRB. Design of study. Performed preliminary experiments with SRB: evaluating Kp-Bt community growth under various glucose conditions in batch culture. Built the mathematical model used in this study (Figure 1). Using the mathematical models predicted state-switching and hysteresis within the community and identified regions of bi-stability with respect to glucose and oxygen input conditions (Figure 2). Designed CSTR experiments. These experiments were performed by RLW with help from TK. Established the protocol for short chain fatty acids measurements (further optimized by RLW). Established the protocol for RNA extraction of CSTR samples for RNA sequencing. Performed the RNA extraction of all CSTR samples for RNA sequencing. Established the bioinformatics pipeline for processing and analyzing the CSTR samples (mixed-species samples). Processed and analyzed the RNA sequencing data (Figure 4). Wrote and made figures for the manuscript.
    - R.L.W.: Performed preliminary plate reader experiments testing state switching with BT/KP and BT/E. coli that determined we would use KP in CSTR experiments. Established the CSTR workflow, optimized media conditions, and performed the CSTR experiments with help from TK (Figure 2). Designed CSTR experiments with TK. Worked with Nathan Dalleska to optimize HPLC for the measurement of SCFAs in CSTR samples. Performed qPCR of all the CSTR samples (Figure 2). Characterized some of the Michaelis Menton constants used in the

mathematical models. This was done through batch experiments for growth of Bt and Kp on various substrates and Bayesian parameter inference. Helped TK in preparing the manuscript.

- R.P.: Performed HCR-FISH and DAPI staining on the bioreactor samples embedded into acrylamide gels and imaged. Created imaging figure.

5. Asher Preska Steinberg, Sujit S. Datta, Thomas Naragon, Justin C. Rolando, **Said R. Bogatyrev**, and Rustem F. Ismagilov. 2019. "High-molecular-weight polymers from dietary fiber drive aggregation of particulates in the murine small intestine." *eLife*. 8:e40387. DOI: 10.7554/eLife.40387
    - **S.R.B. contributions: Investigation, Methodology, Writing – review and editing, Co-performed preliminary experiments; developed fluorescent laser scanning approach appearing in Fig. 1A and 1B; Administered particles to mice in Fig. 1; co-developed approach to extract liquid fraction of murine intestinal contents; co-organized transfer and initial set up of the MUC2KO mutant mouse colony; setup genotyping of MUC2KO mice; helped supervise animal husbandry of MUC2KO colony; helped with interpretation of results.**
    - A.P.S.: Conceptualization, Resources, Data curation, Software, Formal analysis, Funding acquisition, Validation, Investigation, Visualization, Methodology, Writing – original draft, Writing – review and editing, Co-designed all experiments and co-analyzed all experimental results; developed theoretical tools and performed all calculations; co-developed imaging analysis pipeline in ImageJ; developed computational tools for bootstrapping procedure; co-developed microscopy assay (Fig 1C-D;Co-performed, designed, and analyzed data from gavage experiments in Fig 1; performed, designed, and analyzed data from all ex vivo SI aggregation experiments in Figs 2, 3, 5-7; performed, designed, and analyzed data from all GPC measurements in Figs 3, 5-7, and Tables 1-7; performed, designed, and analyzed data from all in vitro PEG aggregation experiments in Fig 4D, Fig 4 - supplements2-3, and with dietary fiber in Fig 7A; developed a

computational approach for theoretical calculations in 4H and 4I and performed all calculations; performed, designed, and analyzed data from Western blots in Figs 5E, 6E, Fig 6- supplements 1-2; helped supervise animal husbandry of MUC2KO colony; performed animal husbandry for WT mice on autoclaved diets in Fig 6; performed animal husbandry for mice on pectin and Fibersol-2 diets in Fig 7; performed, designed, and analyzed all zeta potential measurements in Table 8; performed pH measurements on luminal fluid in Fig 4 - supplement 1; co-interpreted results.

- S.S.D.: Conceptualization, Investigation, Methodology, Writing – review and editing, Conceived and co-planned the project; initially observed the aggregation phenomenon; co-designed and co-analyzed preliminary experiments; performed preliminary ex vivo and in vitro aggregation experiments; co-developed microscopy assay used in Fig 1C and 1D; developed ex vivo/in vitro aggregation assay used in Figs 2-7; co-developed approach to extract liquid fraction of murine intestinal contents; co-developed NMR protocol; organized transfer and initial set up of MUC2KO colony; co-interpreted results.
- T.N.: Data curation, Software, Formal analysis, Methodology, Writing – original draft, Co-developed imaging analysis pipeline in ImageJ; co-analyzed ex vivo aggregation data in Fig 2; co-designed and co-analyzed preliminary ex vivo aggregation experiments with MUC2KO mice; provided useful advice on bootstrapping procedure; co-interpreted results.;
- J.C.R.: Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Developed protocol for NMR measurements on PEG-coated particles, Performed synthesis of particles, Performed NMR measurements in Table 8.

6. Joseph M. Pickard, Corinne F. Maurice, Melissa A. Kinnebrew, Michael C. Abt, Dominik Schenten, Tatyana V. Golovkina, **Said R. Bogatyrev**, Rustem F. Ismagilov, Eric G. Pamer, Peter J. Turnbaugh, and Alexander V. Chervonsky.

"Rapid fucosylation of intestinal epithelium sustains host-commensal symbiosis in sickness." *Nature*. 2014 514:638–641. DOI: 10.1038/nature13823.

- **S.R.B. contributions: Developed a simplified GC-MS method for targeted metabolomics analysis of short-chain fatty acids and performed analysis of short-chain fatty acids in experimental animal samples.**

TABLE OF CONTENTS

Chapter 1

QUANTITATIVE MICROBIOME PROFILING IN LUMENAL AND TISSUE SAMPLES WITH BROAD COVERAGE AND DYNAMIC RANGE VIA A SINGLE ONE-STEP 16S RIBOSOMAL RNA GENE DNA COPY QUANTIFICATION AND AMPLICON BARCODING

Said R. Bogatyrev and Rustem F. Ismagilov

ABSTRACT

Current methods for detecting, accurately quantifying, and profiling complex microbial communities based on the microbial 16S rRNA marker genes are limited by a number of factors, including inconsistent extraction of microbial nucleic acids, amplification interference from contaminants and host DNA, different coverage of PCR primers utilized for quantification and sequencing, and potentially biases in PCR amplification rates among microbial taxa during amplicon barcoding. Here, we describe a method that enables the quantification of microbial 16S rRNA gene DNA copies with wide dynamic range and broad microbial diversity, and simultaneous amplicon barcoding for quantitative profiling of microbiota based on 16S rRNA gene amplicon sequencing. The method is suitable for a variety of sample types and is robust in samples with low microbial abundance, including samples containing high levels of host mammalian DNA, as is common in human clinical samples. We demonstrate that our modification to the Earth Microbiome Project (EMP) V4 16S rRNA gene primers expands their microbial coverage while dramatically reducing non-specific mammalian mitochondrial DNA amplification, thus achieving wide dynamic range in microbial quantification and broad coverage for capturing high microbial diversity in samples with or without high host DNA background. The approach relies only on broadly available hardware (real-time PCR instruments) and standard reagents utilized for conventional 16S rRNA gene amplicon library preparation. Simultaneous 16S rRNA gene DNA copy quantification and amplicon barcoding for multiplexed next-generation sequencing from the same analyzed sample, performed in a combined workflow, reduces time and reagent costs, all of which make the approach amenable for immediate and

widespread adoption. Additionally, we demonstrate that using our modified 16S rRNA gene primers in a digital PCR (dPCR) format enables precise and exact microbial quantification in samples with very high host DNA background levels without the need for quantification standards. Potential future applications of this approach include: (1) quantitative microbiome profiling in human and animal microbiome research; (2) detection of monoinfections and profiling of polymicrobial infections in tissues, stool, and bodily fluids in human and veterinary medicine; (3) environmental sample analyses (e.g., soil and water); and (4) broad-coverage detection of microbial food contamination in products high in mammalian DNA, such as meats. We predict that utilization of this approach primarily for quantitative microbiome profiling will be invaluable to microbiome studies, which have historically been limited to analysis of relative abundances of microbes.

INTRODUCTION

Microbiome analysis has emerged as a prominent research field to improve our understanding of the host-microbiota interactions linked to human disease. Utilization of high-throughput next generation sequencing (NGS) technology in combination with microbial marker gene sequencing (e.g., microbial 16S rRNA gene) has enabled high-diversity and high-depth compositional analyses of microbiomes. NGS-based compositional analyses (relative abundances of the microbiome elements) have dominated the field since their emergence. The limitations of compositional analyses have been gaining broader acknowledgement in the field and a number of quantitative microbiome profiling approaches have been proposed as promising tools for solving the shortcomings of purely compositional analyses. Current quantitative analysis approaches have important limitations: (i) high levels of host DNA interfere with the amplification of target microbial sequences, (ii) coverage of microbial taxa is limited, and (iii) relative quantification cannot provide a complete picture of changes in microbial taxa.

Here, to address the aforementioned limitations of current quantitative analysis methods, we describe an approach that allows simultaneous (with one sample) determination of the

absolute 16S rRNA gene DNA copy loads with broad dynamic range and enables wide-diversity microbiome profiling in a simplified and broadly-adoptable workflow (Fig. 1). The proposed approach for quantitative 16S rRNA gene amplicon profiling is based on the combination of absolute 16S rRNA gene DNA copy quantification and 16S rRNA gene amplicon sequencing utilizing a real-time PCR amplification readout and amplicon barcoding for NGS performed for the variable V4 region of the prokaryotic 16S rRNA gene sequence amplicon.

This approach is optimized for use in samples with high and low levels of mammalian (e.g., mouse) host DNA which enables quantitative 16S rRNA gene amplicon sequencing of clinical samples, such as stool, gastrointestinal contents or lavage fluid, and mucosal biopsies.

"One-step" approach includes the following workflow steps:

A. Total DNA is extracted and purified from such samples using commercially-available kits (Fig. 1A) validated for uniform DNA extraction from complex microbiota [e.g., ZymoBIOMICS] and for quantitative recovery of microbial DNA from samples with microbial loads across multiple orders or magnitude (Fig. S1).

B. PCR reactions are set up using the improved 16S rRNA gene primers and conventional commercial reagents for 16S rRNA gene amplicon library preparation together with the universal 16S rRNA gene primers containing barcodes and Illumina adapters (Fig. 1B). Reactions are replicated to improve the real-time PCR quantification precision and resolution and amplicon barcoding uniformity [1].

C. Amplification and barcoding of the V4 region of the microbial 16S rRNA gene DNA are performed under real-time fluorescence measurements on a real-time PCR instrument (Fig. 1C). We define this approach as "barcoding qPCR" or "BC-qPCR". Real-time fluorescence monitoring enables terminating the amplification of each sample upon reaching the mid-exponential phase to maximize the amplicon yield and minimize the overamplification artifacts [2].

D. Quantitative real-time PCR data (Cq values) are recorded (Fig. 1D) and used to calculate the absolute concentration of the 16S rRNA gene DNA copies in each sample (based on the 16S rRNA gene copy standards included within the same BC-qPCR run) or to calculate the absolute fold-differences in the 16S rRNA gene DNA copy load among the samples (in the absence of the standards).These data are further used to calculate the absolute microbial abundances in the analyzed samples.

E. Barcoded 16S rRNA gene DNA amplicon samples are quantified, pooled, purified, and sequenced on an NGS instrument.

F. NGS sequencing results provide the sequence read and count data from which the microbial identity and relative abundances of the microbial taxa are estimated (Fig. 1F).

G. Microbiota relative abundance profiles (from step "F") are converted to microbiota absolute or absolute fold-difference abundance profiles using the absolute or absolute fold-difference data on 16S rRNA gene DNA loads in the corresponding samples (as measured in the step "D") (Fig. 1G).

To achieve the desired broad dynamic range and coverage of the quantitative 16S rRNA gene amplicon sequencing and its robust performance in samples with high or low host DNA background, we (I) modified the universal 16S rRNA gene primers for gene-copy quantification in qPCR and ddPCR assays and amplicon barcoding in BC-qPCR with high specificity against host DNA; (II) optimized the BC-qPCR parameters to minimize primer dimer formation and host DNA amplification while reducing amplification biases and ensuring uniform amplification of diverse 16S rRNA gene sequences from complex microbiomes; and (III) validated the accuracy of the quantitative 16S rRNA gene amplicon sequencing obtained using the "one-step" BC-qPCR approach compared with the quantitative 16S rRNA gene amplicon sequencing results obtained using real-time and digital PCR.

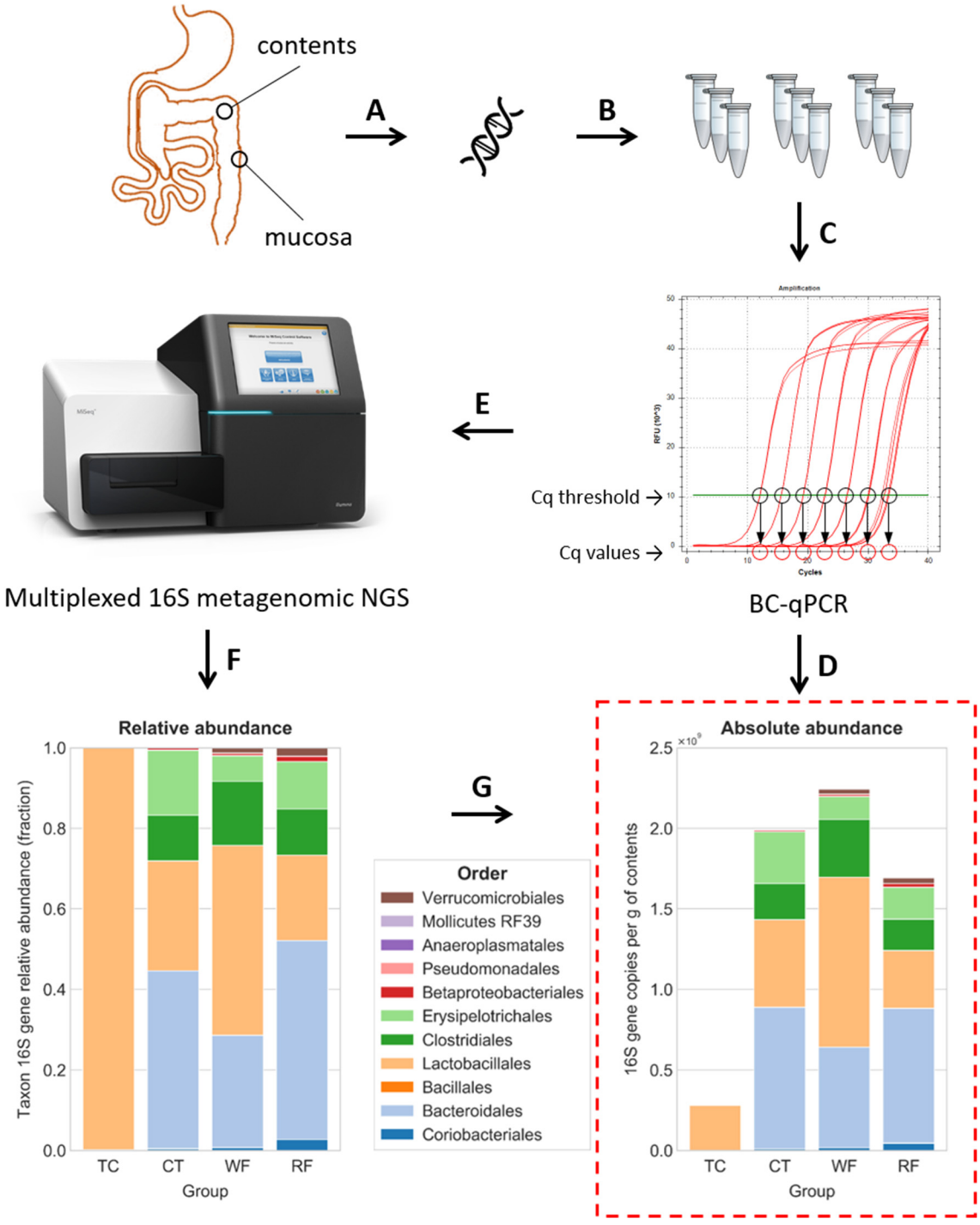

Fig. 1. Schematic of the "one-step" 16S rRNA gene DNA quantification and amplicon barcoding workflow ("BC-qPCR") implementation for quantitative microbiome profiling. (**A**) Sample collection and DNA extraction. (**B**) BC-qPCR reactions are prepared

in replicates for more accurate quantification and uniform amplicon barcoding. (**C**) Amplification and barcoding are performed under real-time fluorescence measurements on a real-time PCR instrument. (**D**) Quantitative PCR data (Cq values) are recorded. (**E**) Barcoded samples are quantified, pooled, purified, and sequenced on an NGS instrument. (**F**) NGS sequencing results provide data on relative abundances of microbial taxa. (**G**) Microbiota relative abundance profiles are converted to microbiota absolute or absolute fold-difference abundance profilies using the absolute or absolute fold-difference data on 16S rRNA gene DNA loads in the corresponding samples measured in step (D).

RESULTS

Optimized primers improve broad-coverage 16S rRNA gene DNA quantification via real-time and digital PCR in the presence of high host DNA background

We first aimed to adapt the Earth Microbiome Project (EMP) 16S rRNA gene amplicon sequencing protocol [1], [3] for quantitative microbiota profiling. This protocol is well-known for having broad microbial coverage and has been widely adopted in the field of basic and clinical microbiome research. We hypothesized that by redesigning the EMP forward primer (designated by us as UN00F0) at its 5′ end to start at the position 519 (UN00F2) of the V4 region of microbial 16S rRNA gene sequence (Fig. 2A) we would either reduce or eliminate its nonspecific annealing to the mouse and human mitochondrial 12S rRNA gene DNA. Such change would increase the primer's specificity for low copy number microbial templates in samples with high content of mouse or human host DNA background. We confirmed the effectiveness of these design considerations by performing qPCR reactions in complex mouse microbiota DNA samples analyzed neat or spiked in with GF mouse small-intestine mucosal DNA at 100 ng/uL. The ~200-bp mithochondrial amplicons were absent in the PCR reactions containing high amounts of mouse DNA and using the modified forward primer UN00F2 (Fig. 2B).

The efficiency of the quantitative PCR reactions set up with the modified forward primer UN00F2 was similar (and high) with and without the presence of 100 ng/µL of mouse DNA in the template sample (Fig. 2C) demonstrating the robust assay performance.

Our qPCR experiments also suggested that the PCR reactions with high host DNA background are intercalating dye-limited: the increase in total fluorescence (Δ-RFU) in each reaction at the end of amplification was lower in samples containing 100 ng/µL of background mouse DNA whereas the total fluorescence levels were similar between samples with and without the background mouse DNA. By combining the use of the new forward primer UN00F2 with the supplementation of commercial reaction mix with additional amounts of intercalating EvaGreen dye improved the digital PCR performance by increasing the separation between negative and positive droplets in the droplet digital PCR (ddPCR) reactions used for quantifying 16S rRNA gene DNA copies in samples with high host DNA background (100 ng/uL) (Fig. 2D). This assay was used to establish or confirm the exact 16S rRNA gene DNA copy numbers in the standard samples, which were further utilized to build the standard curves in the qPCR assays.

Additionally, the modification of the primer set UN00F2 + UN00R0 broadened its taxonomical coverage of the microbial diversity (86.0% Archaea, 87.0% Bacteria) compared with the original EMP primer set UN00F0 + UN00R0 (52.0% Archaea, 87.0% Bacteria) based on the SILVA reference database [4], [5].

Modified barcoded primers and optimized workflow enable simultaneous 16S rRNA gene DNA copy quantification and amplicon barcoding in samples with high host DNA background

We next aimed to evaluate whether the barcoded UN00F2 + UN00R0 primer set would allow the amplification and amplicon barcoding of specific microbial 16S rRNA gene DNA template in the presence of high host DNA background. It is important to note the two essential design principles in the BC-qPCR reaction optimization that guided our work:

1. The amplification and barcoding reaction should be conducted at the lowest possible annealing temperature to maximize the uniformity of amplification of the diverse 16S rRNA gene DNA sequences with degenerate primers (both original and improved EMP) and eliminate the amplification biases.

2. The amplification and barcoding reaction should be conducted at the highest possible annealing temperature to minimize the primer dimer formation and non-specific host mitochondrial DNA amplification both of which would be competing with specific microbial 16S rRNA gene DNA template for reaction resources (dNTPs, primers, polymerase, intercalating dye). Such competing reactions would inevitably have pronounced effects on the samples containing very low levels of specific microbial template and requiring high numbers of amplification cycles.

