“The Human Gut Microbiome in the Context of Disease and Disruption”

Doctoral Thesis

By Aislinn Danielle Rowan-Nash

Graduate Program in Pathobiology

Department of Molecular Microbiology and Immunology

Brown University

© Copyright 2020 by Aislinn D. Rowan-Nash

This dissertation by Aislinn D. Rowan-Nash is accepted in its present form by the Pathobiology Graduate Program as satisfying the dissertation requirement for the degree of Doctor of Philosophy.

Date ______________________ ____
Peter Belenky, Ph.D., Advisor

Recommended to the Graduate Council

Date ________________ ____
Christopher de Graffenried, Ph.D., Reader

Date ________________ ____
Shipra Vaishnava, Ph.D., Reader

Date ________________ ____
Daniel Weinreich, Ph.D., Reader

Date ________________ ____
Vanni Bucci, Ph.D., External Reader

Approved by the Graduate Council

Date ________________ ____
Andrew G. Campbell, Ph.D., Dean of the Graduate School

CURRICULUM VITAE

Education

12.2019	**Doctor of Philosophy, Pathobiology**	Brown University
5.2014	**Bachelor of Arts, Psychology**	Lehigh University
1.2014	**Bachelor of Science, Biology (honors)**	Lehigh University

Research

5.2015-12.2019 **Doctoral Thesis**, Brown University
Department of Molecular Microbiology and Immunology
Laboratory of Peter Belenky

9.2012-5.2014 **Undergraduate Research/Honors Project**, Lehigh University
Department of Biological Sciences
Laboratory of Linda Lowe-Krentz

6.2013-7.2013 **Summer Research**, National University of Ireland - Galway
Department of Microbiology
Laboratory of James O'Gara

5.2012-7.2012 **Summer Research**, Lehigh University
Department of Biological Sciences
Laboratory of Linda Lowe-Krentz

1.2012-5.2013 **Undergraduate Research**, Lehigh University
Department of Psychology
Laboratory of Barbara Malt

8.2011-5.2012 **SEA-PHAGES Research Program**, Lehigh University
Department of Biological Sciences

Supervised by Vassie Ware

Publications

Journal Articles

**Indicates co-primary authorship*

1. Liu RT*, **Rowan-Nash AD***, Sheehan AE, Walsh RFL, Sanzari C, & Belenky P. "Reductions in anti-inflammatory gut bacteria are associated with major depressive disorder in a cohort of American young adults." *In review.*
2. **Rowan-Nash AD**, Araos R, D'Agata EMC, & Belenky P. "Antimicrobial resistance gene prevalence in a population of patients with advanced dementia is related to specific pathobionts." *In review.*
3. Cabral DJ, Penumutchu S, Reinhart EM, Zhang C, Korry BJ, Wurster JI, Nilson R, Guang A, Sano WH, **Rowan-Nash AD**, Li H, & Belenky P. "Microbial metabolism modulates antibiotic susceptibility within the murine gut microbiome." (Sep 2019.) Cell Metabolism, 30: 800-823.
4. Reinhart E, Korry BJ, **Rowan-Nash AD**, & Belenky P. "Defining the distinct skin and gut microbiomes of the northern pike (Esox lucius)". (Sep 2019.) Frontiers in Microbiology, 10:2118.
5. **Rowan-Nash AD**, Korry BK, Mylonakis E, & Belenky P. (January 2019) "Cross-domain and Viral Interactions in the Microbiome." *Microbiology and Molecular Biology Reviews*, 83:e00044-18.
6. Ajibola O*, **Rowan AD***, Ogedengbe CO, Mshelia MB, Cabral DJ, Eze AA, Obaro S., & Belenky P. (January 2019) "Urogenital schistosomiasis is associated with signatures of microbiome dysbiosis in Nigerian adolescents." *Scientific Reports*, 9(1):829.

7. Cabral DJ, Wurster JI, Flokas ME, Alevizakos M, Zabat M, Korry BJ, **Rowan AD**, Sano WH, Andreatos N, Ducharme RB, Chan PA, Mylonakis E, Fuchs BB, & Belenky P. (September 2017.) "The salivary microbiome is consistent between subjects and resistant to impacts of short-term hospitalization." *Scientific Reports*, 8;7(1):11040. PMCID: PMC5591268.

8. **Rowan AD**, Cabral DJ, & Belenky P. (April 2016.) "Bactericidal antibiotics induce programmed metabolic toxicity." *Microbial Cell, 3(4),* 178-180. PMCID: PMC5349092.

9. Pope WH, Bowman CA, Russell DA, Jacobs-Sera D, Asai DJ, Cresawn SG, Jacobs WR, Hendrix RW, Lawrence JG, Hatfull GF, SEA-PHAGES ([…], **Rowan AD**[#], […]), PHIRE, Mycobacterial Genetics Course. (April 2015.) "Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity." *eLife,* 4:e06416. PMCID: PMC4408529. [#]*consortium author*

Genome Annotations

10. Carson AB, Lai PA, **Rowan AD**, Santana MA, Sullivan BJ, Marzillier JY, Kenna MA, Mageeney CM, Ware VC, Bowman CA, Russell DA, Pope WH, Jacobs-Sera D, & Hatfull GF. "Mycobacterium phage Halley, complete genome." (May 2018.) *NCBI GenBank*, Accession Number MH077579.1.

11. Kim BS, Harrison JM, Bowen JP, **Rowan AD**, Mageeney CM, Marziller JY, Kenna MA, Ware VC, Bowman CA, Russell DA, Pope WH, Jacobs-Sera D, & Hatfull GF. (April 2018.) "Mycobacterium phage Tyke, complete genome." *NCBI GenBank,* Accession Number MH051261.1.

12. Bowen JP, Carson AB, Deyang T, Frangos S, Godfried-Sie CP, Govindan J, Hanna S, Kenna MA, Lai PA, Lewis JK, Magier SJ, Marzillier JY, Petersen CL, Robertson ME, **Rowan AD**, Santana MA, Sin DN, Stomel JS, Sullivan BJ, Ware VC, Wang X, Crowell R, Bostrom MA, Burke M, Wright GM, Gregory SG, Colman SD, Bradley KW, Khaja R, Lewis MF, Barker LP, Asai DJ, Bowman CA, Russell DA, Pope WH, Jacobs-Sera D, Hendrix RW & Hatfull GF. "Mycobacterium phage Wanda, complete genome." (July 2014.) *NCBI GenBank*, Accession Number KF006818.1.

Presentations

1. **Rowan-Nash AD**, Araos R, D'Agata E, & Belenky P. (October 2019) "Antibiotic Resistance Gene Burden in Elderly Dementia Patients is Predicted By theAbundance of Three Enteric Bacteria." Poster at Keystone Symposium on Microbiome: Therapeutic Implications, Killarney, Ireland.
2. **Rowan AD**, Ajibola O, & Belenky P. (June 2018) "*Schistosoma haematobium* Infection is Associated with Widespread Changes in the Adolescent Gut Microbiome." Poster at American Society of Microbiology Conference, Atlanta, Georgia, USA.
3. **Rowan AD**, Ajibola O, & Belenky P. (June 2018) "*Schistosoma haematobium* Infection is Associated with Widespread Changes in the Adolescent Gut Microbiome." Poster at Boston Bacterial Meeting, Boston, Massachusetts, USA.
4. **Rowan AD** & Belenky P. (April 2017) "Stress Response and Natural Transformation in *Vibrio cholerae*." Poster at Experimental Biology Conference, with the American Association of Biochemistry and Molecular Biology, Chicago, Illinois, USA.

5. **Rowan AD**, Farwell SL, & Lowe-Krentz LJ. (April 2014). "Stress-activated signaling in endothelial cells". Poster and talk at Experimental Biology Conference, with the American Association of Biochemistry and Molecular Biology, San Diego, California, USA.
6. Bowen J.P., **Rowan AD**, Kenna M.A., Marzillier J.Y., Ware V.C. (June 2012). "Comparative Genomic Analysis of 4 Novel Mycobacteriophages." 4th Annual Howard Hughes Medical Institute SEA-PHAGES Symposium, Janelia Farm Research Campus, Ashburn, Virginia, USA.

Leadership and Institutional Service

2017-2019 **Student Member**, Admissions Committee
Pathobiology Graduate Program

2016-2017 **President**, Graduate Student Council
Brown University

2016-2017 **Standing Member**, Community Council
(*ex officio* as GSC President)
Brown University

2016-2017 **Guest Member**, Corporation Committee on Campus Life
(*ex officio* as GSC President)
Brown University

2016-2017 **Member**, New Alumni Trustee Candidate Advisory Committee
(*ex officio* as GSC President)

2016-2017 **Co-chair**, Finance Committee
(*ex officio* as GSC President)

Brown University Graduate Student Council

2016-2019 **Student Feedback Coordinator**, Recruitment

Pathobiology Graduate Program

2016 **Member**, Search Committee for Dean of the Graduate School

(*ex officio* as GSC President)

Brown University

2016 **Organizer**, Ivy+ Graduate Student Council Summit

(*ex officio* as GSC President)

Graduate Student Council

2015 **Voting Member**, Finance Committee

Brown University Graduate Student Council

2015 **Voting Representative** for Pathobiology Program

Brown University Graduate Student Council

Teaching

2015 **Teaching Assistant**, Eukaryotic Biology of the Cell

Brown University, with Kenneth Miller & Susan Gerbi

2014 **Teaching Assistant**, Biology Core I Laboratory

Lehigh University, with Margaret Kenna

2011-2012 **Undergraduate Teaching Assistant**, General Chemistry I Laboratory

Lehigh University, with Marcos Pires, James Roberts, & Andy Ho

2012-2013 **Undergraduate Teaching Assistant**, Organic Chemistry Laboratories

Lehigh University, with Aliana Lungu

2013-2014 **Tutor**, General Physics I & II, Biology, & Chemistry

Lehigh University, with the Center for Academic Success

Honors and Awards

2014-2017 **Presidential Fellowship**

Brown University

2014 **Student Travel Award**

National Science Foundation & American Society for Biochemistry and Cell Biology

2014 **Langer-Simon Award**

Lehigh University Department of Biological Sciences

2014 **Outstanding Academic Achievement Award**

Lehigh University Department of Psychology

2014 **President's Scholarship**

Lehigh University

2013 **Undergraduate Research Grant**

Lehigh University Department of Biological Sciences

2010-2014 **National Merit Scholarship**

Lehigh University & National Merit Foundation

2010-2014 **Trustee's Scholarship**

Lehigh University

2010-2014 **Dean's List**

Lehigh University

Professional and Honors Societies

2016-present American Society for Biochemistry and Molecular Biology (ASBMB)

2016-present American Chemical Society (ACS)

2016-present American Society for Microbiology (ASM)

2014-present Phi Beta Kappa Honors Society

2012-present National Society of Collegiate Scholars

2012-present National Society of Leadership and Success

2010-present Phi Eta Sigma Honors Society

2010-2014 Lehigh University Eckardt Scholar Honors Program

Classes and Certifications

2015-2016 **Certificate One – Reflective Teaching**, Harriet Sheridan Center for Teaching and Learning, Brown University

2016 **Strategies and Techniques for Analyzing Microbial Population Structures (STAMPS)**, Marine Biological Laboratory, Woods Hole, MA

PREFACE

All experiments, sample preparation, computational analysis, figure generation, and writing described in this thesis were performed by me under the supervision of Peter Belenky, with the following exceptions:

1. The Introduction is adapted from a review article published in *Microbiology and Molecular Biology Reviews*. Figure One was modified from its initial form by illustrator Patrick Lane to conform with the journal's figure style.
2. Chapter One is adapted from a manuscript published in *Scientific Reports*. The study was initiated in collaboration with and conceptualized by Olumide Ajibola and samples were collected by him and his team. He also contributed to the manuscript's introduction and methods, and Damien J. Cabral provided advice during analysis and edited the manuscript.
3. Chapter Two is adapted from a manuscript in review. The study was initiated in collaboration with and conceptualized by Richard Liu and samples were collected by him and his team. He also contributed to the manuscript's introduction and methods and edited the manuscript.
4. Chapter Three is adapted from a manuscript in review. In this study, access to data from the Study of Pathogen Resistance and Antimicrobial use in Dementia (SPREAD) was provided by our collaborators Erika D'Agata and Rafael Araos, who also edited the manuscript.

ACKNOWLEDGEMENTS

To the Graduate Program in Pathobiology and Department of Molecular Microbiology and Immunology of Brown University – It has been nearly six years since I applied to this program, and I will be forever grateful for being given the chance to pursue my doctoral studies here. Everyone, from faculty to students to administrators, has been unfailingly supportive and dedicated to making both program and department a welcoming place for graduate students. While I am excited to pursue my future endeavors, I will miss my time here. I would particularly like to acknowledge Michele Welindt – you have never ceased to amaze me (and everyone I know here) with your above-and-beyond commitment to and enthusiasm for improving the lives of all of the Pathobiology graduate students.

To my thesis committee – My graduate work has taken a number of turns throughout the years, and I deeply appreciate your support and flexibility as my research went in new directions. Chris and Shipra, I am thankful to have had your expertise, advice, and direction from the beginning, whether the focus was on traditional microbiology or the microbiome. Dan, I am grateful that you were able to join my committee, and I'm so glad to have had your voice in this process. Vanni, I deeply appreciate your agreeing to serve as my outside reader and I look forward to hearing your thoughts on my work.

To my PI, Dr. Peter Belenky – No matter how supportive a program might be, being in the right laboratory may be even more important. I'm so thrilled that I had the opportunity to join yours, and I will always feel honored to have been one of your first graduate students. You've been supportive and encouraging throughout all of the uncertainty we've encountered and new territory we've ventured into, and I can't imagine

having done my thesis work anywhere else. Thank you for giving me a chance five years ago, and I wish nothing but the best for you and the fantastic laboratory you've created.

To my colleagues – Even with the best supervisor, coworkers can make or break a work environment. Fortunately, I've gotten to work with you, who have created an incredibly collaborative, supportive, and fun laboratory environment that I will miss so much when I leave. From wide-ranging lab discussions to conference adventures to late-night baking gatherings to movie nights, I can't imagine a better team. In particular, I'd like to acknowledge Damien Cabral and Benjamin Korry, who have been with me on this journey from the very beginning and with whom it has been a privilege to work these last few years.

To my family – You've supported me and my love for science from the very beginning, well before any of us knew that I'd end up pursuing a Ph.D. Mom and Dad, you always trusted me to figure out where I was going with this whole biology thing (even though our family isn't exactly brimming with scientists), and have always been there for me no matter what I needed. Patrick, despite our many differences growing up, we've always believed in each other. In the end, we've turned out a lot more similarly than probably anyone expected, and I look forward to the day that you join me in the "No, not *that* kind of doctor" club.

To my husband Alex – We met shortly after I started graduate school, and I couldn't feel luckier that we did. I can't imagine my graduate career without you in it, and I'll always be so grateful to have had you in my life during this time. Thank you for always being there with a joke to make me smile, distractions from grad student life to keep me sane, and (so much) patience for dealing with the innumerable times that I underestimated how long

something would take me in the lab. Brown was the perfect place for me to pursue my Ph.D. for a host of reasons, but not least because it led me to you.

And last but not least, to our cats, Petri, Sapphire, and Mojo – You will never read this, because, well, you're cats, but my Ph.D. journey has been immeasurably better for having had you in it. From keeping my lap warm when I'm working at my desk to fearlessly protecting our home from the red dot, you've never failed to cheer me up on hard days or make me feel even better on good ones. Thanks for always being there, even when you were deeply affronted that your food was 0.3 seconds late.

TABLE OF CONTENTS

Introduction: The Human Bacterial Microbiome 3

Background 5

Composition of the Bacterial and Archaeal Microbiota 7

Mutualistic Metabolism: Gut Microbes in the Digestive Tract 13

Immunomodulation and Bacterial "Old Friends" 17

Summary 24

Main Figure..25

Chapter One: Urogenital Schistosomiasis in Nigerian Adolescents 26

Abstract 28

Introduction..28

Results...31

Discussion 37

Conclusions..43

Methods..44

Main Figures and Tables...53

Supplementary Materials 59

Chapter Two: Major Depressive Disorder in American Young Adults 74

Abstract 76

Introduction..76

Results...80

Discussion 89

Conclusions..96

Methods....97
Main Figures and Tables....104
Supplementary Materials....111
Chapter Three: Levofloxacin in Elderly Adults with Advanced Dementia....120
Abstract....122
Introduction....123
Results....126
Discussion....137
Conclusions....143
Methods....144
Main Figures and Tables....153
Supplementary Materials....158
Discussion: The Human Gut Microbiota Across Disease and Geography....176
The Gut Microbiota Interacts with Extraintestinal Disorders....177
Signatures of Microbiota Disturbance and Inflammation Across Studies....179
Distinct Microbiota Composition in Nigerian and American Subjects....182
Unanswered Questions and Future Directions....184
Summary....190
Main Figure....192
References....193

LIST OF TABLES

Chapter One: Urogenital Schistosomiasis in Nigerian Adolescents............................4

Supplementary Table 1: Subject Demographics..65

Supplementary Table 2: Subject Lifestyle Data 66

Supplementary Table 3: Differentially Abundant KEGG Pathways 67

Supplementary Table 4: Differentially Abundant KEGG Orthologs 68

Supplementary Table 5: Helminth PCR Conditions and Primers...............................69

Supplementary Table 7: 16S rRNA Amplicon Primers..70

Supplementary Table 8: Sequencing Read Depths...71

Supplementary Table 9: Primers and Conditions for Genus-Specific qPCR 72

Chapter Two: Major Depressive Disorder in American Young Adults 74

Table 1: Subject Demographics..110

Supplementary Table 1: Total Psychotropic Medication Usage...............................116

Supplementary Table 2: Combinations of Psychotropic Medications 117

Supplementary Table 3: Taxa Associated with MDD and Control Cohorts 118

Chapter Three: Levofloxacin in Elderly Adults with Advanced Dementia 120

Supplementary Table 1: Subject Demographics..164

Supplementary Table 2: Sample Collection Timeline 165

Supplementary Table 3: Bin Quality Cutoff Criteria..166

Supplementary Table 4: Bins Selected for Analysis...167

Supplementary Table 5: BioProject IDs for Test Dataset...174

LIST OF ILLUSTRATIONS

Introduction..4

Figure 1: The Human Microbiota 25

Chapter One: Urogenital Schistosomiasis in Nigerian Adolescents...........................26

Figure 1: Alpha Diversity by Schistosomiasis...53

Figure 2: Beta Diversity by Schistosomiasis 54

Figure 3: Differentially Abundant Taxa by Schistosomiasis......................................55

Figure 4: Differentially Abundant Genera by Schistosomiasis 56

Figure 5: qPCR of Differentially Abundant Genera 57

Figure 6: Differentially Abundant Phyla by Schistosomiasis 58

Supplementary Figure 1: Beta Diversity by Gender..59

Supplementary Figure 2: Differentially Abundant Classes by Schistosomiasis 60

Supplementary Figure 3: Differentially Abundant Orders by Schistosomiasis 61

Supplementary Figure 4: Differentially Abundant Families by Schistosomiasis 62

Supplementary Figure 5: Differentially Abundant Lineages by Schistosomiasis 63

Supplementary Figure 6: PCR to Detect Helminth DNA in Samples 64

Chapter Two: Major Depressive Disorder in American Young Adults.....................74

Figure 1: Subject Demographics..104

Figure 2: Significant Alpha and Beta Diversity Metrics by MDD Status 105

Figure 3: All Differentially Abundant Taxa by MDD 106

Figure 4: Differentially Abundant Phyla by MDD 107

Figure 5: Differentially Abundant Genera by MDD...108

Figure 6: Differentially Abundant Taxa by Severity 109

Supplementary Figure 1: Other Alpha and Beta Diversity Metrics...........................111

Supplementary Figure 2: Differentially Abundant Classes, Orders, Families...........112

Supplementary Figure 3: Differentially Abundant Taxa by Drug Usage..................113

Supplementary Figure 4: Overlap of Psychotropic Drugs and Severity....................114

Supplementary Figure 5: Differentially Abundant MetaCyc Pathways 115

Chapter Three: Levofloxacin in Elderly Adults with Advanced Dementia 120

Figure 1: Subject Demographics..153

Figure 2: Alpha and Beta Diversity by Subject 154

Figure 3: Taxonomic and Antimicrobial Resistance Gene Composition 155

Figure 4: Resistance Gene Abundance During Specific Pathobiont Blooms 156

Figure 5: Relationship Between Pathobionts and Resistance Gene Density 157

Supplementary Figure 1: Abundances of Phyla by Subject......................................158

Supplementary Figure 2: Abundances of Species by Subject 159

Supplementary Figure 3: Metagenomics vs 16S rRNA Genus Classification 160

Supplementary Figure 4: Gene Ontology Terms by Subject....................................161

Supplementary Figure 5: Abundances of ARGs by Subject.....................................162

Supplementary Figure 6: MDRO and non-MDRO Bins of Same Species................163

Discussion...176

Figure 1: Differences Between American and Nigerian Control Microbiota............192

ABSTRACT

of "The Human Gut Microbiome in the Context of Disease and Disruption"

by Aislinn D. Rowan-Nash, Ph.D.

Brown University, May 2020

In recent years, it has become increasingly appreciated that the human microbiome plays an important role in health and disease. Optimally, the gut microbiota contributes to digestion, immune regulation, and general gastrointestinal health. When dysregulated or disrupted, however, the gut microbiota has been implicated in a number of disease states, including infectious disease, autoimmune disorders, and even mental health conditions. We have examined the composition and predicted function of the gut microbiota associated with several disease states and potential microbiome disruptions in human subjects. First, we found that the gut microbiome was significantly different between Nigerian adolescents infected with the bloodstream helminth *Schistosoma haematobium* and their uninfected peers, with signatures of potential gastrointestinal dysbiosis associated with infection despite the parasite's location in the vasculature of the urogenital system. Second, we found differences in the gut microbiota of young adults with major depressive disorder (MDD), particularly a reduction in the relative abundance of the genus *Faecalibacterium* and an increase in the abundance of the genus *Flavonifractor* in subjects with MDD. As *Faecalibacterium* has been associated with anti-inflammatory effects and *Flavonifractor* has been previously associated with a number of inflammatory and psychiatric disorders, this may support research suggesting that MDD is associated with chronic low-grade inflammation. Finally, we found that the gut microbiota of institutionalized elderly patients with advanced dementia was temporally unstable and not significantly impacted by

administration of the fluoroquinolone levofloxacin. However, we did find that the level of antibiotic resistance genes (ARGs) in the samples could be predicted by the relative abundance of the three potential pathogens *Escherichia coli*, *Proteus mirabilis*, and *Enterococcus faecalis*, and metagenomic assembly and binning revealed that these species tended to have higher levels of ARGs than other abundant microbiota members. Together, these studies highlight the important role of the gut microbiota as potential contributors to or markers of non-gastrointestinal diseases and as reservoirs of ARGs in the human host.

INTRODUCTION: THE HUMAN BACTERIAL MICROBIOME

Adapted from

"Cross-Domain and Viral Interaction in the Microbiome"

By Aislinn D. Rowan-Nash, Benjamin J. Korry,

Eleftherios Mylonakis, and Peter Belenky

Published in January 2019 in

Microbiology and Molecular Biology Reviews, *83:e00044-18.*

CONTENTS

Background 5

Composition of the Bacterial and Archaeal Microbiota 7

Mutualistic Metabolism: Gut Microbes in the Digestive Tract 13

Immunomodulation and Bacterial "Old Friends" 17

Summary 24

Main Figure 25

Cross-Domain and Viral Interactions in the Microbiome

Aislinn D. Rowan-Nash[1], Benjamin J. Korry[1], Eleftherios Mylonakis[2], & Peter Belenky[1#]

[1]Department of Molecular Microbiology and Immunology, Brown University, Providence, Rhode Island, USA

[2]Infectious Diseases Division, Warren Alpert Medical School of Brown University, Rhode Island Hospital, Providence, Rhode Island, USA

[#]Corresponding author: Peter Belenky, peter_belenky@brown.edu

Background

Over the past several years, the importance of the microbiome to human health and disease has become increasingly recognized. The trillions of microbes, outnumbering even our own cells, that live in and on us can protect us from colonization by pathogens, promote immunoregulation and tolerance by our own immune systems, and digest many of the foods that we ourselves cannot. However, they can also contribute to disease, if their balance is disrupted by antibiotics, immune dysregulation, or other disturbances. The focus of this field has largely been on the bacterial members of the microbiome, as they make up the largest proportion of the living organisms which constitute the microbiota. However, the bacteria exist alongside a diversity of organisms from other domains of life: archaea, fungi, other unicellular eukaryotes, and in some cases helminths, as well as various families of viruses. All of these components can interact with each other and the host to impact health and disease. In this review, we will discuss the various elements of the microbiome, with particular focus on the cross-domain interactions within the microbiota and with the host.

Perhaps the clearest cross-domain interaction related to the microbiome occurs between the commensal bacterial and archaeal microbiota and the eukaryotic host. Colonizing microbes play a number of significant roles in the health of their host, and studies of germ-free animals have revealed that a lack of microbiota results in metabolic and immunological differences in comparison to conventional animals with a normal microbiota. Here, we provide an overview of the composition of the human bacterial and archaeal microbiota and briefly review two major impacts of these commensals on the

human host: liberation of energy and nutrients from food components and stimulation of the immune system to promote a tolerogenic environment.

Of all of the research on the microbiome, it is the bacterial component – sometimes called the "bacteriome" to differentiate it from other members of the microbiota – that has received the lion's share of the attention. Of that work, the majority has examined the gut bacteriome, with publications on that topic dwarfing the combined works on the oral, skin, and urogenital microbiota[1]. A number of robust tools and pipelines have been developed and made available for researchers to assess both the taxonomic classification and function of bacteria at multiple body sites and associated with various disease states[2-4]. Importantly, these methods allow analysis of the bacterial microbiota without the need to culture the species present; researchers can instead extract DNA from samples of interest and use next-generation sequencing technologies to assess composition and/or function of the microbes.

The most common method to analyze the composition of the gut bacteria is marker gene sequencing, generally using the 16S rRNA gene. Universal primers to amplify various regions of the 16S rRNA gene have been developed, and several databases exist to use such amplicons to taxonomically classify the bacterial sequences present within a biological sample[3-11]. This method has the benefit of being relatively simple and inexpensive and has thus been used extensively for bacteriome research. More recently, methods have been developed to predict bacterial metagenomes from 16S rRNA gene sequencing data[12,13]. However, there are several limitations: technological limits on amplicon length have led to the use of various subsections of the 16S rRNA gene rather than its full length, the primers used for each of these subsections may introduce biases for or against certain taxa during amplification, and different bacterial taxa have different numbers of copies of the 16S

rRNA gene[2-4]. Additionally, while this method can also be used to study archaea, the primers are typically optimized to detect bacterial communities and frequently fail to amplify archaeal 16S rRNA gene sequences in useful numbers. Furthermore, the databases for archaeal sequences are less complete, potentially leading to underrepresentation of archaea[14].

Accordingly, there is increasing interest in using shotgun metagenomics to profile the microbiome, as this removes some of the biases of marker gene amplicon sequencing and has the added benefit of assessing the functional potential of all of the genes present in a microbial community[2-4,15,16]. Furthermore, metagenomic approaches can assess the entire breadth of the community of interest, including eukaryotes, archaea, and viruses, rather than simply the bacterial members[2-4,15,16]. Even metagenomics, however, can only provide information about the composition of the community, and tools like multi-organism transcript arrays, metatranscriptomics, metaproteomics, and metabolomics are required to analyze the actual functions being performed by the communities at a given time[17-21]. However, these -omics methods are relatively more expensive, harder to implement, and suffer from a lack of complete and fully-annotated reference databases; as such, the ability to define the contributions of so-called "microbial dark matter" not represented in databases (including many archaea) is limited[14,22,23]. Thus at this time, -omics methods are less common than 16S rRNA gene sequencing, but they are becoming more widespread and are revealing important information about the microbiota[17-19].

Composition of the Bacterial and Archaeal Microbiota

Bacteria and archaea are present along the gastrointestinal tract, with the greatest density present in the colon, and have received much research attention due to their roles

in digestion and immune function[24]. Unsurprisingly, given the largely anaerobic environment of the gastrointestinal tract, the gut microbiota are primarily facultatively or strictly anaerobic[25,26]. The specific taxonomic composition can vary significantly between individuals, impacted by different lifestyles, diets, and ages, although generally they are fairly stable over time within the same individual[27,28]. Insights from metagenomics have led to the conclusion that rather than a set of specific taxa comprising a "core microbiota", there may instead by core functions that can be provided by different bacterial taxa in different individuals[27,29,30]. However, metatranscriptomics suggests that there is still inter-individual variation in transcription levels, which is intermediate between the highly idiosyncratic taxonomic composition and the more conserved functional capacity[31,32]. In the human gut, there appears to be a core metatranscriptome composed largely of housekeeping genes, with a much larger variable metatranscriptome of specialized pathways, suggesting that gut community transcription is context-specific and adaptive to the individual environment[32].

Despite this variation, sequencing, and particularly large-scale efforts including the Human Microbiome Project and Meta-HIT, have revealed some common patterns of bacterial composition[27,33]. The human gut is generally colonized by hundreds of species-level bacterial taxa, which typically are dominated by only a few phyla: primarily Firmicutes, Bacteroidetes, and Proteobacteria, with Actinobacteria and Verrucomicrobia making up smaller proportions (Figure 1)[25,27,30,34]. It should also be noted that the gut microbiota is not a single, homogenous community, but instead displays significant three-dimensional organization. First, the gut is comprised of several unique environments – in particular, the stomach, the small intestine (divided into the duodenum, jejunum, and

ileum), and the large intestine (colon) – which each has different properties and harbors its own community[35]. To date, the vast majority of research has focused on the colon due to the comparative ease of obtaining fecal samples and the fact that it contains by far the greatest density and numbers of bacteria[24]. Second, even within a given compartment, bacteria may differ along the transverse axis, with different populations found in the lumen versus the mucosa[35].

Despite the difficulties in studying the stomach and small intestine, techniques including endoscopy and biopsy have allowed profiling of these microbial communities. In general, the microbial community becomes increasingly anaerobic along the gastrointestinal tract, with the stomach and small intestine containing a greater proportion of facultatively aerobic taxa than the largely anaerobic colon[36]. Work in the stomach has frequently focused on the species *Helicobacter pylori*, given its close association with the gastric mucosa and public health relevance as an organism associated with gastric ulcers and cancers[37-40]. However, the presence and levels of *H. pylori* vary between individuals, and a combination of culturing and amplicon sequencing techniques have revealed that other genera can be found in the gut despite the harshly acidic conditions. While exact findings have differed, *Streptococcus* has been consistently observed in relatively high proportions, along with genera including *Prevotella*, *Lactobacillus*, *Rothia*, *Veillonella*, and *Propionibacterium*[41-45]. Additionally, while the stomach lumen certainly contains transient microbes from the mouth and nose, the gastric community was shown to be distinct from either of these groups[43].

The small intestine also contains a distinct community of bacteria, typically containing the genera *Streptococcus* and *Veillonella*; other frequently-encountered taxa

include *Escherichia*, *Clostridium*, *Turicibacter*, and *Lactobacillus* [46]. Like the stomach, it is less hospitable to bacterial life than the colon, with faster transit time, higher acidity, more antimicrobial molecules, and influx of bile acids, and therefore it is less densely populated along most of its length[46]. Studies of effluent from subjects with ileostomies suggests that the small intestinal bacterial microbiota tends to be more temporally variable than that of the colon, likely due to an increased short-term sensitivity to dietary intake given the small intestine's primary role in host nutrient absorption[47]. Indeed, metatranscriptomics indicates that the maintenance of the small intestinal bacteria is driven by the rapid uptake and utilization of simple carbohydrates, which could make this population particularly sensitive to the composition of ingested food[48]. In particular, the genus *Streptococcus* expressed genes for these functions at high levels, matching their high relative abundance in the population[48]. However, the community is not necessarily consistent along the entire small intestine, and there is evidence that the bacterial composition becomes more similar to that of the colon in the terminal ileum[46].

The colon is more diverse, densely colonized, and anaerobic. Firmicutes and Bacteroidetes make up the majority of bacteria, although Proteobacteria, Actinobacteria, and Verrucomicrobia are typically present in lower proportions. Within these phyla, a number of commonly prevalent bacterial families may be identified, including *Bacteroidaceae*, *Clostridiaceae*, *Prevotellaceae*, *Eubacteriaceae*, *Ruminococcaceae*, *Bifidobacteriaceae*, *Lactobacillaceae*, *Rikenellaceae*, *Verrucomicrobiaceae*, and *Enterobacteriaceae*[1,30,34,46]. Interestingly, within the phylum Bacteroidetes, individuals tend to be dominated by either *Bacteroides* or *Prevotella* based on their diet and lifestyle. Studies of urban subjects eating a "western" diet high in protein and fat tend to be

dominated by *Bacteroides*, while studies of rural communities eating more plant-based, fiber-rich diets are dominated by *Prevotella*[49]. Additionally, the archaeal genus *Methanobrevibacter*, which feeds on metabolites from other gut microbes and produces methane, is typically found in the human gut and is highly active; along with other less dominant methanogenic archaea, these organisms drive bacterial metabolism by removing hydrogen from the local environment and thereby making polysaccharide fermentation more thermodynamically favorable[1,30,32,34,50-53].

Furthermore, it has become increasingly recognized that the gut microbial community displays a transverse organization, with distinct composition in the lumen relative to the mucosa[25,46,51,54] . One reason for this is that luminal and fecal samples contain long-term residents alongside transient bacteria and DNA from the digesta, which is less true for mucosal communities. Another reason is that the intestinal mucus, composed of highly glycosylated mucin proteins, provides a distinct niche for certain microbiome members. In the colon, a continuous mucus barrier covers the epithelium, organized into a dense inner layer that blocks most bacteria and a loose outer layer adjacent to the lumen; in the small intestine, there is only a single layer and it is patchier than in the colon[46,55,56]. The outer layer is home to a number of bacteria, including primarily mucolytic species such as *Akkermansia muciniphila*, mucolysis-capable species such as *B. thetaiotamicron* and some *Bifidobacterium*, and non-mucolytic (and even asaccharolytic) species that can feed on downstream metabolites from this process[46,54,57]. There is also an oxygen gradient in the intestines, with higher oxygen concentrations at the epithelium relative to the largely anaerobic lumen[58]. This gradient tends to favor an enrichment of species that are more aerotolerant closer to the epithelium, including facultative anaerobes and those possessing

mechanisms such as catalase and superoxide dismutase to deal with oxidative stress[58,59]. Finally, some bacteria have adaptations for penetrating the mucus layers and coming in close contact with the epithelium, such as the segmented filamentous bacteria (SFB, sometimes known as *Candidatus* Arthromitus or *Candidatus* Savagella), while others can shelter in the crypts of the small intestine or the folds of the proximal colon[59-61]. As a result of these factors, a number of studies in humans and animal models have found that the communities and transcripts of the lumen or feces are distinct from those associated with the mucus and/or epithelium in the same individual[46,51,57-59,62-65]. Interestingly, even species found in both the lumen and mucus may behave differently based on their location, with work demonstrating differential transcriptional profiles observed between luminal and mucus-associated members of the same species[57].

In contrast to the bacteria in the colon, the oral community displays relatively low inter-individual variation (known as beta diversity), but has comparably high levels of diversity within any given individual (or alpha diversity)[30]. The oral community is frequently dominated by members of the genus *Streptococcus*, but also contains *Prevotella*, *Veillonella*, *Haemophilus*, *Neisseria*, *Corynebacterium*, *Actinomyces*, and *Rothia*, among others; it may also contain archaea including *Methanobrevibacter*[1,27,66]. However, like the gastrointestinal tract, there are several distinct regions within the mouth – including the gingiva, tongue, and teeth – that harbor somewhat distinct communities[27,66]. Similarly, the skin does not harbor a single unified bacterial community, and the composition depends on the characteristics of the site sampled – for example, dry skin, oily (sebaceous) skin, or moist skin[1,27,67-69]. In contrast to the gut, the skin is dominated by Actinobacteria, followed by Firmicutes, Proteobacteria, and Bacteroidetes; common genera include *Staphylococcus*,

Propionibacterium, and *Corynebacterium* [1,27,68-70]. In particular, the lipophilic genus *Propionibacterium* is associated with sebaceous sites[68-70]. Additionally, the skin is colonized by the Thaumarchaeota phylum of archaea, possibly involved in ammonia oxidation[14,71-73]. Finally, the vaginal bacterial community is an interesting demonstration of the fact that grouping bacteria at higher taxonomic levels can hide the diversity at lower levels. The vaginal community in most individuals is dominated by the genus *Lactobacillus*, giving it an apparent low diversity at this level, but the species and strains present are diverse and variable[27]. However, recent work has also revealed that a significant subset individuals possess a more diverse vaginal bacterial microbiota, including members of *Gardnerella*, *Atobium*, *Megasphaera*, *Streptococcus*, and *Prevotella*[74,75].

Mutualistic Metabolism: Gut Microbes in the Digestive Tract

As might be expected given their residence in the gastrointestinal tract, the gut bacterial and archaeal microbiota play an important role in digestion and metabolism. Collectively, gut bacteria possess the ability to extract energy from a wide variety of molecules that are indigestible by the host alone. Generally, these molecules are plant-derived polysaccharides, including fibers and starches, and are broken down into metabolites that can be used by the host or other microbes[76-78]. Indeed, metatranscriptomics studies indicate that carbohydrate transport and metabolism are highly-expressed functions across individual microbiomes, despite taxonomic variation[79]. The importance of including such molecules in the diet is highlighted by studies that suggest that in their absence, gut microbes may instead over-digest the mucus layer, potentially allowing epithelial access to pathobionts[80-82]. Additionally, non-fermentative members of the microbiota may form cooperative metabolic networks with the fermenters; for example, methanogenic archaea

in the gut remove excess hydrogen from the local environment, driving fermentation by increasing the thermodynamic efficiency of the process[52,53]. In fact, metatranscriptomics studies have indicated that methanogens are particularly active relative to some other members of the gut microbiota[31].

The importance of the gut microbiota in harvesting energy from food can be demonstrated by studies in germ-free animals, which lack any microbiota and display metabolic differences from their conventionally-raised counterparts. Germ-free mice are leaner than conventional mice despite consuming more food on a standard diet, but lose this phenotype when they are colonized with the gut microbes of their conventional counterparts[83,84]. This effect arises from the reduced capacity of germ-free mice to extract energy from food, thereby decreasing the caloric intake from the same amount of food, as well as the ability of the gut bacteria to promote fat deposition by the host[84]. Another study found that when on a high-fat diet, germ-free mice actually consumed less food than conventional mice, while also displaying increased lipid excretion and less efficient food utilization. Together, these effects resulted in lower weight gain than in the conventional mice, suggesting a degree of resistance to the ill effects of the high-fat diet[85]. Recent work further confirms this observation, as germ-free mice on a high-fat diet were shown to gain less weight, deposit less epididymal and mesenteric fat, excrete more triglycerides in the stool, and absorb significantly less lipid into the bloodstream than conventional mice; together, this suggests that gut microbes play an important role in lipid digestion and absorption[86,87].

Further studies demonstrated that not all microbiomes are equal. For example, the bacterial microbiota of genetically obese mice have been shown to be more efficient at

extracting dietary energy than those of their lean littermates. Obese mice showed an enrichment in bacterial genes for indigestible polysaccharide breakdown, produced more short-chain fatty acids (SCFA), and had lower fecal energy content, suggesting a greater ability to extract energy from their food. Furthermore, transferring microbiota from an obese mouse to a germ-free mouse resulted in a significantly larger body fat percentage increase compared to transferring microbes from a lean mouse[88]. In fact, transplanting fecal microbiota from humans has a similar effect; mice given microbiota from an obese human gained more weight and fat than mice given microbiota from a lean twin[89]. Interestingly, co-housing both types of mice together led the obese-transplant mice to resemble their lean-transplant counterparts in both bacterial microbiota and body composition, suggesting that the low-fat, high-fiber diet that the mice were provided with selected for the lean-associated microbes[89].

In fact, it is widely recognized that diet is a major factor that influences the makeup and function of the gut microbiota. For example, researchers comparing the gut bacterial microbiota of children in urban Italy and rural Burkina Faso found dramatic differences, including a high prevalence of fiber-digesting taxa and a significantly reduced Firmicutes-Bacteroidetes ratio in the African children compared with their European counterparts. The authors attribute these differences to the high-fiber, low-animal-protein diet of the African cohort, which promotes the growth of bacterial taxa capable of digesting dietary fibers and starches[90]. More experimental studies have further demonstrated the importance of diet in the makeup of the gut microbiota. One study found that a high-fat diet led to a reduction in Bacteroidetes and increases in Firmicutes and Proteobacteria even in an obesity-resistant mouse model[91], while another found that weight loss in obese humans was associated with

increases in Bacteroidetes[92]. Diet may also interact with microbes in the small intestine to regulate lipid absorption; recent work has found that colonizing germ-free mice with jejunal microbiota from mice fed a high-fat diet increases their capacity for lipid absorption even on a low-fat diet, while transferring microbes from mice fed a low-fat diet did not have the same effect[87]. Strikingly, Turnbaugh *et a*l found that switching mice from a low-fat, plant-rich diet to a high-fat, high-sugar diet could change microbial composition and metabolism in as little as a day[93], and the same group demonstrated alterations to the human gut microbiota after only 4 days on a plant-based or animal-based diet[94]. As such, there is significant research interest in the microbial contributions to obesity and metabolic disorders, as well as in whether the gut microbiota present a therapeutic target to treat or prevent these conditions. However, these efforts have been generally complicated by conflicting results and difficulty in finding a consistent signature of metabolic disruption across experiments[95,96].

In addition to simply liberating more energy from the diet, gut bacteria produce important metabolites that may promote host health. Many gut bacteria produce vitamins, particularly vitamin K and several B vitamins, although the amount absorbed by the host relative to the microbiota is unclear. More importantly, many of the gut-resident bacteria produce SCFA – primarily butyrate, acetate, and propionate – as end products of fermentation of undigested fiber, starches, and plant polysaccharides in the colon; in contrast, branched-chain fatty acids including isobutyrate, methylbutyrate, and isovalerate can also be produced as amino acid metabolism byproducts[77]. Acetate is produced by many enteric microbes, including the mucolytic *A. muciniphila*, *Bacteroides* species, and *Bifidobacterium* species. It enters peripheral circulation and is the primary SCFA

detectable in blood, and some functions include serving as a fuel source for the liver and muscles and being used in the synthesis of molecules such as cholesterol[77,97]. Acetate can also be used by other gut microbes to produce butyrate[77]. Propionate, produced by microbes including members of Bacteroidetes and the Negativicutes class of Firmicutes, is almost wholly metabolized in the liver and has impacts on gluconeogenesis[77,97]. Butyrate, which has received a significant amount of research interest, is primarily produced by members of Firmicutes, such as *Faecalibacter prausnitzii*. It is the primary fuel source for the colonic epithelium and has been implicated as an anti-inflammatory influence that helps to maintain intestinal homeostasis[97,98]. The concentrations of SCFA decline along the length of the colon, reaching 70-140 mM in the proximal colon and 20-70 mM in the distal colon, and also form a concentration gradient from the lumen outwards; furthermore, they are present at different molar ratios, with acetate being most abundant, followed by butyrate and propionate at approximately similar fecal levels, although this likely does not accurately represent the ratios in the colon itself due to differences in absorption[77,96,97,99,100]. In addition to their role in host metabolism, SCFA are implicated as important signalling molecules mediating interactions between the gut bacteria and the host immune system, described in more detail below.

Immunomodulation and Bacterial "Old Friends"

In addition to their role in metabolism, the human microbiota play an important role in the immunity of the host, which must be able to differentiate between commensal and/or symbiotic microbes and potentially pathogenic bacteria. Therefore, there is an important balance that develops, involving a limitation of contact between the microbiota and the local mucosa, in addition to immunoregulatory mechanisms allowing beneficial microbes

to persist while preventing autoimmunity or self-damage by the host. The contributions of commensal microbes to immunoregulation form an important part of the "old friends" (formerly "hygiene") hypothesis[101-105]. In short, this hypothesis posits that changes that have occurred in developed nations – including water sanitation, increased usage of antibiotics, higher rates of Caesarean sections, more time indoors, and shifts to a low-fiber "western" diet – have reduced early-life exposure to and colonization with helminths and beneficial microbes ("old friends" that humans co-evolved with) that help to regulate the immune system, thereby leading to increases in autoimmune and allergic disorders in their absence [101-105]. Here, we will focus on the bacterial component of this hypothesis, but we will discuss the contributions of helminths in a later section.

Humans are colonized with commensal microbes during and shortly after birth, and must develop an immune system that can tolerate bacteria at many body sites without losing the ability to defend against pathogens. According to the "old friends" hypothesis, if there is insufficient exposure to diverse commensal or environmental microbes, it can lead to failure to properly train immunological tolerance to harmless stimuli and to subsequent overreactions to allergens or innocuous microbes. At the same time, many beneficial commensal microbes actively regulate the immune response, helping to prevent inappropriate immune activation to both the microbes themselves and other "bystander" antigens[101,106,107]. Without this influence, particularly in early childhood, the risk for diseases of immune hyperreactivity such as asthma, type 1 diabetes, multiple sclerosis, and inflammatory bowel disease (IBD) increases. In fact, adults who immigrate from low- or mid-income nations to high-income nations tend to retain protection against such disorders,

but their children or those who immigrate when very young develop these diseases at higher rates more similar to indigenes of the new country[104,108-114].

There are several mechanisms by which key members of the commensal microbiota modulate the immune response. First, the mere presence of gut microbiota is required for proper immune development; studies in germ-free mice have revealed a number of immunological irregularities. For example, the microbiota are important for the development of the gut-associated lymphoid tissue (GALT), which allows the uptake and presentation of gut antigens to local immune cells. Accordingly, germ-free mice have underdeveloped GALT compared to conventional animals. Specifically, they have small Peyer's patches with fewer germinal centers, reduced numbers of $CD4^+$ T-cells in the lamina propria, and low levels of secretory IgA-producing plasma cells[115-118]. They also show signs of a T_H2-biased immune system, even in peripheral locations such as the spleen, and a decreased ability to develop oral tolerance to ingested antigens[119-123]. They also have increased accumulation of invariant natural killer cells in the colonic lamina propria, though they may be hyporesponsive[124,125]. Generally, such defects can be corrected by colonization with microbes at an early age, but not always in adulthood, supporting the importance of an early-life "critical window" for the microbiota to stimulate normal immune development[125-127]. Even certain single species of bacteria can serve to normalize some aspects of immune function in germ-free mice; for example, *Bacteroides fragilis* monocolonization can correct T_H2 bias and promote immunological balance[123], *Bifidobacterium infantis* can correct oral tolerance defects when administered to neonatal mice[128], and SFB (a lineage within the family *Clostridiaceae*) can direct balanced T-cell maturation comparable to a complete mouse microbiota[129].

The presence of the microbes is also important to the "education" of the adaptive immune system, training it to discriminate between innocuous commensals and harmful pathogens and thereby promoting tolerance of microbiota-derived antigens. A key component of this is the development of forkhead box $P3^+$ ($FoxP3^+$) regulatory T-cells (T_{reg} cells), which can suppress effector $CD4^+$ T-cell subsets and thereby promote immune tolerance. Traditional T_{reg} cells arise from the thymus, with the objective of suppressing self-reactivity by the immune system[130-132]; while these thymic cells (tT_{reg}) play a role in intestinal homeostasis, there is also an important role for naïve T-cells recognizing commensal antigens that are induced to differentiate into T_{reg} cells in the colon (iT_{reg})[126,133-136]. This occurs in part through the action of tolerogenic $CD103^+$ dendritic cells in the epithelium, which preferentially sample the luminal bacteria and favor the differentiation of naïve $CD4^+$ cells into iT_{reg} cells[137-141]. As might be expected, germ-free mice can display defects in their T_{reg} populations, although they do not lack them entirely[142,143]. The presence of colonic iT_{reg} cells with a diverse repertoire of receptors recognizing commensal antigens helps to prevent inappropriate responses to the microbiota and other "bystander" antigens, which has been implicated in the pathogenesis of IBD[144-146]. This is thought to be particularly important during early life; the microbiota of humans is temporally unstable for the first several years, and is theorized to provide a sampling window for the training and development of immunoregulatory responses[122,144-146].

In addition to simply serving to educate the adaptive immune system, several types of commensal microbes have been found to actually direct certain immune responses, often promoting tolerance[147,148]. For example, common gut microbe *B. fragilis* (of the phylum Bacteroidetes) has been found to activate development of T_{reg} cells and increase

immunoregulatory cytokine production via the molecule polysaccharide A (PSA)[149-151]. Accordingly, this molecule has been found to be protective against certain inflammatory diseases in mouse models[149-151]. The related species *B. thetaiotamicron* may be able to downregulate intestinal inflammation, even in the face of inflammatory challenge, by repressing host NFκB signalling[152,153]; some other microbes, including *Lactobacillus* species and non-virulent *Salmonella*, have demonstrated similar capabilities[153,154]. In addition, members of the class Clostridia (of the phylum Firmicutes) can induce expansion of thymic T_{reg} cells and development of colonic T_{reg} cells; this effect is at least partially mediated by the production of SCFA, particularly butyrate[140,155-160]. Specifically, the species *F. prauznitzii* has been found to be anti-inflammatory at least in part via its production of butyrate, inducing T_{reg} and anti-inflammatory cytokine production. Furthermore, it has been suggested to be protective against the development of IBD[161-163]. Further, colonization with altered Schaedler flora – a defined mix of eight commensal bacterial species that robustly colonize mice, including *Lactobacillus*, *Clostridium*, and *Bacteroides* species – has been shown to increase the levels of T_{reg} cells in the colonic lamina propria and promote intestinal immune homeostasis[164,165]. Finally, some *Lactobacillus* species have demonstrated an ability to drive T_{reg} development and subsequent IL-10 production[166]. Contrarily, SFB have been found to associate closely with the mucosa and induce a T_H17 response; the T_H17 response is a generally pro-inflammatory pathway that can help to protect against bacterial pathogens but potentially contribute to autoimmune pathology[60,148,167].

The commensal bacterial microbiota also promote the function of the gut epithelial barrier, the integrity of which is important to preventing inappropriate immune activation

and invasion by pathogens. A barrier of mucus, antimicrobial peptides, and secretory IgA serves to keep most microbes at a safe distance[168-173], although some are able to come in fairly close contact with the epithelium[61,174]. The commensal microbiota appear to serve as a stimulus for increased mucus production, as germ-free animals have been observed to have impaired mucus production which can be rescued via colonization with a normal microbiota or even administration of bacterial products including lipopolysaccharide[56,175-178]. Production of butyrate may contribute to this effect, as it has been demonstrated to promote epithelial production of the major mucus component mucin-2[33,179-181]. Butyrate can also promote epithelial barrier function and integrity[182-184]. Additionally, gut bacteria may stimulate the production of IgA and antimicrobial peptides[120]. While some of these impacts may seem counterproductive to the gut bacteria, they ultimately help both host and microbiota by maintaining a tolerant, anti-inflammatory environment. Furthermore, some commensal bacteria may be able to use host immune factors such as IgA to aid them in stable gut colonization[185].

Finally, commensal bacteria and archaea provide resistance to host infection with pathogens, a phenomenon termed colonization resistance. Commensal microbes occupy the readily-available niches of the sites they colonize and stimulate the local immune system, preventing potential pathogens from effectively establishing infections. They can compete for nutrients, produce antibacterial or inhibitory molecules, or even kill other bacteria through type 6 secretion systems; in contrast, commensal bacteria can also indirectly encourage resistance to pathogenic infection by promoting antimicrobial peptide production, epithelial barrier integrity, and T_H17 responses, as described previously. In one example of such immune-mediated competition, the Gram-negative *B. thetaiotamicron* can

stimulate the production of the anti-microbial peptide RegIIIγ, which primarily acts against Gram-positive bacteria[174,186,187]. Additionally, some interactions require a combination of interbacterial competition and host immune involvement; the probiotic *Escherichia coli* strain Nissle 1917 can antagonize *Salmonella enterica* colonization by competing for iron, but only when the host produces the innate immune molecule lipocalin-2 to limit bacterial iron availability[188]. As might be expected, germ-free mice or antibiotic-treated mice or humans are more susceptible to colonization with certain pathogens, including *S. enterica*, *Clostridium difficile*, *Klebsiella pneumoniae*, and pathogenic *E. coli*[174,189].

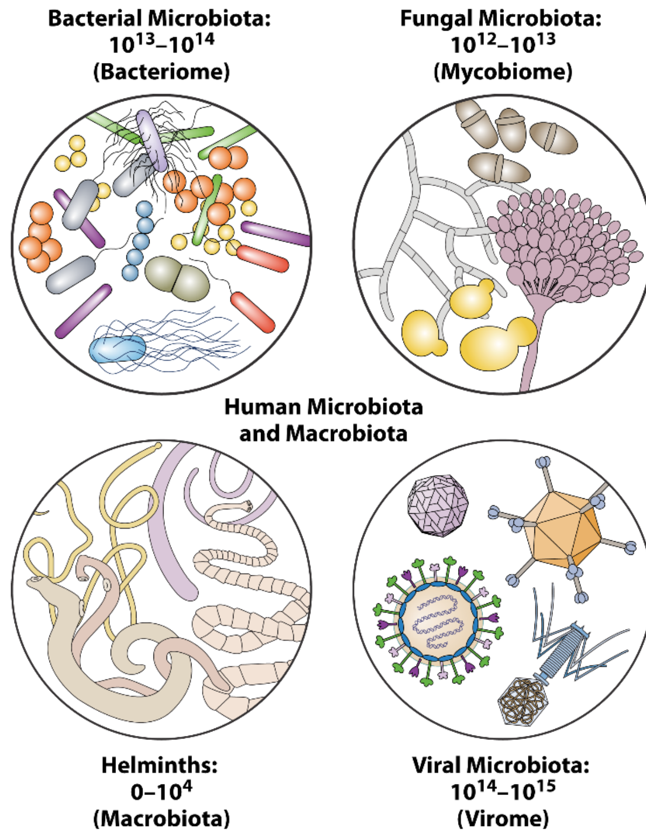
Given the importance of the commensal bacterial microbiota to immune regulation and colonization resistance, there is interest in using probiotics (specific strains or cocktails of bacterial species) and/or prebiotics (food or nutrients, typically fibers, meant to foster the growth of beneficial bacteria) as therapeutic agents. Of greatest interest are lactic acid bacteria, which are generally well-tolerated by humans and are often present in fermented foods. *Lactobacillus* and *Bifidobacterium* species are most commonly studied, although some other microbes including *Streptococcus*, *Lactococcus*, *Enterococcus*, and *E. coli* Nissle 1917 have been studied as well[190]. Both mouse and human studies have examined the potential for probiotics, sometimes in combination with prebiotics, to prevent or alleviate a wide variety of disorders, including antibiotic-related *C. difficile* infection, IBD, *H. pylori* infection, atopic disorders, and necrotizing enterocolitis in preterm infants, among others[191-194]. In this work are potentially promising results, although many studies are of small size or have methodological limitations so it is difficult to draw robust conclusions in some cases. However, large and well-designed studies can demonstrate the potential of pre- and probiotics; for example, Panigrahi, *et al* included over 4,000 subjects

in a randomized, double-blinded, placebo-controlled study of a combined pre- and probiotic (“synbiotic”) that showed a reduced risk of sepsis in full-term infants in rural India[195].

Summary

The bacterial and archaeal microbiota, particularly within the gastrointestinal tract, perform a number of important functions beneficial to the eukaryotic host. Most directly, they play a major role in digestion, allowing the host to extract energy from dietary components that the host does not possess the capacity to break down. In doing so, the gut microbiota produces SCFA, including butyrate, which serves as a primary fuel source for the colonic epithelium. Furthermore, the microbiota and their metabolites have significant impacts on the development and function of the host immune system. They stimulate innate mechanisms to shield the gut epithelium, protect the host against pathogenic colonization, and direct adaptive immune cell populations, particularly $T_{H}17$ and T_{reg} cells; in fact, the lack of a diverse community in early life may contribute to the development of immunological disorders in the genetically susceptible. In return, the eukaryotic host provides its microbial passengers with a sheltered niche and an array of nutrients, maintaining a tolerant environment despite the huge numbers of non-self organisms found in and on its body.

Figure 1: An outline of the major components of the human microbiota, summarized across body sites including the gastrotintestinal tract, oral cavity, vaginal mucosa, and skin. Bacteria (top left) are the most abundant, and include members of the phyla Firmicutes (*Clostridium*, *Faecalibacterium, Lactobacillus*, *Enterococcus*), Bacteroidetes (*Bacteroides*, *Prevotella*), Proteobacteria (*Escherichia*, *Acinetobacter*), Actinobacteria (*Bifidobacterium*), and Verrucomicrobia (*Akkermansia*). Based on metagenomics, human-associated fungi (top right) are significantly outnumbered by the bacteria; they are mainly members of the phylum Ascomycota (*Candida*, *Saccharomyces*, *Aspergillus*, *Malassezia*) but some Basideomycota are detectable. Humans may also be infected with non-fungal eukaryotic pathogens, which are not shown here. Viruses in the human microbiota (bottom right) are primarily bacteriophage and likely outnumber the bacterial population by at least ten-fold. The virome is largely composed of Caudovirales (Siphoviridae, Myoviridae, Podoviridae) and Microviridae, along with some eukaryotic host viruses. Helminths (bottom left) are now typically absent from humans in high-income nations, but still parasitize billions worldwide to varying degrees of severity. They include trematodes (flatworms), nematodes (roundworms), and cestodes (tapeworms).



CHAPTER ONE: UROGENITAL SCHISTOSOMIASIS IS ASSOCIATED WITH SIGNATURES OF MICROBIOME DYSBIOSIS IN NIGERIAN ADOLESCENTS

Adapted from

"Urogenital schistosomiasis is associated with signatures of microbiome dysbiosis in Nigerian adolescents"

By Olumide Ajibola, Aislinn D. Rowan*, Clement O. Ogedengbe, Mari B. Mshelia, Damien J. Cabral, Anthonius A. Eze, Stephen Obaro, and Peter Belenky*

Published in January 2019 in Scientific Reports *9(1):829*

**These authors contributed equally.*

CONTENTS

Abstract....28

Introduction....28

Results....31

Discussion....37

Conclusions....43

Methods....44

Main Figures and Tables....53

Supplementary Materials....59

Urogenital schistosomiasis is associated with signatures of microbiome dysbiosis in Nigerian adolescents

Olumide Ajibola[1#*], Aislinn D. Rowan[2*], Clement O. Ogedengbe[3], Mari B. Mshelia[1], Damien J. Cabral[2], Anthonius A. Eze[3], Stephen Obaro[4,5], & Peter Belenky[2#]

[1] Department of Microbiology, Faculty of Science, Federal University Birnin Kebbi, Birnin Kebbi, Kebbi State, Nigeria

[2] Department of Molecular Microbiology and Immunology, Division of Biology and Medicine, Brown University, Providence, RI, USA

[3] Department of Medical Biochemistry, College of Medicine, University of Nigeria - Enugu Campus, Enugu, Nigeria

[4] Division of Pediatric Infectious Diseases, University of Nebraska Medical Center, Omaha, NE, USA

[5] International Foundation Against Infectious Diseases in Nigeria, Department of Pediatrics, Bayero University Kano, Kano, Nigeria

*These authors contributed equally to this work.

[#]Corresponding authors: Peter Belenky, peter_belenky@brown.edu; Olumida Ajibola, olumide.ajibola@fubk.edu.ng (OA)

Abstract

Urogenital schistosomiasis is a neglected tropical disease caused by the parasite *Schistosoma haematobium*, which resides in the vasculature surrounding the urogenital system. Previous work has suggested that helminthic infections can affect the intestinal microbiome, and we hypothesized that *S. haematobium* infection could result in an alteration of immune system-microbiota homeostasis and impact the composition of the gut microbiota. To address this question, we compared the fecal microbiomes of infected and uninfected schoolchildren from the Argungu Local Government Area of Kebbi State, Nigeria, detecting significant differences in community composition between the two groups. Most remarkably, we observed a decreased abundance of Firmicutes and increased abundance of Proteobacteria – a shift in community structure which has been previously associated with dysbiosis. More specifically, we detected a number of changes in lower taxa reminiscent of inflammation-associated dysbiosis, including decreases in Clostridiales and increases in *Moraxellaceae*, *Veillonellaceae*, *Pasteurellaceae*, and *Desulfovibrionaceae*. Functional potential analysis also revealed an enrichment in orthologs of urease, which has been linked to dysbiosis and inflammation. Overall, our analysis indicates that *S. haematobium* infection is associated with perturbations in the gut microbiota and may point to microbiome disruption as an additional consequence of schistosome infection.

Introduction

Schistosomiasis, or bilharzia, is a parasitic disease that infects hundreds of millions of people each year and is endemic to various tropical regions, notably in Africa[196]. The disease is caused by infection with trematode helminths of the genus *Schistosoma*, which

live and sexually reproduce in the circulatory system of human hosts. Specifically, the species *S. mansoni* and *S. japonicum* live in venules surrounding the gut, while *S. haematobium* lives in the vessels around the urogenital system. There, adult worm pairs produce eggs that migrate through the surrounding tissue to be excreted primarily in the feces or urine, depending on the species, with the ultimate goal of reaching freshwater sources. They then reproduce asexually in their intermediate host – freshwater snails – before infecting humans present in contaminated water, entering through the skin before migrating to the vasculature[197-202]. The disease is typically diagnosed by microscopic examination of feces or urine for the presence of schistosome eggs[198], although some more sensitive techniques have been developed[203]. Treatment of schistosomiasis by administration of the anti-helminth drug praziquantel is the main control strategy employed in endemic areas[204].

The pathology of the disease generally arises from immunological reactions to eggs that become lodged in the tissue surrounding the urogenital system while attempting to migrate to the bladder lumen. The eggs generally provoke a T_H2 immune response, which is characteristic of extracellular insults including helminths and their eggs, leading to granuloma formation and fibrotic lesions that can have severe long-term consequences[199,205-211]. Eventually, the immune response is down-regulated, helping to preserve host health and integrity but allowing the parasite to persist for years[199,212,213]. This altered immune state may interplay with other immune insults, reducing the effectiveness of certain vaccines and altering the course of viral, bacterial, and parasitic co-infections[199,202,214-226]. On the other hand, it may also help to reduce the prevalence or

severity of autoimmune disorders, and there is research interest in the therapeutic potential of helminths or their antigens to treat inflammatory conditions[227-233].

There is evidence that both systemic immunological changes and helminth infection specifically are associated with changes in the gut microbiota. A number of previous studies have indicated that infection with a range of helminths – including gastrointestinal nematodes, tapeworms, tissue flukes, and schistosomes – can have impacts on the composition and function of the gut microbiome, suggesting that alterations to the gut microflora may be an under-recognized side effect of helminth infection[234-245]. However, most of this work has been done in animal models or humans infected with intestinal parasites, making it difficult to separate systemic immunological changes from effects local to the intestinal niche. In contrast, while it can occasionally localize to the enteric system (particularly during heavy infection or co-infection with *S. mansoni*)[246,247], *S. haematobium* primarily lives within the vasculature surrounding the bladder and thus provides an opportunity to study whether helminth infection can impact the microbiome indirectly via systemic immunological or other changes that may disrupt gut homeostasis. Such a link between systemic immunity and the microbiota has been recently proposed in a Ugandan cohort, in which low CD4+ cell counts in HIV patients were associated with significant changes in the gut microbiome[248]; additionally, several studies suggest that immunosuppression can alter the composition and function of the gut microbiota[249-252]. Therefore, we hypothesized that urogenital schistosomiasis may disturb immune-microbial homeostasis and allow for changes in the resident taxa.

In this study, we investigated the impact of *S. haematobium* infection on the intestinal microbiome of adolescents aged 11-15 years in the Argungu Local Government

Area of Kebbi State, Nigeria. As assessed by the Nigerian Federal Ministry of Health, Kebbi State has the highest prevalence of *S. haematobium* infection in the country but a very low prevalence of *S. mansoni*, making it an ideal location to study impacts of urogenital schistosomiasis specifically[253]. Kebbi State also has a low prevalence of soil-transmitted helminths, decreasing the likelihood of coinfections[253]. We chose to focus on adolescent schoolchildren, as children and adolescents are most likely to be infected with *S. haematobium* due to exposure and immunological factors[233,254-260]. Additionally, detection of differences in the human microbiome can be difficult given significant variation between individuals, which can be influenced by age, sex, diet, disease states, and other conditions[261]; to help minimize some such confounding factors, we selected subjects living in the same region, attending the same school, and falling into a relatively narrow age range.

Results

Study Overview and Participants

In order to examine the differences in the gut microbiome of young adolescents infected with *S. haematobium*, we sequenced the fecal microbiomes of 49 adolescent students: 24 individuals infected with *S. haematobium* and 25 controls (Supplementary Table 1). A t-test indicated that the ages of the subjects do not significantly differ between the two groups (p=0.3228), and survey data indicates that important exposure and lifestyle factors are not systematically different (Supplementary Table 2). In both groups, most samples were from male students, as fewer girls attend school in the area and females are less likely to have schistosomiasis both in Kebbi State and elsewhere in Nigeria[253-256,258,262-265]. We performed analyses of community composition between male and female subjects

and found no significant differences or distinct PCoA clustering (Supplementary Figure 1); therefore, males and females were grouped together for overall analyses.

In our analysis, we sequenced the V4 region of the 16S rRNA gene and were able to identify most OTUs down to the genus level using the SILVA 16S database[7]. We analyzed alpha and beta diversity in the infected and uninfected subjects, in addition to examining differences in specific taxa through computational analysis and qPCR. Finally, we used the 16S sequencing results to predict the functional potential of the infected and control gut communities.

Metrics of Diversity Between Infection Groups

We first examined several metrics of alpha diversity, which measures the diversity of taxa within each individual microbial community, of infected and control adolescents (Figure 1). Observed OTUs reflects the taxonomic richness of the community (Figure 1A), the Shannon and Simpson Diversity Indices account for both richness and abundance of taxa (Figure 1B&1C), and Faith's Phylogenetic Diversity also considers the phylogenetic relatedness of the taxa (Figure 1D). Using all four metrics, there was no significant difference in alpha diversity between the schistosomiasis-infected and -uninfected subjects, indicating that infection does not systematically impact the diversity of an individual's gut microbiota.

In contrast, we found significant differences between the microbial communities of infected and uninfected subjects when examining beta diversity, which measures the divergence in community composition between different samples. Again, we tested this using multiple metrics: Bray-Curtis dissimilarity reflects differences in the taxa present independent of their relatedness, unweighted UniFrac distance indicates differences in taxa

while considering their phylogenetic relatedness, and weighted UniFrac also accounts for the abundances of the differential taxa. Using principal coordinate analyses (PCoA), we noted clustering of infected and control samples (Figure 2) and a permutational MANOVA indicated that this difference is statistically significant in all cases. We found the greatest difference between the groups using unweighted UniFrac, suggesting that differences in community composition could be driven by changes in low-abundance taxa.

Significantly Different Genera by Infection Status

Given the significant differences in beta diversity, we examined the differential abundance of taxa between the infected and uninfected subjects. In total, 1,660 unique OTUs were identified across all samples. As most OTUs were not identified down to the species level, we agglomerated our samples at the genus level to perform differential abundance analysis. We detected significant differences in 17 genera: 10 increased (*Megasphaera*, *Dialister*, *Acinetobacter*, *Prevotella*, *Alloprevotella*, *Desulfovibrio*, *Haemophilus*, *Peptococcus*, *Olsenella*, and uncultured *Coriobacteriaceae*) and 7 decreased (*Subdoligranulum*, *Parabacteroides*, uncultured *Erysipelotrichaceae*, *Ruminococcaceae incertae sedis*, *Peptostreptococcaceae incertae sedis*, *Clostridium sensu stricto* 6, and uncultured Mollicutes RF9) in infected adolescents (Figures 3,4). Collectively, these 17 genera comprised an average of 23% of the relative abundance of the microbiota of uninfected subjects, and all have been previously specifically associated with or arise from lineages associated with the human gut microbiota[266-279]. Decreases were mainly found within the phylum Firmicutes, particularly in the class Clostridiales, while increases were mainly found within the phylum Proteobacteria and the family *Veillonellaceae*.

Given that many of the genera that we found to be significant are of low abundance, we decided to use qPCR to independently verify the changes in several of these genera. We designed genus-specific 16S primers, validated their specificity against a mock community, and tested the abundances of each genus relative to total 16S rDNA present in the pooled genomic DNA of schistosomiasis-positive and -negative individuals. Despite differences in primers and methodologies between sequencing and qPCR, we were able to recapitulate differences in the abundances of *Prevotella*, *Peptococcus*, *Megasphaera*, *Olsenella, Dialister, Alloprevotella, Haemophilus,* and *Parabacteroides,* (Figure 5), confirming that these genera did change in abundance in the schistosomiasis-positive individuals. For *Subdoligranulum*, which decreased very slightly in infected adolescents, qPCR did not detect a difference between the groups.

Changes Across Taxonomic Levels

We then began to look at changes in community composition at higher taxonomic levels. At the phylum level, we noted that most phyla decreased in abundance in the schistosomiasis-positive group: we observed significant decreases in Firmicutes, Tenericutes, and Cyanobacteria, and a significant increase in Proteobacteria (Figures 3,6). We then analyzed differential abundances at the class, order, and family levels (Supplementary Figures 2-4) and identified several lineages that show significant differences across multiple taxonomic levels (Supplementary Figure 5).

Within the phylum Proteobacteria, a number of lineages demonstrated significant increases in abundance across multiple taxonomic levels. For example, the lineage from which the genus *Desulfovibrio* arises shows significant increases across all taxonomic levels, including family (*Desulfovibrionaceae*), order (Desulfovibrionales), class

(Deltaproteobacteria), and phylum (Proteobacteria) (Supplementary Figure 5A). However, as Deltaproteobacteria comprise a small proportion of the phylum, the increase in Proteobacteria is in fact largely driven by members of the class Gammaproteobacteria. While there was no significant difference at the class level, there were significant increases in two of its lineages: *Haemophilus*, including family *Pasteurellaceae* and order Pasteurellales, and *Acinetobacter*, including family *Moraxellaceae* and order Pseudomonadales (Supplementary Figure 5B). In the human gut, Proteobacteria are typically found at low abundances relative to the dominant phyla of Firmicutes and Bacteroidetes, but blooms in this phylum have been associated with dysbiosis[280-282].

