Gut and Tissue Microbiome Biogeography and Its Response to Environmental Perturbation

Gut and tissue microbiome biogeography and its response to environmental perturbation

A dissertation presented by

Koji Yasuda

To

The Biological Sciences in Public Health

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biological Sciences in Public Health

Harvard University

Cambridge, Massachusetts

September 2017

Dissertation Advisor: Dr. Curtis Huttenhower Koji Yasuda

Gut and tissue microbiome biogeography and its response to environmental perturbation

ABSTRACT

The symbiotic relationship between the host and its microbiome, which is composed of trillions of bacteria, archaea, viruses, and fungi, is essential for the host to maintain health. In the past two decades, the intestinal microbiota has become one of the most intensely studied microbial ecosystems on the planet; however, there are few studies on much of the microbiota's distribution, the factors shaping its composition both within and outside of the gastrointestinal (GI) tract, and the microbial communities' response to environmental factors, such as fluoride.

To understand the extent of intra-intestinal microbiota composition, its representation in the stool and factors dictating site-specificity of microbial taxa within the gut, we systematically collected stool and paired lumenal and mucosal intestinal samples from ten sites distal to the jejunum from the model organism *Macaca mulatta* (rhesus macaque) and assayed the samples with 16S rRNA amplicon sequencing. We found that stool composition was highly correlated with the microbial composition at the colonic lumen and mucosa, as well as enrichment of oxygen-tolerant bacteria in the mucosa, suggesting that stool is a good representation of distal gut lumenal bacterial communities and that oxygen may be a strong factor in shaping the gut microbial composition.

We then tested the hypothesis that environmental factors, such as fluoride, may affect the oral and gut microbial communities, as fluoride is widely prevalent in drinking water and dental products and may have unexpected effects on health. We modeled human fluoride exposure in mice by administering fluoride daily over a 12-week period. We then assayed oral and stool samples for 16S amplicons and

performed shotgun metagenomic sequencing to assess the effect of fluoride on oral and gut microbiome composition and function. We found that fluoride depletes bacterial taxa belonging to acidogenic bacterial genera (such as *Parabacteroides*, *Bacteroides*, and *Bilophila*) in the oral community. However, fluoride treatment did not induce a significant shift in the composition of the gut microbial community in our mouse model. Although the consequences of fluoride-induced shifts in the oral microbial community on health need further study, fluoride may not affect an established gut microbiome – at least not at the levels added to drinking water and dental products.

Finally, we developed a method to distinguish microbial sequencing reads from those that are introduced during sample processing as contaminants, and tested the hypothesis that microbial DNA may be detectable in low-biomass tissue samples such as intra-abdominal adipose tissues as a result of gut bacterial translocation. For this study, we collected intra-abdominal (mesenteric and omental) and peripheral (subcutaneous) adipose tissues and lymph nodes, along with paired intestinal contents (small and large intestinal contents) from mice, rhesus macaques, and humans and assayed these materials with 16S rRNA gene sequencing. By taking into account the abundance, prevalence and host-uniqueness (for each taxon) data as an input, we were able to distinguish microbial sequencing reads from those that are likely of contamination. Our data show that the majority of bacterial reads identified in human adipose tissues to be contaminants. There were, however, some bacterial sequencing reads identified in fat tissues in macaques and mice resembling those bacterial taxa from the gut. Although the macaque and mice data could possibly support the notion of gut bacterial translocation, the discrepancy with the human data, as well as having significant proportion of sequencing reads in adipose tissues made up of contaminant reads, further studies are needed to clarify whether the bacterial reads commonly found in fat and gut of macaques indeed occurred *in vivo* or post-mortem.

Acknowledgements

First and foremost, I would like to sincerely thank my dissertation advisor, Curtis Huttenhower, for his extraordinary support, astute criticism, and inexhaustible patience throughout my years in the lab. Curtis is a remarkable scientist with unrelenting passion for his work and it has been a source of inspiration for me. He is an outstanding role model as I embark on my journey in a career in science. I am also extremely grateful for the opportunity to be able to work so closely with Xochitl Morgan. Xochitl has admirable skills in both wet-lab and computational work, and I am lucky to have had the opportunity to learn from her.