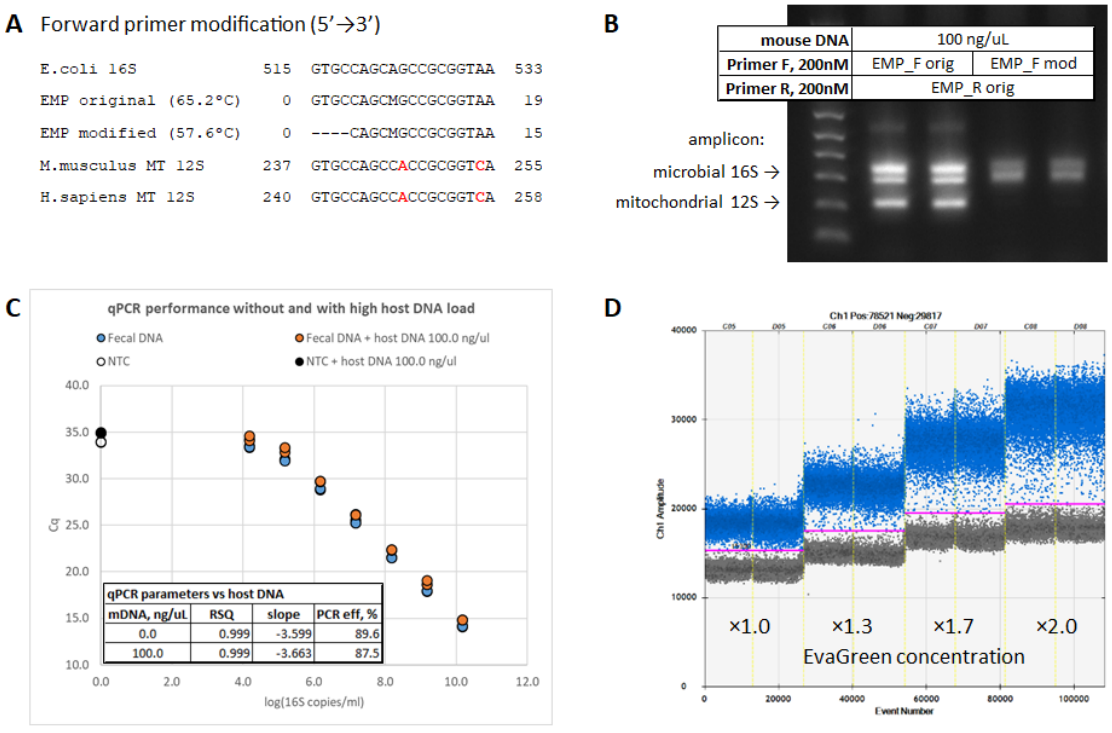

Fig. 2. Optimization of the protocol for microbial 16S rRNA gene DNA copy quantification in samples without and with high mammalian DNA background. (A) Sequence alignment of the original EMP and modified forward primers targeting the V4 region of microbial 16S rRNA gene are shown with the *E. coli* 16S rRNA gene and mouse and human mitochondrial 12S rRNA gene. **(B)** Amplification products of the complex microbiota DNA sample containing 100 ng/uL of GF mouse DNA with the original EMP and modified forward primers. **(C)** Quantitative PCR reaction performance with the serial 10-fold dilutions of the complex microbiota DNA sample with and without 100 ng/μL of mouse DNA. **(D)** Improvement of the 16S rRNA gene DNA copy ddPCR quantification assay performance in the presence of 100 ng/μL of mouse DNA background as a result of the supplementation of intercalating dye to the commercial droplet digital PCR (ddPCR) master mix.

Compared with the improved primer set (UN00F2 + UN00R0), the original EMP primer set (UN00F0 + UN00R0) requires a higher annealing temperature to reduce primer dimer formation and amplification of mouse mitochondrial (MT) DNA. Long "overhangs" (carrying the linker and Illumina adapter sequences) at the 5′ end of the forward primer and non-complimentary to the specific 16S rRNA gene DNA template were not sufficient to prevent the EMP primer set from amplifying the mouse MT DNA. At 54 °C both primer dimers and MT DNA amplification persisted in the reactions using the EMP primers, which suggested that this primer set would require even higher annealing temperatures (>54 °C) to eliminate the amplification artifacts. This in turn will likely introduce amplification biases across a range of specific 16S rRNA gene DNA templates. Using the improved primer set eliminated both artifacts in the reactions conducted at 54 °C (Fig. 3A), while some primer dimer formation was still present in the reactions conducted at 52 °C. Thus, the temperature of 54 °C was selected as optimal for the BC-qPCR reaction.

We next confirmed that the BC-qPCR reaction can provide accurate quantification data for the amount of 16S rRNA gene DNA copy loads in the analyzed samples. The Cq values

obtained based on the real-time fluorescence measurements during the BC-qPCR reaction were in good agreement with the absolute 16S rRNA gene DNA copy values (Fig. 3B) estimated in the same samples using the previously optimized qPCR assay (Fig. 2C).

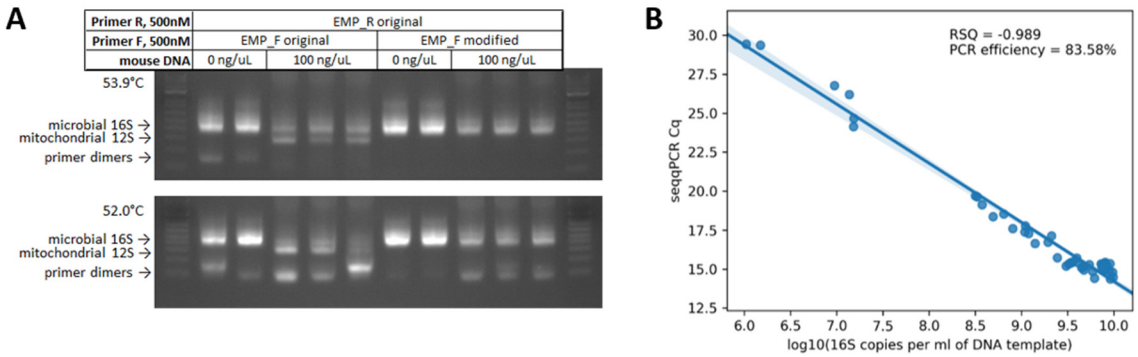

Fig. 3. Optimization of the "one-step" protocol for microbial 16S rRNA gene DNA copy quantification and amplicon barcoding in samples without and with high mammalian DNA background. (A) Amplification products of the complex microbiota DNA sample containing 100 ng/µL of GF mouse DNA with the barcoded original EMP (UN00F0 + UN00R0) and barcoded modified (UN00F2 + UN00R0) primer sets. **(B)** Correlation of the BC-qPCR Cq values (Y-axis) with the absolute 16S rRNA gene DNA copy numbers (X-axis) previously determined in the same set of samples (with and without high host DNA background) using the UN00F2 + UN00R0 qPCR assay.

One-step approach enables absolute or absolute fold-change microbiota profiling

To evaluate the accuracy of the absolute abundances or absolute abundance fold-differences estimated using the "one-step" approach, the BC-qPCR data were validated against the absolute abundances previously obtained using a two-step approach (Fig. 4) on the same set of samples [6]. The BC-qPCR approach provides the fold-differences in absolute microbial

abundances among samples even in the absence of the exact microbial load estimates (i.e., when no standard curve is available).

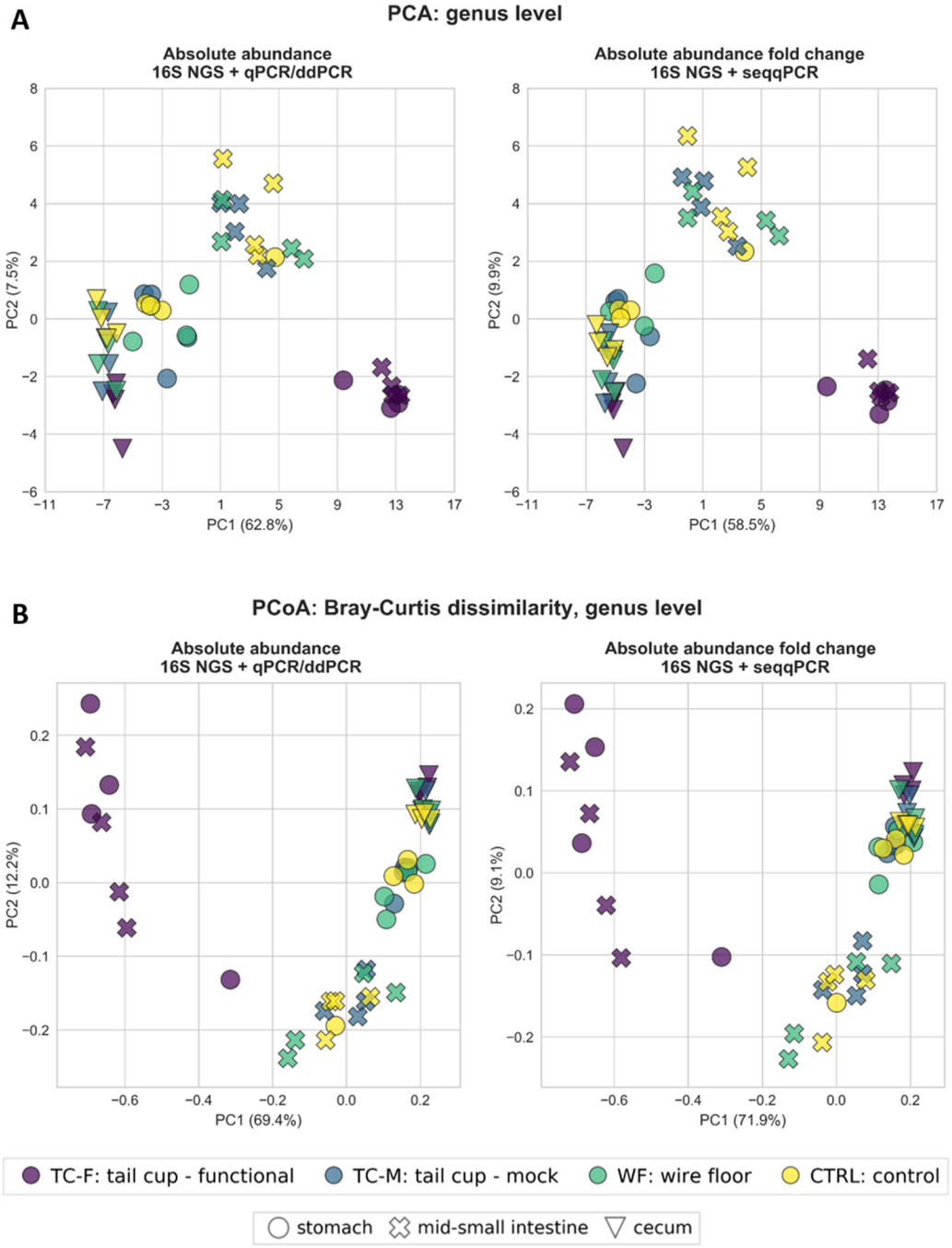

Fig. 4. Exploratory 16S rRNA gene amplicon sequencing data analysis of the absolute complex microbiota profiles (in samples from [6]**) obtained using the standard quantification and sequencing approach or using the "one-step" approach. (A)** Principal component analysis (PCA) of the absolute (left), estimated using the multistep approach, and

absolute fold-difference (right) microbiome profiles, obtained using the "one-step" approach with the assumed BC-qPCR efficiency of 85.0%. (**B**) Principal coordinate analysis (PCoA) of the Bray-Curtis dissimilarity matrices obtained for the same types of data as in panel (A). All values were multiplied by 10^2 to ensure the $\log_{10}$-transformed values of the non-anchored absolute abundances obtained from the BC-qPCR were greater than zero (> 0).

CONCLUSIONS

The "one-step" BC-qPCR approach enables accurate quantification of the number of 16S rRNA DNA gene copies and unbiased absolute abundance profiling of the microbial community structure in samples with microbial loads varying across multiple orders of magnitude and containing high host DNA background. The BC-qPCR approach offers the following advantages over the methods currently used in the field:

- Broader coverage of microbial diversity (87% bacteria, 87% of archaea based on the 16S rRNA marker gene sequences [4], [5]) maximizes the completeness of microbial detection and quantification and richness of diversity profiling.

- Microbial 16S rRNA gene DNA copy quantification demonstrated a broad dynamic range: the lower limit of quantification (LLOQ) – ~$10^{4.83}$ copies/mL and the, upper limit of quantification (ULOQ) – ~$10^{10.95}$ copies/mL.

- Quantification has high resolution – ~1.25-1.67-fold differences in absolute 16S rRNA gene DNA copy concentrations can be distinguished in the demonstrated dynamic range with and without high host DNA background (100 ng/uL).

- "What's quantifiable – is sequenceable, what's sequenceable – is quantifiable": our method maximizes correspondence between the total 16S rRNA gene DNA copy quantification data and 16S rRNA gene amplicon sequencing profiling data as a major advantage over the currently implemented approaches [7]–[11].

- Primer design allows for a good 16S rRNA gene DNA real-time (quantitative) PCR, digital PCR, and amplicon barcoding PCR reaction performance in samples with high mammalian host DNA background. No host DNA depletion is required for accurate microbial quantification and profiling, which is an advantage over the methods currently implemented in the field.

- Optimized “one-step” 16S rRNA gene DNA amplicon barcoding and quantification approach (performed in a single PCR reaction instead of two separate PCR reactions for quantification and barcoding) reduces the reagent and time costs while providing richer absolute or fold-difference microbiota profiles of the analyzed samples.

- Optimized amplicon barcoding PCR reaction chemistry and workflow prevent amplification artifacts and biases [2] that could affect the accuracy of relative abundance measurements across samples with broad range of microbial loads and thus requiring different numbers of amplification cycles.

- The approach eliminates the need in synthetic spike-ins for accurate quantitiative 16S rRNA gene amplicon sequencing. Easily accessible commercial microbiome standards (e.g., ZymoBIOMICS) can be integrated as quantitative standards in the proposed protocol.

- The approach may be applicable in both single (described in this report) and dual-indexing workflows.

- Overall, the proposed “one-step” approach for quantitative 16S rRNA gene amplicon sequencing based on the conventional real-time (quantitative) qPCR workflow allows for broad and immediate adoption of the approach in the field of basic and clinical microbiome research.

METHODS

For methods please refer to the “METHODS” section of the Chapter 2 of this dissertation and in [6].

ACKNOWLEDGEMENTS

This work was supported in part by the Kenneth Rainin Foundation Innovator Award, Army Research Office (ARO) Multidisciplinary University Research Initiative (MURI) contract #W911NF-17-1-0402, and the Jacobs Institute for Molecular Engineering for Medicine. We thank Natasha Shelby for contributions to writing and editing this manuscript.

Chapter II

SELF-REINOCULATION WITH FECAL FLORA CHANGES MICROBIOTA DENSITY AND COMPOSITION LEADING TO AN ALTERED BILE-ACID PROFILE IN THE MOUSE SMALL INTESTINE

Said R. Bogatyrev, Justin C. Rolando, and Rustem F. Ismagilov

ABSTRACT

Alterations to the small-intestine microbiome are implicated in various human diseases, yet the physiological and functional roles of the small-intestine microbiota remain poorly characterized because of sampling complexities. Murine models enable spatial, temporal, compositional, and functional interrogation of the gastrointestinal microbiota, however fecal microbial self-reinoculation (via coprophagy, ubiquitous among rodents) can affect the structure and function of microbiota in the upper gut. Using quantitative 16S rRNA gene amplicon sequencing, quantitative microbial functional gene content inference, and targeted metabolomics, we found that self-reinoculation had profound quantitative and qualitative effects on the mouse small-intestine microbiota, which led to altered bile-acid profiles. The patterns observed in the small intestine of non-coprophagic mice (reduced total microbial load, low abundance of anaerobic microbiota, and bile acids predominantly in the conjugated form) resemble those typically seen in the human small intestine. The implications of our study are likely to be important for future research using mouse models to evaluate gastrointestinal microbial colonization and function in the context of bile-acid and xenobiotic metabolism, diet and probiotics research, and diseases related to small-intestine dysbiosis.

INTRODUCTION

The small intestine is the primary site for enzymatic digestion and nutrient uptake, immune sampling, and drug absorption in the human gastrointestinal system. Its large surface area

vastly exceeds that of the large intestine [12], and thus may serve as a broad interface for host-microbial interactions.

A growing body of scientific evidence highlights the importance of the small-intestine microbiome in normal human physiology and response to dietary interventions [13], [14]. Alterations in the small-intestine microbiome are implicated in a number of human disorders, such as malnutrition [15], [16], obesity, and metabolic disease [17], inflammatory bowel disease (IBD) and irritable bowel syndrome (IBS) [18]–[20], and drug side effects [21]. Despite the apparent importance of the small-intestine microbiome in human health, it remains understudied and poorly characterized largely because of the procedural and logistical complexities associated with its sampling in humans (methods are too invasive and require specialized healthcare facilities). Moreover, microbial composition tends to differ substantially among the small intestine, large intestine, and stool of the same animal or human subject [22], [23], which highlights the importance of targeted sampling of the small intestine for analyses.

Mice are the predominant animal species of model organisms in the field of microbiome research. Compared with other mammalian models, mice have a lower cost of maintenance, their environment and diet can be easily controlled, they are amenable to genetic manipulation, there are numerous genetic mouse models already available, and propagation using inbred colonies reduces inter-individual variability [24]. Additionally, murine germ-free (GF) and gnotobiotic technologies are well established. Using mouse models enables interrogation of the entire gastrointestinal tract (GIT) and examination of the changes in microbiome and host physiology that occur in response to experimental conditions (e.g., dietary modifications, xenobiotic administration, etc.) or microbial colonization (e.g., monocolonization, colonization with defined microbial consortia, human microbiota-associated mice, etc.).

Rodent models also have several well-recognized limitations associated with their genetic, anatomical, and physiological differences with humans [24], [25]. Among these limitations is the persistent tendency of rodents to practice gastrointestinal auto- and allo-reinoculation

with large-intestine microbiota (via fecal ingestion, or coprophagy) in laboratory settings [26]–[28]. This pervasive behavior has been documented in classical studies using observational techniques in both conventional and GF mice [29], in conventional mice maintained on standard and fortified diets [30], in animals with and without access to food [31], and across different mouse strains [27], [32].

Multiple classical studies have attempted to evaluate the effects of self-reinoculation on the structure of the microbiota in the rodent small intestine [33]–[35] and large-intestine and stool [31], [34], [36], [37] using traditional microbiological techniques, but reported conflicting results [34], [36], [37]. This lack of consensus may be attributed to the use of different methods for preventing coprophagy (some of which are ineffective), non-standardized diets, inter-strain or inter-species differences among the animal models, or other unaccounted for experimental parameters. It has been also suggested that repeated self-exposure in mice via coprophagy can promote microbial colonization of the GIT by "exogenous" microbial species, such as *Pseudomonas* spp. [38]. All of these observations highlight the importance of considering self-reinoculation in studies of gastrointestinal microbial ecology in murine models. However, the field currently lacks precise and comprehensive evaluations of the effects of self-reinoculation on the spatial, structural, and functional state of the gut microbiome and its effects on murine host physiology. Current microbiome studies in rodents either do not take self-reinoculation into account, or assume it can be eliminated by single housing of animals or housing them on wire mesh floors (also referred to as "wire screens" or "wire grids") [25]. Despite classical literature suggesting these assumptions can be incorrect [27], [32], [39]–[43], they have not been tested on mice housed in modern facilities using state-of-the-art quantitative tools.

Here, we explicitly test these assumptions about murine self-reinoculation to answer the following three questions relevant to gastrointestinal microbiome research: (1) Do quantitative 16S rRNA gene amplicon sequencing tools detect differences in small-intestine microbial loads between mice known to be coprophagic and non-coprophagic? (2) Does coprophagy impact the microbial composition of the small intestine? (3) Do differences in microbiota density and composition associated with self-reinoculation in mice impact the

microbial function (e.g., alter microbial metabolite production or modifications) in the small intestine?

To answer these questions, we analyzed gastrointestinal samples from mice under conditions known to prevent coprophagy (fitting with "tail" or "fecal collection" cups [27], [34], [37], [41], [44]) and typical laboratory conditions in which mice are known to be coprophagic (housing in standard cages). We also included samples from single-housed mice in standard and wire-floor cages. We analyzed the quantitative and compositional changes in the microbiome along the entire length of the mouse GIT in response to self-reinoculation, computationally inferred the changes in microbial function, and evaluated the microbial function-related metabolite profiles in the corresponding segments of the gut.

RESULTS

We first performed a pilot study to confirm that preventing coprophagy in mice would result in decreased viable microbial load and altered microbiota composition in the small intestine. We used a most probable number (MPN) assay utilizing anaerobic BHI-S broth medium to evaluate the live (culturable) microbial loads along the entire GIT of mice known to be coprophagic (housed in standard cages in groups, N = 5) and mice known to be non-coprophagic (fitted with tail cups and housed in standard cages in groups, N = 5). Consistent with the published, classical literature [31], [35], we found that coprophagic mice had significantly higher loads of culturable microbes in their upper GIT than mice that were non-coprophagic (Fig. S4A). Moreover, the microbial community composition in the proximal GIT, particularly in the stomach, of coprophagic mice more closely resembled the microbial composition of the large intestine (Fig. S4B) as revealed by 16S rRNA gene amplicon sequencing (N = 1 mouse analyzed from each group) and principal components analysis (PCA) of the resulting relative abundance data.

This pilot study confirmed that in our hands tail cups were effective at preventing the self-reinoculation of viable fecal flora in the upper GIT of mice. These results spurred us to design

a rigorous, detailed study (Fig. 1) to answer the three questions posed above using state-of-the-art methods: quantitative 16S rRNA gene amplicon sequencing (to account for both changes in the total microbial load and the unculturable taxa), quantitative functional gene content inference, and targeted bile-acid metabolomics analyses.

The study design (Fig. 1) consisted of six cages of four animals each that were co-housed for 2-6 months and then split into four experimental groups and singly housed for 12-20 days. The four experimental conditions were: animals fitted with functional tail cups (TC-F) and singly housed in standard cages, animals fitted with mock tail cups (TC-M) and singly housed in standard cages, animals singly housed on wire floors (WF), and control animals singly housed in standard conditions (CTRL). At the end of the study, gastrointestinal contents and mucosal samples were collected from all segments of the GIT of each animal and we evaluated total microbial loads (entire GIT) and microbiome composition (stomach (STM), jejunum (SI2), and cecum (CEC)).