Similarly, there are significant increases throughout the taxonomic lineage of *Megasphaera* and *Dialister*, on the family (*Veillonellaceae*), order (Selenomondales), and class (Negativicutes) levels (Supplementary Figure 5C). However, in this case, there is an overall decrease in the parent genus of Firmicutes. This may be related to the fact that Negativicutes, unlike the majority of Gram-positive Firmicutes, are diderms with distinct outer membranes containing lipopolysaccharides that cause them to stain Gram-negative[283,284]. Interestingly, it is hypothesized that these genes may have been laterally acquired from Proteobacteria[285], which also increase in infected adolescents; it is possible that this similarity gives both groups a competitive advantage in the schistosomiasis-associated microbiota.

Additionally, while the *Peptococcaceae* family from which the genus *Peptococcus* stems is significantly increased, the order (Clostridiales), class (Clostridia), and phylum (Firmicutes) are significantly decreased (Supplementary Figure 5D). In fact, *Peptococcus* is the only significant genus within the Clostridiales lineage that increases in infected

individuals, while the several other significant genera all decrease. For example, the related lineage of *Peptostreptococcaceae incertae sedis* shows reductions in abundance at all levels, reflecting the more typical pattern of members of Clostridiales and Firmicutes in general (Supplementary Figure 5D).

Finally, there are two lineages from less-common phyla that demonstrate reductions in abundance in schistomiasis-positive subjects: the Gastranaerophilales-Melainabacteria lineage of Cyanobacteria and the RF9-Mollicutes lineage of Tenericutes. Unlike most Cyanobacteria, the Melainabacteria are non-photosynthetic and rely on fermentation[286,287], while Mollicutes are distinguished from most other bacteria by their lack of a cell wall[288].

Functional Potential of the Microbial Communities

While taxonomic classifications of the microbial communities of the two groups is useful, we were also interested in the functional potential of the gut microbiome and how it might vary between infected and uninfected individuals[93,289]. In general, while there are often significant inter-individual differences in the taxonomic composition of the gut microbiome, the functionality of the resident taxa is relatively stable[27]. Recently, methods have become available to use the 16S content of a microbial community to infer the genomes present, and therefore the potential functionality of that community, in the absence of whole-genome sequencing data[12,13,290]. Importantly, it should be noted that this methodology is based on inference from known genomes, and therefore may not fully recapitulate the existing metagenomic content, relevant strain differences, and the contributions of understudied microbiota members.

We used the web-based tool Piphillin to predict changes in the functional potential of the microbiome by inferring the metagenomes from 16S sequences[12] and the tool MicrobiomeAnalyst to analyze differential abundance of both KEGG Orthologs and Pathways[291]. We identified two KEGG pathways that were enriched in schistosomiasis-positive individuals (Supplementary Table 3), as well as 35 KEGG orthologs that were significantly different between the two groups (Supplementary Table 4). We were particularly interested to see that the top enriched pathway, "atrazine degradation," was populated by the three subunits of bacterial urease (ureA, ureB, and ureC). Furthermore, there were increases in the ureD, ureE, ureF, ureG, and ureH orthologs, all urease accessory proteins, although these were not categorized into any KEGG pathways. Full metagenomics, metatranscriptomics, or functional assays could determine whether the increases observed here reflect a true increase in urease production or function in these microbial communities.

Discussion

We observed a general shift in the gut microbiome of adolescents infected with *S. haematobium* towards a state consistent with dysbiosis, with decreases in the dominant phylum Firmicutes and increases in the prevalence of the minor phylum Proteobacteria (Figure 6). At the genus level, where we focused our analysis, we observed significant changes in sixteen genera collectively comprising over 20% of the gut microflora (Figures 3,4, Data S1). Interestingly, many of the changes we observed have been associated with gut inflammation. This was surprising, as previously helminth infection has been shown to reduce inflammation, and has even been investigated as a therapy to ameliorate symptoms of inflammatory bowel disease[229,230]. The apparent contrast could result from the distant

location of *S. haematobium*; gut-resident helminths may exert local anti-inflammatory effects that are not observed in urogenital schistosomiasis. In general, however, these changes are consistent with our hypothesis that *S. haematobium* infection may impact the gut microbiota.

Several of these changes we found are similar to those observed in other studies of the microbiota of humans infected with helminths. Most directly, a study of Zimbabwean children found significant increases in several OTUs belonging to the genus *Prevotella* in *S. haematobium*-infected subjects[245], a change we also observed on the genus level. Additionally, several studies investigating the impacts of gut-resident soil-transmitted helminths (STH) demonstrated some of the same taxonomic changes we observed. For example, a study in a Malaysian population infected with multiple STH found increases in the order Bacteroidales[235]. Similarly, we observed increases in the *Prevotella* and *Alloprevotella* genera within Bacteroidales, although we also observed a decrease in the *Parabacteroides* genus in this order. Another study found increases in *Olsenella* and the *Desulfovibrio* lineage in individuals infected with STH in both Indonesian and Liberian populations, as well as associations between STH infection and the *Dialister* lineage in the Indonesian group and *Megasphaera* and *Peptococcus* in the Liberian group[234]. However, this study also generally found increased abundances of Clostridiales members associated with helminths, which is largely contrary to what we observed[234]. Finally, researchers studying an Ecuadorean population generally found minimal differences in the microbiota of children with STH infections, but did find significant reductions in members of Clostridiales in children with mixed *Trichuris trichiura* and *Ascaris lumbricoides* infections[236].

In addition to similarities with other studies of helminths and the human microbiome, we noted that some of the changes we observed were reminiscent of those seen in dysbiosis and inflammation. On the phylum level, decreases in the prevalence of Firmicutes have previously been associated with gut inflammation[144,280,292]. Firmicutes typically make up a significant proportion of the human gut microbiota, and some members are associated with immunoregulatory impacts. Firmicutes – particularly Clostridia – are associated with regulatory T-cell activation[142,156,159], which is important for the prevention of intestinal inflammation. Additionally, some members of the phylum – such as *Faecalibacterium prauznitzii*, a member of the Clostridiales-*Ruminococcaceae* lineage – have been shown to have anti-inflammatory effects due to production of the short-chain fatty acid butyrate and have been negatively correlated with inflammatory bowel disease[161].

Furthermore, increased levels of Proteobacteria have been associated with gut inflammation in a number of studies[280-282], although whether they are causative or symptomatic remains unclear. Inflammation is associated with increased levels of oxygen and production of nitrate by the gut epithelium; Proteobacteria are generally aerotolerant and some have the capacity to utilize nitrate, potentially allowing them to outcompete other members of the microbiota – such as Clostridia – and bloom during inflammatory conditions[281,282,293,294]. In addition to thriving in an inflammatory environment, Proteobacteria themselves may contribute to inflammation. Relevant to our study, *Desulfovibrio* and other sulfate-reducing bacteria have been associated with gut inflammation and colitis, potentially through their production of cytotoxic hydrogen sulfide[295-300]. Additionally, Gram-negative bacteria, such as Proteobacteria and

Negativicutes, can exacerbate existing gut inflammation through lipopolysaccharide infiltration into circulation[301,302]. Finally, it was recently shown that urease producers, potentially enriched in schistosomiasis-infected subjects, may contribute to a dysbiotic environment, favoring Proteobacteria at the expense of Clostridia and potentially promoting inflammation through increased nitrogen flux[303].

In addition to these phylum-level shifts, we observed some changes on lower taxonomic levels that were also associated with gut inflammation. A large study of the gut mucosal and stool microbiota in new-onset pediatric Crohn's disease patients revealed a number of changes that were similar to what we observed, including reductions in the order Clostridiales and the family *Erysipelotrichaceae* and increases in the families *Veillonellaceae* and *Pasteurellaceae* in patients with disease[304]. Similarly, we saw reductions in Clostridiales and a genus within *Erysipelotrichaceae* and increases in both *Veillonellaceae* and *Pasteurellaceae;* the exception is *Peptococcus*, which we saw increased within the Clostridiales family (Figure 3). Additionally, researchers observed reductions in the order Bacteroidales and increases in the family *Fusobacteriaceae* in the Crohn's disease patients[304]; we noted both increases (*Prevotella* and *Alloprevotella* within *Prevotellaceae*) and decreases (*Parabacteroides* within *Porphyromonadaceae*) within Bacteroidales in infected subjects, as well as increases in higher taxonomic levels (Fusobacteriia, Fusobacteriales) of the *Fusobacteriaceae* lineage (Figure 3).

In another study, researchers compared the microbiota associated with inflamed mucosa with normal tissue in ulcerative colitis patients, finding that inflamed mucosa was enriched in Proteobacteria and reduced in Firmicutes. Furthermore, these changes were driven largely by increases in the abundance of the Pseudomonadales-*Moraxellaceae*-

Acinetobacter lineage of Proteobacteria and decreases in the Clostridia-Clostridiales lineage, particularly *Ruminococcaceae*[305]. Reductions in *Ruminococcaceae* have also been observed in other studies of inflammatory bowel disease[306,307]. This is quite similar to our observations in the microbiota of schistosomiasis-infected adolescents, where we saw an enrichment in the *Acinetobacter* lineage (Figures 3, S5B) and reductions in many members of Clostridiales, including two *Ruminococcaceae* (*Subdoligranulum* and an *incertae sedis*) (Figure 3). Taken together, these results suggest that the gut microbiota of *S. haematobium*-infected adolescents may reflect an inflammatory environment.

Importantly, it should be noted that while some of our observations have been seen in inflammation-related contexts in other individuals, they are not diagnostic of inflammation and it is unknown whether urogenital schistosomiasis-infected adolescents actually experience intestinal inflammation. In the future, it may be prudent to profile gut inflammation in this population in conjunction with microbiome analysis, potentially through measuring fecal biomarkers such as calprotectin[308]. If, in fact, there is intestinal inflammation associated with schistosomiasis and microbiome alterations, the directionality of this effect would remain unclear; infection-mediated immunological shifts might allow a bloom of pro-inflammatory microflora or might cause inflammation that allows dysbiotic microbes to proliferate.

Additionally, a potential confounder is the presence of co-infection with enteric helminths. While we selected our region of study due to its low rates of these infections[253] and ruled out subjects with gastrointestinal symptoms, we also used PCR to check for the presence of these organisms in extracted fecal DNA. Using previously published species-specific primers[309,310], we did not detect *Ascaris* spp, *Ancyclostoma* spp, *Necator*

americanus, *Trichuris trichiura*, or *S. mansoni* in samples from either infected or uninfected subjects (Supplementary Figure 6). Additionally, while *S. haematobium* primarily excretes eggs through the bladder, in a small percentage of cases it can also take up residence in the enteric system and extrude eggs through the intestinal wall[246,247]. Therefore, we also used PCR to check for the presence of *S. haematobium* in the fecal samples, and did not detect it in samples from either group (Supplementary Figure 6). Importantly, while more sensitive than microscopic methods such as Kato-Katz[203,311], even PCR is not perfectly sensitive on a single stool sample for detection of infection with gastrointestinal helminths or schistosomes; thus, while unlikely, it is still possible that some subjects may have those underlying infections and that a portion of our dysbiotic signal may originate from such intestinal morbidity.

Additionally, while we hypothesized that *S. haematobium* infection could lead to alterations in the gut microbiota due to its impacts on immune function, we cannot discount the possibility that adolescents with pre-existing dysbiosis may be more susceptible to successful schistosome infections, potentially due to immunological changes mediated by the gut microflora. For example, there is evidence that the gut microbiome influences the course of infection with *S. mansoni*, potentially via immunoregulatory effects; abolishing the gut microbiome of mice infected with *S. mansoni* reduces gut inflammation and egg excretion, although this may be due to local interactions as this parasite lives proximal to the gut itself[312]. It is even possible that the "uninfected" microbiome reflects the status of individuals who have acquired immunity to reinfection. In this observational study, it is not possible to determine whether *S. haematobium* infection is antecedent to changes in gut microbiota. We envision that a longitudinal study of the microbiome of children in

schistosomiasis-endemic areas that profiles of the same individuals before, during, and after clearance of *S. haematobium* infection via praziquantel treatment, could help to elucidate cause and effect in the system. In addition, our study was relatively small and subjects were recruited from a single site. It would be prudent to replicate our results in a larger, multi-site study to determine whether our findings are applicable to a wider community. Similarly, profiling the microbiota in a younger cohort may also be sensible, as the potential for gut inflammation could contribute to malnutrition and growth inhibition observed in infected children[313].

Conclusions

In general, we have found that the adolescent gut microbiome may be shifted towards a dysbiotic state by infection with *S. haematobium*, with some similarities to prior observations of the gut microbiota in inflammatory contexts. Such a broad dysbiosis would be an interesting observation in urogenital schistosomiasis, building on the increased abundance of *Prevotella* OTUs associated with infection previously observed by Kay *et al*[245]. Given the endemic nature of infection in tropical and subtropical regions, it is important to assess how potential dysbiosis may contribute to disease morbidity. In particular, infected adolescents should be assessed for the presence of intestinal inflammation to determine whether these observed microbiome changes truly reflect an inflammatory state or are associated with any of the known morbidities of urogenital schistosomiasis. Additionally, even the schistosomiasis-negative individuals in the current study are likely to have had been infected in the past, due to the endemic nature of the infection, but their microbiomes were significantly different than those of currently-infected individuals. Therefore, it would be useful to track subjects long-term after curative

praziquantel administration to see whether and how quickly their microbiota returns to an uninfected state; this would additionally help to clarify the causality of the observed changes.

Methods

Ethics Statement

This study was approved by the Kebbi State Ministry of Health and permission to visit the Gotomo primary school in the Argungu Local Government Area was obtained from the local government education department. All research was undertaken in accordance with the the relevant guidelines and regulations of the Kebbi State Ministry of Health. The study and its risks were explained to the students, who then verbally assented to participation if they were interested. To reduce the risk of co-infection with gastrointestinal helminths, any potential subjects who reported recent gastrointestinal distress were excluded. Prior to sample collection, the parents/guardians of students who had assented to participate were informed of the study as well as the associated risks. Parents of assented students gave informed approval for their child's participation by signing the study consent form; in the case of illiteracy, thumbprinting was used, as approved by the State Ministry of Health Ethics Committee. All identified cases of schistosomiasis were reported to the Department of Neglected Tropical Diseases at the Kebbi State Ministry of Health. All children at the school were treated with 40 mg/kg praziquantel by the state government as part of the routine national schistosomiasis control program within two weeks after our study.

Subject Characteristics

All samples were collected from Gotomo Primary School, which draws students from seven local villages in rural Nigeria. Six of the villages are within 1.5 km of the school, while the seventh is within 4 km; as is common in this rural area, most students walk to the school. Compared to other Nigerian states, this area is less developed and less influenced by Western culture and diet; the communities surrounding the schools are primarily small-scale farmers of low socioeconomic status. As a result, the residents of this area share similar lifestyle and dietary characteristics, which reduces potential confounders compared with more developed regions of the country.

90 subjects were screened for *S. haematobium*, with 40 (44.4%) identified as infected. 50 adolescents aged 11-15 years were included in the study (Supplementary Table 1). All samples were collected between July and August 2017. Questionnaires were administered to all participants, covering questions on demographics including age, biological sex, maternal occupation, drinking water source, and exposure to river water (Supplementary Table 2). All biological samples collected were immediately transported to the Federal University Birnin Kebbi Microbiology Laboratory for analysis.

Sample Processing

Urine samples were collected between 10 AM and 2 PM in labelled sample containers and placed in black polyethylene bags. The sedimentation technique was applied for examination of *S. haematobium* eggs in the urine. A minimum of 7 mL of urine was collected per subject. Urine was spun down at 1,000x*g* for 5 minutes, the supernatant was decanted, and the sediments were examined by an experienced technician under the 40X objective of a brightfield microscope (Olympus, USA) to identify *S. haematobium* eggs, which are characterized by a terminal spine. The number of eggs in each sample was

divided by the provided volume and multiplied by 10 to obtain normalized counts of eggs/10 mL (Supplementary Table 1). The presence of eggs in the urine was used to identify cases of adolescents with urogenital schistosomiasis; for subjects with no eggs in the urine, a second sample was obtained and assessed the following day by a different technician to confirm the lack of eggs and reduce the risk of false negatives. The first 25 samples collected in each group were used in the study; cases and controls were not otherwise matched.

Stool samples were also collected from each child that provided a urine sample; at the school, stool was delivered on sterile paper, collected with sterile plastic spatulas, and stored in sterile bottles. All samples were frozen at -20 degrees Celsius within one hour of production until DNA extraction. Microbial DNA was extracted from stool samples of 25 adolescents infected with urinary schistosomiasis and 25 uninfected controls. 1g of stool was removed from the center of defrosted fecal samples and was processed following the manufacturer's protocol using the ZR Fungal/Bacterial DNA Kit™ ZR D6005 (Zymo Research), which utilizes robust mechanical lysis. All samples were extracted at once using the same kit. Extracted DNA was shipped frozen to Brown University for 16S sequencing.

While the prevalence of confounding co-infections in Kebbi State is low[253], species-specific PCR was nevertheless used to help rule out the presence of gastrointestinal helminths in our subjects. We also used PCR to assess whether *S. haematobium* eggs were present in the stool, which can occasionally occur due to unusual placement of adult worms. Extracted fecal DNA was pooled in equimolar amounts by infection status, and PCR for *S. mansoni*, *S. haematobium*, *Ascaris* species, *Ancyclostoma* species, *T. trichiura*, and *N. americanus* was performed using previously published primers (Supplementary

Figure 6)[309,310]. For positive controls, parasite genomic DNA was used; *S. mansoni* and *S. haematobium* DNA was obtained from BEI Resources, while *A. lumbricoides*, *A. duodenale*, *T. trichiura*, and *N. americanus* DNA was graciously gifted to us by the Williams Laboratory at Smith College. Primer sequences and PCR conditions are listed in Supplementary Table 5.

16S rRNA Amplicon Sequencing

Extracted genomic DNA was quantified using a Qubit2 (Invitrogen) to ensure sufficient quantity for amplification. Amplification was performed in triplicate according to the Earth Microbiome Project protocols[314], using a library of barcoded adaptor primers (515F) and a single reverse primer (806R) to amplify the V4 region of the 16S gene (Supplementary Table 6)[315]. 240 ng of each amplicon was pooled together for sequencing.

Sequencing was performed on the Illumina MiSeq platform using the paired-end 2x250 bp protocol. Sample 42 was of poor quality, producing only 14 reads and was therefore removed from further analysis. This left 49 samples for analysis. A total of 1,939,065 reads were obtained across all samples, with a t-test (Prism 7) indicating no significant difference in read depth between infected and uninfected samples (Supplementary Table 7, p=0.112).

Data Processing

Reads were demultiplexed using the idemp utility, allowing for 2 barcode mismatches. Demultiplexed reads were imported into the software package Quantitative Insights Into Microbial Ecology 2 (QIIME2 version 2017-8)[5,316]. Within QIIME2, sequences were quality-filtered and denoised using the Divisive Amplicon Denoising Algorithm 2 (DADA2) pipeline[317]. A total of 1,660 ribosomal sequence variants (RSVs)

were identified across all samples. Taxonomy was assigned using the 99% identity SILVA (release 119) V4 classifier[7]. RSVs are analogous to the Operational Taxonomic Units (OTUs) generated through traditional clustering methods, and we have used this more familiar terminology throughout the paper. The OTU table, rooted phylogenetic tree, representative sequences, and metadata from QIIME2 were then exported for further analysis in R (V3.3.1). Demultiplexed reads, metadata, and code are available from the Brown Digital Repository (doi: https://doi.org/10.7301/Z0K35RVK) and the NCBI Short Read Archive under BioProject ID PRJNA526732.

Diversity Analyses

Alpha diversity metrics were calculated using the phyloseq (V1.19.1) package (Shannon and Simpson Diversity indices and Observed OTUs) and btools (V0.0.1) packages (Faith's Phylogenetic Diversity)[318,319]. Two-tailed Welch's t-tests (Prism 7) were used to determine the significance of differences in alpha diversity between infection groups. Rarefaction curves were generated to ensure that potential differences in OTU counts were not attributable to increased read depth (Fig S2). Beta diversity (Bray-Curtis dissimilarity and weighted and unweighted UniFrac) was analyzed using the VEGAN (V2.4-4) package[320], and PERMANOVA was used to analyze the significance of differences in beta diversity. Principal coordinate analysis (PCoA) was then performed on the beta diversity distance matrices to visualize any relationships between microbiome composition and infection status.

Differential Abundance Analyses

To reduce noise, data was trimmed to include only genera that were present in at least two samples. Differential abundance analysis between infection groups was

conducted using the DESeq2 (V1.14.1) package in R with agglomeration at various taxonomic levels (Data S1)[321]. To account for multiple hypothesis testing, a Benjamini-Hochberg correction was applied to obtain the false discovery rate (FDR), and taxa with FDR values below 0.05 were considered significant. Abundances by sample at all taxonomic levels is provided in Data S1.

Inferred Metagenomics

To predict the functional potential of the positive and negative communities, we used the web-based tool Piphillin, which infers metagenomes from 16S content[12] and matches them to orthologs and pathways from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. The OTU abundance table and representative sequences file were exported from QIIME2 and uploaded to Piphillin with required formatting, and the analysis was run using the following parameters: Database – **KEGG**; Database Version – **KEGG May 2017**; % Identity Cutoff – **99**. The Features output from Piphillin was formatted and uploaded to the web-based tool MicrobiomeAnalyst to identify differentially abundant KEGG orthologs and pathways[291]. The following MicrobiomeAnalyst parameters were used to analyze differential abundance in orthologs: Data filtering – **None**; Data Normalization – **Relative Log Expression**; Analysis Overview – **RNASeq Methods (DESeq2).** To assess pathway enrichment, the **Network Mapping** function on the DESeq2 output was used. In both cases, a FDR less than 0.05 was considered significant.

Confirmatory qPCR of Specific Genus-Level Changes

Genus-specific primers were used to confirm that the changes seen in differential abundance analysis were reflected in the original templates. Many of the significant genera were taxonomically classified as "uncultured" or "*incertae sedis*" members of a higher

taxonomic level, and therefore these were excluded. To design genus-specific primers, at least one 16S sequence from each of several major species in each significant genus was downloaded from the National Center for Biotechnology Information, as well as 16S sequences of representative species from other genera in the same family. When a large number of other genera were present, those also present in the samples were prioritized. For each genus, the relevant species were aligned using Muscle in UGENE (V1.28)[322]. Alignments were visually scanned for regions where the species within the relevant genus were very similar but were different from species in related genera.

After selecting a few potential regions, primer pairs were tested in NCBI's Primer-BLAST, which combines Primer3 and BLAST[323-325]. Briefly, the 16S sequence of a representative species from the relevant genus was used as the template, and the potential primer pairs were input with the following parameters: Search Mode – **Automatic**; Database – **Refseq Representative Genomes**; Organism – **Bacteria (taxid:2)**; Primer must have at least **3** total mismatches to unintended targets; At least **2** mismatches within the last **5** bps; Ignore targets that have **6** or more mismatches to the primer. Only primer pairs with unintended targets that matched with other species in the genera but not with species in related genera present in the sample were accepted. Primer pairs were validated by robust amplification from schistosomiasis-positive gDNA samples compared to a mock community that did not contain the genera of interest (ZymoBIOMICS Microbial Community DNA Standard). The exception is *Desulfovibrio*, which despite apparent strong specificity in PrimerBLAST did show some amplification from the mock community. Primers are listed in Supplementary Table 7.

Equivalent amounts of genomic DNA from all schistosomiasis-positive and schistosomiasis-negative samples were pooled into two samples for analysis. All qPCR was run on a Roche Lightcycler 480, using the SYBRGreen-based 2X Fast Start Essential DNA Green Master Mix in the following preparation: 7.5 µL Master Mix, 6.35 µL H2O, 0.075 µL each primer, and 1 ng of template gDNA. Each qPCR run was performed in triplicate technical replicates on pooled positive gDNA, pooled negative gDNA, a negative control mock community, and a no template control. Reactions were performed in parallel. Cycling conditions are listed in Supplementary Table 8. Changes were calculated using the ΔΔCT method, using total 16S DNA amplified from the pooled samples with universal primers to normalize data.

Declarations

Ethics

This study was approved by the Kebbi State Ministry of Health and permission to visit the Gotomo primary school in the Argungu Local Government Area was obtained from the local government education department. All research was undertaken in accordance with the the relevant guidelines and regulations of the Kebbi State Ministry of Health.

Data Availability

Underlying sequencing data can be found at the NCBI Short Read Archive under BioProject ID PRJNA526732. Taxonomic composition data can be found in Supplementary Data 1.

Acknowledgements

We wish to acknowledge the Gotomo primary school headmaster, teachers, and parents for their support during the study. We also thank the Director of Public Health of the Kebbi

State Ministry of Health for his support and guidance during the project. We acknowledge the University of Rhode Island Genomics and Sequencing Center, which performed the quality control and 16S sequencing, and the Brown University Statistical Consulting Group and the Brown COBRE Computational Biology Core for their advice on statistics. We also thank the Williams Laboratory at Smith College for their generous gift of genomic DNA from several helminth species.

Funding

This work was supported by the National Institute of Allergy and Infectious Diseases (grant number AI097493 awarded to SO), the Bill and Melinda Gates Foundation (grant number OPP1034619 awarded to SO), and the National Institute of General Medical Science (IDeA grant PM20GM109035, Center for Computational Biology of Human Disease, awarded to PB). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication. Opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the funders. The authors declare no competing interests.

Author Contributions

OA & PB designed and oversaw the study. SO advised during the study. OA, COO, MBM, and AAE contributed to participant recruitment and sample collection and processing. ADR performed data analysis and generated figures. DJC advised during data analysis. ADR & OA wrote the manuscript. All authors reviewed and revised the manuscript.

Figure 1: Measures of Alpha Diversity in Schistosomiasis-positive and -negative Individuals
(A) Observed OTUs: p=0.12. (B) Shannon's Index of diversity: p=0.77. (C) Simpson's Index of diversity: p=0.69. (D) Faith's Phylogenetic Diversity: p=0.49. Statistics: two-tailed t-test with Welch's correction, error bars indicate SEM.

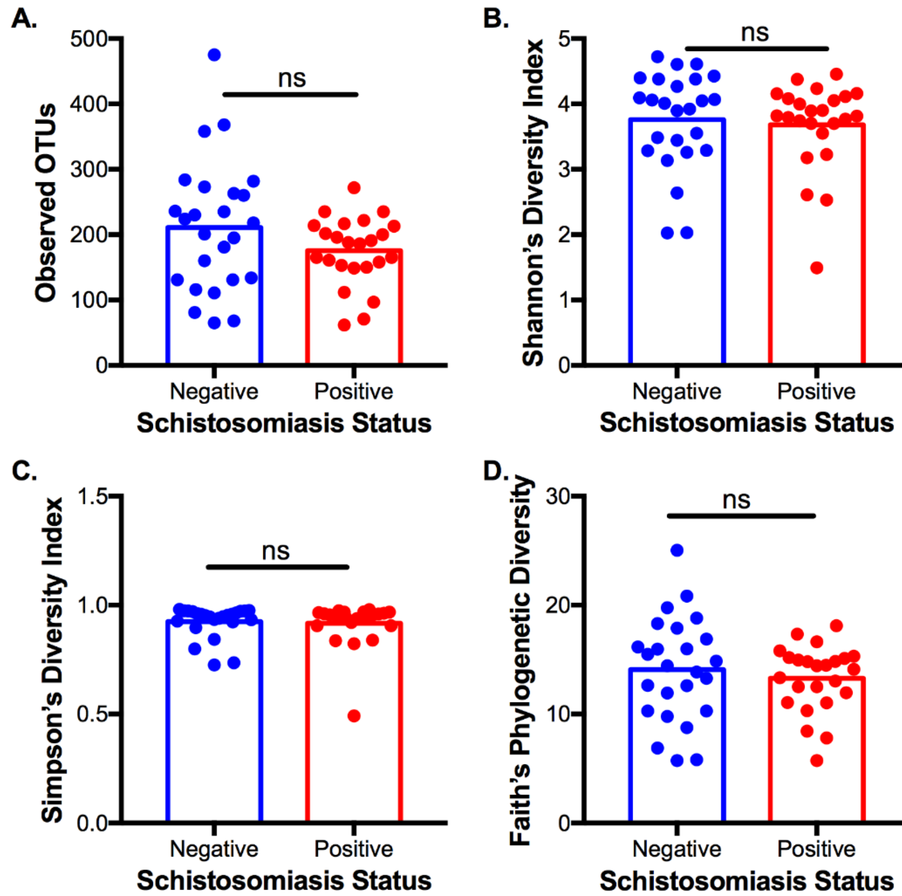


Figure 2: Principal Coordinate Analysis of community similarity by schistosomiasis infection status
Distance matrices were calculated using (A) Bray-Curtis Dissimilarity: p=0.005. (B) Unweighted UniFrac: p=0.003. (C) Weighted UniFrac: p=0.012. Statistics: PERMANOVA through vegan package in R, * p<0.05, ** p<0.01

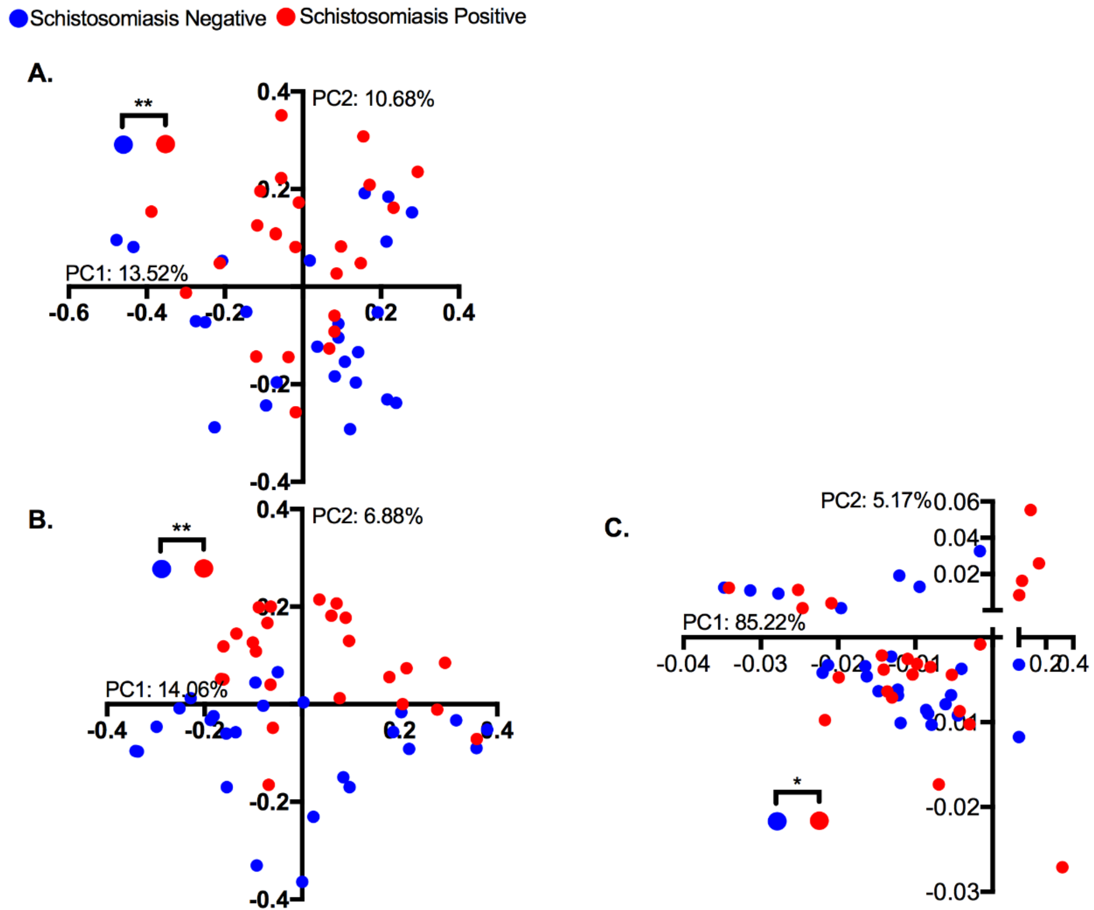


Figure 3: Taxa and associated lineages that significantly changed in schistosomiasis-positive subjects Taxa that increased significantly in infected subjects are shown in orange cells, while taxa that decreased significantly are shown in blue cells. Taxa that did not change or with changes that did not reach significance are shown in white cells. Fold change values (shown in **bold**) were calculated from the log_2(fold change) value output from DESeq2, and FDR values (shown in *italics*) were obtained from the same DESeq2 output.

Phylum	Class	Order	Family	Genus
Actinobacteria	Actinobacteria **+3.35**, *0.018*	Bifidobacteriales	Bifidobacteriaceae **+15.54**, *0.047*	
	Coriobacteriia	Coriobacteriales	Coriobacteriaceae	*Olsenella* **+10.33**, *0.00078*
				Uncultured bacterium **+3.53**, *0.028*
Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae **+2.16**, *0.030*	*Alloprevotella* **+1.50**, *0.01*
				Prevotella **+2.47**, *0.013*
			Porphyromonadaceae **-7.04**, *0.018*	*Parabacteroides* **-7.39**, *0.028*
Cyanobacteria **-3.72**, *0.023*	Melainabacteria **-4.41**, *0.04*	Gastranaerophilales **-5.48**, *0.037*	Uncultured rumen bacterium **-19.58**, *0.029*	
Firmicutes **-1.58**, *0.00094*	Clostridia **-1.73**, *0.0065*	Clostridiales **-1.73**, *0.0070*	Peptostreptococcaceae **-2.33**, *0.0017*	*Incertae Sedis* **2.47**, *0.0028*
			Peptococcaceae **+10.70**, *0.0013*	*Peptococcus* **+31.81**, *5.0E-6*
			Clostridiaceae 1	*Clostridium ss* 6 **-22.08**, *0.01*
			Ruminococcaceae	*Incertae Sedis* **-3.24**, *0.028*
				Subdoligranulum **-1.59**, *0.037*
	Erysipelotrichia	Erysipelotrichales	Erysipelotrichaceae	Uncultured bacterium **-3.85**, *0.0062*
	Negativicutes **+4.55**, *0.00049*	Selenomonadales **+5.60**, *0.00013*	Veillonellaceae **+6.99**, *0.00029*	*Megasphaera* **+217.1**, *2.34E-8*
				Dialister **+5.86**, *0.0062*
Fusobacteria	Fusobacteriia **+8.07**, *0.034*	Fusobacteriales **+22.31**, *0.017*		
Proteobacteria **+2.58**, *0.012*	Betaproteobacteria **-5.75**, *0.04*			
	Deltaproteobacteria **+8.78**, *0.011*	Desulfovibrionales **+10.69**, *0.017*	Desulfovibrionaceae **+18.58**, *0.0010*	*Desulfovibrio* **+30.7**, *0.00032*
	Gammaproteobacteria	Pasteurellales **+4.93**, *0.037*	Pasteurellaceae **+4.79**, *0.035*	*Haemophilus* **+5.35**, *0.025*
		Pseudomonadales **+15.00**, *0.017*	Moraxellaceae **+12.94**, *0.018*	*Acinetobacter* **+10.69**, *0.028*
Tenericutes **-3.84**, *0.0055*	Mollicutes **-4.15**, *0.0083*	RF9 **-4.41**, *0.017*	Uncultured bacterium **-3.98**, *0.020*	Uncultured bacterium **-4.02**, *0.019*

Figure 4: Differences in Relative Abundances of Genera Between Schistosomiasis-positive and -negative Subjects
(A) Average relative abundances of all genera, with genera showing significant differences between positive and negative samples highlighted in color. (B-R) Genera that changed in infected adolescents, with negative and positive samples interleaved by ranked abundance of each taxon and dotted lines representing the average relative abundance by group. Statistics: Wald test of differential abundance through DESeq2 package in R, * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001, error bars indicate SEM. Exact corrected p-values (FDR) can be found in Figure 3.

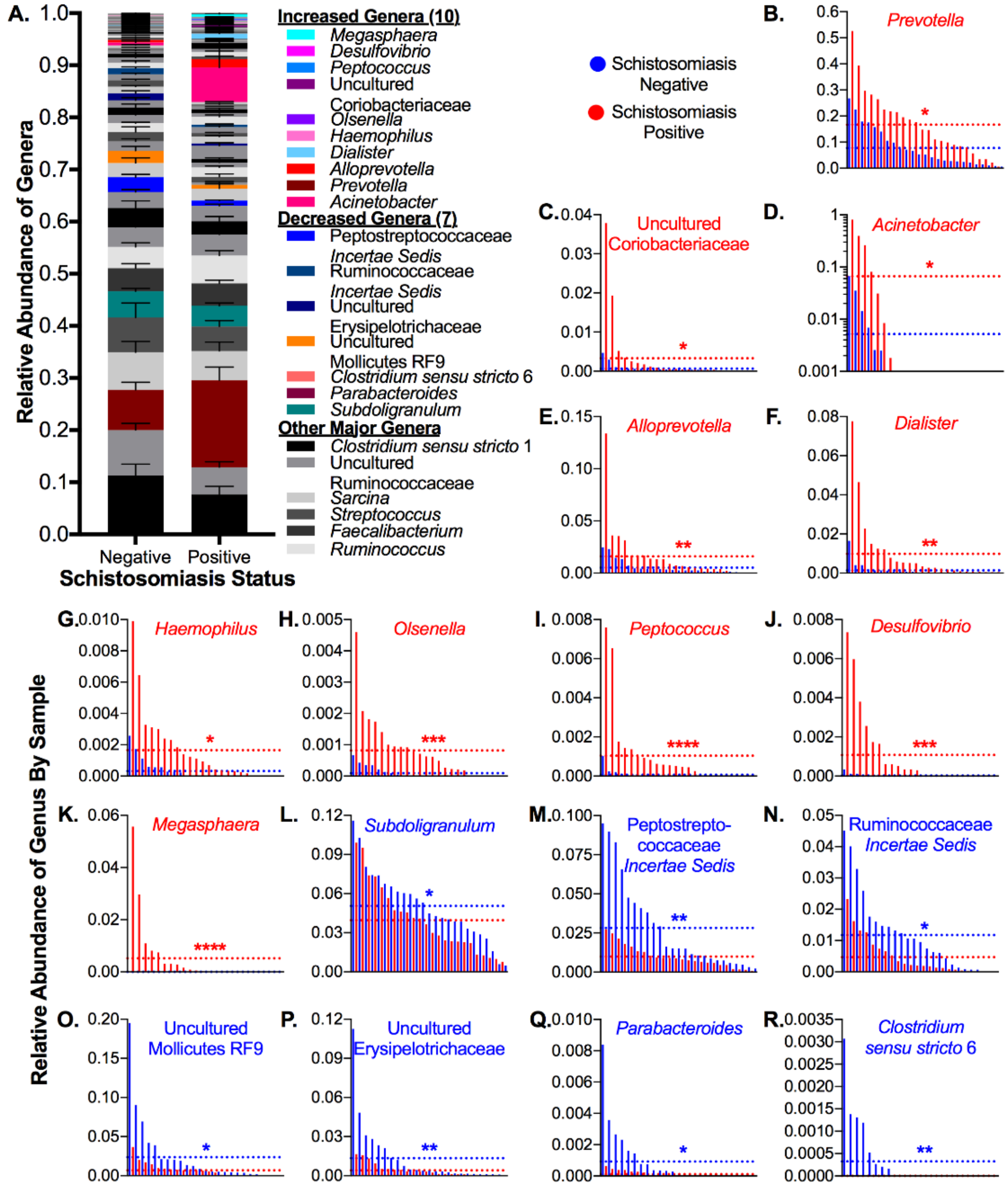


Figure 5: -$\Delta\Delta C_T$ Values of Significant Genera Obtained from qPCR Using Genus-specific Primers
$\Delta\Delta C_T$, used to allow all genera to be shown on the same scale, is the corrected raw difference in C_T values between infected and uninfected samples, and the sign change causes positive -$\Delta\Delta C_T$ values to indicate a positive fold change. Error bars indicate SEM of technical replicates.

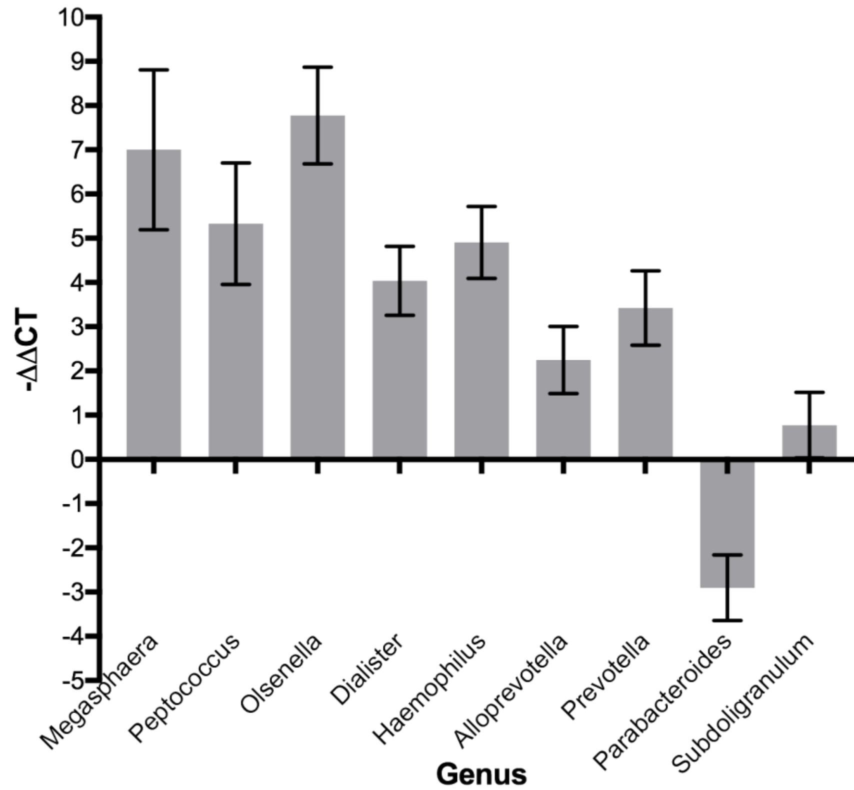
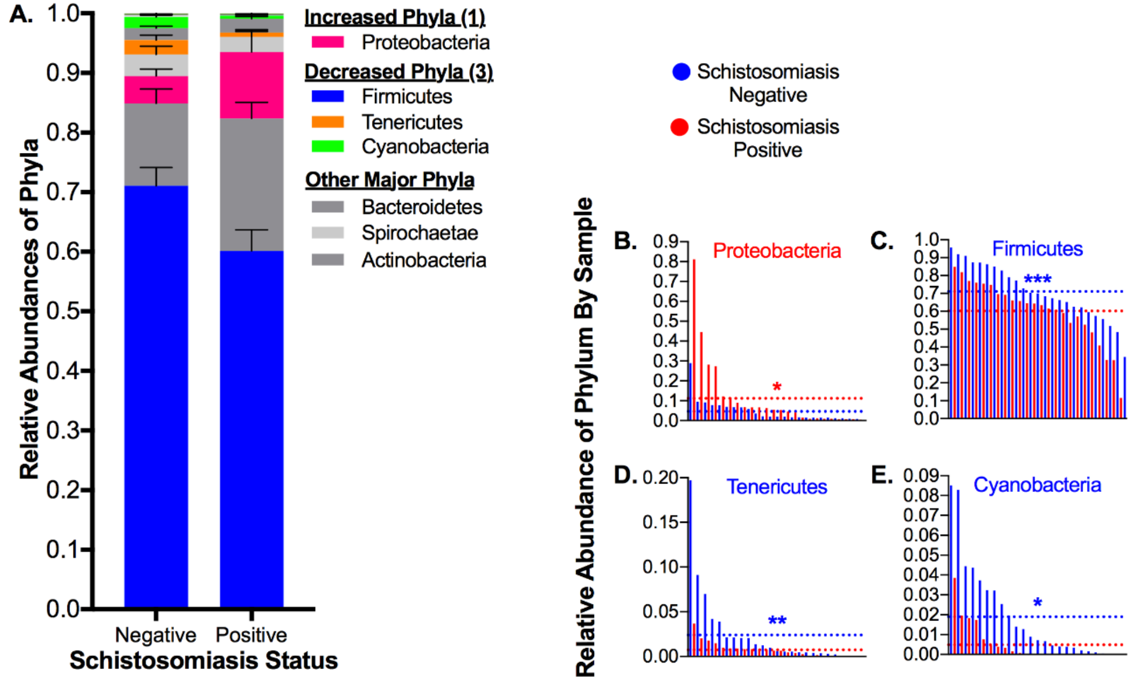
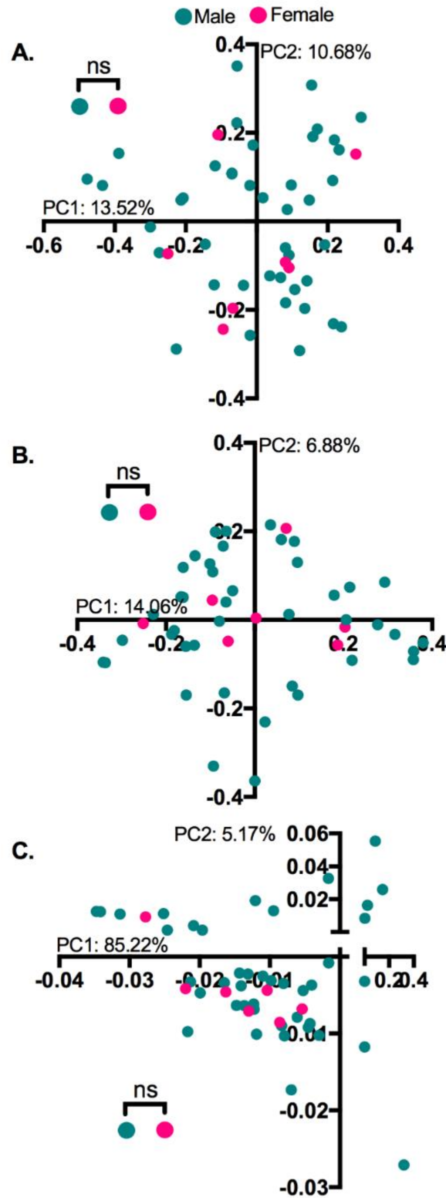


Figure 6: Differences in Relative Abundances of Phyla between Schistosomiasis-positive and -negative Subjects

(A) Average relative abundances of all phyla, with phyla showing significant differences between positive and negative samples highlighted in color. (B-E) Phyla that changed in infected adolescents, with negative and positive samples interleaved by ranked abundance of each taxon and dotted lines representing the average relative abundance by group. Statistics: Wald test of differential abundance through DESeq2 package in R, * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001, error bars indicate SEM. Exact corrected p-values (FDR) can be found in Figure 3.

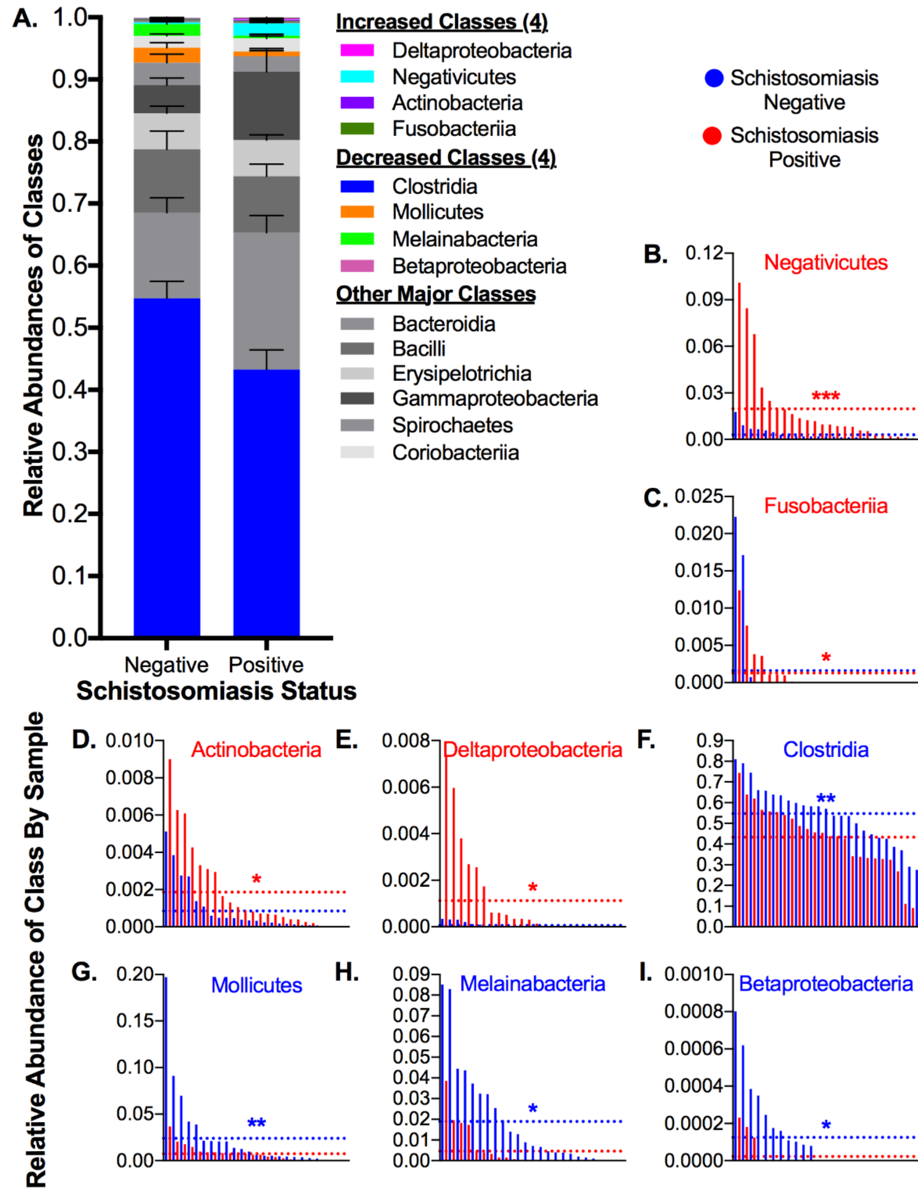


Supplementary Figure 1: Principal Coordinate Analysis of Community Similarity by Gender
Distance matrices were calculated using (A) Bray-Curtis (p=0.853), (B) unweighted UniFrac (p=0.589), and (C) weighted UniFrac (p=0.53). Statistics: PERMANOVA through vegan package in R.



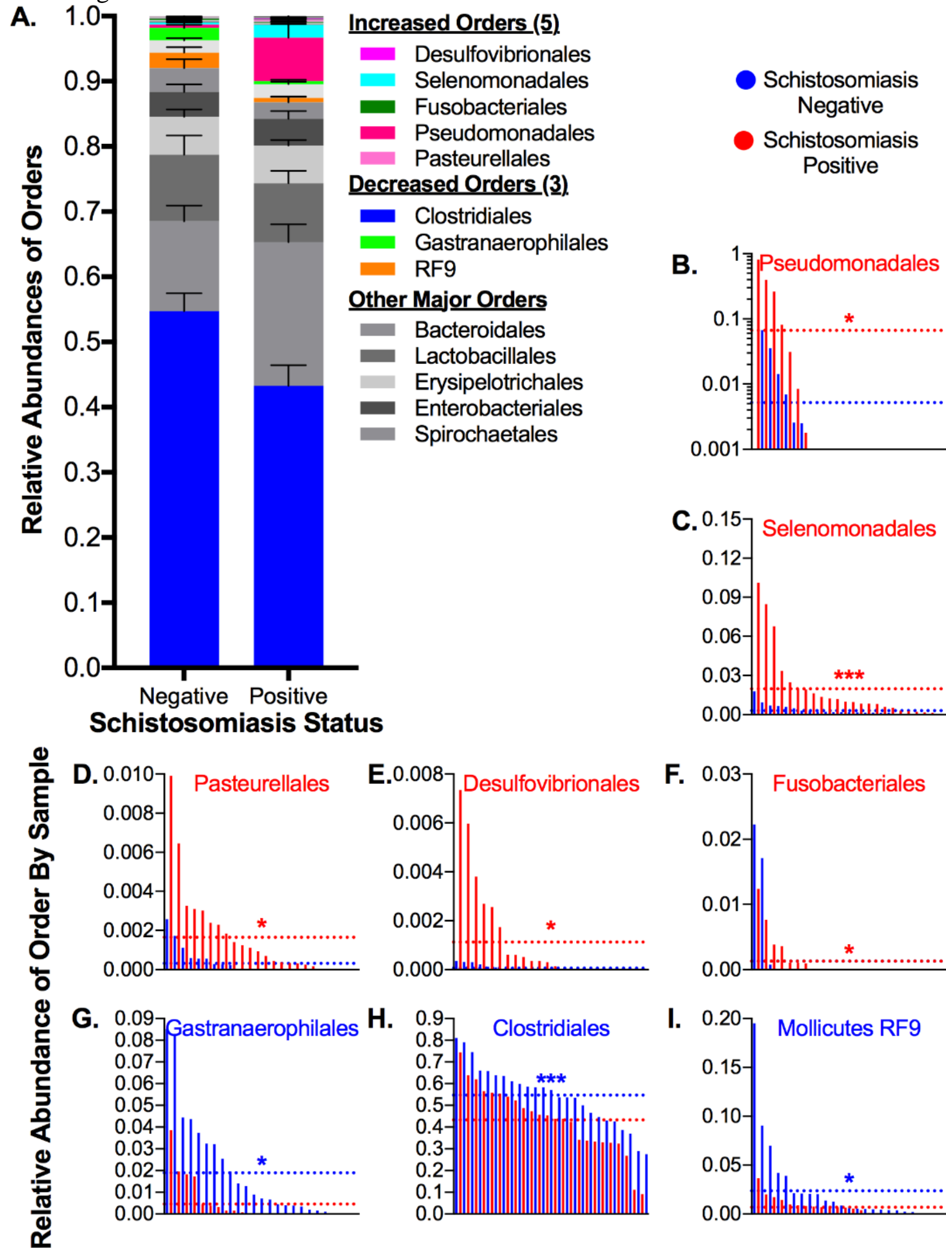
Supplementary Figure 2: Differences in Relative Abundance of Classes between Schistosomiasis-positive and -negative Subjects

(A) Average relative abundances of all classes, with classes showing significant differences between positive and negative samples highlighted in color. (B-I) Classes that changed in infected adolescents, with negative and positive samples interleaved by ranked abundance of each taxon and dotted lines representing the average relative abundance by group. Statistics: Wald test of differential abundance through DESeq2 package in R, * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001, error bars indicate SEM. Exact corrected p-values (FDR) can be found in Figure 3.



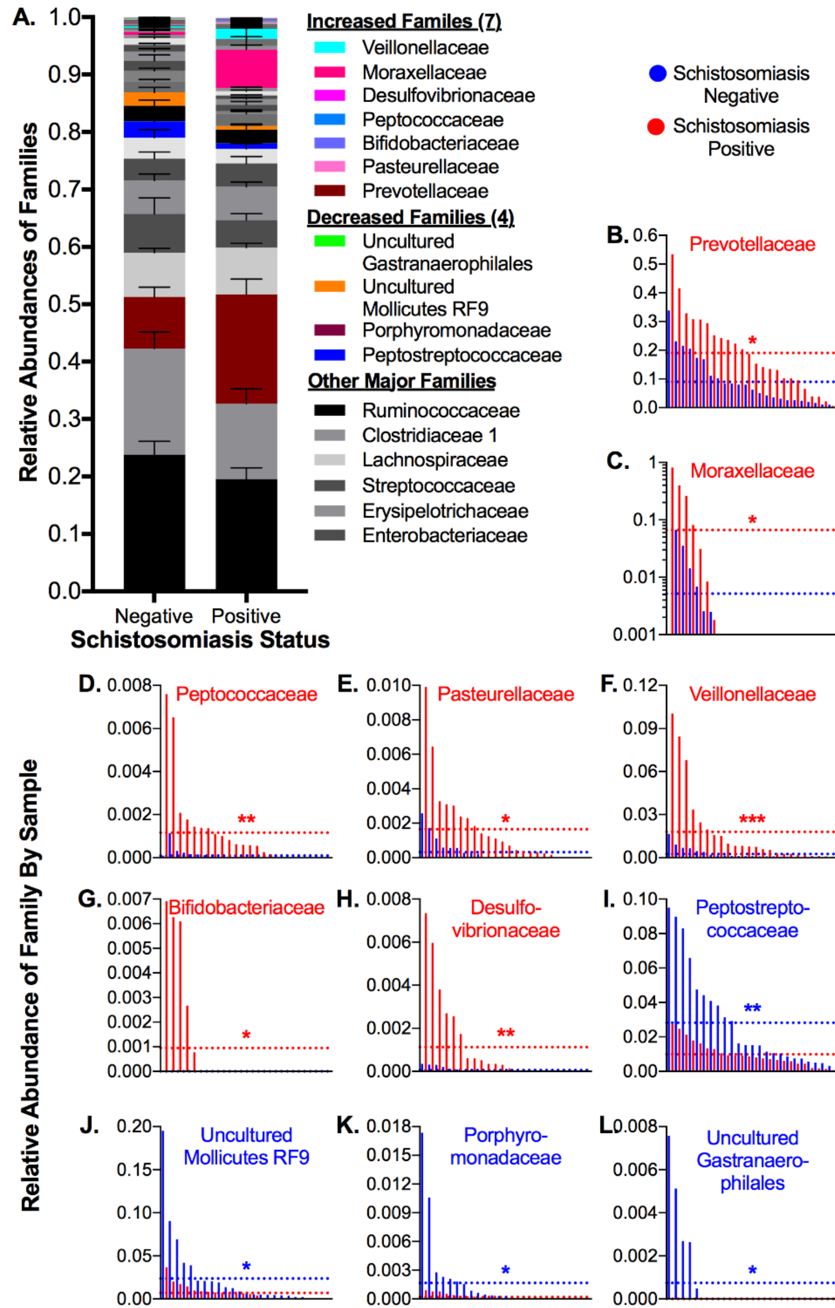
Supplementary Figure 3: Differences in Relative Abundance of Orders between Schistosomiasis-positive and -negative Subjects

(A) Average relative abundances of all orders, with orders showing significant differences between positive and negative samples highlighted in color. (B-I) Orders that changed in infected adolescents, with negative and positive samples interleaved by ranked abundance of each taxon and dotted lines representing the average relative abundance by group. Statistics: Wald test of differential abundance through DESeq2 package in R, * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001, error bars indicate SEM. Exact corrected p-values (FDR) can be found in Figure 3.



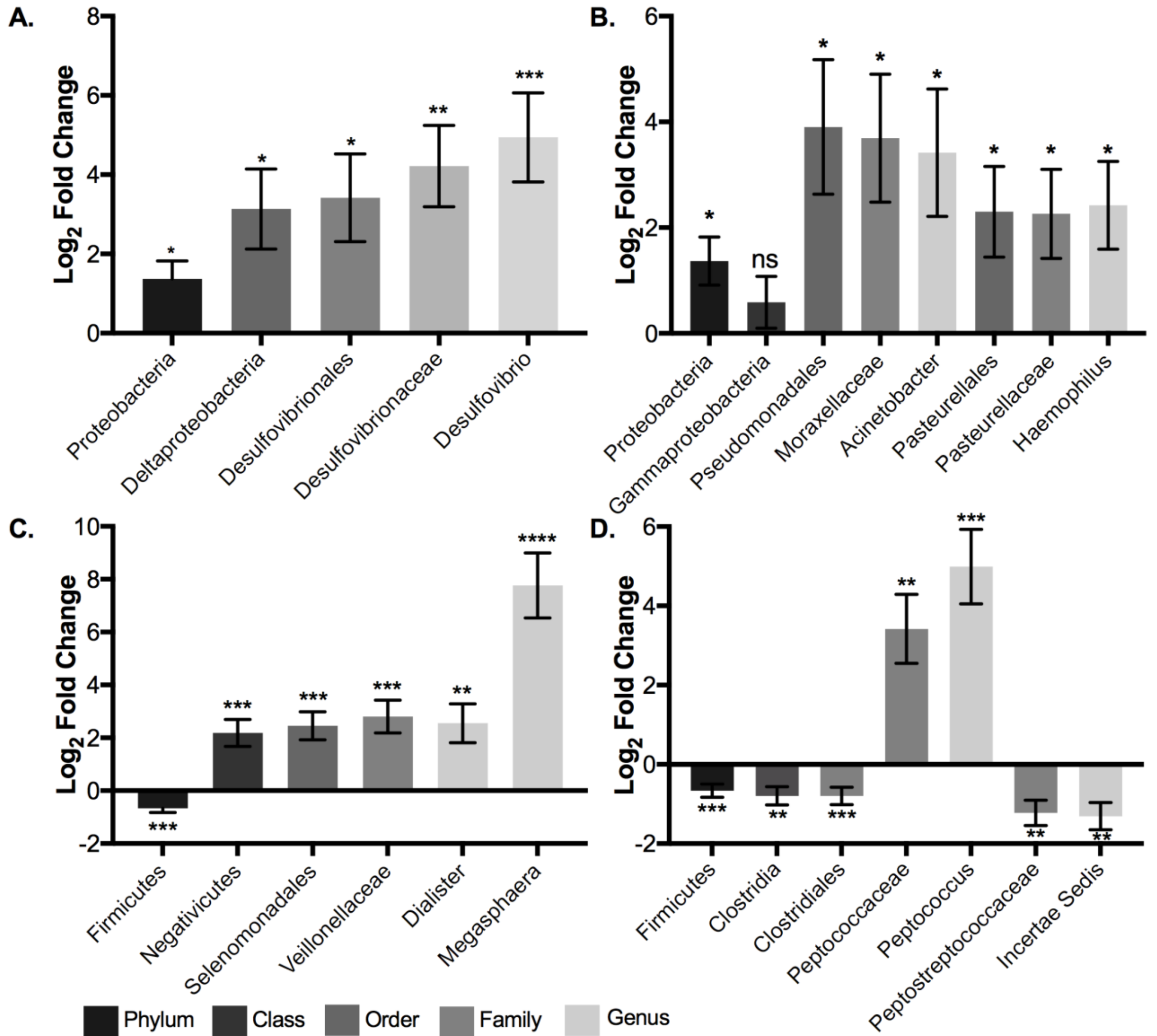
Supplementary Figure 4: Differences in Relative Abundance of Families between Schistosomiasis-positive and -negative Subjects.

(A) Average relative abundances of all families, with families showing significant differences between positive and negative samples highlighted in color. (B-L) Families that changed in infected adolescents, with negative and positive samples interleaved by ranked abundance of each taxon and dotted lines representing the average by group. Statistics: Wald test of differential abundance through DESeq2 package in R, * $p<0.05$, ** $p<0.01$ error bars indicate SEM. Exact corrected p-values (FDR) can be found in Figure 3.



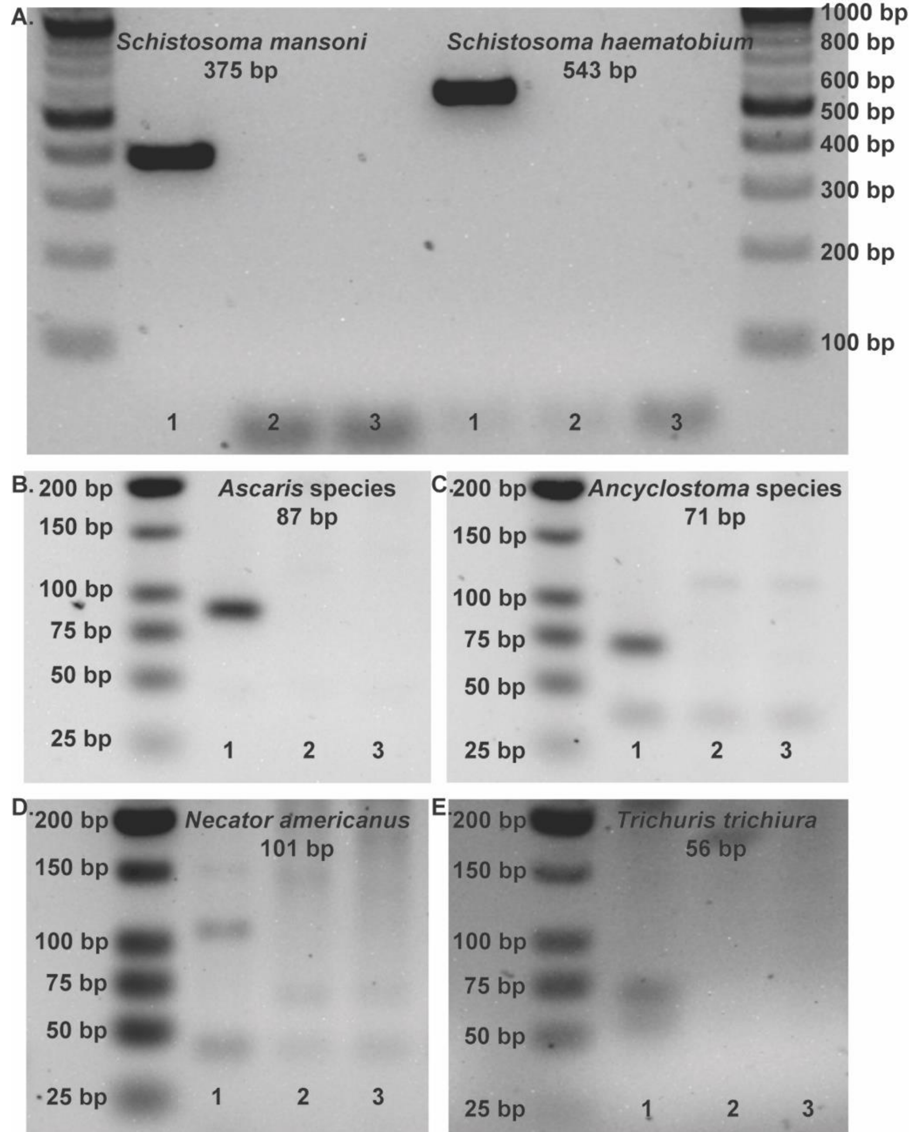
Supplementary Figure 5: Fold Changes of Differentially-Abundant Lineages Between Schistosomiasis-positive and -negative Groups

(A) *Desulfovibrio* lineage (B) *Acinetobacter* and *Haemophilus* lineage (C) *Megasphaera* and *Dialister* lineage (D) *Peptococcus* and *Peptostreptococcaceae incertae sedis* lineage. Statistics: Wald test of differential abundance through DESeq2 package in R, * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$, error bars indicate standard error of $\log_2$(Fold Change).



Supplementary Figure 6: PCR to Detect Helminth DNA in Fecal Samples

(A) PCR for *S. mansoni* and *S. haematobium*. (B) PCR for *Ascaris* spp. (C) PCR for *Ancyclostoma* spp. (D) PCR for *Necator americanus*. (E) PCR for *Trichuris trichiura*. Lane Labels: (1) Positive Control gDNA, (2) Pooled Schistosomiasis-Negative Fecal DNA, (3) Pooled Schistosomiasis-Positive Fecal DNA. All gels were photographed using the auto-exposure setting of the Gel Doc EZ-Imager (BioRad).



Supplementary Table 1: Demographic Data on Study Participants

This table includes subject number, biological sex, age, schistosomiasis infection status, and urine egg counts per 10 mL for infected participants. Subject 42 (indicated with *) was not included in analysis due to low-quality amplification of 16S rDNA.

Subject	Sex	Age	Status	Eggs	Subject	Sex	Age	Status	Eggs
1	Male	11	Negative	0	26	Male	12	Positive	20
2	Male	11	Negative	0	27	Male	11	Positive	250
3	Male	12	Negative	0	28	Male	11	Positive	20
4	Male	11	Negative	0	29	Male	13	Positive	10
5	Male	11	Negative	0	30	Male	13	Positive	40
6	Male	12	Negative	0	31	Male	12	Positive	380
7	Male	13	Negative	0	32	Male	11	Positive	100
8	Male	12	Negative	0	33	Male	11	Positive	10
9	Male	11	Negative	0	34	Male	15	Positive	20
10	Male	11	Negative	0	35	Male	11	Positive	80
11	Male	12	Negative	0	36	Male	11	Positive	110
12	Male	12	Negative	0	37	Male	12	Positive	20
13	Male	12	Negative	0	38	Male	13	Positive	90
14	Male	12	Negative	0	39	Male	15	Positive	30
15	Male	15	Negative	0	40	Male	11	Positive	10
16	Male	11	Negative	0	41	Male	15	Positive	40
17	Male	11	Negative	0	42*	Male	11	Positive	100
18	Male	15	Negative	0	43	Male	15	Positive	20
19	Male	15	Negative	0	44	Male	13	Positive	20
20	Male	13	Negative	0	45	Male	13	Positive	30
21	Female	12	Negative	0	46	Male	15	Positive	50
22	Female	12	Negative	0	47	Male	12	Positive	50
23	Female	11	Negative	0	48	Female	12	Positive	50
24	Female	11	Negative	0	49	Female	11	Positive	30
25	Female	11	Negative	0	50	Female	11	Positive	30

Supplementary Table 2: Lifestyle Data on Study Participants

This table includes the source of drinking water, weekly exposure to river water, and maternal occupation of all subjects. Subjects did not significantly differ on any metric by infection group, as analyzed by chi-square analysis (p = 0.732, p = 0.340, p = 0.958). To meet the requirements for chi-square analysis, the "tap" and "river" responses were combined and the "housewife" and "none" responses were combined in the contingency tables for "Drinking Water Source" and "Maternal Occupation", respectively.