I would next like to thank the members of my Dissertation Advisory Committee for their many helpful discussions and valuable guidance: Wendy Garrett, Chih-Hao Lee, Michael Gilmore, and Keith Mansfield. Wendy in particular is not only a brilliant scientist but has also been so generous with her time and has shown me continuous support in my science and my career throughout my PhD training. I also want to thank the defense committee members for taking the time out of busy schedule to review the dissertation and to be at the defense: Wendy Garrett, Eric Rubin, Sarah Fortune, and John Leung.

I am indebted to Dirk Gevers, who supported my interest in pursuing a PhD in microbiome research through our conversation over coffee while I was a veterinary pathology resident at the New England Primate Research Center. As a result of our conversations, the macaque-biogeography work that I started at the primate center has now been published in collaboration with the Broad and is a part of my thesis. I am very grateful for the friendship of all of the present and past members of the Huttenhower lab: Eric Franzosa, Timothy Tickle, Levi Waldron, Nicola Segata, Tiffany Hsu, Emma Schwager, Siyuan Ma, Boyu Ren, Galeb Abu-Ali, Casey DuLong, Ali Rahnavard, Jason Lioyd-Price, Randall Schwager, Yan Yan, Melanie Schrimer, Tommi Vatanen, Long Nguyen, Hera Vlamakis, Cesar Arze, Himel Mallick, Bahar Sayoldin, Lauren Mclver, George Weingart, Alex Kostic, Alexandra Sirota-

Madi, Afrah Shafquat, Regina Joice, Kevin Oh, and others. I am also grateful for the members of the Garrett Lab and Lee Lab for their input on experimental design, setting up assays, and friendship: Carey Ann Gallini, Leslie Wardwell, Michelle Rooks, Michael Howitt, Eunyoung Chung, Sonia Ballal, Lior Lobel, Diogo da Fonseca Pereira, Sydney Lavoie, Lathrin Fenn, Caitlin Brennan and Monia Michaud and the Lee lab: Nelson Knudsen, and Ryan Alexander.

I also owe so much gratitude towards our adipose tissue microbiome collaborators: Allison Goldfine from Joslin Diabetes Center and Ashley Vernon from Brigham and Women's Hospital, who found time in their busy clinical schedules to collaborate on this project. Additionally, without Gemma Carter and Will Pettee, our research assistants, we would not have been able to carryout such a logistically difficult project and I am grateful to them for their enthusiasm and dedication to this work.

I would like to thank the Biological Sciences in Public Health and the Department of Biostatistics at Harvard School of Public Health for their extraordinary support behind the scenes. I especially want to thank Nicole Levesque and Alice Winkle for administrative and grant writing support. I could not have accomplished this work without their support. I am also extremely grateful to Marianne Wessling-Resnick, Eric Rubin, Wendy Garrett, Phyllis Kanki and Max Essex for taking the time with me before I was even accepted to the program to encourage me to pursue my PhD at Harvard.

Finally, I would like to thank my wife, Jess. Her support, encouragement, patience and unwavering love were undeniably the foundation of my professional and personal life for our years together. I am so grateful to her for lifting me up when I was feeling defeated, and for celebrating my victories together however small they might have been. I also would like to thank my parents, Sanae and Hidemi Yasuda, for their unending support and encouraging me to be as ambitious as I wanted and to always follow my dreams. I also want to thank my brother, Junya and his wife, Michiko, and Jess's family Lori, John, Sarah, and Quin for their continuous support.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AC	ascending colon
ATM	adipose tissue macrophages
CNS	central nervous system
DC	descending colon
FDR	false discovery rate
GI	gastrointestinal tract
HUMAnN	The HMP Unified Metabolic Analysis Network
Ile	ileum
Jej	jejunum
KEGG	Kyoto Encyclopedia of Genes and Genomes
KOs	KEGG orthologs
LDA	Linear Discriminate Analysis
LeFSe	Linear discriminant analysis Effect Size
LI	large intestine
LN	lymph nodes
LPS	lipopolysaccharide
MaAsLin	Multivariate Associations by Linear models
MES	mesenteric fat
MetaPhlAn2	Metagenomic Phylogenetic Analysis
NEPRC	New England Primate Research Center
NSTI	nearest sequenced taxon index
OM	omental fat
OTU	operational taxonomic unit
PCoA	principle coordinate analysis
PCR	polymerase chain reaction
PICRUSt	Phylogenetic Investigation of Communities by Reconstruction of Unobserved States
QIIME	Quantitative Insights into Microbial Ecology
SCFA	short chain fatty acid
SI	small intestine
SQ	subcutaneous fat
TC	transverse colon
V4	hypervariable region iv of 16S ribosomal RNA gene
WGS	whole genome sequencing
WT	wild-type
16S	16S rRNA gene