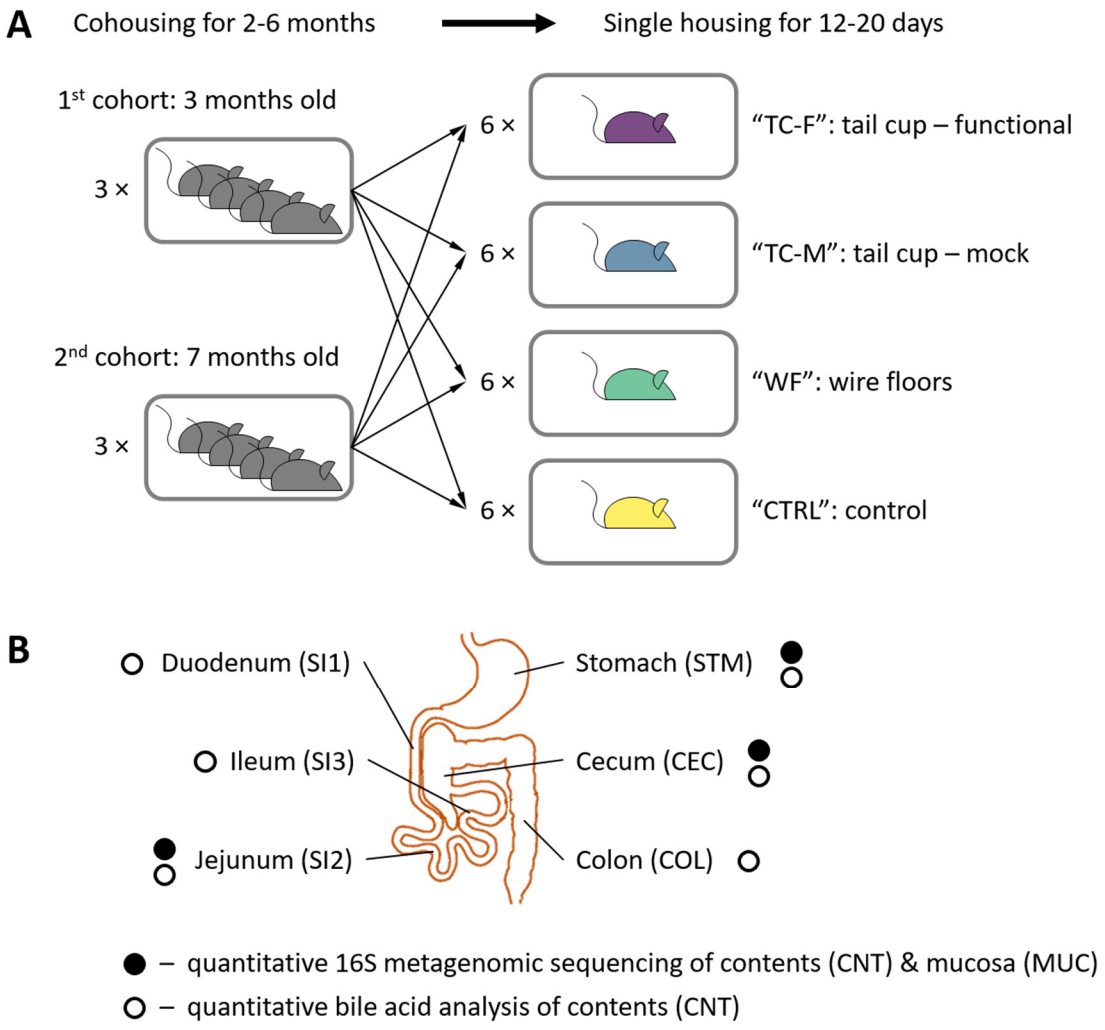

Fig. 2.1. An overview of the study design and timeline. (A) Mice from two age cohorts (3-months-old and 7-months-old) were raised co-housed (four mice to a cage) for 2-6 months. One mouse from each cage was then assigned to one of the four experimental conditions: (functional tail cups (TC-F), mock tail cups (TC-M), housing on wire floors (WF), and controls housed in standard conditions (CTRL). All mice were singly housed and maintained on each treatment for 12-20 days (N = 24, 6 mice per group). **(B)** Samples were taken from six sites throughout the gastrointestinal tract. Each sample was analyzed by quantitative 16S rRNA gene amplicon sequencing of lumenal contents (CNT) and mucosa (MUC) and/or quantitative bile-acid analyses of CNT. Panel B is adapted from [24], [45]).

We chose the cecum segment of the large intestine for quantitative 16S rRNA gene amplicon sequencing because the analysis of the contents of this section can provide a complete snapshot of the large-intestine and fecal microbial diversity in response to environmental factors [46]–[48]. Cecal contents also enabled us to collect a more consistent amount of sample from all animals across all experimental conditions (whereas defecation may be inconsistent among animals at the time of terminal sampling).

Self-reinoculation increases microbial loads in the upper gut

To answer our first question (Can quantitative sequencing tools detect the difference in 16S rRNA gene DNA copy load in the upper GIT of mice known to be coprophagic and non-coprophagic?), we analyzed total quantifiable microbial loads across the GIT using 16S rRNA gene DNA quantitative PCR (qPCR) and digital PCR (dPCR). Preventing self-reinoculation in mice equipped with functional tail cups dramatically decreased the lumenal microbial loads in the upper GIT but not in the lower GIT (Fig. 2A). Total quantifiable microbial loads in the upper GIT were reduced only in mice equipped with functional tail cups. All other experimental groups of singly-housed animals (those equipped with mock tail cups, housed on wire floors, or housed on standard woodchip bedding) that retained access to fecal matter and practiced self-reinoculation had similarly high microbial loads in the upper GIT, as expected from the published literature [27], [32], [39]–[43].

Across all test groups, mucosal microbial loads in the mid-small intestine demonstrated high correlation (Pearson's R = 0.84, $P = 2.8 \times 10^{-7}$) with the microbial loads in the lumenal contents (Fig. 2B).

Stomach (STM) and small-intestine (SI1, SI2, and SI3) samples from one (out of six) of the TC-F mice showed higher microbial loads compared with the other TC-F mice. The total microbial load in the upper GIT in this TC-F mouse was similar to mice from all other groups

(TC-M, WF, CTRL), which emphasizes the crucial importance of performing analyses of both microbial load and composition (discussed below) on the same samples.

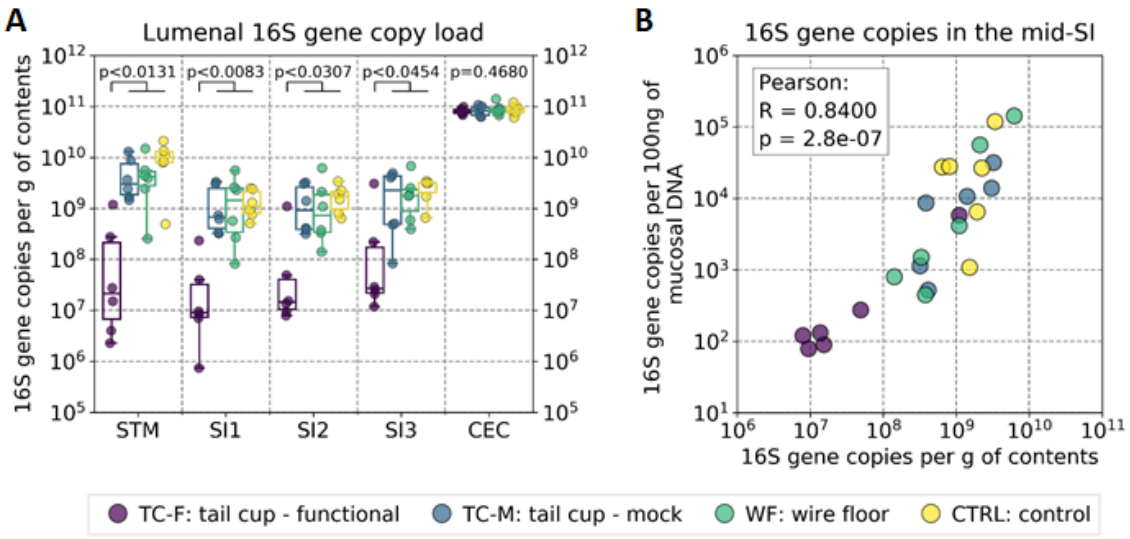

Fig. 2.2. Quantification of microbial loads in lumenal contents and mucosa of the gastrointestinal tracts (GIT) of mice in the four experimental conditions: (functional tail cups (TC-F), mock tail cups (TC-M), housing on wire floors (WF), and controls housed in standard conditions (CTRL). (A) Total 16S rRNA gene DNA copy loads, a proxy for total microbial loads, were measured along the GIT of mice of all groups (STM = stomach; SI1 = upper third of the small intestine (SI), SI2 = middle third or the SI, SI3 = lower third of the SI roughly corresponding to the duodenum, jejunum, and ileum respectively; CEC = cecum; COL = colon). Multiple comparisons were performed using a Kruskal–Wallis test, followed by pairwise comparisons using the Wilcoxon–Mann–Whitney test with false-discovery rate (FDR) correction. Individual data points are overlaid onto box-and-whisker plots; whiskers extend from the quartiles (Q2 and Q3) to the last data point within 1.5 × interquartile range (IQR). **(B)** Correlation between the microbial loads in the lumenal contents (per g total contents) and in the mucosa (per 100 ng of mucosal DNA) of the mid-SI. N = 6 mice per experimental group.

Self-reinoculation substantially alters the microbiota composition in the upper gut but has less pronounced effects in the large intestine

To answer our second question (does self-reinoculation with fecal microbiota impact upper GIT microbial composition?), we performed quantitative 16S rRNA gene amplicon sequencing [49], [50] on stomach (STM), jejunum (SI2), and cecum (CEC) samples. Qualitative sequencing revealed dramatic overall changes in the upper GIT microbiota caused by self-reinoculation (Fig. 3). An exploratory PCA performed on the multidimensional absolute microbial abundance profiles highlights the unique and distinct composition of the upper GIT microbiome of non-coprophagic mice (Fig. 3A). It is noteworthy that the stomach (STM) and small-intestine (SI2) microbiota in all coprophagic mice clustered closer to the large-intestine microbiota, suggesting the similarity was due to persistent self-reinoculation with the large-intestine microbiota (Fig. 3A).

Self-reinoculation had differential effects across microbial taxa (Fig. 3C), which could be classified into three main categories depending on the pattern of their change:

1. "Fecal taxa" (e.g., *Clostridiales*, *Bacteroidales, Erysipelotrichales*) that either dropped significantly or disappeared (fell below the lower limit of detection [LLOD] of the quantitative sequencing method [49], [50]) in the upper GIT of non-coprophagic mice;

2. "True small-intestine taxa" (e.g., *Lactobacillales*) that remained relatively stable in the upper GIT in non-coprophagic mice;

3. Taxa that had lower absolute abundance in the cecum (e.g., *Bacteroidales*, *Erysipelotrichales*, *Betaproteobacteriales*) of non-coprophagic (compared with coprophagic) mice.

Overall, the composition of the small-intestine microbiota of coprophagic mice was consistent with that previously reported in literature [46]. The upper-GIT microbiota in non-

coprophagic mice was dominated by *Lactobacilli* (Fig. 3C), known to be a prominent microbial taxon in human small-intestine microbiota [14], [51], [52]. Importantly, the compositional analysis showed that the single TC-F mouse that had high microbial loads in its stomach and small intestine had a microbial composition in those segments of the GIT similar (i.e., dominated by *Lactobacillales*) to all other TC-F mice, and very distinct from all coprophagic mice (Fig. 3B,C). The PCA showed that the stomach and mid-small intestine of this mouse clustered with the stomach and mid-small intestine of all other TC-F mice (Fig. 3A).

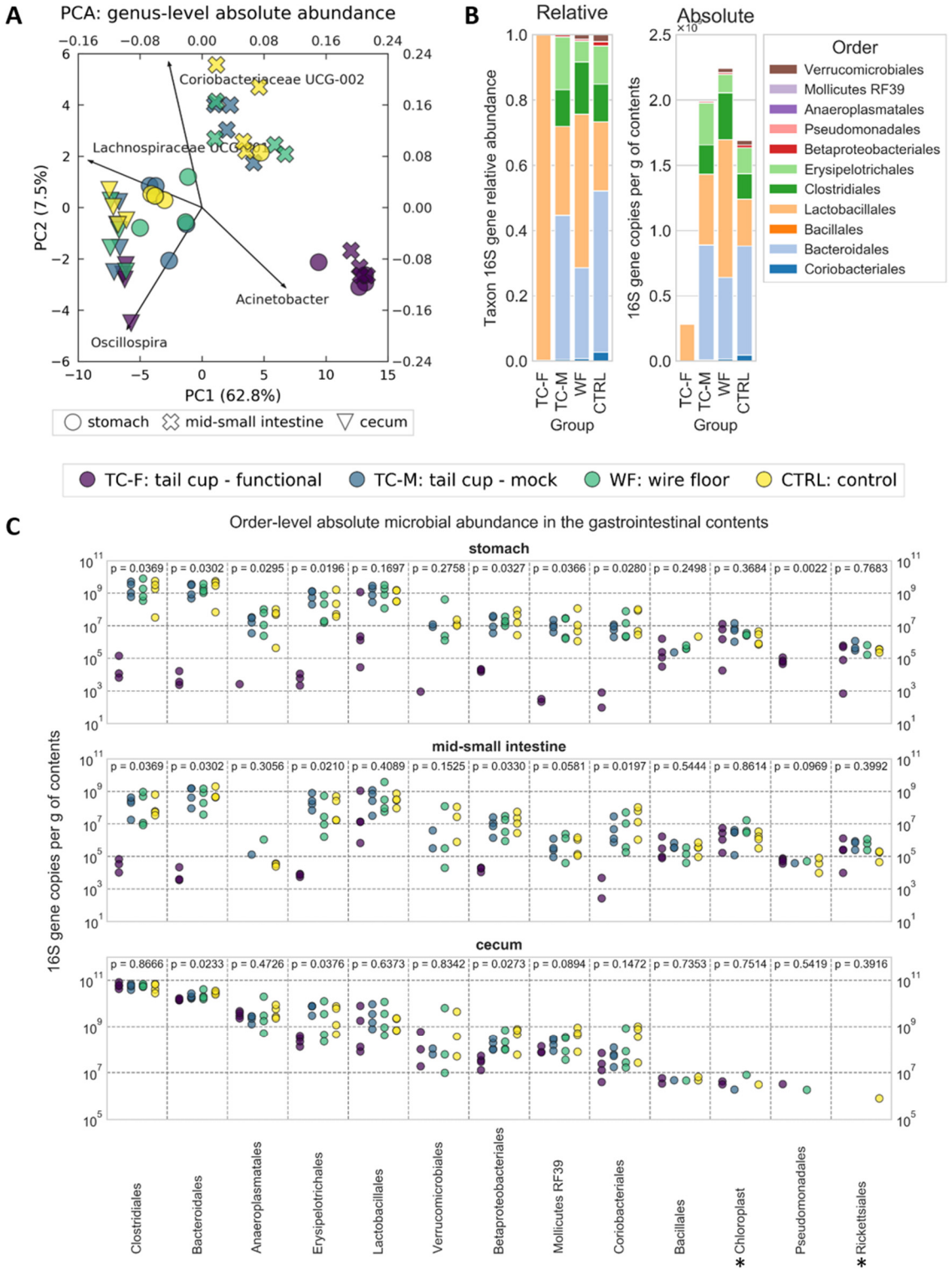

Fig. 2.3. Compositional and quantitative 16S rRNA gene amplicon sequencing of the gut microbiota. (A) Principal components analysis (PCA) of the $\log_{10}$-transformed and

standardized (mean = 0, S.D. = 1) absolute microbial abundance profiles in the stomach, mid-small intestine, and cecum. Loadings of the top contributing taxa are shown for each principal component. **(B)** Mean relative and absolute abundance profiles of microbiota in the mid-SI (order-level) for all experimental conditions. Functional tail cups (TC-F), mock tail cups (TC-M), housing on wire floors (WF), and controls housed in standard conditions (CTRL). N = 6 mice per experimental group, 4 of which were used for sequencing. **(C)** Absolute abundances of microbial taxa (order-level) compared between coprophagic and non-coprophagic mice along the mouse GIT. *Chloroplast and *Richettsiales (mitochondria) represent 16S rRNA gene DNA amplicons from food components of plant origin. Multiple comparisons were performed using the Kruskal–Wallis test.

Changes in the small-intestine microbiota lead to differences in inferred microbial functional gene content

We hypothesized that the quantitative and qualitative changes in the small-intestine microbiota induced by self-reinoculation may result in altered microbial function [53], [54] and an altered metabolite profile, either indirectly, as a result of functional changes in the microbiota, or directly via re-ingestion of fecal metabolites. To understand how such alterations to microbiota would impact microbial function in the small intestine, we next aimed to predict how the absolute abundances of functional microbial genes would be affected. We coupled the pipeline for microbial functional inference based on the 16S rRNA marker gene sequences (PICRUSt2) [55], [56] with our quantitative 16S rRNA gene amplicon sequencing approach [49], [50]. We focused our analysis on microbial functions that would be highly relevant to small-intestine physiology: microbial conversion of host-derived bile acids and microbial modification of xenobiotics.

We found that the inferred absolute abundances of a number of microbial gene orthologs implicated in enzymatic hydrolysis of conjugated bile acids (bile salt hydrolase, BSH [57]–[59]) and xenobiotic conjugates (e.g., beta-glucuronidase, arylsulfatase [60], [61]) in the stomach and the small intestine of coprophagic mice were dramatically higher (in some cases

by several orders of magnitude) than in non-coprophagic mice (Fig. 4). This difference was not observed in the cecum.

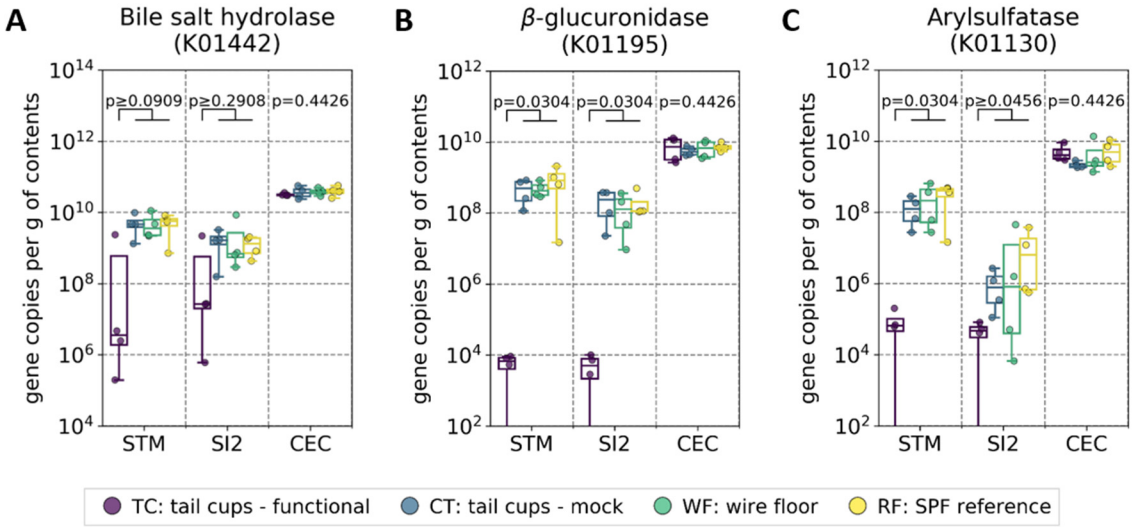

Fig. 2.4. Inference of microbial genes involved in bile-acid and xenobiotic conjugate modification along the GIT of coprophagic and non-coprophagic mice. Inferred absolute abundance of the microbial genes encoding (A) bile salt hydrolases (cholylglycine hydrolases), (B) beta-glucuronidases, and (C) arylsulfatases throughout the GIT (STM = stomach; SI2 = middle third of the small intestine (SI) roughly corresponding to the jejunum; CEC = cecum). KEGG orthology numbers are given in parentheses for each enzyme. In all plots, individual data points are overlaid onto box-and-whisker plots; whiskers extend from the quartiles (Q2 and Q3) to the last data point within 1.5 × interquartile range (IQR). Multiple comparisons were performed using the Kruskal–Wallis test; pairwise comparisons were performed using the Wilcoxon–Mann–Whitney test with FDR correction. N = 4 mice per group.

Changes in the small-intestine microbiota induced by self-reinoculation alter the bile acid profile

Bile acids are a prominent class of host-derived compounds with multiple important physiological functions and effects on the host and its gut microbiota [62], [63]. These host-derived molecules are highly amenable to microbial modification in both the small and large intestine [64]. The main microbial bile-acid modifications in the GIT include deconjugation, dehydrogenation, dehydroxilation, and epimerization [63]. Thus, we next performed quantitative bile acid profiling along the entire GIT to evaluate the effects of self-reinoculation on bile acid composition.

The small intestine is the segment of the GIT that harbors the highest levels of bile acids (up to 10 mM) and where they function in lipid emulsification and absorption [65]–[67]. Given these high concentrations of bile acid substrates, we specifically wished to analyze whether the differences we observed in small-intestine microbiota (Fig. 2, 3) between coprophagic and non-coprophagic mice would result in pronounced effects on microbial deconjugation of bile acids. We also wished to test whether any differences in bile-acid deconjugation were in agreement with the differences in the absolute BSH gene content we inferred (Fig. 4A) from the absolute microbial abundances (Fig. 3C).

We first confirmed that in all four experimental groups, total bile acids levels (conjugated and unconjugated; primary and secondary) across all sections of the GIT were highest in the small intestine (Fig. 5A). We then compared the levels of conjugated and unconjugated (Fig. 5B) as well as primary (host-synthesized) and secondary (microbe-modified) bile acids (Fig. S5) between coprophagic and non-coprophagic mice.

Across all sections of the GIT and in bile, non-coprophagic mice (TC-F) had significantly lower levels of unconjugated bile acids compared with coprophagic mice (Fig. 5B). Consistent with the computational inference in Fig. 4A (performed on mid-SI samples only), in all three sections of the small intestine of non-coprophagic mice (TC-F), the levels of unconjugated bile acids were substantially lower than in coprophagic mice. Almost 100% of

the total bile acid pool remained in a conjugated form in the small intestine of non-coprophagic mice.