Schistosomiasis-Negative				Schistosomiasis-Positive			
Subject	Drinking Water Source	Weekly River Contact	Maternal Occupation	Subject	Drinking Water Source	Weekly River Contact	Maternal Occupation
1	Tap	Once	Farmer	26	Well	Once	None
2	Tap	Never	Farmer	27	Well	Once	Office worker
3	Well	Once	Farmer	28	Tap	Twice	None
4	Well	Once	Farmer	29	Well	Twice	Farmer
5	Well	Daily	Farmer	30	Well	Never	None
6	Well	Once	Housewife	31	Well	Once	Farmer
7	Well	Once	None	32	Tap	Twice	None
8	Well	Twice	None	33	Well	Twice	Farmer
9	Well	Daily	Office worker	34	Tap	Daily	None
10	Well	Daily	None	35	Well	Once	Farmer
11	Well	Once	Farmer	36	Well	Once	Farmer
12	Well	Once	Farmer	37	Well	Once	None
13	Well	Never	Office worker	38	Well	Never	None
14	Well	Daily	None	39	Well	Once	Farmer
15	Well	Twice	None	40	Tap	Never	Farmer
16	Tap	Daily	None	41	Well	Never	Office worker
17	Well	Once	Farmer	42	Well	Once	None
18	Well	Once	Farmer	43	River	Daily	Farmer
19	Tap	Once	None	44	Well	Twice	Farmer
20	Well	Twice	Farmer	45	Well	Twice	None
21	Well	Daily	Farmer	46	Well	Twice	None
22	Tap	Never	None	47	Well	Daily	None
23	Tap	Never	None	48	Well	Once	Farmer
24	Well	Never	None	49	Well	Once	None
25	Well	Daily	None	50	Well	Once	None
Summary (Negative)	19 Well 6 Tap 0 River	5 Never 10 Once 3 Twice 7 Daily	11 Farmer 1 Housewife 2 Office Worker 11 None	Summary (Positive)	20 Well 4 Tap 1 River	4 Never 11 Once 7 Twice 3 Daily	10 Farmer 0 Housewife 2 Office Worker 13 None

Supplementary Table 3: Differentially-abundant KEGG Pathways

This table shows the two KEGG pathways that were significantly enriched in schistosomiasis-positive predicted metagenomes, as well as significantly-enriched orthologs within those pathways.

KEGG Pathway	Included Orthologs	FDR
Atrazine degradation	ureC (urease subunit alpha) ureB (urease subunit beta) ureA (urease subunit gamma)	0.000337
Arginine and proline metabolism	prdB (D-proline reductase) prdF (proline racemase)	0.0156

Supplementary Table 4: Differentially-abundant KEGG Orthologs

This table shows KEGG orthologs that were differentially abundant in the schistosomiasis-positive and -negative groups. Orthologs in bold are members of significantly enriched pathways. Fold changes reflect the abundance in schistosomiasis-positive relative to -negative subjects.

KEGG Ortholog	Name	Fold Change	FDR
K07006	uncharacterized protein	6.825808785	9.31E-05
K05346	deoR; deoxyribonucleoside regulator	6.93453911	9.31E-05
K03929	pnbA; para-nitrobenzyl esterase	5.402398268	0.0012918
K07454	putative restriction endonuclease	5.698570258	0.0021704
K01501	nitrilase	5.716769016	0.0021704
K00019	bdh; 3-hydroxybutyrate dehydrogenase	4.781953164	0.0023841
K08365	merR; MerR family transcriptional regulator, mercuric resistance operon regulatory protein	4.478141143	0.0051743
K07276	uncharacterized protein	5.018652309	0.0051743
K10794	**prdB; D-proline reductase (dithiol)**	**4.987787494**	**0.0051743**
K10811	thiamine pyridinylase	4.987787494	0.0051743
K20626	lcdA; lactoyl-CoA dehydratase subunit alpha	4.987787494	0.0051743
K20627	lcdB; lactoyl-CoA dehydratase subunit beta	4.987787494	0.0051743
K01777	**prdF; proline racemase**	**4.927650723**	**0.0059194**
K18923	stbD; antitoxin StbD	4.322601978	0.010191
K13928	mdcR; LysR family transcriptional regulator, malonate utilization transcriptional regulator	4.250698698	0.011485
K07267	oprB; porin	4.48124623	0.011485
K00529	hcaD; 3-phenylpropionate/trans-cinnamate dioxygenase ferredoxin reductase component	2.901904416	0.012016
K00480	salicylate hydroxylase	2.91925452	0.012272
K05819	mhpT; MFS transporter, AAHS family, 3-hydroxyphenylpropionic acid transporter	2.870893873	0.01367
K01428	**ureC; urease subunit alpha**	**3.343594748**	**0.01367**
K12542	lapC; membrane fusion protein, adhesin transport system	2.841987088	0.01367
K05710	hcaC; 3-phenylpropionate/trans-cinnamate dioxygenase ferredoxin component	2.841002301	0.01367
K01430	**ureA; urease subunit gamma**	**3.428315777**	**0.01367**
K02077	ABC.ZM.S; zinc/manganese transport system substrate-binding protein	2.928780351	0.013924
K03188	ureF; urease accessory protein	3.311993864	0.013924
K03189	ureG; urease accessory protein	3.311993864	0.013924
K03190	ureD, ureH; urease accessory protein	3.311075712	0.013924
K13818	mobAB; molybdopterin-guanine dinucleotide biosynthesis protein	4.304662112	0.013924
K03187	ureE; urease accessory protein	3.305113922	0.013924
K01429	**ureB; urease subunit beta**	**3.333181782**	**0.026514**
K01692	paaF, echA; enoyl-CoA hydratase	3.966317037	0.030301
K11103	dctA; aerobic C4-dicarboxylate transport protein	2.571303666	0.030301
K08728	nucleoside deoxyribosyltransferase	3.963019318	0.030301
K07783	uhpC; MFS transporter, OPA family, sugar phosphate sensor protein	0.413769727	0.030301
K00004	BDH, butB; (R,R)-butanediol dehydrogenase / meso-butanediol dehydrogenase / diacetyl reductase	3.043222884	0.04899

Supplementary Table 5: Primers and Conditions Used for PCR Detection of Helminth DNA in Fecal Samples

Organism/Gene	Primers	Cycle Conditions	Positive Control DNA
Schistosoma mansoni Cox1	FWD: TTTTTTGGTCATCCTGAGGTGTAT REV: TGCAGATAAAGCCACCCCTGTG	98C for 30 seconds 61C for 45 seconds 72C for 1 minute	*Schistosoma mansoni*, adult worm, male and female genomic DNA (mixed) (BEI)
Schistosoma haematobium Cox1	FWD: TTTTTTGGTCATCCTGAGGTGTAT REV: TGATAATCAATGACCCTGCAATAA	98C for 30 seconds 64C for 45 seconds 72C for 1 minute	*Schistosoma haematobium*, adult worm, male and female genomic DNA (mixed) (BEI)
Ascaris spp. ITS1	FWD: GTAATAGCAGTCGGCGGTTTCTT REV: GCCCAACATGCCACCTATTC	98C for 10 seconds 60C for 10 seconds 72C for 15 seconds	*Ascaris lumbricoides* genomic DNA (Williams lab)
Ancyclostoma spp. ITS1	FWD: GAATGACAGCAAACTCGTTGTTG REV: ATACTAGCCACTGCCGAAACGT	98C for 10 seconds 60C for 10 seconds 72C for 15 seconds	*Ancyclostoma duodenale* genomic DNA (Williams lab)
Necator americanus ITS2	FWD: CTGTTTGTCGAACGGTACTTGC REV: ATAACAGCGTGCACATGTTGC	98C for 10 seconds 57C for 10 seconds 72C for 15 seconds	*Necator americanus* genomic DNA (Williams lab)
Trichuris trichiura ITS1	FWD: TCCGAACGGCGGATCA REV: CTCGAGTGTCACGTCGTCCTT	98C for 10 seconds 57C for 10 seconds 72C for 15 seconds	*Trichuris trichuria* genomic DNA (Williams lab)

Supplementary Table 6: Primers Used for Amplicon Generation

This table includes the standard 5' Illumina Adapter, Pad, Linker, and 515F primer segments as well as the variable barcodes used for sample identification and demultiplexing. It also includes the 806R primer, with the associated Pad, Linker, and Adapter segments. Primer design was obtained from the Earth Microbiome Project protocols.

Primer Structure			
Forward Primer	5' Illumina Adapter – Barcode – Pad – Linker – 515FB Primer AATGATACGGCGACCACCGAGATCTACACGCT-BARCODE-TATGGTAATT-GT-GTGYCAGCMGCCGCGGTAA		
Reverse Primer	CAAGCAGAAGACGGCATACGAGAT-AGTCAGCCAG-CC-GGACTACNVGGGTWTCTAAT		
Subject	Barcode Sequence	Subject	Barcode Sequence
AR001	AGCCTTCGTCGC	AR026	CGGGACACCCGA
AR002	TCCATACCGGAA	AR027	CTGTCTATACTA
AR003	AGCCCTGCTACA	AR028	TATGCCAGAGAT
AR004	CCTAACGGTCCA	AR029	CGTTTGGAATGA
AR005	CGCGCCTTAAAC	AR030	AAGAACTCATGA
AR006	TATGGTACCCAG	AR031	TGATATCGTCTT
AR007	TACAATATCTGT	AR032	CGGTGACCTACT
AR008	AATTTAGGTAGG	AR033	AATGCGCGTATA
AR009	GACTCAACCAGT	AR034	CTTGATTCTTGA
AR010	GCCTCTACGTCG	AR035	GAAATCTTGAAG
AR011	ACTACTGAGGAT	AR036	GAGATACAGTTC
AR012	AATTCACCTCCT	AR037	GTGGAGTCTCAT
AR013	CGTATAAATGCG	AR038	ACCTTACACCTT
AR014	ATGCTGCAACAC	AR039	TAATCTCGCCGG
AR015	ACTCGCTCGCTG	AR040	ATCTAGTGGCAA
AR016	TTCCTTAGTAGT	AR041	ACGCTTAACGAC
AR017	CGTCCGTATGAA	AR042	TACGGATTATGG
AR018	ACGTGAGGAACG	AR043	ATACATGCAAGA
AR019	GGTTGCCCTGTA	AR044	CTTAGTGCAGAA
AR020	CATATAGCCCGA	AR045	AATCTTGCGCCG
AR021	GCCTATGAGATC	AR046	AGGATCAGGGAA
AR022	CAAGTGAAGGGA	AR047	AATAACTAGGGT
AR023	CACGTTTATTCC	AR048	TATTGCAGCAGC
AR024	TAATCGGTGCCA	AR049	TGATGTGCTAAG
AR025	TGACTAATGGCC	AR050	GTAGTAGACCAT

Supplementary Table 7: Read Depths by Sample
Read depths did not differ significantly by sample.

Sample (Negative)	Reads	Sample (Positive)	Reads
AR001	177382	AR026	25868
AR002	34580	AR027	29106
AR003	109962	AR028	43414
AR004	46723	AR029	36513
AR005	29127	AR030	45852
AR006	36463	AR031	42730
AR007	43240	AR032	33965
AR008	39282	AR033	13797
AR009	50435	AR034	29927
AR010	35716	AR035	43143
AR011	29119	AR036	40509
AR012	35924	AR037	27851
AR013	39648	AR038	31072
AR014	40182	AR039	25288
AR015	22118	AR040	31453
AR016	41157	AR041	27389
AR017	27675	AR042	14 (not analyzed)
AR018	38916	AR043	27188
AR019	25422	AR044	41881
AR020	48118	AR045	31704
AR021	27700	AR046	57774
AR022	36180	AR047	31301
AR023	39116	AR048	32661
AR024	33414	AR049	39150
AR025	35564	AR050	26365

Supplementary Table 8: Genus-specific Primers and Cycle Conditions Used for qPCR Confirmation of Genus Changes

Genus	PCR Primers	Cycle Conditions
Megasphaera	Forward: AGAGACTGCCGCAGACAATGCGGAGG Reverse: TTTGGGGTTTGCTCCGGATCGCTCCTT	98C for 10 seconds 74C for 30 seconds (2-step)
Dialister	Forward: GGAAACTGGGAAGCTGGAGTATC Reverse: TTAATCTTGCGATCGTACTTCCCAGG	98C for 10 seconds 66C for 10 seconds 72C for 10 seconds
Peptococcus	Forward: AGTGGGGAATAACAGTGAGAAATCA Reverse: TCTCTTGGATGAGGACAGAGTTTT	98C for 10 seconds 65C for 10 seconds 72C for 10 seconds
Prevotella	Forward: CTATGGGTTGTAAACTGCT Reverse: ACATTTCACAACACGCTTA	98C for 10 seconds 56C for 10 seconds 72C for 10 seconds
Olsenella	Forward: GGTGAAGCGGCGGAGACGCCGTGGCCG Reverse: GGTCTCGCATGGGTGCCCGGCCGAA	98C for 10 seconds 74C for 30 seconds (2-step)
Alloprevotella	Forward: AGAAAAAGGACCGGCTAATT Reverse: AGTTTCAACTGCA	98C for 10 seconds 59C for 10 seconds 72C for 10 seconds
Haemophilus	Forward: ATAACTACGGGAAACTGTAGCTAAT Reverse: ACACCTCACTTAAGTCACCG	98C for 10 seconds 58C for 10 seconds 72C for 10 seconds
Parabacteroides	Forward: ACCCGGGTTTGAACG Reverse: CAGCTTACGCTGGCAGTC	98C for 10 seconds 60C for 10 seconds 72C for 10 seconds
Subdoligranulum	Forward: GGCATCGGATTGAGGGAAA Reverse: TGTCTCAGTCCCAATGTGGC	98C for 10 seconds 61C for 10 seconds 72C for 10 seconds
Total 16S	Forward: CCAGCAGCYGCGGTAAN Reverse: GGACTACHVGGGTWTCTAATCC	98C for 10 seconds 55C for 10 seconds 72C for 10 seconds

Supplementary Data 1: Taxonomic Classifications from 16S rRNA Sequencing with QIIME2 Analysis
Tab 1 includes phylum-level data, Tab 2 includes class-level data, Tab 3 includes order-level data, Tab 4 includes family-level data, and Tab 5 includes genus-level data.
This file is available at the Brown Digital Repository at https://doi.org/10.26300/vmxr-rm64.

All Supplementary Data files for this thesis can be found at the Brown Digital Repository at https://doi.org/10.26300/enej-vt18.

CHAPTER TWO: REDUCTIONS IN ANTI-INFLAMMATORY GUT BACTERIA ARE ASSOCIATED WITH MAJOR DEPRESSIVE DISORDER IN AMERICAN YOUNG ADULTS

Adapted from

"Reductions in anti-inflammatory gut bacteria are associated with major depressive disorder in American young adults"

By Richard Liu, Aislinn D. Rowan-Nash*, Ana E. Sheehan, Rachel F.L. Walsh, Christina M. Sanzari, and Peter Belenky*

Manuscript in Preparation

**These authors contributed equally.*

CONTENTS

Abstract 76

Introduction 76

Results 80

Discussion 89

Conclusions 96

Methods 97

Main Figures and Tables 104

Supplementary Materials 111

Reductions in anti-inflammatory gut bacteria are associated with major depressive disorder in American young adults

Richard T. Liu[1#*], Aislinn D. Rowan-Nash[2*], Ana E. Sheehan[3], Rachel F.L. Walsh[4], Christina M. Sanzari[1], and Peter Belenky[2]

[1] Department of Psychiatry and Human Behavior, Alpert Medical School of Brown University, Providence, RI, USA

[2] Department of Molecular Microbiology and Immunology, Brown University, Providence, RI, USA

[3] Department of Psychological and Brain Sciences, University of Delaware, Newark, DE, USA

[4] Department of Psychology, Temple University, Philadelphia, PA, USA

*These authors contributed equally

Corresponding author: Richard Liu, richard_liu@brown.edu

Abstract

We assessed the gut microbiota of 90 American young adults, comparing 43 participants with major depressive disorder (MDD) and 47 healthy controls, and found that the MDD cohort had significantly different gut microbiota compared to the healthy controls at multiple taxonomic levels. At the phylum level, participants with MDD had lower levels of Firmicutes and higher levels of Bacteroidetes, with similar trends in the at the class (Clostridia and Bacteroidia) and order (Clostridiales and Bacteroidales) levels. At the genus level, the MDD cohort had lower levels of *Faecalibacterium* and other related members of the family *Ruminococcaceae*, which was also reduced relative to healthy controls. Contrarily, the genus *Flavonifractor* was enriched in participants with MDD. Accordingly, predicted functional differences between the two cohorts include a reduced abundance of short-chain fatty acid production pathways in the MDD group. We also demonstrated that the magnitude of taxonomic changes was associated with the severity of depressive symptoms in many cases, and that most changes were present regardless of whether depressed participants were taking prescription psychotropic medications. Overall, our results support a link between MDD and lower levels of anti-inflammatory, butyrate-producing bacteria, and may support a connection between the gut microbiota and the chronic, low-grade inflammation often observed in MDD patients.

Introduction

There is increasing recognition of the fact that the gut microbiota is associated with a wide range of health conditions and disease states in human and animal hosts. Alterations to gut microbiome composition or function have been linked to gastrointestinal disorders[326,327], autoimmune disorders[328-333], and metabolic and cardiovascular disease[333-

[335]. Perhaps most surprisingly, given the physiological distance and the presence of the blood-brain barrier, the gut microbiota has also been implicated in psychiatric disorders or syndromes, including anxiety disorders [336-338], bipolar disorder[339-342], and major depressive disorder (MDD)[336,343-351].

Links between the gut microbiota and the brain are likely mediated in part through the gut-brain axis (GBA), a proposed series of complex communication pathways between the gastrointestinal tract and the central nervous system[352-358]. The GBA includes neural, immune, endocrine, and metabolic pathways involved in the regulation of hunger and satiety, stress, immunity, and intestinal motility; the gut microbiota are believed to play a direct role in some of these. In particular, microbes or their metabolites such as short-chain fatty acids (SCFA) can stimulate afferent inputs to the vagus nerve[359-361], induce enteroendocrine cells to produce neuropeptides and activate afferent nerve pathways[362,363], promote normal stress responses and hypothalamic-pituitary-adrenal (HPA) axis development and function[364,365], and participate in local neurotransmitter production and systemic regulation via tryptophan metabolism or direct secretion[354,366-368]. In the other direction, the vagus nerve can promote anti-inflammatory responses and decrease intestinal permeability[359,360,369,370], stress-induced glucocorticoid induction through the HPA axis can lead to microbial changes and increased gut barrier permeability[371-377], and the central nervous system can influence the gut environment through release of signalling molecules, changes to mucus secretion, and regulation of intestinal motility[353,357,378-380].

In terms of depression, links to the microbiome have been established in both human and animal models[381-384]. In one particularly compelling case, fecal transplants from humans with MDD resulted in the development of depressive symptoms in a germ-free

mouse model[345]. Similarly, fecal transplants from depressed human subjects into a germ-free rat model induced development of depressive symptoms, including anhedonia and anxiety, in addition to changes in tryptophan metabolism[385]. Additionally, certain microbes including *Faecalibacterium prausnitzii*, *Lactobacillus* spp., and *Bifidobacterium* spp. have been found to ameliorate the onset of anxiety and depressive symptoms that rodents develop when subjected to chronic unpredictable mild stress, maternal separation, or chemically-induced colitis[386-388]. In several cases, although not all, these effects were dependent on the presence of an intact vagus nerve[387,388], emphasizing the importance of this pathway in communication from the gut and microbiota to the brain.

In humans, studies have consistently indicated that the gut microbiota of adults with MDD are different from those of their healthy counterparts, although specific differences have varied between studies. Some studies have found that the phylum Bacteroidetes is underrepresented in subjects with depression, while Firmicutes are overrepresented[345,346,349-351], although other studies have found the opposite trend[343,344]. Multiple studies have linked higher abundance of the genera *Alistipes, Oscillibacter,* and *Flavonifractor* and the family *Enterobacteriaceae* to MDD and low quality of life scores[343,345-350], while *Faecalibacterium*, *Dialister*, *Coprococcus*, and *Prevotella* have been found to be lower in subjects with depression and/or low quality of life scores[343,345,347,348,385]. In at least one study, *Faecalibacterium* has even been found to negatively correlate with depression severity[343], and this species has also been found to be negatively associated with both bipolar disorder and generalized anxiety disorder[337,339].

A potential link between the gut microbiota and MDD is the low-grade, chronic inflammation that has previously been observed in a substantial proportion of depressed

individuals[389-391]. Significant subsets of depressed subjects have been associated with higher levels of circulating inflammatory cytokines, particularly IL-6 and TNF-α[392-397], in addition to hypercortisolism and dysregulation of the HPA axis[398-400]. Furthermore, a few studies have demonstrated that combining antidepressants with anti-inflammatory drugs improved response rates[401,402], and inflammasome signalling has been linked to induction of anxiety and depressive behaviors in mice[403-407]. Human patients with chronic inflammatory illnesses have higher levels of depression than the general population[408-412], and administration of inflammatory cytokines or immune-provoking stimuli such as lipopolysaccharide (LPS) leads to the development of "sickness behavior" and depressive symptoms in both animal models and human patients[413-421]. Mechanistically, inflammatory cytokines may increase blood-brain barrier permeability and in some cases cross it[421-427], activate vagus nerve afferents[370,426-430], impact neurotransmitter levels in the brain[426,427,431-433], contribute to hyperactive dysregulation of the HPA axis[434-439], and affect serotonergic neurotransmission by promoting enzymatic metabolism of the precursor tryptophan[421,440-445].

The gut microbiota may contribute to such an effect through their capacity to either promote or protect against inflammation. For example, loss of bacteria that produce the anti-inflammatory, barrier-strengthening molecule butyrate, such as *Faecalibacterium* or *Coprococcus*, could lead to a loss of protection against epithelial inflammation and gut barrier disruption. Combined with increases in LPS-producing bacterial groups such as Proteobacteria or potentially pro-inflammatory species such as *Flavonifractor*, this could lead to increased translocation of immunogenic bacterial products and activation of low-grade systemic inflammation. In fact, studies have found that depressed subjects have

increased levels of bacterial DNA in circulation and increased antibody responses to LPS[446-450].

In this study, we analyzed the gut microbiota of American young adults with major depressive disorder and healthy controls. We hypothesized that we would observe potential signatures of inflammation, including either reductions in protective, butyrate-producing bacterial taxa or increases in pro-inflammatory taxa. Our study differs from previous studies in terms of demographics, as most previous studies have been performed on subjects of Chinese heritage[343-345,347,349-351] and a few have examined European subjects[346,348,385], but to our knowledge none have so far assessed differences in American subjects. Furthermore, most previous studies have examined older age cohorts, with age ranges in the thirties and forties, while we utilized a cohort of participants aged 18-25. Finally, our study analyzed a large cohort which included both a notable subset of MDD participants who were not taking psychotropic medications as well as MDD subjects with a range of symptom severities. This allowed us to assess the potential contributions of these characteristics to changes in the gut microbiota observed in participants with MDD. Overall, we observed that the MDD cohort exhibited lower levels of potentially protective taxa, including *Faecalibacterium* and *Subdoligranulum*, and higher levels of potentially pro-inflammatory taxa, including *Flavonifractor* and Gammaproteobacteria; furthermore, many of these changes track with symptom severity.

Results

Participant Characteristics

All participants were young adults with an average age of 21.7 in the control cohort and 22.7 in the MDD cohort (Table 1, Figure 1A). The average PROMIS score recorded at

the in-person assessment was 9.3 in the control group and 25.0 in the MDD group (Table 1, Figure 1B). The participants as a whole were predominantly assigned female at birth (80.0%), though the MDD group was slightly more female-biased than the control group (88.4% vs 72.3%) (Table 1, Figure 1C); this is unsurprising, given the higher incidence of MDD in women than in men[451,452]. Additionally, when considering gender identity, the percentage of female-identifying subjects was more equal between the two groups (74.4% vs 72.3%) (Table 1, Figure 1D), which can be attributed to several assigned-female-at-birth subjects in the MDD group who did not identify as female. This aligns with research demonstrating higher incidence of depression in transgender or gender-nonconforming youths than in their cisgender peers[453-455]. The cohort was also predominantly white, with little difference between the two groups (80.1% in controls and 76.7% in MDD) (Table 1, Figure 1E), and primarily non-Hispanic (93.3% in controls and 86.0% in MDD) (Table 1, Figure 1F).

Of the subjects in the MDD cohort, 15 (34.9%) were not actively taking prescribed psychotropic medications, while 28 (62.8%) were taking at least one such medication; specifically, 30.0% were taking a single medication, while 34.9% were taking two or more medications (Supplementary Table 1-2, Figure 1G). These medications were quite varied, and included a range of both anti-depressants, anxiolytics, and stimulants; in the control cohort, only one subject was taking a psychotropic medication, specifically an amphetamine for ADHD (Supplementary Table 2, Figure 1H).

Alpha and Beta Diversity

We started by analyzing the alpha diversity of the control and MDD cohorts. We first utilized the Observed ASVs metric, which assesses the richness of ASVs found in the

samples without considering their phylogenetic relatedness, and Shannon's Diversity Index, which reflects both the richness and evenness of the samples. On both of these measures, there was no significant difference between the two groups (Supplementary Figure 1A-B). However, when using Faith's Phylogenetic Diversity, a biodiversity metric which analyzes phylogenetic tree branch length to incorporate relatedness of taxa, we observed a slight but significant decrease in this metric in the MDD cohort (Figure 2A). Furthermore, this diversity metric was inversely related to the severity of depressive symptoms (Figure 2B), and was not impacted by the usage of psychotropic medications (Figure 2C).

We then utilized metrics of beta diversity to assess whether the healthy and MDD microbiomes were different at a whole-community level. First, we found that there were statistically significant differences in community composition between the two groups based on Bray-Curtis Dissimilarity, which does not consider phylogenetic relatedness of taxa in a sample. Similarly, there was a significant difference in community composition based on Unifrac Distance, which does take relatedness into account. However, the permANOVA indicates that the condition (control or MDD) explains very little of the discrimination between samples ($R^2=0.018$ *and* $R^2=0.014$, respectively); accordingly, when plotting these metrics using a Principal Coordinate Analysis (PCoA), we did not observe clear separation of the two groups (Figure 2D-E). In the case of Bray-Curtis Dissimilarity, there was a clear clustering of several samples in both groups separately from the majority, which could be attributed to the dominance of *Prevotella 9* rather than *Bacteroides* as the predominant genus-level taxon of the phylum Bacteroidetes in those samples (Supplementary Figure 1C). This pattern was not observed in the PCoA for

Unifrac Distance (Supplementary Figure 1D), consistent with the fact that this metric accounts for the two genera's taxonomic relatedness. Finally, using Weighted Unifrac Distance, which accounts for both phylogenetic relatedness and abundance of taxa, there was no significant difference between the groups and no clear clustering by condition (Supplementary Figure 1E).

Taxonomic Composition

We next analyzed the composition of the samples at multiple taxonomic levels to assess whether there were differences in the communities of depressed and healthy subjects. We utilized the Linear Discriminant Analysis Effect Size (LEfSe) tool to identify taxa that were biomarkers of each group and found a range of discriminating taxa across all taxonomic levels (Figure 3, Supplementary Table 4). At the phylum level, controls were enriched in Firmicutes and the MDD cohort was enriched in Bacteroidetes (Figure 4A-C, Supplementary Table 4). Similar findings were obtained at the class and order levels (Supplementary Figure 2A-B), with the class Clostridia and order Clostridiales of Firmicutes associated with controls while the class Bacteroidia and order Bacteroidales of Bacteroidetes were associated with depressed subjects (Supplementary Figure 2C-F, Supplementary Table 4). Additionally, within the phylum Proteobacteria, the order Rhodospirillales of the class Alphaproteobacteria was associated with controls, while the class Gammaproteobacteria was associated with the MDD cohort (Supplementary Figure 2G-H). At the family level (Supplementary Figure 2I, Supplementary Table 4), the Clostridiales families *Ruminococcaceae* and *Christensenellaceae* were associated with the control group (Supplementary Figure 2J-K), as well as the Bacteroidetes family *Barnesiellaceae* and an uncultured family of the order Rhodospirillales (Supplementary

Figure 2L-M), while the family *Enterococcaceae* (of the Bacilli-Lactobacillales lineage of Firmicutes) was associated with the MDD cohort (Supplementary Figure 2N).

A number of trends at the family level were also represented in their subordinate genera, although there were some contrasting patterns as well (Figure 5A, Supplementary Table 4). First, the four most abundant genera within the Firmicutes family *Ruminococcaceae* (*Faecalibacterium*, *Subdoligranulum*, *[Eubacterium] coprostanoligenes* group, and *Ruminococcus 1*) were associated with healthy controls (Figure 5B-E), as well as the less-abundant genus-level taxon CAG-352 (Figure 5F). On the contrary, only a single *Ruminococcaceae* genus, *Flavonifractor*, was associated with MDD subjects (Figure 5G). Additionally, while the related Clostridiales family *Lachnospiraceae* itself was not associated with either group, a number of its member genera were associated with healthy controls, including *Fusicatenibacter*, *Tyzzerella 3*, and *[Eubacterium] ventriosum* group (Figure 5H-J). In contrast, the *Lachnospiraceae* genus *Sellimonas* was associated with the MDD cohort (Figure 5K). The genus-level taxon {*Lachnospiraceae*} UCG-001 was also called as associated with the MDD cohort, but a closer examination reveals that this taxon is in fact more highly present and abundant in the control cohort, and the misidentification is likely due to one extreme outlier in the MDD cohort (Figure 5L).

Other discriminatory genus-level taxa within the phylum Firmicutes included an uncultured organism of the Clostridiales vadinBB60 family and the R-7 group of *Christensenellaceae*, both associated with controls (Figure 5M-N), and *Enterococcus*, which was associated with MDD subjects (Figure 5O). Within the phylum Bacteroidetes, *Barnesiella* and an uncultured bacterium of the *Muribaculaceae* family were associated

with controls (Figure 5P-Q), while no genus-level taxa were associated with MDD subjects. Finally, within Proteobacteria, the Deltaproteobacteria genus *Desulfovibrio* was associated with healthy controls (Figure 5R). As discussed in greater depth in the discussion, these changes generally appear to reflect a loss of protective bacteria and an increase in pro-inflammatory bacteria in the MDD cohort.

Impact of Psychotropic Medication

We then examined whether any of the associations of taxa with depressed subjects were driven by the consumption of psychotropic medication. Of our MDD cohort of 43 subjects, 15 were taking no prescribed psychotropic medication, 13 were taking a single drug, and 15 were taking multiple. We compared the taxa that were identified as significant between the healthy controls, depressed subjects taking no medication, and depressed subjects taking one or more medications. Unfortunately, as the number and types of medications varied significantly, we could not assess the impacts of specific classes or combinations of medications.

Generally, we found that the trends observed in the depressed subjects were present in both the medicated and unmedicated groups, although the effect was sometimes stronger in one group than the other (Supplementary Figure 3). Specifically, the MDD-associated reductions observed in the phylum Firmicutes, the class Clostridia, the order Clostridiales, the family *Ruminococcaceae*, and the genus-level taxa *Faecalibacterium*, and *[Eubacterium] coprostanoligenes* group, were slightly stronger in the medicated group (Supplementary Figure 3A-F). Similarly, the depression-associated increases in the phylum Bacteroidetes, the class Bacteroidia, and the order Bacteroidales, were somewhat stronger in this group (Supplementary Figure 3G-I). In particular, the changes in the genera

Flavonifractor and *Sellimonas* appeared to be driven primarily by the medicated group (Supplementary Figure 3J-K). On the other hand, the MDD-associated reductions in the families *Christensenellaceae* and *Barnesiellaceae*, and genera *Christensenellaceae* R-7 group, *Barnesiella*, and an uncultured organism of the Clostridiales vadinBB60 family-level taxon appear somewhat stronger in the unmedicated group (Supplementary Figure 3L-P). In the case of the genus *Fusicatenibacter*, reductions in the depressed group appeared to be primarily driven by the unmedicated subjects (Supplementary Figure 3Q). There were also a number of cases in which there were no apparent differences between the depressed subjects based on medication (Supplementary Figure 3R-AA). Finally, in the case of the family *Enterococcaceae* and its daughter genus *Enterococcus*, levels of these taxa were found only in medicated subjects, suggesting that this group may drive the effect (Supplementary Figure 3AB-AC); however, as they were detected in so few samples, it is difficult to make this judgment.

Overall, the differences between depressed and healthy subjects could not be specifically attributed to psychotropic medications, as changes were typically present in both medicated and unmedicated subjects, although they may play a role in some cases. Additionally, we were unable to examine the impacts of specific classes of psychotropic drugs, given the variety and combinations of medications taken by participants.

Impact of Depression Severity

While psychotropic medication use may have had some impact on taxonomic trends, we also found that depression severity could significantly confound this interpretation. Based on PROMIS Depression scores, the proportion of subjects taking psychotropic medication increased based on their symptom severity: 46.7% of the 15

subjects with mild symptoms (PROMIS < 23) were taking psychotropic medication, compared with 68.2% of the 22 subjects with moderate symptoms (PROMIS 23-32) and 83.3% of the 6 subjects with severe symptoms (PROMIS > 32) (Supplementary Figure 4A). Accordingly, the proportions of subjects with mild symptoms was higher in the unmedicated group (53.3% vs. 28.6%), while the proportion with severe symptoms was lower (6.7% vs. 17.9%) (Supplementary Figure 4B). Therefore, we decided to assess whether any of the observed trends tracked with symptom severity, which might explain differences better than whether or not depressed subjects were taking one or more of a wide range of psychotropic drugs with different mechanisms of action.

In fact, despite interindividual variation, the observed changes were more exaggerated in subjects with higher depressive symptom severity scores in the majority of taxa (Figure 6A); this trend was significant in most of these cases, although a few fell just short of statistical significance (Supplementary Table 3). Specifically, the phylum Firmicutes, class Clostridia, order Clostridiales, family *Ruminococcaceae* and its member genera *Faecalibacterium*, *[Eubacterium] coprostanoligenes* group, *Subdoligranulum*, and *Ruminococcus 1*, and family *Christensenellaceae* and member genus-level taxon *Christensenellaceae* R-7 group were more reduced in subjects with more severe symptoms (Figure 6B-K). In the case of *Fusicatenibacter,* reductions were most pronounced in the subjects with severe symptoms, although there was not a consistent trend in the mild and moderate symptom groups (Figure 6L). In the other direction, the phylum Bacteroidetes, classes Bacteroidia and Gammaproteobacteria, order Bacteroidales, and genera *Flavonifractor* and *Sellimonas* were more abundant in subjects with more severe symptoms (Figure 6M-R). There were also a few cases in which such a trend was not present or where

there were very few depressed samples with detectable levels of a given taxon, making it difficult to assess whether their abundance aligns with symptom severity. (Figure 6S-AD).

In general, these results suggest that symptom severity tracks with changes in a number of discriminatory taxa. While the impact of psychotropic drug usage on this pattern cannot be ruled out due to the higher levels of medication utilization in subjects with more severe depressive symptoms (Supplementary Figure 4A-B), the high degree of variability in the number and classes of drugs used (Figure 1, Supplementary Tables 2-3) tends to suggest that the effect is more likely related to symptom severity.

Functional Predictions

We utilized Phylogenetic Investigation of Communities by Reconstruction of Unobserved States 2 (PICRUSt2) to predict the functional potential of the healthy and depressed microbial communities based on the 16S rRNA gene content. It should be noted that this is only a prediction based on the inference from the 16S rRNA content and cannot definitively measure functional potential or transcriptional activity. When assessing function at the MetaCyc pathway level, LEfSe detected a number of pathways that were associated with each cohort. Pathways associated with the MDD cohort tended to be related to vitamin (folate and thiamine) biosynthesis, LPS biosynthesis, and long-chain fatty acid biosynthesis. On the other hand, pathways associated with the healthy cohort tended to be related to fermentation to short chain fatty acids, phospholipid biosynthesis, nucleic acid metabolism, and aliphatic amino acid biosynthesis (Supplementary Figure 5). Of particular note is the association with healthy controls of PWY 5676, "acetyl CoA fermentation to butanoate II", as fermentation of acetyl-CoA is the dominant pathway by which the gut microbiota, including *Faecalibacterium prausnitzii*, produce the anti-inflammatory short-

chain fatty acid butyrate[456,457]. The control cohort was also enriched in the PWY5100, "pyruvate fermentation to acetate and lactate II", which produces another major microbial SCFA, acetate. Additionally, the MDD cohort's association with LPS production (PWY1269 – "CMP-3-deoxy-D-manno-octulosonate biosynthesis", NAGLIPASYN – "lipid IVA biosynthesis [*E. coli*]", PWY6467 – "KDO transfer to lipid IVA III [*Chlamydia*]", PWY7323 – "superpathway of GDP-mannose-derived O-antigen building blocks"), likely due to the enrichment in the Gram-negative phylum Bacteroidetes and class Gammaproteobacteria, is notable, as Proteobacteria-derived LPS in particular is known to be immunogenic and has been linked to chronic inflammation[458,459].

Discussion

Depression has been previously been found to be associated with differences in the gut microbiota, and our study adds to this body of work. One of the key concepts in microbiome research is the heterogeneity of microbial populations between various groups of people. Thus, as we compare our conclusions to previous work, it is important to consider the factors that make this study unique relative to other work. First, the majority of previous MDD-microbiome studies have been undertaken in Chinese populations[343-345,349-351], with a few others studying European cohorts[346,348,385], while our study focused on an American population. Given that the underlying microbiome can differ significantly by geography[90,460-463], it is important to study the impacts of disease in a range of populations. Second, most previous studies have studied a wider and older age range than our own, which recruited only subjects between the ages of 18 and 25. As the microbiome can change through the lifespan[464], recruitment of a narrow age range can limit underlying noise and increase power to detect differences between groups. Third, we were able to

recruit a notable subset of MDD participants who were not taking psychotropic medication, which allowed us to compare this group to the larger group of participants who were using these drugs, which most previous studies were not able to do. Finally, we were able to recruit MDD participants with a range of depressive symptom severities, which allowed us to assess whether the changes we observed in the MDD cohort were related to this metric. Based on this design, we observed a number of notable differences between MDD and control participants, and further found that many of these changes may track with symptom severity.

Despite participant differences, many of our results do align with observations made previously in other populations. Most notably, we found that *Faecalibacterium* levels were reduced in subjects with MDD, supporting previous work linking lower levels of this genus to depression or bipolar disorder and lower quality of life[339,343-345,347,348]. We also found this pattern in a number of related genera within the family *Ruminococcaceae*, including *Subdoligranulum*, *Ruminococcus 1*, and *[Eubacterium] coprostanoligenes* group. There was a member of this family – *Flavonifractor* – which was instead more abundant in the MDD subjects, supporting previous work finding that higher levels of this genus are associated with depression, bipolar disorder, and lower quality of life[340,343,348]. Remarkably, we found that many of differences that we observed were exacerbated in subjects with more severe depressive symptoms, which had previously only been observed for *Faecalibacterium* in major depressive disorder and bipolar disorder[339,343]. In general, our results, particularly the negative correlation between *Faecalibacterium* levels and depressive symptoms and the association of MDD with increased levels of *Flavonifractor* align most closely with those of Jiang *et al*[343]. This may relate to the fact that this study

utilized a participant group closer in age to our own, with average age in the mid-twenties, compared with other studies where the average ages ranged from mid-thirties to late forties. Further work to examine whether microbiome alterations in depressed subjects are related to age may be warranted.

There were also a few trends that were contradictory to prior data. In particular, while several studies have found lower levels of the phylum Bacteroidetes and higher levels of the phylum Firmicutes in depressed subjects[345-347,349-351], we found the opposite trend. Furthermore, while Jiang *et al* did find that Bacteroidetes were higher in depressed subjects, this was in fact driven by increases in the families *Rikenellaceae* and *Porphyromonadaceae*, while *Bacteroidaceae* was actually slightly reduced[343]; this contrasts with our data, in which increases in the levels of Bacteroidetes in our MDD cohort were largely driven by increases in *Bacteroidaceae* and *Prevotellaceae*, although these differences were not significant at the family level. Additionally, while some studies have found increases in the Gammaproteobacteria family of *Enterobacteriaceae* in depressed subjects[343], we did not observe such a change. While we found that Gammaproteobacteria was associated with the MDD cohort, an examination of the data suggests that this increase was driven by the genus *Parasutterella* of the family *Burkholderiaceae*. Finally, while lower levels of *Fusicatenibacter* were found in a previous study to be associated with lower quality-of-life scores and potentially with depression, the relationship with depression was not found in the non-medicated subset of subjects[348]. This is in contrast to our results, where we instead found that low levels of this genus were primarily found in the non-medicated MDD cohort. Importantly, as noted previous, our study is fairly demographically distinct

from previous studies of the gut microbiota in depression, which could be responsible for some of the disparities between our results and those of prior studies.

In light of the links between MDD and chronic inflammation, a number of microbiota differences observed in our study are notable. Perhaps most interesting is the relationship between MDD subjects and lower levels of the family *Ruminococcaceae* and its daughter genera *Faecalibacterium*, *Subdoligranulum*, *Ruminococcus 1*, and *[Eubacterium] coprostanoligenes* group. The genus *Faecalibacterium* includes only one named species, *F. prausnitzii*, which has been demonstrated to have anti-inflammatory properties[161-163,465]. Importantly, it produces the short-chain fatty acid butyrate[466], which serves as a colonic fuel source, fosters immunoregulation, and promotes epithelial barrier integrity[156,160,467-472]. In fact, lower levels of this genus and species have been associated with IBD, *Clostridiodes difficile* colitis, autoimmune disorders, and atherosclerotic cardiovascular disease[161,304,326,330,331,335,473-476], in addition to mental health disorders including depression, bipolar disorder, and generalized anxiety disorder[337,339,343-345]. Similarly, *Subdoligranulum* includes only a single named species, *S. variable*, which is also known to produce butyrate[466] and has been negatively associated with IBD and type I diabetes[330,474,477]. Additionally, while not previously linked to MDD specifically, it was found to be less abundant in subjects with generalized anxiety disorder and correlated negatively with depressive symptoms in this population[337,338]. The genus-level taxon *Ruminococcus 1* includes the species *R. albus* and *R. callidus*, both of which have also been negatively linked to IBD[476], although they do not themselves produce butyrate[466].

The relationship between *[Eubacterium] coprostanoligenes* group and MDD is less clear, as this genus-level taxon has not been specifically linked to depression or

inflammation. However, like *Subdoligranulum*, its abundance was decreased in subjects with generalized anxiety disorder and negatively correlated with depression and anxiety scale scores in this cohort[338]. It is named for primary component species *E. coprostanoligenes*, which is known for its ability to reduce cholesterol to coprostanol, which is less-well-absorbed by the host[478,479]. Fecal coprostanol levels or the ratio of fecal coprostanol/coprostanone to cholesterol have been found to be reduced in Crohn's disease, ulcerative colitis, and *Clostridioides difficile* colitis[480-484], suggesting that a reduction in *E. coprostanoligenes* or related species could contribute to inflammation through increased colonic cholesterol levels. Evidence suggests that this may relate to serum cholesterol as well, as administration of *E. coprostanoligenes* to mice and rabbits decreased serum cholesterol levels[485,486] and a human study found that serum cholesterol was inversely related to fecal coprostanol:cholesterol ratios[487]. However, while higher levels of cholesterol have generally been thought to be negative for health, there is some evidence that very low cholesterol levels are associated with severe depression or suicidality, although the evidence is mixed and may depend on gender and the type of cholesterol (HDL vs LDL)[488-490].

Flavonifractor, the sole member of *Ruminococcaceae* that was associated with the depressed cohort, has previously been linked to lower quality of life scores and MDD[343,348] as well as bipolar disorder[340] and generalized anxiety disorder[337]. This genus, which currently includes the single named species *F. plautii* (formerly *Clostridium orbiscindens* and *Eubacterium plautii*)[491], has also previously been linked to various autoimmune disorders[328,332,492,493], chronic kidney disease[494], and colorectal cancer[495,496]. Furthermore, *F. plautii* was demonstrated *in vitro* to have epithelial invasive potential[497]. Therefore, there

is significant evidence to suggest that unlike its generally anti-inflammatory relatives, *Flavonifractor* may be associated with disease despite its ability to produce butyrate under some conditions[491,498]. As suggested in other work[340,496], this is possibly related to the genus' eponymous capacity for cleaving flavonoids that reach the colon, including antioxidants such as quercetin, although disentangling the impacts of flavonoids and their microbial breakdown products on inflammation *in vivo* is difficult[491,499-504].

In addition to the *Ruminococcaceae*, some other members of Firmicutes were altered in the depressed cohort, including a few members of the family *Lachnospiraceae*. The genus-level taxon *Tyzzerella 3* was associated with the control cohort, although it has previously been linked to generalized anxiety disorder[338]; however, it was also linked to healthy controls in a study of chronic kidney disease[494]. *[Eubacterium] ventriosum* group was also associated with controls, and this genus-level taxon has been found to produce butyrate and inversely correlate with the inflammatory cytokines IL-6 and IL-8[505,506]. Finally, *Fusicatenibacter* was also more abundant in the control cohort, and was particularly reduced in the subjects with severe symptoms; its primary member species *F. saccharivorans* has previously been found to be reduced in subjects with active IBD and colorectal cancer[507,508]. On the other hand, the genus *Sellimonas* was associated with the MDD cohort, and has been previously linked to rheumatoid arthritis[492] and chronic kidney disease[494].

Previous work has demonstrated that the genera *Barnesiella* and *Christensenellaceae R-7* were associated with healthy controls in comparison to various inflammation-related gut diseases, including IBD, colorectal cancer, and *C. difficile* colitis[493]; however, these taxa have not previously been linked to depression. Finally, in

the family Proteobacteria, the genus *Desulfovibrio* was associated with the healthy control cohort, although in previous studies it has been found to be associated with IBD and experimental colitis models[300,509,510]; however, in a study of generalized anxiety disorder, its source family of *Desulfovibrionaceae* was associated with healthy controls[337]. Additionally, the class Gammaproteobacteria was associated with MDD, which was largely driven by increases in the family *Burkholderiaceae* and genus *Parasutterella*, which has itself been linked to MDD[343]. Relatedly, we also observed a predicted enrichment in LPS biosynthesis pathways in the MDD cohort, which was also observed by Huang *et al*[344]; the LPS of members of Gammaproteobacteria is known to be immunogenic relative to that of some commensal Gram-negative bacteria such as *Bacteroides* species[458,459], and depressed subjects have been found to have increased serum immunoglobulin A and M responses against the LPS of members of this class[446,447].

Some limitations of this study must be acknowledged. First, while there are clear trends that significantly discriminatory taxa became increasingly divergent with increasing symptom severity, we cannot completely disentangle this phenomenon from the taking of psychotropic drugs given the strong degree of overlap between symptom severity and medication usage. Future work specifically focusing on newly-diagnosed subjects who have not previously taken psychotropic drugs would be beneficial in understanding the relationship between the microbiota, depressive symptom severity, and medication. Additionally, we utilized 16S rRNA sequencing, which can generally only identify taxa down to the genus level, so we may be missing important species- or strain-level differences between the communities. Furthermore, while we can use taxonomic composition to predict the gene content of the communities, full metagenomics and

transcriptomics would be required to comment further on changes in the functional potential or transcriptional activity, such as potentially reduced capacity for SCFA production in depressed subjects. Finally, we cannot account for complex confounding variables that may cause microbiome differences between the two groups not specifically related to disease pathology, such as differential dietary habits either before or after the onset of depressive symptoms.

Conclusions

The microbiomes of American young adults with major depressive disorder were found to be significantly different from those of healthy controls. At high taxonomic levels, depressed subjects had lower levels of the phylum Firmicutes, class Clostridia, and order Clostridiales, and correspondingly higher levels of the phylum Bacteroidetes, classes Bacteroidia and Gammaproteobacteria, and order Bacteroidales. Most notably, subjects with MDD had lower levels of the families *Ruminococcaceae* (including the genera *Faecalibacterium*, *Subdoligranulum*, *Ruminococcus 1*, and *[Eubacterium] coprostanoligenes* group), *Christensenellaceae* (including the genus-level taxon *R-7* group), and *Barnesiellaceae* (including the genus *Barnesiella*). These subjects also had higher levels of the *Ruminococcaceae* genus *Flavonifractor* and the *Lachnospiraceae* genus *Sellimonas*. Additionally, we found that the majority of notable taxonomic changes in the depressed cohort were more pronounced in subjects with higher scores on a depressive symptom scale, although we cannot rule out the impact of psychotropic medication due to significant overlap of usage with symptom severity. Overall, our findings align with previous studies of the gut microbiota in subjects with depression, particularly that depressed subjects have lower levels of *Faecalibacterium* and higher

levels of *Flavonifractor*. In general, the differences that we observed are supportive of an inflammatory state in subjects with MDD, as these subjects tended to have lower levels of butyrate-producing, anti-inflammatory bacteria such as *Faecalibacterium* and *Subdoligranulum* and higher levels of taxa previously associated with inflammatory disorders such as *Flavonifractor* and Gammaproteobacteria. Importantly, there was significant overlap in the proportions of the discriminant taxa between the control and MDD cohorts in most cases, aligning with previous observations that inflammation may play a role in the etiology of depression in a significant subset of patients but is neither necessary nor sufficient to cause its onset.

Methods

Measures for Assessing Depression and Symptoms in Participants

The PROMIS (Patient-Reported Outcomes Measurement Information System) Depression – Short Form[511] was created by National Institutes of Health as part of the Roadmap for Medical Research initiative to use item-response-theory methodology to develop psychometrically advanced self-report measures of health outcomes. The adult depression short form consists of an eight-item questionnaire that assesses depressive symptoms over the past seven days. Response options were on a five-point Likert scale from “never” to “always.” Higher scores indicate greater depressive symptom severity.

The Structured Clinical Interview for DSM-5 Disorders (SCID-V)[512] is a semi-structured diagnostic interview of current and lifetime DSM-5 psychopathology. For the current study, only the depression module was administered. Research staff were trained by doctoral level clinicians and certified by the first author in the research procedures.

The Columbia-Suicide Severity Rating Scale (C-SSRS)[513] and Self-Injurious Thoughts and Behavior Interview (SITBI)[514] are semi-structured interviews for assessing lifetime history of suicidal thoughts and behaviors and non-suicidal self-injury. All interviewers received extensive training and supervision from the first author in the administration of this interview and rating of its data. A rigorous protocol developed by the first author was implemented, with an average training period of three to four months before interviewers administered the measure independently. Interviewers conferred with the first author whenever coding questions arose.

Participant Recruitment Procedures

Young adults were recruited from the community and local psychiatric clinics through flyers and social media advertisements. To participate in the study, prospective participants were required to be 18-25 years old, meet eligibility criteria for either the major depressive disorder (MDD) group or the healthy control group, and not be subject to any exclusion criteria. To be eligible for the healthy control group, subjects were required to have a PROMIS Depression score below 13, no lifetime history of major depression, no lifetime history of suicidal ideation, suicide attempts, or non-suicidal self-injury as assessed by the C-SSRS and SITBI. To be eligible for the MDD group, participants needed to meet diagnostic criteria for a current major depressive episode and have PROMIS Depression scores < 21. Individuals were excluded if they had smoked cigarettes or cigars in the past 12 months, were vegan, had gastrointestinal illness in the past six months, had diarrhea in the past two weeks, used anti-diarrhea medication in the past six weeks, or used antibiotics in the past three months.

Prospective participants first completed an online screener which included the PROMIS Depression scale and questions regarding the exclusion criteria. Those who remained potentially eligible based on the online screener then completed a phone screener, in which they answered the SCID-V, C-SSRS, and SITBI. Those who were eligible were invited to attend a one-hour in-person assessment, at which they were re-administered the PROMIS Depression questionnaire. At this visit, participants were provided with an OMNIgene•GUT stool collection kit (DNA Genotek) to collect a fecal sample for microbiome analysis, which was then mailed to Brown University.

Ultimately, we recruited 90 participants for this study. The healthy control group (n=47) had a mean age of 21.7±2.1, mean PROMIS Depression scores of 9.3±1.4, and was 72.3% assigned female at birth, 72.3% identifying as female, 76.7% white, and 86.0% non-Hispanic (Table 1, Figure 1A-F). The MDD group (n=43) had a mean age of 22.7±1.8, mean PROMIS Depression scores of 25.0±6.9, and was 88.4% assigned female at birth, 74.4% identifying as female, 80.1% white, and 93.3% non-Hispanic (Table 1, Figure 1A-F).

Extraction and Preparation

Upon receipt at Brown University, fecal samples were stored at -80°C until all samples had been collected. Samples were then thawed, and 300 µL of fecal suspension from each sample was transferred into two plates of the ZymoBIOMICS 96 DNA Kit (Zymo Research) to extract DNA. Samples from the two cohorts were randomized across the two 96-well plates. Extraction was performed according to manufacturer's protocols, and extracted DNA was measured using the Qubit 3.0 system with Broad-Range DNA reagents (Thermo Fisher Scientific).

Sequencing

Amplicons of the V4 hypervariable regions of the 16S rRNA gene were generated according to the Earth Microbiome Protocol[314]. In brief, 10 μg of extracted DNA from each sample was used as template for triplicate PCR reactions utilizing individually barcoded 515F forward primers (GTGYCAGCMGCCGCGGTAA) with Illumina adapters and the 806R reverse primer (GGACTACNVGGGTWTCTAAT) with Illumina adapters. Triplicates were combined and measured using the Qubit 3.0 system with Broad-Range DNA Reagents (Thermo Fisher Scientific). Samples were pooled in equimolar concentrations and sent out for 2x250 paired-end sequencing utilizing an Illumina MiSeq system at the University of Rhode Island. We obtained a total of 3,806,054 quality-filtered sequences across all 90 samples. The average sequencing depth was 41,509 reads in the control cohort and 44,169 in the MDD cohort. Sequences can be found at the NCBI Short Read Archive under BioProject ID PRJNA591924.

Data Analysis

Data was initially processed utilizing the QIIME2 (v2019.7) pipeline[515]. In brief, samples were imported using the *tools* plugin, demultiplexed using the *demux* plugin, and denoised using the *dada2* plugin to obtain amplicon sequence variants (ASVs)[317]. Phylogenetic trees were generated using the *phylogeny* plugin and taxonomy was assigned using the *feature-classifier* plugin and the Silva (release 132) 99% identity V4 classifier. Additionally, functional potential was predicted using the *picrust2* plugin[516]. The feature table, representative sequences, rooted phylogenetic tree, and taxonomy QIIME2 artifacts were exported, and the feature table, taxonomy, and sample metadata were merged into a unified biom file using the *add-metadata* function of the biom-format package[517].

The exported biom file, phylogenetic tree, and representative sequences were imported into the phyloseq package (v1.28.0)[318,319] in R (v3.6.1) using the *import_biom* function. Alpha diversity metrics were calculated using the *estimate_richness* function of phyloseq (Shannon's Diversity Index, Simpson's Diversity Index, Observed ASVs) and the *estimate_pd* function of the btools package (v0.0.1) (Faith's Phylogenetic Diversity). Beta diversity (Bray Curtis Dissimilarity and both Unweighted and Weighted Unifrac Distances) was calculated using the *phyloseq:distance* function of the vegan package (v2.5-6)[320], statistically analyzed using the *adonis* function, and subjected to Principal Coordinates Analysis using the *ordinate* function of vegan.

ASV tables were agglomerated at the phylum, class, order, family, and genus levels and exported as relative abundance tables for plotting and analysis. The genus-level table was reformatted to conform to the requirements of the Linear Discriminant Analysis Effect Size (LEfSe) web-based tool[518], and analysis to identify group biomarkers was performed according to default parameters. Similarly, the MetaCyc pathways output from picrust2 was formatted for and analyzed with LEfSe. All figures were generated using GraphPad Prism v8e, with the exception of Figure 3B, which is a modified LEfSe output. Statistics were performed in R for beta diversity metrics, LEfSe for differential abundance, and GraphPad Prism for alpha diversity metrics and tests for trends based on psychotropic drug usage or symptom severity. Throughout the text and figures, in cases where discriminant taxa had unclear or nonspecific names at that particular level, the next-higher taxon was included in curly brackets to indicate its provenance (for example, {*Muribaculaceae*} uncultured bacterium).

Declarations

Ethics

This study was approved by the Institutional Review Board at Lifespan.

Data Availability

Underlying sequencing data can be found at the NCBI Short Read Archive under BioProject ID PRJNA591924. Taxonomic composition data can be found in Supplementary Data 1.

Acknowledgments

We would like to acknowledge the University of Rhode Island Genomics and Sequencing Center for their sequencing expertise.

Funding

This study was funded by the Brown Institute for Brain Science/Norman Prince Neurosciences Institute (RTL) and the National Institutes of Health, through the National Institute of Mental Health awards RF1 MH120830, R01MH101138, R01MH115905, and R21MH112055 (RTL) and the National Institute of General Medical Sciences institutional development award P20GM121344 for the COBRE Center for Antimicrobial Resistance and Therapeutic Discovery at Brown University (PB). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication. Opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the funders. The authors declare no competing interests.

Author Contributions

RTL and PB conceptualized the study. RTL, AES, CMS, and RFLW recruited subjects and collected samples. ADR prepared the samples, analyzed data, and generated figures. ADR and RTL wrote the manuscript. PB edited the manuscript.

Figure 1: Demographic Characteristics of Control and MDD Cohorts
A) Age. B) PROMIS Depression scores. C) Sex assigned at birth. D) Gender Identity. E) Race. F) Ethnicity. G) Number of psychotropic drugs taken. H) Categories of psychotropic drugs taken. *In A-B, bars represent group averages and error bars indicate standard deviation. In H, the sum of all bars in the MDD group may exceed 43 due to subjects taking multiple medications.*

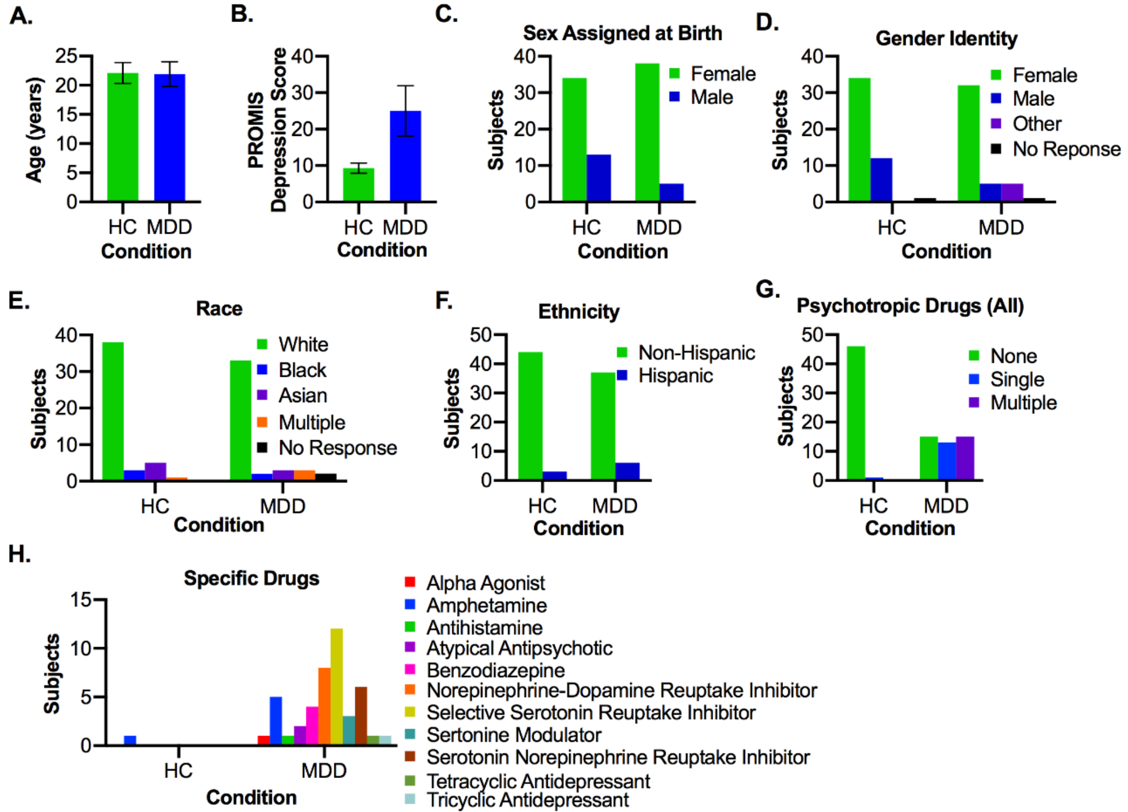


Figure 2: Alpha and Beta Diversity Metrics
A) Faith's Phylogenetic Diversity in control and MDD cohorts (*, *p=0.0313, t-test with Welch's correction*). B) Faith's Phylogenetic Diversity in healthy controls and depressed subjects with mild, moderate, and severe symptoms according to PROMIS depression scores (*, *p=0.0242, ANOVA post-test for linear trend of column means*). C) Faith's Phylogenetic Diversity in depressed subjects taking no psychotropic medication and depressed subjects taking one or more psychotropic drugs (*ns, p=0.613, t-test with Welch's correction*). D) Principal Coordinates 1 and 2 of Bray-Curtis Dissimilarity, with points colored according to their source sample's condition group (*, *p=0.027,* R^2*=0.01, permANOVA*). E) Principal Coordinates 1 and 2 of Unweighted Unifrac Distance, with points colored according to their source sample's condition group (*, *p = 0.041,* R^2*=0.0145, permANOVA*). *In A-C, the central line indicates the group average and error bars indicate standard deviation.*

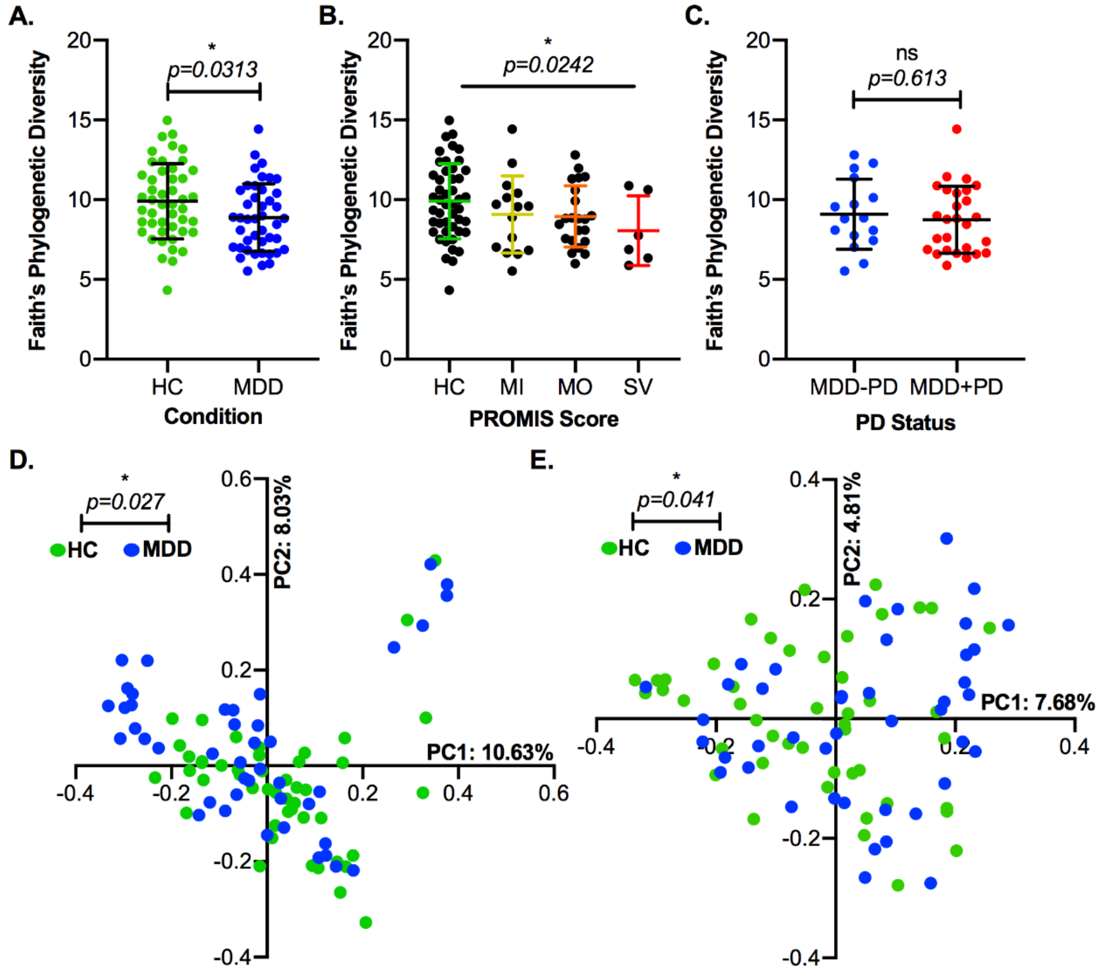
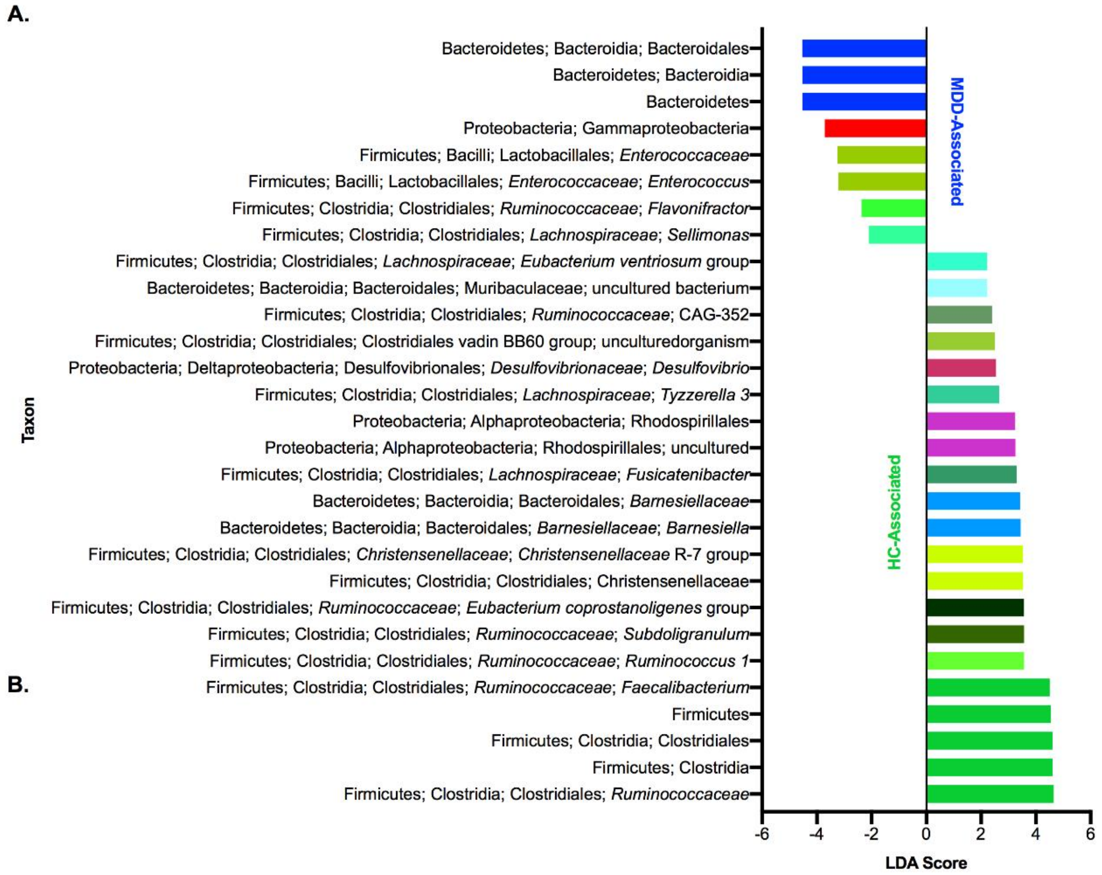


Figure 3: Taxa Associated with Control and Depressed Samples
A) Taxa identified by LEfSe as biomarkers of samples from the control or MDD cohorts (cutoffs were LDA≥2 and p-value≤0.05). B) Cladogram indicating the phylogenetic relatedness of the discriminant taxa.



B.

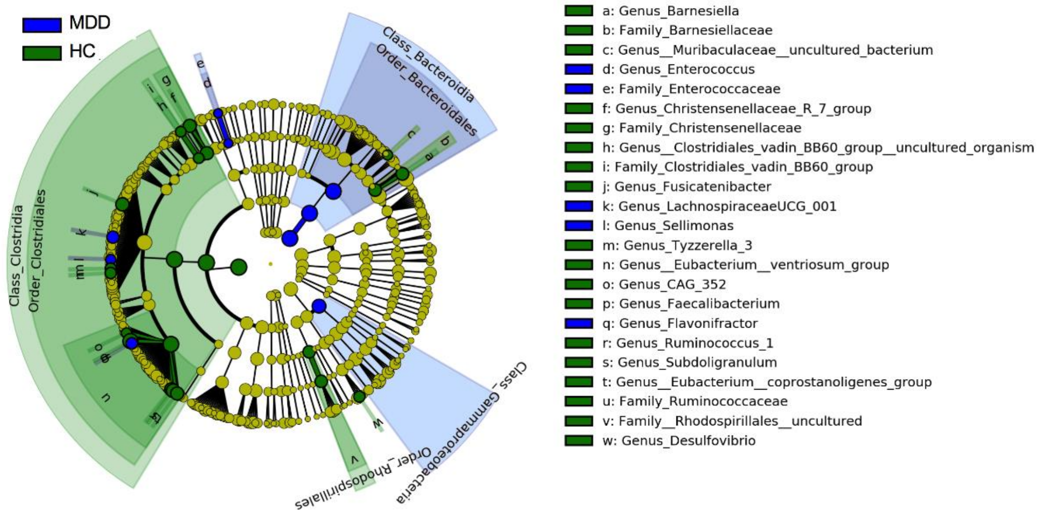


Figure 4: Differences in Relative Abundance of Discriminant Phyla
A) Stacked bar plot indicating the average relative abundance of phyla within the control and MDD cohorts, with discriminant phyla identified by LEfSe highlighted in color. B-C) Phyla that were discriminant of the cohorts, with control and MDD samples interleaved by ranked abundance of each taxon and dotted lines indicating the average relative abundance by group. In A, error bars indicate standard error of the mean. In B-C, the text color indicates the cohort that the phylum was associated with, and p-values are from LEfSe.