This page is intentionally left blank.

INTRODUCTION

Characterizing Intestinal and Tissue Microbiome in the Era of High-Throughput Genomic Sequencing

Overview

The symbiotic relationship between the host and its microbiome, which is composed of trillions of bacteria, archaea, viruses, fungi, and parasites, is essential for the host to maintain health [1-6]. The microbes provide otherwise inaccessible nutrients [7], prime the host's immune system [8-12] and prevent colonization by potential pathogens [13, 14]. An imbalance in this system, termed dysbiosis, has been indicated in diseases of the intestine such as inflammatory bowel disease [15-17], gastric cancer [18] and colorectal cancer [19-21], as well as diseases outside the intestine such as systemic autoimmune diseases [12, 22, 23], and even pathology of the central nervous system [24-26]. It is clear that understanding the role that the gut microbiome may play in health and disease is not only critical to a better understanding of ourselves, but also to developing potential breakthroughs in diagnostics and therapeutics.

In the past two decades, the intestinal microbiota has become one of the most intensely studied microbial ecosystems on the planet; however, much of its distribution and migration pattern within the gastrointestinal (GI) tract, as well as its relation to stool microbial composition, have remained under-studied. Our first area of study centered on the basic question of the distribution of mucosal, lumenal, and stool-associated microbiota and how these three compartments relate to each other along various points in the gastrointestinal tract.

In addition, recent studies have shown that various environmental factors affect the composition and function of the gut microbiome, including antibiotics [27] [28, 29], prebiotics

[30] [31], and diet [32-34]. In recent years, compounds such as triclosan [35, 36] and arsenic [37], which are included in various cosmetic and hygiene products, have also been shown to influence the microbiome. One other ubiquitously present element in our environment is fluoride, an abiotic trace element that has been widely added to drinking water and dental products since the 1940s. Direct health effects of fluoride have been studied in the contexts of dental, musculoskeletal, reproductive, and other organ systems, but its potential effects on the gut and oral microbiota have not been explored. Thus, our second area of study aimed to investigate the effects of fluoride on oral and gut microbial community structure and function.

Finally, we tested the hypothesis that microbial DNA may be detectable in intra-abdominal adipose tissues as a result of gut bacterial translocation. Notably, there are adipose tissue microbiota previously reported by others in the subcutaneous fat [38-40]. Given its proximity to the gut and the possible mechanism of translocation, the intra-abdominal adipose tissue may also support unique microbial communities that potentially originate from the gut. In addition, the possible presence of resident microbiota would provide a completely different mechanism of understanding the tissue inflammation associated with obesity and type 2 diabetes as well as other diseases associated with chronic low-grade inflammation. Our third study is therefore aimed to probe the existence of a visceral adipose tissue microbiome.

In this introduction, I will first review the known facts about the bacterial composition of the gut microbiota and their biogeographical distribution across the different segments of the GI tract in humans and animal models. I will then discuss environmental perturbations and their influence on host-associated microbiota. Finally, I will explore the current evidence exploring the presence of the tissue microbiota and difficulties in studying these tissues with the currently available techniques.