In all groups of coprophagic mice (TC-M, WF, and CTRL) the fraction of unconjugated bile acids gradually increased from the proximal to distal end of the small intestine. Gallbladder bile-acid profiling (Fig. 5B) confirmed that bile acids were secreted into the duodenum predominantly in the conjugated form in all coprophagic mice. This pattern is consistent with the hypothesis that the exposure of bile acids to microbial deconjugation activity increases as they transit down a small intestine with high microbial loads (Fig. 2A) [65].

In the large intestine, non-coprophagic (TC-F) mice carried a smaller fraction of unconjugated bile acids compared with all coprophagic experimental groups (Fig. 5B).

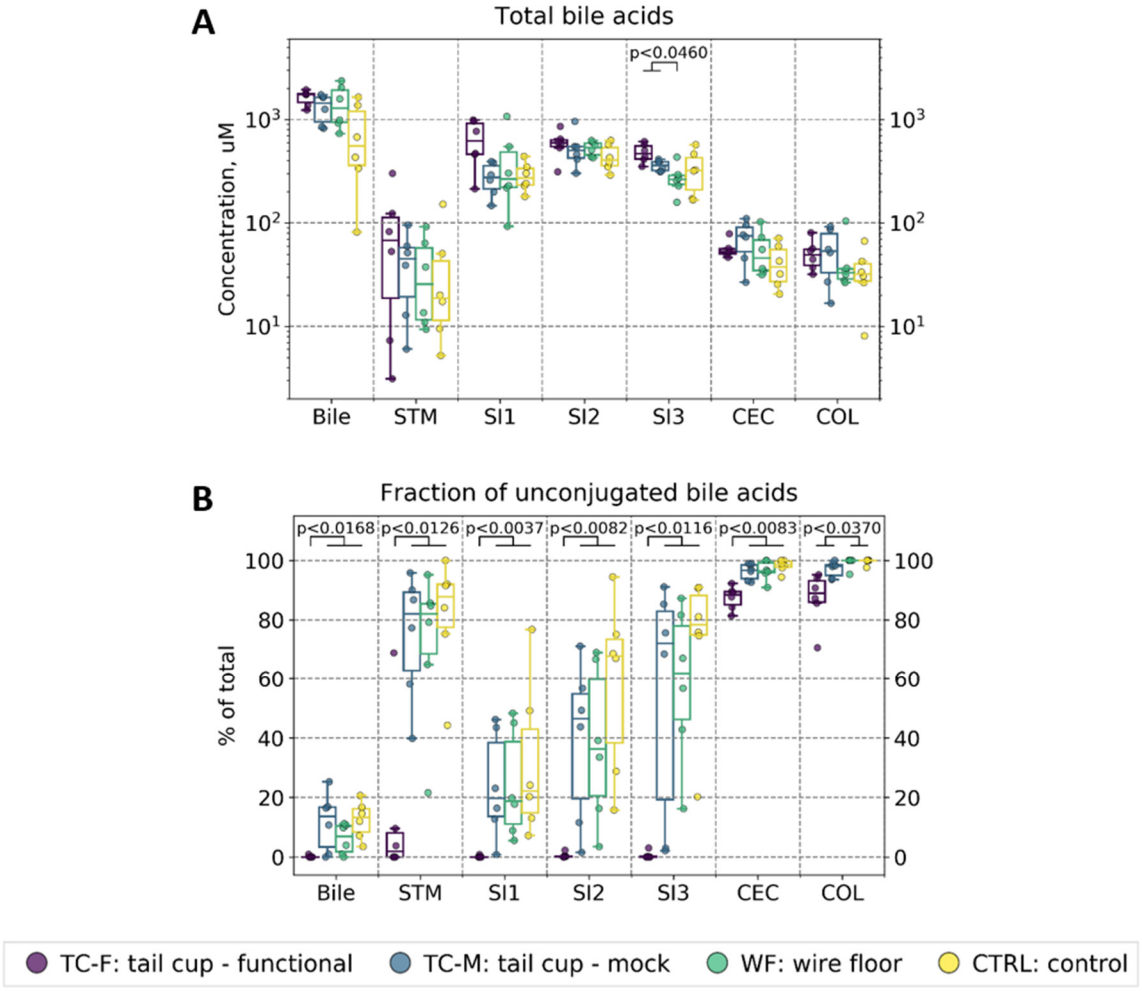

Fig. 2.5. Bile acid profiles in gallbladder bile and in lumenal contents along the entire GIT. (A) Total bile acid levels (conjugated and unconjugated; primary and secondary) and **(B)** the fraction of unconjugated bile acids in gallbladder bile and throughout the GIT (STM = stomach; SI1 = upper third of the small intestine (SI), SI2 = middle third or the SI, SI3 = lower third of the SI roughly corresponding to the duodenum, jejunum, and ileum respectively; CEC = cecum; COL = colon). In all plots, individual data points are overlaid onto box-and-whisker plots; whiskers extend from the quartiles (Q2 and Q3) to the last data point within 1.5 × interquartile range (IQR). Multiple comparisons were performed using the Kruskal–Wallis test; pairwise comparisons were performed using the Wilcoxon–Mann–Whitney test with FDR correction. N = 6 mice per group.

Bile acid deconjugation in the small intestine of coprophagic mice was uniform for all glyco- and tauro-conjugates of all primary and secondary bile acids measured in our study, suggesting a broad-specificity BSH activity was provided by a complex fecal flora in the small intestine of those animals.

In gallbladder bile and across all segments of the GIT from the stomach to the cecum, non-coprophagic mice had a statistically significantly lower fraction of total secondary bile acids (conjugated and unconjugated) than coprophagic mice (Fig. S5). This change was uniform for the entire secondary bile acid pool of those analyzed. The only segment of the gut in which the difference in the fraction of secondary bile acids was not statistically significant between coprophagic and non-coprophagic mice was the colon. In fact, the differences in the fractions of total unconjugated and total secondary bile-acids between coprophagic and non-coprophagic mice would have gone largely undetected had we only analyzed colonic contents or stool. These findings further highlight the importance of the comprehensive spatial interrogation of the complex crosstalk between the microbiota and bile acids in the gastrointestinal tract.

Discussion

In this study, we used modern tools for quantitative microbiota profiling and showed that when self-reinoculation with fecal flora is prevented, the mouse small intestine harbors dramatically lower densities of microbiota and an altered microbial profile. Consistent with published literature [27], [32], [39]–[43], we confirmed that single housing on wire floors failed to prevent mice from practicing coprophagy and that only functional tail cups reliably prevented the self-reinoculation with fecal flora.

Despite its effectiveness, the tail cup approach has limitations. Tail cups in their current design may not be suitable for female rodents due to anatomical differences leading to urine entering and remaining inside the devices [68]. Animals need to be singly housed to prevent them from gnawing on each other's tail cups and causing device failure or injury. The tail

cup approach may be hard to implement in younger and actively growing mice (e.g., before or around weaning). Some mice in our study developed self-inflicted skin lesions from over-grooming at the location where the tail cups come in contact with the body at the animal's hind end. Thus, we concluded that the approach in its current implementation is limited to 2-3 weeks in adult animals.

Our device design reduced the risk of tail injury and necrosis described in previous works [44] and allows for emptying the cups only once every 24 hours to reduce handling stress. Because host stress can affect the microbiota [69] and other physiological parameters, we included a mock tail-cup group. Both TC-F and TC-M mice demonstrated a similar degree of weight loss (Fig. S3A) when compared with the WF and CTRL mice despite similar food intake rates across all four groups (Fig. S3B). Mice fitted with mock tail cups (TC-M) had microbial patterns and bile acids profiles similar to control mice (CTRL), thus the effects we observed in non-coprophagic mice are not attributable to stress.

We believe that the tail cup approach is implementable in gnotobiotic settings (e.g., flexible film isolators and individually ventilated cages), which can aid studies that involve association of mice with defined microbial communities or with human-derived microbiota.

The non-coprophagic mouse model may be more relevant to humans

Using quantitative microbiota profiling, our study demonstrated that preventing self-reinoculation dramatically reduced the total levels of several prominent taxonomical groups of obligate anaerobes (e.g., *Clostridiales*, *Bacteroidales*, *Erysipelotrichale*) in the upper gastrointestinal microbiota of conventional mice. Despite these differences in taxa, levels of *Lactobacillales* in the small intestine and cecum, but not in the stomach, remained similar between coprophagic and non-coprophagic animals (Fig. 3C). The physiological significance of the maintained persistent population of *Lactobacillales* in the upper gastrointestinal tract (e.g., stomach or small intestine) and their overall consistent presence along the entire GIT [25], [70] for the host is not fully understood. However, *Lactobacilli*

colonization in the stomach and small intestine has been shown to promote resistance to colonization by pathogens (reviewed in [71], [72]).

Compared with conventional (coprophagic) mice, the non-coprophagic mice displayed features of the small-intestine microbiota and bile acid profiles that are more similar to the patterns seen in the small intestine of humans: orders of magnitude lower microbiota density, reduced abundance of obligate anaerobic flora and dominance of *Lactobacillales*, and a higher ratio of conjugated bile acids. These findings highlight the need to understand and control self-reinoculation in mouse models used to answer questions relevant to host-microbiota interactions in human health.

Self-reinoculation and microbial ecology in the mouse GIT

We observed that within the approximately two-week timeframe of our study, the taxonomical diversity of the mouse large-intestine microbiome was stable in the absence of persistent microbial self-reinoculation: all taxonomical groups at the order level observed in the cecum of coprophagic mice were present in the cecum of non-coprophagic mice, and vice versa.

The trending changes in the absolute abundances of several taxa in the large intestine of non-coprophagic mice may be the result of eliminated self-reinoculation and/or the consequence of the altered profile of bile acids entering the cecum from the small intestine. It has been previously suggested that the degree of bile acid deconjugation may alter the microbiota profile [57].

Stability of complex microbiomes in response to perturbations with and without continuous species reintroduction is an important subject of research in microbial ecology [73], [74]. Eliminating fecal ingestion provides a way to study stability and recovery of the mouse gut microbiota (e.g., in response to dietary change or antibiotic exposure [75]) in a way more

relevant to modern humans. Thus, non-coprophagic mouse model can significantly aid such research.

Self-reinoculation with fecal flora leads to altered bile acid profiles in the GIT

We demonstrated that changes to small-intestine microbiota density and composition had pronounced effects on microbial function resulting in increased bile acid deconjugation in that segment of the GIT.

Bile acid deconjugation is a microbiota-mediated process that in healthy humans is conventionally believed to take place in the distal small intestine (ileum) and in the large intestine [76] such that sufficient lipid emulsification (with conjugated bile acids) and absorption can take place in the small intestine by the time digesta reaches the ileum [77]. As a result of the much higher bile acid concentrations in the small intestine compared with the large intestine, altered deconjugation of bile acids in the small intestine may have more wide-ranging effects on the entire enterohepatic system. Our data indicate that bile acid deconjugation can take place in any segment of the small intestine of conventional healthy mice as a function of the microbial density and composition (Fig. 2A, 3, 5B), which is consistent with previous findings in animal models and in humans with small-intestinal microbial overgrowth (SIBO) [78]–[82].

Strikingly, the very low degree of bile acid deconjugation in the small intestine of non-coprophagic mice in our study resembles profiles seen in germ-free animals [83]–[85], gnotobiotic animals colonized only with microbes incapable of deconjugating bile acids [86]–[89], and antibiotic-treated animals [90]–[92]. Our observations suggest a mechanistic link between the small-intestine microbiota density and composition and the bile acid modification in this segment of the GIT. The small intestine of healthy human subjects is believed to harbor bile acids predominantly in the conjugated form [93], which further substantiates that (compared with coprophagic mice) the small intestine of non-coprophagic mice is more similar to the small intestine of a healthy human.

Although microbiota density and composition in the large intestine of coprophagic and non-coprophagic mice were largely similar, non-coprophagic mice had a higher fraction of bile acids that remained in the conjugated form in the large intestine (Fig. 4B), likely as a result of the bile acids entering the large intestine from the ileum predominantly in a conjugated form. Additionally, across all study groups, the total concentrations of bile acids in the small intestine were ~10-fold greater than in the large intestine. We therefore infer that in coprophagic mice a greater absolute amount of bile acids underwent deconjugation in the small intestine than in the large intestine, i.e., in coprophagic mice, the small intestine contaminated with high loads of fecal flora was the primary site of bile acid deconjugation.

Regulation of bile acid deconjugation activity in the gut is considered a potential health-promoting modality in a number of contexts, including lowering blood cholesterol levels (reviewed in [94]–[96]). BSH-active probiotics can be a promising delivery vehicle for promoting increased bile acid deconjugation in the gut. Our study emphasizes the importance of controlling for self-reinoculation when using mice to study the effects of BSH-active microbial strains or probiotics [59], [97]–[102] (especially those with high selectivity for particular bile acid conjugates [58], [86], [89]) because conventional (coprophagic) mice already have pronounced BSH activity in their small intestines. A non-coprophagic mouse may be a better animal model in such studies**.**

Our findings also have implications for the use of conventional (coprophagic) mice in diet studies. Deconjugated bile acids are less effective than conjugated at lipid emulsification and fat micelle formation [78], [103]. Increased bile acid deconjugation in the small intestine of animals and humans can lead to lipid malabsorption and fat-soluble vitamin deficiency and in extreme scenarios even to steatorrhea [81], [104]. Previous research has shown that the small-intestine microbiota plays an important role in mediating the effect of high fat diets on the host [105]; our results suggest that future studies of the microbiota-mediated effects of high fat diets need to consider increased microbial bile acid deconjugation in the mouse intestine due to self-reinoculation with fecal flora.

Bile acid deconjugation is considered to be obligatory [88], [106], [107] before the secondary bile acid metabolism (believed to be predominantly occurring in the large intestine [76]) can take place. These reactions in many cases are carried out by different members of the microbiota. Thus, the reduction of the deconjugation activity in the small intestine of non-coprophagic mice and consequently lower availability of free primary bile acids to further microbial modification can explain the decrease in the secondary bile acid fraction (percentage of all bile acids) in the bile acid pool across the GIT and gallbladder bile of non-coprophagic mice in our study. A similar but more pronounced trend has been observed in rabbits [108]. Reduced oral intake and recycling of fecal secondary bile acids as a result of eliminating coprophagy may also be a contributing factor to the lower fraction of secondary bile acids in the total bile acid pool in the enterohepatic circulation in these animals.

Total bile acid levels in the stomach were similar in coprophagic and non-coprophagic mice (and agree with literature [108], [109]), however bile acid profiles (including the fraction of total unconjugated and total secondary bile acids) were substantially different. Surprisingly, in all coprophagic mice the fraction of unconjugated bile acids in the stomach appeared to be intermediate between the profiles in the small intestine and in the large intestine (Fig. 5B), suggesting that the bile acids in the stomach of coprophagic mice could be accumulating from bile acids re-ingested in feces and bile acids refluxed from the duodenum. This pattern was not observed in non-coprophagic mice, suggesting that coprophagy may alter the bile acid profile in the upper GIT both directly (via re-ingestion of fecal metabolites) and indirectly (via altered microbiota function).

Inferences about microbial function in bile acid and drug modification

Our quantitative functional gene inference analysis predicted differential absolute abundance of the BSH orthologs between the small intestine of coprophagic and non-coprophagic mice (Fig. 4A). This approach has limitations associated with incomplete gene annotations, limited ability to infer metagenomes from the marker gene sequences when multiple microbial strains with similar 16S rRNA gene sequences exist [55], [56], difficulty to predict the exact gene expression and enzyme activity and specificity. To test our prediction about the BSH

we employed the targeted bile acid metabolomic analysis of mouse gastrointestinal samples and observed the differences in the small-intestine bile acid deconjugation between coprophagic and non-coprophagic mice (Fig. 5B) that were in agreement with the differences in the inferred BSH gene abundances in the small intestine of those two types of animals (Fig. 4A).

We next explored the effects of self-reinoculation on the absolute abundance of microbial gene orthologs implicated in xenobiotic modification [110] in the small intestine as microbiota-dependent drug modification and toxicity in the small intestine have been previously observed in rodents [111]–[121]. Many drugs administered to humans and mice both via enteral and parenteral routes after reaching the systemic circulation are transformed by the liver into conjugates (e.g., glucuronic acid-, sulphate-, or glutathione-conjugates) and excreted with bile into the GIT lumen. Such transformations are believed to reduce the small-intestine reabsorption of xenobiotics and promote their excretion from the body with stool. Alterations in the small-intestine microbiota may also lead to increased hydrolysis of such conjugates by microbial enzymes and promote the local toxicity of the drug and enable its re-uptake from the small intestine (i.e., undergo enterohepatic circulation) [21], [119], resulting in an increase in the xenobiotic flux through the liver [122], [123] and to an overall microbiota-dependent change in drug pharmacokinetics.

As with the inferred differential BSH absolute abundances (correlating activity of which we confirmed with the bile acid deconjugation measurements), our analysis predicted differences in the absolute abundance (Fig. 4B, C) of the microbial gene orthologs responsible for drug conjugate hydrolysis (e.g., beta-glucuronidases, sulfohydrolases) between the small intestine of coprophagic and non-coprophagic mice. If this prediction is further experimentally confirmed, it would imply that self-reinoculation must be controlled for or taken into account when investigating the drug pharmacology in mice.

Relevance of self-reinoculation in probiotics research

Many studies on probiotics and their effects on host animal physiology rely on repeated oral administration of live probiotic microorganisms to rodents. Our study suggests that self-

reinoculation with live fecal flora in laboratory mice could both interfere with and introduce inconsistencies in live probiotic administration regimens. As has been stated earlier, particular attention should be given to self-reinoculation and its effects on the small-intestine bile acid profile in studies aiming to evaluate the health effects of probiotics and other therapeutic modalities [59], [94]–[102] targeting bile acid deconjugation and metabolism.

Relevance of mouse models in human microbiota research

The role of mouse models in human microbiota research remains a subject of a debate [24], [25], [124]. At the same time, the field is recognizing the importance of reproducibility in gut microbiota research that uses mouse models [69], [124]. Several recent studies have highlighted the variability in lab-mouse microbiota related to animal strains and sources of origin [47], [125]–[129]. Others have attempted to catalog "normal" or "core" gut microbiome [130], [131] and its spatial organization [46], [47] and function [132] in laboratory and wild mice. Recently, the small-intestine microbiome has become the focus of studies conducted in mice in the context of host physiology [105] and disease [15], [133]. Yet, little attention has been given to the impact of self-reinoculation on the gut microbiota spatial structure and function or to how study outcomes might be affected by controlling (or not controlling) for this experimental parameter in mouse models.

Self-reinoculation in rodents may affect not only their native microbiota, but also individual microbial colonizers [35] (e.g., in gnotobiotic animals) and complex xenomicrobiota (e.g., in human microbiota-associated (HMA) mice). HMA mice have emerged as an important research model for dissecting the mechanistic connection between the gut microbiota and the host phenotype in health and disease, even though the field acknowledges its limitations [134], [135]. Compositional differences between the small-intestine and large-intestine microbiomes in primates and humans [23], [51], [52] appear to be more substantial than those reported for laboratory mice [46], [132]. Our study emphasizes that the compositional similarity between small- and large-intestine microbiota in conventional laboratory mice can be a result of self-reinoculation with fecal flora. Thus, the effects of self-reinoculation on the

spatial organization and function of human microbiota in HMA mice warrant future exploration.

In conclusion, this study uses modern tools to demonstrate the importance of self-reinoculation in the context of microbial ecology and function within the mammalian gastrointestinal system. Our work highlights the importance of recognizing and properly controlling for self-reinoculation when murine studies analyzing small-intestine microbiota and its function intend to draw parallels with human physiology and pathophysiology. Additionally, spatial interrogation of the gut microbiota and its function in mouse models is important because even dramatic changes in the small-intestine microbiome profile, function, and metabolome may be overlooked if only large-intestine and stool samples are analyzed.

METHODS

Experimental animals

All animal handling and procedures were performed in accordance with the California Institute of Technology (Caltech) Institutional Animal Care and Use Committee (IACUC).

C57BL/6J male specific-pathogen-free (SPF) mice were obtained at the age of 7-8 weeks from Jackson Laboratory (Sacramento, CA, USA) and housed four mice per cage. Two cohorts of animals were used: the first cohort was allowed to acclimate in the Caltech animal facility for 2 months and mice were 4 months old at the start of the study; the second cohort acclimated for 6 months and mice were 8 months old at the start of the study.

All animals were maintained on chow diet (PicoLab Rodent Diet 20 5053, LabDiet, St. Louis, MO, USA) and autoclaved water ad lib and subjected to a daily 13:11 light:dark cycle during acclimation and throughout the entire study. Mice were given measured amounts of food, and food intake during the experiment was measured by weighing the food during weekly cage changes and at the end time point for each animal. Body weight was measured at the start of the experiment, during weekly cage changes, and at the end time point.

Animal housing conditions

During the experiment, all mice were singly housed in autoclaved cages (Super Mouse 750, Lab Products, Seaford, DE, USA). The mice in the control (CTRL), mock tail cup (TC-M) and functional tail cup (TC-F) treatments were housed on heat-treated hardwood chip bedding (Aspen Chip Bedding, Northeastern Products, Warrensburg, NY, USA) and provided with tissue paper (Kleenex, Kimberly-Clark, Irving, TX, USA) nesting material. The mice in the wire-floor (WF) treatment were housed on raised wire floors with a mesh size of 3 × 3 per square inch (#75016, Lab Products) and provided with floorless paper huts (#91291, Shepherd Specialty Papers, Watertown, TN, USA). A thin layer of woodchip bedding was added under the wire floors to absorb liquid waste from the animals (Fig. S1D).

Tail cup design and mounting

We designed the tail cups based on published literature [41], [136]–[138], including the locking mechanism [41]. Each cup was locked in place around the hind end of animals by anchoring to a tail sleeve designed with a perpendicular groove. Such tail sleeves allow for the cup to be held snugly against the animal so that the total weight of the tail cup is distributed along a large surface area of the tail skin, which minimizes complications. When mounted, the tail cups can freely rotate along the longitudinal axis, which ensures the locking mechanism does not strangulate the tail.