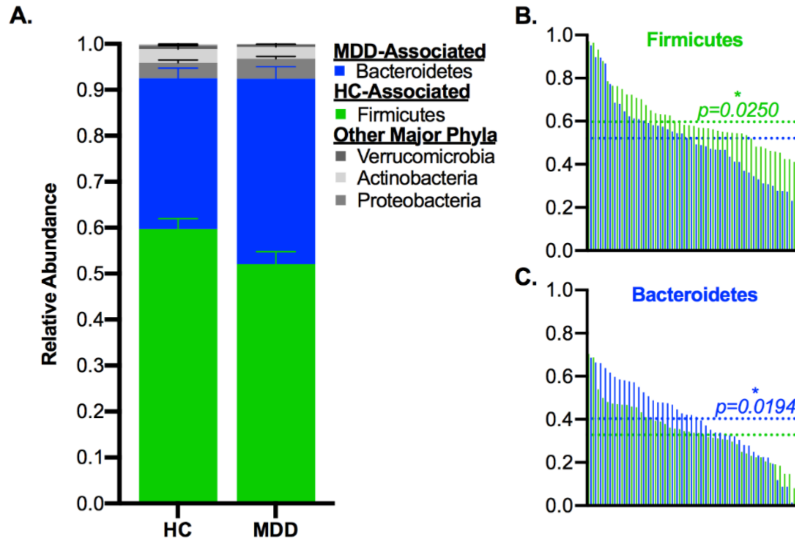


Figure 5: Differences in Relative Abundance of Discriminant Genera

A) Stacked bar plot indicating the average relative abundance of genera within the control and MDD cohorts, with discriminant phyla identified by LEfSe highlighted in color. B-R) Genera that were discriminant of the cohorts, with control and MDD samples interleaved by ranked abundance of each taxon and dotted lines indicating the average relative abundance by group. In A, stacking is done by phylogeny, so all genera are grouped by their higher taxonomic ranks. Error bars indicate standard error of the mean. In B-R, the text color indicates the cohort that the phylum was associated with, and p-values are from LEfSe.

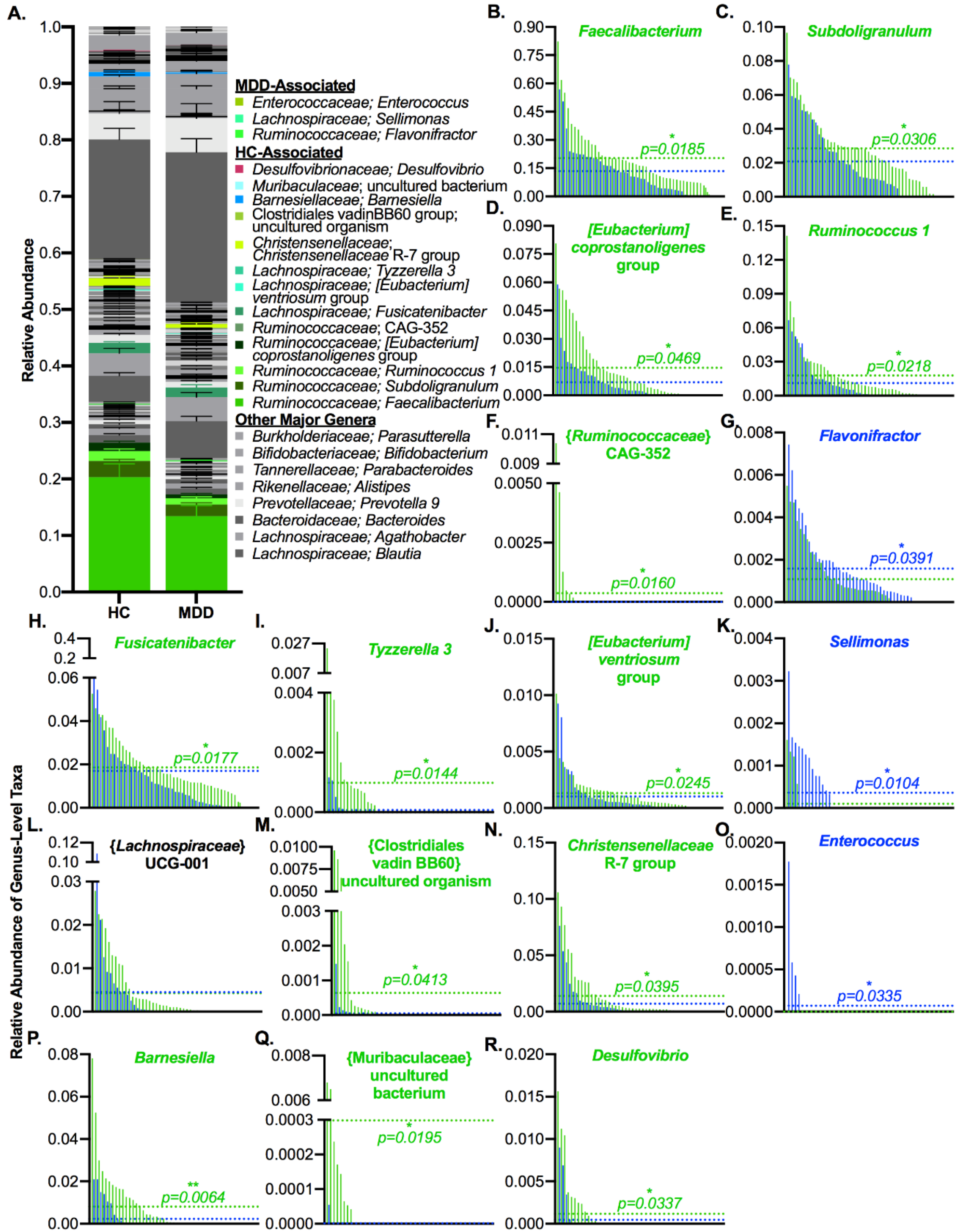


Figure 6: Differences in Relative Abundance of Discriminant Taxa by Symptom Severity

A) Stacked bar plot indicating the average relative abundance of genera within the control subjects and MDD subjects with mild, moderate, or severe symptoms according to PROMIS depression scores, with genera that discriminated the control and MDD cohorts identified by LEfSe highlighted in color. B-AD) Relative abundances of discriminant taxa in the control, MDD-mild symptoms, MDD-moderate symptoms, and MDD-severe symptoms groups. In A, stacking is done by phylogeny, so all genera are grouped by their higher taxonomic ranks. Error bars indicate standard error of the mean. In B-AD, central lines indicate the group mean and error bars indicate standard deviation, p-values are from the ANOVA post-test for linear trend of column means, and arrows indicate the direction of the trend.

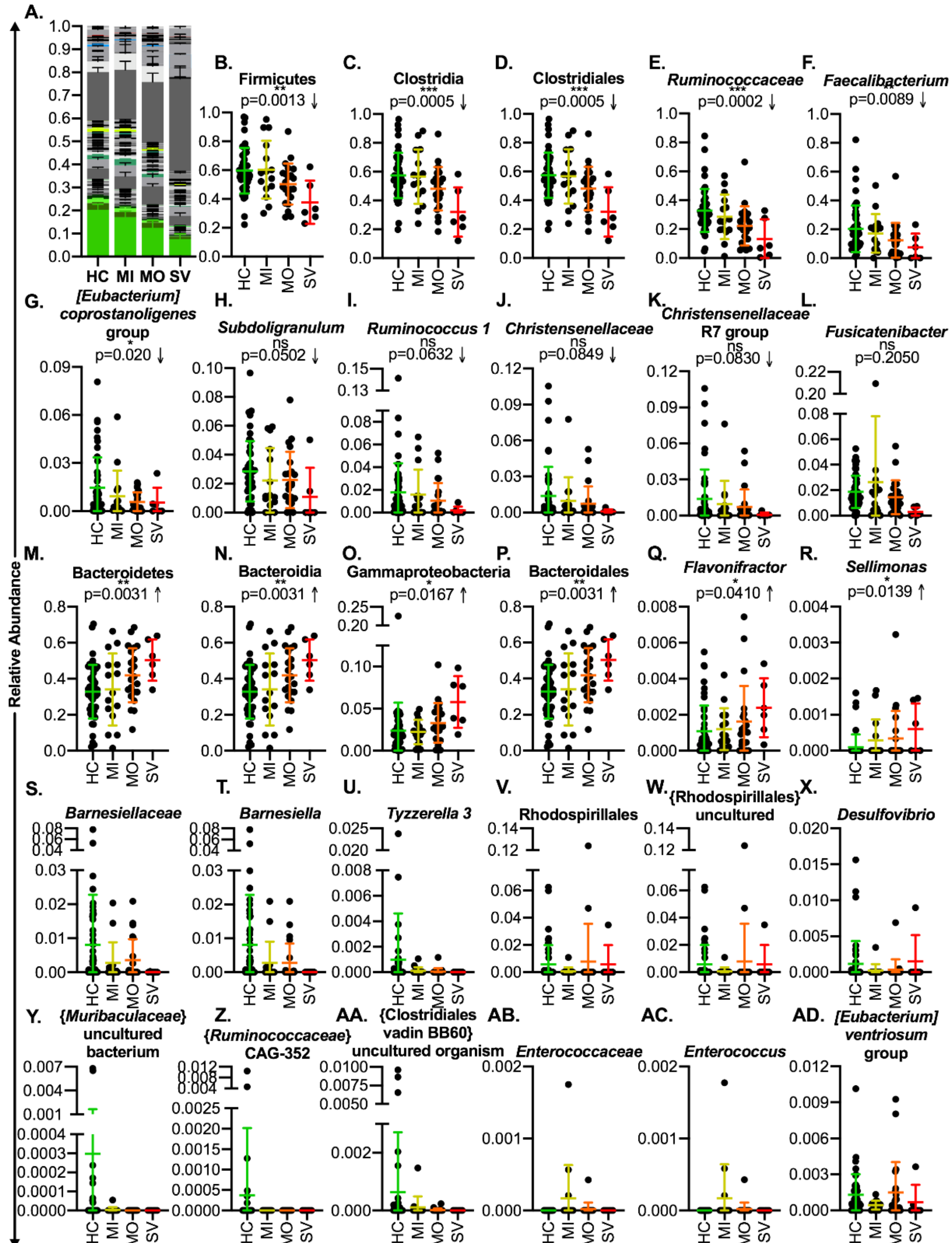


Table 1: Demographic Information

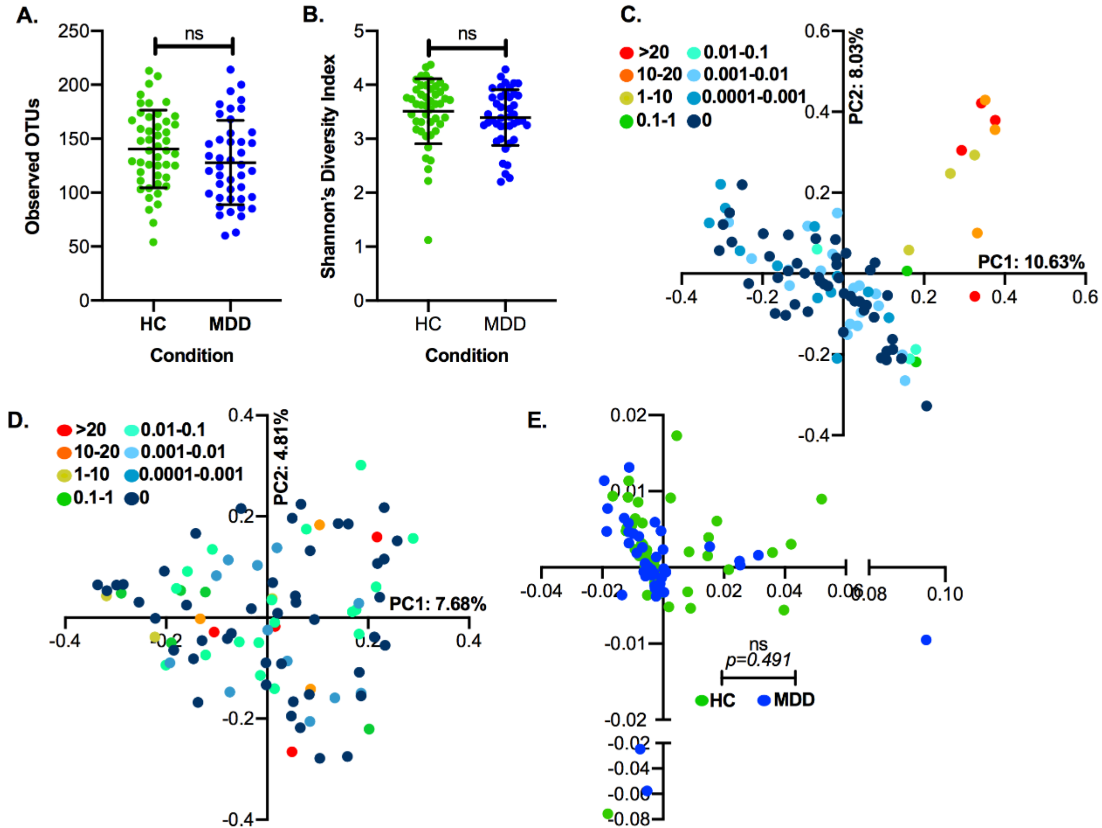
This table shows the demographic information in both the healthy control and MDD cohorts, including sex assigned at birth, gender identity, age, ethnicity, race, PROMIS Depression scores, and prescription of psychoactive drugs. HC = healthy control cohort, MDD = major depressive disorder cohort, SD = standard deviation.

Characteristic	Specific Descriptor	HC	MDD
Sex Assigned at Birth	Female	34 (72.3%)	38 (88.4%)
	Male	13 (27.7%)	5 (11.6%)
Gender Identity	Female	34 (72.3%)	32 (74.4%)
	Male	12 (25.5%)	5 (11.6%)
	Other*	0 (0%)	5 (11.6%)
	No Response	1 (2.1%)	1 (2.3%)
Age (years)	Mean (SD)	22.1 (1.8)	21.9 (2.1)
Ethnicity	Hispanic	3 (6.4%)	6 (14.0%)
	Non-Hispanic	44 (93.3%)	37 (86.0%)
Race	White	38 (80.1%)	33 (76.7%)
	Black	3 (6.4%)	2 (4.7%)
	Asian	5 (10.6%)	3 (7.0%)
	Multiple	1 (2.1%)	3 (7.0%)
	No Response	0 (0%)	2 (4.7%)
PROMIS Depression Score	Mean (SD)	9.3 (1.4)	25.0 (6.9)
Prescribed Psychoactive Drugs	Single	1 (2.1%)	13 (30.0%)
	Multiple	0 (0%)	15 (34.9%)
	No	46 (97.9%)	15 (34.9%)

* “Other” includes: nonbinary, genderqueer, genderfluid, and not sure

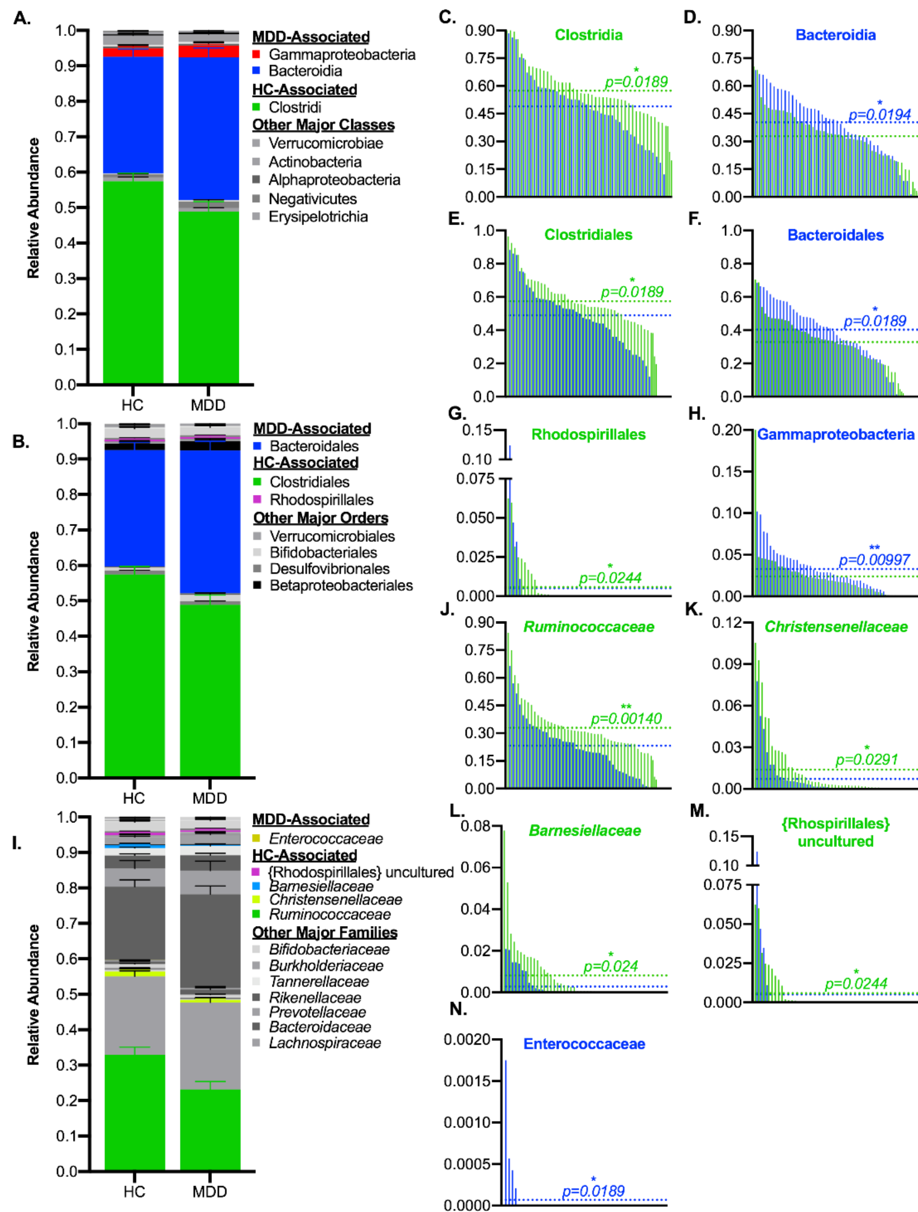
Supplementary Figure 1: Additional Alpha and Beta Diversity Metrics

A) Observed Amplicon Sequence Variants in control and MDD cohorts (ns, p=0.117, t-test with Welch's correction). B) Shannon's Diversity Index in control and MDD cohorts (ns, p=0.323, t-test with Welch's correction). C) Principal Coordinates 1 and 2 of Bray-Curtis Dissimilarity, with points colored according to their source sample's ratio of *Prevotella 9* to *Bacteroides*. D) Principal Coordinates 1 and 2 of Unweighted Unifrac Distance, with points colored according to their source sample's ratio of *Prevotella 9* to *Bacteroides*. E) Principal Coordinates 1 and 2 of Weighted Unifrac Distance, with points colored according to their source sample's condition group (ns, p=0.491, permANOVA).

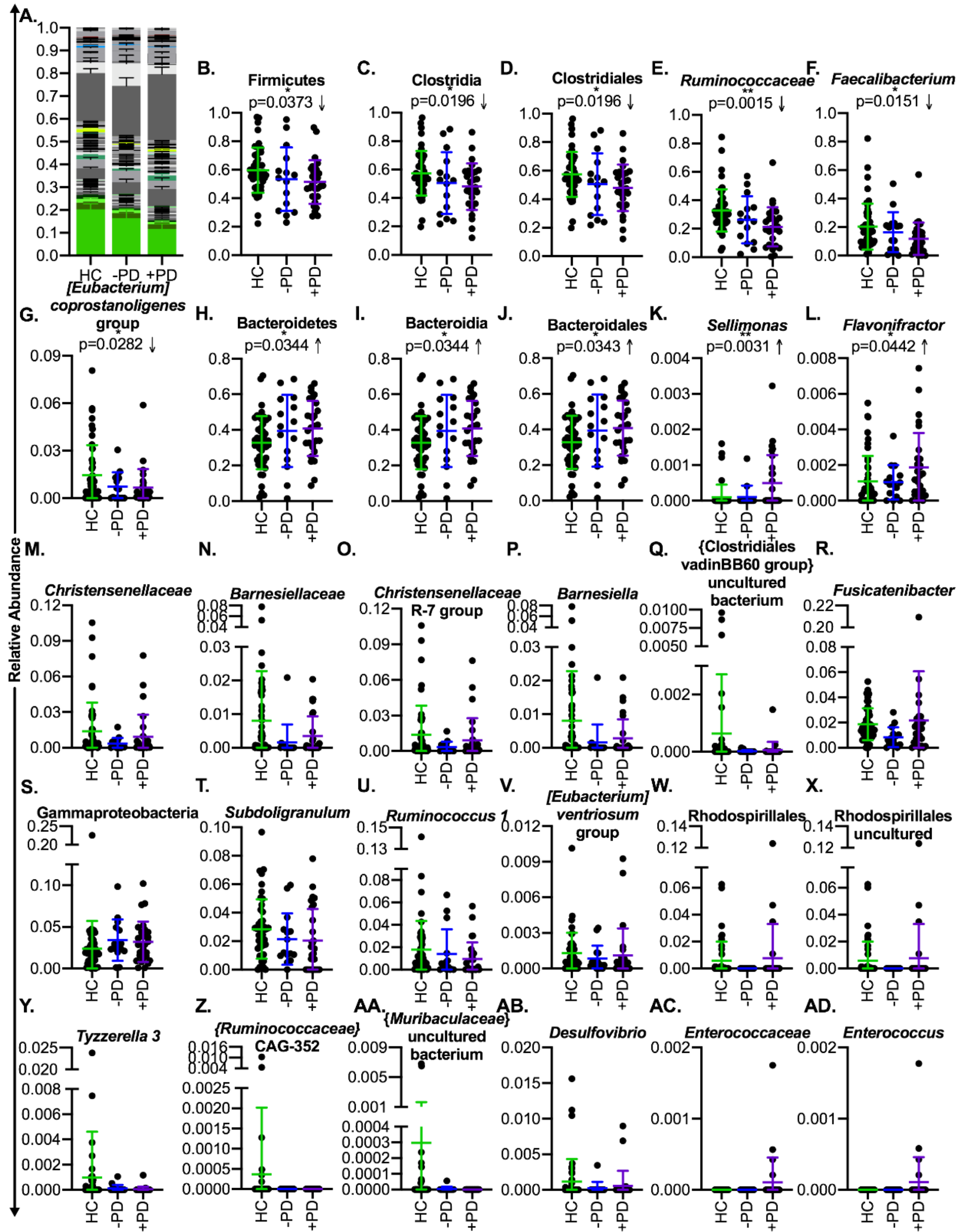


Supplementary Figure 2: Differences in Relative Abundance of Discriminant Classes, Orders, and Families

A) Stacked bar plot indicating the average relative abundance of classes within the control and MDD cohorts, with discriminant phyla identified by LEfSe highlighted in color. B) Stacked bar plot indicating the average relative abundance of orders within the control and MDD cohorts, with discriminant phyla identified by LEfSe highlighted in color. C-H) Classes and orders that were discriminant of the cohorts, with control and MDD samples interleaved by ranked abundance of each taxon and dotted lines indicating the average relative abundance by group. I) Stacked bar plot indicating the average relative abundance of families within the control and MDD cohorts, with discriminant phyla identified by LEfSe highlighted in color. J-N) Families that were discriminant of the cohorts, with control and MDD samples interleaved by ranked abundance of each taxon and dotted lines indicating the average relative abundance by group. *In A-B and I, stacking is done by phylogeny, so all classes, orders, and families are grouped by their higher taxonomic ranks. Error bars indicate standard error of the mean. In C-H and J-N, the text color indicates the cohort that the phylum was associated with, and p-values are from LEfSe.*

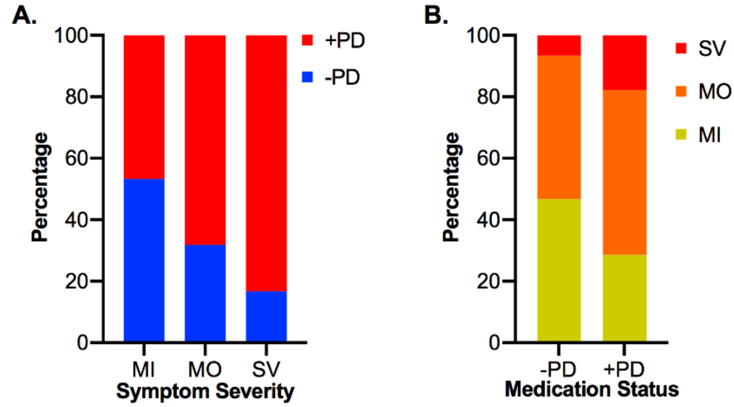


Supplementary Figure 3: Differences in Relative Abundance of Discriminant Taxa by Medication
A) Stacked bar plot indicating the average relative abundance of genera within the control, MDD with no psychotropic drugs, and MDD with psychotropic drugs cohorts, with genera that discriminated the control and MDD cohorts identified by LEfSe highlighted in color. B-AD) Relative abundances of discriminant taxa in the control, MDD with no psychotropic drugs, and MDD with psychotropic drugs. *In A, stacking is done by phylogeny, so all genera are grouped by their higher taxonomic ranks. Error bars indicate standard error of the mean. In B-AD, central lines indicate the group mean and error bars indicate standard deviation, p-values are from the ANOVA post-test for linear trend of column means, and arrows indicate the direction of the trend.*



Supplementary Figure 4: Overlap of Psychotropic Drug Usage and Symptom Severity

A) Stacked bar plot indicating the percentage of subjects with mild, moderate, and severe symptoms according to PROMIS depression scores who are taking or not taking psychotropic medications. B) Stacked bar plot indicating the percentage of subjects taking or not taking psychotropic medicate who have mild, moderate, or severe symptoms according to PROMIS depression scores.



Supplementary Figure 5: MetaCyc Pathways Associated with Control and MDD Cohorts
Taxa identified by LEfSe as biomarkers of samples from the control or MDD cohorts (cutoffs were LDA≥2 and p-value≤0.05). *Colors indicate the general metabolic class the pathways belong to.*

Supplementary Table 1: Psychotropic Medication Total Usage

This table lists the total number of subjects taking each drug class and specific psychotropic medication, and well as the indication for which the medication was being taken.

Drug		Subjects Taking Drug For:					
Type	**Name**	**Depression**	**Anxiety**	**ADHD/LD**	**Sleep**	**Total**	**Type Total**
Alpha Agonist (AA)	Clonidine	0	0	0	1	1	1
Amphetamine (AMP)	Dexamphetamines	0	0	4	0	4	5
	Lisdexamfetamine	0	0	1	0	1	
Antihistamine (AH)	Hydroxyzine	0	1	0	0	1	1
Atypical Antipsychotic (AAP)	Lurasidone	2 (1*)	0	0	0	2	2
Benzodiazepine (BZD)	Alprazolam	0	1	0	0	1	4
	Clonazepam	0	2	0	0	2	
	Lorazepam	0	1	0	0	1	
Norepinephrine-Dopamine Reuptake Inhibitor (NDRI)	Bupropion	6	1	0	0	7	8
	Methylphenidate	0	1	0	0	1	
Selective Serotonin Reuptake Inhibitor (SSRI)	Citalopram	1	0	0	0	1	12
	Escitalopram	0	4	0	0	4	
	Fluoxetine	2	1	0	0	3	
	Sertraline	3	1	0	0	4	
Serotonin Modulator (SM)	Vortioxetine	2	1	0	0	3	3
Serotonin-Norepinephrine Reuptake Inhibitor (SNRI)	Duloxetine	2	0	0	0	2	6
	Venlafaxine	4	0	0	0	4	
Tetracyclic Antidepressant (TetAD)	Mirtazapine	1	0	0	0	1	1
Tricyclic Antidepressant (TriAD)	Clomipramine	0	1	0	0	1	1

*bipolar depression

Supplementary Table 2: Psychotropic Medication Usage Combinations
This table lists the combinations of psychotropic medications being taken by MDD participants in the study, both by drug class, specific medication, and indication for which the medication was being taken.

Drugs Taken	Subjects	Drug Type	Drug Name	Reason for Taking
None	15	N/A	N/A	N/A
Single	13	4 SNRI	1 duloxetine	1 depression
			3 venlafaxine	3 depression
		1 SM	1 vortioxetine	1 depression
		5 SSRI	3 fluoxetine	2 depression
				1 anxiety
			2 escitalopram	2 anxiety
		1 AMP	1 lisdexamfetamine	1 ADHD
		1 NDRI	1 bupropion	1 depression
		1 TriAD	1 clomipramine	1 anxiety
Multiple	15	1 AA + TetAD	1 clonidine + mirtazapine	1 sleep/depression
		1 AAP + AMP	1 lurasidone + adderall	1 depression*/ADHD
		1 AMP + SM	1 adderall + vortioxetine	1 ADHD/depression
		1 AAP + SNRI	1 lurasidone + duloxetine	1 depression/depression
		1 AMP + NDRI	1 adderall + bupropion	1 ADHD/depression
		1 BZD + SM	1 lorazepam + vortioxetine	1 anxiety/anxiety
		1 BZD + SNRI	1 clonazepam + duloxetine	1 depression/anxiety
		2 BZD + SSRI	1 alprazolam + sertraline	1 anxiety/depression
			1 citalopram + clonazepam	1 depression/anxiety
		1 NDRI + SNRI	1 bupropion + venlafaxine	1 depression/depression
		3 NDRI + SSRI	2 bupropion + sertraline	1 depression/depression
				1 anxiety/depression
			1 methylphenidate + sertraline	1 anxiety/anxiety
		1 NDRI+SSRI	1 bupropion + escitalopram	1 depression/anxiety
		1 NDRI+SSRI+AH	1 bupropion + escitalopram + hydroxyzine	1 depression/anxiety/anxiety

*bipolar depression

Supplementary Table 3: Table of Discriminant Taxa

This table lists all taxa that are associated with either the control or MDD cohorts, with their p-values from LEfSe. Where relevant, the p-value for the ANOVA post-test for linear trend of column means for symptom severity (SV) and/or psychotropic drugs (PD) is also included. *Taxa colored blue are associated with the MDD cohort, while taxa colored green are associated with controls.*

Phylum	Class	Order	Family	Genus
Bacteroidetes MDD: *, 0.0194 SV: **, 0.0031 PD: *, 0.0344	**Bacteroidia** MDD: *, 0.0194 SV: **, 0.0031 PD: *, 0.0344	**Bacteroidales** MDD: *, 0.0194 SV: **, 0.0031 PD: *, 0.0343	***Barnesiellaceae*** MDD: *, 0.024	***Barnesiella*** MDD: **, 0.0064
			Muribaculaceae	**Uncultured bacterium** MDD: *, 0.0195
Firmicutes MDD: *, 0.0250 SV: **, 0.0013 PD: *, 0.0373	**Clostridia** MDD: *, 0.0189 SV: ***, 0.0005 PD: *, 0.0196	**Clostridiales** MDD: *, 0.0189 SV: ***, 0.0005 PD: *, 0.0196	***Ruminococcaceae*** MDD: **, 0.00140 SV: ***, 0.0002 PD: **, 0.0015	***Faecalibacterium*** MDD: *, 0.0185 SV: **, 0.00849 PD: *, 0.0151
				Subdoligranulum MDD: *, 0.0306 SV: ns, 0.0520 PD: ns, 0.0957
				Ruminococcus 1 MDD: *, 0.0218 SV: ns, 0.0631 PD: ns, 0.1198
				[*Eubacterium*] *coprostanoligenes* group MDD: *, 0.0469 SV: *, 0.0200 PD: *, 0.0282
				CAG-352 MDD: *, 0.0160
				Flavonifractor MDD: *, 0.0391 SV: *, 0.041 PD: *, 0.0422
			Christensenellaceae MDD: *, 0.0291 SV: ns, 0.049	**R-7 group** MDD: *, 0.0395 SV: ns, 0.0830
			Lachnospiraceae	***Fusicatenibacter*** MDD: *, 0.0177 SV: ns, 0.2050 PD: ns, 0.7239
				Tyzzerella 3 MDD: *, 0.0144
				[*Eubacterium*] *ventriosum* group MDD: *, 0.0245
				Sellimonas MDD: *, 0.0104 SV: *, 0.0139 PD: **, 0.0031
			Clostridiales vadin BB60 group	**Uncultured organism** MDD: *, 0.0413
	Bacilli	Lactobacillales	***Enterococcaceae*** MDD: *, 0.0335	***Enterococcus*** MDD: *, 0.0335
Proteobacteria	Alphaproteobacteria	**Rhodospiralles** MDD: *, 0.0244	**Uncultured** MDD: *, 0.0244	
	Deltaproteobacteria	Desulfovibrionales	*Desulfovibrionaceae*	*Desulfovibrio* MDD: *, 0.0337
	Gammaproteobacteria MDD: **, 0.00997 SV: *, 0.0167 PD: ns, 0.2121			

Supplementary Data 1: Taxonomic Classifications from 16S rRNA Sequencing with QIIME2
This file contains the relative abundances of taxa in all samples at the phylum (tab 1), class (tab 2), order (tab 3), family (tab 4), and genus (tab 5) levels.
This file is available at the Brown Digital Repository at https://doi.org/10.26300/cvpj-xw35.

Supplementary Data 2: Predicted MetaCyc Pathways from 16S rRNA Sequencing with picrust2 Analysis
This file includes the MetaCyc pathway abundances predicted from the 16S rRNA sequencing data using picrust2 within QIIME2.
This file is available at the Brown Digital Repository at https://doi.org/10.26300/9nx7-8d78.

All Supplementary Data files for this thesis can be found at the Brown Digital Repository at https://doi.org/10.26300/enej-vt18.

CHAPTER THREE: ANTIMICROBIAL RESISTANCE GENE PREVALENCE IN A POPULATION OF PATIENTS WITH ADVANCED DEMENTIA IS RELATED TO SPECIFIC PATHOBIONTS

Adapted from

"Antimicrobial resistance gene prevalence in a population of patients with advanced dementia is related to specific pathobionts"

By Aislinn D. Rowan-Nash, Rafael Araos, Erika M.C. D'Agata, and Peter Belenky

Manuscript Submitted

CONTENTS

Abstract 122

Introduction 123

Results 126

Discussion 137

Conclusions 143

Methods 144

Main Figures and Tables 153

Supplementary Materials 158

Antimicrobial resistance gene prevalence in a population of patients with advanced dementia is related to specific pathobionts

Aislinn D. Rowan-Nash[1], Rafael Araos[2,3,4], Erika M.C. D'Agata[5], & Peter Belenky[1#]

[1] Department of Molecular Microbiology and Immunology, Brown University, Providence RI, USA

[2] Instituto de Ciencias e Innovación en Medicina (ICIM), Facultad de Medicina Clinica Alemana Universidad del Desarrollo, Chile

[3] Millenium Nucleus for Collaborative Research on Bacterial Resistance (MICROB-R), Chile

[4] Advanced Center for Chronic Diseases (ACCDiS), Facultad de Medicina Clinica Alemana Univarsidad del Desarrollo, Santiago Chile.

[5] Infectious Diseases Division, Rhode Island Hospital, Warren Alpert Medical School of Brown University, Providence, RI, USA

[#] Corresponding author: Peter Belenky, peter_belenky@brown.edu

Abstract

The issue of antimicrobial resistance continues to grow worldwide, and long-term care facilities are significant reservoirs of antimicrobial-resistant organisms, in part due to high frequency of antimicrobial use. Patients with advanced dementia are particularly vulnerable to multidrug-resistant organism acquisition and antimicrobial overuse, which has negative consequences for the gut microbiome and can contribute to the selection and propagation of antimicrobial resistance genes. In this study, we longitudinally examined a group of advanced dementia patients treated with the fluoroquinolone antimicrobial levofloxacin. We observed significant inter- and intra-subject heterogeneity in the composition of the microbiota of the longitudinal levofloxacin cohort, suggesting temporal instability. Within this dataset, we did not find significant impacts of levofloxacin on the diversity, composition, function, or resistome of the gut microbiota of this population. However, we were able to link the antimicrobial resistance gene burden in a sample with the relative abundance of several pathobionts – particularly *Escherichia coli*, *Proteus mirabilis*, and *Enterococcus faecalis*, as well as less-prevalent species including *Providencia stuartii* and *Staphylococcus haemolyticus*. Furthermore, we used metagenomic assembly and binning to demonstrate that these species had higher genomic resistance gene levels than common gut commensals, and we were able to predict antimicrobial resistance gene burden from the relative abundances of these species in a separate, larger cohort from the same population.Given the high frequency with which these species were found at high levels in this population and the underlying vulnerability to infection with multidrug resistant organisms of advanced dementia patients, attention to microbial blooms of these species may be warranted. Additionally, in this study, we were

able to utilize genomic assembly from metagenomic data to more definitively associate antimicrobial resistance gene levels with specific assembled species; as this technology continues to develop, assembly could prove to be a valuable method to monitor both specific resistance genes and blooms of multidrug-resistant organisms.

Introduction

It is well-recognized that there is a growing threat of antimicrobial-resistant (AMR) bacterial strains that threaten the health and lives of millions worldwide. In the United States alone, the Centers for Disease Control and Prevention estimates that at least 2 million people get an AMR infection each year, and at least 23,000 die as a result[519]. A number of factors have driven the rise in AMR bacteria worldwide, including overprescription of antibiotics in the healthcare setting, over-the-counter access to antibiotics in some countries, and widespread use of antibiotics in animal husbandry for non-veterinary purposes[520-522]. Concerningly, hospitals and other medical institutions are frequent sites of AMR bacteria acquisition, where patients may already be ill or immunocompromised, antimicrobial use is common, and patient-to-patient transmission of AMR isolates can occur via inadequate hygiene or environmental contamination[523-526]. For example, AMR bacteria are highly prevalent in nursing homes, with estimates that over 35% of nursing home residents are colonized with multidrug resistant organisms (MDROs)[527-532]. This is particularly problematic in light of the fact that elderly patients in long-term care facilities may be frequently hospitalized, potentially serving as a mode of bidirectional transport of MDROs between healthcare facilities[533-535]. They are also prone to infections and are frequently treated with antimicrobials[536-538], which has long been associated with acquisition of MDROs and may not always be indicated[530,535,538-546].

The problem of MDRO prevalence and inappropriate antimicrobial use is of particular relevance in elderly subjects with advanced dementia, a population which receives extensive antimicrobial treatment which becomes more frequent closer to death, calling its benefit and effectiveness into question[546,547]. Accordingly, advanced dementia specifically has been shown to be a risk factor of MDRO colonization[531,548]. To examine this issue, the Study of Pathogen Resistance and Exposure to Antimicrobials in Dementia (SPREAD) was undertaken from 2009-2012 in order to analyze MDRO acquisition and appropriateness of antimicrobial prescription in elderly adults with advanced dementia residing in nursing homes[549]. Supporting the widespread nature of MDRO carriage in this population, analysis of SPREAD subjects revealed that there were significant baseline levels and new acquisitions of MDROs, and that there was notable spread of MDRO strains within and even between nursing home facilities[546,550].

In addition to potential facilitation of MDRO acquisition or spread, antimicrobial overuse may also have negative impacts on the diversity, composition, or function of the gut microbiota, which may already be vulnerable in elderly populations. Healthy younger adults tend to have a fecal microbiome characterized by relatively high diversity of species and populated primarily by members of the phyla *Bacteroidetes* and *Firmicutes*, largely obligate anaerobes which exist in homeostasis with the intestinal epithelium[1,27,33,551,552]. However, it has been found that during senescence, the gut tends to have higher levels of *Bacteroidetes* and *Proteobacteria* and harbors higher levels of facultative aerobes and potential pathobionts, including *Enterobacterales* such as *E. coli* [464,551,553-559]. These changes become more pronounced as aging progresses, and several studies have indicated that age-related alterations to the gut microbiota are relatively minor in septuagenarians,

but become more pronounced over time and are clear in centenarians and supercentenarians[464,506,557,560,561]. This is likely due to a number of factors, including the decline of immune function, onset of age-related diseases (including metabolic disorders), changes to diet and mobility, and the increased likelihood of medication utilization and/or hospitalization[556,562]. However, lifestyle of elderly adults has an important impact, as research suggests that community-resident elderly subjects have a distinct and more diverse microbiome compared with those of their hospitalized or institutionalized peers, which was suggested to be at least in part due to nutritional differences[562,563]. Furthermore, reduced microbiome diversity has been associated with increased frailty of elderly subjects[562,564]. Accordingly, given that the microbiomes of institutionalized elderly patients are perhaps already at risk, understanding the impacts of antimicrobial use and MDRO acquisition on this population is of importance.

We analyzed the gut microbiomes of eleven subjects from SPREAD to examine the impact of antimicrobial use on the gut microbiota composition, function, and antimicrobial resistance gene (ARG) profile of elderly dementia patients. These subjects were chosen as they were the largest cohort who had received a single antimicrobial (levofloxacin) during the collection period, and we anticipated that this intervention could have an impact on the already-vulnerable microbiota of this elderly, institutionalized cohort. Levofloxacin is an antimicrobial of the fluoroquinolone class with high oral bioavailability[565-567] which has been found to reduce levels of Gram-negative aerobic bacteria – including *Proteobacteria* and particularly *Enterobacterales* – in the fecal microbiota[568-574], although fluoroquinolone resistance among this taxon has been spreading[575-582]. A maximum of five rectal swab samples, collected every three months, were taken from each subject, and both 16S rRNA

amplicon and shotgun metagenomics sequencing were performed. We analyzed alpha and beta diversity, taxonomic composition, functional potential, and antimicrobial resistance gene profiles before and after administration of levofloxacin, but were unable to detect specific impact of levofloxacin on any of these measures. However, we did find an association between blooms of particular enteric species and ARG burden, including in samples where MDRO were not detected by culture, suggesting that certain pathobionts carrying high ARG burdens may frequently colonize this population and that metagenomics may allow detection of resistant bacteria not flagged by culture-based methods.

Results

Overview of Subjects

Elderly patients in long-term care facilities, and particularly patients with advanced dementia, are frequently exposed to antimicrobials and are at high-risk of acquisition and carriage of MDRO[527-531,536-538,545-548,550]. From within the SPREAD cohort, we selected the largest group of subjects who had been administered a single antimicrobial during their participation in the study. This gave us a group of eleven subjects who had been given the fluoroquinolone levofloxacin, one of the most commonly prescribed antimicrobials. We analyzed up to five rectal swabs, taken every three months over the course of a year, from these eleven subjects in the SPREAD cohort[549], using both 16S rRNA and shotgun metagenomics sequencing (Figure 1A). During their participation in the study, these subjects had received only a single course of levofloxacin (average course of eight days), which has previously been shown to decrease the proportion of the *Enterobacterales* order of *Proteobacteria*[568-574]. Of the eleven subjects, all but Subject I were female and all but

Subject G were white. They ranged in age from 72 to 101, and six members of the cohort did not survive for the full year of the study (Supplementary Table 1). All but two subjects (C and G) resided in different nursing homes. Overall, there were 38 samples for metagenomics sequencing (Supplementary Table 2). Culture-based methods indicated that four of the eleven subjects acquired a MDRO during the study: Subject A acquired methicillin-resistant *S. aureus* (MRSA) at the 12-month timepoint, Subject B acquired multidrug-resistant *E. coli* at the 3-month timepoint, and Subjects C and D both acquired multidrug-resistant *P. mirabilis* at the 3-month timepoint (Supplementary Table 1).

Alpha and Beta Diversity Metrics

Before focusing on antimicrobial resistance, we first wanted to assess the composition of the community throughout the longitudinal timeframe. We initially used the metagenomic sequencing data to compare the alpha diversity, or the diversity within each sample, of samples collected before and after levofloxacin administration. According to Shannon's Diversity Index, which incorporates both richness and evenness of samples, there was no significant difference between the pre- and post-levofloxacin samples (Figure 1B). Furthermore, the alpha diversity was variable over time even within the same subject, and there was no clear trend of recovery in alpha diversity after antibiotic cessation. This suggests a degree of temporal instability, in which the richness and/or evenness of the samples varies changes over time.

We then examined beta diversity, or the diversity between samples. We utilized the Bray-Curtis Dissimilarity metric, which considers the identity and abundance of taxa shared between samples. Plotting this metric in a principal coordinate analysis (PCoA) revealed no apparent pattern of clustering based on either subject or sample collection point

relative to levofloxacin, and in fact, samples from the same subject were often located quite distantly from one another (Figure 1C). We then compared the within-subjects dissimilarity of sequential samples within a subject when both were pre-levofloxacin, both were post-levofloxacin, or one sample was pre- and one was post-levofloxacin; there was no significant difference between any of the groups (Figure 1D), suggesting that levofloxacin was not associated with community disruption. Furthermore, while within-subject dissimilarity was lower than between-subjects dissimilarity, the effect size was low (0.7013 vs. 0.7712, respectively; Figure 1E).

Taxonomic Composition

We utilized Kraken2 in conjunction the with Bayesian Reestimation of Abundance with KrakEN2 (Bracken2) pipeline to assign taxonomy to our metagenomic sequencing samples[583,584]. Corresponding to the high between-subjects beta-diversity, the taxonomic composition of the gut microbiome varied significantly between subjects. As is typical for the human gut microbiome, most bacteria belonged to the five major phyla of *Firmicutes*, *Bacteroidetes*, *Proteobacteria*, *Actinobacteria*, and *Verrucomicrobia*. However, consistent with the high within-subjects beta diversity, the dominant phyla varied greatly even between samples from the same subject (Supplementary Figure 1); for example, the most abundant phylum in Subject E was *Bacteroidetes* at two timepoints, *Proteobacteria* at two timepoints, and *Firmicutes* at one (Supplementary Figure 1F). Overall, the most abundant phylum was *Actinobacteria* in three samples, *Bacteroidetes* in seventeen samples, *Firmicutes* in seven samples, and *Proteobacteria* in eleven samples (Supplementary Figure 1A-L); averaging across all samples, *Bacteroidetes* was highest at 34.2%, followed by *Proteobacteria* (26.9%), *Firmicutes* (23.3%), and *Actinobacteria* (11.2%) (Supplementary

Figure 1A). Qualitatively, many of the samples from this population represent highly divergent and dysbiotic microbiomes compared with what is typically seen with younger subjects, in which *Proteobacteria* in particular make up a much smaller proportion of the microbiome than in these elderly dementia subjects[27].

The genus- and species-level taxonomic composition was also variable. Blooms of potential pathogens[585], including *Campylobacter ureolyticus*[586], *Corynebacterium urealyticum*[587], *Enterococcus faecalis*[588,589], *Escherichia coli* [590,591], *Oligella urethralis*[592-595], *Proteus mirabilis*[596,597], *Providencia stuartii*[598,599], *Pseudomonas aeruginosa*[600,601], *Staphylococcus aureus*[602-604], and *Staphylococcus haemolyticus*[605-607], were fairly common, both before and after levofloxacin administration (Figure 2A, Supplementary Figure 2). Across subjects, even baseline samples varied in composition, as expected from beta-diversity analysis. Averaging across all samples, the single most-abundant species was *E. coli*, further supporting the qualitatively dysbiotic nature of the gut microbiome of this cohort (Figure 2A). Despite the high proportion of members of *Enterobacterales* in this cohort, Linear Discriminant Analysis Effect Size (LEfSe) analysis[518] did not reveal biomarkers for pre- or post-levofloxacin samples at the phylum, genus, or species level. Full data on taxonomic composition at the phylum and species levels can be found in Supplementary Data 1.

As we had access to full 16S rRNA and shotgun metagenomics data for our samples, we compared their taxonomic identifications at the genus level. The two methods of analysis were generally consistent, and blooms of prominent genera (including *Escherichia, Proteus, Enterococcus, Providencia, Staphylococcus*, and *Bacteroides*) were generally detected by both analysis pipelines (Supplementary Figure 3A). Metagenomics

analysis was unsurprisingly able to detect more distinct genera, and of the genera that were called by both pipelines, LEfSe analysis revealed biases in both methods. For example, metagenomics analysis by Kraken2 and Bracken2 detected higher levels of *Bacteroides*, while 16S rRNA analysis with Quantitative Insights Into Microbial Ecology 2 (QIIME2)[5] detected higher levels of *Ruminiclostridium*. Full data on taxonomic abundances at the genus level can be found in Supplementary Data 1 for metagenomics and Supplementary Data 2 for 16S rRNA.

Functional Potential

We used the Human Microbiome Project Unified Metabolic Analysis Network 2 (HUMAnN2) pipeline[608] to analyze the genetic content of the metagenomic samples. We utilized LEfSe to compare community function at the Kyoto Encyclopedia of Genes and Genomes (KEGG) ortholog, Gene Ontology (GO) term, and MetaCyc pathway levels. As in the taxonomic analysis, there were no significant biomarkers of either pre-or post-levofloxacin administration samples. However, while the taxonomic profile of the samples varied greatly, the functional capacity of the samples was fairly consistent across samples (Supplementary Figure 4). Full data on functional potential can be found in Supplementary Data 3.

Antimicrobial Resistance Gene Profile

We used the DeepARG machine-learning program[609] to detect resistance genes in the metagenomic samples. Across all samples, the most abundant class of ARG was "multidrug", followed by "macrolide-lincosamide-streptogramin" (MLS), and "tetracycline". The most common specific gene detected was the multidrug resistance *rpoB2* variant of the RNA polymerase beta subunit, followed by the MLS resistance gene

macB and a multidrug ABC transporter (Figure 2B). LEfSe analysis revealed no ARG biomarkers of either pre- or post-levofloxacin samples. Full data on ARG composition can be found in Supplementary Data 4.

However, we were able to detect changes in specific ARG classes and genes that corresponded with the detection of antimicrobial-resistant organisms in two subjects. Subject A acquired MRSA at the 12-month timepoint, and a bloom of this species to 25.0% could be detected in the metagenomic taxonomic data (Figure 3A, Supplementary Figure 2B). While the overall level of ARGs did not notably increase at this timepoint, there was a clear expansion in beta-lactam resistance genes (Figures 3B, Supplementary Figure 5B), including the *mecA*/*mecR1*/*mecI* operon, which regulates expression of the low-affinity penicillin-binding protein *mecA* (PBP-2A)[610-613] (Figure 3C). This operon is characteristic of MRSA strains[610-613], supporting the culture-based classification of this *S. aureus* isolate as MRSA.

Similarly, Subject B acquired multidrug-resistant *E. coli* (resistant to the beta-lactams ampicillin/sulbactam, cefazolin, ceftazidime, and ceftriaxone and to the fluoroquinolone ciprofloxacin) at the three-month timepoint, and the proportion of this species expanded to 47.3% of the population in the corresponding sample (Figure 3D, Supplementary Figure 2C). Accordingly, this sample showed a notable increase in the relative abundance of ARGs, which was in large part driven by an increase in a number of multidrug resistance genes (Figure 3E); there was also a clear increase in several beta-lactam resistance genes, including the low-affinity penicillin-binding protein genes *PBP-1A, PBP-1B,* and *penA* (*PBP2*) as well as class C beta-lactamase genes[614-619], and several

fluoroquinolone resistance genes, including the transporters *patA* and *mdtK* [620-624] (Figure 3F-G).

Despite the acquisition of multidrug-resistant *P. mirabilis* at the three-month timepoint in Subjects C and D, there was no corresponding increase in ARGs. ARG levels stayed approximately the same in Subject C (0.372% at baseline and 0.384% at three months) and decreased in Subject D from 0.482% at baseline to 0.364% at the three-month timepoint (Figures 2B, Supplementary Figure 5D-E). However, this corresponds to the taxonomic data; levels of *P. mirabilis* were low and stable in Subject C (0.55% at baseline and 0.61% three months later), and while *P. mirabilis* made up 13.8% of the population at baseline in Subject D, it underwent a reduction to 2.3% at the three-month timepoint (Figures 2A, Supplementary Figure 2D-E). Taken together, this data indicates that our metagenomics pipeline can detect blooms of AMR pathogens and that the corresponding change in ARG levels aligns with culture-based detection of MDROs. At the same time, metagenomic analysis of some samples found blooms of pathogens and ARGs that were not associated with culture-based MDRO detection.

Attribution of ARG Density to Specific Species

While total ARG density within samples did not vary by levofloxacin administration, there was significant variability between samples. In fact, most samples had similar baseline levels of ARGs of 0.3% to 0.4% of the total reads, while only a few samples rose above this value to between 0.6% and 0.8%. Close inspection of the taxonomic composition of the samples revealed that samples with higher levels of ARGs tended to have blooms of one or more of the *Proteobacteria* species *E. coli* and *P. mirabilis* and the *Firmicutes* species *E. faecalis*, strains of which are common pathobionts[625-631]

(Figure 4A). Confirming this association, correlation analysis between ARG levels and the sum of the relative abundances of these three species showed a strong and significant positive correlation ($r = 0.791$, $R^2 = 0.6254$, $p < 0.0001$, Pearson's correlation; Figure 4B). This suggests that in samples with higher-than-baseline ARG levels, ARG abundance is being driven by high relative abundance of these three species.

However, there were two notable exceptions: Samples E9 and H6 had high levels of ARGs without corresponding blooms of these three species. However, *P. stuartii* bloomed to 41.9% relative abundance in Sample E9 and *S. haemolyticus* bloomed to 36.9% in Sample H6 (Figure 2A, Supplementary Figure 2F&I). Both species have long been associated with AMR phenotypes[606,607,632-638] and were not found at high levels in other samples, but could explain the higher ARG abundance in these samples (Figure 4A). Supporting this possibility, an examination of the ARGs in Sample H6 showed a distinct profile relative to other samples, with high levels of staphylococcal resistance genes including fluoroquinolone resistance gene *norB* and macrolide-streptogramin resistance gene *msrA* [639-642](Figures 2B, 3I-L). Accordingly, addition of *P. stuartii* and *S. haemolyticus* abundances to the analysis resulted in a stronger correlation ($r = 0.933$, $R^2 = 0.8706$, $p < 0.0001$, Pearson's correlation; Figure 4C).

To more rigorously examine the relationship between the species of interest and ARG levels, we performed metagenomic assembly and binning to compare the levels of ARGs in these organisms to levels in other common and abundant species, including likely commensals and potential pathogens (Figure 4D). Specifically, we analyzed bins that passed various quality controls (Supplementary Table 3) and corresponded to species

identified by Kraken2/Bracken2 to make up greater than 0.1% of their source samples (Supplementary Table 4).

As anticipated, we found that the levels of ARGs in bins from *E. coli* and *P. mirabilis* were consistently high compared to other species analyzed. In fact, *E. coli* had the highest average ARG density of any species analyzed, while *P. mirabilis* was the fifth-highest. Notably, the ARG composition of the bins of these species from samples in which MDROs were detected (B3, C3, and D3) did not appear to be different from those of other samples (Supplementary Figure 6A-B), although it is possible that some resistance genes were carried on plasmids that were not assembled into genomes. *P. stuartii* had the second-highest average ARG density, reflecting the expansion of ARGs detected in sample E9, where this species bloomed to 41.9% of the population. The third and fourth positions were taken by the single bins constructed for *Klebsiella oxytoca* and *Morganella morganii*, other *Proteobacteria* with pathogenic potential[643-645]. *P. aeruginosa* bins rounded out the top six, with similar levels to the other top species. However, as *K. oxytoca* and *M. morganii* were never present at greater than 3% and *P. aeruginosa* bloomed in only two samples, they did not significantly contribute to overall ARG density in the cohort. Importantly, high ARG density was not a universal feature of *Proteobacteria*, or even of pathogenic *Proteobacteria*; bins constructed for the *Campylobacter* species *C. hominis* and *C. ureolyticus* had universally low ARG levels. Additionally, while we could not construct a high-quality bin for *O. urethralis*, the low ARG densities in the samples in which this species bloomed (C0 and C3) suggests that it also has low genomic ARG content. This suggests that high ARG density among the *Proteobacteria* analyzed was restricted to the

Gammaproteobacteria class, primarily of the order *Enterobacterales* but also including *Pseudomonadales*.

We were only able to construct two good-quality bins for *E. faecalis*, which varied in their ARG levels, particularly on the basis of bacitracin resistance. On average, while the two bins did not have ARG levels as high as the *Proteobacteria* of interest, they did rank among the highest of the *Firmicutes* bins tested. We were also able to create a single bin for *S. haemolyticus* from Sample H6 in which it made up 36.9% of the population. This bin had an ARG density higher than the average for any other non-*Proteobacteria* species, supporting its role in the high ARG levels found in the corresponding sample. As expected from the analysis of the total ARG population of that sample (Figure 3G), the staphylococcal resistance genes *norB* and *msrA* were found in this bin. We were also able to create two bins for *S. aureus*, including from sample A12 where MRSA was detected. The A12 bin contained the characteristic MRSA gene *mecA* while the H6 bin did not, suggesting that the *S. aureus* strain found in H6 was not MRSA (Supplementary Figure 6C). In general, bins from the phyla *Actinobacteria* (including *Bifidobacterium* and *Corynebacterium* species) and *Bacteroidetes* (including *Bacteroides* and *Parabacteroides* species) had low ARG levels. Full data on the ARGs and classes found in species-level bins can be found in Supplementary Data 4.

Prediction of ARG Density from Species Abundances

Our initial analysis only considered the eleven subjects for whom we had longitudinal metagenomics data due to their receiving levofloxacin. We also had access to a larger dataset: shotgun metagenomics had been performed on a further 67 samples for a related study. In this case, the data was not longitudinal and encompassed an array of

antibiotic treatment conditions across 67 subjects, providing a diverse set of taxonomic and ARG data on which to test whether the relationship between *E. coli*, *P. mirabilis*, and *E. faecalis* and ARG density held true. As an initial test, we performed the same correlation analyses between species of interest and ARG levels as on the levofloxacin dataset, finding that both the simple and complex models showed significant correlation ($r = 0.7179$, $r^2 = 0.5154$, $p < 0.0001$ and $r = 0.7627$, r2 = 0.5817, $p<0.0001$, respectively; Pearson's correlation; Figure 5A-B). This provided initial support for the trend being present in the wider dataset.

We then created a multiple linear regression model to predict ARG density using the relative abundances (RA) of the three main species of interest in the initial levofloxacin dataset, with the following equation: (ARG density) = 0.003482 + 0.006221(*E. coli* RA) + 0.006248(*P. mirabilis* RA) + 0.006920(*E. faecalis* RA) (Figure 5B). We then used this equation to predict the ARG density in the larger metagenomics dataset and found that it was able to accurately predict the true ARG level of those samples, with predicted and actual values correlating significantly ($r = 0.7139$, $r^2 = 0.5096$, $p<0.0001$; Pearson's correlation; Figure 5C). As before, there were a few notable outliers with higher ARG levels than predicted by the model; those three samples contained high levels of *P. stuartii*, *P. aeruginosa*, or *Klebsiella pneumoniae* This maps well to the fact that we observed high levels of ARGs in bins constructed from *P. stuartii*, *P. aeruginosa*, and the related species *K. oxytoca* (Figure 4D).

We also created a multiple linear regression model that incorporated the relative abundances of *P. stuartii* and *S. haemolyticus*, which caused outliers from the original species-ARG correlation: (ARG density) = 0.003253 + 0.006715(*E. coli* RA) +

0.006748(*P. mirabilis* RA) + 0.003461(*E. faecalis* RA) + 0.01123(*S. haemolyticus* RA) + 0.007569(*P. stuartii* RA) (Figure 5E). As before, we tested this equation against the larger dataset, and found that it slightly increased the accuracy of the predictions; specifically, it eliminated the outlier which had high *P. stuartii* levels and slightly improved the correlation between predicted and actual ARG levels ($r = 0.7753$, $r^2 = 0.6012$, $p<0.0001$; Pearson's correlation; Figure 5F). However, the simpler model is more broadly applicable, as blooms of *P. stuartii* and *S. haemolyticus* are relatively uncommon. Similarly, while *Klebsiella* spp. and *P. aeruginosa* may also contribute to high ARG density in samples, they do not bloom as commonly in this cohort as the core predictive species of *E. coli*, *P. mirabilis*, and *E. faecalis*.

These results indicate that in this population, levels of only a few key species could predict the majority of ARG abundance beyond background levels. Both the core predictive species (*E. coli*, *P. mirabilis*, *E. faecalis*) and others that are associated with high ARG levels in samples (*P. stuartii*, *S. haemolyticus*, *P. aeruginosa*, *Klebsiella* spp.) are pathogens and/or pathobionts. Monitoring levels of these species may be helpful in elderly, institutionalized populations, as these patients may be vulnerable to developing or transmitting AMR infections from high-level carriage of these species.

Discussion

Overall, we found that the microbial composition of the gut microbiome of elderly patients with advanced dementia was quite variable, both between subjects and over time within the same subject. Even in the absence of antimicrobial treatment, there was notable fluctuation in the abundance of a number of species, including pathobionts such as *E. coli*, *P. mirabilis*, and *E. faecalis*. When comparing the taxonomic composition, functional

potential, and resistome of pre- and post-levofloxacin samples, we did not observe any significant differences. One potential reason for this finding is that oral levofloxacin is well-absorbed by the host, with greater than 99% bioavailability[565,566,646-648], and therefore may not be directly available to the luminal microbiota of the lower gastrointestinal tract at high levels. Furthermore, other studies have suggested that levofloxacin has a relatively minor impact on the gut microbiome, primarily reducing levels of *Enterobacterales*[568-574], and it may be less-associated with *Clostridiodes difficile*-associated diarrhea outbreaks than other antimicrobials, including other fluoroquinolones[649].

Additionally, in this cohort, levofloxacin was typically administered at least two weeks prior to collected timepoints, potentially allowing sufficient time for the microbiome to recover from or shift away from its immediately post-antibiotic state. Furthermore, the impacts of levofloxacin on the gut microbiome may be dependent upon the initial state upon administration. If the microbiome is initially relatively diverse and healthy, antibiotic administration may be disruptive and allow blooms of atypical dominant species such as members of *Proteobacteria*; such an occurrence might be observed in Subject F, where a diverse *Bacteroides*-dominated microbiome was overtaken by several *Enterobacterales* after levofloxacin treatment (Supplementary Figure 2G). Alternatively, if the microbiome is initially dominated by one or more pathogens, antimicrobial administration may correct such blooms and allow for the restoration of a diverse community, as might have occurred in Subject E as a *P. stuartii* bloom was eliminated (Supplementary Figure 2F).

Finally, since the pre-existing temporal instability of this community was high, levofloxacin-related changes may not be detectable through the noise of this cohort's general microbiome instability. In contrast to our observations, studies in adults have

generally found that the within-subjects dissimilarity is much lower than between-subjects dissimilarity, in line with the fact that the gut microbiome tends to be relatively stable within the same subject over time – including in an elderly cohort[1,27,555,650,651]. This suggests that the gut microbiomes of the subjects in this study were less stable than that of other cohorts, potentially suggesting that this institutionalized population with advanced functional impairment is more prone to infections or has weaker immune systems than young healthy adults or even community-resident elderly adults. Interestingly, despite the taxonomic variability, the functional composition of the cohort was relatively similar across samples and subjects. This is in line with previous studies of the human gut microbiome, which suggest that variable taxa can fill the same functional niches, resulting in a more similar functional composition across individuals despite inter-individual differences in the taxonomic composition[27,33,93].

As all of the subjects had been given an antibiotic, we were particularly interested in the antibiotic resistance profile of the subjects before and after levofloxacin administration. However, as observed in the taxonomic and functional data, there was no apparent association of any ARG genes or classes with either pre- or post-levofloxacin status. This may be due to the fact that levofloxacin did not have any specific impacts on the resistome of this cohort, or due to the factors that may have concealed any impacts of levofloxacin, as discussed above. However, we were particularly intrigued by the finding that ARG density in a particular sample could be linked to the abundance of a few key species. *E. coli*, *P. mirabilis*, and *E. faecalis* are all pathobionts that are often found at low levels in a healthy microbiome, but bloomed frequently at various timepoints across a majority of our subjects. All three species can cause severe illness, have been previously

observed to colonize nursing home residents, and include well-known multidrug-resistant strains[529-531,546,548,550,581,588,625-631,652]. In fact, three of the subjects (B, C, and D) are known to have acquired multidrug-resistant strains of *E. coli* and *P. mirabilis* during the study. However, we observed an association between these three species and ARG levels across the entire sample set (Figure 4B), and the ARG composition of the bins of *E. coli* and *P. mirabilis* from samples where MDROs were detected were not distinct from their other bins (Supplementary Figure 6A-B). This suggests that metagenomic sequencing may allow the identification of antimicrobial-resistant organisms that escape detection via culture-based techniques, although it is also possible that the multidrug-resistant isolates contained ARG-carrying plasmids that were not captured by our assembly and binning strategy.

A major implication of this finding is that metagenomic analysis could be a particularly useful tool to track antimicrobial resistance in institutions like nursing homes and hospitals, particularly with the capability to construct contigs and bins that allow examination of specific genomes. In this case, it has allowed us to connect the high levels of ARGs in certain samples with correspondingly high levels of specific pathobionts, which had high proportions of ARGs within their genomes even in samples where MDROs were not detected. In a vulnerable population already prone to infections and carriage of MDROs, metagenomics could be a useful surveillance tool to assess the prevalence or transmission of ARGs in long-term care facilities.

Importantly, all of the subjects in our study were institutionalized in nursing homes, and there exists significant potential for transfer of bacteria between patients. As all but two of our subjects (C and G) lived in different homes, we could not directly examine this possibility ourselves, but it is possible that the high abundance of pathobionts and/or ARGs

in our cohort is related to the spread of isolates within nursing homes. This also raises the possibility that we would not find a similar association between pathobionts and ARG levels in a healthy or community-based elderly cohort, who might be less likely to harbor or transmit such high levels of these bacteria. However, if an association between particular "sentinel" species and ARGs holds true in other elderly institutionalized populations, qPCR detection of the loads of these such pathobionts may allow for prediction of resistant bacterial outbreaks before they occur.

In addition to the increased potential for spread of resistant strains through institutions, there are some other potential explanations for the association between ARGs and these particular species. In particular, all of the species that we found to be associated with ARG density are potential human pathogens, can be grown *in vitro*, and have been previously associated with AMR phenotypes. ARGs, as well as mobile genetic elements carrying them, from these species may be better-studied than those from organisms less likely to pose a threat to human health, including gut commensals. If ARGs from these organisms are well-represented in databases, it could potentially bias analyses based on these databases toward detecting pathogen- over commensal-associated ARGs. However, there has been significant work done on the resistome of the human commensal microbiome, including functional metagenomics to detect new ARGs. These have found that commensal anaerobes may serve as significant reservoirs of ARGs, and may in some cases contribute to the transfer of resistance to pathobionts[653-660]. Commensal carriage of antimicrobial resistance genes may correspond to the baseline level of 0.3-0.4% ARGs observed in samples without pathobiont dominance.

Some limitations to the findings of this study must be acknowledged. First, as for all database-based methodologies, we are limited by accuracy and completeness of those databases. While the human gut microbiome is fairly well-characterized, there may be so-called microbial dark matter that is not well-represented in the taxonomic database used for species identification. We also used a database composed of bacterial and archaeal genomes, excluding consideration of bacteriophage and microbial eukaryotes from our analyses. As mentioned, database representation is particularly relevant for our ARG analysis, as the genes in this database may be skewed towards easily-culturable and pathogenic source species, and our analysis may have missed ARGs found in commensal or unculturable gut species. Additionally, critics have noted that some genes found in ARG databases used have unclear links to resistance phenotypes, and may perform regulatory, efflux, or other functions not always related to antimicrobial resistance[653,661].

Second, we were limited by the original SPREAD population, in which few subjects received only a single antimicrobial during the course of the study; this makes it difficult to say whether the temporal variability we observed was widespread in the cohort, although the fact that there were frequently high pathobiont levels observed in the larger metagenomics dataset we used to test our multiple linear regression suggests that this may be the case. Third, in this study we worked with rectal swabs, which are similar but not identical to fecal samples and may be susceptible to cross-contamination from urinary pathogens or skin flora, particularly in incontinent advanced dementia patients[662-665]. Fourth, metagenomic assembly has limitations. It cannot create bins of all species found in a given sample, genome reconstruction is based on the isolates present in the database used, and analysis of assembled genomes may exclude consideration of plasmids – which are

often sources of ARGs. Finally, as we analyzed metagenomic data, we cannot comment on the actual antimicrobial resistance phenotypes of the communities or individual bacteria that we studied.

Conclusions

The gut microbiome was highly variable both between and within subjects, with frequent blooms and reductions of bacterial species both before and after levofloxacin treatment. We did not observe a consistent impact of levofloxacin on specific taxa or functions, levels of antimicrobial resistance genes, or overall microbiome diversity in these subjects. However, while we could not link levofloxacin to antimicrobial resistance gene levels, there were a number of samples that had higher relative abundances of these genes. In our original metagenomics dataset, we were able to identify that levels of these genes could be linked to blooms of specific bacterial species, including *E. coli*, *P. mirabilis*, and *E. faecalis*. We were able to build a model to predict total ARG levels in a sample from the relative abundance of these species, and confirm the validity of this model in a larger metagenomics dataset from the rest of the SPREAD study, including subjects taking a range of antibiotics. Furthermore, use of metagenomic assembly and binning allowed us to confirm that our species of interest carry greater ARG densities than other abundant members of the microbiome, even in subjects where MDROs were not detected by culturing.

This demonstrates that there is a significant amount of information that can be obtained from metagenomic assembly and binning. With sufficient depth, powerful computational tools allow whole genomes to be assembled from short-read metagenomic sequencing, which permits interrogation of the likely features of species of interest in

complex microbial communities. In our case, we were able to confirm the association between pathobiont blooms and ARG levels in the gut, showing that the genomes of pathobionts contained a greater proportion of ARGs than gut commensals such as *Bacteroides* and *Bifidobacterium* species. This suggests that while the commensal microbiota are known to serve as reservoirs of antimicrobial resistance, in this cohort blooms of pathobionts may serve as the driver of ARG levels in the gut microbiome. Given how frequently these blooms occurred, special attention should be paid to these species in dementia patients in long-term care facilities, a vulnerable group which is often immunocompromised, frequently administered medication including antimicrobials, and may carry MDRO at relatively high levels.