The Forces Shaping the Gut Microbiota

Host factors, such as gastrointestinal tract physiology and anatomy, and non-host factors such as microbe-microbe interactions, influence the composition of the intestinal microbiota. One of the most well described host-derived factors affecting the microbial mass and composition of the GI tract is the intestinal pH. For example, bacterial counts are relatively high in the oral cavity (10^{8-10} CFU per gram (g) of content), and they become fewer in number (10^{2-3} CFU/g) as well as less diverse in the acidic environment of the stomach (pH < 4.0) [41]. Pancreatic-derived bicarbonates drive the number of bacteria as well as the pH gradually up in the small intestine, where the bacterial counts and pH reach approximately 10^{2-4} CFU/g and 7.0, respectively [41]. The number of bacteria significantly increases in the colon, where the counts can reach as high as 10^{10-14} CFU/g and where the pH becomes neutral to slightly acidic (pH = 6.5) due to various acids produced by microbial fermentation [41]. Therefore, the intestinal pH is an important factor in affecting the number of bacteria along the GI tract.

In addition to pH, numerous other host-derived factors are increasingly understood to influence gut microbiota composition. These factors include fatty acids and bile acids [42, 43], host-derived enzymes such as lactase [44], oxygen content [45, 46] [47], signals from the intestinal immune cells [9, 10, 48] and surgical procedures such as gastric bypass surgery [49] [50, 51] [52]. Although it is perhaps underappreciated, mucin composition is a critical factor in determining the microbial composition of the gut, as certain bacterial species (e.g. *Akkermansia muciniphila*) preferentially hydrolyze polysaccharides found in mucin as an energy source [2, 53-58]. The complexities of the local environment are important determinants of the microbial community structure.

Microbe-microbe interaction is another force that shapes the compositional and functional profiles of the gut microbiome. Certain symbionts directly fend off neighboring bacteria by the secretion of antimicrobial factors that are only effective at killing phylogenetically distant bacterial taxa and are protective to those that are within the same species [59, 60]. Other bacteria use strategies such as using secreting protein siderophores for example to sequester Fe^{3+}, which is essential for DNA synthesis in most bacteria [61] and mammals. Other bacteria compete metabolically for the same nutrient sources [13, 14, 62] or promote host resistance by triggering immune-mediated protection by stimulating the production of host cellular IFNγ production [63]. The symbionts also possess unique features that are absent in pathogens to evade host immune-protection such as the modification of surface lipopolysaccharide (LPS) structures [64]. Thus, microbial interactions are critical in determining the compositional structure and function of the microbial community.

Certain unknown host- and/or microbial factors may allow specific taxa of bacteria to persist over time. Several studies have shown that different strains of bacteria can persist over 2 years in an individual [65], even after such harsh perturbations as antibiotic treatment [29, 66] or following fecal microbial transplantation to a different host [67]. There are clearly numerous factors from both the host and the microbiota that influence the microbial composition of the gut and that in turn impact the health and disease of both the host and the microbial residents within.

The Composition of the Mammalian Gut Microbiota

The "healthy" human distal gut is dominated by two major bacterial phyla, the Firmicutes and Bacteroidetes, with smaller contributions from Proteobacteria, and Actinobacteria, and rare representative of the Verrucomicrobia, Fusobacteria, Cyanobacteria and Euryarchaeota [1, 5]. At

the phylum level, the bacterial composition within humans shares similarities with other mammals such as mice and nonhuman primates; however, there are notable differences at the species level. For example, some species such as *Allobaculum*, *Oscillospira*, Rikeneraceae, and *Odoribacter* are at relatively high abundance in the distal gut in mice [17, 68, 69] whereas there is a high abundance of *Helicobacter*, *Treponema*, and Spirochaetes in the colonic mucosa of macaques; these species are unique to macaques [47, 70].

Unlike the gut, the oral cavity is predominantly colonized by taxa belonging to *Streptococcus, Veillonella, Granulicatella, Gemella, Actinomyces, Corynebacterium, Rothia, Fusobacterium, Porphyromonas, Prevotella, Capnocytophaga, Neisseria, Haemophilis, Treponema, Lactobacterium, Eikenella, Leptotrichia, Peptostreptococcus, Staphylococcus, Eubacteria, and Propionibacterium* [5, 71]. The number of bacterial species reaches approximately 500-700 operational taxonomic units (OTUs) in the oral cavity [5, 72-76]. Despite such diversity found in the oral site, only a small portion (<10 species) of the oral bacterial OTUs is detected in the stool when a paired oral-stool samples from the same individual are compared for their sequences similarity [71]. Oral bacteria are known to form multispecies community biofilms on tooth surfaces [77] [78], which protect these microbial communities from environmental factors [79]. The oral microbiota have been associated with diseases both within and beyond the oral cavity including the development of dental caries [80, 81] and periodontitis [72], inflammatory bowel disease [82-84], pancreatic cancer [85], atherosclerosis [86] and autoimmune disease such as celiac disease [87, 88].