We hand-made the tail cups from 20 mL syringes (#4200.000V0 Norm-Ject 20 mL Luer-Lock, Henke-Sass Wolf GmbH, Tuttlingen, Germany) as depicted on Fig. S1A-C. Multiple perforations were designed to accelerate desiccation of the captured fecal pellets. Lateral slits allowed for increasing the diameter of the locking edge; pressing on the slits with two fingers allowed tail cups to be quickly unfastened from tail sleeves. Mock tail cups were modified with wide gaps in the walls to allow the fecal pellets to fall out of the cup.

To prevent mice from gnawing on the plastic parts of the tail cups (which could create a jagged edge and lead to a subsequent injury), they were reinforced with metal flared rings made from stainless steel grommets (#72890, SS-4, C.S. Osborne, Harrison, NJ, USA) that were modified to reduce their size and weight. Metal rings were attached to tail cups using 4 mm-wide rubber rings cut from latex tubing (Amber Latex Rubber Tubing #62996-688, 1/2” ID, 3/4” OD; VWR, Radnor, PA, USA).

Tail sleeves were made from high-purity silicone tubing (HelixMark 60-411-51, 1/8" ID, 1/4" OD; Helix Medical, Carpinteria, CA, USA). The tubing was split longitudinally and a 2.0 mm wide strip of the wall was removed to accommodate for variable tail diameters among animals and along the tail length, to prevent uneven tail compression, and to facilitate uniform application of the tissue adhesive. The perpendicular tail-cup mounting groove was made using a rotary tool (Craftsman #572.610530, Stanley Black & Decker, New Britain,

CT, USA) equipped with a cutting disc (RD1, Perma-Grit Tools, Lincolnshire, UK). Each tail cup and sleeve together weighed approximately 4.12 g empty.

Before mounting the tail cups, animals were anesthetized with 10 min isoflurane and placed on a heating pad to maintain body temperature. Sleeves were de-greased on the inside using 70% ethanol and a veterinary tissue adhesive (GLUture Topical Adhesive #32046, Abbott Laboratories, Lake Bluff, IL, USA) was applied to the tail base. The adhesive was allowed to cure for 5 min and then tail cups were mounted. Mice were returned back to their cages and allowed to recover from the anesthesia and ambulate.

Tail cups were emptied of fecal pellets daily at 08:00 AM. Mice were prompted to enter a restrainer [139] made from a black polypropylene 50 mL conical tube (TB5000 LiteSafe, Cole-Parmer, Vernon Hills, IL, USA) and the tail cups were unclipped and quickly emptied. Any residue on the tail cup was cleaned using a paper towel and Rescue solution (Virox Technologies, Oakville, ON, Canada) prior to the cups being remounted. Animals fitted with the mock tail cups were subjected to the identical procedure to match the handling conditions.

Tail cups were mounted in animals for a duration of between 12 and 20 days. All TC-F animals were time-matched with TC-M animals, (i.e., each animal from the TC-F group had a time-matched animal from the TC-M group handled and euthanized at the same time).

Sample collection and treatment

All mice were euthanized as approved by the Caltech IACUC in accordance with the American Veterinary Medical Association Guidelines on Euthanasia [140]. Mice were euthanized while under isoflurane anesthesia (delivered via a calibrated gas vaporizer in an induction chamber followed by maintenance on a nose cone) via cardiac puncture followed by cervical dislocation. Blood was collected using a 1 mL syringe (#309659, Becton Dickinson) and 21G × 1” needle (#26414, EXELINT International, Redondo Beach, CA, USA).

Blood was immediately placed into K_2EDTA plasma separation tubes (MiniCollect 450480, Greiner Bio-One GmbH, Kremsmünster, Austria), gently mixed, and stored on ice for up 1 h prior to centrifugation. Bile and urine were collected directly from the gall and urinary bladders respectively using a 1-mL syringe (#4010.200V0 Norm-Ject 1 mL Tuberculin Luer, Henke-Sass Wolf GmbH) and 27G × 1/2" needle (#26400, EXELINT International) and stored on ice.

Fecal samples were collected if present at the time of euthanasia. The entire gastrointestinal tract was excised from the gastro-esophageal junction to the anal sphincter and stored on ice during processing.

Plasma separation:

Blood samples were centrifuged in the plasma separation tubes at 2000 RCF for 5 min at 4 °C. Plasma was separated and stored at -80 °C.

Processing of GIT contents

To prepare samples for the main experimental analyses (Fig. 2-4), each mouse GIT was split into stomach, three equal-length thirds of the small intestine, cecum, and colon. Contents from each segment of the GIT were flushed out using 2-5 mL of cold (4 °C) sterile autoclaved saline solution (0.9% NaCl (#S5886, Sigma-Aldrich) in ultrapure water (Milli-Q, MilliporeSigma, Burlington, MA, USA)) followed by very gentle squeezing with tweezers to avoid mucosal damage. All samples were stored on ice during processing.

An aliquot of each sample diluted in saline was concentrated by centrifugation at 25000 RCF for 10 min at 4*C. The supernatant was removed and the pellet was reconstituted in 9 volumes of 1× DNA/RNA Shield (DRS) solution (R1100-250, Zymo Research, Irvine, CA,

USA), mixed by vortexing and stored at -80 °C for future DNA extraction. Separate aliquots of each sample were stored at -80 °C for the metabolomic (bile acid) analysis.

Preparation of GIT contents for the MPN-based microbial quantification and 16S rRNA gene amplicon sequencing (pilot study; Fig. S4B) was the same as above, but conducted inside a vinyl anaerobic chamber (Coy Laboratory Products, Grass Lake, MI, USA) in an atmosphere of 5% hydrogen, 10% carbon dioxide, and 85% nitrogen. All samples were maintained on ice and immediately processed for the culture-based assay.

Preparation of GIT mucosa

After flushing its contents, each segment of the GIT was gently rinsed in sterile cold (~4 °C) saline, cut longitudinally, and placed flat on a glass slide. The mucosa was scraped from the tissue gently using a second clean glass slide. Glass slides (VistaVision #16004-422, VWR) were sterilized by dry heat sterilization at 200 °C for at least 2 h. Mucosal scrapings were collected and combined with 9 volumes of DRS solution, mixed by vortexing, and stored at -80 °C in preparation for DNA and RNA extraction.

Most probable number (MPN) assay

For the pilot study (Fig. S4A), the MPN assays (adapted from [141]–[145]) were performed on each GIT section (stomach, three sub-sections of the small intestine, cecum, and colon) from five mice fitted with functional tail cups and five control mice. The growth medium was brain-heart infusion broth (Bacto BHI, #237500, Becton Dickinson, Franklin Lakes, NJ, USA), prepared in ultrapure water (Milli-Q), sterilized by autoclaving, allowed to cool to room temperature, and supplemented with 1.0 mg/L vitamin K_1 (#L10575, Alfa Aesar, Haverhill, MA, USA), 5 mg/L hematin (#H3281, Sigma-Aldrich St. Louis, MO, USA), and 0.25 g/L L-cysteine (#168149, Sigma-Aldrich). The medium was allowed to equilibrate inside the anaerobic chamber for at least 24 hours before use.

MPN assays were performed in clear, sterile, non-treated polystyrene 384-well plates (Nunc 265202, Thermo Fisher Scientific, Waltham, MA, USA). Two series of eight consecutive 10-fold serial dilutions were prepared from each sample in sterile autoclaved saline solution (equilibrated inside the anaerobic chamber for at least 24 h) on clear sterile non-treated polystyrene 96-well plates (Corning Costar 3370, Corning, NY, USA). We injected 10 µL of each serial dilution from each series into four (eight total per dilution) culture-medium replicates (wells) filled with 90 µL of the BHI-S broth medium.

Plates were sealed with a breathable membrane (Breath-Easy BEM-1, Diversified Biotech) and incubated for 5 d at 37.0 °C inside the anaerobic chamber. The plates were lidless for the first 24 h to facilitate uniform gas equilibration, then from 24 h to the end of the incubation period (120 h), a plastic lid was kept over the plates.

At the end of the incubation, the plates were scanned using a flatbed scanner (HP ScanJet 8250, Hewlett-Packard, Palo Alto, CA, USA) in the reflective mode with black background at 300 dpi resolution. The positive wells (replicates) were called by visually observing each acquired high-resolution image. The MPN for each sample was calculated using Microsoft Excel with the "Calc_MPN" macro [146].

DNA extraction

DNA was extracted from thawed GIT contents and mucosal sample aliquots preserved in DRS solution with the ZymoBIOMICS DNA Miniprep Kit (D4300, Zymo Research) according to the manufacturer's instructions. Samples were homogenized on a bead-beater (MiniBeadBeater-16, Model 607, Bio Spec Products, Bartlesville, OK, USA) for 5 min at the default speed of 3450 RPM. Quantitative recovery of DNA across multiple orders of microbial loads in the samples was previously verified in [49].

DNA yield and purity in the extracts was evaluated via light absorbance (NanoDrop 2000c, Thermo Fisher Scientific) and via a fluorometric assay (Qubit dsDNA HS Assay Kit Q32854, Thermo Fisher Scientific) on a fluorometer (Invitrogen Qubit 3, Thermo Fisher Scientific).

Quantitative PCR (qPCR) for 16S rRNA gene DNA copy enumeration

The qPCR reactions were set up in triplicates for each DNA sample. A single replicate reaction volume of 15 µL contained 1.5 µL of the DNA extracts combined with the qPCR master mix (SsoFast EvaGreen Supermix, #172-5200, Bio-Rad Laboratories, Hercules, CA, USA), forward and reverse primers (synthesized by Integrated DNA Technologies, San Diego, CA, USA; Table S1) at a final concentration of 500 nM, and ultrapure water (Invitrogen UltraPure DNase/RNase-Free Distilled Water 10977-015, Thermo Fisher Scientific). Reactions were set up in white 96-well PCR plates (#HSP9655, Bio-Rad Laboratories) sealed with a PCR tape (#MSB1001, Bio-Rad Laboratories).

The standard curve was built for each qPCR run based on the included series of 10-fold dilutions of the "standard" SPF mouse fecal DNA extract (with the quantified absolute concentration of 16S rRNA gene copies using digital PCR).

Amplification was performed with real-time fluorescence measurements (CFX96 Real-Time PCR Detection System, Bio-Rad Laboratories). Thermocycling conditions were used according to Table S2. The qPCR data files were analyzed using Bio-Rad CFX Manager 3.1 (#1845000, Bio-Rad Laboratories) and the Cq data were exported to Microsoft Excel for further processing.

Digital PCR (ddPCR) for absolute 16S rRNA gene DNA copy enumeration

Droplet digital PCR (ddPCR) reactions were set up according to [49]. Single replicate reaction volume of 20 µL contained 2.0 µL of the DNA extracts combined with the ddPCR

master mix (QX200 ddPCR EvaGreen Supermix, #1864033, Bio-Rad Laboratories), forward and reverse primers (synthesized by Integrated DNA Technologies; Table S1) at final concentration of 500 nM each, and ultrapure water (Thermo Fisher Scientific).

Droplets were generated using DG8 cartriges (#1864008, Bio-Rad Laboratories), droplet generation oil (#1864006, Bio-Rad Laboratories), and DG8 gaskets (#1863009, Bio-Rad Laboratories) on a QX200 droplet generator (#1864002, Bio-Rad Laboratories) and analyzed using a QX200 Droplet Digital PCR System (#1864001, Bio-Rad Laboratories) using droplet reader oil (#1863004, Bio-Rad Laboratories). The ddPCR data files were analyzed using QuantaSoft Software (#1864011, Bio-Rad Laboratories) and the raw data were exported to Microsoft Excel for further processing.

Thermocycling conditions were used according to [49] and Table S3. Amplification was performed in PCR plates (#0030133374, Eppendorf, Hauppauge, NY, USA) sealed with pierceable heat seals (#1814040, Bio-Rad Laboratories) using PCR plate sealer (PX1, #1814000, Bio-Rad Laboratories) on a 96-deep well thermocycler (C1000 Touch, # 1841100), Bio-Rad Laboratories).

16S rRNA gene DNA amplicon barcoding for next generation sequencing (NGS)

PCR reactions was set up according to [49], in triplicates for each DNA sample. Single-replicate reaction volumes of 30 µL contained 3 µL of the DNA extracts combined with the PCR master mix (5PRIME HotMasterMix, #2200400, Quantabio, Beverly, MA, USA), DNA intercalating dye (EvaGreen, #31000, Biotium, Fremont, CA, USA) at the suggested by the manufacturer concentration (×1), barcoded forward and reverse primers (synthesized by Integrated DNA Technologies; Table S1) at final concentration of 500 nM each, and ultrapure water (Thermo Fisher Scientific).

Reactions were set up in 0.2 mL white PCR tubes (#TLS0851) with flat optical caps (#TCS0803, Bio-Rad Laboratories).

Thermocycling conditions were used according to [49] and Table S4. Amplification was performed with real-time fluorescence measurements (CFX96 Real-Time PCR Detection System, Bio-Rad Laboratories) and samples were amplified for a variable number of cycles until the mid-exponential (logarithmic) phase to maximize the amplicon yield and minimize artifacts related to over-amplification [2].

Digital PCR (ddPCR) for Illumina library quantification

Single replicate reaction volume of 20 uL contained 2.0 uL of the diluted amplicon sample ligated with the Illumina adapters, 10 uL of ddPCR master mix (QX200 ddPCR EvaGreen Supermix, #186-4033, Bio-Rad Laboratories), forward and reverse primers (synthesized by Integrated DNA Technologies; Table S1) targeting the Illumina P5 and P7 adapters respectively at the final concentration of 125 nM each, and ultrapure water (Invitrogen).

Thermocycling conditions were used according to Table S5. PCR amplification and droplet analysis were performed as above.

Barcoded sample quantification, pooling, library purification and quality control

Triplicates of each barcoded amplicon sample were combined. Each samples was diluted × 10^5-10^7-fold and the molar concentration of barcoded amplicons was quantified using a home-brew ddPCR library quantification assay and KAPA SYBR FAST Universal qPCR Library Quantification Kit (#KK4824, Kapa Biosystems, Wilmington, MA, USA) according to the manufacturer's instructions (the qPCR reaction was set up same as above).

Barcoded samples were pooled in equimolar amounts. Pooled library was purified using Agencourt AMPure XP beads (#A63880, Beckman Coulter, Brea, CA, USA) according to the manufacturer's instructions and eluted with ultrapure water (Invitrogen).

The purified library was confirmed to have the 260 nm to 280 nm light absorbance ratio of >1.8 using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific). The average amplicon size of approximately ~400 bp was confirmed with a High Sensitivity D1000 ScreenTape System (##5067-5584 and 5067-5585, Agilent Technologies, Santa Clara, CA, USA) using a 2200 TapeStation instrument (Agilent Technologies) and the Agilent 2200 TapeStation Software A02.01. (Agilent Technologies).

The molar concentration of the pooled library was measured using the ddPCR and KAPA qPCR assays and the library was submitted for next generation sequencing (NGS) with the sequencing primers described in Table S1.

Next generation sequencing

The library was sequenced on a MiSeq instrument (Illumina, San Diego, CA, USA) in a 300-base paired-end mode using a MiSeq Reagent Kit v3 (#MS-102-3003, Illumina). PhiX control spike-in was added at 15%.

PCR primer oligonucleotides (Table S1)

Same universal microbial 16S rRNA gene V4 primers (modified from [1], [3] and validated in [49], [50]) targeting the V4 region of the 16S rRNA gene from the 519 to 806 positions were used for 16S rRNA gene DNA copy quantification and multiplexed microbial community profiling based on 16S rRNA gene amplicon sequencing. Reverse barcoded primers for 16S rRNA gene DNA amplicon barcoding were according to [3].

Primers targeting the P5 and P7 Illumina adapters for barcoded amplicon and pooled library quantification using the ddPCR assay were according to [1], [3], [147]–[149].

Sequencing read processing

Demultiplexed 2 × 300 reads were processed using the Qiime2-2019.01 pipeline [150]. DADA2 plugin [151] was used to filter (forward trimming – 5, forward truncation – 230, reverse trimming – 5, reverse truncation – 160), denoise, merge the paired-end sequences, and remove the chimeras. Taxonomic sequence (amplicon sequence variant, ASV) classification was performed using the classifier (available for download from [152]) trained [153] on the V4 515-806 bp regions of 16S rRNA gene sequences from the Silva rRNA reference database, release 132 [4] (available for download from [154]).

Functional gene inference analysis with the PICRUSt2 [55], [56] was performed on the ASVs within the Qiime2 environment. Absolute and relative abundances of ASVs were normalized using the inferred 16S rRNA gene DNA copy counts. Obtained predicted metagenome data were used to calculate the normalized relative and absolute abundances of the gene orthologs of interest using Python tools (described below).

Sequencing data processing

Data handling, calculations, and statistical analyses were performed using Microsoft Excel with the Real Statistics Resource Pack [155], and the Python packages NumPy [156], Pandas [157], SciPy [158], Statsmodels [159]. Plotting was performed with Matplotlib [160] and Seaborn [161]. All Python packages were run using IPython [162] within Jupyter notebooks [163] distributed with the Anaconda environment [164].

Frequency data for the 16S rRNA gene ASVs assigned to taxa in each sample were converted to relative abundances for each sample. Relative abundances then were converted to absolute abundances using the corresponding values of total 16S rRNA gene DNA loads obtained from the qPCR and ddPCR assays for each sample.

Absolute abundance data were then collapsed to the genus (Fig. 3A) or order (Fig. 3B,C) taxonomical levels using a custom made Python function (confirmed to yield identical results

to the "collapse" method of the Qiime2 "Taxa" plugin [150]). We defined contaminating taxa (from sample handling during collection or from the DNA extraction kit or PCR reagents) using two methods: taxa that were not present in at least 1 out of 16 cecum contents samples (4 mice out of 6 from each group × 4 groups), and taxa identified with a frequency-based contaminant identification [165] implemented by us in Python. Data for chloroplasts and mitochondria of plant origin (likely from the chow diet) were kept in the dataset for Fig. 3A and 3C and removed for Fig. 3B. Mean absolute abundances of taxa for each group were calculated, converted to relative abundances, and plotted in Fig. 3B.

Principle component analysis (PCA) of the relative abundance data (Fig. S4B) was performed on centered log-ratio (CLR)-transformed [166], [167] (after a pseudocount equal to the minimal non-zero sequence count in the dataset was added to all zero values) genus-level relative abundance data using the Python Scikit-learn package [168].

PCA of the absolute abundance data (Fig. 3A) was performed on $\log_{10}$-transformed and centered-standardized (converted to normally-distributed data with mean = 0 and standard deviation = 1) [169] genus-level absolute abundance data using the Python Scikit-learn package [168].

Bile acid analysis

Reagents:

TαMCA, TβMCA, TωMCA, THCA, αMCA, βMCA, ωMCA, HCA, HDCA, MCA, GCDCA, GDCA, and GCA (Table S6) were obtained from Steraloids (Newport, Rhode Island, USA).

TCA, CA, DCA, TCDCA, TDCA, TUDCA, TLCA, CDCA, UDCA, LCA, D_4-TCA, D_4-DCA, D_4-CA, D_4-TDCA, D_4-GLCA, D_4-GUDCA, D_4-GCDCA, D_4-GCA, and D_4-GDCA (Table S6) were obtained from Isosciences (Ambler, PA, USA).

LC/MS grade acetonitrile (#A955-500), water (#W6500), and formic acid (#A117-50) were obtained from Thermo Fisher Scientific.

Sample preparation

To overcome sample buffering (pH issues), samples were extracted (using a protocol adapted and modified from [87]–[89]) in 9 volumes of ethanol with 0.5% formic acid and nine different heavy isotope (D_4) internal standards at 5 µM. D_4 internal standards were taurocholic acid (TCA), cholic acid (CA), deoxycholic acid (DCA), taurodeoxycholic acid (TDCA), glycocholic acid (GCA), glycolithocholic acid (GLCA), glycoursodeoxycholic acid (GUDCA), glycochenodeoxycholic acid (GCDCA), and glycodeoxycholic acid (GDCA). Samples were heated for one hour at 70°C with orbital shaking at 900 RPM. Solids were precipitated by centrifugation at 17000 RCF for 15 minutes at 4°C. Supernatants were decanted as 10% of the original sample (e.g. 100 µL of a 1 mL extraction sample) and evaporated at approximately 100 mTorr at RT on a rotovap (Centrivap Concentrator #7810016, Labconco, Kansas City, MO, USA). The evaporated samples were reconsistuted at 100x dilution from the original sample (e.g. 100 µL decanted solution is resuspended at 1 mL) in 20% acetonitrile, 80% water with 0.1% formic acid.

Due to small volumes, gall bladder bile samples were first diluted in 10 volumes of 100% ethanol (#3916EA, Decon Labs, King of Prussia, PA, USA). The ethanol-based dilutions were combined with 9 volumes of ultrapure water (Invitrogen) and subjected to extraction as above.

Each 10 µL extracted and reconsistuted sample injection was analyzed on a Waters Acquity UPLC coupled to a Xevo-qTOF Mass Spectrometer (Waters, Manchester, UK) using an Acquity UPLC HSS T3 1.8 micron, 2.1 × 100 mm column (# 186003539) and Acquity UPLC HSS T3 1.8 micron Guard Column (# 186003976). Needle wash was two parts isopropanol, one part water, and one part acetonitrile. Purge solvent was 5% acetonitrile in water. A pooled quality control sample was run every 8 injections to correct for drift in response.

Mass spectrometer instrument parameters were as follows: Capillary Voltage 2.4 kV, Collision Energy 6.0 eV, Sampling Cone 90V, Source Offset 40 V, Source 120 °C, desolvation gas temperature 550 °C, cone gas 50 L/Hr, and desolvation Gas 900 L/Hr. Time-of-flight mass spectra were collected in resolution mode, corresponding to 30000 m/Δm. The mass axis was calibrated with sodium formate clusters and locked using leucine enkephalin.