Methods

Subject Selection

Eleven subjects were chosen from the SPREAD cohort based on the following inclusion criteria: at least two consecutive rectal swabs were collected from the subject during the study, subjects had received a single oral course of levofloxacin during the study (average course of 8 days), and subjects received no other antimicrobials during the study or in the 3 months prior to the first sample collection. Of the 11 subjects, 10 were female and 10 were white, while ages ranged from 72 to 101. Five subjects lived through the entire sample collection period, while the other six passed away at some point prior to the final collection; between this attrition, one sample that was not collected, and three samples that were not well-sequenced, we had a total of 38 usable metagenomic samples (Figure 1A; Supplementary Tables 1-2). All samples were collected under SPREAD, which was approved by the Institutional Review Board at Hebrew Life[549].

Sample Collection

Samples were collected by insertion of sterile double-tipped swabs (Starswab II; Starplex Scientific Inc., Ontario, Canada) into the anus of the subject. The first swab was used to identify multidrug-resistant organisms (including methicillin-resistant *S. aureus*, vancomycin-resistant enterococci, and multidrug-resistant Gram-negative organisms such as *E. coli*, *P. mirabilis*, *P. aeruginosa*, or *P. stuartii*) via culturing techniques as described previously[666]. The second swab was frozen in 20% glycerol at -80°C for DNA extraction and sequencing.

Sample Processing

Frozen rectal swabs were thawed and placed into 96-well plates, before extraction using the PowerSoil DNA Isolation Kit (MOBIO, West Carlsbad, CA) according to the manufacturer's instructions. DNA concentrations were measured using a Nanodrop 1000 (Thermo Scientific, Waltham, MA) and extracted DNA was stored at -20°C until further use.

16S rRNA Amplicon Sequencing

The V4 hypervariable region of the 16S rRNA gene was amplified according to Earth Microbiome Project protocols. Amplification was performed using Illumina-adapted universal 16S primers 515F and 806R under the following conditions: 3 minutes at 94°C, 45 cycles of [45 seconds at 94°C, 60 seconds at 50°C, 90 seconds at 72°C], 10 minutes at 72°C. All reactions were prepared using 5 PRIME polymerase 1X HotMasterMix (5PRIME, Gaithersburg, MD) and run in triplicate to alleviate primer bias. Triplicates were pooled before cleaning with a PCR Purification Kit (Qiagen). These products were quantified using the Qubit dsDNA High Sensitivity Assay Kit (Invitrogen, Eugene, OR)

and samples were pooled in equimolar amounts. Sequencing was performed using the Illumina MiSeq platform located at the New York University Langone Medical Center Genome Technology Core. Sequences can be found under the BioProject accession number PRJNA573963 (http://www.ncbi.nlm.nih.gov/bioproject/573963).

16S rRNA Amplicon Data Processing

Data processing was performed using the QIIME2 (v 2019.1) pipeline[5]. The Divisive Amplicon Denoising Algorithm 2 (DADA2) method was used to quality-filter sequences and categorize amplicon sequence variants (ASVs)[317], and the SILVA (release 132) 99% identity V4 classifier was used to assign taxonomy to ASVs[7]. See Supplementary File 1 for more information. Taxonomic relative abundances were exported at the genus level for further analysis. Output data can be found in Supplementary Data 2.

Shotgun Sequencing

Extracted DNA (2 ng DNA in 50 uL buffer) was sheared to 450bp using a Covaris LE220 system. Library preparation was performed using a Biomek FXP Automated Liquid Handling Workstation (Beckman Coulter) with the KAPA HyperPrep Kit (Roche), with 12 cycles of PCR. Final libraries were normalized and pooled, with 20 samples per poor. Each pool was run on 2 lanes of an Illumina HiSeq 4000 using the paired-end 2x150bp protocol. Library preparation and sequencing was performed at the New York University Langone Medical Center Genome Technology Core. Sequences can be found under the BioProject accession number PRJNA573963 (http://www.ncbi.nlm.nih.gov/bioproject/573963) for the levofloxacin dataset and under the BioProject accession number PRJNA531921 (https://www.ncbi.nlm.nih.gov/bioproject/531921) for the test dataset.

Shotgun Processing

Raw shotgun sequencing reads were processed using Kneaddata (v0.6.1) to remove contaminating human sequences from the dataset[667]. Briefly, the *kneaddata* function was used with the Bowtie2 *Homo sapiens* database (v0.1)[668] to remove contaminating host reads from the sequencing files. See Supplementary File 1 for more information.

Shotgun Sequencing Taxonomic Classification

Kraken2, a taxonomic classifier that maps shotgun sequencing k-mers to genomic databases, was used to assign taxonomy to kneaddata-processed shotgun sequencing reads[583]. Briefly, the *kraken2-build* function was used to create a custom database containing the "bacteria" and "archaea" from NCBI libraries, and the *kraken2* function was used to run the kneaddata-filtered shotgun sequencing reads against this database and assign taxonomy. While Kraken2 does not estimate species abundances, Bracken2 (Bayesian Reestimation of Abundance with KrakEN) uses the taxonomy assigned by Kraken2 to estimate the number of reads per sample that originate from individual species[584]. The Kraken2 database was used to create a Bracken-compatible database using the *bracken-build* function, and the Kraken2 report files for each sample were run against the Bracken database using the *bracken* function for the phylum, genus, and species levels. Phylum- and species-level relative abundance outputs were formatted for biomarker discovery using LEfSe. The *kraken-biom* function was used to convert the Bracken report files into a biom file for import into R. Output data can be found in Supplementary Data 1. Relative abundance plots were generated in GraphPad Prism v8.

Shotgun Sequencing Taxonomic Diversity Analysis

Alpha and beta diversity analyses were performed using the phyloseq (v1.27.2)[318,319] and vegan (v2.5-4)[320] packages in R (v3.4.3). Briefly, the biom file was

imported into a phyloseq object. The phyloseq *estimate_richness* function was used to obtain Shannon's Diversity Index values for all samples, while the vegan *phyloseq::distance* and *ordinate* functions were used to generate a Bray-Curtis matrix and PCoA values. See Supplementary File 1 for more information. Data was exported as csv files for formatting, and plotting was performed in GraphPad Prism v8.

Shotgun Sequencing Gene and Pathways Analysis

The Human Microbiome Project Unified Metabolic Analysis Network 2 (HUMAnN2) pipeline was used to profile the presence and abundance of genetic pathways in our samples[608]. Briefly, the *HUMAnN2* function was used with the kneaddata-filtered metagenomic sequences to estimate genes and MetaCyc pathways present in the samples based on the UniRef90 database, files were joined using the *HUMAnN2_join_tables* function and the full tables were de-leveled using the *HUMAnN2_split_stratified_table* function. The unstratified gene-level abundances were converted to both GO terms and KEGG orthologs using the *HUMAnN2_regroup_table* function, and the *HUMAnN2_renorm_table* function was used to normalize the MetaCyc pathway, GO term, and KEGG ortholog tables by computing relative abundance. These relative abundance tables were formatted for biomarker discovery with LEfSe. Additionally, the, and LEfSe was also used to analyze pre- and post-treated samples using both outputs. See Supplementary File 1 for more information. Output data can be found in Supplementary Data 3. Plots were generated in Graphpad Prism 8.

Shotgun Sequencing Resistome Analysis

The ARG content of the samples was analyzed using DeepARG-SS, a deep learning model that can predict ARGs from short-read metagenomic data[609]. We first analyzed the

data using the *deeparg* function with the *-reads* flag. The mapped ARGs output was then imported into R, where it was processed to obtain tables of the ARGs detected per sample at both the specific gene and antibiotic class levels. The ARGs detected were normalized to the number of reads per sample.

Additionally, after metagenomic assembly and binning was performed (see below), individual bins were analyzed using DeepARG-LS, a deep learning model optimized to predict ARGS from gene-level input. The *DNA_features* output from selected bins was analyzed using the *deeparg* function with the *-genes* flag to analyze whether the levels or identity of ARGs could be linked to specific species of interest. The ARGs detected were normalized to the number of features per bin. All output data can be found in Supplementary Data 4. Plots were generated in GraphPad Prism 8.

Shotgun Sequencing Metagenomic Assembly and Binning

To further examine the ARGs present in the samples, kneaddata-filtered reads were uploaded to the web-based Pathosystems Resource Integration Center (PATRIC)[669]. Reads were assembled into contigs using the *auto* option of the Assembly service, which provides both raw output contigs from specific assembly algorithms and contigs of the "best" assembly as judged by the in-house PATRIC script ARAST. We ran the assembly using two different inputs: reads that had been processed by *kneaddata* as pairs, which has the advantage of utilizing mate-pairing information for longer total reads, and reads that had been processed by *kneaddata* after pairs were concatenated into a single file, which has the advantage of keeping reads whose mates failed trimming.

Both the raw SPAdes assembly algorithm contigs[670] and the best assembly contigs were then processed using the Metagenomics Binning service, which assigns contigs to

specific organisms and annotates the bin's genome. Quality measures were used to define bins as either "good", "acceptable", or "bad" according to the criteria in Supplementary Table 3, and only "good" or "acceptable" bins were used moving forward. When more than one binning strategy (paired assembly or single assembly, SPAdes contigs or best contigs) called a particular bin as "good" or "acceptable", quality measures from the four strategies were compared and the highest-quality bin for a given species of interest was chosen for ARG analysis. Finally, only bins of species present at 0.1% or greater relative abundance in the corresponding sample were selected for further analysis. A list of bins used, their source, and quality measures can be found in Supplementary Table 4.

Taxonomic Biomarker Analysis

LEfSe was used to identify potential biomarkers distinguishing levofloxacin-treated samples[518]. In all cases (taxonomic abundances, MetaCyc pathways, KEGG orthologs, GO terms, ARGs), data was formatted into csv files and uploaded to the Galaxy webserver. LEfSe was run under default parameters for biomarker detection, comparing either all pre-levofloxacin to all post-levofloxacin or immediately pre-levofloxacin to immediately post-levofloxacin. LEfSe was also used to compare genus-level taxonomic abundance outputs from Kraken2/Bracken2 and QIIME2, again under default parameters.

Declarations

Ethics Approval and Consent to Participate

Written information about SPREAD was mailed to the healthcare proxies of all eligible subjects. Proxies were then telephoned two weeks later to solicit participation and verbally obtain informed consent for the participation of themselves and the subjects. Approval for

SPREAD, including the consent procedures, was obtained from the Institutional Review Board committee at Hebrew SeniorLife.

Data Availability

The underlying sequencing data for the current study are available in the NCBI Short Read Archive repository; 16S rRNA and shotgun metagenomics data for the levofloxacin samples can be found under BioProject ID PRJNA573963 and shotgun metagenomics data for the test dataset can be found under BioProject ID PRJNA531921. Analysis code and data generated from this study can be found in Supplementary File 1 and Supplementary Data 1-4, respectively.

Acknowledgements

The authors acknowledge the laboratory of Dr. Martin Blaser and the New York University Langone Medical Center Genome Technology Core, which prepared metagenomic libraries and performed all 16S rRNA and shotgun metagenomics sequencing for this study.

Funding

This study was funded by: the National Institute of General Medical Sciences institutional development award P20GM121344 for the COBRE Center for Antimicrobial Resistance and Therapeutic Discovery at Brown University (PB), the Centers for Disease Control and Prevention award 200-2016-91939 (EMCD), the National Institutes of Allergy and Infectious Diseases award K24AI119158 (EMCD), the National Institute of Aging award R01AG032982 (EMCD), and the Millennium Nucleus for Collaborative Research on Bacterial Resistance (MICROB-R) supported by the Millennium Scientific Initiative of the Ministry of Economy, Development and Tourism (Chile) (RA). The funders had no role in

study design, data collection and interpretation, or the decision to submit the work for publication. Opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the funders. The authors declare no competing interests.

Author's Contributions

ADR conceptualized the project, performed analysis, generated figures, and wrote the manuscript. RA collected the data and contributed to manuscript preparation. EMCD collected the data and contributed to manuscript preparation. PB conceptualized the project and wrote the manuscript. All authors read and approved the final manuscript.

Figure 1: Subject Overview and Diversity Metrics

(A) Metagenomics sequencing was performed on longitudinal samples from eleven subjects from SPREAD who had received a single course of levofloxacin during their participation in the study. Points represent collection of samples, at intervals of approximately 3 months, relative to administration of levofloxacin. (B) Shannon diversity over time of all subjects. The dashed line indicates administration of levofloxacin. p = 0.175 for immediately pre-levofloxacin vs. immediately post-levofloxacin samples and p = 0.1006 for all pre-levofloxacin vs. all post-levofloxacin samples; Mann-Whitney test. (C) PCoA analysis of Bray-Curtis Dissimilarity. Solid arrows connect immediately pre- with immediately post-levofloxacin samples, dashed arrows connect other sequential samples, and dotted arrows connect samples where an intermediate sample is missing. (D) Within-subjects Bray-Curtis Dissimilarity of sequential samples. p = 0.6248 between pre-levofloxacin samples, post-levofloxacin samples, or pre-post levofloxacin samples; ANOVA). (E) Overall within-subjects, T0 between-subjects, and overall between-subjects Bray-Curtis Dissimilarity. p = 0.0262 for overall within-subjects vs. T0 between-subjects and p = 0.0175 for overall within-subjects vs. overall between-subjects; t-test with Welch's correction.

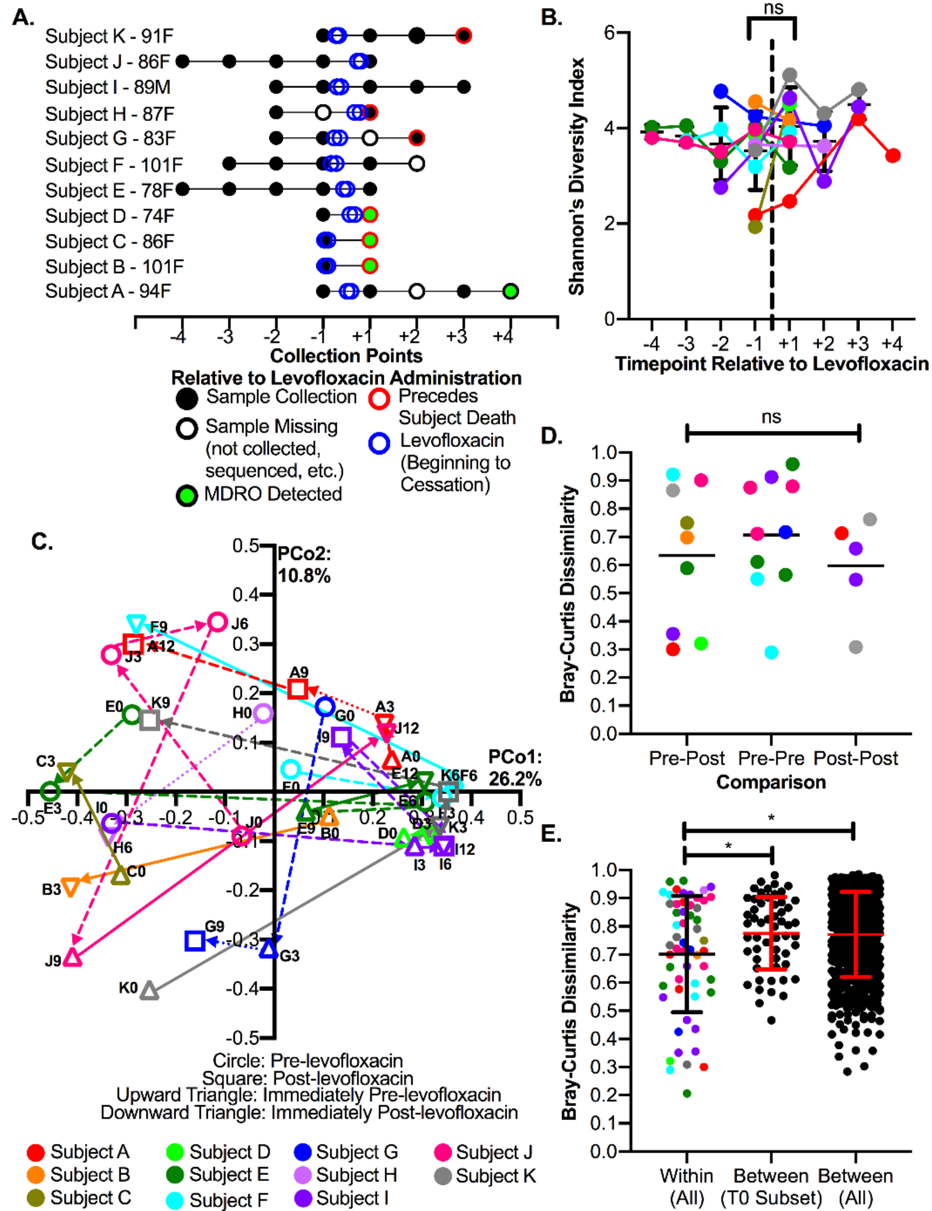


Figure 2: Relative Abundances of Species and Antimicrobial Resistance Genes

(A) Relative abundance of the most-abundant species across all samples, with all other species grouped in the "other" category. Species are grouped by genus and phylum, and are ranked within those levels by average relative abundance across all samples. Broad color categories distinguish phylum (*Proteobacteria* are red, *Bacteroidetes* are blue, *Firmicutes* are green, and *Actinobacteria* are purple), while different species of the same genus are given the same specific background color. Red lines indicate levofloxacin administration; dashed lines indicate usage between consecutive timepoints, while dotted lines indicate usage where the immediately post-levofloxacin sample is missing (B) Relative abundance of the most-abundant antimicrobial resistance genes (ARGs) across all samples. Specific ARGs are grouped by the class of antimicrobials they provide resistance to. Broad color categories distinguish class (Multidrug RGs are blue, MLS RGs are red, etc.), while related gene categories (ex: the *mec* operon or *mex* efflux proteins) are given the same specific background color. All ARGs were normalized to the total number of reads.

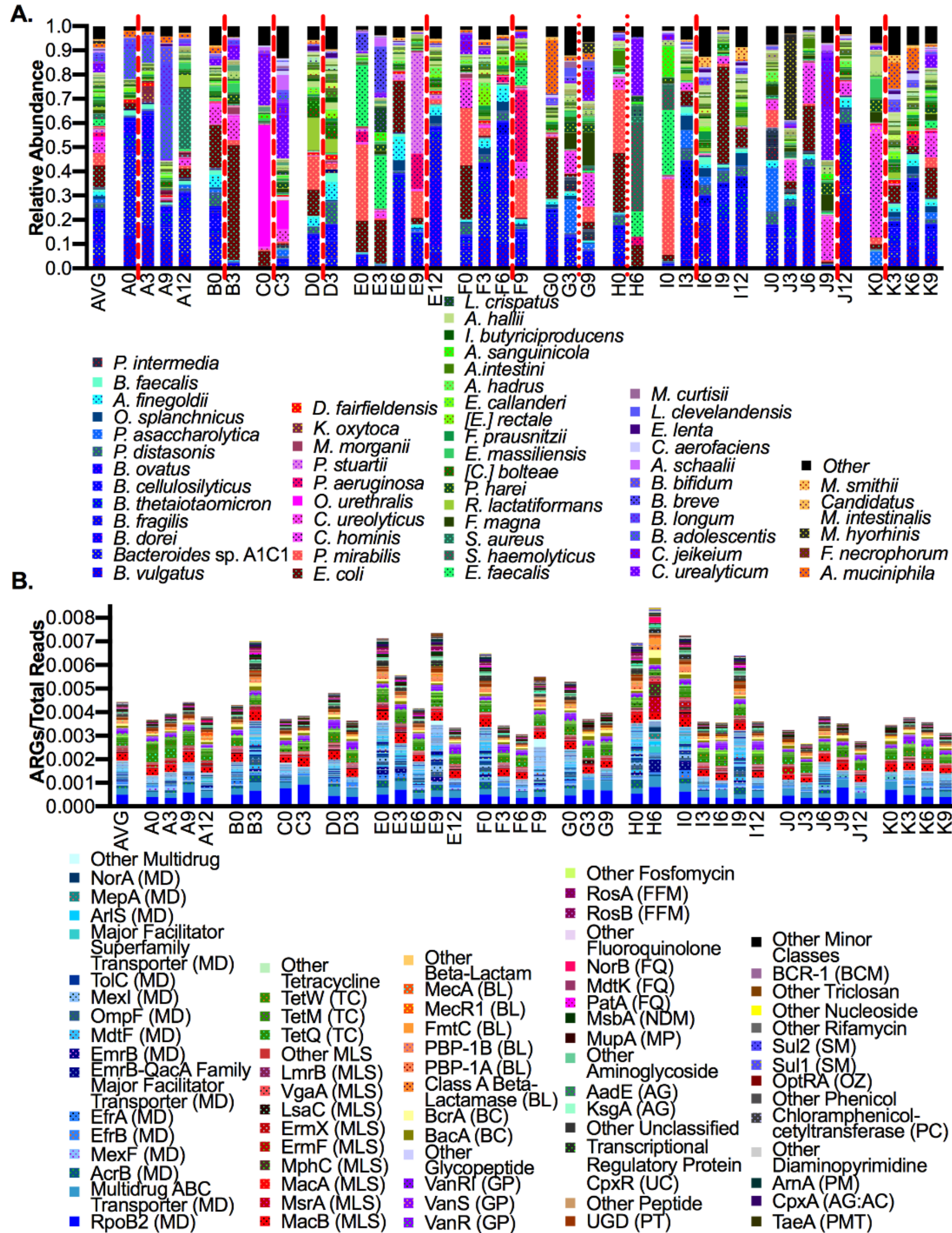


Figure 3: Antimicrobial Resistance Gene Profiles Reflect Taxonomic Observations

(A) Relative abundance of species in Subject A, showing a bloom in *S. aureus* at T12. (B) Relative abundance of ARG classes in Subject A, showing an expansion in beta-lactam resistance genes at T12. (C) Relative abundance of beta-lactam resistance genes in Subject A, showing increases in the *mecA/mecI/mecRI* operon at T12. (D) Relative abundance of species in Subject B, showing a bloom in *E. coli* at T3. (E) Relative abundance of ARG classes in Subject B, showing an expansion in multidrug, beta-lactam, and fluoroquinolone resistance genes at T3. (F) Relative abundance of multidrug resistance genes in Subject B, showing increases in various ARGs at T3. (G) Relative abundance of fluoroquinolone resistance genes in Subject B, showing increases in genes including *patA* and *mdtK* at T3. (H) Relative abundance of beta-lactam resistance genes in Subject B, showing increases in genes including penicillin-binding proteins and class C beta-lactamase at T3. (I) Relative abundance of species in Subject H, showing a bloom in *S. haemolyticus* at T6. (J) Relative abundance of ARG classes in Subject H, showing increases in MLS and fluoroquinolone resistance genes. (K) Relative abundance of MLS resistance genes in Subject H, showing an increase in staphylococcal resistance gene *msrA* and others at T6. (L) Relative abundance of fluoroquinolone resistance genes in Subject H, showing an increase in staphylococcal resistance gene *norB* and others at T6.

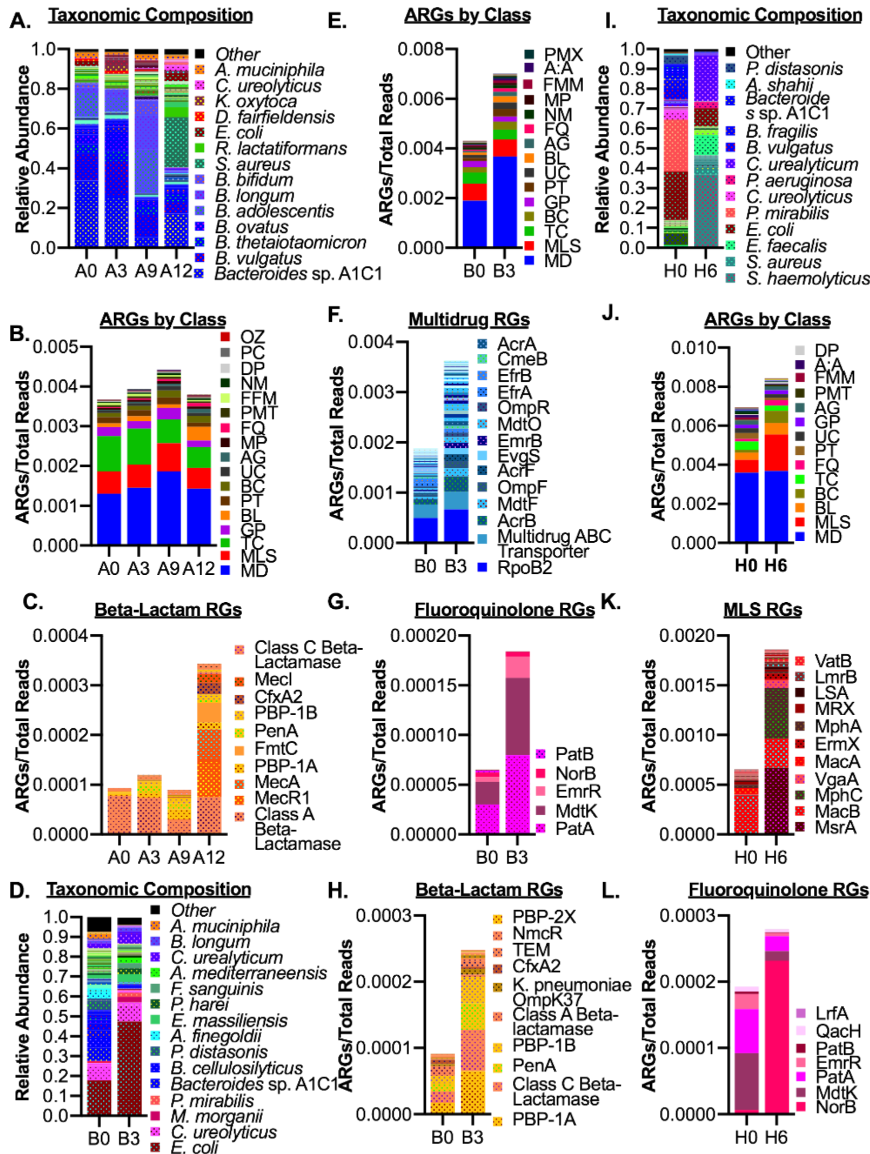


Figure 4: Relationship of ARG Levels to the Relative Abundance of Specific Pathobionts

(A) Correspondence between the relative abundances of key species of interest (*E. coli*, *P. mirabilis*, *E. faecalis*, *P. stuartii*, and *S. haemolyticus*) and total ARG density in each sample. (B) Correlation between the sum of the relative abundances of *E. coli*, *P. mirabilis*, and *E. faecalis* and the total ARG density in each sample (r = 0.791, $R^2 = 0.6254$, $p < 0.0001$; Pearson's correlation). (C) Correlation between the sum of the relative abundances of *E. coli*, *P. mirabilis*, *E. faecalis*, *P. stuartii*, and *S. haemolyticus* and the total ARG density in each sample (r = 0.933, $R^2 = 0.8706$, $p < 0.0001$; Pearson's correlation). (D) Average ARG density in bins of species across all samples in which we were able to construct a bin for that species. Specific genes are grouped and colored by their ARG class, and bins are grouped by phylum and ranked by their total average ARG density within that phylum.

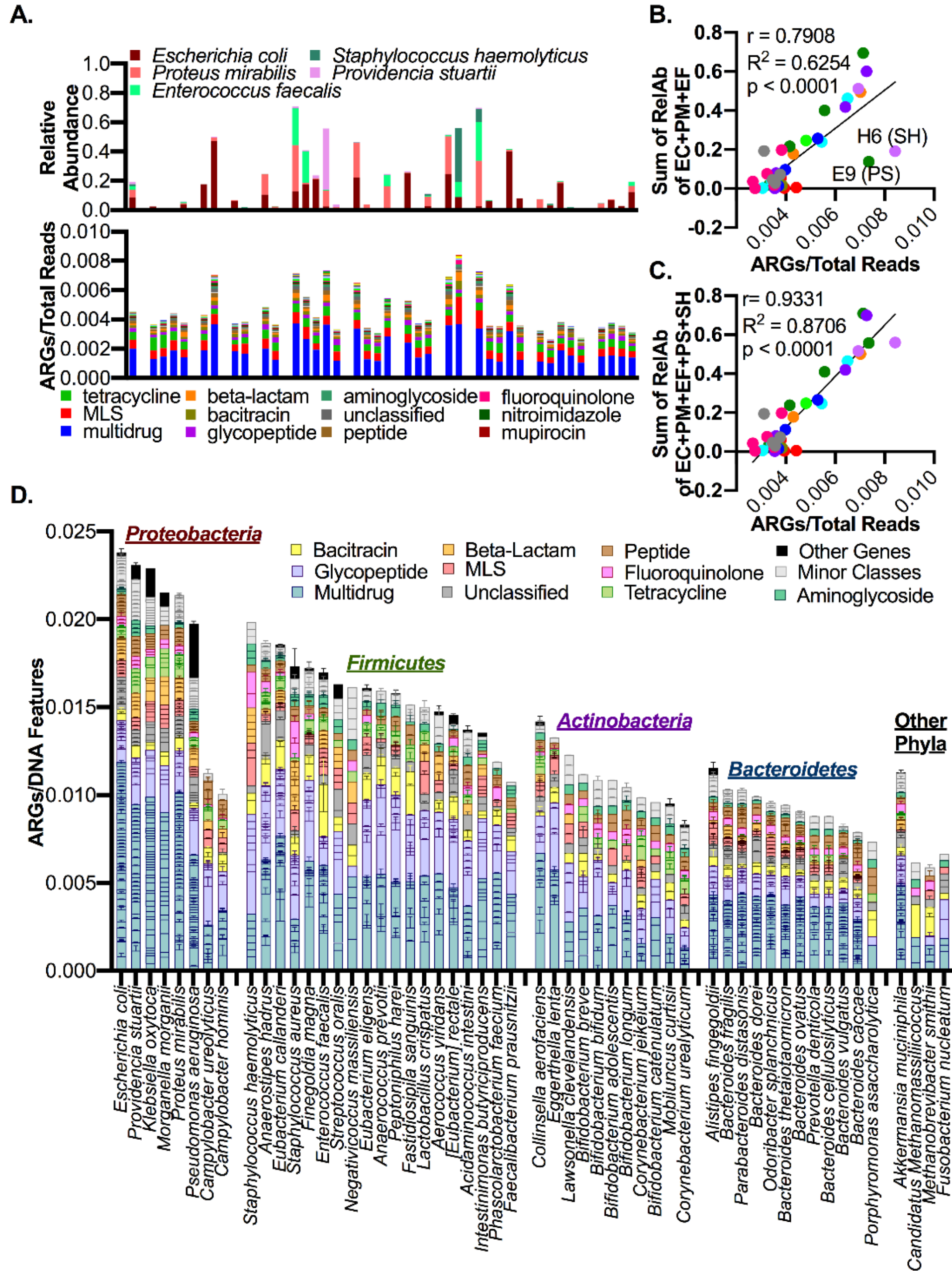
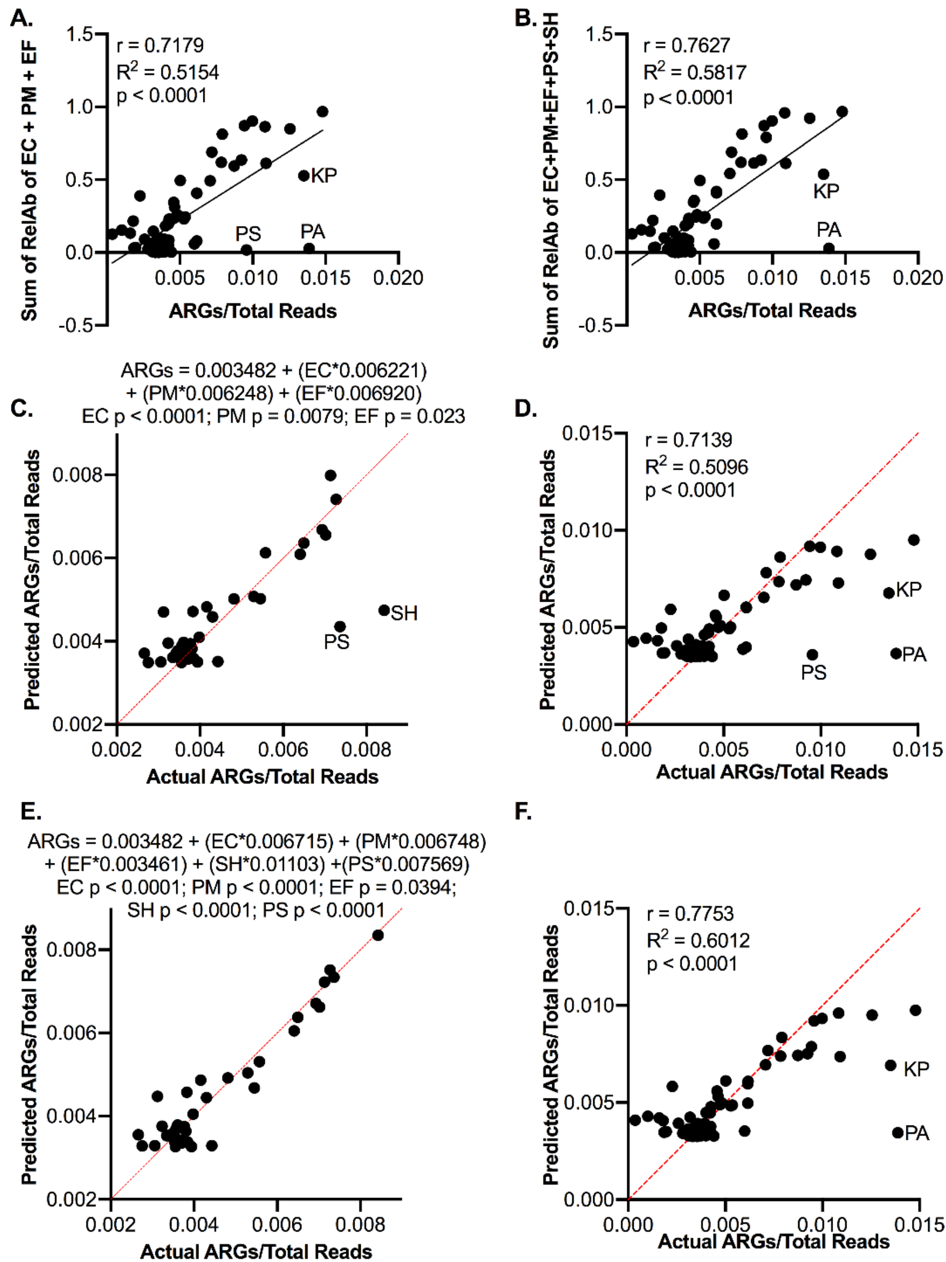
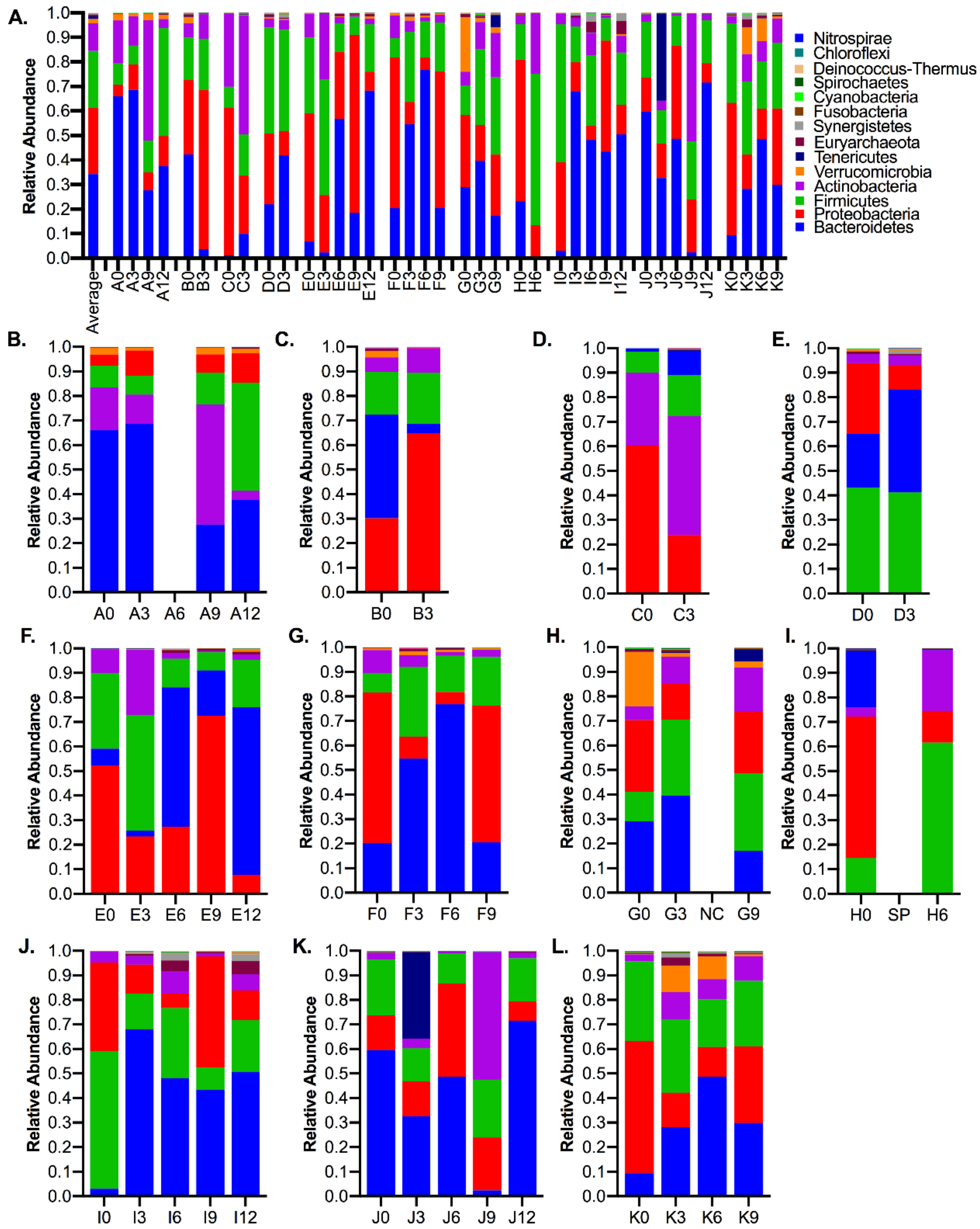


Figure 5: Prediction of ARG Density From Relative Abundance of Specific Pathobionts

(A) Correlation between the sum of the relative abundances of *E. coli*, *P. mirabilis*, and *E. faecalis* and the total ARG density in each sample in the test dataset ($r = 0.7139$, $r^2 = 0.5096$, $p<0.0001$; Pearson's correlation). (B) Correlation between the sum of the relative abundances of *E. coli*, *P. mirabilis*, *E. faecalis*, *P. stuartii*, and *S. haemolyticus* and the total ARG density in each sample in the test dataset ($r = 0.7753$, $r^2 = 0.6012$, $p<0.0001$; Pearson's correlation). (C) Multiple linear regression of relative abundances of *E. coli*, *P. mirabilis*, and *E. faecalis* to ARG density in samples in the levofloxacin dataset (38 samples). (D) Correlation between the predicted ARG density and actual ARG density in the test dataset (67 samples) based on the relative abundances of *E. coli*, *P. mirabilis*, and *E. faecalis*. (E) Multiple linear regression of relative abundances of *E. coli*, *P. mirabilis*, *E. faecalis*, *P. stuartii*, and *S. haemolyticus* to ARG density in samples in the levofloxacin dataset (38 samples). (F) Correlation between the predicted ARG density and actual ARG density in the test dataset (67 samples) based on the relative abundances of *E. coli*, *P. mirabilis*, *E. faecalis*, *P. stuartii*, and *S. haemolyticus*.

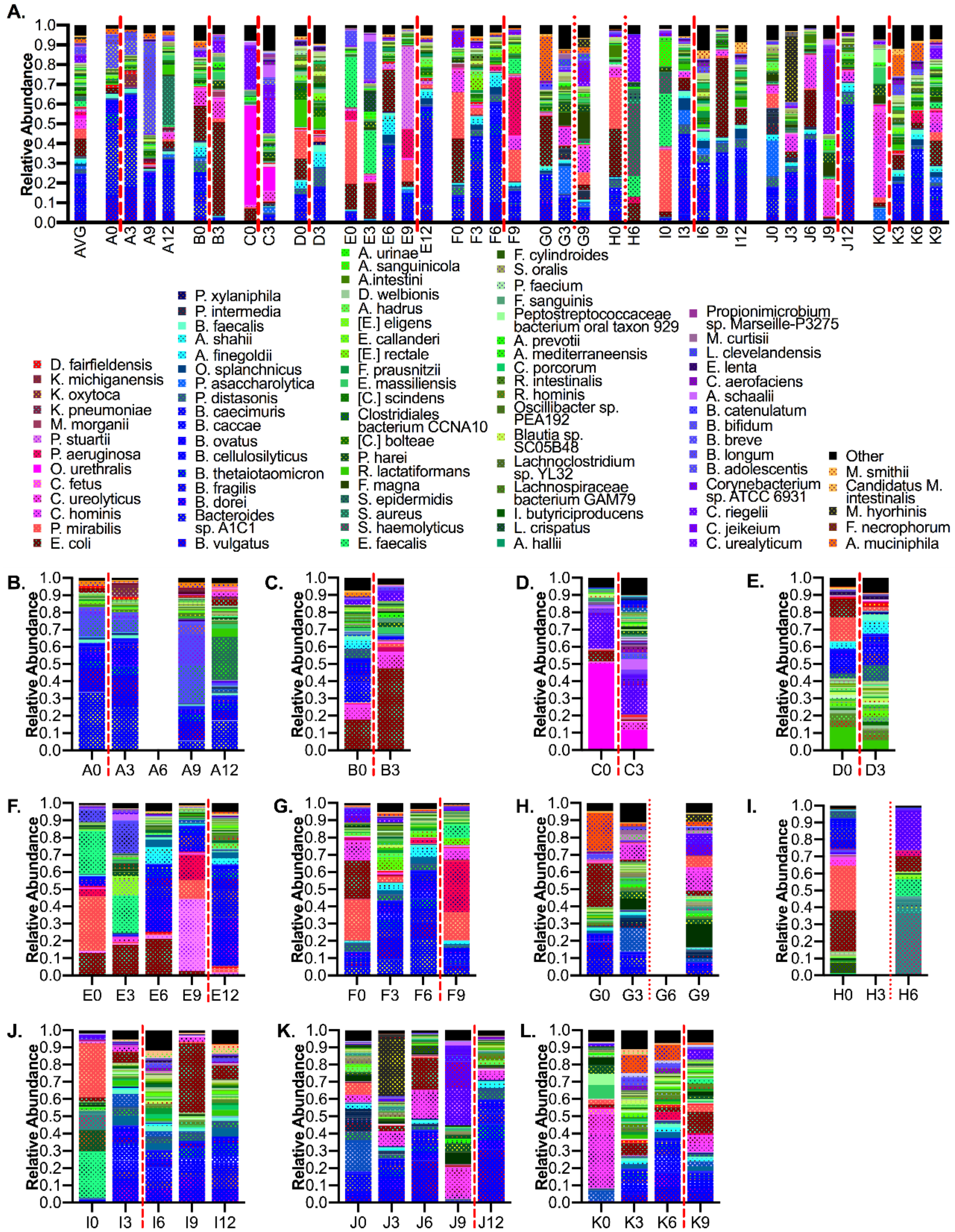


Supplementary Figure 1: Relative Abundances of Phyla Across and Within Subjects
(A) Relative abundance of phyla in all samples, ranked by average across all samples. (B-L) Relative abundances of phyla by subject, ranked by average within each subject.



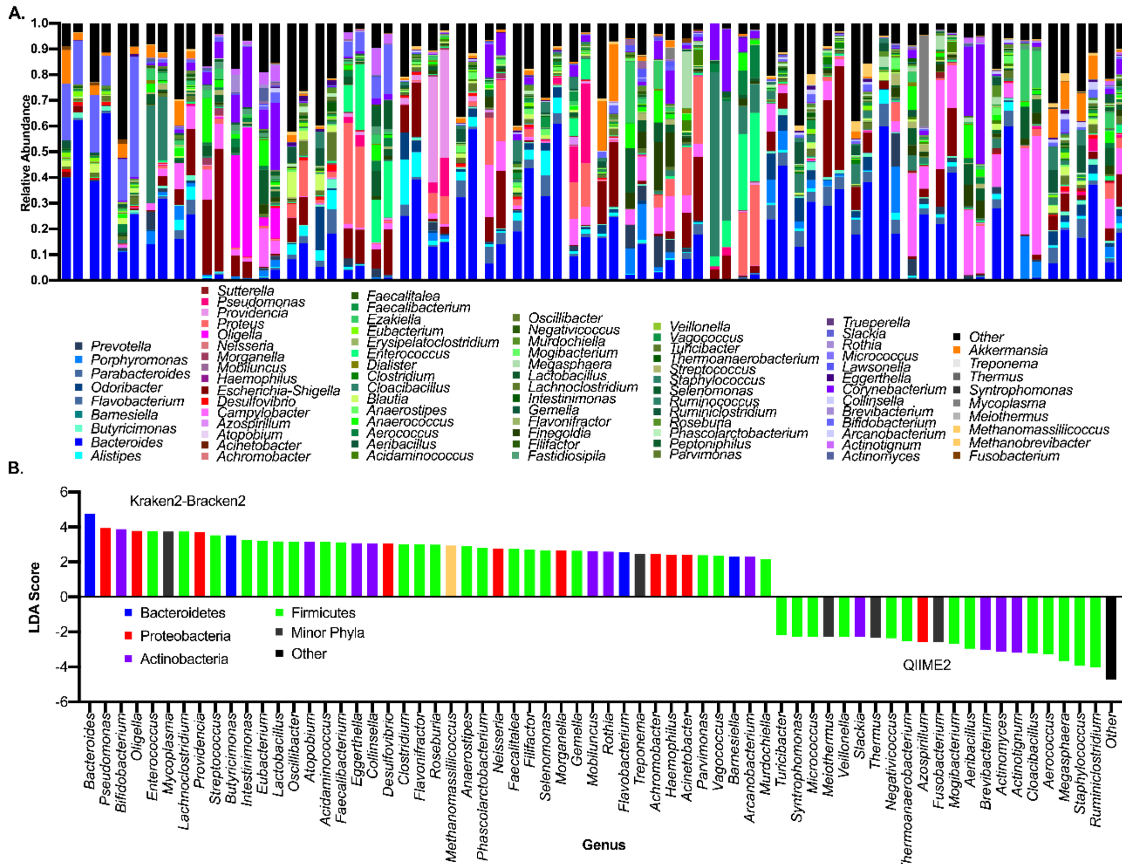
Supplementary Figure 2: Relative Abundances of Species Across and Within Subjects

(A) Relative abundance of species in all samples, grouped by genus and phylum and ranked within those levels by average relative abundance across all samples. (B-L) Relative abundances of phyla by subject, grouped by genus and phylum ranked within those levels by average within each subject. Coloring is the same as in Figure 2A.

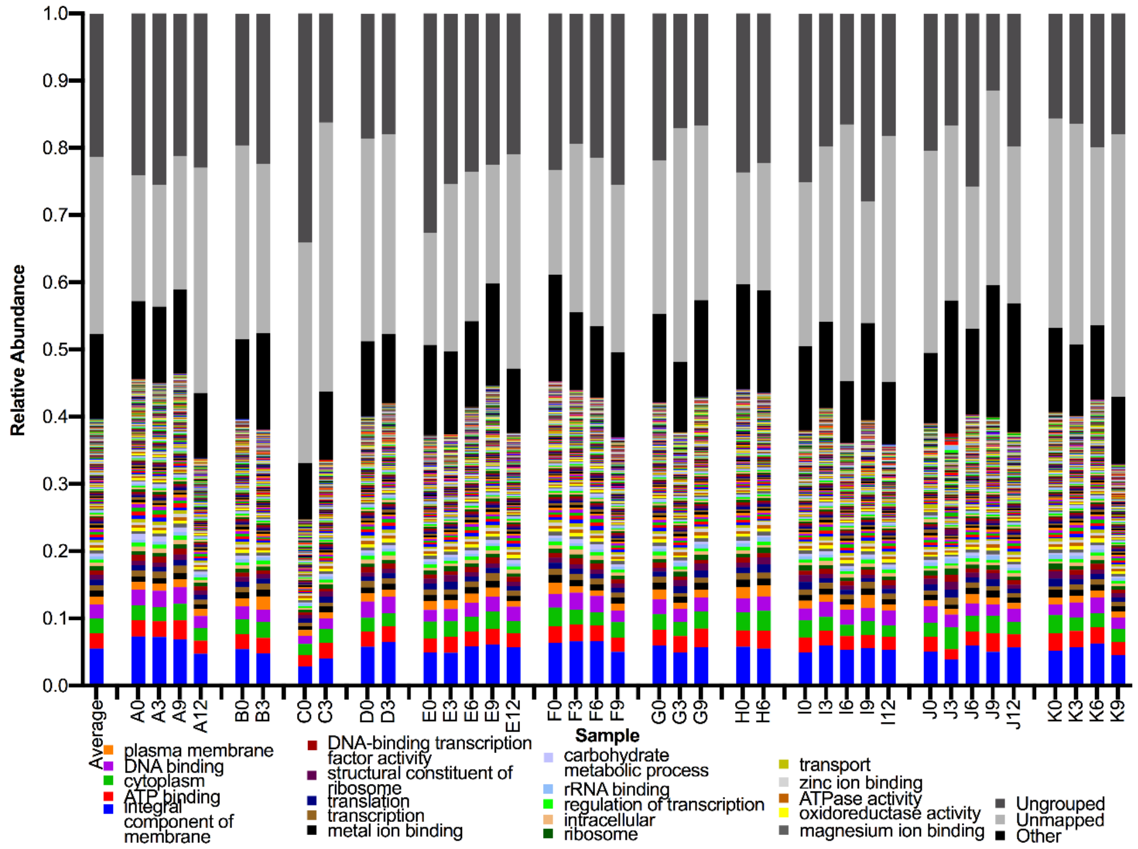


Supplementary Figure 3: Comparison of Genus-level Classifications by Metagenomics and 16S rRNA Analysis

(A) Relative abundances of genera called by both QIIME2 and Kraken2/Bracken2, where pairs of stacked bars indicate the same sample as measured by both methods. (B) Genera called by LEfSe as associated with either QIIME2 or Kraken2/Bracken2. Each genus is colored according to its source phylum.

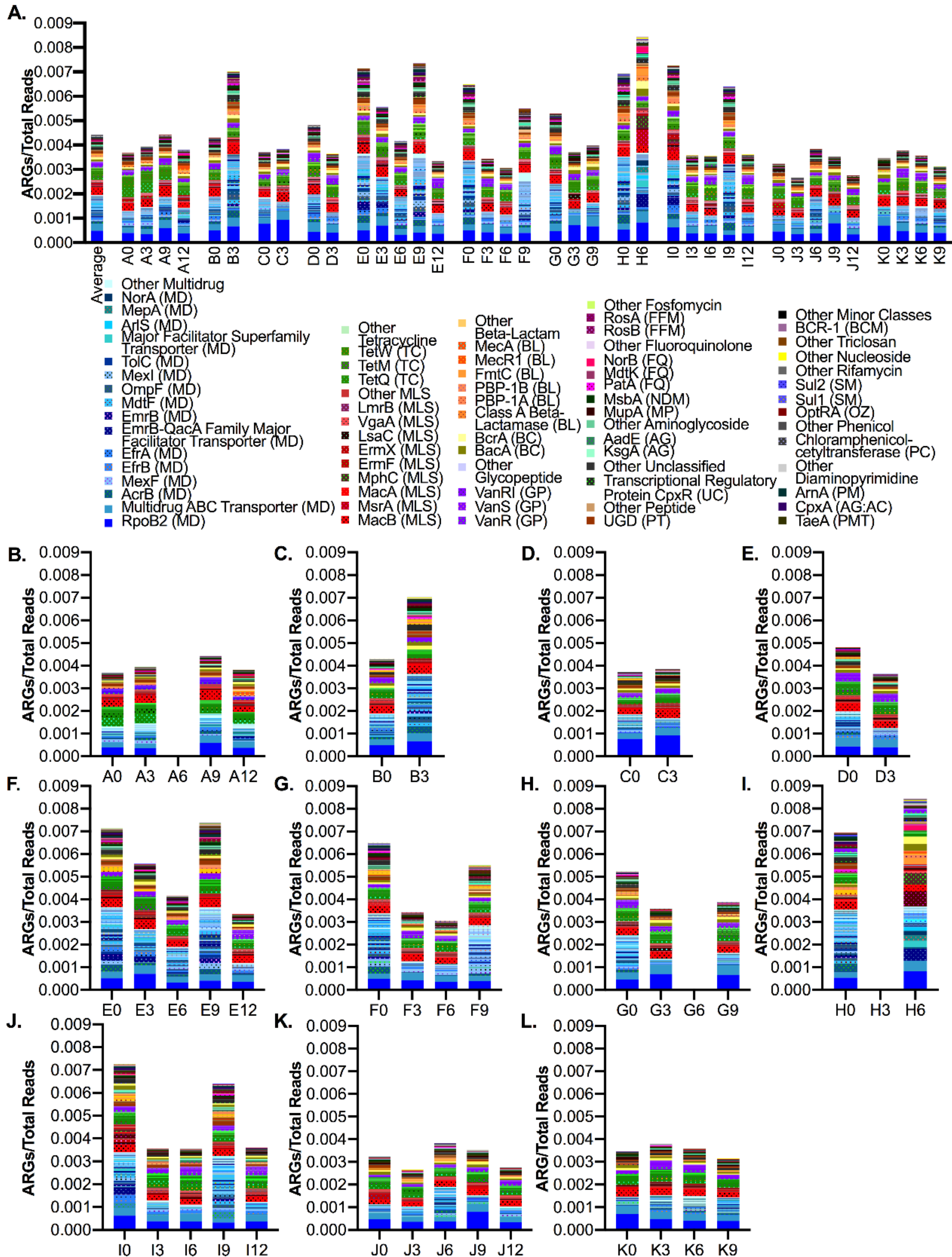


Supplementary Figure 4: Relative Abundance of Gene Ontology Terms Across All Samples
(A) Relative abundances of the top 250 most-abundant GO terms, representing broad functional categories, across all samples. A significant proportion are "unmapped" or "ungrouped", as not all UniRef90 gene families can be mapped to a GO term.

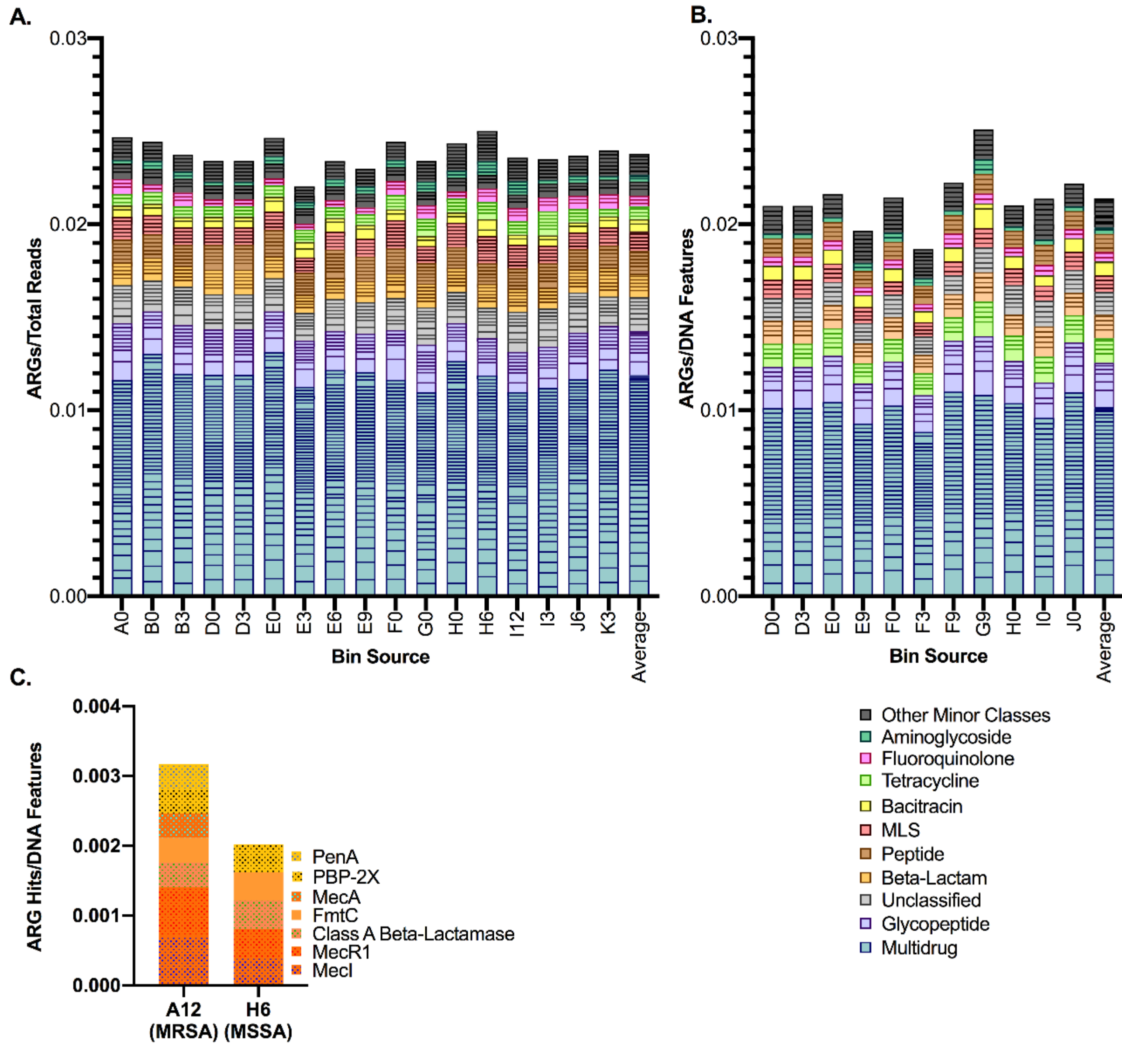


Supplementary Figure 5: Relative Abundance of Antimicrobial Resistance Genes Within and Across Subjects

(A) Relative abundance of species in all samples, grouped and ranked within class by average relative abundance across all samples. (B-L) Relative abundances of ARGs by subject, grouped and ranked within class by average relative abundance within each subject. Coloring is the same as in Figure 2B.



Supplementary Figure 6: Comparison of MDRO and non-MDRO Bins of the Same Species
(A) ARG density in all *E. coli* bins across samples. (B) ARG density in all *P. mirabilis* bins across samples. (C) Beta-lactam ARG density in all *S. aureus* bins across samples.



Supplementary Table 1: Metadata on Levofloxacin Cohort from SPREAD

This table lists the age, biological sex, and race of all subjects, whether a multidrug-resistant organism (MDRO) was detected in the subject at any timepoint, the duration of levofloxacin administration, and the reason for which they were administered levofloxacin. For MDROs, the specific organism detected and the antimicrobial agents it was found to be resistant to are also listed.

Subject	Sex	Age	Race	MDRO Detected	Levofloxacin Duration	Reason for Levofloxacin Administration
A	F	94	W	Yes (*S. aureus*; methicillin)	7 days	Urinary tract infection
B	F	101	W	Yes (*E. coli*; ampicillin/sulbactam, cefazolin, ceftazidime, ceftriaxone, ciprofloxacin)	6 days	Upper respiratory tract infection
C	F	88	W	Yes (*P. mirabilis*; ampicillin/sulbactam, ciprofloxacin, gentamicin)	7 days	Urinary tract infection
D	F	74	W	Yes (*P. mirabilis*; ampicillin/sulbactam, ciprofloxacin, gentamicin)	10 days	Upper respiratory tract infection
E	F	78	W	No	7 days	Upper respiratory tract infection
F	F	101	W	No	10 days	Upper respiratory tract infection
G	F	83	NW	No	11 days	Fever of unknown source
H	F	87	W	No	10 days	Upper respiratory tract infection
I	M	89	W	No	8 days	Upper respiratory tract infection
J	F	86	W	No	7 days	Fever of unknown source
K	F	91	W	No	6 days	Upper respiratory tract infection

Supplementary Table 2: Overview of Longitudinal Sample Collection from Levofloxacin Cohort from SPREAD

This table lists all samples from the levofloxacin cohort that were collected, sequenced, or analyzed in this study. Samples that were successfully analyzed are marked with a "yes", while samples that could not be collected, sequenced, or analyzed are marked with a "no". For samples that were not analyzed, a reason is also provided according to the following key: SD = subject deceased at this timepoint, NC = sample was not collected, NS = sample was not sequenced, SP = sample sequenced poorly.

Subject	T0	T3	T6	T9	T12
A	Yes	Yes	No – NS	Yes	Yes
B	Yes	Yes	No – SD	No – SD	No – SD
C	Yes	Yes	No – SD	No – SD	No – SD
D	Yes	Yes	No - SD	No - SD	No - SD
E	Yes	Yes	Yes	Yes	Yes
F	Yes	Yes	Yes	Yes	No - NS
G	Yes	Yes	No - NC	Yes	No – SD
H	Yes	No - SP	Yes	No - SD	No – SD
I	Yes	Yes	Yes	Yes	Yes
J	Yes	Yes	Yes	Yes	Yes
K	Yes	Yes	Yes	Yes	No - SD

Supplementary Table 3: Bin Selection Quality Cutoffs

This table lists the cutoffs used to determine whether a bin was "good" or "acceptable" to be used in further analysis, or "bad" enough to be discarded. Briefly, "good" bins had to meet the "good" cutoffs for all five criteria measured, "acceptable" bins could have a maximum of two "acceptable" criteria as long as all others were "good", and "bad" bins contained any criterion below the "bad" cutoffs.

Bin Type	Coarse Consistency	Complete-ness	Fine Consistency	Contamin-ation	Single PheS	Number of Criteria to be Met
Good	87%+	87%+	87%+	<10%	Yes	All 5
Acceptable	80-86.9%	80-86.9%	80-64.9%	10.1-20%	No	1-2, if all others are "good"
Bad	<80%	<80%	<80%	20%+	No	Any "bad" criterion or 3+ "acceptable" criteria

Supplementary Table 4: Bins Selected for DeepARG Analysis
This table lists all of the bins generated by PATRIC that were selected based on the criteria in Supplementary Table 3 to be analyzed using DeepARG. It includes all quality scores used to assess bin quality, as well as the PATRIC reference genome used to annotate the bin.

Read Type	Binning Strategy	Source Sample	PATRIC Score	Species Identification	Reference Genome	FC	CP	CC	CM	PheS
Single	Best	A0	1528	*Bacteroides ovatus*	28116.191	98.2	96.4	92.3	8.9	Yes
Paired	All	A0	1257	*Bacteroides thetaiotaomicron*	818.2	96	93.7	82.1	11.7	Yes
Single	Best	A0	1829	*Bacteroides vulgatus*	821.509	98.5	97.5	96.2	3.9	Yes
Single	All	A0	1978	*Bifidobacterium adolescentis*	1680.104	99.3	98.4	98.6	1.6	Yes
Single	Best	A0	1766	*Bifidobacterium bifidum*	500634.3	97.9	96.6	94.9	4.7	Yes
Single	Best	A0	1648	*Bifidobacterium longum*	1298922.3	98.8	96.6	98.6	7.8	Yes
Single	Best	A0	2017	*Escherichia coli*	562.30402	99.5	98.4	100	1.1	Yes
Single	Best	A0	1583	*Eubacterium callanderi*	53442.4	99	94.4	100	8.9	Yes
Single	Best	A3	1711	*Bacteroides ovatus*	28116.191	99	97.3	98.1	6.6	Yes
Paired	Best	A3	1311	*Bacteroides thetaiotaomicron*	818.2	96.5	94.8	80.8	10.6	Yes
Single	Best	A3	1789	*Bacteroides vulgatus*	821.509	98.5	97.7	99	5.3	Yes
Single	Best	A3	1935	*Bifidobacterium adolescentis*	1680.104	99	98.2	98.5	2.4	Yes
Paired	Best	A3	1729	*Bifidobacterium longum*	1298922.3	99.6	96.7	100	6.5	Yes
Single	Best	A3	1844	*Klebsiella oxytoca*	571.142	99.6	94.9	100	3.8	Yes
Single	All	A9	1119	*Bacteroides ovatus*	28116.191	95.6	89.6	85.3	14.2	Yes
Single	Best	A9	1590	*Bacteroides thetaiotaomicron*	818.2	98.2	95.2	92.3	7.4	Yes
Single	Best	A9	1678	*Bacteroides vulgatus*	821.509	98.7	97.1	99.5	7.5	Yes
Single	Best	A9	2029	*Bifidobacterium adolescentis*	1680.104	99.6	99	100	1	Yes
Single	Best	A9	1985	*Bifidobacterium bifidum*	500634.3	99.3	98.9	99.7	1.8	Yes
Paired	Best	A9	1672	*Bifidobacterium longum*	1298922.3	99.3	97	100	7.7	Yes
Single	Best	A12	2055	*Staphylococcus aureus*	1280.1372	100	99.6	100	0.6	Yes
Single	Best	B0	1264	*Corynebacterium urealyticum*	43771.9	96.7	92.4	99.2	14.7	Yes
Single	Best	B0	2039	*Escherichia coli*	562.23379	99.9	98.6	100	0.7	Yes
Single	Best	B0	2012	*Parabacteroides distasonis*	823.7	99.3	97.5	100	1	Yes
Single	All	B0	1806	*Methanobrevibacter smithii*	1263088.3	99.4	98	98.3	4.9	Yes
Single	All	B0	1455	*Porphyromonas asaccharolytica*	879243.3	96.5	93.4	98.7	11	Yes
Paired	Best	B0	1199	*Alistipes finegoldii*	214856.4	98.4	92.6	99.5	16.1	Yes
Single	Best	B3	1416	*Corynebacterium urealyticum*	43771.9	94.3	92.1	87.8	9.3	Yes
Paired	Best	B3	1877	*Escherichia coli*	562.17562	99.7	97	100	3.6	No
Paired	Best	B3	1770	*Fastidiosipila sanguinis*	236753.3	93.3	92.8	86.9	2.2	Yes
Single	Best	C3	1488	*Corynebacterium urealyticum*	43771.9	95.1	92.2	95.4	9.4	Yes
Single	Best	D0	1252	*[Eubacterium] eligens*	39485.21	99	94.3	100	15.5	Yes

Single	Best	D0	1891	*Bacteroides fragilis*	817.199	96.6	95.6	93.9	1.8	Yes
Single	Best	D0	1357	*Bacteroides ovatus*	28116.19	99.3	94.5	98.7	13.2	Yes
Single	Best	D0	1108	*Phascolarctobacterium faecium*	1122957.3	95.6	91.4	98.7	17.5	Yes
Single	Best	D0	2081	*Methanobrevibacter smithii*	420247.28	100	99.7	100	0.1	Yes
Paired	Best	D0	1919	*Eggerthella lenta*	84112.24	98.3	95.3	96.6	1.7	Yes
Paired	Best	D0	2072	*Escherichia coli*	562.28676	99.6	98.4	100	0	Yes
Paired	Best	D0	1604	*Parabacteroides distasonis*	823.7	97	93.3	92.3	6.7	Yes
Paired	Best	D0	2012	*Proteus mirabilis*	584.9	98.7	97	100	0.9	Yes
Paired	Best	D3	1876	*Bacteroides fragilis*	817.199	96.7	95.2	93.9	2	Yes
Paired	Best	D3	1078	*Bacteroides ovatus*	28116.19	98.7	94.4	94	17.8	Yes
Paired	Best	D3	1999	*Methanobrevibacter smithii*	420247.28	99.9	99.5	100	1.7	Yes
Paired	Best	D3	1604	*Parabacteroides distasonis*	823.7	97	93.3	92.3	6.7	Yes
Paired	Best	D3	2012	*Proteus mirabilis*	584.9	98.7	97	100	0.9	Yes
Paired	Best	D3	2072	*Escherichia coli*	562.28676	99.6	98.4	100	0	Yes
Paired	Best	D3	1919	*Eggerthella lenta*	84112.24	98.3	95.3	96.6	1.7	Yes
Single	Best	D3	1116	*Eubacterium rectale*	1263079.3	98.7	93.1	96.7	17.3	Yes
Single	Best	D3	1293	*Intestinimonas butyriciproducens*	1297617.27	93.9	89.1	88.2	11.2	Yes
Single	Best	E0	1802	*Bacteroides dorei*	997877.5	99.8	97.9	100	5.3	Yes
Single	Best	E0	2079	*Bifidobacterium breve*	1385939.3	99.5	99	100	0	Yes
Single	Best	E0	2012	*Escherichia coli*	562.22574	99.7	97.9	100	1.1	Yes
Single	Best	E0	1972	*Mobiluncus curtisii*	887899.3	96.3	95.6	100	1.4	Yes
Single	Best	E0	2017	*Peptoniphilus harei*	54005.3	97.6	96.6	100	0.7	Yes
Single	Best	E0	2012	*Proteus mirabilis*	584.293	98.8	97	100	0.9	Yes
Single	Best	E0	2047	*Pseudomonas aeruginosa*	1402503.3	99.6	98	99.4	0.3	Yes
Paired	Best	E0	1880	*Enterococcus faecalis*	1351.868	99.6	97.3	100	3.6	Yes
Single	Best	E3	2065	*Bifidobacterium breve*	1385939.3	99.6	99.1	100	0.3	Yes
Single	Best	E3	1339	*Escherichia coli*	562.22574	98.4	87.6	100	12.3	Yes
Single	Best	E3	1669	*Fastidiosipila sanguinis*	236753.3	92.9	92.5	87.6	4.3	Yes
Single	Best	E3	2072	*Lactobacillus crispatus*	47770.179	99.7	98.4	100	0	Yes
Single	Best	E3	1131	*Peptoniphilus harei*	54005.3	97.5	92.3	100	17.5	Yes
Single	Best	E6	1292	*Anaerostipes hadrus*	649756.24	97	92.2	99.3	14.1	Yes
Single	Best	E6	2034	*Bacteroides dorei*	997877.5	98.7	98.4	96.2	0	Yes
Single	Best	E6	2088	*Methanobrevibacter smithii*	420247.28	99.9	99.9	100	0	Yes
Single	Best	E6	1912	*Providencia stuartii*	588.6	98.7	94.2	98.6	2	Yes
Paired	Best	E6	2060	*Bacteroides fragilis*	817.25	98.8	97.6	99.7	0	Yes
Paired	Best	E6	1732	*Bacteroides thetaiotaomicron*	818.2	98.4	96.2	94.9	5.3	Yes
Paired	Best	E6	2032	*Escherichia coli*	562.22574	99.8	97.9	100	0.7	Yes

Paired	Best	E6	2069	*Odoribacter splanchnicus*	28118.6	99.1	98.5	99.6	0	Yes
Paired	Best	E9	2012	*Acidaminococcus intestini*	1120921.3	97.3	96.7	99.3	0.7	Yes
Paired	Best	E9	2017	*Escherichia coli*	562.22574	99.7	97.5	100	0.9	Yes
Paired	Best	E9	1866	*Bacteroides dorei*	997877.5	99.8	99	98.7	4	Yes
Paired	Best	E9	1883	*Bacteroides fragilis*	817.25	99	97.6	100	3.6	Yes
Paired	Best	E9	1755	*Bacteroides thetaiotaomicron*	818.2	98.4	97.2	95.1	5.1	Yes
Paired	Best	E9	2010	*Campylobacter hominis*	360107.7	99.7	99.3	97.3	0.9	Yes
Paired	Best	E9	1937	*Finegoldia magna*	1260.9	99.9	98.4	100	2.7	Yes
Paired	Best	E9	2056	*Proteus mirabilis*	584.293	98.7	96.9	100	0	Yes
Paired	Best	E9	1986	*Providencia stuartii*	588.6	99	96.5	100	1.3	Yes
Paired	Best	E9	2057	*Pseudomonas aeruginosa*	287.2475	99.7	98.7	99.7	0.3	Yes
Single	Best	E9	1797	*Alistipes finegoldii*	214856.4	98.9	97	100	5.2	Yes
Single	Best	E9	1704	*Campylobacter ureolyticus*	883165.3	96.5	94.5	93.9	5.3	Yes
Single	Best	E9	1800	*Peptoniphilus harei*	54005.3	97.7	95	100	4.7	Yes
Single	All	E9	1376	*Fastidiosipila sanguinis*	236753.3	93.1	91.5	87.9	10	Yes
Single	All	E9	2083	*Methanobrevibacter smithii*	420247.28	99.9	99.4	100	0	Yes
Paired	Best	E12	1980	*Odoribacter splanchnicus*	28118.6	99.5	97.7	99.6	1.6	Yes
Single	Best	E12	1758	*Akkermansia muciniphila*	239935.94	98.1	94.4	100	5.4	Yes
Single	Best	E12	1581	*Anaerostipes hadrus*	649756.24	98.4	95.3	99.3	9	Yes
Single	Best	E12	1957	*Bacteroides dorei*	997877.5	99.9	99.3	100	2.5	Yes
Single	Best	E12	1701	*Bacteroides fragilis*	817.25	99.3	97.6	98.3	6.9	Yes
Single	Best	E12	1765	*Bacteroides thetaiotaomicron*	818.2	98	95.9	93.5	4.3	Yes
Single	All	F0	1380	*Akkermansia muciniphila*	239935.96	99.2	93.2	100	12.7	No
Single	Best	F0	1618	*Anaerostipes hadrus*	649756.5	98.2	95.3	100	8.4	Yes
Single	Best	F0	1967	*Proteus mirabilis*	584.299	98.8	97	100	1.8	Yes
Single	Best	F0	1751	*Morganella morganii*	582.171	93.6	91.3	87.1	2.3	Yes
Paired	Best	F0	1906	*Campylobacter hominis*	360107.7	99.8	99.4	97.3	3	Yes
Paired	Best	F0	1707	*Corynebacterium jeikeium*	38289.3	97	94.3	100	6.4	Yes
Paired	Best	F0	1626	*Escherichia coli*	562.23379	99	93.3	100	7.8	Yes
Paired	Best	F0	2079	*Finegoldia magna*	1260.9	99.4	99	100	0	Yes
Paired	Best	F0	1950	*Parabacteroides distasonis*	823.7	99.3	96.7	99.1	1.9	Yes
Paired	Best	F0	1296	*Peptoniphilus harei*	54005.7	97.3	94	99.7	14.5	Yes
Paired	Best	F3	1474	*[Eubacterium] eligens*	39485.21	96.7	94.8	89.1	9	Yes
Paired	Best	F3	1457	*Anaerostipes hadrus*	649756.24	93.1	90.6	81	6.8	Yes
Paired	Best	F3	1845	*Bacteroides cellulosilyticus*	246787.5	97.7	95.3	92.2	2.3	Yes
Paired	Best	F3	1754	*Eubacterium rectale*	657317.3	99.3	97.4	99.3	6	Yes
Single	All	F3	1091	*Bacteroides caccae*	47678.6	95.6	88.2	83	14	Yes
Single	All	F3	1292	*Campylobacter hominis*	360107.7	99.3	95.8	97.3	14.5	Yes

Single	All	F3	1395	*Campylobacter ureolyticus*	883165.3	93.8	91.2	88.6	9.7	Yes
Single	Best	F3	1763	*Akkermansia muciniphila*	239935.96	99.2	97.6	97.5	5.5	Yes
Single	Best	F3	1310	*Alistipes finegoldii*	1263035.3	96.8	93.3	91.9	12.5	Yes
Single	Best	F3	1345	*Bacteroides fragilis*	817.199	93.6	90.5	89.4	10.7	Yes
Single	Best	F3	1195	*Bacteroides thetaiotaomicron*	818.287	98	94.5	90.5	14.8	Yes
Single	Best	F3	1655	*Corynebacterium urealyticum*	43771.9	96.9	95	99.5	7.5	Yes
Single	Best	F3	1872	*Odoribacter splanchnicus*	28118.4	99.6	97.9	100	3.9	Yes
Single	Best	F3	2030	*Parabacteroides distasonis*	823.7	98.2	95	100	0.1	Yes
Single	Best	F3	1998	*Proteus mirabilis*	584.299	99	97.1	100	1.2	Yes
Single	Best	F6	994	*Alistipes finegoldii*	1263035.3	95.9	92.4	89.7	18.2	Yes
Single	Best	F6	1418	*Anaerostipes hadrus*	649756.24	98.5	93.9	100	12.1	Yes
Single	Best	F6	1637	*Bacteroides caccae*	47678.6	94.6	91.7	81.8	3.6	Yes
Single	Best	F6	1152	*Bacteroides thetaiotaomicron*	818.287	97.6	91.8	90.7	15.1	Yes
Single	Best	F6	1048	*Faecalibacterium prausnitzii*	853.27	98.3	93.9	98	19.1	Yes
Paired	Best	F6	1391	*Bacteroides cellulosilyticus*	246787.5	97.7	93.7	91.5	10.9	Yes
Paired	Best	F6	1764	*Bacteroides fragilis*	817.199	95.8	93.8	92.2	3.6	Yes
Paired	Best	F6	2015	*Odoribacter splanchnicus*	28118.4	99.6	98.7	100	1.2	Yes
Paired	Best	F6	2056	*Parabacteroides distasonis*	823.7	99.1	96.9	100	0	Yes
Single	Best	F9	1903	*Campylobacter hominis*	360107.7	99.6	99.1	97.3	3	Yes
Single	All	F9	1093	*Proteus mirabilis*	584.299	98.1	86.1	100	16.9	Yes
Single	All	F9	1620	*Pseudomonas aeruginosa*	287.3918	98.9	93.9	99.7	8	Yes
Paired	Best	G0	1871	*Akkermansia muciniphila*	239935.85	99.9	99.2	100	4.2	Yes
Single	Best	G0	2022	*Bacteroides fragilis*	1339290.3	99.9	98.4	100	1	Yes
Single	Best	G0	1950	*Bacteroides vulgatus*	821.84	99.7	99.2	99.4	2.5	Yes
Single	Best	G0	1765	*Bifidobacterium catenulatum*	1686.6	99	97.3	100	5.9	Yes
Single	Best	G0	2049	*Escherichia coli*	562.23525	99.6	98.1	100	0.4	Yes
Single	Best	G0	1994	*Methanobrevibacter smithii*	911133.5	98.8	98.6	100	1.6	Yes
Paired	Best	G3	1448	*Finegoldia magna*	1260.9	99.6	97.1	100	12.2	Yes
Paired	Best	G3	1934	*Prevotella denticola*	28129.7	98.5	97.2	100	2.5	Yes
Single	All	G3	1670	*Akkermansia muciniphila*	239935.85	100	97.3	100	7.8	Yes
Single	Best	G3	1875	*Campylobacter hominis*	360107.7	99.3	98.4	97.3	3.4	Yes
Single	Best	G3	1950	*Lawsonella clevelandensis*	1528099.5	95.9	95.7	90.7	0	Yes
Single	Best	G3	1858	*Mobiluncus curtisii*	887899.3	95.9	94.8	100	3.5	Yes
Single	Best	G3	1765	*Negativicoccus massiliensis*	1702287.3	96.6	95	100	5.4	Yes
Paired	Best	G9	1826	*Akkermansia muciniphila*	239935.85	100	97.8	100	4.8	Yes
Paired	Best	G9	1945	*Anaerococcus prevotii*	879305.3	98.9	98.2	100	2.5	Yes
Paired	Best	G9	2009	*Campylobacter hominis*	360107.7	99.7	99.2	97.3	0.9	Yes

Paired	Best	G9	1365	*Campylobacter ureolyticus*	827.23	96.6	94	95.6	12.3	Yes
Paired	Best	G9	1550	*Corynebacterium jeikeium*	38289.26	96.7	93.6	99.5	9.3	Yes
Paired	Best	G9	1627	*Corynebacterium urealyticum*	43771.9	96.7	93.6	98.7	7.6	Yes
Paired	Best	G9	1982	*Finegoldia magna*	1260.9	99.8	99.3	100	2	Yes
Paired	Best	G9	1952	*Lawsonella clevelandensis*	1528099.5	96.1	95.9	90.7	0	Yes
Paired	Best	G9	1952	*Mobiluncus curtisii*	887899.3	96.2	95.6	100	1.8	Yes
Paired	Best	G9	942	*Peptoniphilus harei*	54005.7	95.8	91.4	93.1	19.7	Yes
Paired	Best	G9	1800	*Proteus mirabilis*	584.9	98.5	94.1	100	4.5	Yes
Paired	Best	H0	2009	*Escherichia coli*	562.28156	99.6	98.1	100	1.2	Yes
Single	All	H0	1007	*Fusobacterium nucleatum*	469603.3	98.2	90.2	99.4	19.4	Yes
Single	Best	H0	1423	*Acidaminococcus intestini*	1120921.3	97	93.7	98.7	11.7	Yes
Single	Best	H0	1268	*Aerococcus viridans*	1377.13	96.3	93	93	13.5	Yes
Single	Best	H0	1148	*Anaerococcus prevotii*	879305.3	98.8	93.4	100	17.4	Yes
Single	Best	H0	1800	*Bacteroides fragilis*	1339290.3	98.4	95.9	100	4.9	Yes
Single	Best	H0	1939	*Bacteroides vulgatus*	821.509	98.2	97	96.2	1.6	Yes
Single	Best	H0	1886	*Finegoldia magna*	1260.1	99.6	98.3	100	3.7	Yes
Single	Best	H0	2063	*Proteus mirabilis*	1125694.3	98.8	97.6	100	0	Yes
Paired	All	H6	1777	*Aerococcus viridans*	1377.13	97.2	95	93.7	3.9	Yes
Paired	All	H6	2068	*Enterococcus faecalis*	1158622.3	99.7	98	100	0	Yes
Paired	Best	H6	1734	*Escherichia coli*	562.28156	99.3	93.1	100	5.6	Yes
Single	Best	H6	1581	*Lactobacillus crispatus*	575597.3	96.2	94.3	89.9	6.9	Yes
Single	Best	H6	1566	*Staphylococcus aureus*	1280.10924	95.6	92.8	93	7.5	Yes
Single	Best	H6	1759	*Staphylococcus haemolyticus*	1283.114	94.8	93.2	95.9	4.3	Yes
Paired	Best	I0	1618	*Finegoldia magna*	1260.14	99.2	96.2	100	8.6	Yes
Paired	Best	I0	1864	*Peptoniphilus harei*	54005.7	97.4	95.8	100	3.6	Yes
Paired	Best	I0	1862	*Proteus mirabilis*	584.9	98.8	94.7	100	3.4	Yes
Paired	Best	I3	1732	*Odoribacter splanchnicus*	28118.4	99.5	96.1	100	6.3	Yes
Single	All	I3	1745	*Methanobrevibacter smithii*	2173.7	99.6	97.3	100	6.3	Yes
Single	Best	I3	1486	*Bacteroides caccae*	47678.171	97.6	94.2	89	8.6	Yes
Single	Best	I3	1945	*Campylobacter hominis*	360107.7	99.6	99.3	97.3	2.2	Yes
Single	Best	I3	2008	*Escherichia coli*	562.30949	99.6	97.9	99.6	1.1	Yes
Single	Best	I3	1554	*Parabacteroides distasonis*	823.236	99.4	94	100	9.4	No
Single	Best	I6	1659	*Collinsella aerofaciens*	74426.49	97.1	92.9	99.2	6.9	Yes
Single	Best	I6	2088	*Methanobrevibacter smithii*	2173.7	99.9	99.9	100	0	Yes
Paired	Best	I9	1991	*Campylobacter hominis*	360107.7	99.6	99.4	97.3	1.3	Yes
Single	Best	I9	1524	*Bacteroides thetaiotaomicron*	818.2	99.6	96.8	99.9	10.6	Yes
Single	Best	I9	2021	*Parabacteroides distasonis*	823.236	99.2	97.8	100	0.9	Yes

Paired	Best	I12	2081	*Odoribacter splanchnicus*	28118.4	99.9	99.2	100	0	Yes
Paired	Best	I12	1862	*Parabacteroides distasonis*	823.236	99.3	95.6	100	3.6	Yes
Paired	Best	I12	1891	*Eggerthella lenta*	84112.14	98.9	96.9	100	3.3	Yes
Paired	Best	I12	2069	*Escherichia coli*	562.23525	99.6	98.1	100	0	Yes
Single	All	I12	1188	*Phascolarctobacterium faecium*	1122957.3	94.9	88.2	98.7	15.2	Yes
Single	Best	I12	1747	*Bacteroides caccae*	47678.171	98.9	97.2	95.8	5.4	Yes
Single	Best	I12	2068	*Bacteroides fragilis*	1339290.3	99	98	100	0	Yes
Single	Best	I12	1880	*Bacteroides thetaiotaomicron*	818.2	99	97.5	96.3	2.9	Yes
Single	Best	I12	867	*Bifidobacterium bifidum*	1681.45	95.6	90.8	85.8	19.6	Yes
Single	Best	I12	2003	*Candidatus Methanomassiliicoccus*	1295009.4	97	96.2	95.5	0	Yes
Single	Best	I12	2088	*Methanobrevibacter smithii*	2173.7	99.9	99.9	100	0	Yes
Single	Best	J0	1929	*Campylobacter ureolyticus*	827.18	96.8	96.7	96.5	1.8	Yes
Single	Best	J0	1233	*Prevotella denticola*	28129.7	98.4	95.3	100	16.1	Yes
Single	Best	J0	1941	*Proteus mirabilis*	584.664	97.2	95.6	94.9	1	Yes
Single	Best	J0	1460	*Streptococcus oralis*	1303.283	99.1	95.6	99.8	11.6	Yes
Single	Best	J3	1187	*Campylobacter hominis*	360107.7	99.5	95.6	93	15.7	Yes
Single	Best	J6	1912	*Acidaminococcus intestini*	1120921.3	97.4	96.3	99.3	2.6	Yes
Single	Best	J6	1944	*Campylobacter hominis*	360107.7	99.6	99.2	97.3	2.2	Yes
Single	Best	J6	1786	*Escherichia coli*	749531.3	99.6	94.2	100	4.8	Yes
Paired	Best	J9	2007	*Campylobacter hominis*	360107.7	99.6	99	97.3	0.9	Yes
Paired	Best	J9	1157	*Corynebacterium urealyticum*	43771.9	97	93.2	99.2	17	Yes
Paired	Best	J12	1461	*Parabacteroides distasonis*	1339341.3	98.9	94.2	99	11.1	Yes
Single	All	J12	1266	*Campylobacter ureolyticus*	827.18	95.7	92.6	92.2	13.3	Yes
Single	All	J12	1286	*Collinsella aerofaciens*	74426.49	97.4	92.9	96.9	13.9	Yes
Single	Best	J12	1882	*Campylobacter hominis*	360107.7	99.5	99	97.3	3.4	Yes
Paired	Best	K0	1923	*Campylobacter hominis*	360107.7	99.6	99.1	97.3	2.6	Yes
Single	Best	K0	1611	*Peptoniphilus harei*	54005.7	96.2	94.2	95.5	7.4	Yes
Paired	Best	K3	1083	*Bacteroides vulgatus*	821.509	99.6	93.9	100	18.8	Yes
Paired	Best	K3	1891	*Bifidobacterium bifidum*	1681.55	99.3	98.1	99.7	3.5	Yes
Paired	Best	K3	1723	*Bifidobacterium longum*	216816.147	99.4	97.9	98.6	6.6	Yes
Single	Best	K3	2002	*Candidatus Methanomassiliicoccus*	1295009.4	97	96.1	95.5	0	Yes
Single	Best	K3	2030	*Escherichia coli*	562.28156	99.9	98.1	99.6	0.7	Yes
Single	Best	K3	1575	*Eubacterium callanderi*	53442.4	98.9	96.4	98	9.1	Yes
Single	Best	K3	2088	*Methanobrevibacter smithii*	2173.71	99.9	99.9	100	0	Yes
Single	Best	K6	1898	*Bacteroides caccae*	47678.175	98.7	97.4	96.7	2.6	Yes

Single	Best	K6	1976	*Odoribacter splanchnicus*	28118.38	99.7	98.8	100	2	Yes
Single	Best	K9	1725	*Campylobacter hominis*	360107.7	99.3	97.9	97.3	6.3	Yes
Single	Best	K9	1375	*Campylobacter ureolyticus*	827.23	95.6	92.8	93.9	11.5	Yes

Supplementary Table 5: BioProject Sample Identifiers for Test Dataset

This table lists the sample names used in this study, the SPREAD IDs, and the BioProject PRJNA531921 sample names for the shotgun metagenomics sequencing files of the 67-sample dataset used to test the multiple linear regression developed from the levofloxacin dataset.

Sample	SPREAD ID	BioSample ID	Sample	SPREAD ID	BioSample ID
1_A2	02/007/3/6/R	S02_007_3_6_R	**2_E5**	10/010/6/6/R	S10_010_6_6_R
1_A8	02/021/6/8/R	S02_021_6_8_R	**2_F1**	10/012/6/6/R	S10_012_6_6_R
1_A11	02/023/3/6/R	S02_023_3_6_R	**2_F11**	13/030/3/6/R	S13_030_3_6_R
1_B3	02/032/B/7/R	S02_032_B_7_R	**2_G6**	13/035/12/7/R	S13_035_12_7_R
1_B5	02/041/3/7/R	S02_041_3_7_R	**2_G9**	13/080/6/7/R	S13_080_6_7_R
1_B11	04/003/6/5/R	S04_003_6_5_R	**2_H2**	19/009/6/5/R	S19_009_6_5_R
1_C6	04/011/12/5/R	S04_011_12_5_R	**2_H4**	19/031/B/5/R	S19_031_B_5_R
1_D2	04/059/6/9/R	S04_059_6_9_R	**2_H10**	21/012/12/7/R	S21_012_12_7_R
1_D8	06/007/9/6/R	S06_007_9_6_R	**3_A7**	21/037/6/7/R	S21_037_6_7_R
1_D12	06/027/6/6/R	S06_027_6_6_R	**3_A12**	21/060/12/7/R	S21_060_12_7_R
1_E4	06/040/6/7/R	S06_040_6_7_R	**3_B4**	23/025/9/9/R	S23_025_9_9_R
1_E9	06/048/9/5/R	S06_048_9_5_R	**3_B8**	26/031/6/5/R	S26_031_6_5_R
1_F1	06/060/6/5/R	S06_060_6_5_R	**3_B11**	26/038/3/9/R	S26_038_3_9_R
1_F8	06/068/6/5/R	S06_068_6_5_R	**3_C2**	29/013/6/9/R	S29_013_6_9_R
1_F10	06/071/B/6/R	S06_071_B_6_R	**3_C4**	31/039/B/7/R	S31_039_B_7_R
1_G3	06/083/9/9/R	S06_083_9_9_R	**3_C8**	32/019/6/5/R	S32_019_6_5_R
1_G9	06/085/12/9/R	S06_085_12_9_R	**3_D2**	32/022/12/5/R	S32_022_12_5_R
1_G10	06/102/B/9/R	S06_102_B_9_R	**3_D8**	32/052/9/9/R	S32_052_9_9_R
1_H3	06/107/B/9/R	S06_107_B_9_R	**3_D10**	34/009/B/2/R	S34_009_B_2_R
1_H6	06/108/3/9/R	S06_108_3_9_R	**3_E4**	35/010/B/5/R	S35_010_B_5_R
2_A2	07/020/3/7/R	S07_020_3_7_R	**3_E6**	35/031/B/9/R	S35_031_B_9_R
2_A6	07/056/3/7/R	S07_056_3_7_R	**3_E8**	36/007/B/7/R	S36_007_B_7_R
2_B2	07/059/6/7/R	S07_059_6_7_R	**3_F4**	38/001/3/5/R	S38_001_3_5_R
2_B4	09/018/B/6/R	S09_018_B_6_R	**3_F6**	38/004/3/5/R	S38_004_3_5_R
2_B11	09/048/9/5/R	S09_048_9_5_R	**3_F10**	38/017/B/5/R	S38_017_B_5_R
2_C1	09/085/3/5/R	S09_085_3_5_R	**3_G3**	38/024/9/9/R	S38_024_9_9_R
2_C2	09/086/B/5/R	S09_086_B_5_R	**3_G7**	39/008/6/7/R	S39_008_6_7_R
2_C5	09/099/9/9/R	S09_099_9_9_R	**3_G11**	39/011/6/7/R	S39_011_6_7_R
2_C9	09/138/9/9/R	S09_138_9_9_R	**3_H6**	40/038/9/9/R	S40_038_9_9_R
2_C12	09/143/9/9/R	S09_143_9_9_R	**4_A3**	40/044/6/9/R	S40_044_6_9_R
2_D2	09/153/3/9/R	S09_153_3_9_R	**4_A12**	42/002/9/7R/2	S42_002_9_7R_2
2_D6	09/187/3/9/R	S09_187_3_9_R	**4_B4**	42/014/6/7/R	S42_014_6_7_R
2_D10	09/192/6/9/R	S09_192_6_9_R	**4_B7**	42/015/3/7/R	S42_015_3_7_R
2_E1	09/214/6/9/R	S09_214_6_9_R			

Supplementary Data 1: Taxonomic Classifications from Shotgun Metagenomics Sequencing with Kraken2 and Bracken2 Analysis

This file includes the relative abundances of the taxonomic classifications at the phylum, genus, and species level for both the initial levofloxacin-treated dataset (tabs 1, 2, and 3) and the second, larger test dataset (tabs 4, 5, and 6). Supplementary Table 5 links the sample names used for the test dataset in this study with their identifiers in BioProject PRJNA531921.

This file is available at the Brown Digital Repository at https://doi.org/10.26300/d74v-de34.

Supplementary Data 2: Taxonomic Classifications from 16S rRNA Sequencing with QIIME2 Analysis

This file includes the relative abundances of the taxonomic classifications at the phylum (tab 1) and genus (tab 2) level for the initial levofloxacin-treated dataset.

This file is available at the Brown Digital Repository at https://doi.org/10.26300/kd0n-8561.

Supplementary Data 3: Metagenomic Classifications from Shotgun Metagenomics Sequencing with HUMAnN2 Analysis

This file includes the relative abundances of the MetaCyc pathway (tab 1), KEGG ortholog (tab 2), and GO term (tab 3) outputs for the initial levofloxacin-treated dataset.

This file is available at the Brown Digital Repository at https://doi.org/10.26300/maj5-b696.

Supplementary Data 4: Antimicrobial Resistance Gene Profiles from Shotgun Metagenomics Sequencing with DeepARG Analysis

This file includes the relative abundances of antimicrobial class and specific resistance genes for the initial levofloxacin-treated dataset (tabs 1 and 2), the second, larger test dataset (tabs 3 and 4), and the bins generated by PATRIC (tabs 5 and 6). Supplementary Table 5 links the sample names used for the test dataset in this study with their identifiers in BioProject PRJNA531921.

This file is available at the Brown Digital Repository at https://doi.org/10.26300/j66h-1p28.

Supplementary File 1: Code Used for Analyses

This file includes all of the analysis code used for QIIME2, Kraken2 and Bracken, Phyloseq, HUMAnN2, and DeepARG.

This file is available at the Brown Digital Repository at https://doi.org/10.26300/yhyv-5b03.

All Supplementary Data files for this thesis can be found at the Brown Digital Repository at https://doi.org/10.26300/enej-vt18.

DISCUSSION: THE HUMAN GUT MICROBIOTA ACROSS DISEASE AND GEOGRAPHY

CONTENTS

The Gut Microbiota Can Interact with Extraintestinal Disorders 177

Signatures of Microbiota Disturbance and Inflammation Across Studies 179

Distinct Microbiota Composition in African and American Subjects 182

Unanswered Questions and Future Directions 184

Conclusions 190

Main Figure 192

The Gut Microbiota Can Interact with Extraintestinal Disorders

These chapters describe three quite distinct projects, focusing on the human gut microbiota in very dissimilar disease states and disruptions. At first glance, studies about a neglected tropical parasitic disease, a mental health disorder, and antimicrobial administration in the elderly may not seem to have very much in common. However, together they demonstrate an important point about the myriad bacteria that live in the human gastrointestinal tract: they do not only have local impacts on the gut but can potentially influence or be influenced by perturbations far from their domicile.

First, in the case of urogenital schistosomiasis, we observed that Nigerian adolescents infected with the parasitic trematode *S. haematobium* had clear differences in their gut microbiota relative to their healthy peers – despite the fact that this helminth resides in the venules surrounding the urogenital system, and has no clear mechanism through which to directly interact with the gut microbiota. We observed differences between *S. haematobium*-infected and control adolescents across all taxonomic levels, including an overall schistosomiasis-associated reduction in the Firmicutes-Clostridia-Clostridiales lineage, including the genus-level taxa {*Ruminococccaceae*} *incertae sedis* and *Subdoligranulum*. Contrarily, we observed that the adolescents with schistosomiasis were enriched in the phylum Proteobacteria and a number of specific taxa within it, including the sulfate-reducing *Desulfovibrio* lineage and several lineages of Gammaproteobacteria. While we cannot demonstrate causality in such a study, we hypothesize that systemic immunological alterations that occur during schistosome infection – particularly T_H2 polarization and immunoregulation[199,205,209,212,213,671]– disrupt

typical gut immune-microbe homeostasis; based on our observations, this may allow a shifting of the gut microbiota towards a more locally pro-inflammatory state.

Second, we also observed differences in the gut microbiota of American young adults with major depressive disorder when compared to their healthy peers. Despite the physiological distance between the gastrointestinal tract and the brain, a significant body of research implicates the gut microbiota in a range of mental health disorders[336-340,343,374,381,382]. Reflecting previous research on gut bacteria and systemic inflammation in MDD, we found that the microbiomes of young adults with MDD were distinct from those of controls and seemed to suggest a loss of protective, anti-inflammatory species. Specifically, subjects with MDD exhibited lower levels of the butyrate-producing *Ruminococcaceae* genera *Faecalibacterium* and *Subdoligranulum*, among others, alongside higher levels of potentially pro-inflammatory *Flavonifractor* and Gammaproteobacteria. As with our study of schistosomiasis, we cannot determine the directionality of causation, although a pro-inflammatory microbiome may contribute to the low-grade, chronic inflammation that is thought to contribute to some manifestations of depression[389-391,397,427,435,450].

Finally, we found that the gut microbiota of an elderly population with advanced dementia was extremely temporally unstable and characterized by blooms of pathobionts carrying resistance genes. In particular, we observed frequent blooms of the Proteobacteria species *E. coli* and *P. mirabilis* and the Firmicutes species *E. faecalis*, which correlated with overall levels of antimicrobial resistance genes; we further observed that the Proteobacteria species *P. stuartii* and the Firmicutes species *S. haemolyticus* significantly contributed to resistance gene burden in specific samples. Microbial and resistome

composition could not be linked to antimicrobial administration and may instead represent the impacts of advanced age and institutionalization, including immune system decline, repetitive diet, high rates of medication and hospitalization, and increased contact with care workers.

Signatures of Microbiota Disturbance and Inflammation Across Studies

In all three cases, we observed perturbations consistent with a disturbed microbial ecosystem and potential inflammation. In the case of the elderly subjects, for whom we had longitudinal data, their temporally-unstable microbiota was often overtaken by Proteobacteria – to the point that it was the second-most abundant phylum across all samples. These Proteobacteria species were generally pathogens or pathobionts, possibly reflecting or contributing to the chronic "inflammaging" that occurs in the elderly[672-676]. Additionally, the phylum Firmicutes made up a much lesser proportion of the gut microbiota than is typical for American adults, and Bacteroidetes was the most-abundant phylum across all samples. Increased levels of Bacteroidetes has previously been observed in the elderly[464,555] and may represent general effects of aging on the gut microbiome. However, as this study included only subjects who were institutionalized with advanced dementia, we cannot determine how their microbiota composition compares to that of community-resident elderly adults or nursing-home residents without advanced dementia.

In contrast, for our studies on urogenital schistosomiasis and major depressive disorder, we were able to compare the gut microbiota between subjects with and without active disease. In both cases, we observed trends that are suggestive of a shift to a more pro-inflammatory state in the disease cohorts, including a loss of potentially protective bacterial lineages and increased abundance of potentially pro-inflammatory bacteria. In

particular, both disease groups exhibited a reduction in the Firmicutes-Clostridia-Clostridiales lineage relative to their control groups, with concomitant (although non-significant) increases in the Bacteroidetes-Bacteroidia-Bacteroidales lineage. This is notable given that members of Clostridia are known to produce butyrate and provide an anti-inflammatory influence in the gut. Indeed, in both studies, members of the Clostridiales family *Ruminococcaceae*, which contains many butyrate-producing taxa[456,466,677], were reduced in the disease cohorts; in particular, the genus *Subdoligranulum*, which contains the single named species *S. variabile*, was reduced in both groups. In the MDD cohort, the *Ruminococcaceae* genus *Flavonifractor* was instead enriched, but this genus (and its sole named species *F. plautii*) has been associated with a host of autoimmune, gastrointestinal, and mental health disorders[328,332,340,496,497], suggesting that it may have pro-inflammatory potential.

Additionally, both groups demonstrated increases in their levels of certain Proteobacteria, although the specific differences were dissimilar. In particular, the subjects with urogenital schistosomiasis displayed increased levels of the phylum Proteobacteria overall, including the *Moraxellaceae-Acinetobacter* and *Pasteurellaceae-Haemophilus* lineages of Gammaproteobacteria. Both *Acinetobacter* and *Haemophilus* contain potentially pathogenic species that may promote inflammation[678-682]. While the subjects with MDD did not display increases in these bacterial genera, they did show increases in the class Gammaproteobacteria that were driven by increases in the *Burkholderiaceae-Parasutterella* lineage, which has been associated with inflammatory diseases of the gut as well as MDD[343,683-685].

The subjects with schistosomiasis also had increases in the *Desulfovibrionaceae-Desulfovibrio* lineage of Deltaproteobacteria, which has been previously linked to gastrointestinal inflammation through its production of the toxic gas hydrogen sulfide[297,299,300,509]. Interestingly, *Desulfovibrio* was instead found to be reduced in subjects with MDD, which may seem counter to the largely pro-inflammatory changes observed in this population. This disparity may stem from underlying differences in the communities. The control Nigerian subjects had extremely low levels of this genus, which expanded significantly in the infected cohort from an average of 0.0023% to an average of 0.11% of the microbial population; such an expansion could significantly increase hydrogen sulfide production and contribute to disruption in an ecosystem not adapted to levels of this microbe or molecule In contrast, the American control subjects began with an average abundance of 0.12% *Desulfovibrio*, which reduced to 0.046% in the MDD cohort; in this case, the American gut ecosystem may be more adapted for the presence of *Desulfovibrio*, perhaps due to dietary differences, and the loss of its sulfate-reducing capacity may be problematic in some gut communities. However, it should be noted that in all four populations, some subjects had undetectable levels of this genus at our sequencing depth, underscoring the fact that there is significant intra-subject heterogeneity in community composition and membership.

Distinct Microbiota Composition in Nigerian and American Subjects

As suggested by the underlying differences in *Desulfovibrio*, the fact that we observed any commonalities between these two studies is actually fairly notable in light of the fact that the microbiota composition of the two cohorts is extremely divergent across all taxonomic levels. Significant previous research has demonstrated that gut microbial

populations can differ significantly by geography, especially between industrialized and non-industrialized regions, likely due to differences in diet and lifestyle[90,460-463,686]. In our studies, the Nigerian control adolescents appear to have a generally more diverse gut microbiome, with higher levels of minor phyla including Proteobacteria, Actinobacteria, Tenericutes, and Cyanobacteria than the American control young adults (Figure 1A). In fact, while a direct comparison cannot be made due to differences in sample collection and preparation and changes made to the Silva taxonomy database between the two studies, it is potentially telling that the control Nigerian subjects had an average observed ASVs count of 211 while the control American subjects had an average of only 140, suggesting a greater richness of bacterial taxa in the African cohort.

In addition to differential abundances of phyla overall, the composition of the major phyla Firmicutes and Bacteroidetes is quite different in the two populations. Reflecting previous work comparing African and western cohorts, the predominant member of Bacteroidetes in the Nigerian subjects was the genus *Prevotella*, while it was *Bacteroides* in the American subjects. These two genera exist on a gradient, but often within a particular sample one will predominate to the exclusion of the other; *Prevotella* tends to thrive in subjects living in rural areas and consuming a high-fiber diet, while *Bacteroides* instead proliferates in subjects consuming more animal protein and fats[49,90,460-463]. Consistent with the fact that Americans tend to consume less fiber[687,688], we observed that while there were a few subjects in the MDD study who had *Prevotella*-dominated microbiota, they made up a minority of the overall cohort (Figure 1B). Interestingly, in the Nigerian subjects, levels of *Prevotella* were significantly increased in the disease cohort, perhaps suggesting that

endemic helminth infections could further contribute to high levels of this taxon in rural African populations.

Similarly, the proportions of members of Firmicutes were quite distinct between these two populations. While the family *Ruminococcaceae* was the largest family in both groups, the abundance was notably higher in the American cohort than in the Nigerian subjects (32.9% and 23.7%, respectively). Furthermore, the second-most abundant family was *Clostridiaceae 1* in the Nigerian cohort at 18.5%, with *Lachnospiraceae* making up a distant third at 7.7%; in contrast, *Lachnospiraceae* was the second-most abundant family in the American subjects at 22.1%, while *Clostridiacaeae 1* made up a very small proportion of the community at 0.11%. Drilling down further to the genus level, in the American cohort, the most abundant Firmicutes were three members of *Ruminococcaceae* (*Faecalibacterium* [20.3%], *Subdoligranulum* [2.9%], and *Ruminococcus 1* [1.8%]) and two members of *Lachnospiraceae* (*Blautia* [4.6%] and *Agathobacter* [3.9%]) (Figure 1B). In the Nigerian cohort, the most abundant genera were two members of the family *Clostridiacaeae 1* (*Clostridium sensu stricto 1* [11.3%] and *Sarcina* [7.2%]) and three members of *Ruminococcaceae* (uncultured *Ruminococcaceae* [8.7%], *Subdoligranulum* [5.1%], and *Faecalibacterium* [4.4%]) (Figure 1B).

These quite substantial differences in community composition between our American and Nigerian cohorts demonstrate the importance of considering context in microbiome studies. Changes in the relative abundance of specific microbes may not be replicable across different cohorts, and there may even be cases where taxa change in opposite directions in the same disease model based on the underlying microbial community. Therefore, it is important to conduct microbiome studies in a wide range of

cohorts, to account for underlying variability that may impact results. For example, MDD-related changes in the gut microbiota of North Americans may not resemble MDD-related changes in an African cohort. While this does present a challenge, particularly in recruiting subjects from rural and less-developed regions, it also presents an opportunity to find common mechanisms driving disease-related changes across different bacterial taxa.

Unanswered Questions and Future Directions

While our studies on schistosomiasis and MDD provide an important examination of the gut microbiota in various human disease states, they have two drawbacks: they are observational, so they cannot demonstrate causality, and they utilize the 16S rRNA marker gene to infer community composition, so they cannot demonstrate functional changes. Therefore, while both interesting and informative, they leave a clear path for future work to take to further understand the links between these disorders and the gut microbiota.

Regarding our study on urogenital schistosomiasis, there are a few clear next steps for future work. First and foremost, the issue of causality should be addressed, which can be approached in multiple ways. The standard methodology would be to perform a prospective cohort study, in which a large sample of initially-uninfected subjects in a *S. haematobium*-endemic area periodically contribute fecal samples for microbiome analysis and are tested for active urogenital schistosomiasis. This would allow discernment of whether microbiota differences follow schistosomiasis, suggesting that active infection leads to microbiome alterations. Alternatively, microbiota differences might precede schistosomiasis, suggesting that they in some way make subjects more vulnerable to successful infection. However, this type of study would be both expensive and difficult to implement, given the need to follow a large group over time without knowing ahead of

time how many will become infected. An alternative approach is to identify a cohort of subjects with active infections, treat them with the antihelminthic drug praziquantel to eradicate their schistosomiasis, and track their microbiota over time to see whether the community composition returns to a baseline state. While not as robust as a prospective cohort study, this would provide an indication of whether active infection is the driver of observed microbiome differences. Additionally, this approach could also be implemented alongside a prospective study, by administering praziquantel to subjects who become infected and retaining them in the cohort after treatment.

Regardless of causality, further work would be needed to understand the mechanistic link between infection and changes to the gut microbiota, as well as whether the microbiome shifts have a discernible impact on gastrointestinal health. Unfortunately, there is not a standard murine model for *S. haematobium* infection that fully recapitulates human disease, making mechanistic investigations more difficult. However, some aspects of human disease can be modeled by injecting schistosome eggs into the murine urogenital system[689,690], which could provide a basis for study of how egg deposition and associated immune responses specifically impact the gut microbiota.

Additionally, human studies are also possible through utilization of non-invasive methods based on fecal and blood sampling to assess gastrointestinal inflammation, gut barrier dysfunction, circulating cytokine levels, and T_H cell polarization. In particular, it would be interesting to link microbiome shifts to preceding changes in systemic or gastrointestinal immunity, and to identify whether the potentially dysbiotic changes in the gut community actually disrupt barrier integrity or cause localized inflammation. Additionally, including metagenomic, metatranscriptomic, and metabolomic analyses of

the fecal microbiota will also be important to understand how the taxonomic changes that we observed relate to functional alterations of the gut community.

As for schistosomiasis, additional work is required to understand potential causation regarding links between MDD and the gut microbiota. Again, the ideal scenario would be to design and implement a prospective cohort study, to identify when microbiota changes and/or inflammatory responses occur in relation to the onset of depressive symptoms. This is perhaps even more difficult to do than for schistosomiasis, however: as MDD is not an infectious disease and a study can't be geographically targeted to a region where the etiological agent is endemic, it would likely require enrollment and retention of a very large number of subjects. Partnerships with initiatives such as the National Institute of Health's All of Us Program, which aims to recruit one-million diverse Americans to allow for the study of disease across a range of cohort characteristics[691,692], may help, but it is still a relatively daunting prospect. There are also other complications, including the fact that subjects with MDD are likely to be taking psychotropic medications, which may themselves impact the gut microbiota in some cases[693-695], and that MDD cannot be simply cured like an infectious disease to examine potential returns to baseline.

There are some alternative approaches that might be more practical in the absence of resources to undertake such a prospective study. First, there are some murine models used to induce a depression-like phenotype, including learned helplessness, repeated social defeat, and chronic mild stress protocols[696-703], although assessing a mental health disorder in animals is somewhat difficult and subjective. However, given that previous work using fecal transplants has demonstrated that depression-associated microbiota may be causative[345,385], these models may be useful to test the contributions of particular microbes

to the onset of symptoms in the face of a precipitating stressful event. For example, germ-free mice could be colonized with a neutral defined microbial consortium (such as altered Schaedler flora) with potentially protective (such as *F. prausnitzii*) or potentially pathogenic (such as *F. plautii*) species added. These mice could then be subjected to a depression-induction protocol and the onset of depressive symptoms (anhedonia, lethargy, loss of appetite, anxiety behaviors, despair behaviors, etc.) then compared between mice colonized with neutral, protective, or pathogenic consortia. Furthermore, analysis of inflammatory markers, gut barrier function, and microbial metatranscriptomes and metabolomes could reveal information about potential mechanisms by which gut microbes interact with the immune system and central nervous system.

Additionally, there are a number of directions for non-prospective human studies. First, and most simply, future studies of the gut microbiota in MDD should move towards incorporation of metagenomic, metatranscriptomic, and metabolomic analysis. This will allow a more detailed interrogation of the gut microbial community in the context of depression and allow researchers to answer questions regarding how the gut microbiota might contribute to protection against or promotion of inflammation and depressive symptoms. Second, our work demonstrated that severity of depression may be related to the degree of alterations in the gut microbiota, but this effect could not be fully disentangled from the overlapping influence of psychotropic medication. Future studies focusing on new-onset, treatment-naïve MDD subjects could help to separate these two effects, as could a study focused on recruitment of sufficient numbers of treated and treatment-naïve subjects with mild, moderate, and severe symptoms to compare the influence of severity in the presence and absence of treatment. Finally, studies of the microbiome in depression

have been primarily carried out in populations of primarily Asian and European descent[343-351,385], and future work should strive to include other ethnic groups which might have distinct microbial responses in the context of MDD.

Our study of the gut microbiota in an elderly population with advanced dementia was quite distinct from our other two studies and leaves a number of avenues for future work. First, while we analyzed the potential impacts of antimicrobial administration on the gut microbiota of the elderly, the underlying study from which we took the data[549] was not designed for this purpose. In particular, samples were collected at pre-determined timepoints and were not selected in relation to antimicrobial administration, and therefore there were often weeks between administration and sample collection. To better get at this question, a future study could instead recruit subjects who were going to be administered a given antimicrobial, take a sample before the drug was given, and continue taking samples periodically for a few weeks afterwards; this would allow a more targeted study of how a particular antimicrobial would disrupt the microbiota of these subjects, and how quickly they might recover. Additionally, given the inter- and intra-subject variability we observed, a larger sample size would be warranted.

Additionally, our study found that our subject population had a high degree of temporal variability in their gut microbiota over the course of a year, with frequent blooms of various pathobionts that frequently carried high levels of antimicrobial resistance genes. In future work, it would be relevant to compare our population of institutionalized subjects with advanced dementia to other elderly populations, including community-resident elderly, institutionalized elderly without dementia diagnoses, and elderly with milder dementia. Such comparisons would help to identify whether the instability and high rate of

pathobiont expansion found in this population is linked primarily to advanced age, nursing home residence, high need for contact with care workers, or other factors.

Finally, our study uncovered a link between the overall abundance of antimicrobial resistance genes in the gut microbiota and expansion of specific pathobionts, which carried high densities of these genes based on metagenomic assembly. However, this link may only have been apparent due to the fact that these pathobionts reached high levels in a large number of samples in both our initial levofloxacin cohort and our larger confirmatory cohort, while in a younger, healthy population this might rarely occur. It would be interesting to first replicate this finding in other cohorts of elderly subjects with advanced dementia who might be likely to have a similar rate of pathobiont blooms; while we had subjects from a range of nursing homes, all were located in the same geographic area, making generalizability an outstanding question. Furthermore, it would be helpful to identify if elderly subjects naturally carry strains of pathobionts with particularly high resistance gene densities, if advanced dementia patients in nursing homes are more prone to novel acquisition of high-resistance strains, or if any population would have a similar trend if their pathobiont levels were high enough. To answer this question, it would be useful to compare the antimicrobial resistance gene density of pathobiont isolates from both younger adults and community-resident elderly to those of nursing home residents. Finally, while we were able to leverage deep metagenomic sequencing to examine resistance genes at the community and species levels, we cannot comment on expression of these genes or actual resistance phenotypes and there may be significant heterogeneity at the strain level. Metatranscriptomics and *in vitro* strain characterization would be helpful

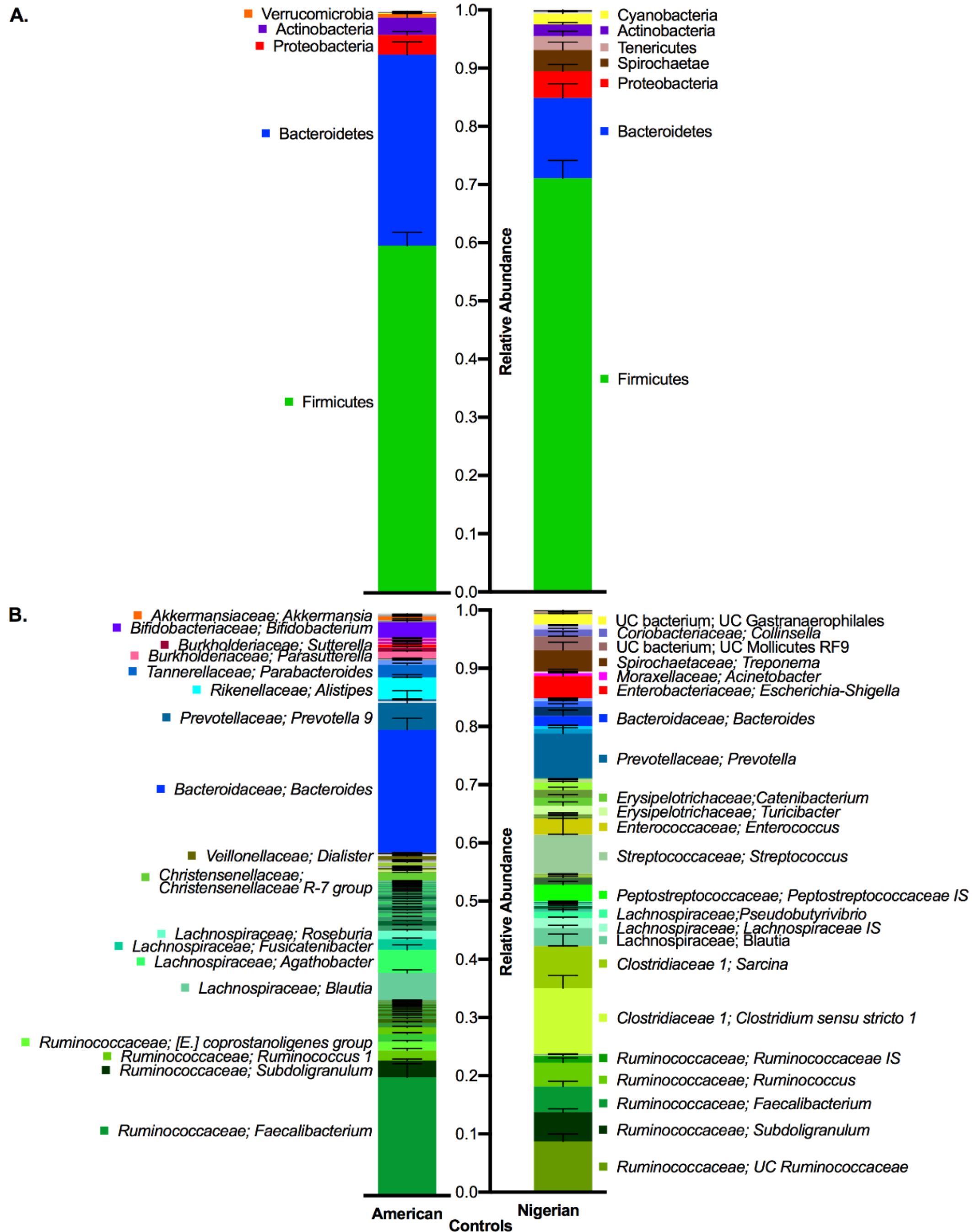
in understanding whether these communities or specific pathobiont strains have the capacity to resist antimicrobial treatment.

Summary

Despite their breadth of subject matter, these three studies together illustrate some important points. First, the human gut microbiota is not only relevant in gastrointestinal diseases, but instead can be linked to parasitic infection in the urogenital vasculature, a mental health disorder characterized by neurotransmitter disruptions in the brain, and potentially to the general inflammaging immune disruption that occurs in the elderly. Second, in all three of these cases, there are indications of enrichment in pro-inflammatory bacteria and/or reductions in protective bacteria, suggesting a promotion of gastrointestinal inflammation and reinforcing the fact that disruptions to the gut microbiota can have important consequences for host health; in the case of major depressive disorder, this inflammation may actually contribute to the development of disease. Third, these studies demonstrate the importance of considering background and context in the study of the human microbiota. It is clear that the fecal microbiota of American young adults is highly distinct from that of Nigerian adolescents, likely due to significant dietary and lifestyle differences. This reinforces the importance of considering underlying microbiome differences when interpreting the impacts of disease or disruption on the microbiota, as well as the fact that what our conceptions of a “normal” microbiome should not be wholly based on studies of western populations. Additionally, even within American subjects, there were quite distinct differences between young adult and elderly subjects, even allowing for differences between fecal and rectal microbiota. Finally, in all of these cases,

there is more work to be done, particularly in uncovering the mechanisms of interaction between extraintestinal manifestations of disease and the human gut microbiota.

Figure 1: Comparison of Taxonomic Composition of American and Nigerian Control Cohorts
(A) Phylum-level average relative abundance bar plots for the control cohorts of the Nigerian subjects from Chapter One's study on urogenital schistosomiasis and the American subjects from Chapter Two's study on major depressive disorder. (B) Genus-level average relative abundance bar plots for the control cohorts of the Nigerian subjects from Chapter One's study on urogenital schistosomiasis and the American subjects from Chapter Two's study on major depressive disorder. *In A, taxa are stacked in order of relative abundance. In B, taxa are stacked in order of relative abundance within their higher taxa to demonstrate the phylogenetic relationships between genera. Due to updates in the Silva database between analyses, there are some categorization differences, particularly the splitting of the genera* Ruminococcus *and* Prevotella *into sub-groups (*Ruminococcus 1, Prevotella 9, *etc.)*



REFERENCES

1 Lloyd-Price, J., Abu-Ali, G. & Huttenhower, C. The healthy human microbiome. *Genome Med* **8**, 51, doi:10.1186/s13073-016-0307-y (2016).

2 Kim, Y., Koh, I. & Rho, M. Deciphering the human microbiome using next-generation sequencing data and bioinformatics approaches. *Methods* **79-80**, 52-59, doi:10.1016/j.ymeth.2014.10.022 (2015).

3 Kuczynski, J. *et al.* Experimental and analytical tools for studying the human microbiome. *Nat Rev Genet* **13**, 47-58, doi:10.1038/nrg3129 (2011).

4 Bikel, S. *et al.* Combining metagenomics, metatranscriptomics and viromics to explore novel microbial interactions: towards a systems-level understanding of human microbiome. *Comput Struct Biotechnol J* **13**, 390-401, doi:10.1016/j.csbj.2015.06.001 (2015).

5 Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**, 335-336, doi:10.1038/nmeth.f.303 (2010).

6 Cole, J. R. *et al.* Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* **42**, D633-642, doi:10.1093/nar/gkt1244 (2014).

7 Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**, D590-596, doi:10.1093/nar/gks1219 (2013).

8 DeSantis, T. Z. *et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* **72**, 5069-5072, doi:10.1128/AEM.03006-05 (2006).

9 Weisburg, W. G., Barns, S. M., Pelletier, D. A. & Lane, D. J. 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* **173**, 697-703 (1991).

10 Case, R. J. *et al.* Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. *Appl Environ Microbiol* **73**, 278-288, doi:10.1128/AEM.01177-06 (2007).

11 Lane, D. J. *et al.* Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proc Natl Acad Sci U S A* **82**, 6955-6959 (1985).

12 Iwai, S. *et al.* Piphillin: Improved Prediction of Metagenomic Content by Direct Inference from Human Microbiomes. *PLoS One* **11**, e0166104, doi:10.1371/journal.pone.0166104 (2016).

13 Langille, M. G. *et al.* Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* **31**, 814-821, doi:10.1038/nbt.2676 (2013).

14 Koskinen, K. *et al.* First Insights into the Diverse Human Archaeome: Specific Detection of Archaea in the Gastrointestinal Tract, Lung, and Nose and on Skin. *MBio* **8**, doi:10.1128/mBio.00824-17 (2017).

15 Jovel, J. *et al.* Characterization of the Gut Microbiome Using 16S or Shotgun Metagenomics. *Front Microbiol* **7**, 459, doi:10.3389/fmicb.2016.00459 (2016).

16 Segata, N. *et al.* Computational meta'omics for microbial community studies. *Mol Syst Biol* **9**, 666, doi:10.1038/msb.2013.22 (2013).

17 Bashiardes, S., Zilberman-Schapira, G. & Elinav, E. Use of Metatranscriptomics in Microbiome Research. *Bioinform Biol Insights* **10**, 19-25, doi:10.4137/BBI.S34610 (2016).
18 Turnbaugh, P. J. & Gordon, J. I. An invitation to the marriage of metagenomics and metabolomics. *Cell* **134**, 708-713, doi:10.1016/j.cell.2008.08.025 (2008).
19 Aguiar-Pulido, V. *et al.* Metagenomics, Metatranscriptomics, and Metabolomics Approaches for Microbiome Analysis. *Evol Bioinform Online* **12**, 5-16, doi:10.4137/EBO.S36436 (2016).
20 Barczak, A. K. *et al.* RNA signatures allow rapid identification of pathogens and antibiotic susceptibilities. *Proc Natl Acad Sci U S A* **109**, 6217-6222, doi:10.1073/pnas.1119540109 (2012).
21 Grahl, N. *et al.* Profiling of Bacterial and Fungal Microbial Communities in Cystic Fibrosis Sputum Using RNA. *mSphere* **3**, doi:10.1128/mSphere.00292-18 (2018).
22 Breitweiser, F. P., Lu, J. & Salzberg, S. L. A review of methods and databases for metagenomic classification and assembly. *Brief Bioinform* **bbx120**, doi:https://doi.org/10.1093/bib/bbx120 (2017).
23 Hugenholtz, P. Exploring prokaryotic diversity in the genomic era. *Genome Biol* **3**, REVIEWS0003 (2002).
24 Sender, R., Fuchs, S. & Milo, R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol* **14**, e1002533, doi:10.1371/journal.pbio.1002533 (2016).
25 Thursby, E. & Juge, N. Introduction to the human gut microbiota. *Biochem J* **474**, 1823-1836, doi:10.1042/BCJ20160510 (2017).
26 Roediger, W. E. Anaerobic bacteria, the colon and colitis. *Aust N Z J Surg* **50**, 73-75 (1980).
27 Human Microbiome Project, C. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207-214, doi:10.1038/nature11234 (2012).
28 Faith, J. J. *et al.* The long-term stability of the human gut microbiota. *Science* **341**, 1237439, doi:10.1126/science.1237439 (2013).
29 Backhed, F. *et al.* Defining a healthy human gut microbiome: current concepts, future directions, and clinical applications. *Cell Host Microbe* **12**, 611-622, doi:10.1016/j.chom.2012.10.012 (2012).
30 Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K. & Knight, R. Diversity, stability and resilience of the human gut microbiota. *Nature* **489**, 220-230, doi:10.1038/nature11550 (2012).
31 Franzosa, E. A. *et al.* Relating the metatranscriptome and metagenome of the human gut. *Proc Natl Acad Sci U S A* **111**, E2329-2338, doi:10.1073/pnas.1319284111 (2014).
32 Abu-Ali, G. S. *et al.* Metatranscriptome of human faecal microbial communities in a cohort of adult men. *Nat Microbiol* **3**, 356-366, doi:10.1038/s41564-017-0084-4 (2018).
33 Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59-65, doi:10.1038/nature08821 (2010).
34 Eckburg, P. B. *et al.* Diversity of the human intestinal microbial flora. *Science* **308**, 1635-1638, doi:10.1126/science.1110591 (2005).

35 Tropini, C., Earle, K. A., Huang, K. C. & Sonnenburg, J. L. The Gut Microbiome: Connecting Spatial Organization to Function. *Cell Host Microbe* **21**, 433-442, doi:10.1016/j.chom.2017.03.010 (2017).

36 Gu, S. *et al.* Bacterial community mapping of the mouse gastrointestinal tract. *PLoS One* **8**, e74957, doi:10.1371/journal.pone.0074957 (2013).

37 Hooi, J. K. Y. *et al.* Global Prevalence of Helicobacter pylori Infection: Systematic Review and Meta-Analysis. *Gastroenterology* **153**, 420-429, doi:10.1053/j.gastro.2017.04.022 (2017).

38 Plummer, M., Franceschi, S., Vignat, J., Forman, D. & de Martel, C. Global burden of gastric cancer attributable to Helicobacter pylori. *Int J Cancer* **136**, 487-490, doi:10.1002/ijc.28999 (2015).

39 Graham, D. Y. History of Helicobacter pylori, duodenal ulcer, gastric ulcer and gastric cancer. *World J Gastroenterol* **20**, 5191-5204, doi:10.3748/wjg.v20.i18.5191 (2014).

40 Lanas, A. & Chan, F. K. L. Peptic ulcer disease. *Lancet* **390**, 613-624, doi:10.1016/S0140-6736(16)32404-7 (2017).

41 Engstrand, L. & Lindberg, M. Helicobacter pylori and the gastric microbiota. *Best Pract Res Clin Gastroenterol* **27**, 39-45, doi:10.1016/j.bpg.2013.03.016 (2013).

42 Nardone, G. & Compare, D. The human gastric microbiota: Is it time to rethink the pathogenesis of stomach diseases? *United European Gastroenterol J* **3**, 255-260, doi:10.1177/2050640614566846 (2015).

43 Delgado, S., Cabrera-Rubio, R., Mira, A., Suarez, A. & Mayo, B. Microbiological survey of the human gastric ecosystem using culturing and pyrosequencing methods. *Microb Ecol* **65**, 763-772, doi:10.1007/s00248-013-0192-5 (2013).

44 Bik, E. M. *et al.* Molecular analysis of the bacterial microbiota in the human stomach. *Proc Natl Acad Sci U S A* **103**, 732-737, doi:10.1073/pnas.0506655103 (2006).

45 Li, X. X. *et al.* Bacterial microbiota profiling in gastritis without Helicobacter pylori infection or non-steroidal anti-inflammatory drug use. *PLoS One* **4**, e7985, doi:10.1371/journal.pone.0007985 (2009).

46 Donaldson, G. P., Lee, S. M. & Mazmanian, S. K. Gut biogeography of the bacterial microbiota. *Nat Rev Microbiol* **14**, 20-32, doi:10.1038/nrmicro3552 (2016).

47 Booijink, C. C. *et al.* High temporal and inter-individual variation detected in the human ileal microbiota. *Environ Microbiol* **12**, 3213-3227, doi:10.1111/j.1462-2920.2010.02294.x (2010).

48 Zoetendal, E. G. *et al.* The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *ISME J* **6**, 1415-1426, doi:10.1038/ismej.2011.212 (2012).

49 Gorvitovskaia, A., Holmes, S. P. & Huse, S. M. Interpreting Prevotella and Bacteroides as biomarkers of diet and lifestyle. *Microbiome* **4**, 15, doi:10.1186/s40168-016-0160-7 (2016).

50 Samuel, B. S. *et al.* Genomic and metabolic adaptations of Methanobrevibacter smithii to the human gut. *Proc Natl Acad Sci U S A* **104**, 10643-10648, doi:10.1073/pnas.0704189104 (2007).

51 Koropatkin, N. M., Cameron, E. A. & Martens, E. C. How glycan metabolism shapes the human gut microbiota. *Nat Rev Microbiol* **10**, 323-335, doi:10.1038/nrmicro2746 (2012).

52 Stams, A. J. Metabolic interactions between anaerobic bacteria in methanogenic environments. *Antonie Van Leeuwenhoek* **66**, 271-294 (1994).

53 Stams, A. J., Oude Elferink, S. J. & Westermann, P. Metabolic interactions between methanogenic consortia and anaerobic respiring bacteria. *Adv Biochem Eng Biotechnol* **81**, 31-56 (2003).

54 Sicard, J. F., Le Bihan, G., Vogeleer, P., Jacques, M. & Harel, J. Interactions of Intestinal Bacteria with Components of the Intestinal Mucus. *Front Cell Infect Microbiol* **7**, 387, doi:10.3389/fcimb.2017.00387 (2017).

55 Ermund, A., Schutte, A., Johansson, M. E., Gustafsson, J. K. & Hansson, G. C. Studies of mucus in mouse stomach, small intestine, and colon. I. Gastrointestinal mucus layers have different properties depending on location as well as over the Peyer's patches. *Am J Physiol Gastrointest Liver Physiol* **305**, G341-347, doi:10.1152/ajpgi.00046.2013 (2013).

56 Johansson, M. E. *et al.* The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc Natl Acad Sci U S A* **105**, 15064-15069, doi:10.1073/pnas.0803124105 (2008).

57 Li, H. *et al.* The outer mucus layer hosts a distinct intestinal microbial niche. *Nat Commun* **6**, 8292, doi:10.1038/ncomms9292 (2015).

58 Albenberg, L. *et al.* Correlation between intraluminal oxygen gradient and radial partitioning of intestinal microbiota. *Gastroenterology* **147**, 1055-1063 e1058, doi:10.1053/j.gastro.2014.07.020 (2014).

59 Pedron, T. *et al.* A crypt-specific core microbiota resides in the mouse colon. *MBio* **3**, doi:10.1128/mBio.00116-12 (2012).

60 Ivanov, II *et al.* Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell* **139**, 485-498, doi:10.1016/j.cell.2009.09.033 (2009).

61 Atarashi, K. *et al.* Th17 Cell Induction by Adhesion of Microbes to Intestinal Epithelial Cells. *Cell* **163**, 367-380, doi:10.1016/j.cell.2015.08.058 (2015).

62 Jiang, Y., Xiong, X., Danska, J. & Parkinson, J. Metatranscriptomic analysis of diverse microbial communities reveals core metabolic pathways and microbiome-specific functionality. *Microbiome* **4**, 2, doi:10.1186/s40168-015-0146-x (2016).

63 Nava, G. M., Friedrichsen, H. J. & Stappenbeck, T. S. Spatial organization of intestinal microbiota in the mouse ascending colon. *ISME J* **5**, 627-638, doi:10.1038/ismej.2010.161 (2011).

64 Lavelle, A. *et al.* Spatial variation of the colonic microbiota in patients with ulcerative colitis and control volunteers. *Gut* **64**, 1553-1561, doi:10.1136/gutjnl-2014-307873 (2015).

65 Jakobsson, H. E. *et al.* The composition of the gut microbiota shapes the colon mucus barrier. *EMBO Rep* **16**, 164-177, doi:10.15252/embr.201439263 (2015).

66 Krishnan, K., Chen, T. & Paster, B. J. A practical guide to the oral microbiome and its relation to health and disease. *Oral Dis* **23**, 276-286, doi:10.1111/odi.12509 (2017).

67 Perez Perez, G. I. *et al.* Body Site Is a More Determinant Factor than Human Population Diversity in the Healthy Skin Microbiome. *PLoS One* **11**, e0151990, doi:10.1371/journal.pone.0151990 (2016).

68 Costello, E. K. *et al.* Bacterial community variation in human body habitats across space and time. *Science* **326**, 1694-1697, doi:10.1126/science.1177486 (2009).

69 Grice, E. A. *et al.* Topographical and temporal diversity of the human skin microbiome. *Science* **324**, 1190-1192, doi:10.1126/science.1171700 (2009).

70 Oh, J. *et al.* Biogeography and individuality shape function in the human skin metagenome. *Nature* **514**, 59-64, doi:10.1038/nature13786 (2014).

71 Probst, A. J., Auerbach, A. K. & Moissl-Eichinger, C. Archaea on human skin. *PLoS One* **8**, e65388, doi:10.1371/journal.pone.0065388 (2013).

72 Horz, H. P. Archaeal Lineages within the Human Microbiome: Absent, Rare or Elusive? *Life (Basel)* **5**, 1333-1345, doi:10.3390/life5021333 (2015).

73 Moissl-Eichinger, C. *et al.* Human age and skin physiology shape diversity and abundance of Archaea on skin. *Sci Rep* **7**, 4039, doi:10.1038/s41598-017-04197-4 (2017).

74 Ma, B., Forney, L. J. & Ravel, J. Vaginal microbiome: rethinking health and disease. *Annu Rev Microbiol* **66**, 371-389, doi:10.1146/annurev-micro-092611-150157 (2012).

75 Ravel, J. *et al.* Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A* **108 Suppl 1**, 4680-4687, doi:10.1073/pnas.1002611107 (2011).

76 Flint, H. J., Bayer, E. A., Rincon, M. T., Lamed, R. & White, B. A. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nat Rev Microbiol* **6**, 121-131, doi:10.1038/nrmicro1817 (2008).

77 Koh, A., De Vadder, F., Kovatcheva-Datchary, P. & Backhed, F. From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites. *Cell* **165**, 1332-1345, doi:10.1016/j.cell.2016.05.041 (2016).

78 Rowland, I. *et al.* Gut microbiota functions: metabolism of nutrients and other food components. *Eur J Nutr* **57**, 1-24, doi:10.1007/s00394-017-1445-8 (2018).

79 Gosalbes, M. J. *et al.* Metatranscriptomic approach to analyze the functional human gut microbiota. *PLoS One* **6**, e17447, doi:10.1371/journal.pone.0017447 (2011).

80 Desai, M. S. *et al.* A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility. *Cell* **167**, 1339-1353 e1321, doi:10.1016/j.cell.2016.10.043 (2016).

81 Sonnenburg, J. L. *et al.* Glycan foraging in vivo by an intestine-adapted bacterial symbiont. *Science* **307**, 1955-1959, doi:10.1126/science.1109051 (2005).

82 Earle, K. A. *et al.* Quantitative Imaging of Gut Microbiota Spatial Organization. *Cell Host Microbe* **18**, 478-488, doi:10.1016/j.chom.2015.09.002 (2015).

83 Wostmann, B. S., Larkin, C., Moriarty, A. & Bruckner-Kardoss, E. Dietary intake, energy metabolism, and excretory losses of adult male germfree Wistar rats. *Lab Anim Sci* **33**, 46-50 (1983).

84 Backhed, F. *et al.* The gut microbiota as an environmental factor that regulates fat storage. *Proc Natl Acad Sci U S A* **101**, 15718-15723, doi:10.1073/pnas.0407076101 (2004).

85 Rabot, S. *et al.* Germ-free C57BL/6J mice are resistant to high-fat-diet-induced insulin resistance and have altered cholesterol metabolism. *FASEB J* **24**, 4948-4959, doi:10.1096/fj.10-164921 (2010).
86 Chang, E. B. & Martinez-Guryn, K. Small intestinal microbiota: the neglected stepchild needed for fat digestion and absorption. *Gut Microbes*, 1-6, doi:10.1080/19490976.2018.1502539 (2018).
87 Martinez-Guryn, K. *et al.* Small Intestine Microbiota Regulate Host Digestive and Absorptive Adaptive Responses to Dietary Lipids. *Cell Host Microbe* **23**, 458-469 e455, doi:10.1016/j.chom.2018.03.011 (2018).
88 Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027-1031, doi:10.1038/nature05414 (2006).
89 Ridaura, V. K. *et al.* Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science* **341**, 1241214, doi:10.1126/science.1241214 (2013).
90 De Filippo, C. *et al.* Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc Natl Acad Sci U S A* **107**, 14691-14696, doi:10.1073/pnas.1005963107 (2010).
91 Hildebrandt, M. A. *et al.* High-fat diet determines the composition of the murine gut microbiome independently of obesity. *Gastroenterology* **137**, 1716-1724 e1711-1712, doi:10.1053/j.gastro.2009.08.042 (2009).
92 Ley, R. E., Turnbaugh, P. J., Klein, S. & Gordon, J. I. Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022-1023, doi:10.1038/4441022a (2006).
93 Turnbaugh, P. J. *et al.* The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci Transl Med* **1**, 6ra14, doi:10.1126/scitranslmed.3000322 (2009).
94 David, L. A. *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559-563, doi:10.1038/nature12820 (2014).
95 Walters, W. A., Xu, Z. & Knight, R. Meta-analyses of human gut microbes associated with obesity and IBD. *FEBS Lett* **588**, 4223-4233, doi:10.1016/j.febslet.2014.09.039 (2014).
96 den Besten, G. *et al.* The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J Lipid Res* **54**, 2325-2340, doi:10.1194/jlr.R036012 (2013).
97 Wong, J. M., de Souza, R., Kendall, C. W., Emam, A. & Jenkins, D. J. Colonic health: fermentation and short chain fatty acids. *J Clin Gastroenterol* **40**, 235-243 (2006).
98 Tan, J. *et al.* The role of short-chain fatty acids in health and disease. *Adv Immunol* **121**, 91-119, doi:10.1016/B978-0-12-800100-4.00003-9 (2014).
99 Bergman, E. N. Energy contributions of volatile fatty acids from the gastrointestinal tract in various species. *Physiol Rev* **70**, 567-590, doi:10.1152/physrev.1990.70.2.567 (1990).
100 Ohira, H., Tsutsui, W. & Fujioka, Y. Are Short Chain Fatty Acids in Gut Microbiota Defensive Players for Inflammation and Atherosclerosis? *J Atheroscler Thromb* **24**, 660-672, doi:10.5551/jat.RV17006 (2017).

101 Rook, G. A. & Brunet, L. R. Microbes, immunoregulation, and the gut. *Gut* **54**, 317-320, doi:10.1136/gut.2004.053785 (2005).

102 Kondrashova, A., Seiskari, T., Ilonen, J., Knip, M. & Hyoty, H. The 'Hygiene hypothesis' and the sharp gradient in the incidence of autoimmune and allergic diseases between Russian Karelia and Finland. *APMIS* **121**, 478-493, doi:10.1111/apm.12023 (2013).