A seven point external calibration curve was collected three times within the run from 0.05 to 30 μM of the bile acid standards [0.05, 0.1, 0.5, 1, 5, 10, 30 μM]. External standards were taurocholic acid (TCA), tauro-alpha-muricholic acid (TαMCA), tauro-beta-muricholic acid (TβMCA), tauro-omega-muricholic acid (TωMCA), tauro-hyocholic acid (THCA), tauro-deoxycholic acid (TDCA), tauro-ursodeoxycholic acid (TUDCA), tauro-chenodeoxycholic acid (TCDCA), taurolithocholic acid (TLCA), glyco-cholic acid (GCA), glyco-hyocholic acid (GHCA), glyco-deoxycholic acid (GDCA), glyco-hyodeoxycholic acid (GHDCA), cholic acid (CA), alpha-muricholic acid (αMCA), beta-muricholic acid (βMCA), omega-muricholic acid (ωMCA), hyocholic acid (HCA, also known as γ-muricholic acid), deoxycholic acid (DCA), chenodeoxycholic acid (CDCA), ursodeoxycholic acid (UDCA), hyodeoxycholic acid (HDCA), murocholic acid (murideoxycholic acid, MDCA), lithocholic acid (LCA), glycolithocholic acid (GLCA), glycourosodeoxycholic acid (GUDCA), and glycochenodeoxycholic acid (GCDCA). It was not possible to resolve UDCA and HDCA; so the sum was reported.

Integrated areas of extracted ion chromatograms were obtained using QuanLynx (Waters, Milford, MA, USA) and a mass extraction window of 10 mDa. Final corrections accounting for drift in instrumental sensitivity were performed in Microsoft Excel.

Elution Gradient

Samples were eluted using the following gradient of water with 0.1% formic acid (“A”) and balance of acetonitrile with 0.1% formic acid:

1. 0 min, 0.55 mL/min at 68% A
2. 2 min, 0.55 mL/min at 60% A, 10 curve
3. 5 min, 0.55 mL/min at 40% A, 5 curve
4. 6 min, 1.1 mL/min at 0% A, 10 curve
5. 6.2 min, 1.2 mL/min at 0% A, 6 curve
6. 6.5 min, 1.47 mL/min at 0% A, 6 curve
7. 8.9 min, 1.5 mL/min at 0% A, 6 curve
8. 9.0 min, 0.9 mL/min at 68% A, 6 curve
9. 10 min, 0.55 mL/min at 68% A, 6 curve

Bile acid data processing:

Bile acid data analysis was performed using the tools described in "Sequencing data processing."

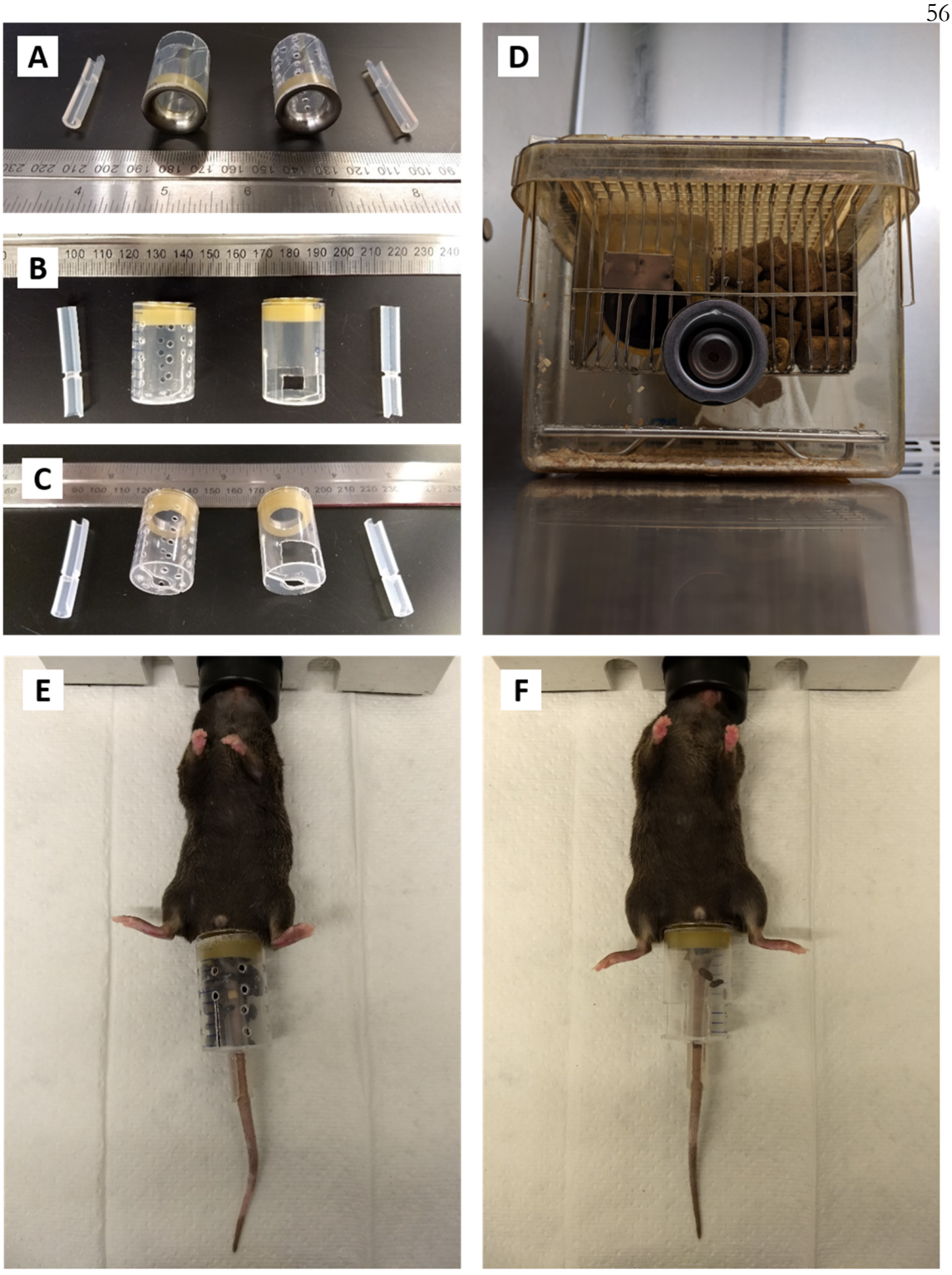

Fig. 2S1. Tail cup design and experimental setup for preventing coprophagy. (A, B, C) Functional (TC-F, left) and mock (TC-M, right) tail cups as viewed from different perspectives. (D) The standard cages with wire mesh floors used in this study (WF). (E, F)

Ventral view of the functional (TC-F; left) and mock (TC-M, right) tail cups 24 hours after emptying (TC-F) or mock emptying (TC-M).

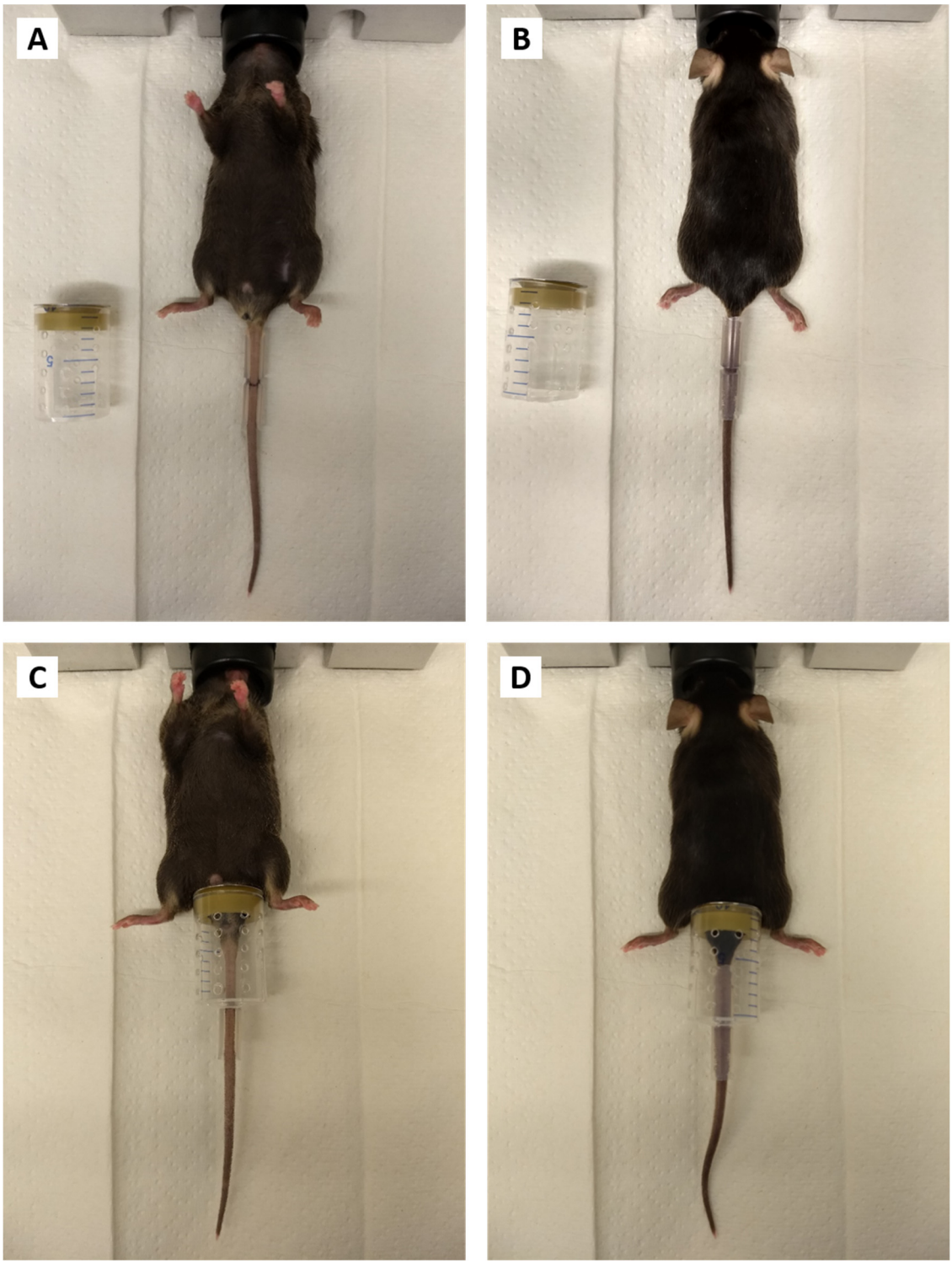

Fig. 2S2. Mounting of functional tail cups onto mice. (A, B) Ventral and dorsal view of the tail sleeve mounted at the tail base. **(C, D)** Ventral and dorsal view of the functional tail cup installed and locked in place using the tail sleeve.

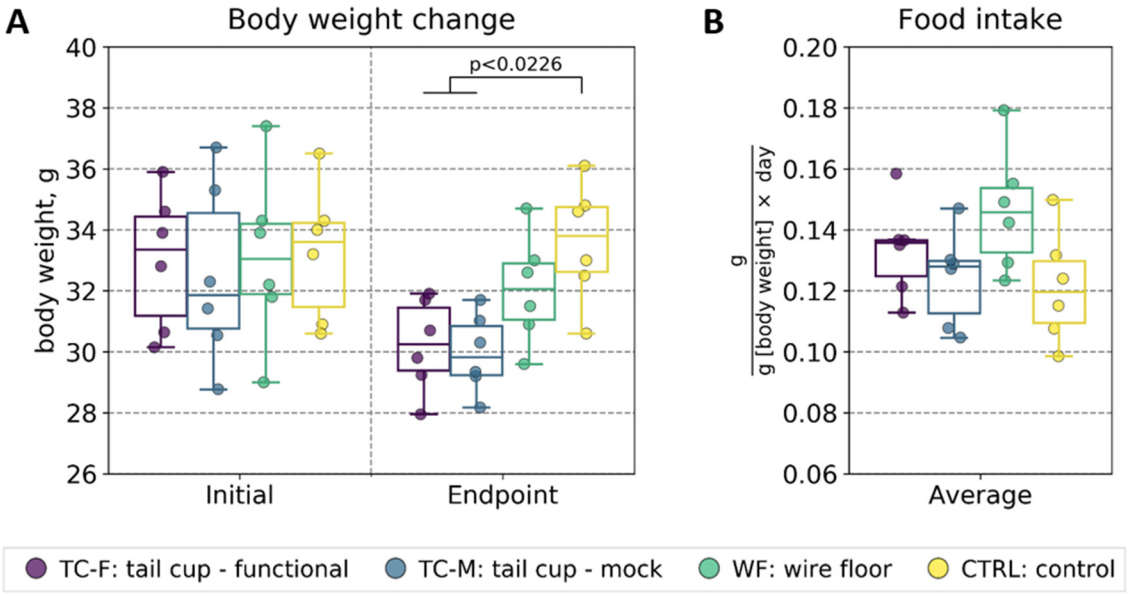

Fig. 2S3. Body weight changes across all groups of mice in relation to food intake over the course of the study. (A) Body weights of each individual animal at the beginning and at the endpoint of the study. **(B)** Normalized food intake per gram of body weight per day measured over the entire duration of the study. Multiple comparisons of the normally-distributed homoscedastic data were performed using one-way ANOVA; pairwise comparisons were performed using the Student's *t*-test with FDR correction. N = 6 mice per group.

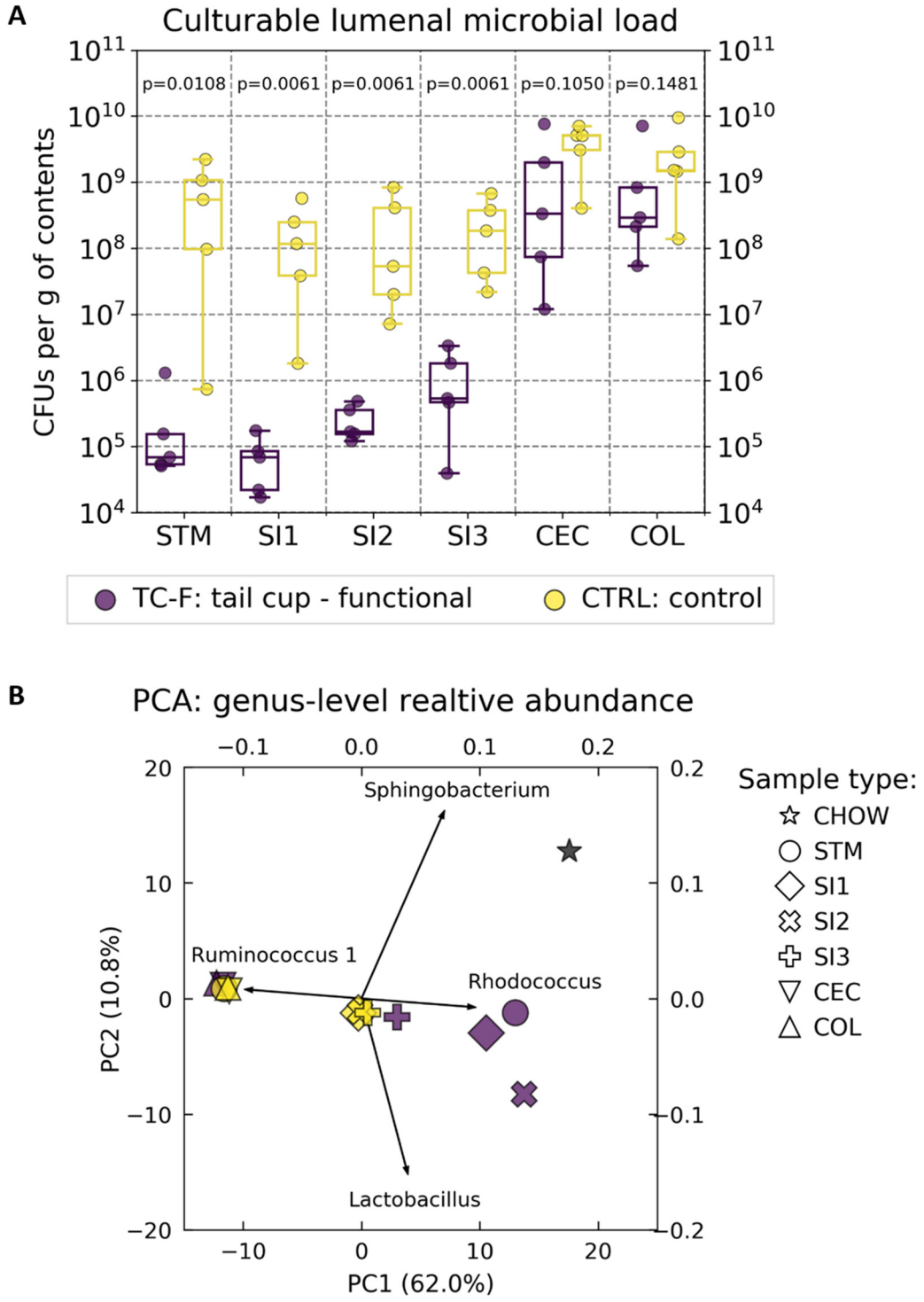

Fig. 2S4. Quantification of the culturable microbial load and microbiota profile along the entire GIT of mice fitted with functional tail cups (TC-F) and control mice (CTRL).

(A) Culturable microbial loads in contents along the gastrointestinal tract were evaluated using the most probable number (MPN) assay performed in anaerobic BHI-S broth (N = 5 mice per group, *P*-values were calculated using the Wilcoxon–Mann–Whitney test). **(B)** PCA analysis of the CLR-transformed relative microbial abundance profiles (16S rRNA gene amplicon sequencing) along the entire GIT in TC and CT mice (N = 1 mouse from each group).

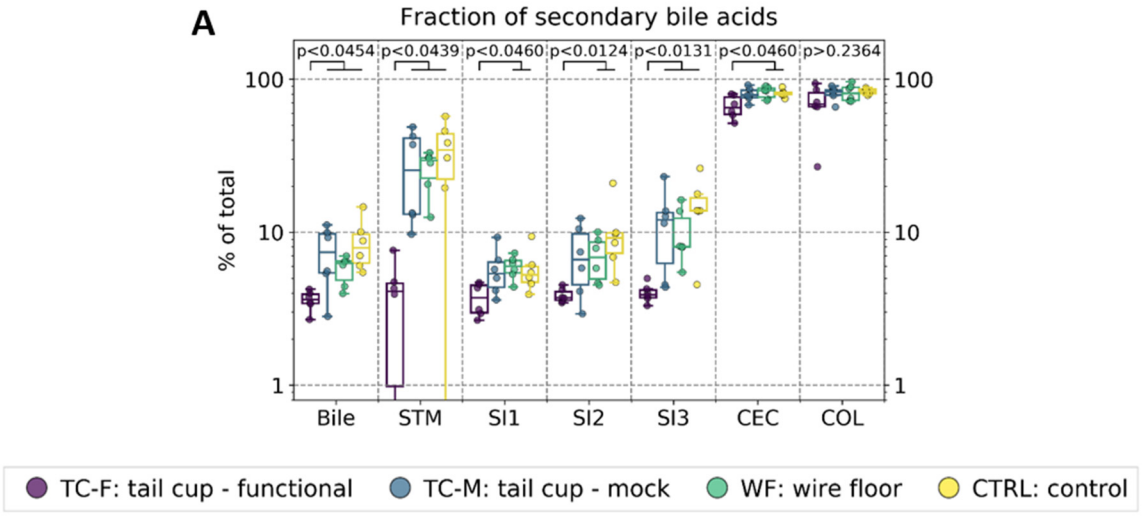

Fig. 2S5. Bile acid profiles in gallbladder bile and in lumenal contents along the entire GIT. (A) The fraction of secondary bile acids (conjugated + unconjugated) in gallbladder bile and throughout the GIT (STM = stomach; SI1 = upper third of the small intestine (SI), SI2 = middle third or the SI, SI3 = lower third of the SI roughly corresponding to the duodenum, jejunum, and ileum respectively; CEC = cecum; COL = colon). In all plots, individual data points are overlaid onto box-and-whisker plots; whiskers extend from the quartiles (Q2 and Q3) to the last data point within 1.5 × interquartile range (IQR). Multiple comparisons were performed using the Kruskal–Wallis test; pairwise comparisons were performed using the Wilcoxon–Mann–Whitney test with FDR correction. N = 6 mice per group.

Table 2S1. Primer oligonucleotide sequences used in the study. [NNNNNNNNNNNN] – 12-base barcode sequences "806rcbc" according to [3].

Primer	Oligonucleotide sequence	Assay	Reference
UN00F2	CAGCMGCCGCGGTAA	16S rRNA gene DNA qPCR	[49]
UN00R0	GGACTACHVGGGTWTCTAAT	16S rRNA gene DNA ddPCR	[1], [3]
UN00F2_BC	AATGATACGGCGACCACCGA GATCTACACTATGGTAATTGT CAGCMGCCGCGGTAA	16S rRNA gene DNA amplicon barcoding	[49]
UN00R0_BC	CAAGCAGAAGACGGCATACGAGAT [NNNNNNNNNNNN] AGTCAGTCAGCC GGACTACHVGGGTWTCTAAT		[1], [3]
ILM00F(P5)	AATGATACGGCGACCACCGA	Barcoded amplicon and NGS library quantification ddPCR	[1], [3], [147]–[149]
ILM00R(P7)	CAAGCAGAAGACGGCATACGA		
Seq_UN00F2_Read_1	TATGGTAATTGTCAGCMGCCGCGGTAA	MiSeq read 1	[49]
Seq_UN00R0_Read_2	AGTCAGTCAGCCGGACTACHVGGGTWTCTAAT	MiSeq read 2	[1], [3]
Seq_UN00R0_RC_Index	ATTAGAWACCCBDGTAGTCCGGCTGACTGACT	MiSeq index read	[1], [3]

Table 2S2. Thermocycling parameters for the quantitative PCR (qPCR) assay for 16S rRNA gene DNA copy quantification.

Step	Repeats	Temperature, °C	Time, sec
Initial denaturation	× 1	95	120
Cycle	× 40	95	15
		53	10
		68	45

Table 2S3. Thermocycling parameters for the digital PCR (dPCR) assay for absolute 16S rRNA gene DNA copy quantification.

Step	Repeats	Temperature, °C	Time, sec	Ramp, °C/sec
Initial denaturation	× 1	95	300	2.0
Cycle	× 40	95	30	2.0
		52	30	2.0
		68	60	2.0
Dye stabilization	× 1	4	300	2.0
		90	300	2.0
		12	∞	2.0

Table 2S4. Thermocycling parameters for the 16S rRNA gene DNA amplicon barcoding PCR reaction for next generation sequencing (NGS).

Step	Repeats	Temperature, °C	Time, sec
Initial denaturation	× 1	94	180
Cycle	× var.	94	45
		54	60
		72	105
Final extension	× 1	72	600

Table 2S5. Thermocycling parameters for the digital PCR (dPCR) assay for barcoded amplicon and Illumina NGS library quantification.

Step	Repeats	Temperature, °C	Time, sec	Ramp, °C/sec
Initial denaturation	× 1	95	300	2.0
Cycle	× 40	95	30	2.0
		60	90	2.0
Dye stabilization	× 1	4	300	2.0
		90	300	2.0
		12	∞	2.0

Table 2S6. Reagents and chemical standards used in the bile acid metabolomics assay.

Bile acid	Reference #	Vendor	LOT
TαMCA	C1893-000	Steraloids	B1439
TβMCA	C1899-000	Steraloids	B1594
TωMCA	C1889-000	Steraloids	B1731
THCA	C1887-000	Steraloids	B1621
αMCA	C1890-000	Steraloids	B1529
βMCA	C1895-000	Steraloids	B1725
ωMCA	C1888-000	Steraloids	B1710
HCA (gMCA)	C1850-000	Steraloids	B0696
HDCA	C0860-000	Steraloids	B0684
MCA	C0910-000	Steraloids	B1711
GDCA	C1087-000	Steraloids	B2122
GCA	C1927-000	Steraloids	
GHDCA	C0865-000	Steraloids	B1667
GHCA	C1860-000	Steraloids	L1105
TCA	13232UNL	Isosciences	EH1-2015-111A1
CA	13098UNL	Isosciences	EH1-2014-075A1
DCA	13100UNL	Isosciences	EH1-2014-076A1
TCDCA	13105UNL	Isosciences	EH1-2015-110A1
TDCA	13225UNL	Isosciences	EH1-2015-112A1
TUDCA	13106UNL	Isosciences	EH1-2014-027A1
TLCA	13230UNL	Isosciences	EH1-2014-077A1
CDCA	13101UNL	Isosciences	PG1-2014-149A1
UDCA	13102UNL	Isosciences	EH1-2015-113A1
LCA	13099UNL	Isosciences	EH1-2014-030A1
D4-TCA	13232	Isosciences	SJ5-2015-035A1
D4-DCA	13100	Isosciences	RS6-2014-168A1
D4-CA	13098	Isosciences	SJ5-2015-100A1

D4-TDCA	13225	Isosciences	SJ5-2015-034A1
D4-GLCA	13231	Isosciences	SR3-2015-203A1
D4-GUDCA	13224	Isosciences	SJ5-2017-206A1
D4-GCDCA	13104	Isosciences	SJ4-2012-070A1
D4-GCA	13443	Isosciences	SJ5-2015-118A1
D4-GDCA	13226	Isosciences	SJ5-2015-033A1

AVAILABILITY OF DATA AND MATERIAL

Sequencing data (paired end reads in FASTQ) and a manifest file for analysis in Qiime2 are available under a CC-BY license via CaltechDATA: http://dx.doi.org/10.22002/D1.1295.

Supplementary Information includes a zip file containing all sequencing sample metadata, numerical microbial quantification data (16S copies from the main study + MPN from the pilot study), Qiime2 sequencing output data, PICRUSt2 output data, numerical bile acid analysis data, numerical body weight data, numerical food intake data, and analytical scripts (iPython Notebooks) for all figures and statistical analyses in the manuscript.

ACKNOWLEDGEMENTS

This work was supported in part by the Kenneth Rainin Foundation Innovator Award, Army Research Office (ARO) Multidisciplinary University Research Initiative (MURI) contract #W911NF-17-1-0402, and the Jacobs Institute for Molecular Engineering for Medicine. We thank Karen Lencioni, Janet Baer, the Caltech Office of Laboratory Animal Resources, veterinary technicians at the Church Animal Facility for experimental resources. We thank Liang Ma for the introduction into 16S rRNA gene amplicon sequencing, Heidi Klumpe for the assistance with the preliminary MPN experiments, and Justin Bois for the introduction to data analysis in Python. S.R.B. would like to thank Kimberly Zhou for the personal feedback on the project and inspiration. This project benefited from the use of instrumentation made available by the Caltech Environmental Analysis Center and technical support from Nathan Dalleska. We thank Natasha Shelby for contributions to writing and editing this manuscript.

BIBLIOGRAPHY

[1] J. G. Caporaso *et al.*, “Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 108 Suppl, pp. 4516–22, Mar. 2011.

[2] M. F. Polz and C. M. Cavanaugh, “Bias in template-to-product ratios in multitemplate PCR.,” *Appl. Environ. Microbiol.*, vol. 64, no. 10, pp. 3724–30, Oct. 1998.

[3] J. G. Caporaso *et al.*, “Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms.,” *ISME J.*, vol. 6, no. 8, pp. 1621–4, Aug. 2012.

[4] E. Pruesse *et al.*, “SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB.,” *Nucleic Acids Res.*, vol. 35, no. 21, pp. 7188–96, 2007.

[5] A. Klindworth *et al.*, “Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies.,” *Nucleic Acids Res.*, vol. 41, no. 1, p. e1, Jan. 2013.

[6] S. R. Bogatyrev, J. C. Rolando, and R. F. Ismagilov, “Self-reinoculation with fecal flora changes microbiota density and composition leading to an altered bile-acid profile in the mouse small intestine,” 2019.

[7] C. Jian, P. Luukkonen, H. Yki-Jarvinen, A. Salonen, and K. Korpela, “Quantitative PCR provides a simple and accessible method for quantitative microbiome profiling,” *bioRxiv*, p. 478685, Jan. 2018.

[8] D. Vandeputte *et al.*, “Quantitative microbiome profiling links gut community variation to microbial load.,” *Nature*, vol. 551, no. 7681, pp. 507–511, 2017.

[9] R. Props *et al.*, “Absolute quantification of microbial taxon abundances.,” *ISME J.*, vol. 11, no. 2, pp. 584–587, Feb. 2017.

[10] A. Tkacz, M. Hortala, and P. S. Poole, “Absolute quantitation of microbiota abundance in environmental samples.,” *Microbiome*, vol. 6, no. 1, p. 110, 2018.

[11] D. M. O'Sullivan *et al.*, "Assessing the accuracy of quantitative molecular microbial profiling.," *Int. J. Mol. Sci.*, vol. 15, no. 11, pp. 21476–91, Nov. 2014.

[12] H. F. Helander and L. Fändriks, "Surface area of the digestive tract - revisited.," *Scand. J. Gastroenterol.*, vol. 49, no. 6, pp. 681–9, Jun. 2014.

[13] S. El Aidy, B. van den Bogert, and M. Kleerebezem, "The small intestine microbiota, nutritional modulation and relevance for health.," *Curr. Opin. Biotechnol.*, vol. 32, pp. 14–20, Apr. 2015.

[14] G. B. Saffouri *et al.*, "Small intestinal microbial dysbiosis underlies symptoms associated with functional gastrointestinal disorders.," *Nat. Commun.*, vol. 10, no. 1, p. 2012, 2019.

[15] E. M. Brown *et al.*, "Diet and specific microbial exposure trigger features of environmental enteropathy in a novel murine model.," *Nat. Commun.*, vol. 6, p. 7806, 2015.

[16] J. R. Donowitz and W. A. Petri, "Pediatric small intestine bacterial overgrowth in low-income countries.," *Trends Mol. Med.*, vol. 21, no. 1, pp. 6–15, Jan. 2015.

[17] E. Ierardi *et al.*, "Macronutrient intakes in obese subjects with or without small intestinal bacterial overgrowth: an alimentary survey.," *Scand. J. Gastroenterol.*, vol. 51, no. 3, pp. 277–80, Mar. 2016.

[18] E. J. Giamarellos-Bourboulis, E. Pyleris, C. Barbatzas, A. Pistiki, and M. Pimentel, "Small intestinal bacterial overgrowth is associated with irritable bowel syndrome and is independent of proton pump inhibitor usage.," *BMC Gastroenterol.*, vol. 16, no. 1, p. 67, Jul. 2016.

[19] A. Shah, M. Morrison, and G. J. Holtmann, "Gastroduodenal 'Dysbiosis': a New Clinical Entity.," *Curr. Treat. Options Gastroenterol.*, vol. 16, no. 4, pp. 591–604, Dec. 2018.

[20] A. Shah *et al.*, "Systematic review with meta-analysis: the prevalence of small intestinal bacterial overgrowth in inflammatory bowel disease.," *Aliment. Pharmacol. Ther.*, vol. 49, no. 6, pp. 624–635, Mar. 2019.

[21] M. Muraki *et al.*, “Role of small intestinal bacterial overgrowth in severe small intestinal damage in chronic non-steroidal anti-inflammatory drug users.,” *Scand. J. Gastroenterol.*, vol. 49, no. 3, pp. 267–73, Mar. 2014.

[22] H. Hayashi, R. Takahashi, T. Nishi, M. Sakamoto, and Y. Benno, “Molecular analysis of jejunal, ileal, caecal and recto-sigmoidal human colonic microbiota using 16S rRNA gene libraries and terminal restriction fragment length polymorphism.,” *J. Med. Microbiol.*, vol. 54, no. Pt 11, pp. 1093–101, Nov. 2005.

[23] K. Yasuda *et al.*, “Biogeography of the intestinal mucosal and lumenal microbiome in the rhesus macaque.,” *Cell Host Microbe*, vol. 17, no. 3, pp. 385–391, Mar. 2015.

[24] T. L. A. Nguyen, S. Vieira-Silva, A. Liston, and J. Raes, “How informative is the mouse for human gut microbiota research?,” *Dis. Model. Mech.*, vol. 8, no. 1, pp. 1–16, Jan. 2015.

[25] F. Hugenholtz and W. M. de Vos, “Mouse models for human intestinal microbiota research: a critical evaluation.,” *Cell. Mol. Life Sci.*, vol. 75, no. 1, pp. 149–160, 2018.

[26] G. J. Kenagy and D. F. Hoyt, “Reingestion of feces in rodents and its daily rhythmicity.,” *Oecologia*, vol. 44, no. 3, pp. 403–409, Jan. 1979.

[27] K. Y. Ebino, “Studies on coprophagy in experimental animals.,” *Jikken Dobutsu.*, vol. 42, no. 1, pp. 1–9, Jan. 1993.

[28] O. Soave and C. D. Brand, “Coprophagy in animals: a review.,” *Cornell Vet.*, vol. 81, pp. 357–364, 1991.

[29] K. Y. Ebino, H. Amao, T. Suwa, Y. Kuwabara, T. R. Saito, and K. W. Takahashi, “[Coprophagy in the germfree mouse].,” *Jikken Dobutsu.*, vol. 36, no. 1, pp. 33–7, Jan. 1987.

[30] K. Y. Ebino, T. Suwa, Y. Kuwabara, T. R. Saito, and K. W. Takahashi, “[Analyses of constituents of feces and the effect of a vitamin B12 fortified diet on coprophagy in the mouse].,” *Jikken Dobutsu.*, vol. 35, no. 4, pp. 381–6, Oct. 1986.

[31] H. W. Smith, “Observations on the flora of the alimentary tract of animals and factors affecting its composition,” *J. Pathol. Bacteriol.*, vol. 89, no. 1, pp. 95–122, Jan. 1965.

[32] K. W. Takahashi, K. Y. Ebino, T. R. Saito, and T. Imamichi, “Strain Difference in Coprophagous Behavior in Laboratory Mice (Mus musculus),” *Zoolog. Sci.*, vol. 2, no. 2, pp. p249-255, 1985.

[33] H. L. B. M. Klaasen, J. P. Koopman, P. M. Scholten, M. E. Van Den Brink, and A. G. M. Theeuwes, “Effect of Preventing Coprophagy on Colonisation by Segmented Filamentous Bacteria in the Small Bowel of Mice,” *Microb. Ecol. Health Dis.*, vol. 3, no. 2, pp. 99–103, Jan. 1990.

[34] M. Lev, R. H. Alexander, and S. M. Levenson, “Stability of the Lactobacillus population in feces and stomach contents of rats prevented from coprophagy.,” *J. Bacteriol.*, vol. 92, no. 1, pp. 13–6, Jul. 1966.

[35] S. A. Syed, G. D. Abrams, and R. Freter, “Efficiency of various intestinal bacteria in assuming normal functions of enteric flora after association with germ-free mice.,” *Infect. Immun.*, vol. 2, no. 4, pp. 376–86, Oct. 1970.

[36] R. J. Fitzgerald, B. E. Gustafsson, and E. G. McDaniel, “Effects of Coprophagy Prevention on Intestinal Microflora in Rats,” *J. Nutr.*, vol. 84, no. 2, pp. 155–160, Oct. 1964.

[37] B. E. Gustafsson and R. J. Fitzgerald, “Alteration in Intestinal Microbial Flora of Rats with Tail Cups to Prevent Coprophagy,” *Exp. Biol. Med.*, vol. 104, no. 2, pp. 319–322, Jun. 1960.

[38] S. E. George, M. J. Kohan, L. D. Claxton, and D. B. Walsh, “Acute colonization study of polychlorinated biphenyl-degrading pseudomonads in the mouse intestinal tract: Comparison of single and multiple exposures,” *Environ. Toxicol. Chem.*, vol. 8, no. 2, pp. 123–131, Feb. 1989.

[39] M. H. Roscoe, “Spontaneous Cures in Rats reared upon a Diet devoid of Vitamin B and Antineuritic Vitamin.,” *J. Hyg. (Lond).*, vol. 27, no. 1, pp. 103–107, 1927.

[40] R. H. Barnes and G. Fiala, “Effects of the prevention of coprophagy in the rat. I. Growth studies.,” *J. Nutr.*, vol. 64, no. 4, pp. 533–40, Apr. 1958.

[41] R. H. Barnes, G. Fiala, B. McGehee, and A. Brown, “Prevention of Coprophagy in

the Rat," *J. Nutr.*, vol. 63, no. 4, pp. 489–498, Dec. 1957.

[42] R. H. Barnes, "Nutritional implications of coprophagy.," *Nutr. Rev.*, vol. 20, pp. 289–91, Oct. 1962.

[43] H. Hörnicke and G. Björnhag, "Coprophagy and related strategies for digesta utilization," in *Digestive Physiology and Metabolism in Ruminants: Proceedings of the 5th International Symposium on Ruminant Physiology, held at Clermont --- Ferrand, on 3rd--7th September, 1979*, Y. Ruckebusch and P. Thivend, Eds. Dordrecht: Springer Netherlands, 1980, pp. 707–730.

[44] K. Y. Ebino, K. Yoshinaga, T. R. Saito, and K. W. Takahashi, "A simple method for prevention of coprophagy in the mouse.," *Lab. Anim.*, vol. 22, no. 1, pp. 1–4, Jan. 1988.

[45] N. Zmora *et al.*, "Personalized Gut Mucosal Colonization Resistance to Empiric Probiotics Is Associated with Unique Host and Microbiome Features.," *Cell*, vol. 174, no. 6, pp. 1388-1405.e21, 2018.

[46] S. Gu *et al.*, "Bacterial community mapping of the mouse gastrointestinal tract.," *PLoS One*, vol. 8, no. 10, p. e74957, 2013.

[47] T. A. Suzuki and M. W. Nachman, "Spatial Heterogeneity of Gut Microbial Composition along the Gastrointestinal Tract in Natural Populations of House Mice.," *PLoS One*, vol. 11, no. 9, p. e0163720, Sep. 2016.

[48] A. C. Ericsson, J. Gagliardi, D. Bouhan, W. G. Spollen, S. A. Givan, and C. L. Franklin, "The influence of caging, bedding, and diet on the composition of the microbiota in different regions of the mouse gut.," *Sci. Rep.*, vol. 8, no. 1, p. 4065, 2018.

[49] S. R. Bogatyrev and R. F. Ismagilov, "Quantitative profiling of lumenal and mucosal microbiota via a single-step 16S rRNA gene copy quantification and amplicon barcoding," 2019.

[50] J. T. Barlow, S. R. Bogatyrev, and R. F. Ismagilov, "A Quantitative sequencing framework for absolute abundance measurements of mucosal and lumenal microbial

communities," 2019.

[51] A. M. Seekatz *et al.*, "Spatial and Temporal Analysis of the Stomach and Small-Intestinal Microbiota in Fasted Healthy Humans.," *mSphere*, vol. 4, no. 2, 2019.

[52] F. Vuik *et al.*, "Composition of the mucosa-associated microbiota along the entire gastrointestinal tract of human individuals.," *United Eur. Gastroenterol. J.*, vol. 7, no. 7, pp. 897–907, Aug. 2019.

[53] Z. Xu, D. Malmer, M. G. I. Langille, S. F. Way, and R. Knight, "Which is more important for classifying microbial communities: who's there or what they can do?," *ISME J.*, vol. 8, no. 12, pp. 2357–9, Dec. 2014.

[54] E. J. Contijoch *et al.*, "Gut microbiota density influences host physiology and is shaped by host and microbial factors.," *Elife*, vol. 8, Jan. 2019.

[55] M. G. I. Langille *et al.*, "Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences.," *Nat. Biotechnol.*, vol. 31, no. 9, pp. 814–21, Sep. 2013.

[56] G. M. Douglas *et al.*, "PICRUSt2: An improved and extensible approach for metagenome inference," *bioRxiv*, p. 672295, Jan. 2019.

[57] M. H. Foley, S. O'Flaherty, R. Barrangou, and C. M. Theriot, "Bile salt hydrolases: Gatekeepers of bile acid metabolism and host-microbiome crosstalk in the gastrointestinal tract.," *PLoS Pathog.*, vol. 15, no. 3, p. e1007581, 2019.

[58] Z. Dong and B. H. Lee, "Bile salt hydrolases: Structure and function, substrate preference, and inhibitor development.," *Protein Sci.*, vol. 27, no. 10, pp. 1742–1754, 2018.

[59] S. A. Joyce, F. Shanahan, C. Hill, and C. G. M. Gahan, "Bacterial bile salt hydrolase in host metabolism: Potential for influencing gastrointestinal microbe-host crosstalk.," *Gut Microbes*, vol. 5, no. 5, pp. 669–74, 2014.

[60] *Role of the Gut Flora in Toxicity and Cancer*. Elsevier, 1988.

[61] C. D. Klaassen and J. Y. Cui, "Review: Mechanisms of How the Intestinal Microbiota Alters the Effects of Drugs and Bile Acids.," *Drug Metab. Dispos.*, vol. 43, no. 10,

pp. 1505–21, Oct. 2015.

[62] K. M. Schneider, S. Albers, and C. Trautwein, “Role of bile acids in the gut-liver axis.,” *J. Hepatol.*, vol. 68, no. 5, pp. 1083–1085, 2018.

[63] A. Wahlström, S. I. Sayin, H.-U. Marschall, and F. Bäckhed, “Intestinal Crosstalk between Bile Acids and Microbiota and Its Impact on Host Metabolism.,” *Cell Metab.*, vol. 24, no. 1, pp. 41–50, 2016.

[64] J. M. Ridlon and J. S. Bajaj, “The human gut sterolbiome: bile acid-microbiome endocrine aspects and therapeutics.,” *Acta Pharm. Sin. B*, vol. 5, no. 2, pp. 99–105, Mar. 2015.

[65] T. C. Northfield and I. McColl, “Postprandial concentrations of free and conjugated bile acids down the length of the normal human small intestine.,” *Gut*, vol. 14, no. 7, pp. 513–8, Jul. 1973.

[66] P. Tso, J. A. Balint, M. B. Bishop, and J. B. Rodgers, “Acute inhibition of intestinal lipid transport by Pluronic L-81 in the rat.,” *Am. J. Physiol.*, vol. 241, no. 6, pp. G487-97, Dec. 1981.

[67] C. T. Phan and P. Tso, “Intestinal lipid absorption and transport.,” *Front. Biosci.*, vol. 6, pp. D299-319, Mar. 2001.

[68] V. C. Metta, L. Nash, and B. C. Johnson, “A Tubular Coprophagy-Preventing Cage for the Rat,” *J. Nutr.*, vol. 74, no. 4, pp. 473–476, 1961.

[69] C. L. Franklin and A. C. Ericsson, “Microbiota and reproducibility of rodent models.,” *Lab Anim. (NY).*, vol. 46, no. 4, pp. 114–122, Mar. 2017.

[70] G. W. Tannock, “The Lactic Microflora of Pigs, Mice and Rats,” in *The Lactic Acid Bacteria Volume 1*, Boston, MA: Springer US, 1992, pp. 21–48.

[71] Z.-K. Wang and Y.-S. Yang, “Upper gastrointestinal microbiota and digestive diseases.,” *World J. Gastroenterol.*, vol. 19, no. 10, pp. 1541–50, Mar. 2013.

[72] V. Liévin-Le Moal and A. L. Servin, “Anti-infective activities of lactobacillus strains in the human intestinal microbiota: from probiotics to gastrointestinal anti-infectious biotherapeutic agents.,” *Clin. Microbiol. Rev.*, vol. 27, no. 2, pp. 167–99, Apr. 2014.