103 Rook, G. A. 99th Dahlem conference on infection, inflammation and chronic inflammatory disorders: darwinian medicine and the 'hygiene' or 'old friends' hypothesis. *Clin Exp Immunol* **160**, 70-79, doi:10.1111/j.1365-2249.2010.04133.x (2010).

104 Rook, G. A., Lowry, C. A. & Raison, C. L. Microbial 'Old Friends', immunoregulation and stress resilience. *Evol Med Public Health* **2013**, 46-64, doi:10.1093/emph/eot004 (2013).

105 Rook, G. A. A Darwinian View of the Hygiene or "Old Friends" Hypothesis. *Microbe Magazine*, 173 - 180 (2012).

106 Thornton, A. M. & Shevach, E. M. Suppressor effector function of CD4+CD25+ immunoregulatory T cells is antigen nonspecific. *J Immunol* **164**, 183-190 (2000).

107 Rook, G. A. Review series on helminths, immune modulation and the hygiene hypothesis: the broader implications of the hygiene hypothesis. *Immunology* **126**, 3-11, doi:10.1111/j.1365-2567.2008.03007.x (2009).

108 Eldeirawi, K. *et al.* Associations of doctor-diagnosed asthma with immigration status, age at immigration, and length of residence in the United States in a sample of Mexican American School Children in Chicago. *J Asthma* **46**, 796-802 (2009).

109 Hjern, A., Rasmussen, F. & Hedlin, G. Age at adoption, ethnicity and atopic disorder: a study of internationally adopted young men in Sweden. *Pediatr Allergy Immunol* **10**, 101-106 (1999).

110 Pereg, D. *et al.* Prevalence of asthma in a large group of Israeli adolescents: influence of country of birth and age at migration. *Allergy* **63**, 1040-1045, doi:10.1111/j.1398-9995.2008.01661.x (2008).

111 Li, X., Sundquist, J., Hemminki, K. & Sundquist, K. Risk of inflammatory bowel disease in first- and second-generation immigrants in Sweden: a nationwide follow-up study. *Inflamm Bowel Dis* **17**, 1784-1791, doi:10.1002/ibd.21535 (2011).

112 Ahlgren, C., Oden, A. & Lycke, J. A nationwide survey of the prevalence of multiple sclerosis in immigrant populations of Sweden. *Mult Scler* **18**, 1099-1107, doi:10.1177/1352458511433062 (2012).

113 Soderstrom, U., Aman, J. & Hjern, A. Being born in Sweden increases the risk for type 1 diabetes - a study of migration of children to Sweden as a natural experiment. *Acta Paediatr* **101**, 73-77, doi:10.1111/j.1651-2227.2011.02410.x (2012).

114 Carr, I. & Mayberry, J. F. The effects of migration on ulcerative colitis: a three-year prospective study among Europeans and first- and second- generation South Asians in Leicester (1991-1994). *Am J Gastroenterol* **94**, 2918-2922, doi:10.1111/j.1572-0241.1999.01438.x (1999).

115 Macpherson, A. J. & Harris, N. L. Interactions between commensal intestinal bacteria and the immune system. *Nat Rev Immunol* **4**, 478-485, doi:10.1038/nri1373 (2004).

116 Falk, P. G., Hooper, L. V., Midtvedt, T. & Gordon, J. I. Creating and maintaining the gastrointestinal ecosystem: what we know and need to know from gnotobiology. *Microbiol Mol Biol Rev* **62**, 1157-1170 (1998).

117 Round, J. L. & Mazmanian, S. K. The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol* **9**, 313-323, doi:10.1038/nri2515 (2009).

118 Forchielli, M. L. & Walker, W. A. The role of gut-associated lymphoid tissues and mucosal defence. *Br J Nutr* **93 Suppl 1**, S41-48 (2005).

119 Dobber, R., Hertogh-Huijbregts, A., Rozing, J., Bottomly, K. & Nagelkerken, L. The involvement of the intestinal microflora in the expansion of CD4+ T cells with a naive phenotype in the periphery. *Dev Immunol* **2**, 141-150 (1992).

120 Moreau, M. C. & Corthier, G. Effect of the gastrointestinal microflora on induction and maintenance of oral tolerance to ovalbumin in C3H/HeJ mice. *Infect Immun* **56**, 2766-2768 (1988).

121 Artis, D. Epithelial-cell recognition of commensal bacteria and maintenance of immune homeostasis in the gut. *Nat Rev Immunol* **8**, 411-420, doi:10.1038/nri2316 (2008).

122 Kirjavainen, P. V. & Gibson, G. R. Healthy gut microflora and allergy: factors influencing development of the microbiota. *Ann Med* **31**, 288-292 (1999).

123 Mazmanian, S. K., Liu, C. H., Tzianabos, A. O. & Kasper, D. L. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* **122**, 107-118, doi:10.1016/j.cell.2005.05.007 (2005).

124 Wingender, G. *et al.* Intestinal microbes affect phenotypes and functions of invariant natural killer T cells in mice. *Gastroenterology* **143**, 418-428, doi:10.1053/j.gastro.2012.04.017 (2012).

125 Olszak, T. *et al.* Microbial exposure during early life has persistent effects on natural killer T cell function. *Science* **336**, 489-493, doi:10.1126/science.1219328 (2012).

126 Nutsch, K. *et al.* Rapid and Efficient Generation of Regulatory T Cells to Commensal Antigens in the Periphery. *Cell Rep* **17**, 206-220, doi:10.1016/j.celrep.2016.08.092 (2016).

127 Cahenzli, J., Koller, Y., Wyss, M., Geuking, M. B. & McCoy, K. D. Intestinal microbial diversity during early-life colonization shapes long-term IgE levels. *Cell Host Microbe* **14**, 559-570, doi:10.1016/j.chom.2013.10.004 (2013).

128 Sudo, N. *et al.* The requirement of intestinal bacterial flora for the development of an IgE production system fully susceptible to oral tolerance induction. *J Immunol* **159**, 1739-1745 (1997).

129 Gaboriau-Routhiau, V. *et al.* The key role of segmented filamentous bacteria in the coordinated maturation of gut helper T cell responses. *Immunity* **31**, 677-689, doi:10.1016/j.immuni.2009.08.020 (2009).

130 Wing, K. & Sakaguchi, S. Regulatory T cells exert checks and balances on self tolerance and autoimmunity. *Nat Immunol* **11**, 7-13, doi:10.1038/ni.1818 (2010).

131 Li, Z., Li, D., Tsun, A. & Li, B. FOXP3+ regulatory T cells and their functional regulation. *Cell Mol Immunol* **12**, 558-565, doi:10.1038/cmi.2015.10 (2015).
132 Sakaguchi, S., Miyara, M., Costantino, C. M. & Hafler, D. A. FOXP3+ regulatory T cells in the human immune system. *Nat Rev Immunol* **10**, 490-500, doi:10.1038/nri2785 (2010).
133 Cebula, A. *et al.* Thymus-derived regulatory T cells contribute to tolerance to commensal microbiota. *Nature* **497**, 258-262, doi:10.1038/nature12079 (2013).
134 Lathrop, S. K. *et al.* Peripheral education of the immune system by colonic commensal microbiota. *Nature* **478**, 250-254, doi:10.1038/nature10434 (2011).
135 Grayson, M. H. *et al.* Intestinal Microbiota Disruption Reduces Regulatory T Cells and Increases Respiratory Viral Infection Mortality Through Increased IFNgamma Production. *Front Immunol* **9**, 1587, doi:10.3389/fimmu.2018.01587 (2018).
136 Russler-Germain, E. V., Rengarajan, S. & Hsieh, C. S. Antigen-specific regulatory T-cell responses to intestinal microbiota. *Mucosal Immunol* **10**, 1375-1386, doi:10.1038/mi.2017.65 (2017).
137 Farache, J. *et al.* Luminal bacteria recruit CD103+ dendritic cells into the intestinal epithelium to sample bacterial antigens for presentation. *Immunity* **38**, 581-595, doi:10.1016/j.immuni.2013.01.009 (2013).
138 Annacker, O. *et al.* Essential role for CD103 in the T cell-mediated regulation of experimental colitis. *J Exp Med* **202**, 1051-1061, doi:10.1084/jem.20040662 (2005).
139 Matteoli, G. *et al.* Gut CD103+ dendritic cells express indoleamine 2,3-dioxygenase which influences T regulatory/T effector cell balance and oral tolerance induction. *Gut* **59**, 595-604, doi:10.1136/gut.2009.185108 (2010).
140 Sun, C. M. *et al.* Small intestine lamina propria dendritic cells promote de novo generation of Foxp3 T reg cells via retinoic acid. *J Exp Med* **204**, 1775-1785, doi:10.1084/jem.20070602 (2007).
141 Ai, T. L., Solomon, B. D. & Hsieh, C. S. T-cell selection and intestinal homeostasis. *Immunol Rev* **259**, 60-74, doi:10.1111/imr.12171 (2014).
142 Atarashi, K. *et al.* Induction of colonic regulatory T cells by indigenous Clostridium species. *Science* **331**, 337-341, doi:10.1126/science.1198469 (2011).
143 Ostman, S., Rask, C., Wold, A. E., Hultkrantz, S. & Telemo, E. Impaired regulatory T cell function in germ-free mice. *Eur J Immunol* **36**, 2336-2346, doi:10.1002/eji.200535244 (2006).
144 Lane, E. R., Zisman, T. L. & Suskind, D. L. The microbiota in inflammatory bowel disease: current and therapeutic insights. *J Inflamm Res* **10**, 63-73, doi:10.2147/JIR.S116088 (2017).
145 Hansen, J. J. & Sartor, R. B. Therapeutic Manipulation of the Microbiome in IBD: Current Results and Future Approaches. *Curr Treat Options Gastroenterol* **13**, 105-120, doi:10.1007/s11938-014-0042-7 (2015).
146 Kennedy, R. J. *et al.* Interleukin 10-deficient colitis: new similarities to human inflammatory bowel disease. *Br J Surg* **87**, 1346-1351, doi:10.1046/j.1365-2168.2000.01615.x (2000).
147 Zeng, H. & Chi, H. Metabolic control of regulatory T cell development and function. *Trends Immunol* **36**, 3-12, doi:10.1016/j.it.2014.08.003 (2015).

148 Furusawa, Y., Obata, Y. & Hase, K. Commensal microbiota regulates T cell fate decision in the gut. *Semin Immunopathol* **37**, 17-25, doi:10.1007/s00281-014-0455-3 (2015).

149 Round, J. L. & Mazmanian, S. K. Inducible Foxp3+ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. *Proc Natl Acad Sci U S A* **107**, 12204-12209, doi:10.1073/pnas.0909122107 (2010).

150 Ochoa-Reparaz, J. *et al.* A polysaccharide from the human commensal Bacteroides fragilis protects against CNS demyelinating disease. *Mucosal Immunol* **3**, 487-495, doi:10.1038/mi.2010.29 (2010).

151 Mazmanian, S. K., Round, J. L. & Kasper, D. L. A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature* **453**, 620-625, doi:10.1038/nature07008 (2008).

152 Kelly, D. *et al.* Commensal anaerobic gut bacteria attenuate inflammation by regulating nuclear-cytoplasmic shuttling of PPAR-gamma and RelA. *Nat Immunol* **5**, 104-112, doi:10.1038/ni1018 (2004).

153 Collier-Hyams, L. S., Sloane, V., Batten, B. C. & Neish, A. S. Cutting edge: bacterial modulation of epithelial signaling via changes in neddylation of cullin-1. *J Immunol* **175**, 4194-4198 (2005).

154 Tien, M. T. *et al.* Anti-inflammatory effect of Lactobacillus casei on Shigella-infected human intestinal epithelial cells. *J Immunol* **176**, 1228-1237 (2006).

155 Cording, S. *et al.* Commensal microbiota drive proliferation of conventional and Foxp3(+) regulatory CD4(+) T cells in mesenteric lymph nodes and Peyer's patches. *Eur J Microbiol Immunol (Bp)* **3**, 1-10, doi:10.1556/EuJMI.3.2013.1.1 (2013).

156 Furusawa, Y. *et al.* Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* **504**, 446-450, doi:10.1038/nature12721 (2013).

157 Schwarz, A., Bruhs, A. & Schwarz, T. The Short-Chain Fatty Acid Sodium Butyrate Functions as a Regulator of the Skin Immune System. *J Invest Dermatol* **137**, 855-864, doi:10.1016/j.jid.2016.11.014 (2017).

158 Allali, I. *et al.* A comparison of sequencing platforms and bioinformatics pipelines for compositional analysis of the gut microbiome. *BMC Microbiol* **17**, 194, doi:10.1186/s12866-017-1101-8 (2017).

159 Smith, P. M. *et al.* The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science* **341**, 569-573, doi:10.1126/science.1241165 (2013).

160 Arpaia, N. *et al.* Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. *Nature* **504**, 451-455, doi:10.1038/nature12726 (2013).

161 Sokol, H. *et al.* Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci U S A* **105**, 16731-16736, doi:10.1073/pnas.0804812105 (2008).

162 Qiu, X., Zhang, M., Yang, X., Hong, N. & Yu, C. Faecalibacterium prausnitzii upregulates regulatory T cells and anti-inflammatory cytokines in treating TNBS-induced colitis. *J Crohns Colitis* **7**, e558-568, doi:10.1016/j.crohns.2013.04.002 (2013).

163 Zhou, L. *et al.* Faecalibacterium prausnitzii Produces Butyrate to Maintain Th17/Treg Balance and to Ameliorate Colorectal Colitis by Inhibiting Histone Deacetylase 1. *Inflamm Bowel Dis*, doi:10.1093/ibd/izy182 (2018).

164 Geuking, M. B. *et al.* Intestinal bacterial colonization induces mutualistic regulatory T cell responses. *Immunity* **34**, 794-806, doi:10.1016/j.immuni.2011.03.021 (2011).

165 Geuking, M. B., McCoy, K. D. & Macpherson, A. J. The continuum of intestinal CD4+ T cell adaptations in host-microbial mutualism. *Gut Microbes* **2**, 353-357, doi:10.4161/gmic.18604 (2011).

166 Smits, H. H. *et al.* Selective probiotic bacteria induce IL-10-producing regulatory T cells in vitro by modulating dendritic cell function through dendritic cell-specific intercellular adhesion molecule 3-grabbing nonintegrin. *J Allergy Clin Immunol* **115**, 1260-1267, doi:10.1016/j.jaci.2005.03.036 (2005).

167 Farkas, A. M. *et al.* Induction of Th17 cells by segmented filamentous bacteria in the murine intestine. *J Immunol Methods* **421**, 104-111, doi:10.1016/j.jim.2015.03.020 (2015).

168 Okumura, R. & Takeda, K. Roles of intestinal epithelial cells in the maintenance of gut homeostasis. *Exp Mol Med* **49**, e338, doi:10.1038/emm.2017.20 (2017).

169 Johansson, M. E., Larsson, J. M. & Hansson, G. C. The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. *Proc Natl Acad Sci U S A* **108 Suppl 1**, 4659-4665, doi:10.1073/pnas.1006451107 (2011).

170 Gallo, R. L. & Hooper, L. V. Epithelial antimicrobial defence of the skin and intestine. *Nat Rev Immunol* **12**, 503-516, doi:10.1038/nri3228 (2012).

171 Vaishnava, S. *et al.* The antibacterial lectin RegIIIgamma promotes the spatial segregation of microbiota and host in the intestine. *Science* **334**, 255-258, doi:10.1126/science.1209791 (2011).

172 Meyer-Hoffert, U. *et al.* Secreted enteric antimicrobial activity localises to the mucus surface layer. *Gut* **57**, 764-771, doi:10.1136/gut.2007.141481 (2008).

173 Lievin-Le Moal, V. & Servin, A. L. The front line of enteric host defense against unwelcome intrusion of harmful microorganisms: mucins, antimicrobial peptides, and microbiota. *Clin Microbiol Rev* **19**, 315-337, doi:10.1128/CMR.19.2.315-337.2006 (2006).

174 Buffie, C. G. & Pamer, E. G. Microbiota-mediated colonization resistance against intestinal pathogens. *Nat Rev Immunol* **13**, 790-801, doi:10.1038/nri3535 (2013).

175 Petersson, J. *et al.* Importance and regulation of the colonic mucus barrier in a mouse model of colitis. *Am J Physiol Gastrointest Liver Physiol* **300**, G327-333, doi:10.1152/ajpgi.00422.2010 (2011).

176 Szentkuti, L., Riedesel, H., Enss, M. L., Gaertner, K. & Von Engelhardt, W. Pre-epithelial mucus layer in the colon of conventional and germ-free rats. *Histochem J* **22**, 491-497 (1990).

177 Comelli, E. M. *et al.* Multifaceted transcriptional regulation of the murine intestinal mucus layer by endogenous microbiota. *Genomics* **91**, 70-77, doi:10.1016/j.ygeno.2007.09.006 (2008).

178 Enss, M. L. *et al.* Effects of perorally applied endotoxin on colonic mucins of germfree rats. *Scand J Gastroenterol* **31**, 868-874 (1996).

179 Mack, D. R., Michail, S., Wei, S., McDougall, L. & Hollingsworth, M. A. Probiotics inhibit enteropathogenic E. coli adherence in vitro by inducing intestinal mucin gene expression. *Am J Physiol* **276**, G941-950 (1999).
180 Hatayama, H., Iwashita, J., Kuwajima, A. & Abe, T. The short chain fatty acid, butyrate, stimulates MUC2 mucin production in the human colon cancer cell line, LS174T. *Biochem Biophys Res Commun* **356**, 599-603, doi:10.1016/j.bbrc.2007.03.025 (2007).
181 Willemsen, L. E., Koetsier, M. A., van Deventer, S. J. & van Tol, E. A. Short chain fatty acids stimulate epithelial mucin 2 expression through differential effects on prostaglandin E(1) and E(2) production by intestinal myofibroblasts. *Gut* **52**, 1442-1447 (2003).
182 Peng, L., He, Z., Chen, W., Holzman, I. R. & Lin, J. Effects of butyrate on intestinal barrier function in a Caco-2 cell monolayer model of intestinal barrier. *Pediatr Res* **61**, 37-41, doi:10.1203/01.pdr.0000250014.92242.f3 (2007).
183 Zheng, L. *et al.* Microbial-Derived Butyrate Promotes Epithelial Barrier Function through IL-10 Receptor-Dependent Repression of Claudin-2. *J Immunol* **199**, 2976-2984, doi:10.4049/jimmunol.1700105 (2017).
184 Kelly, C. J. *et al.* Crosstalk between Microbiota-Derived Short-Chain Fatty Acids and Intestinal Epithelial HIF Augments Tissue Barrier Function. *Cell Host Microbe* **17**, 662-671, doi:10.1016/j.chom.2015.03.005 (2015).
185 Donaldson, G. P. *et al.* Gut microbiota utilize immunoglobulin A for mucosal colonization. *Science* **360**, 795-800, doi:10.1126/science.aaq0926 (2018).
186 Sonnenburg, J. L., Chen, C. T. & Gordon, J. I. Genomic and metabolic studies of the impact of probiotics on a model gut symbiont and host. *PLoS Biol* **4**, e413, doi:10.1371/journal.pbio.0040413 (2006).
187 Cash, H. L., Whitham, C. V., Behrendt, C. L. & Hooper, L. V. Symbiotic bacteria direct expression of an intestinal bactericidal lectin. *Science* **313**, 1126-1130, doi:10.1126/science.1127119 (2006).
188 Deriu, E. *et al.* Probiotic bacteria reduce salmonella typhimurium intestinal colonization by competing for iron. *Cell Host Microbe* **14**, 26-37, doi:10.1016/j.chom.2013.06.007 (2013).
189 Vollaard, E. J., Clasener, H. A. & Janssen, A. J. Co-trimoxazole impairs colonization resistance in healthy volunteers. *J Antimicrob Chemother* **30**, 685-691 (1992).
190 Nagpal, R. *et al.* Probiotics, their health benefits and applications for developing healthier foods: a review. *FEMS Microbiol Lett* **334**, 1-15, doi:10.1111/j.1574-6968.2012.02593.x (2012).
191 McFarland, L. V. Probiotics for the Primary and Secondary Prevention of C. difficile Infections: A Meta-analysis and Systematic Review. *Antibiotics (Basel)* **4**, 160-178, doi:10.3390/antibiotics4020160 (2015).
192 Zajac, A. E., Adams, A. S. & Turner, J. H. A systematic review and meta-analysis of probiotics for the treatment of allergic rhinitis. *Int Forum Allergy Rhinol* **5**, 524-532, doi:10.1002/alr.21492 (2015).
193 Aceti, A. *et al.* Probiotics for prevention of necrotizing enterocolitis in preterm infants: systematic review and meta-analysis. *Ital J Pediatr* **41**, 89, doi:10.1186/s13052-015-0199-2 (2015).

194 Shen, N. T. *et al.* Timely Use of Probiotics in Hospitalized Adults Prevents Clostridium difficile Infection: A Systematic Review With Meta-Regression Analysis. *Gastroenterology* **152**, 1889-1900 e1889, doi:10.1053/j.gastro.2017.02.003 (2017).

195 Panigrahi, P. *et al.* A randomized synbiotic trial to prevent sepsis among infants in rural India. *Nature* **548**, 407-412, doi:10.1038/nature23480 (2017).

196 Vos, T. *et al.* Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* **380**, 2163-2196, doi:10.1016/S0140-6736(12)61729-2 (2012).

197 Leiper, R. T. *Researches on Egyptian bilharziosis (a report to the War Office on the results of the Bilharzia Mission in Egypt, 1915)*. (Bale & Danielsson, 1918).

198 Gryseels, B., Polman, K., Clerinx, J. & Kestens, L. Human schistosomiasis. *Lancet* **368**, 1106-1118, doi:10.1016/S0140-6736(06)69440-3 (2006).

199 Pearce, E. J. & MacDonald, A. S. The immunobiology of schistosomiasis. *Nat Rev Immunol* **2**, 499-511, doi:10.1038/nri843 (2002).

200 Colley, D. G., Bustinduy, A. L., Secor, W. E. & King, C. H. Human schistosomiasis. *Lancet* **383**, 2253-2264, doi:10.1016/S0140-6736(13)61949-2 (2014).

201 Dew, H. R. Observations on the Pathology of Schistosomiasis (*S. haematobium* and *S. mansoni*) in the Human Subject. *J Pathol Bacteriol* **26**, 27-39 (1923).

202 King, C. H. & Dangerfield-Cha, M. The unacknowledged impact of chronic schistosomiasis. *Chronic Illn* **4**, 65-79, doi:10.1177/1742395307084407 (2008).

203 Weerakoon, K. G., Gobert, G. N., Cai, P. & McManus, D. P. Advances in the Diagnosis of Human Schistosomiasis. *Clin Microbiol Rev* **28**, 939-967, doi:10.1128/CMR.00137-14 (2015).

204 Montresor, A. *et al.* Preventive chemotherapy and the fight against neglected tropical diseases. *Expert Rev Anti Infect Ther* **10**, 237-242, doi:10.1586/eri.11.165 (2012).

205 Pearce, E. J., Caspar, P., Grzych, J. M., Lewis, F. A. & Sher, A. Downregulation of Th1 cytokine production accompanies induction of Th2 responses by a parasitic helminth, Schistosoma mansoni. *J Exp Med* **173**, 159-166 (1991).

206 Everts, B. *et al.* Omega-1, a glycoprotein secreted by Schistosoma mansoni eggs, drives Th2 responses. *J Exp Med* **206**, 1673-1680, doi:10.1084/jem.20082460 (2009).

207 Girgis, N. M., Gundra, U. M. & Loke, P. Immune regulation during helminth infections. *PLoS Pathog* **9**, e1003250, doi:10.1371/journal.ppat.1003250 (2013).

208 Maizels, R. M. & McSorley, H. J. Regulation of the host immune system by helminth parasites. *J Allergy Clin Immunol* **138**, 666-675, doi:10.1016/j.jaci.2016.07.007 (2016).

209 McSorley, H. J. & Maizels, R. M. Helminth infections and host immune regulation. *Clin Microbiol Rev* **25**, 585-608, doi:10.1128/CMR.05040-11 (2012).

210 Phillips, S. M. & Lammie, P. J. Immunopathology of granuloma formation and fibrosis in schistosomiasis. *Parasitol Today* **2**, 296-302 (1986).

211 Davies, S. J. & McKerrow, J. H. in *Biology of Parasitism* (eds C Tschudi & E. J. Pearce) 273-289 (Springer US, 2000).

212 Hoffmann, K. F., Cheever, A. W. & Wynn, T. A. IL-10 and the dangers of immune polarization: excessive type 1 and type 2 cytokine responses induce distinct forms of lethal immunopathology in murine schistosomiasis. *J Immunol* **164**, 6406-6416 (2000).

213 Colley, D. G. & Secor, W. E. Immunology of human schistosomiasis. *Parasite Immunol* **36**, 347-357, doi:10.1111/pim.12087 (2014).

214 Sabin, E. A., Araujo, M. I., Carvalho, E. M. & Pearce, E. J. Impairment of tetanus toxoid-specific Th1-like immune responses in humans infected with Schistosoma mansoni. *J Infect Dis* **173**, 269-272 (1996).

215 Malhotra, I. *et al.* Helminth- and Bacillus Calmette-Guerin-induced immunity in children sensitized in utero to filariasis and schistosomiasis. *J Immunol* **162**, 6843-6848 (1999).

216 Marshall, A. J. *et al.* Toxoplasma gondii and Schistosoma mansoni synergize to promote hepatocyte dysfunction associated with high levels of plasma TNF-alpha and early death in C57BL/6 mice. *J Immunol* **163**, 2089-2097 (1999).

217 Actor, J. K. *et al.* Helminth infection results in decreased virus-specific CD8+ cytotoxic T-cell and Th1 cytokine responses as well as delayed virus clearance. *Proc Natl Acad Sci U S A* **90**, 948-952 (1993).

218 Melhem, R. F. & LoVerde, P. T. Mechanism of interaction of Salmonella and Schistosoma species. *Infect Immun* **44**, 274-281 (1984).

219 LoVerde, P. T., Amento, C. & Higashi, G. I. Parasite-parasite interaction of Salmonella typhimurium and Schistosoma. *J Infect Dis* **141**, 177-185 (1980).

220 Feldmeier, H., Krantz, I. & Poggensee, G. Female genital schistosomiasis as a risk-factor for the transmission of HIV. *Int J STD AIDS* **5**, 368-372, doi:10.1177/095646249400500517 (1994).

221 Mbabazi, P. S. *et al.* Examining the relationship between urogenital schistosomiasis and HIV infection. *PLoS Negl Trop Dis* **5**, e1396, doi:10.1371/journal.pntd.0001396 (2011).

222 Kallestrup, P. *et al.* Schistosomiasis and HIV-1 infection in rural Zimbabwe: effect of treatment of schistosomiasis on CD4 cell count and plasma HIV-1 RNA load. *J Infect Dis* **192**, 1956-1961, doi:10.1086/497696 (2005).

223 Brown, M. *et al.* Schistosoma mansoni, nematode infections, and progression to active tuberculosis among HIV-1-infected Ugandans. *Am J Trop Med Hyg* **74**, 819-825 (2006).

224 DiNardo, A. R. *et al.* Schistosome Soluble Egg Antigen Decreases Mycobacterium tuberculosis-Specific CD4+ T-Cell Effector Function With Concomitant Arrest of Macrophage Phago-Lysosome Maturation. *J Infect Dis* **214**, 479-488, doi:10.1093/infdis/jiw156 (2016).

225 Monin, L. *et al.* Helminth-induced arginase-1 exacerbates lung inflammation and disease severity in tuberculosis. *J Clin Invest* **125**, 4699-4713, doi:10.1172/JCI77378 (2015).

226 El-Awady, M. K. *et al.* Soluble egg antigen of Schistosoma Haematobium induces HCV replication in PBMC from patients with chronic HCV infection. *BMC Infect Dis* **6**, 91, doi:10.1186/1471-2334-6-91 (2006).

227 Sipahi, A. M. & Baptista, D. M. Helminths as an alternative therapy for intestinal diseases. *World J Gastroenterol* **23**, 6009-6015, doi:10.3748/wjg.v23.i33.6009 (2017).

228 Rook, G. A., Raison, C. L. & Lowry, C. A. Microbial 'old friends', immunoregulation and socioeconomic status. *Clin Exp Immunol* **177**, 1-12, doi:10.1111/cei.12269 (2014).

229 Helmby, H. Human helminth therapy to treat inflammatory disorders - where do we stand? *BMC Immunol* **16**, 12, doi:10.1186/s12865-015-0074-3 (2015).

230 Maruszewska-Cheruiyot, M., Donskow-Lysoniewska, K. & Doligalska, M. Helminth therapy: Advances in the use of parasitic worms against Inflammatory Bowel Diseases and its challenges. *Helminthologia* **55**, 1-11, doi:10.1515/helm-2017-0048 (2018).

231 Bloomfield, S. F. *et al.* Time to abandon the hygiene hypothesis: new perspectives on allergic disease, the human microbiome, infectious disease prevention and the role of targeted hygiene. *Perspect Public Health* **136**, 213-224, doi:10.1177/1757913916650225 (2016).

232 Stiemsma, L. T., Reynolds, L. A., Turvey, S. E. & Finlay, B. B. The hygiene hypothesis: current perspectives and future therapies. *Immunotargets Ther* **4**, 143-157, doi:10.2147/ITT.S61528 (2015).

233 Mutapi, F. *et al.* Age-related and infection intensity-related shifts in antibody recognition of defined protein antigens in a schistosome-exposed population. *J Infect Dis* **198**, 167-175, doi:10.1086/589511 (2008).

234 Rosa, B. A. *et al.* Differential human gut microbiome assemblages during soil-transmitted helminth infections in Indonesia and Liberia. *Microbiome* **6**, 33, doi:10.1186/s40168-018-0416-5 (2018).

235 Lee, S. C. *et al.* Helminth colonization is associated with increased diversity of the gut microbiota. *PLoS Negl Trop Dis* **8**, e2880, doi:10.1371/journal.pntd.0002880 (2014).

236 Cooper, P. *et al.* Patent human infections with the whipworm, Trichuris trichiura, are not associated with alterations in the faecal microbiota. *PLoS One* **8**, e76573, doi:10.1371/journal.pone.0076573 (2013).

237 Holm, J. B. *et al.* Chronic Trichuris muris Infection Decreases Diversity of the Intestinal Microbiota and Concomitantly Increases the Abundance of Lactobacilli. *PLoS One* **10**, e0125495, doi:10.1371/journal.pone.0125495 (2015).

238 Jenkins, T. P. *et al.* Infections by human gastrointestinal helminths are associated with changes in faecal microbiota diversity and composition. *PLoS One* **12**, e0184719, doi:10.1371/journal.pone.0184719 (2017).

239 Li, R. W. *et al.* The effect of helminth infection on the microbial composition and structure of the caprine abomasal microbiome. *Sci Rep* **6**, 20606, doi:10.1038/srep20606 (2016).

240 Li, R. W. *et al.* Alterations in the porcine colon microbiota induced by the gastrointestinal nematode Trichuris suis. *Infect Immun* **80**, 2150-2157, doi:10.1128/IAI.00141-12 (2012).

241 Plieskatt, J. L. *et al.* Infection with the carcinogenic liver fluke Opisthorchis viverrini modifies intestinal and biliary microbiome. *FASEB J* **27**, 4572-4584, doi:10.1096/fj.13-232751 (2013).

242 Broadhurst, M. J. *et al.* Therapeutic helminth infection of macaques with idiopathic chronic diarrhea alters the inflammatory signature and mucosal microbiota of the colon. *PLoS Pathog* **8**, e1003000, doi:10.1371/journal.ppat.1003000 (2012).
243 Simcock, D. C. *et al.* Hypergastrinaemia, abomasal bacterial population densities and pH in sheep infected with Ostertagia circumcincta. *Int J Parasitol* **29**, 1053-1063 (1999).
244 McKenney, E. A. *et al.* Alteration of the rat cecal microbiome during colonization with the helminth Hymenolepis diminuta. *Gut Microbes* **6**, 182-193, doi:10.1080/19490976.2015.1047128 (2015).
245 Kay, G. L. *et al.* Differences in the Faecal Microbiome in Schistosoma haematobium Infected Children vs. Uninfected Children. *PLoS Negl Trop Dis* **9**, e0003861, doi:10.1371/journal.pntd.0003861 (2015).
246 Blair, D. M. The routes of schistosome egg passage from the human body. *Cent Afr J Med* **11**, 243-249 (1965).
247 Cunin, P., Tchuem Tchuente, L. A., Poste, B., Djibrilla, K. & Martin, P. M. Interactions between Schistosoma haematobium and Schistosoma mansoni in humans in north Cameroon. *Trop Med Int Health* **8**, 1110-1117 (2003).
248 Monaco, C. L. *et al.* Altered Virome and Bacterial Microbiome in Human Immunodeficiency Virus-Associated Acquired Immunodeficiency Syndrome. *Cell Host Microbe* **19**, 311-322, doi:10.1016/j.chom.2016.02.011 (2016).
249 Bhat, M. *et al.* Impact of Immunosuppression on the Metagenomic Composition of the Intestinal Microbiome: a Systems Biology Approach to Post-Transplant Diabetes. *Sci Rep* **7**, 10277, doi:10.1038/s41598-017-10471-2 (2017).
250 Zhang, Z. *et al.* Immunosuppressive effect of the gut microbiome altered by high-dose tacrolimus in mice. *Am J Transplant* **18**, 1646-1656, doi:10.1111/ajt.14661 (2018).
251 Xu, X. & Zhang, X. Effects of cyclophosphamide on immune system and gut microbiota in mice. *Microbiol Res* **171**, 97-106, doi:10.1016/j.micres.2014.11.002 (2015).
252 Tourret, J. *et al.* Immunosuppressive Treatment Alters Secretion of Ileal Antimicrobial Peptides and Gut Microbiota, and Favors Subsequent Colonization by Uropathogenic Escherichia coli. *Transplantation* **101**, 74-82, doi:10.1097/TP.0000000000001492 (2017).
253 Nigeria, F. M. o. H. o. Report on the Epidemiological Mapping of Schistosomiasis and Soil Transmitted Helminthiases in 19 States and the FCT, Nigeria. (Federal Ministry of Health of Nigeria, 2015).
254 Okwori, A. E. J. *et al.* Prevalence of Schistosomiasis among Primary School Children in Gadabuke District, Toto LGA, North Central Nigeria. *Brit Microbiol Res J* **4**, 255-261 (2014).
255 Nduka, F. O., Ajaero, C. M. & Nwoke, B. E. Urinary schistosomiasis among school children in an endemic community in south-eastern Nigeria. *Appl Parasitol* **36**, 34-40 (1995).
256 Dawaki, S. *et al.* Prevalence and Risk Factors of Schistosomiasis among Hausa Communities in Kano State, Nigeria. *Rev Inst Med Trop Sao Paulo* **58**, 54, doi:10.1590/S1678-9946201658054 (2016).

257 Kapito-Tembo, A. P. *et al.* Prevalence distribution and risk factors for Schistosoma hematobium infection among school children in Blantyre, Malawi. *PLoS Negl Trop Dis* **3**, e361, doi:10.1371/journal.pntd.0000361 (2009).
258 Abdullahi, M. & Saidu, T. B. Prevalence of Urinary Schistosomiasis among School-Aged Children in Wushishi Local Government Area of Niger State, Nigeria. *Bayero J Pure Appl Sci* **4**, 53-55 (2012).
259 Mitchell, K. M., Mutapi, F., Savill, N. J. & Woolhouse, M. E. Protective immunity to Schistosoma haematobium infection is primarily an anti-fecundity response stimulated by the death of adult worms. *Proc Natl Acad Sci U S A* **109**, 13347-13352, doi:10.1073/pnas.1121051109 (2012).
260 Mutapi, F. *et al.* Chemotherapy accelerates the development of acquired immune responses to Schistosoma haematobium infection. *J Infect Dis* **178**, 289-293 (1998).
261 Healey, G. R., Murphy, R., Brough, L., Butts, C. A. & Coad, J. Interindividual variability in gut microbiota and host response to dietary interventions. *Nutr Rev* **75**, 1059-1080, doi:10.1093/nutrit/nux062 (2017).
262 Ingang-Etoh, P. C., Essien, U. C., Amama, S. A. & Useh, M. F. Prevalence of urinary schistosomiasis among school children in Ukwelo-Obudu and Abini communities in Cross River State, Nigeria. *Port Harcourt Medical Journal* **3** (2009).
263 Okoli, E. I. & Odaibo, A. B. Urinary schistosomiasis among schoolchildren in Ibadan, an urban community in south-western Nigeria. *Trop Med Int Health* **4**, 308-315 (1999).
264 Fatiregun, A. A., Osungbade, K. O. & Olumide, A. E. Cost-effectiveness of screening methods for urinary schistosomiasis in a school-based control programme in Ibadan, Nigeria. *Health Policy* **89**, 72-77, doi:10.1016/j.healthpol.2008.05.004 (2009).
265 Ivoke, N. *et al.* Prevalence and transmission dynamics of Schistosoma haematobium infection in a rural community of southwestern Ebonyi State, Nigeria. *Trop Biomed* **31**, 77-88 (2014).
266 Agans, R. *et al.* Distal gut microbiota of adolescent children is different from that of adults. *FEMS Microbiol Ecol* **77**, 404-412, doi:10.1111/j.1574-6941.2011.01120.x (2011).
267 Xu, Y. *et al.* Bacterial Diversity of Intestinal Microbiota in Patients with Substance Use Disorders Revealed by 16S rRNA Gene Deep Sequencing. *Sci Rep* **7**, 3628, doi:10.1038/s41598-017-03706-9 (2017).
268 Kvasnovsky, C. L. *et al.* Clinical and symptom scores are significantly correlated with fecal microbiota features in patients with symptomatic uncomplicated diverticular disease: a pilot study. *Eur J Gastroenterol Hepatol* **30**, 107-112, doi:10.1097/MEG.0000000000000995 (2018).
269 Lopetuso, L. R., Scaldaferri, F., Petito, V. & Gasbarrini, A. Commensal Clostridia: leading players in the maintenance of gut homeostasis. *Gut Pathog* **5**, 23, doi:10.1186/1757-4749-5-23 (2013).
270 Suau, A. *et al.* Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. *Appl Environ Microbiol* **65**, 4799-4807 (1999).

271 Belizario, J. E. & Napolitano, M. Human microbiomes and their roles in dysbiosis, common diseases, and novel therapeutic approaches. *Front Microbiol* **6**, 1050, doi:10.3389/fmicb.2015.01050 (2015).

272 Bressa, C. *et al.* Differences in gut microbiota profile between women with active lifestyle and sedentary women. *PLoS One* **12**, e0171352, doi:10.1371/journal.pone.0171352 (2017).

273 Perez-Cobas, A. E. *et al.* Structural and functional changes in the gut microbiota associated to Clostridium difficile infection. *Front Microbiol* **5**, 335, doi:10.3389/fmicb.2014.00335 (2014).

274 Hakansson, A. & Molin, G. Gut microbiota and inflammation. *Nutrients* **3**, 637-682, doi:10.3390/nu3060637 (2011).

275 Bhute, S. *et al.* Molecular Characterization and Meta-Analysis of Gut Microbial Communities Illustrate Enrichment of Prevotella and Megasphaera in Indian Subjects. *Front Microbiol* **7**, 660, doi:10.3389/fmicb.2016.00660 (2016).

276 Carbonero, F. & Gaskins, H. R. in *Encyclopedia of Metagenomics* (eds S.K. Highlander, F. Rodriguez-Valera, & B.A. White) (Springer, Boston, MA, 2015).

277 Ricaboni, D. *et al. Olsenella provencensis* sp. nov., *Olsenella phocaeensis* sp. nov., and *Olsenella mediterranea* sp. nov. isolated from the human colon. *Hum Microbiome J* **4**, 22-23 (2017).

278 Rey, F. E. *et al.* Metabolic niche of a prominent sulfate-reducing human gut bacterium. *Proc Natl Acad Sci U S A* **110**, 13582-13587, doi:10.1073/pnas.1312524110 (2013).

279 Clavel, T. *et al.* Intestinal microbiota in metabolic diseases: from bacterial community structure and functions to species of pathophysiological relevance. *Gut Microbes* **5**, 544-551, doi:10.4161/gmic.29331 (2014).

280 Major, G. & Spiller, R. Irritable bowel syndrome, inflammatory bowel disease and the microbiome. *Curr Opin Endocrinol Diabetes Obes* **21**, 15-21, doi:10.1097/MED.0000000000000032 (2014).

281 Rizzatti, G., Lopetuso, L. R., Gibiino, G., Binda, C. & Gasbarrini, A. Proteobacteria: A Common Factor in Human Diseases. *Biomed Res Int* **2017**, 9351507, doi:10.1155/2017/9351507 (2017).

282 Shin, N. R., Whon, T. W. & Bae, J. W. Proteobacteria: microbial signature of dysbiosis in gut microbiota. *Trends Biotechnol* **33**, 496-503, doi:10.1016/j.tibtech.2015.06.011 (2015).

283 Gupta, R. S. Origin of diderm (Gram-negative) bacteria: antibiotic selection pressure rather than endosymbiosis likely led to the evolution of bacterial cells with two membranes. *Antonie Van Leeuwenhoek* **100**, 171-182, doi:10.1007/s10482-011-9616-8 (2011).

284 Marchandin, H. *et al.* Negativicoccus succinicivorans gen. nov., sp. nov., isolated from human clinical samples, emended description of the family Veillonellaceae and description of Negativicutes classis nov., Selenomonadales ord. nov. and Acidaminococcaceae fam. nov. in the bacterial phylum Firmicutes. *Int J Syst Evol Microbiol* **60**, 1271-1279, doi:10.1099/ijs.0.013102-0 (2010).

285 Campbell, C., Sutcliffe, I. C. & Gupta, R. S. Comparative proteome analysis of Acidaminococcus intestini supports a relationship between outer membrane

biogenesis in Negativicutes and Proteobacteria. *Arch Microbiol* **196**, 307-310, doi:10.1007/s00203-014-0964-4 (2014).

286 Soo, R. M. *et al.* An expanded genomic representation of the phylum cyanobacteria. *Genome Biol Evol* **6**, 1031-1045, doi:10.1093/gbe/evu073 (2014).

287 Di Rienzi, S. C. *et al.* The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria. *eLife* **2**, e01102, doi:10.7554/eLife.01102 (2013).

288 Brown, D. R. & Bradbury, J. M. in *Mollicutes: Molecular Biology and Pathogenesis* (eds Browning G.F. & C. Citti) (Caister Academic Press, 2014).

289 Cho, I. & Blaser, M. J. The human microbiome: at the interface of health and disease. *Nat Rev Genet* **13**, 260-270, doi:10.1038/nrg3182 (2012).

290 Asshauer, K. P., Wemheuer, B., Daniel, R. & Meinicke, P. Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data. *Bioinformatics* **31**, 2882-2884, doi:10.1093/bioinformatics/btv287 (2015).

291 Dhariwal, A. *et al.* MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Res* **45**, W180-W188, doi:10.1093/nar/gkx295 (2017).

292 Hold, G. L. *et al.* Role of the gut microbiota in inflammatory bowel disease pathogenesis: what have we learnt in the past 10 years? *World J Gastroenterol* **20**, 1192-1210, doi:10.3748/wjg.v20.i5.1192 (2014).

293 Rigottier-Gois, L. Dysbiosis in inflammatory bowel diseases: the oxygen hypothesis. *ISME J* **7**, 1256-1261, doi:10.1038/ismej.2013.80 (2013).

294 Winter, S. E. & Baumler, A. J. Dysbiosis in the inflamed intestine: chance favors the prepared microbe. *Gut Microbes* **5**, 71-73, doi:10.4161/gmic.27129 (2014).

295 Rooks, M. G. *et al.* Gut microbiome composition and function in experimental colitis during active disease and treatment-induced remission. *ISME J* **8**, 1403-1417, doi:10.1038/ismej.2014.3 (2014).

296 van den Elsen, L. W., Poyntz, H. C., Weyrich, L. S., Young, W. & Forbes-Blom, E. E. Embracing the gut microbiota: the new frontier for inflammatory and infectious diseases. *Clin Transl Immunology* **6**, e125, doi:10.1038/cti.2016.91 (2017).

297 Loubinoux, J., Bronowicki, J. P., Pereira, I. A., Mougenel, J. L. & Faou, A. E. Sulfate-reducing bacteria in human feces and their association with inflammatory bowel diseases. *FEMS Microbiol Ecol* **40**, 107-112, doi:10.1111/j.1574-6941.2002.tb00942.x (2002).

298 Roediger, W. E., Duncan, A., Kapaniris, O. & Millard, S. Reducing sulfur compounds of the colon impair colonocyte nutrition: implications for ulcerative colitis. *Gastroenterology* **104**, 802-809 (1993).

299 Roediger, W. E., Moore, J. & Babidge, W. Colonic sulfide in pathogenesis and treatment of ulcerative colitis. *Dig Dis Sci* **42**, 1571-1579 (1997).

300 Rowan, F. *et al.* Desulfovibrio bacterial species are increased in ulcerative colitis. *Dis Colon Rectum* **53**, 1530-1536, doi:10.1007/DCR.0b013e3181f1e620 (2010).

301 Guo, S., Al-Sadi, R., Said, H. M. & Ma, T. Y. Lipopolysaccharide causes an increase in intestinal tight junction permeability in vitro and in vivo by inducing enterocyte membrane expression and localization of TLR-4 and CD14. *Am J Pathol* **182**, 375-387, doi:10.1016/j.ajpath.2012.10.014 (2013).

302 Guo, S. *et al.* Lipopolysaccharide Regulation of Intestinal Tight Junction Permeability Is Mediated by TLR4 Signal Transduction Pathway Activation of FAK and MyD88. *J Immunol* **195**, 4999-5010, doi:10.4049/jimmunol.1402598 (2015).

303 Ni, J. *et al.* A role for bacterial urease in gut dysbiosis and Crohn's disease. *Sci Transl Med* **9**, doi:10.1126/scitranslmed.aah6888 (2017).

304 Gevers, D. *et al.* The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* **15**, 382-392, doi:10.1016/j.chom.2014.02.005 (2014).

305 Leung, J. M. *et al.* IL-22-producing CD4+ cells are depleted in actively inflamed colitis tissue. *Mucosal Immunol* **7**, 124-133, doi:10.1038/mi.2013.31 (2014).

306 Halfvarson, J. *et al.* Dynamics of the human gut microbiome in inflammatory bowel disease. *Nat Microbiol* **2**, 17004, doi:10.1038/nmicrobiol.2017.4 (2017).

307 Morgan, X. C. *et al.* Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol* **13**, R79, doi:10.1186/gb-2012-13-9-r79 (2012).

308 Summerton, C. B., Longlands, M. G., Wiener, K. & Shreeve, D. R. Faecal calprotectin: a marker of inflammation throughout the intestinal tract. *Eur J Gastroenterol Hepatol* **14**, 841-845 (2002).

309 Sady, H. *et al.* Detection of Schistosoma mansoni and Schistosoma haematobium by Real-Time PCR with High Resolution Melting Analysis. *Int J Mol Sci* **16**, 16085-16103, doi:10.3390/ijms160716085 (2015).

310 Llewellyn, S. *et al.* Application of a Multiplex Quantitative PCR to Assess Prevalence and Intensity Of Intestinal Parasite Infections in a Controlled Clinical Trial. *PLoS Negl Trop Dis* **10**, e0004380, doi:10.1371/journal.pntd.0004380 (2016).

311 Gordon, C. A., Gray, D. J., Gobert, G. N. & McManus, D. P. DNA amplification approaches for the diagnosis of key parasitic helminth infections of humans. *Mol Cell Probes* **25**, 143-152, doi:10.1016/j.mcp.2011.05.002 (2011).

312 Holzscheiter, M. *et al.* Lack of host gut microbiota alters immune responses and intestinal granuloma formation during schistosomiasis. *Clin Exp Immunol* **175**, 246-257, doi:10.1111/cei.12230 (2014).

313 McGarvey, S. T. Schistosomiasis: Impact on childhood and adolescent growth, malnutrition, and morbidity. *Semin Pediatr Infect Dis* **11**, 269-274 (2000).

314 Thompson, L. R. *et al.* A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551**, 457-463, doi:10.1038/nature24621 (2017).

315 Walters, W. *et al.* Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. *mSystems* **1**, doi:10.1128/mSystems.00009-15 (2016).

316 Navas-Molina, J. A. *et al.* Advancing our understanding of the human microbiome using QIIME. *Methods Enzymol* **531**, 371-444, doi:10.1016/B978-0-12-407863-5.00019-8 (2013).

317 Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**, 581-583, doi:10.1038/nmeth.3869 (2016).

318 McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**, e61217, doi:10.1371/journal.pone.0061217 (2013).

319 McMurdie, P. J. & Holmes, S. Phyloseq: a bioconductor package for handling and analysis of high-throughput phylogenetic sequence data. *Pac Symp Biocomput*, 235-246 (2012).
320 Dixon, P. VEGAN, a package of R functions for community ecology. *J Veg Sci* **14**, 927-930 (2009).
321 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
322 Okonechnikov, K., Golosova, O., Fursov, M. & team, U. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* **28**, 1166-1167, doi:10.1093/bioinformatics/bts091 (2012).
323 Rozen, S. & Skaletsky, H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**, 365-386 (2000).
324 Kumar, A. & Chordia, N. *In silico* PCR Primer Designing and Validation. *Methods Mol Biol* **1275**, 143-151 (2015).
325 Coordinators, N. R. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **42**, D7-17, doi:10.1093/nar/gkt1146 (2014).
326 Kostic, A. D., Xavier, R. J. & Gevers, D. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* **146**, 1489-1499, doi:10.1053/j.gastro.2014.02.009 (2014).
327 Ahn, J. *et al.* Human gut microbiome and risk for colorectal cancer. *J Natl Cancer Inst* **105**, 1907-1911, doi:10.1093/jnci/djt300 (2013).
328 He, Z., Shao, T., Li, H., Xie, Z. & Wen, C. Alterations of the gut microbiome in Chinese patients with systemic lupus erythematosus. *Gut Pathog* **8**, 64, doi:10.1186/s13099-016-0146-9 (2016).
329 Scher, J. U., Littman, D. R. & Abramson, S. B. Microbiome in Inflammatory Arthritis and Human Rheumatic Diseases. *Arthritis Rheumatol* **68**, 35-45, doi:10.1002/art.39259 (2016).
330 Brown, C. T. *et al.* Gut microbiome metagenomics analysis suggests a functional model for the development of autoimmunity for type 1 diabetes. *PLoS One* **6**, e25792, doi:10.1371/journal.pone.0025792 (2011).
331 de Paiva, C. S. *et al.* Altered Mucosal Microbiome Diversity and Disease Severity in Sjogren Syndrome. *Sci Rep* **6**, 23561, doi:10.1038/srep23561 (2016).
332 Huang, S., Mao, J., Zhou, L., Xiong, X. & Deng, Y. The imbalance of gut microbiota and its correlation with plasma inflammatory cytokines in pemphigus vulgaris patients. *Scand J Immunol* **90**, e12799, doi:10.1111/sji.12799 (2019).
333 Kasselman, L. J., Vernice, N. A., DeLeon, J. & Reiss, A. B. The gut microbiome and elevated cardiovascular risk in obesity and autoimmunity. *Atherosclerosis* **271**, 203-213, doi:10.1016/j.atherosclerosis.2018.02.036 (2018).
334 Arora, T. & Backhed, F. The gut microbiota and metabolic disease: current understanding and future perspectives. *J Intern Med* **280**, 339-349, doi:10.1111/joim.12508 (2016).
335 Jie, Z. *et al.* The gut microbiome in atherosclerotic cardiovascular disease. *Nat Commun* **8**, 845, doi:10.1038/s41467-017-00900-1 (2017).

336 Foster, J. A. & McVey Neufeld, K. A. Gut-brain axis: how the microbiome influences anxiety and depression. *Trends Neurosci* **36**, 305-312, doi:10.1016/j.tins.2013.01.005 (2013).

337 Jiang, H. Y. *et al.* Altered gut microbiota profile in patients with generalized anxiety disorder. *J Psychiatr Res* **104**, 130-136, doi:10.1016/j.jpsychires.2018.07.007 (2018).

338 Chen, Y. H. *et al.* Association between fecal microbiota and generalized anxiety disorder: Severity and early treatment response. *J Affect Disord* **259**, 56-66, doi:10.1016/j.jad.2019.08.014 (2019).

339 Evans, S. J. *et al.* The gut microbiome composition associates with bipolar disorder and illness severity. *J Psychiatr Res* **87**, 23-29, doi:10.1016/j.jpsychires.2016.12.007 (2017).

340 Coello, K. *et al.* Gut microbiota composition in patients with newly diagnosed bipolar disorder and their unaffected first-degree relatives. *Brain Behav Immun* **75**, 112-118, doi:10.1016/j.bbi.2018.09.026 (2019).

341 Dickerson, F., Severance, E. & Yolken, R. The microbiome, immunity, and schizophrenia and bipolar disorder. *Brain Behav Immun* **62**, 46-52, doi:10.1016/j.bbi.2016.12.010 (2017).

342 Nguyen, T. T., Kosciolek, T., Eyler, L. T., Knight, R. & Jeste, D. V. Overview and systematic review of studies of microbiome in schizophrenia and bipolar disorder. *J Psychiatr Res* **99**, 50-61, doi:10.1016/j.jpsychires.2018.01.013 (2018).

343 Jiang, H. *et al.* Altered fecal microbiota composition in patients with major depressive disorder. *Brain Behav Immun* **48**, 186-194, doi:10.1016/j.bbi.2015.03.016 (2015).

344 Huang, Y. *et al.* Possible association of Firmicutes in the gut microbiota of patients with major depressive disorder. *Neuropsychiatr Dis Treat* **14**, 3329-3337, doi:10.2147/NDT.S188340 (2018).

345 Zheng, P. *et al.* Gut microbiome remodeling induces depressive-like behaviors through a pathway mediated by the host's metabolism. *Mol Psychiatry* **21**, 786-796, doi:10.1038/mp.2016.44 (2016).

346 Naseribafrouei, A. *et al.* Correlation between the human fecal microbiota and depression. *Neurogastroenterol Motil* **26**, 1155-1162, doi:10.1111/nmo.12378 (2014).

347 Chen, Z. *et al.* Comparative metaproteomics analysis shows altered fecal microbiota signatures in patients with major depressive disorder. *Neuroreport* **29**, 417-425, doi:10.1097/WNR.0000000000000985 (2018).

348 Valles-Colomer, M. *et al.* The neuroactive potential of the human gut microbiota in quality of life and depression. *Nat Microbiol* **4**, 623-632, doi:10.1038/s41564-018-0337-x (2019).

349 Lai, W. T. *et al.* Shotgun metagenomics reveals both taxonomic and tryptophan pathway differences of gut microbiota in major depressive disorder patients. *Psychol Med*, 1-12, doi:10.1017/S0033291719003027 (2019).

350 Rong, H. *et al.* Similarly in depression, nuances of gut microbiota: Evidences from a shotgun metagenomics sequencing study on major depressive disorder versus bipolar disorder with current major depressive episode patients. *J Psychiatr Res* **113**, 90-99, doi:10.1016/j.jpsychires.2019.03.017 (2019).

351 Lin, P. *et al.* Prevotella and Klebsiella proportions in fecal microbial communities are potential characteristic parameters for patients with major depressive disorder. *J Affect Disord* **207**, 300-304, doi:10.1016/j.jad.2016.09.051 (2017).
352 Sharon, G., Sampson, T. R., Geschwind, D. H. & Mazmanian, S. K. The Central Nervous System and the Gut Microbiome. *Cell* **167**, 915-932, doi:10.1016/j.cell.2016.10.027 (2016).
353 Carabotti, M., Scirocco, A., Maselli, M. A. & Severi, C. The gut-brain axis: interactions between enteric microbiota, central and enteric nervous systems. *Ann Gastroenterol* **28**, 203-209 (2015).
354 O'Mahony, S. M., Clarke, G., Borre, Y. E., Dinan, T. G. & Cryan, J. F. Serotonin, tryptophan metabolism and the brain-gut-microbiome axis. *Behav Brain Res* **277**, 32-48, doi:10.1016/j.bbr.2014.07.027 (2015).
355 Kelly, J. R., Clarke, G., Cryan, J. F. & Dinan, T. G. Brain-gut-microbiota axis: challenges for translation in psychiatry. *Ann Epidemiol* **26**, 366-372, doi:10.1016/j.annepidem.2016.02.008 (2016).
356 Foster, J. A., Rinaman, L. & Cryan, J. F. Stress & the gut-brain axis: Regulation by the microbiome. *Neurobiol Stress* **7**, 124-136, doi:10.1016/j.ynstr.2017.03.001 (2017).
357 Mayer, E. A., Tillisch, K. & Gupta, A. Gut/brain axis and the microbiota. *J Clin Invest* **125**, 926-938, doi:10.1172/JCI76304 (2015).
358 Dinan, T. G., Stilling, R. M., Stanton, C. & Cryan, J. F. Collective unconscious: how gut microbes shape human behavior. *J Psychiatr Res* **63**, 1-9, doi:10.1016/j.jpsychires.2015.02.021 (2015).
359 Bonaz, B., Sinniger, V. & Pellissier, S. Vagus Nerve Stimulation at the Interface of Brain-Gut Interactions. *Cold Spring Harb Perspect Med* **9**, doi:10.1101/cshperspect.a034199 (2019).
360 Forsythe, P., Bienenstock, J. & Kunze, W. A. Vagal pathways for microbiome-brain-gut axis communication. *Adv Exp Med Biol* **817**, 115-133, doi:10.1007/978-1-4939-0897-4_5 (2014).
361 Caspani, G., Kennedy, S., Foster, J. A. & Swann, J. Gut microbial metabolites in depression: understanding the biochemical mechanisms. *Microb Cell* **6**, 454-481, doi:10.15698/mic2019.10.693 (2019).
362 Raybould, H. E. Gut chemosensing: interactions between gut endocrine cells and visceral afferents. *Auton Neurosci* **153**, 41-46, doi:10.1016/j.autneu.2009.07.007 (2010).
363 Sarkar, A. *et al.* Psychobiotics and the Manipulation of Bacteria-Gut-Brain Signals. *Trends Neurosci* **39**, 763-781, doi:10.1016/j.tins.2016.09.002 (2016).
364 O'Mahony, S. M. *et al.* Early life stress alters behavior, immunity, and microbiota in rats: implications for irritable bowel syndrome and psychiatric illnesses. *Biol Psychiatry* **65**, 263-267, doi:10.1016/j.biopsych.2008.06.026 (2009).
365 Sudo, N. *et al.* Postnatal microbial colonization programs the hypothalamic-pituitary-adrenal system for stress response in mice. *J Physiol* **558**, 263-275, doi:10.1113/jphysiol.2004.063388 (2004).
366 Wikoff, W. R. *et al.* Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc Natl Acad Sci U S A* **106**, 3698-3703, doi:10.1073/pnas.0812874106 (2009).

367 Yano, J. M. *et al.* Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis. *Cell* **161**, 264-276, doi:10.1016/j.cell.2015.02.047 (2015).

368 Barrett, E., Ross, R. P., O'Toole, P. W., Fitzgerald, G. F. & Stanton, C. gamma-Aminobutyric acid production by culturable bacteria from the human intestine. *J Appl Microbiol* **113**, 411-417, doi:10.1111/j.1365-2672.2012.05344.x (2012).

369 Borovikova, L. V. *et al.* Vagus nerve stimulation attenuates the systemic inflammatory response to endotoxin. *Nature* **405**, 458-462, doi:10.1038/35013070 (2000).

370 Johnston, G. R. & Webster, N. R. Cytokines and the immunomodulatory function of the vagus nerve. *Br J Anaesth* **102**, 453-462, doi:10.1093/bja/aep037 (2009).

371 Park, A. J. *et al.* Altered colonic function and microbiota profile in a mouse model of chronic depression. *Neurogastroenterol Motil* **25**, 733-e575, doi:10.1111/nmo.12153 (2013).

372 de Punder, K. & Pruimboom, L. Stress induces endotoxemia and low-grade inflammation by increasing barrier permeability. *Front Immunol* **6**, 223, doi:10.3389/fimmu.2015.00223 (2015).

373 Vicario, M. *et al.* Chronological assessment of mast cell-mediated gut dysfunction and mucosal inflammation in a rat model of chronic psychosocial stress. *Brain Behav Immun* **24**, 1166-1175, doi:10.1016/j.bbi.2010.06.002 (2010).

374 Kelly, J. R. *et al.* Breaking down the barriers: the gut microbiome, intestinal permeability and stress-related psychiatric disorders. *Front Cell Neurosci* **9**, 392, doi:10.3389/fncel.2015.00392 (2015).

375 Rodino-Janeiro, B. K. *et al.* Role of Corticotropin-releasing Factor in Gastrointestinal Permeability. *J Neurogastroenterol Motil* **21**, 33-50, doi:10.5056/jnm14084 (2015).

376 Moussaoui, N. *et al.* Changes in intestinal glucocorticoid sensitivity in early life shape the risk of epithelial barrier defect in maternal-deprived rats. *PLoS One* **9**, e88382, doi:10.1371/journal.pone.0088382 (2014).

377 Soderholm, J. D. *et al.* Neonatal maternal separation predisposes adult rats to colonic barrier dysfunction in response to mild stress. *Am J Physiol Gastrointest Liver Physiol* **283**, G1257-1263, doi:10.1152/ajpgi.00314.2002 (2002).

378 Rhee, S. H., Pothoulakis, C. & Mayer, E. A. Principles and clinical implications of the brain-gut-enteric microbiota axis. *Nat Rev Gastroenterol Hepatol* **6**, 306-314, doi:10.1038/nrgastro.2009.35 (2009).

379 Martin, C. R., Osadchiy, V., Kalani, A. & Mayer, E. A. The Brain-Gut-Microbiome Axis. *Cell Mol Gastroenterol Hepatol* **6**, 133-148, doi:10.1016/j.jcmgh.2018.04.003 (2018).

380 Mayer, E. A. Gut feelings: the emerging biology of gut-brain communication. *Nat Rev Neurosci* **12**, 453-466, doi:10.1038/nrn3071 (2011).

381 Huang, T. T. *et al.* Current Understanding of Gut Microbiota in Mood Disorders: An Update of Human Studies. *Front Genet* **10**, 98, doi:10.3389/fgene.2019.00098 (2019).

382 Dinan, T. G. & Cryan, J. F. Melancholic microbes: a link between gut microbiota and depression? *Neurogastroenterol Motil* **25**, 713-719, doi:10.1111/nmo.12198 (2013).

383 Evrensel, A. & Ceylan, M. E. The Gut-Brain Axis: The Missing Link in Depression. *Clin Psychopharmacol Neurosci* **13**, 239-244, doi:10.9758/cpn.2015.13.3.239 (2015).
384 Lach, G., Schellekens, H., Dinan, T. G. & Cryan, J. F. Anxiety, Depression, and the Microbiome: A Role for Gut Peptides. *Neurotherapeutics* **15**, 36-59, doi:10.1007/s13311-017-0585-0 (2018).
385 Kelly, J. R. *et al.* Transferring the blues: Depression-associated gut microbiota induces neurobehavioural changes in the rat. *J Psychiatr Res* **82**, 109-118, doi:10.1016/j.jpsychires.2016.07.019 (2016).
386 Hao, Z., Wang, W., Guo, R. & Liu, H. Faecalibacterium prausnitzii (ATCC 27766) has preventive and therapeutic effects on chronic unpredictable mild stress-induced depression-like and anxiety-like behavior in rats. *Psychoneuroendocrinology* **104**, 132-142, doi:10.1016/j.psyneuen.2019.02.025 (2019).
387 Bravo, J. A. *et al.* Ingestion of Lactobacillus strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. *Proc Natl Acad Sci U S A* **108**, 16050-16055, doi:10.1073/pnas.1102999108 (2011).
388 Bercik, P. *et al.* The anxiolytic effect of Bifidobacterium longum NCC3001 involves vagal pathways for gut-brain communication. *Neurogastroenterol Motil* **23**, 1132-1139, doi:10.1111/j.1365-2982.2011.01796.x (2011).
389 Felger, J. C. & Lotrich, F. E. Inflammatory cytokines in depression: neurobiological mechanisms and therapeutic implications. *Neuroscience* **246**, 199-229, doi:10.1016/j.neuroscience.2013.04.060 (2013).
390 Amodeo, G. T., M.A.; Fagiolini, A. Depression and inflammation: Disentangling a clear yet complex and multifaceted link. *Neuropsychiatry* **7**, 448-457 (2017).
391 Miller, A. H. & Raison, C. L. The role of inflammation in depression: from evolutionary imperative to modern treatment target. *Nat Rev Immunol* **16**, 22-34, doi:10.1038/nri.2015.5 (2016).
392 Dowlati, Y. *et al.* A meta-analysis of cytokines in major depression. *Biol Psychiatry* **67**, 446-457, doi:10.1016/j.biopsych.2009.09.033 (2010).
393 Kohler, C. A. *et al.* Peripheral cytokine and chemokine alterations in depression: a meta-analysis of 82 studies. *Acta Psychiatr Scand* **135**, 373-387, doi:10.1111/acps.12698 (2017).
394 Haapakoski, R., Mathieu, J., Ebmeier, K. P., Alenius, H. & Kivimaki, M. Cumulative meta-analysis of interleukins 6 and 1beta, tumour necrosis factor alpha and C-reactive protein in patients with major depressive disorder. *Brain Behav Immun* **49**, 206-215, doi:10.1016/j.bbi.2015.06.001 (2015).
395 Liu, Y., Ho, R. C. & Mak, A. Interleukin (IL)-6, tumour necrosis factor alpha (TNF-alpha) and soluble interleukin-2 receptors (sIL-2R) are elevated in patients with major depressive disorder: a meta-analysis and meta-regression. *J Affect Disord* **139**, 230-239, doi:10.1016/j.jad.2011.08.003 (2012).
396 Goldsmith, D. R., Rapaport, M. H. & Miller, B. J. A meta-analysis of blood cytokine network alterations in psychiatric patients: comparisons between schizophrenia, bipolar disorder and depression. *Mol Psychiatry* **21**, 1696-1709, doi:10.1038/mp.2016.3 (2016).

397 Himmerich, H., Patsalos, O., Lichtblau, N., Ibrahim, M. A. A. & Dalton, B. Cytokine Research in Depression: Principles, Challenges, and Open Questions. *Front Psychiatry* **10**, 30, doi:10.3389/fpsyt.2019.00030 (2019).

398 Burke, H. M., Davis, M. C., Otte, C. & Mohr, D. C. Depression and cortisol responses to psychological stress: a meta-analysis. *Psychoneuroendocrinology* **30**, 846-856, doi:10.1016/j.psyneuen.2005.02.010 (2005).

399 Carroll, B. J. *et al.* Pathophysiology of hypercortisolism in depression. *Acta Psychiatr Scand Suppl*, 90-103, doi:10.1111/j.1600-0447.2007.00967.x (2007).

400 Parker, K. J., Schatzberg, A. F. & Lyons, D. M. Neuroendocrine aspects of hypercortisolism in major depression. *Horm Behav* **43**, 60-66, doi:10.1016/s0018-506x(02)00016-8 (2003).

401 Mendlewicz, J. *et al.* Shortened onset of action of antidepressants in major depression using acetylsalicylic acid augmentation: a pilot open-label study. *Int Clin Psychopharmacol* **21**, 227-231, doi:10.1097/00004850-200607000-00005 (2006).

402 Muller, N. *et al.* The cyclooxygenase-2 inhibitor celecoxib has therapeutic effects in major depression: results of a double-blind, randomized, placebo controlled, add-on pilot study to reboxetine. *Mol Psychiatry* **11**, 680-684, doi:10.1038/sj.mp.4001805 (2006).

403 Wong, M. L. *et al.* Inflammasome signaling affects anxiety- and depressive-like behavior and gut microbiome composition. *Mol Psychiatry* **21**, 797-805, doi:10.1038/mp.2016.46 (2016).

404 Zhang, Y. *et al.* NLRP3 Inflammasome Mediates Chronic Mild Stress-Induced Depression in Mice via Neuroinflammation. *Int J Neuropsychopharmacol* **18**, doi:10.1093/ijnp/pyv006 (2015).

405 Alcocer-Gomez, E. *et al.* Stress-Induced Depressive Behaviors Require a Functional NLRP3 Inflammasome. *Mol Neurobiol* **53**, 4874-4882, doi:10.1007/s12035-015-9408-7 (2016).

406 Iwata, M., Ota, K. T. & Duman, R. S. The inflammasome: pathways linking psychological stress, depression, and systemic illnesses. *Brain Behav Immun* **31**, 105-114, doi:10.1016/j.bbi.2012.12.008 (2013).

407 Kaufmann, F. N. *et al.* NLRP3 inflammasome-driven pathways in depression: Clinical and preclinical findings. *Brain Behav Immun* **64**, 367-383, doi:10.1016/j.bbi.2017.03.002 (2017).

408 Barnard, K. D., Skinner, T. C. & Peveler, R. The prevalence of co-morbid depression in adults with Type 1 diabetes: systematic literature review. *Diabet Med* **23**, 445-448, doi:10.1111/j.1464-5491.2006.01814.x (2006).

409 Dickens, C., McGowan, L., Clark-Carter, D. & Creed, F. Depression in rheumatoid arthritis: a systematic review of the literature with meta-analysis. *Psychosom Med* **64**, 52-60, doi:10.1097/00006842-200201000-00008 (2002).

410 Figueiredo-Braga, M. *et al.* Depression and anxiety in systemic lupus erythematosus: The crosstalk between immunological, clinical, and psychosocial factors. *Medicine (Baltimore)* **97**, e11376, doi:10.1097/MD.0000000000011376 (2018).

411 Byrne, G. *et al.* Prevalence of Anxiety and Depression in Patients with Inflammatory Bowel Disease. *Can J Gastroenterol Hepatol* **2017**, 6496727, doi:10.1155/2017/6496727 (2017).

412 Caneo, C., Marston, L., Bellon, J. A. & King, M. Examining the relationship between physical illness and depression: Is there a difference between inflammatory and non inflammatory diseases? A cohort study. *Gen Hosp Psychiatry* **43**, 71-77, doi:10.1016/j.genhosppsych.2016.09.007 (2016).

413 Dantzer, R. Cytokine-induced sickness behaviour: a neuroimmune response to activation of innate immunity. *Eur J Pharmacol* **500**, 399-411, doi:10.1016/j.ejphar.2004.07.040 (2004).