[73] C. A. Lozupone, J. I. Stombaugh, J. I. Gordon, J. K. Jansson, and R. Knight, "Diversity, stability and resilience of the human gut microbiota.," *Nature*, vol. 489, no. 7415, pp. 220–30, Sep. 2012.

[74] K. Z. Coyte, J. Schluter, and K. R. Foster, "The ecology of the microbiome: Networks, competition, and stability.," *Science*, vol. 350, no. 6261, pp. 663–6, Nov. 2015.

[75] K. M. Ng *et al.*, "Recovery of the gut microbiota after antibiotics depends on host diet and environmental reservoirs," *bioRxiv*, 2019.

[76] J. M. Ridlon, D.-J. Kang, and P. B. Hylemon, "Bile salt biotransformations by human intestinal bacteria.," *J. Lipid Res.*, vol. 47, no. 2, pp. 241–59, 2006.

[77] E. F. Enright, B. T. Griffin, C. G. M. Gahan, and S. A. Joyce, "Microbiome-mediated bile acid modification: Role in intestinal drug absorption and metabolism.," *Pharmacol. Res.*, vol. 133, pp. 170–186, 2018.

[78] Y. S. Kim, N. Spritz, M. Blum, J. Terz, and P. Sherlock, "The role of altered bile acid metabolism in the steatorrhea of experimental blind loop.," *J. Clin. Invest.*, vol. 45, no. 6, pp. 956–62, Jun. 1966.

[79] S. Tabaqchali, J. Hatzioannou, and C. C. Booth, "Bile-salt deconjugation and steatorrhoea in patients with the stagnant-loop syndrome.," *Lancet (London, England)*, vol. 2, no. 7558, pp. 12–6, Jul. 1968.

[80] T. C. Northfield, "Intraluminal precipitation of bile acids in stagnant loop syndrome.," *Br. Med. J.*, vol. 2, no. 5869, pp. 743–5, Jun. 1973.

[81] R. M. Donaldson, "Studies on the pathogenesis of steatorrhea in the blind loop syndrome.," *J. Clin. Invest.*, vol. 44, no. 11, pp. 1815–25, Nov. 1965.

[82] K. Shindo, M. Machida, M. Fukumura, K. Koide, and R. Yamazaki, "Omeprazole induces altered bile acid metabolism.," *Gut*, vol. 42, no. 2, pp. 266–71, Feb. 1998.

[83] S. I. Sayin *et al.*, "Gut microbiota regulates bile acid metabolism by reducing the levels of tauro-beta-muricholic acid, a naturally occurring FXR antagonist.," *Cell Metab.*, vol. 17, no. 2, pp. 225–35, Feb. 2013.

[84] D. Madsen, M. Beaver, L. Chang, E. Bruckner-Kardoss, and B. Wostmann, "Analysis

of bile acids in conventional and germfree rats.," *J. Lipid Res.*, vol. 17, no. 2, pp. 107–11, 1976.

[85] S. P. Claus *et al.*, "Systemic multicompartmental effects of the gut microbiome on mouse metabolic phenotypes.," *Mol. Syst. Biol.*, vol. 4, p. 219, 2008.

[86] L. Yao *et al.*, "A selective gut bacterial bile salt hydrolase alters host metabolism.," *Elife*, vol. 7, 2018.

[87] S. Narushima *et al.*, "Deoxycholic acid formation in gnotobiotic mice associated with human intestinal bacteria.," *Lipids*, vol. 41, no. 9, pp. 835–43, Sep. 2006.

[88] S. Narushima, K. Itoh, F. Takamine, and K. Uchida, "Absence of cecal secondary bile acids in gnotobiotic mice associated with two human intestinal bacteria with the ability to dehydroxylate bile acids in vitro.," *Microbiol. Immunol.*, vol. 43, no. 9, pp. 893–7, 1999.

[89] T. Chikai, H. Nakao, and K. Uchida, "Deconjugation of bile acids by human intestinal bacteria implanted in germ-free rats.," *Lipids*, vol. 22, no. 9, pp. 669–71, Sep. 1987.

[90] B. E. Gustafsson, J. Gustafsson, and B. Carlstedt-Duke, "Prolonged induction of germfree bile acid pattern in conventional rats by antibiotics.," *Acta Med. Scand.*, vol. 201, no. 3, pp. 155–60, 1977.

[91] S. D. Feighner and M. P. Dashkevicz, "Subtherapeutic levels of antibiotics in poultry feeds and their effects on weight gain, feed efficiency, and bacterial cholyltaurine hydrolase activity.," *Appl. Environ. Microbiol.*, vol. 53, no. 2, pp. 331–6, Feb. 1987.

[92] H. Kuribayashi, M. Miyata, H. Yamakawa, K. Yoshinari, and Y. Yamazoe, "Enterobacteria-mediated deconjugation of taurocholic acid enhances ileal farnesoid X receptor signaling.," *Eur. J. Pharmacol.*, vol. 697, no. 1–3, pp. 132–8, Dec. 2012.

[93] S. Tabaqchali, O. A. Okubadejo, G. Neale, and C. C. Booth, "Influence of abnormal bacterial flora on small intestinal function.," *Proc. R. Soc. Med.*, vol. 59, no. 12, pp. 1244–6, Dec. 1966.

[94] M. Begley, C. Hill, and C. G. M. Gahan, "Bile salt hydrolase activity in probiotics.," *Appl. Environ. Microbiol.*, vol. 72, no. 3, pp. 1729–38, 2006.

[95] M. L. Jones, C. Tomaro-Duchesneau, C. J. Martoni, and S. Prakash, “Cholesterol lowering with bile salt hydrolase-active probiotic bacteria, mechanism of action, clinical evidence, and future direction for heart health applications.,” *Expert Opin. Biol. Ther.*, vol. 13, no. 5, pp. 631–42, May 2013.

[96] W. Geng and J. Lin, “Bacterial bile salt hydrolase: an intestinal microbiome target for enhanced animal health.,” *Anim. Heal. Res. Rev.*, vol. 17, no. 2, pp. 148–158, 2016.

[97] G. W. Tannock, M. P. Dashkevicz, and S. D. Feighner, “Lactobacilli and bile salt hydrolase in the murine intestinal tract.,” *Appl. Environ. Microbiol.*, vol. 55, no. 7, pp. 1848–51, 1989.

[98] S. A. Joyce *et al.*, “Regulation of host weight gain and lipid metabolism by bacterial bile acid modification in the gut.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 20, pp. 7421–6, May 2014.

[99] J. Jeun *et al.*, “Hypocholesterolemic effects of Lactobacillus plantarum KCTC3928 by increased bile acid excretion in C57BL/6 mice.,” *Nutrition*, vol. 26, no. 3, pp. 321–30, Mar. 2010.

[100] C. Li *et al.*, “Cholesterol-lowering effect of Lactobacillus plantarum NCU116 in a hyperlipidaemic rat model,” *J. Funct. Foods*, vol. 8, pp. 340–347, May 2014.

[101] X. Guan, Q. Xu, Y. Zheng, L. Qian, and B. Lin, “Screening and characterization of lactic acid bacterial strains that produce fermented milk and reduce cholesterol levels.,” *Braz. J. Microbiol.*, vol. 48, no. 4, pp. 730–739, 2017.

[102] G. Wang, W. Huang, Y. Xia, Z. Xiong, and L. Ai, “Cholesterol-lowering potentials of Lactobacillus strain overexpression of bile salt hydrolase on high cholesterol diet-induced hypercholesterolemic mice.,” *Food Funct.*, vol. 10, no. 3, pp. 1684–1695, Mar. 2019.

[103] A. F. Hofmann, “Clinical implications of physicochemical studies on bile salts.,” *Gastroenterology*, vol. 48, no. 4, pp. 484–94, Apr. 1965.

[104] A. M. Dawson, “Bile salts and fat absorption.,” *Gut*, vol. 8, no. 1, pp. 1–3, Feb. 1967.

[105] K. Martinez-Guryn *et al.*, “Small Intestine Microbiota Regulate Host Digestive and

Absorptive Adaptive Responses to Dietary Lipids.," *Cell Host Microbe*, vol. 23, no. 4, pp. 458-469.e5, 2018.

[106] A. K. Batta, G. Salen, R. Arora, S. Shefer, M. Batta, and A. Person, "Side chain conjugation prevents bacterial 7-dehydroxylation of bile acids.," *J. Biol. Chem.*, vol. 265, no. 19, pp. 10925–8, 1990.

[107] J. M. Lambert, R. J. Siezen, W. M. de Vos, and M. Kleerebezem, "Improved annotation of conjugated bile acid hydrolase superfamily members in Gram-positive bacteria.," *Microbiology*, vol. 154, no. Pt 8, pp. 2492–500, 2008.

[108] K. Yahiro, T. Setoguchi, and T. Katsuki, "Effect of coprophagy on bile acid metabolism in the rabbit.," *Gastroenterol. Jpn.*, vol. 14, no. 6, pp. 545–52, Dec. 1979.

[109] K. Hellström and J. Sjövall, "Turnover of deoxycholic acid in the rabbit," *J. Lipid Res.*, vol. 3, no. 4, pp. 397–404, Oct. 1962.

[110] I. D. Wilson and J. K. Nicholson, "Gut microbiome interactions with drug metabolism, efficacy, and toxicity.," *Transl. Res.*, vol. 179, pp. 204–222, 2017.

[111] A. LoGuidice, B. D. Wallace, L. Bendel, M. R. Redinbo, and U. A. Boelsterli, "Pharmacologic targeting of bacterial β-glucuronidase alleviates nonsteroidal anti-inflammatory drug-induced enteropathy in mice.," *J. Pharmacol. Exp. Ther.*, vol. 341, no. 2, pp. 447–54, May 2012.

[112] S. Seitz and U. A. Boelsterli, "Diclofenac acyl glucuronide, a major biliary metabolite, is directly involved in small intestinal injury in rats.," *Gastroenterology*, vol. 115, no. 6, pp. 1476–82, Dec. 1998.

[113] Z.-Y. Zhong *et al.*, "Ciprofloxacin blocked enterohepatic circulation of diclofenac and alleviated NSAID-induced enteropathy in rats partly by inhibiting intestinal β-glucuronidase activity.," *Acta Pharmacol. Sin.*, vol. 37, no. 7, pp. 1002–12, Jul. 2016.

[114] K. S. Saitta, C. Zhang, K. K. Lee, K. Fujimoto, M. R. Redinbo, and U. A. Boelsterli, "Bacterial β-glucuronidase inhibition protects mice against enteropathy induced by indomethacin, ketoprofen or diclofenac: mode of action and pharmacokinetics.," *Xenobiotica.*, vol. 44, no. 1, pp. 28–35, Jan. 2014.

[115] R. W. Blackler *et al.*, “Hydrogen sulphide protects against NSAID-enteropathy through modulation of bile and the microbiota.,” *Br. J. Pharmacol.*, vol. 172, no. 4, pp. 992–1004, Feb. 2015.

[116] S. Fiorucci *et al.*, “Activation of the farnesoid-X receptor protects against gastrointestinal injury caused by non-steroidal anti-inflammatory drugs in mice.,” *Br. J. Pharmacol.*, vol. 164, no. 8, pp. 1929–38, Dec. 2011.

[117] S. A. Mayo *et al.*, “Indomethacin injury to the rat small intestine is dependent upon biliary secretion and is associated with overgrowth of enterococci.,” *Physiol. Rep.*, vol. 4, no. 6, Mar. 2016.

[118] X. Liang, K. Bittinger, X. Li, D. R. Abernethy, F. D. Bushman, and G. A. FitzGerald, “Bidirectional interactions between indomethacin and the murine intestinal microbiota.,” *Elife*, vol. 4, p. e08973, 2015.

[119] J. L. Wallace *et al.*, “Proton pump inhibitors exacerbate NSAID-induced small intestinal injury by inducing dysbiosis.,” *Gastroenterology*, vol. 141, no. 4, pp. 1314–22, 1322.e1–5, Oct. 2011.

[120] H. Satoh, P. H. Guth, and M. I. Grossman, “Role of bacteria in gastric ulceration produced by indomethacin in the rat: cytoprotective action of antibiotics.,” *Gastroenterology*, vol. 84, no. 3, pp. 483–9, Mar. 1983.

[121] C. Scarpignato, “NSAID-induced intestinal damage: are luminal bacteria the therapeutic target?,” *Gut*, vol. 57, no. 2, pp. 145–8, Feb. 2008.

[122] G. L. Larsen, “Deconjugation of Biliary Metabolites by Microfloral β-glucuronidases, Sulphatases and Cysteine Conjugate β-lyases and their Subsequent Enterohepatic Circulation,” in *Role of the Gut Flora in Toxicity and Cancer*, Elsevier, 1988, pp. 79–107.

[123] A. F. Hofmann, “Targeting drugs to the enterohepatic circulation: Lessons from bile acids and other endobiotics,” *J. Control. Release*, vol. 2, pp. 3–11, Nov. 1985.

[124] P. V Turner, “The role of the gut microbiota on animal model reproducibility.,” *Anim. Model. Exp. Med.*, vol. 1, no. 2, pp. 109–115, Jun. 2018.

[125] I. I. Ivanov *et al.*, “Induction of intestinal Th17 cells by segmented filamentous bacteria.,” *Cell*, vol. 139, no. 3, pp. 485–98, Oct. 2009.

[126] K. Buhnik-Rosenblau, Y. Danin-Poleg, and Y. Kashi, “Predominant effect of host genetics on levels of Lactobacillus johnsonii bacteria in the mouse gut.,” *Appl. Environ. Microbiol.*, vol. 77, no. 18, pp. 6531–8, Sep. 2011.

[127] A. M. Snijders *et al.*, “Influence of early life exposure, host genetics and diet on the mouse gut microbiome and metabolome.,” *Nat. Microbiol.*, vol. 2, p. 16221, Nov. 2016.

[128] L. Xiao *et al.*, “A catalog of the mouse gut metagenome.,” *Nat. Biotechnol.*, vol. 33, no. 10, pp. 1103–8, Oct. 2015.

[129] S. P. Rosshart *et al.*, “Wild Mouse Gut Microbiota Promotes Host Fitness and Improves Disease Resistance.,” *Cell*, vol. 171, no. 5, pp. 1015-1028.e13, Nov. 2017.

[130] J. Shin *et al.*, “Analysis of the mouse gut microbiome using full-length 16S rRNA amplicon sequencing.,” *Sci. Rep.*, vol. 6, p. 29681, Jul. 2016.

[131] J. B. Treweek *et al.*, “Whole-body tissue stabilization and selective extractions via tissue-hydrogel hybrids for high-resolution intact circuit mapping and phenotyping.,” *Nat. Protoc.*, vol. 10, no. 11, pp. 1860–96, Nov. 2015.

[132] Y. Yamamoto *et al.*, “A Metabolomic-Based Evaluation of the Role of Commensal Microbiota throughout the Gastrointestinal Tract in Mice.,” *Microorganisms*, vol. 6, no. 4, Sep. 2018.

[133] C. Llorente *et al.*, “Gastric acid suppression promotes alcoholic liver disease by inducing overgrowth of intestinal Enterococcus.,” *Nat. Commun.*, vol. 8, no. 1, p. 837, 2017.

[134] H. Chung *et al.*, “Gut immune maturation depends on colonization with a host-specific microbiota.,” *Cell*, vol. 149, no. 7, pp. 1578–93, Jun. 2012.

[135] M.-C. Arrieta, J. Walter, and B. B. Finlay, “Human Microbiota-Associated Mice: A Model with Challenges.,” *Cell Host Microbe*, vol. 19, no. 5, pp. 575–8, May 2016.

[136] K. Schaarschmidt *et al.*, “[Improved model of a fecal collection device for the

prevention of coprophagia in the rat].," *J. Exp. Anim. Sci.*, vol. 34, no. 2, pp. 67–71, 1991.

[137] R. E. Smyth, "Fecal cup for collection of feces in male rats.," *Lab. Anim. Sci.*, vol. 29, no. 5, pp. 677–8, Oct. 1979.

[138] D. L. Frape, J. Wilkinson, and L. G. Chubb, "A simplified metabolism cage and tail cup for young rats.," *Lab. Anim.*, vol. 4, no. 1, pp. 67–73, Apr. 1970.

[139] J. Hoff, "Methods of blood collection in the mouse," *Lab Anim. (NY).*, vol. 29, no. 10, pp. 47–53, Nov. 2000.

[140] "AVMA Guidelines for the Euthanasia of Animals: 2013 Edition," 2013.

[141] J. L. Oblinger and J. A. Koburger, "Understanding and Teaching the Most Probable Number Technique," *J. Milk Food Technol.*, vol. 38, no. 9, pp. 540–545, Sep. 1975.

[142] R. Rowe, R. Todd, and J. Waide, "Microtechnique for most-probable-number analysis.," *Appl. Environ. Microbiol.*, vol. 33, no. 3, pp. 675–80, Mar. 1977.

[143] L. Kuai, A. A. Nair, and M. F. Polz, "Rapid and simple method for the most-probable-number estimation of arsenic-reducing bacteria.," *Appl. Environ. Microbiol.*, vol. 67, no. 7, pp. 3168–73, Jul. 2001.

[144] C. Y. Chen, G. W. Nace, and P. L. Irwin, "A 6 x 6 drop plate method for simultaneous colony counting and MPN enumeration of Campylobacter jejuni, Listeria monocytogenes, and Escherichia coli.," *J. Microbiol. Methods*, vol. 55, no. 2, pp. 475–9, Nov. 2003.

[145] R. Ozkanca, F. Saribiyik, K. Isik, N. Sahin, E. Kariptas, and K. P. Flint, "Resuscitation and quantification of stressed Escherichia coli K12 NCTC8797 in water samples.," *Microbiol. Res.*, vol. 164, no. 2, pp. 212–20, 2009.

[146] R. Blodgett, "BAM Appendix 2: Most Probable Number from Serial Dilutions," 2010. [Online]. Available: https://www.fda.gov/food/laboratory-methods-food/bam-appendix-2-most-probable-number-serial-dilutions.

[147] L. Aigrain, Y. Gu, and M. A. Quail, "Quantitation of next generation sequencing library preparation protocol efficiencies using droplet digital PCR assays - a

systematic comparison of DNA library preparation kits for Illumina sequencing.," *BMC Genomics*, vol. 17, p. 458, 2016.

[148] "KAPA Library Quantification Technical Guide v2.19," 2019.

[149] "Illumina Adapter Sequences 1000000002694 v11," 2019.

[150] E. Bolyen *et al.*, "Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2.," *Nat. Biotechnol.*, vol. 37, no. 8, pp. 852–857, Aug. 2019.

[151] B. J. Callahan, P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes, "DADA2: High-resolution sample inference from Illumina amplicon data.," *Nat. Methods*, vol. 13, no. 7, pp. 581–3, 2016.

[152] "Taxonomy classifiers for use with q2-feature-classifier." [Online]. Available: https://docs.qiime2.org/2019.1/data-resources/#taxonomy-classifiers-for-use-with-q2-feature-classifier.

[153] N. A. Bokulich *et al.*, "Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin.," *Microbiome*, vol. 6, no. 1, p. 90, 2018.

[154] "SILVA rRNA database project." [Online]. Available: https://www.arb-silva.de/download/archive/qiime.

[155] C. Zaiontz, "Real Statistics Resource Pack (Release 6.2)." 2019.

[156] S. van der Walt, S. C. Colbert, and G. Varoquaux, "The NumPy Array: A Structure for Efficient Numerical Computation," *Comput. Sci. Eng.*, vol. 13, no. 2, pp. 22–30, Mar. 2011.

[157] W. McKinney, "Data Structures for Statistical Computing in Python," in *Proceedings of the 9th Python in Science Conference*, 2010, pp. 51–56.

[158] E. Jones, T. Oliphant, P. Peterson, and others, "{SciPy}: Open source scientific tools for {Python}." .

[159] S. Seabold and J. Perktold, "Statsmodels: Econometric and statistical modeling with python," in *9th Python in Science Conference*, 2010.

[160] J. D. Hunter, “Matplotlib: A 2D Graphics Environment,” *Comput. Sci. Eng.*, vol. 9, no. 3, pp. 90–95, 2007.

[161] M. Waskom *et al.*, “mwaskom/seaborn: v0.9.0 (July 2018).” Jul-2018.

[162] F. Perez and B. E. Granger, “IPython: A System for Interactive Scientific Computing,” *Comput. Sci. Eng.*, vol. 9, no. 3, pp. 21–29, 2007.

[163] T. Kluyver *et al.*, “Jupyter Notebooks-a publishing format for reproducible computational workflows.,” in *ELPUB*, 2016, pp. 87–90.

[164] Anaconda, “Anaconda Software Distribution. Version 4.7.11.” 2019.

[165] N. M. Davis, D. M. Proctor, S. P. Holmes, D. A. Relman, and B. J. Callahan, “Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data.,” *Microbiome*, vol. 6, no. 1, p. 226, 2018.

[166] G. B. Gloor, J. R. Wu, V. Pawlowsky-Glahn, and J. J. Egozcue, “It’s all relative: analyzing microbiome data as compositions.,” *Ann. Epidemiol.*, vol. 26, no. 5, pp. 322–9, 2016.

[167] D. Lovell, V. Pawlowsky-Glahn, J. J. Egozcue, S. Marguerat, and J. Bähler, “Proportionality: a valid alternative to correlation for relative data.,” *PLoS Comput. Biol.*, vol. 11, no. 3, p. e1004075, Mar. 2015.

[168] F. Pedregosa *et al.*, “Scikit-learn: Machine Learning in Python,” *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Jan. 2012.

[169] R. A. van den Berg, H. C. J. Hoefsloot, J. A. Westerhuis, A. K. Smilde, and M. J. van der Werf, “Centering, scaling, and transformations: improving the biological information content of metabolomics data.,” *BMC Genomics*, vol. 7, p. 142, Jun. 2006.