414 Raison, C. L. & Miller, A. H. Do cytokines really sing the blues? *Cerebrum* **2013**, 10 (2013).

415 Miyaoka, H. *et al.* Depression from interferon therapy in patients with hepatitis C. *Am J Psychiatry* **156**, 1120, doi:10.1176/ajp.156.7.1120 (1999).

416 O'Connor, J. C. *et al.* Lipopolysaccharide-induced depressive-like behavior is mediated by indoleamine 2,3-dioxygenase activation in mice. *Mol Psychiatry* **14**, 511-522, doi:10.1038/sj.mp.4002148 (2009).

417 Ohgi, Y., Futamura, T., Kikuchi, T. & Hashimoto, K. Effects of antidepressants on alternations in serum cytokines and depressive-like behavior in mice after lipopolysaccharide administration. *Pharmacol Biochem Behav* **103**, 853-859, doi:10.1016/j.pbb.2012.12.003 (2013).

418 Zhang, Y. *et al.* Involvement of inflammasome activation in lipopolysaccharide-induced mice depressive-like behaviors. *CNS Neurosci Ther* **20**, 119-124, doi:10.1111/cns.12170 (2014).

419 Fu, X. *et al.* Central administration of lipopolysaccharide induces depressive-like behavior in vivo and activates brain indoleamine 2,3 dioxygenase in murine organotypic hippocampal slice cultures. *J Neuroinflammation* **7**, 43, doi:10.1186/1742-2094-7-43 (2010).

420 Anisman, H., Merali, Z., Poulter, M. O. & Hayley, S. Cytokines as a precipitant of depressive illness: animal and human studies. *Curr Pharm Des* **11**, 963-972, doi:10.2174/1381612053381701 (2005).

421 Schiepers, O. J., Wichers, M. C. & Maes, M. Cytokines and major depression. *Prog Neuropsychopharmacol Biol Psychiatry* **29**, 201-217, doi:10.1016/j.pnpbp.2004.11.003 (2005).

422 Rochfort, K. D., Collins, L. E., Murphy, R. P. & Cummins, P. M. Downregulation of blood-brain barrier phenotype by proinflammatory cytokines involves NADPH oxidase-dependent ROS generation: consequences for interendothelial adherens and tight junctions. *PLoS One* **9**, e101815, doi:10.1371/journal.pone.0101815 (2014).

423 Farkas, G. *et al.* Experimental acute pancreatitis results in increased blood-brain barrier permeability in the rat: a potential role for tumor necrosis factor and interleukin 6. *Neurosci Lett* **242**, 147-150, doi:10.1016/s0304-3940(98)00060-3 (1998).

424 de Vries, H. E. *et al.* The influence of cytokines on the integrity of the blood-brain barrier in vitro. *J Neuroimmunol* **64**, 37-43, doi:10.1016/0165-5728(95)00148-4 (1996).

425 Plotkin, S. R., Banks, W. A. & Kastin, A. J. Comparison of saturable transport and extracellular pathways in the passage of interleukin-1 alpha across the blood-brain barrier. *J Neuroimmunol* **67**, 41-47, doi:10.1016/0165-5728(96)00036-7 (1996).
426 Watkins, L. R., Maier, S. F. & Goehler, L. E. Cytokine-to-brain communication: a review & analysis of alternative mechanisms. *Life Sci* **57**, 1011-1026, doi:10.1016/0024-3205(95)02047-m (1995).
427 Miller, A. H., Maletic, V. & Raison, C. L. Inflammation and its discontents: the role of cytokines in the pathophysiology of major depression. *Biol Psychiatry* **65**, 732-741, doi:10.1016/j.biopsych.2008.11.029 (2009).
428 Pavlov, V. A. & Tracey, K. J. The vagus nerve and the inflammatory reflex--linking immunity and metabolism. *Nat Rev Endocrinol* **8**, 743-754, doi:10.1038/nrendo.2012.189 (2012).
429 Steinberg, B. E. *et al.* Cytokine-specific Neurograms in the Sensory Vagus Nerve. *Bioelectron Med* **3**, 7-17 (2016).
430 Zanos, T. P. *et al.* Identification of cytokine-specific sensory neural signals by decoding murine vagus nerve activity. *Proc Natl Acad Sci U S A* **115**, E4843-E4852, doi:10.1073/pnas.1719083115 (2018).
431 Galic, M. A., Riazi, K. & Pittman, Q. J. Cytokines and brain excitability. *Front Neuroendocrinol* **33**, 116-125, doi:10.1016/j.yfrne.2011.12.002 (2012).
432 De Simoni, M. G. & Imeri, L. Cytokine-neurotransmitter interactions in the brain. *Biol Signals Recept* **7**, 33-44, doi:10.1159/000014526 (1998).
433 Miller, A. H., Haroon, E., Raison, C. L. & Felger, J. C. Cytokine targets in the brain: impact on neurotransmitters and neurocircuits. *Depress Anxiety* **30**, 297-306, doi:10.1002/da.22084 (2013).
434 Gadek-Michalska, A. T., J.; Rachwalska, P.; Bugjaski, J. Cytokines, prostaglandins and nitric oxide in the regulation of stress-response systems. *Pharmacological Reports* **65**, 1655-1662 (2013).
435 Dantzer, R., Wollman, E., Vitkovic, L. & Yirmiya, R. Cytokines and depression: fortuitous or causative association? *Mol Psychiatry* **4**, 328-332, doi:10.1038/sj.mp.4000572 (1999).
436 Zunszain, P. A., Anacker, C., Cattaneo, A., Carvalho, L. A. & Pariante, C. M. Glucocorticoids, cytokines and brain abnormalities in depression. *Prog Neuropsychopharmacol Biol Psychiatry* **35**, 722-729, doi:10.1016/j.pnpbp.2010.04.011 (2011).
437 Humphreys, D., Schlesinger, L., Lopez, M. & Araya, A. V. Interleukin-6 production and deregulation of the hypothalamic-pituitary-adrenal axis in patients with major depressive disorders. *Endocrine* **30**, 371-376 (2006).
438 Pace, T. W., Hu, F. & Miller, A. H. Cytokine-effects on glucocorticoid receptor function: relevance to glucocorticoid resistance and the pathophysiology and treatment of major depression. *Brain Behav Immun* **21**, 9-19, doi:10.1016/j.bbi.2006.08.009 (2007).
439 Miller, A. H., Pariante, C. M. & Pearce, B. D. Effects of cytokines on glucocorticoid receptor expression and function. Glucocorticoid resistance and relevance to depression. *Adv Exp Med Biol* **461**, 107-116, doi:10.1007/978-0-585-37970-8_7 (1999).

440 Capuron, L. *et al.* Association between decreased serum tryptophan concentrations and depressive symptoms in cancer patients undergoing cytokine therapy. *Mol Psychiatry* **7**, 468-473, doi:10.1038/sj.mp.4000995 (2002).
441 Dantzer, R., O'Connor, J. C., Lawson, M. A. & Kelley, K. W. Inflammation-associated depression: from serotonin to kynurenine. *Psychoneuroendocrinology* **36**, 426-436, doi:10.1016/j.psyneuen.2010.09.012 (2011).
442 Maes, M. *et al.* Serotonin-immune interactions in major depression: lower serum tryptophan as a marker of an immune-inflammatory response. *Eur Arch Psychiatry Clin Neurosci* **247**, 154-161, doi:10.1007/bf03033069 (1997).
443 Myint, A. M. & Kim, Y. K. Cytokine-serotonin interaction through IDO: a neurodegeneration hypothesis of depression. *Med Hypotheses* **61**, 519-525, doi:10.1016/s0306-9877(03)00207-x (2003).
444 Riedel, W. J., Klaassen, T. & Schmitt, J. A. Tryptophan, mood, and cognitive function. *Brain Behav Immun* **16**, 581-589, doi:10.1016/s0889-1591(02)00013-2 (2002).
445 Heyes, M. P. *et al.* Quinolinic acid and kynurenine pathway metabolism in inflammatory and non-inflammatory neurological disease. *Brain* **115 (Pt 5)**, 1249-1273, doi:10.1093/brain/115.5.1249 (1992).
446 Maes, M., Kubera, M. & Leunis, J. C. The gut-brain barrier in major depression: intestinal mucosal dysfunction with an increased translocation of LPS from gram negative enterobacteria (leaky gut) plays a role in the inflammatory pathophysiology of depression. *Neuro Endocrinol Lett* **29**, 117-124 (2008).
447 Maes, M., Kubera, M., Leunis, J. C. & Berk, M. Increased IgA and IgM responses against gut commensals in chronic depression: further evidence for increased bacterial translocation or leaky gut. *J Affect Disord* **141**, 55-62, doi:10.1016/j.jad.2012.02.023 (2012).
448 Maes, M. *et al.* In depression, bacterial translocation may drive inflammatory responses, oxidative and nitrosative stress (O&NS), and autoimmune responses directed against O&NS-damaged neoepitopes. *Acta Psychiatr Scand* **127**, 344-354, doi:10.1111/j.1600-0447.2012.01908.x (2013).
449 Keri, S., Szabo, C. & Kelemen, O. Expression of Toll-Like Receptors in peripheral blood mononuclear cells and response to cognitive-behavioral therapy in major depressive disorder. *Brain Behav Immun* **40**, 235-243, doi:10.1016/j.bbi.2014.03.020 (2014).
450 Berk, M. *et al.* So depression is an inflammatory disease, but where does the inflammation come from? *BMC Med* **11**, 200, doi:10.1186/1741-7015-11-200 (2013).
451 Kuehner, C. Gender differences in unipolar depression: an update of epidemiological findings and possible explanations. *Acta Psychiatr Scand* **108**, 163-174, doi:10.1034/j.1600-0447.2003.00204.x (2003).
452 Weissman, M. M. *et al.* Cross-national epidemiology of major depression and bipolar disorder. *JAMA* **276**, 293-299 (1996).
453 Connolly, M. D., Zervos, M. J., Barone, C. J., 2nd, Johnson, C. C. & Joseph, C. L. The Mental Health of Transgender Youth: Advances in Understanding. *J Adolesc Health* **59**, 489-495, doi:10.1016/j.jadohealth.2016.06.012 (2016).

454 Reisner, S. L. *et al.* Mental health of transgender youth in care at an adolescent urban community health center: a matched retrospective cohort study. *J Adolesc Health* **56**, 274-279, doi:10.1016/j.jadohealth.2014.10.264 (2015).
455 Olson, J., Schrager, S. M., Belzer, M., Simons, L. K. & Clark, L. F. Baseline Physiologic and Psychosocial Characteristics of Transgender Youth Seeking Care for Gender Dysphoria. *J Adolesc Health* **57**, 374-380, doi:10.1016/j.jadohealth.2015.04.027 (2015).
456 Vital, M., Howe, A. C. & Tiedje, J. M. Revealing the bacterial butyrate synthesis pathways by analyzing (meta)genomic data. *MBio* **5**, e00889, doi:10.1128/mBio.00889-14 (2014).
457 Anand, S., Kaur, H. & Mande, S. S. Comparative In silico Analysis of Butyrate Production Pathways in Gut Commensals and Pathogens. *Front Microbiol* **7**, 1945, doi:10.3389/fmicb.2016.01945 (2016).
458 d'Hennezel, E., Abubucker, S., Murphy, L. O. & Cullen, T. W. Total Lipopolysaccharide from the Human Gut Microbiome Silences Toll-Like Receptor Signaling. *mSystems* **2**, doi:10.1128/mSystems.00046-17 (2017).
459 Vatanen, T. *et al.* Variation in Microbiome LPS Immunogenicity Contributes to Autoimmunity in Humans. *Cell* **165**, 842-853, doi:10.1016/j.cell.2016.04.007 (2016).
460 Gupta, V. K., Paul, S. & Dutta, C. Geography, Ethnicity or Subsistence-Specific Variations in Human Microbiome Composition and Diversity. *Front Microbiol* **8**, 1162, doi:10.3389/fmicb.2017.01162 (2017).
461 Yatsunenko, T. *et al.* Human gut microbiome viewed across age and geography. *Nature* **486**, 222-227, doi:10.1038/nature11053 (2012).
462 Ou, J. *et al.* Diet, microbiota, and microbial metabolites in colon cancer risk in rural Africans and African Americans. *Am J Clin Nutr* **98**, 111-120, doi:10.3945/ajcn.112.056689 (2013).
463 Lin, A. *et al.* Distinct distal gut microbiome diversity and composition in healthy children from Bangladesh and the United States. *PLoS One* **8**, e53838, doi:10.1371/journal.pone.0053838 (2013).
464 Odamaki, T. *et al.* Age-related changes in gut microbiota composition from newborn to centenarian: a cross-sectional study. *BMC Microbiol* **16**, 90, doi:10.1186/s12866-016-0708-5 (2016).
465 Quevrain, E. *et al.* Identification of an anti-inflammatory protein from Faecalibacterium prausnitzii, a commensal bacterium deficient in Crohn's disease. *Gut* **65**, 415-425, doi:10.1136/gutjnl-2014-307649 (2016).
466 Louis, P. & Flint, H. J. Formation of propionate and butyrate by the human colonic microbiota. *Environ Microbiol* **19**, 29-41, doi:10.1111/1462-2920.13589 (2017).
467 Hamer, H. M. *et al.* Review article: the role of butyrate on colonic function. *Aliment Pharmacol Ther* **27**, 104-119, doi:10.1111/j.1365-2036.2007.03562.x (2008).
468 Leonel, A. J. & Alvarez-Leite, J. I. Butyrate: implications for intestinal function. *Curr Opin Clin Nutr Metab Care* **15**, 474-479, doi:10.1097/MCO.0b013e32835665fa (2012).

469 Parada Venegas, D. *et al.* Short Chain Fatty Acids (SCFAs)-Mediated Gut Epithelial and Immune Regulation and Its Relevance for Inflammatory Bowel Diseases. *Front Immunol* **10**, 277, doi:10.3389/fimmu.2019.00277 (2019).
470 VanHook, A. M. Butyrate benefits the intestinal barrier. *Science Signaling* **8**, ec135 (2015).
471 Wang, H. B., Wang, P. Y., Wang, X., Wan, Y. L. & Liu, Y. C. Butyrate enhances intestinal epithelial barrier function via up-regulation of tight junction protein Claudin-1 transcription. *Dig Dis Sci* **57**, 3126-3135, doi:10.1007/s10620-012-2259-4 (2012).
472 Zhang, M. *et al.* Butyrate inhibits interleukin-17 and generates Tregs to ameliorate colorectal colitis in rats. *BMC Gastroenterol* **16**, 84, doi:10.1186/s12876-016-0500-x (2016).
473 Kabeerdoss, J., Sankaran, V., Pugazhendhi, S. & Ramakrishna, B. S. Clostridium leptum group bacteria abundance and diversity in the fecal microbiota of patients with inflammatory bowel disease: a case-control study in India. *BMC Gastroenterol* **13**, 20, doi:10.1186/1471-230X-13-20 (2013).
474 Quince, C. *et al.* Extensive Modulation of the Fecal Metagenome in Children With Crohn's Disease During Exclusive Enteral Nutrition. *Am J Gastroenterol* **110**, 1718-1729; quiz 1730, doi:10.1038/ajg.2015.357 (2015).
475 Takahashi, K. *et al.* Reduced Abundance of Butyrate-Producing Bacteria Species in the Fecal Microbial Community in Crohn's Disease. *Digestion* **93**, 59-65, doi:10.1159/000441768 (2016).
476 Kang, S. *et al.* Dysbiosis of fecal microbiota in Crohn's disease patients as revealed by a custom phylogenetic microarray. *Inflamm Bowel Dis* **16**, 2034-2042, doi:10.1002/ibd.21319 (2010).
477 Papa, E. *et al.* Non-invasive mapping of the gastrointestinal microbiota identifies children with inflammatory bowel disease. *PLoS One* **7**, e39242, doi:10.1371/journal.pone.0039242 (2012).
478 Freier, T. A., Beitz, D. C., Li, L. & Hartman, P. A. Characterization of Eubacterium coprostanoligenes sp. nov., a cholesterol-reducing anaerobe. *Int J Syst Bacteriol* **44**, 137-142, doi:10.1099/00207713-44-1-137 (1994).
479 Ren, D., Li, L., Schwabacher, A. W., Young, J. W. & Beitz, D. C. Mechanism of cholesterol reduction to coprostanol by Eubacterium coprostanoligenes ATCC 51222. *Steroids* **61**, 33-40, doi:10.1016/0039-128x(95)00173-n (1996).
480 Sundqvist, T. S., L.; Tjellstrom, B.; Magnusson, K.; Midtvedt, T.; Norin, E.; Hogberg, L. Evidence of disturbed gut microbial metabolic activity in pediatric Crohn’s disease. *Crohn’s & Colitis 360* **1**, otz010 (2019).
481 Reddy, B. S., Martin, C. W. & Wynder, E. L. Fecal bile acids and cholesterol metabolites of patients with ulcerative colitis, a high-risk group for development of colon cancer. *Cancer Res* **37**, 1697-1701 (1977).
482 Santoru, M. L. *et al.* Cross sectional evaluation of the gut-microbiome metabolome axis in an Italian cohort of IBD patients. *Sci Rep* **7**, 9523, doi:10.1038/s41598-017-10034-5 (2017).
483 Antharam, V. C. *et al.* An Integrated Metabolomic and Microbiome Analysis Identified Specific Gut Microbiota Associated with Fecal Cholesterol and

Coprostanol in Clostridium difficile Infection. *PLoS One* **11**, e0148824, doi:10.1371/journal.pone.0148824 (2016).

484 Midtvedt, T. *et al.* Increase of faecal tryptic activity relates to changes in the intestinal microbiome: analysis of Crohn's disease with a multidisciplinary platform. *PLoS One* **8**, e66074, doi:10.1371/journal.pone.0066074 (2013).

485 Li, L., Batt, S. M., Wannemuehler, M., Dispirito, A. & Beitz, D. C. Effect of feeding of a cholesterol-reducing bacterium, Eubacterium coprostanoligenes, to germ-free mice. *Lab Anim Sci* **48**, 253-255 (1998).

486 Li, L., Buhman, K. K., Hartman, P. A. & Beitz, D. C. Hypocholesterolemic effect of Eubacterium coprostanoligenes ATCC 51222 in rabbits. *Lett Appl Microbiol* **20**, 137-140, doi:10.1111/j.1472-765x.1995.tb00410.x (1995).

487 Sekimoto, H., Shimada, O., Makanishi, M., Nakano, T. & Katayama, O. Interrelationship between serum and fecal sterols. *Jpn J Med* **22**, 14-20, doi:10.2169/internalmedicine1962.22.14 (1983).

488 Deisenhammer, E. A. *et al.* No evidence for an association between serum cholesterol and the course of depression and suicidality. *Psychiatry Res* **121**, 253-261, doi:10.1016/j.psychres.2003.09.007 (2004).

489 Kim, Y. K. & Myint, A. M. Clinical application of low serum cholesterol as an indicator for suicide risk in major depression. *J Affect Disord* **81**, 161-166, doi:10.1016/S0165-0327(03)00166-6 (2004).

490 Sansone, R. A. Cholesterol quandaries: relationship to depression and the suicidal experience. *Psychiatry (Edgmont)* **5**, 22-34 (2008).

491 Carlier, J. P., Bedora-Faure, M., K'Ouas, G., Alauzet, C. & Mory, F. Proposal to unify Clostridium orbiscindens Winter et al. 1991 and Eubacterium plautii (Seguin 1928) Hofstad and Aasjord 1982, with description of Flavonifractor plautii gen. nov., comb. nov., and reassignment of Bacteroides capillosus to Pseudoflavonifractor capillosus gen. nov., comb. nov. *Int J Syst Evol Microbiol* **60**, 585-590, doi:10.1099/ijs.0.016725-0 (2010).

492 Sun, Y. C., Q.; Lin, P.; Xu, R.; He, D.; Ji, W.; Bian, Y.; Shen, Y.; Li, Q.; Liu, C.; Dong, K.; Tang, Y.; Pei, Z.; Yang, L.; Lu, H.; Guo, X.; Xiao, L. Characteristics of Gut Microbiota in Patients with Rheumatoid Arthritis in Shanghai, China. *Frontiers in Cellular and Infection Microbiology* **9**, 1-11 (2019).

493 Mancabelli, L. *et al.* Identification of universal gut microbial biomarkers of common human intestinal diseases by meta-analysis. *FEMS Microbiol Ecol* **93**, doi:10.1093/femsec/fix153 (2017).

494 Lun, H. *et al.* Altered gut microbiota and microbial biomarkers associated with chronic kidney disease. *Microbiologyopen* **8**, e00678, doi:10.1002/mbo3.678 (2019).

495 Ai, D. *et al.* Using Decision Tree Aggregation with Random Forest Model to Identify Gut Microbes Associated with Colorectal Cancer. *Genes (Basel)* **10**, doi:10.3390/genes10020112 (2019).

496 Gupta, A. *et al.* Association of Flavonifractor plautii, a Flavonoid-Degrading Bacterium, with the Gut Microbiome of Colorectal Cancer Patients in India. *mSystems* **4**, doi:10.1128/mSystems.00438-19 (2019).

497 Armstrong, H. *et al.* Host immunoglobulin G selectively identifies pathobionts in pediatric inflammatory bowel diseases. *Microbiome* **7**, 1, doi:10.1186/s40168-018-0604-3 (2019).
498 Klaring, K. *et al.* Intestinimonas butyriciproducens gen. nov., sp. nov., a butyrate-producing bacterium from the mouse intestine. *Int J Syst Evol Microbiol* **63**, 4606-4612, doi:10.1099/ijs.0.051441-0 (2013).
499 Winter, J., Popoff, M. R., Grimont, P. & Bokkenheuser, V. D. Clostridium orbiscindens sp. nov., a human intestinal bacterium capable of cleaving the flavonoid C-ring. *Int J Syst Bacteriol* **41**, 355-357, doi:10.1099/00207713-41-3-355 (1991).
500 Schoefer, L., Mohan, R., Schwiertz, A., Braune, A. & Blaut, M. Anaerobic degradation of flavonoids by Clostridium orbiscindens. *Appl Environ Microbiol* **69**, 5849-5854, doi:10.1128/aem.69.10.5849-5854.2003 (2003).
501 Tunon, M. J., Garcia-Mediavilla, M. V., Sanchez-Campos, S. & Gonzalez-Gallego, J. Potential of flavonoids as anti-inflammatory agents: modulation of pro-inflammatory gene expression and signal transduction pathways. *Curr Drug Metab* **10**, 256-271, doi:10.2174/138920009787846369 (2009).
502 Loke, W. M. *et al.* Metabolic transformation has a profound effect on anti-inflammatory activity of flavonoids such as quercetin: lack of association between antioxidant and lipoxygenase inhibitory activity. *Biochem Pharmacol* **75**, 1045-1053, doi:10.1016/j.bcp.2007.11.002 (2008).
503 Peng, X. *et al.* In vitro catabolism of quercetin by human fecal bacteria and the antioxidant capacity of its catabolites. *Food Nutr Res* **58**, doi:10.3402/fnr.v58.23406 (2014).
504 Jaganath, I. B., Mullen, W., Lean, M. E., Edwards, C. A. & Crozier, A. In vitro catabolism of rutin by human fecal bacteria and the antioxidant capacity of its catabolites. *Free Radic Biol Med* **47**, 1180-1189, doi:10.1016/j.freeradbiomed.2009.07.031 (2009).
505 Barcenilla, A. *et al.* Phylogenetic relationships of butyrate-producing bacteria from the human gut. *Appl Environ Microbiol* **66**, 1654-1661, doi:10.1128/aem.66.4.1654-1661.2000 (2000).
506 Biagi, E. *et al.* Through ageing, and beyond: gut microbiota and inflammatory status in seniors and centenarians. *PLoS One* **5**, e10667, doi:10.1371/journal.pone.0010667 (2010).
507 Takeshita, K. *et al.* A Single Species of Clostridium Subcluster XIVa Decreased in Ulcerative Colitis Patients. *Inflamm Bowel Dis* **22**, 2802-2810, doi:10.1097/MIB.0000000000000972 (2016).
508 Zhang, Y. *et al.* Changes in gut microbiota and plasma inflammatory factors across the stages of colorectal tumorigenesis: a case-control study. *BMC Microbiol* **18**, 92, doi:10.1186/s12866-018-1232-6 (2018).
509 Figliuolo, V. R. *et al.* Sulfate-reducing bacteria stimulate gut immune responses and contribute to inflammation in experimental colitis. *Life Sci* **189**, 29-38, doi:10.1016/j.lfs.2017.09.014 (2017).
510 Millien, V., Rosen, D., Hou, J. & Shah, R. Proinflammatory Sulfur-Reducing Bacteria Are More Abundant in Colonic Biopsies of Patients with Microscopic

Colitis Compared to Healthy Controls. *Dig Dis Sci* **64**, 432-438, doi:10.1007/s10620-018-5313-z (2019).

511 Pilkonis, P. A. *et al.* Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS(R)): depression, anxiety, and anger. *Assessment* **18**, 263-283, doi:10.1177/1073191111411667 (2011).

512 First, M. W., JBW; Karg, RS; Spitzer, RL. *Structured Clinical Interview for DSM-5—Research Version (SCID-5 for DSM-5, Research Version; SCID-5-RV).* (American Psychiatric Association, 2015).

513 Posner, K. *et al.* The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry* **168**, 1266-1277, doi:10.1176/appi.ajp.2011.10111704 (2011).

514 Nock, M. K., Holmberg, E. B., Photos, V. I. & Michel, B. D. Self-Injurious Thoughts and Behaviors Interview: development, reliability, and validity in an adolescent sample. *Psychol Assess* **19**, 309-317, doi:10.1037/1040-3590.19.3.309 (2007).

515 Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**, 852-857, doi:10.1038/s41587-019-0209-9 (2019).

516 Douglas, G. M., Beiko, R. G. & Langille, M. G. I. Predicting the Functional Potential of the Microbiome from Marker Genes Using PICRUSt. *Methods Mol Biol* **1849**, 169-177, doi:10.1007/978-1-4939-8728-3_11 (2018).

517 McDonald, D. *et al.* The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome. *Gigascience* **1**, 7, doi:10.1186/2047-217X-1-7 (2012).

518 Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol* **12**, R60, doi:10.1186/gb-2011-12-6-r60 (2011).

519 Prevention, C. f. D. C. a. (ed U.S. Department of Health and Human Services) (Atlanta, GA, U.S.A., 2013).

520 Aslam, B. *et al.* Antibiotic resistance: a rundown of a global crisis. *Infect Drug Resist* **11**, 1645-1658, doi:10.2147/IDR.S173867 (2018).

521 Michael, C. A., Dominey-Howes, D. & Labbate, M. The antimicrobial resistance crisis: causes, consequences, and management. *Front Public Health* **2**, 145, doi:10.3389/fpubh.2014.00145 (2014).

522 Ventola, C. L. The antibiotic resistance crisis: part 1: causes and threats. *P T* **40**, 277-283 (2015).

523 Paterson, D. L. The role of antimicrobial management programs in optimizing antibiotic prescribing within hospitals. *Clin Infect Dis* **42 Suppl 2**, S90-95, doi:10.1086/499407 (2006).

524 Struelens, M. J. The epidemiology of antimicrobial resistance in hospital acquired infections: problems and possible solutions. *BMJ* **317**, 652-654, doi:10.1136/bmj.317.7159.652 (1998).

525 Mulvey, M. R. & Simor, A. E. Antimicrobial resistance in hospitals: how concerned should we be? *CMAJ* **180**, 408-415, doi:10.1503/cmaj.080239 (2009).

526 Cookson, B. Clinical significance of emergence of bacterial antimicrobial resistance in the hospital environment. *J Appl Microbiol* **99**, 989-996, doi:10.1111/j.1365-2672.2005.02693.x (2005).
527 Cassone, M. & Mody, L. Colonization with Multi-Drug Resistant Organisms in Nursing Homes: Scope, Importance, and Management. *Curr Geriatr Rep* **4**, 87-95, doi:10.1007/s13670-015-0120-2 (2015).
528 Aliyu, S., Smaldone, A. & Larson, E. Prevalence of multidrug-resistant gram-negative bacteria among nursing home residents: A systematic review and meta-analysis. *Am J Infect Control* **45**, 512-518, doi:10.1016/j.ajic.2017.01.022 (2017).
529 O'Fallon, E., Gautam, S. & D'Agata, E. M. Colonization with multidrug-resistant gram-negative bacteria: prolonged duration and frequent cocolonization. *Clin Infect Dis* **48**, 1375-1381, doi:10.1086/598194 (2009).
530 O'Fallon, E., Kandel, R., Schreiber, R. & D'Agata, E. M. Acquisition of multidrug-resistant gram-negative bacteria: incidence and risk factors within a long-term care population. *Infect Control Hosp Epidemiol* **31**, 1148-1153, doi:10.1086/656590 (2010).
531 Pop-Vicas, A., Mitchell, S. L., Kandel, R., Schreiber, R. & D'Agata, E. M. Multidrug-resistant gram-negative bacteria in a long-term care facility: prevalence and risk factors. *J Am Geriatr Soc* **56**, 1276-1280, doi:10.1111/j.1532-5415.2008.01787.x (2008).
532 Trick, W. E. *et al.* Colonization of skilled-care facility residents with antimicrobial-resistant pathogens. *J Am Geriatr Soc* **49**, 270-276, doi:10.1046/j.1532-5415.2001.4930270.x (2001).
533 van den Dool, C., Haenen, A., Leenstra, T. & Wallinga, J. The Role of Nursing Homes in the Spread of Antimicrobial Resistance Over the Healthcare Network. *Infect Control Hosp Epidemiol* **37**, 761-767, doi:10.1017/ice.2016.59 (2016).
534 Trick, W. E. *et al.* Regional dissemination of vancomycin-resistant enterococci resulting from interfacility transfer of colonized patients. *J Infect Dis* **180**, 391-396, doi:10.1086/314898 (1999).
535 Morrill, H. J., Caffrey, A. R., Jump, R. L., Dosa, D. & LaPlante, K. L. Antimicrobial Stewardship in Long-Term Care Facilities: A Call to Action. *J Am Med Dir Assoc* **17**, 183 e181-116, doi:10.1016/j.jamda.2015.11.013 (2016).
536 Bonomo, R. A. Multiple antibiotic-resistant bacteria in long-term-care facilities: An emerging problem in the practice of infectious diseases. *Clin Infect Dis* **31**, 1414-1422, doi:10.1086/317489 (2000).
537 Daneman, N. *et al.* Antibiotic use in long-term care facilities. *J Antimicrob Chemother* **66**, 2856-2863, doi:10.1093/jac/dkr395 (2011).
538 van Buul L.W., v. d. S. J. T., Veenhuizen R.B., Achterberg W.P., Schellevis F.G., Essink R.T.G.M., van Benthem B.H.B., Natsch S., Hertogh C.M.P.M. Antibiotic Use and Resistance in Long Term Care Facilities. *Journal of the American Medical Directors Association* **12**, 568.e561-568.e513 (2012).
539 McGowan, J. E., Jr. Antimicrobial resistance in hospital organisms and its relation to antibiotic use. *Rev Infect Dis* **5**, 1033-1048, doi:10.1093/clinids/5.6.1033 (1983).
540 Nicolle, L. E., Bentley, D. W., Garibaldi, R., Neuhaus, E. G. & Smith, P. W. Antimicrobial use in long-term-care facilities. SHEA Long-Term-Care

Committee. *Infect Control Hosp Epidemiol* **21**, 537-545, doi:10.1086/501798 (2000).

541 Jones, S. R., Parker, D. F., Liebow, E. S., Kimbrough, R. C., 3rd & Frear, R. S. Appropriateness of antibiotic therapy in long-term care facilities. *Am J Med* **83**, 499-502, doi:10.1016/0002-9343(87)90761-3 (1987).

542 Dyar, O. J., Pagani, L. & Pulcini, C. Strategies and challenges of antimicrobial stewardship in long-term care facilities. *Clin Microbiol Infect* **21**, 10-19, doi:10.1016/j.cmi.2014.09.005 (2015).

543 Peron, E. P., Hirsch, A. A., Jury, L. A., Jump, R. L. & Donskey, C. J. Another setting for stewardship: high rate of unnecessary antimicrobial use in a veterans affairs long-term care facility. *J Am Geriatr Soc* **61**, 289-290, doi:10.1111/jgs.12099 (2013).

544 Loeb, M. Antibiotic use in long-term-care facilities: many unanswered questions. *Infect Control Hosp Epidemiol* **21**, 680-683, doi:10.1086/501713 (2000).

545 D'Agata, E., Loeb, M. B. & Mitchell, S. L. Challenges in assessing nursing home residents with advanced dementia for suspected urinary tract infections. *J Am Geriatr Soc* **61**, 62-66, doi:10.1111/jgs.12070 (2013).

546 Mitchell, S. L. *et al.* Infection management and multidrug-resistant organisms in nursing home residents with advanced dementia. *JAMA Intern Med* **174**, 1660-1667, doi:10.1001/jamainternmed.2014.3918 (2014).

547 D'Agata, E. & Mitchell, S. L. Patterns of antimicrobial use among nursing home residents with advanced dementia. *Arch Intern Med* **168**, 357-362, doi:10.1001/archinternmed.2007.104 (2008).

548 Snyder, G. M., O'Fallon, E. & D'Agata, E. M. Co-colonization with multiple different species of multidrug-resistant gram-negative bacteria. *Am J Infect Control* **39**, 506-510, doi:10.1016/j.ajic.2010.09.012 (2011).

549 Mitchell, S. L., Shaffer, M. L., Kiely, D. K., Givens, J. L. & D'Agata, E. The study of pathogen resistance and antimicrobial use in dementia: study design and methodology. *Arch Gerontol Geriatr* **56**, 16-22, doi:10.1016/j.archger.2012.08.001 (2013).

550 D'Agata, E. M., Habtemariam, D. & Mitchell, S. Multidrug-Resistant Gram-Negative Bacteria: Inter- and Intradissemination Among Nursing Homes of Residents With Advanced Dementia. *Infect Control Hosp Epidemiol* **36**, 930-935, doi:10.1017/ice.2015.97 (2015).

551 O'Toole P.W., C. M. J. Gut microbiota: Changes throughout the lifespan from infancy to elderly. *International Dairy Journal* **20**, 281-291 (2010).

552 Rowan-Nash, A. D., Korry, B. J., Mylonakis, E. & Belenky, P. Cross-Domain and Viral Interactions in the Microbiome. *Microbiol Mol Biol Rev* **83**, doi:10.1128/MMBR.00044-18 (2019).

553 Nagpal, R. *et al.* Gut microbiome and aging: Physiological and mechanistic insights. *Nutr Healthy Aging* **4**, 267-285, doi:10.3233/NHA-170030 (2018).

554 Wang, F. *et al.* Gut Microbiota Community and Its Assembly Associated with Age and Diet in Chinese Centenarians. *J Microbiol Biotechnol* **25**, 1195-1204, doi:10.4014/jmb.1410.10014 (2015).

555 Claesson, M. J. *et al.* Composition, variability, and temporal stability of the intestinal microbiota of the elderly. *Proc Natl Acad Sci U S A* **108 Suppl 1**, 4586-4591, doi:10.1073/pnas.1000097107 (2011).

556 Salazar, N., Valdes-Varela, L., Gonzalez, S., Gueimonde, M. & de Los Reyes-Gavilan, C. G. Nutrition and the gut microbiome in the elderly. *Gut Microbes* **8**, 82-97, doi:10.1080/19490976.2016.1256525 (2017).

557 Biagi, E. *et al.* Gut Microbiota and Extreme Longevity. *Curr Biol* **26**, 1480-1485, doi:10.1016/j.cub.2016.04.016 (2016).

558 Alfa, M. J. *et al.* A randomized trial to determine the impact of a digestion resistant starch composition on the gut microbiome in older and mid-age adults. *Clin Nutr* **37**, 797-807, doi:10.1016/j.clnu.2017.03.025 (2018).

559 Park, S. H. *et al.* Comparative analysis of gut microbiota in elderly people of urbanized towns and longevity villages. *BMC Microbiol* **15**, 49, doi:10.1186/s12866-015-0386-8 (2015).

560 Rampelli, S. *et al.* Functional metagenomic profiling of intestinal microbiome in extreme ageing. *Aging (Albany NY)* **5**, 902-912, doi:10.18632/aging.100623 (2013).

561 Santoro, A. *et al.* Gut microbiota changes in the extreme decades of human life: a focus on centenarians. *Cell Mol Life Sci* **75**, 129-148, doi:10.1007/s00018-017-2674-y (2018).

562 Claesson, M. J. *et al.* Gut microbiota composition correlates with diet and health in the elderly. *Nature* **488**, 178-184, doi:10.1038/nature11319 (2012).

563 Ticinesi, A. *et al.* Gut microbiota composition is associated with polypharmacy in elderly hospitalized patients. *Sci Rep* **7**, 11102, doi:10.1038/s41598-017-10734-y (2017).

564 van Tongeren, S. P., Slaets, J. P., Harmsen, H. J. & Welling, G. W. Fecal microbiota composition and frailty. *Appl Environ Microbiol* **71**, 6438-6442, doi:10.1128/AEM.71.10.6438-6442.2005 (2005).

565 Fish, D. N. & Chow, A. T. The clinical pharmacokinetics of levofloxacin. *Clin Pharmacokinet* **32**, 101-119, doi:10.2165/00003088-199732020-00002 (1997).

566 Chien, S. C. *et al.* Pharmacokinetic profile of levofloxacin following once-daily 500-milligram oral or intravenous doses. *Antimicrob Agents Chemother* **41**, 2256-2260 (1997).

567 Anderson, V. R. & Perry, C. M. Levofloxacin : a review of its use as a high-dose, short-course treatment for bacterial infection. *Drugs* **68**, 535-565, doi:10.2165/00003495-200868040-00011 (2008).

568 Inagaki, Y., Nakaya, R., Chida, T. & Hashimoto, S. The effect of levofloxacin, an optically-active isomer of ofloxacin, on fecal microflora in human volunteers. *Jpn J Antibiot* **45**, 241-252 (1992).

569 Edlund, C. & Nord, C. E. Effect of quinolones on intestinal ecology. *Drugs* **58 Suppl 2**, 65-70, doi:10.2165/00003495-199958002-00013 (1999).

570 Edlund, C. & Nord, C. E. Effect on the human normal microflora of oral antibiotics for treatment of urinary tract infections. *J Antimicrob Chemother* **46 Suppl A**, 41-48 (2000).

571 Edlund, C., Sjostedt, S. & Nord, C. E. Comparative effects of levofloxacin and ofloxacin on the normal oral and intestinal microflora. *Scand J Infect Dis* **29**, 383-386, doi:10.3109/00365549709011835 (1997).

572 Bhalodi, A. A., van Engelen, T. S. R., Virk, H. S. & Wiersinga, W. J. Impact of antimicrobial therapy on the gut microbiome. *J Antimicrob Chemother* **74**, i6-i15, doi:10.1093/jac/dky530 (2019).

573 Ziegler, M. *et al.* Impact of Levofloxacin for the Prophylaxis of Bloodstream Infection on the Gut Microbiome in Patients With Hematologic Malignancy. *Open Forum Infect Dis* **6**, ofz252, doi:10.1093/ofid/ofz252 (2019).

574 Sullivan, A., Edlund, C. & Nord, C. E. Effect of antimicrobial agents on the ecological balance of human microflora. *Lancet Infect Dis* **1**, 101-114, doi:10.1016/S1473-3099(01)00066-4 (2001).

575 de Lastours, V., Chau, F., Roy, C., Larroque, B. & Fantin, B. Emergence of quinolone resistance in the microbiota of hospitalized patients treated or not with a fluoroquinolone. *J Antimicrob Chemother* **69**, 3393-3400, doi:10.1093/jac/dku283 (2014).

576 Lautenbach, E. *et al.* Longitudinal trends in fluoroquinolone resistance among Enterobacteriaceae isolates from inpatients and outpatients, 1989-2000: differences in the emergence and epidemiology of resistance across organisms. *Clin Infect Dis* **38**, 655-662, doi:10.1086/381549 (2004).

577 Dalhoff, A. Global fluoroquinolone resistance epidemiology and implictions for clinical use. *Interdiscip Perspect Infect Dis* **2012**, 976273, doi:10.1155/2012/976273 (2012).

578 Spellberg, B. & Doi, Y. The Rise of Fluoroquinolone-Resistant Escherichia coli in the Community: Scarier Than We Thought. *J Infect Dis* **212**, 1853-1855, doi:10.1093/infdis/jiv279 (2015).

579 Acar, J. F. & Goldstein, F. W. Trends in bacterial resistance to fluoroquinolones. *Clin Infect Dis* **24 Suppl 1**, S67-73, doi:10.1093/clinids/24.supplement_1.s67 (1997).

580 Nordmann, P. & Poirel, L. Emergence of plasmid-mediated resistance to quinolones in Enterobacteriaceae. *J Antimicrob Chemother* **56**, 463-469, doi:10.1093/jac/dki245 (2005).

581 Paterson, D. L. Resistance in gram-negative bacteria: enterobacteriaceae. *Am J Med* **119**, S20-28; discussion S62-70, doi:10.1016/j.amjmed.2006.03.013 (2006).

582 Ruppe, E., Woerther, P. L. & Barbier, F. Mechanisms of antimicrobial resistance in Gram-negative bacilli. *Ann Intensive Care* **5**, 61, doi:10.1186/s13613-015-0061-0 (2015).

583 Wood, D. E. & Salzberg, S. L. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* **15**, R46, doi:10.1186/gb-2014-15-3-r46 (2014).

584 Lu, J. B., F.P.; Thielen P.; Salzberg, S.L. Bracken: estimating species abundance in metagenomics data. *PeerJ Computer Science*, 3:e104 (2017).

585 Weiner, L. M. *et al.* Antimicrobial-Resistant Pathogens Associated With Healthcare-Associated Infections: Summary of Data Reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention,

2011-2014. *Infect Control Hosp Epidemiol* **37**, 1288-1301, doi:10.1017/ice.2016.174 (2016).

586 O'Donovan, D., Corcoran, G. D., Lucey, B. & Sleator, R. D. Campylobacter ureolyticus: a portrait of the pathogen. *Virulence* **5**, 498-506, doi:10.4161/viru.28776 (2014).

587 Salem, N., Salem, L., Saber, S., Ismail, G. & Bluth, M. H. Corynebacterium urealyticum: a comprehensive review of an understated organism. *Infect Drug Resist* **8**, 129-145, doi:10.2147/IDR.S74795 (2015).

588 Agudelo Higuita, N. I. & Huycke, M. M. in *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection* (eds M. S. Gilmore, D. B. Clewell, Y. Ike, & N. Shankar) (2014).

589 Fiore, E., Van Tyne, D. & Gilmore, M. S. Pathogenicity of Enterococci. *Microbiol Spectr* **7**, doi:10.1128/microbiolspec.GPP3-0053-2018 (2019).

590 Conway, T. & Cohen, P. S. Commensal and Pathogenic Escherichia coli Metabolism in the Gut. *Microbiol Spectr* **3**, doi:10.1128/microbiolspec.MBP-0006-2014 (2015).

591 Woodward, S. E., Krekhno, Z. & Finlay, B. B. Here, there, and everywhere: How pathogenic E. coli sense and respond to gastrointestinal biogeography. *Cell Microbiol*, e13107, doi:10.1111/cmi.13107 (2019).

592 Baqi, M. & Mazzulli, T. Oligella infections: Case report and review of the literature. *Can J Infect Dis* **7**, 377-379, doi:10.1155/1996/153512 (1996).

593 Graham, D. R., Band, J. D., Thornsberry, C., Hollis, D. G. & Weaver, R. E. Infections caused by Moraxella, Moraxella urethralis, Moraxella-like groups M-5 and M-6, and Kingella kingae in the United States, 1953-1980. *Rev Infect Dis* **12**, 423-431, doi:10.1093/clinids/12.3.423 (1990).

594 Pugliese, A., Pacris, B., Schoch, P. E. & Cunha, B. A. Oligella urethralis urosepsis. *Clin Infect Dis* **17**, 1069-1070, doi:10.1093/clinids/17.6.1069 (1993).

595 Wilmer, A. W., P.; Press, N.; Leung, V.; Romney, M.; Champagne, S. Oligella urethralis as a Cause of Urosepsis. *Clinical Microbiology Newsletter* **35**, 84-85 (2013).

596 Schaffer, J. N. & Pearson, M. M. Proteus mirabilis and Urinary Tract Infections. *Microbiol Spectr* **3**, doi:10.1128/microbiolspec.UTI-0017-2013 (2015).

597 Chen, C. Y. *et al.* Proteus mirabilis urinary tract infection and bacteremia: risk factors, clinical presentation, and outcomes. *J Microbiol Immunol Infect* **45**, 228-236, doi:10.1016/j.jmii.2011.11.007 (2012).

598 Wie, S. H. Clinical significance of Providencia bacteremia or bacteriuria. *Korean J Intern Med* **30**, 167-169, doi:10.3904/kjim.2015.30.2.167 (2015).

599 Kurmasheva, N., Vorobiev, V., Sharipova, M., Efremova, T. & Mardanova, A. The Potential Virulence Factors of Providencia stuartii: Motility, Adherence, and Invasion. *Biomed Res Int* **2018**, 3589135, doi:10.1155/2018/3589135 (2018).

600 Bassetti, M., Vena, A., Croxatto, A., Righi, E. & Guery, B. How to manage Pseudomonas aeruginosa infections. *Drugs Context* **7**, 212527, doi:10.7573/dic.212527 (2018).

601 Mittal, R., Aggarwal, S., Sharma, S., Chhibber, S. & Harjai, K. Urinary tract infections caused by Pseudomonas aeruginosa: a minireview. *J Infect Public Health* **2**, 101-111, doi:10.1016/j.jiph.2009.08.003 (2009).

602 Gordon, R. J. & Lowy, F. D. Pathogenesis of methicillin-resistant Staphylococcus aureus infection. *Clin Infect Dis* **46 Suppl 5**, S350-359, doi:10.1086/533591 (2008).
603 Lowy, F. D. Staphylococcus aureus infections. *N Engl J Med* **339**, 520-532, doi:10.1056/NEJM199808203390806 (1998).
604 Naimi, T. S. *et al.* Comparison of community- and health care-associated methicillin-resistant Staphylococcus aureus infection. *JAMA* **290**, 2976-2984, doi:10.1001/jama.290.22.2976 (2003).
605 Becker, K., Heilmann, C. & Peters, G. Coagulase-negative staphylococci. *Clin Microbiol Rev* **27**, 870-926, doi:10.1128/CMR.00109-13 (2014).
606 Czekaj, T., Ciszewski, M. & Szewczyk, E. M. Staphylococcus haemolyticus - an emerging threat in the twilight of the antibiotics age. *Microbiology* **161**, 2061-2068, doi:10.1099/mic.0.000178 (2015).
607 Froggatt, J. W., Johnston, J. L., Galetto, D. W. & Archer, G. L. Antimicrobial resistance in nosocomial isolates of Staphylococcus haemolyticus. *Antimicrob Agents Chemother* **33**, 460-466, doi:10.1128/aac.33.4.460 (1989).
608 Franzosa, E. A. *et al.* Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods* **15**, 962-968, doi:10.1038/s41592-018-0176-y (2018).
609 Arango-Argoty, G. *et al.* DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Microbiome* **6**, 23, doi:10.1186/s40168-018-0401-z (2018).
610 Hiramatsu, K. Molecular evolution of MRSA. *Microbiol Immunol* **39**, 531-543, doi:10.1111/j.1348-0421.1995.tb02239.x (1995).
611 Hiramatsu, K., Cui, L., Kuroda, M. & Ito, T. The emergence and evolution of methicillin-resistant Staphylococcus aureus. *Trends Microbiol* **9**, 486-493, doi:10.1016/s0966-842x(01)02175-8 (2001).
612 Tsubakishita, S., Kuwahara-Arai, K., Sasaki, T. & Hiramatsu, K. Origin and molecular evolution of the determinant of methicillin resistance in staphylococci. *Antimicrob Agents Chemother* **54**, 4352-4359, doi:10.1128/AAC.00356-10 (2010).
613 Enright, M. C. *et al.* The evolutionary history of methicillin-resistant Staphylococcus aureus (MRSA). *Proc Natl Acad Sci U S A* **99**, 7687-7692, doi:10.1073/pnas.122108599 (2002).
614 Smith, A. M. & Klugman, K. P. Alterations in PBP 1A essential-for high-level penicillin resistance in Streptococcus pneumoniae. *Antimicrob Agents Chemother* **42**, 1329-1333 (1998).
615 Zapun, A., Contreras-Martel, C. & Vernet, T. Penicillin-binding proteins and beta-lactam resistance. *FEMS Microbiol Rev* **32**, 361-385, doi:10.1111/j.1574-6976.2007.00095.x (2008).
616 Brannigan, J. A., Tirodimos, I. A., Zhang, Q. Y., Dowson, C. G. & Spratt, B. G. Insertion of an extra amino acid is the main cause of the low affinity of penicillin-binding protein 2 in penicillin-resistant strains of Neisseria gonorrhoeae. *Mol Microbiol* **4**, 913-919, doi:10.1111/j.1365-2958.1990.tb00664.x (1990).
617 Dowson, C. G., Jephcott, A. E., Gough, K. R. & Spratt, B. G. Penicillin-binding protein 2 genes of non-beta-lactamase-producing, penicillin-resistant strains of

Neisseria gonorrhoeae. *Mol Microbiol* **3**, 35-41, doi:10.1111/j.1365-2958.1989.tb00101.x (1989).
618 Thulin, S., Olcen, P., Fredlund, H. & Unemo, M. Total variation in the penA gene of Neisseria meningitidis: correlation between susceptibility to beta-lactam antibiotics and penA gene heterogeneity. *Antimicrob Agents Chemother* **50**, 3317-3324, doi:10.1128/AAC.00353-06 (2006).
619 Beceiro, A. B., G. Class C Beta-Lactamases: An Increasing Problem Worldwide. *Reviews in Medical Microbiology* **15**, 141-152 (2004).
620 Garvey, M. I., Baylay, A. J., Wong, R. L. & Piddock, L. J. Overexpression of patA and patB, which encode ABC transporters, is associated with fluoroquinolone resistance in clinical isolates of Streptococcus pneumoniae. *Antimicrob Agents Chemother* **55**, 190-196, doi:10.1128/AAC.00672-10 (2011).
621 El Garch, F. *et al.* Fluoroquinolones induce the expression of patA and patB, which encode ABC efflux pumps in Streptococcus pneumoniae. *J Antimicrob Chemother* **65**, 2076-2082, doi:10.1093/jac/dkq287 (2010).
622 Hernando-Amado, S. *et al.* Multidrug efflux pumps as main players in intrinsic and acquired resistance to antimicrobials. *Drug Resist Updat* **28**, 13-27, doi:10.1016/j.drup.2016.06.007 (2016).
623 Sato, T. *et al.* Fluoroquinolone resistance mechanisms in an Escherichia coli isolate, HUE1, without quinolone resistance-determining region mutations. *Front Microbiol* **4**, 125, doi:10.3389/fmicb.2013.00125 (2013).
624 Li, X. Z., Plesiat, P. & Nikaido, H. The challenge of efflux-mediated antibiotic resistance in Gram-negative bacteria. *Clin Microbiol Rev* **28**, 337-418, doi:10.1128/CMR.00117-14 (2015).
625 Chow, J., Tang, H. & Mazmanian, S. K. Pathobionts of the gastrointestinal microbiota and inflammatory disease. *Curr Opin Immunol* **23**, 473-480, doi:10.1016/j.coi.2011.07.010 (2011).
626 Dzutsev, A. & Trinchieri, G. Proteus mirabilis: The Enemy Within. *Immunity* **42**, 602-604, doi:10.1016/j.immuni.2015.04.004 (2015).
627 Butto, L. F., Schaubeck, M. & Haller, D. Mechanisms of Microbe-Host Interaction in Crohn's Disease: Dysbiosis vs. Pathobiont Selection. *Front Immunol* **6**, 555, doi:10.3389/fimmu.2015.00555 (2015).
628 Mirsepasi-Lauridsen, H. C. V., B.A.; Krogfelt, K.A.; Petersen, A.M. Escherichia coli pathobionts associated with inflammatory bowel disease. *Clinical Microbiology Reviews*, 32:e00060-00018 (2019).
629 Yang, H. M.-L., H.C.; Struve, C.; Allaire, J.M.; Bosman, E.; Sivignon, A.; Vogl, W.; Ma, C.; Reid, G.; Li, X.; Petersen, A.M.; Jacobson, K.; Gouin, S.; Barnich, N.; Yu, H.; Krogfelt, K.A.; Vallance, B.A. Ulcerative Colitis-associated E. coli Pathobionts Potentiate Colitis in Susceptible Hosts. *The Journal of Immunology* **202**, 192.193 (2019).
630 Hamilton, A. L., Kamm, M. A., Ng, S. C. & Morrison, M. Proteus spp. as Putative Gastrointestinal Pathogens. *Clin Microbiol Rev* **31**, doi:10.1128/CMR.00085-17 (2018).
631 Archambaud, C., Derre-Bobillot, A., Lapaque, N., Rigottier-Gois, L. & Serror, P. Intestinal translocation of enterococci requires a threshold level of enterococcal

overgrowth in the lumen. *Sci Rep* **9**, 8926, doi:10.1038/s41598-019-45441-3 (2019).
632 Overturf, G. D., Wilkins, J. & Ressler, R. Emergence of resistance of Providencia stuartii to multiple antibiotics: speciation and biochemical characterization of Providencia. *J Infect Dis* **129**, 353-357, doi:10.1093/infdis/129.3.353 (1974).
633 McHale, P. J., Keane, C. T. & Dougan, G. Antibiotic resistance in Providencia stuartii isolated in hospitals. *J Clin Microbiol* **13**, 1099-1104 (1981).
634 Hawkey, P. M. Providencia stuartii: a review of a multiply antibiotic-resistant bacterium. *J Antimicrob Chemother* **13**, 209-226, doi:10.1093/jac/13.3.209 (1984).
635 Warren, J. W. Providencia stuartii: a common cause of antibiotic-resistant bacteriuria in patients with long-term indwelling catheters. *Rev Infect Dis* **8**, 61-67, doi:10.1093/clinids/8.1.61 (1986).
636 Oikonomou, O. *et al.* Providencia stuartii Isolates from Greece: Co-Carriage of Cephalosporin (blaSHV-5, blaVEB-1), Carbapenem (blaVIM-1), and Aminoglycoside (rmtB) Resistance Determinants by a Multidrug-Resistant Outbreak Clone. *Microb Drug Resist* **22**, 379-386, doi:10.1089/mdr.2015.0215 (2016).
637 Barros, E. M., Ceotto, H., Bastos, M. C., Dos Santos, K. R. & Giambiagi-Demarval, M. Staphylococcus haemolyticus as an important hospital pathogen and carrier of methicillin resistance genes. *J Clin Microbiol* **50**, 166-168, doi:10.1128/JCM.05563-11 (2012).
638 Maleki, A., Ghafourian, S., Taherikalani, M. & Soroush, S. Alarming and Threatening Signals from Health Centers About Multi Drug Resistance Staphylococcus haemolyticus. *Infect Disord Drug Targets* **19**, 118-127, doi:10.2174/1871526518666180911142806 (2019).
639 Costa, S. S., Viveiros, M., Amaral, L. & Couto, I. Multidrug Efflux Pumps in Staphylococcus aureus: an Update. *Open Microbiol J* **7**, 59-71, doi:10.2174/1874285801307010059 (2013).
640 Truong-Bolduc, Q. C., Dunman, P. M., Strahilevitz, J., Projan, S. J. & Hooper, D. C. MgrA is a multiple regulator of two new efflux pumps in Staphylococcus aureus. *J Bacteriol* **187**, 2395-2405, doi:10.1128/JB.187.7.2395-2405.2005 (2005).
641 Lina, G. *et al.* Distribution of genes encoding resistance to macrolides, lincosamides, and streptogramins among staphylococci. *Antimicrob Agents Chemother* **43**, 1062-1066 (1999).
642 Ross, J. I. *et al.* Inducible erythromycin resistance in staphylococci is encoded by a member of the ATP-binding transport super-gene family. *Mol Microbiol* **4**, 1207-1214, doi:10.1111/j.1365-2958.1990.tb00696.x (1990).
643 Singh, L., Cariappa, M. P. & Kaur, M. Klebsiella oxytoca: An emerging pathogen? *Med J Armed Forces India* **72**, S59-S61, doi:10.1016/j.mjafi.2016.05.002 (2016).
644 Moradigaravand, D., Martin, V., Peacock, S. J. & Parkhill, J. Population structure of multidrug resistant Klebsiella oxytoca within hospitals across the UK and Ireland identifies sharing of virulence and resistance genes with K. pneumoniae. *Genome Biol Evol*, doi:10.1093/gbe/evx019 (2017).

645 Liu, H., Zhu, J., Hu, Q. & Rao, X. Morganella morganii, a non-negligent opportunistic pathogen. *Int J Infect Dis* **50**, 10-17, doi:10.1016/j.ijid.2016.07.006 (2016).
646 Noel, G. J. A Review of Levofloxacin for the Treatment of Bacterial Infections. *Clinical Medicine Insights: Therapeutics* **1** (2009).
647 Bush, L. M., Chaparro-Rojas, F., Okeh, V. & Etienne, J. Cumulative clinical experience from over a decade of use of levofloxacin in urinary tract infections: critical appraisal and role in therapy. *Infect Drug Resist* **4**, 177-189, doi:10.2147/IDR.S15610 (2011).
648 Croom, K. F. & Goa, K. L. Levofloxacin: a review of its use in the treatment of bacterial infections in the United States. *Drugs* **63**, 2769-2802, doi:10.2165/00003495-200363240-00008 (2003).
649 Deshpande, A., Pant, C., Jain, A., Fraser, T. G. & Rolston, D. D. Do fluoroquinolones predispose patients to Clostridium difficile associated disease? A review of the evidence. *Curr Med Res Opin* **24**, 329-333, doi:10.1185/030079908x253735 (2008).
650 Mehta, R. S. *et al.* Stability of the human faecal microbiome in a cohort of adult men. *Nat Microbiol* **3**, 347-355, doi:10.1038/s41564-017-0096-0 (2018).
651 Liskiewicz, P. *et al.* Fecal Microbiota Analysis in Patients Going through a Depressive Episode during Treatment in a Psychiatric Hospital Setting. *J Clin Med* **8**, doi:10.3390/jcm8020164 (2019).
652 Davies, E. V. *et al.* Temperate phages both mediate and drive adaptive evolution in pathogen biofilms. *Proc Natl Acad Sci U S A* **113**, 8266-8271, doi:10.1073/pnas.1520056113 (2016).
653 van Schaik, W. The human gut resistome. *Philos Trans R Soc Lond B Biol Sci* **370**, 20140087, doi:10.1098/rstb.2014.0087 (2015).
654 Francino, M. P. Antibiotics and the Human Gut Microbiome: Dysbioses and Accumulation of Resistances. *Front Microbiol* **6**, 1543, doi:10.3389/fmicb.2015.01543 (2015).
655 Penders, J., Stobberingh, E. E., Savelkoul, P. H. & Wolffs, P. F. The human microbiome as a reservoir of antimicrobial resistance. *Front Microbiol* **4**, 87, doi:10.3389/fmicb.2013.00087 (2013).
656 Hu, Y. *et al.* Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. *Nat Commun* **4**, 2151, doi:10.1038/ncomms3151 (2013).
657 Jackson, R. W., Vinatzer, B., Arnold, D. L., Dorus, S. & Murillo, J. The influence of the accessory genome on bacterial pathogen evolution. *Mob Genet Elements* **1**, 55-65, doi:10.4161/mge.1.1.16432 (2011).
658 Scott, K. P. The role of conjugative transposons in spreading antibiotic resistance between bacteria that inhabit the gastrointestinal tract. *Cell Mol Life Sci* **59**, 2071-2082 (2002).
659 Salyers, A. A., Gupta, A. & Wang, Y. Human intestinal bacteria as reservoirs for antibiotic resistance genes. *Trends Microbiol* **12**, 412-416, doi:10.1016/j.tim.2004.07.004 (2004).

REFERENCES

660 Kazimierczak, K. A. & Scott, K. P. Antibiotics and resistance genes: influencing the microbial ecosystem in the gut. *Adv Appl Microbiol* **62**, 269-292, doi:10.1016/S0065-2164(07)62009-7 (2007).

661 Martinez, J. L., Coque, T. M. & Baquero, F. What is a resistance gene? Ranking risk in resistomes. *Nat Rev Microbiol* **13**, 116-123, doi:10.1038/nrmicro3399 (2015).

662 Budding, A. E. *et al.* Rectal swabs for analysis of the intestinal microbiota. *PLoS One* **9**, e101344, doi:10.1371/journal.pone.0101344 (2014).

663 Bassis, C. M. *et al.* Comparison of stool versus rectal swab samples and storage conditions on bacterial community profiles. *BMC Microbiol* **17**, 78, doi:10.1186/s12866-017-0983-9 (2017).

664 Biehl, L. M. *et al.* Usability of rectal swabs for microbiome sampling in a cohort study of hematological and oncological patients. *PLoS One* **14**, e0215428, doi:10.1371/journal.pone.0215428 (2019).

665 Fair, K. *et al.* Rectal Swabs from Critically Ill Patients Provide Discordant Representations of the Gut Microbiome Compared to Stool Samples. *mSphere* **4**, doi:10.1128/mSphere.00358-19 (2019).

666 Snyder, G. M. & D'Agata, E. M. Diagnostic accuracy of surveillance cultures to detect gastrointestinal colonization with multidrug-resistant gram-negative bacteria. *Am J Infect Control* **40**, 474-476, doi:10.1016/j.ajic.2011.06.011 (2012).

667 McIver, L. J. *et al.* bioBakery: a meta'omic analysis environment. *Bioinformatics* **34**, 1235-1237, doi:10.1093/bioinformatics/btx754 (2018).

668 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25, doi:10.1186/gb-2009-10-3-r25 (2009).

669 Wattam, A. R. *et al.* Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Res* **45**, D535-D542, doi:10.1093/nar/gkw1017 (2017).

670 Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. metaSPAdes: a new versatile metagenomic assembler. *Genome Res* **27**, 824-834, doi:10.1101/gr.213959.116 (2017).

671 Wilson, M. S. *et al.* Immunopathology of schistosomiasis. *Immunol Cell Biol* **85**, 148-154, doi:10.1038/sj.icb.7100014 (2007).

672 Mangiola, F., Nicoletti, A., Gasbarrini, A. & Ponziani, F. R. Gut microbiota and aging. *Eur Rev Med Pharmacol Sci* **22**, 7404-7413, doi:10.26355/eurrev_201811_16280 (2018).

673 Fransen, F. *et al.* Aged Gut Microbiota Contributes to Systemical Inflammaging after Transfer to Germ-Free Mice. *Front Immunol* **8**, 1385, doi:10.3389/fimmu.2017.01385 (2017).

674 Franceschi, C., Garagnani, P., Parini, P., Giuliani, C. & Santoro, A. Inflammaging: a new immune-metabolic viewpoint for age-related diseases. *Nat Rev Endocrinol* **14**, 576-590, doi:10.1038/s41574-018-0059-4 (2018).

675 Franceschi, C. & Campisi, J. Chronic inflammation (inflammaging) and its potential contribution to age-associated diseases. *J Gerontol A Biol Sci Med Sci* **69 Suppl 1**, S4-9, doi:10.1093/gerona/glu057 (2014).

676 Franceschi, C., Garagnani, P., Vitale, G., Capri, M. & Salvioli, S. Inflammaging and 'Garb-aging'. *Trends Endocrinol Metab* **28**, 199-212, doi:10.1016/j.tem.2016.09.005 (2017).
677 Vital, M., Karch, A. & Pieper, D. H. Colonic Butyrate-Producing Communities in Humans: an Overview Using Omics Data. *mSystems* **2**, doi:10.1128/mSystems.00130-17 (2017).
678 Garcia-Patino, M. G., Garcia-Contreras, R. & Licona-Limon, P. The Immune Response against Acinetobacter baumannii, an Emerging Pathogen in Nosocomial Infections. *Front Immunol* **8**, 441, doi:10.3389/fimmu.2017.00441 (2017).
679 Antunes, L. C., Visca, P. & Towner, K. J. Acinetobacter baumannii: evolution of a global pathogen. *Pathog Dis* **71**, 292-301, doi:10.1111/2049-632X.12125 (2014).
680 Frankard, J., Rodriguez-Villalobos, H., Struelens, M. J. & Jacobs, F. Haemophilus parainfluenzae: an underdiagnosed pathogen of biliary tract infections? *Eur J Clin Microbiol Infect Dis* **23**, 46-48, doi:10.1007/s10096-003-1050-z (2004).
681 Murphy, T. F. & Apicella, M. A. Nontypable Haemophilus influenzae: a review of clinical aspects, surface antigens, and the human immune response to infection. *Rev Infect Dis* **9**, 1-15, doi:10.1093/clinids/9.1.1 (1987).
682 Van Eldere, J., Slack, M. P., Ladhani, S. & Cripps, A. W. Non-typeable Haemophilus influenzae, an under-recognised pathogen. *Lancet Infect Dis* **14**, 1281-1292, doi:10.1016/S1473-3099(14)70734-0 (2014).
683 Ricanek, P. *et al.* Gut bacterial profile in patients newly diagnosed with treatment-naive Crohn's disease. *Clin Exp Gastroenterol* **5**, 173-186, doi:10.2147/CEG.S33858 (2012).
684 Chen, Y. J. *et al.* Parasutterella, in association with irritable bowel syndrome and intestinal chronic inflammation. *J Gastroenterol Hepatol* **33**, 1844-1852, doi:10.1111/jgh.14281 (2018).
685 Zhang, Q. *et al.* Accelerated dysbiosis of gut microbiota during aggravation of DSS-induced colitis by a butyrate-producing bacterium. *Sci Rep* **6**, 27572, doi:10.1038/srep27572 (2016).
686 Gomez, A. *et al.* Gut Microbiome of Coexisting BaAka Pygmies and Bantu Reflects Gradients of Traditional Subsistence Patterns. *Cell Rep* **14**, 2142-2153, doi:10.1016/j.celrep.2016.02.013 (2016).
687 Rehm, C. D., Penalvo, J. L., Afshin, A. & Mozaffarian, D. Dietary Intake Among US Adults, 1999-2012. *JAMA* **315**, 2542-2553, doi:10.1001/jama.2016.7491 (2016).
688 Dahl, W. J. & Stewart, M. L. Position of the Academy of Nutrition and Dietetics: Health Implications of Dietary Fiber. *J Acad Nutr Diet* **115**, 1861-1870, doi:10.1016/j.jand.2015.09.003 (2015).
689 Fu, C. L., Odegaard, J. I., Herbert, D. R. & Hsieh, M. H. A novel mouse model of Schistosoma haematobium egg-induced immunopathology. *PLoS Pathog* **8**, e1002605, doi:10.1371/journal.ppat.1002605 (2012).
690 Richardson, M. L. *et al.* A new mouse model for female genital schistosomiasis. *PLoS Negl Trop Dis* **8**, e2825, doi:10.1371/journal.pntd.0002825 (2014).
691 Kaiser, J. BIOMEDICINE. NIH opens precision medicine study to nation. *Science* **349**, 1433, doi:10.1126/science.349.6255.1433 (2015).

692 Collins, F. S. & Varmus, H. A new initiative on precision medicine. *N Engl J Med* **372**, 793-795, doi:10.1056/NEJMp1500523 (2015).
693 Cussotto, S. *et al.* Differential effects of psychotropic drugs on microbiome composition and gastrointestinal function. *Psychopharmacology (Berl)* **236**, 1671-1685, doi:10.1007/s00213-018-5006-5 (2019).
694 Cussotto, S., Clarke, G., Dinan, T. G. & Cryan, J. F. Psychotropics and the Microbiome: a Chamber of Secrets. *Psychopharmacology (Berl)* **236**, 1411-1432, doi:10.1007/s00213-019-5185-8 (2019).
695 Flowers, S. A., Evans, S. J., Ward, K. M., McInnis, M. G. & Ellingrod, V. L. Interaction Between Atypical Antipsychotics and the Gut Microbiome in a Bipolar Disease Cohort. *Pharmacotherapy* **37**, 261-267, doi:10.1002/phar.1890 (2017).
696 Cryan, J. F. & Mombereau, C. In search of a depressed mouse: utility of models for studying depression-related behavior in genetically modified mice. *Mol Psychiatry* **9**, 326-357, doi:10.1038/sj.mp.4001457 (2004).
697 Chourbaji, S. *et al.* Learned helplessness: validity and reliability of depressive-like states in mice. *Brain Res Brain Res Protoc* **16**, 70-78, doi:10.1016/j.brainresprot.2005.09.002 (2005).
698 Anisman, H. & Merali, Z. Rodent models of depression: learned helplessness induced in mice. *Curr Protoc Neurosci* **Chapter 8**, Unit 8 10C, doi:10.1002/0471142301.ns0810cs14 (2001).
699 Yan, H. C., Cao, X., Das, M., Zhu, X. H. & Gao, T. M. Behavioral animal models of depression. *Neurosci Bull* **26**, 327-337, doi:10.1007/s12264-010-0323-7 (2010).
700 Golden, S. A., Covington, H. E., 3rd, Berton, O. & Russo, S. J. A standardized protocol for repeated social defeat stress in mice. *Nat Protoc* **6**, 1183-1191, doi:10.1038/nprot.2011.361 (2011).
701 Iniguez, S. D. *et al.* Social defeat stress induces depression-like behavior and alters spine morphology in the hippocampus of adolescent male C57BL/6 mice. *Neurobiol Stress* **5**, 54-64, doi:10.1016/j.ynstr.2016.07.001 (2016).
702 Mineur, Y. S., Belzung, C. & Crusio, W. E. Effects of unpredictable chronic mild stress on anxiety and depression-like behavior in mice. *Behav Brain Res* **175**, 43-50, doi:10.1016/j.bbr.2006.07.029 (2006).
703 Farooq, R. K. *et al.* Is unpredictable chronic mild stress (UCMS) a reliable model to study depression-induced neuroinflammation? *Behav Brain Res* **231**, 130-137, doi:10.1016/j.bbr.2012.03.020 (2012).