Similarly to the oral site, the major phyla found in the stomach with sequencing-based studies consist of Actinobacteria, Bacteroides, Firmicutes, Proteobacteria, and Fusobacteria [89, 90]. However, the majority of these oral microbiota are not viable in the stomach [91, 92] but

rather represent remnant DNA from bacteria that colonize the oral cavity [74]. Certain bacterial taxa such as *Helicobacter pylori,* possess unique features such as cytoplasmic ureases that can convert urea into carbon dioxide and ammonia to locally neutralize the acidic condition of the stomach, thereby enabling survival in the hostile acidic environment of the stomach [93]. *H. pylori* is among the most intensely studied gastric bacteria in relation to disease due to its relatively recently discovered association with gastric ulcers [94] [95] [96] and gastric carcinoma [18, 97] [98]. Notably, the number of bacteria in the stomach is the lowest among the gastrointestinal tract [99]; the bacterial diversity measured by the Shannon index is the lowest in the stomach [100].

Given its difficult-to-access location, the "healthy" microbial composition of the small intestine has particularly been particularly under-studied. A few studies using endoscopy and colonoscopy [1] [100], both of which require cleansing that itself perturbs microbial composition, revealed distinct microbial communities characterized by higher abundance of Proteobacteria and Fusobacteria in the small intestine compared to the distal gut. Most likely, the most comprehensive study in terms of number of subjects (n = 17 subjects vs. n = 3 by Eckburg et al, vs. n = 4 by Stearn et al.) that profiled the ileal microbial composition was one by Hartman et al. showing that the normal community of the ileum is dominated by strict anaerobes such as Bacteroides and Clostridia [101]. The study also illustrated the dominance of facultative anaerobes such as *Lactobacillus* and *Enterobacteria* in the ileal effluent obtained from patients with ileostomies, where local concentrations of oxygen were hypothesized to be higher than typical given the presence of the ileostomy which interfaces with the external environment. This finding suggested that oxygen is a critical ecological determinant shaping the gut microbiota [101].

Comprehensive characterization of the "healthy" human microbiota has been appreciated in recent years by two independent consortia: the National Institute of Health (NIH) led Human Microbiome Project (HMP) [5] and the European led Metagenomic of the Human Intestinal Tract (MetaHIT) [102]. The HMP project comprehensively characterized the microbial community structures and functions at 15 male and 18 female body sites from 242 healthy adults in the United States at three different time points. The MetaHIT study focused solely on stool samples from 146 European individuals using metagenomic shotgun sequencing. The major findings from these studies were that the gut and sites within the oral cavity showed the greatest between-subject microbial beta-diversity but also the lowest between-visit variability, whereas the skin had lower between-subject diversity but higher between-visit variability. This suggests that the composition of the microbiota varies greatly across individuals, although it is temporally stable within a single person, and the functionality of the microbiota at the genus-level is highly conserved across individuals [5]. One of the subsequent MetaHIT studies reported that individuals, regardless of gender, geography or race, can be grouped into one of the three "enterotypes" characterized by variation in the levels of three dominant bacterial genera: *Bacteroides*, *Prevotella*, *Ruminococcus* [102]. However, recent studies suggest a gradient model, rather than this discrete "enterotype" model [103, 104].

Tissue Microbiome

The interface between the host and its environment clearly introduces many factors that affect the microbiome structure and function of the microbiome at various locations along the gut. It is less clear if there is a potential for bacteria to reside outside of locations of

environmental interface (such as the gut) and instead in internal tissues that were previously thought sterile. During bacterial infection of tissues and as a result of inflammation of the gut (e.g., colitis), viable bacteria are often cultured and their DNA fragments detected from the site of infection [105] and from associated lymph nodes [17, 20]. There are also bacterial taxa that are known to have the ability to survive within host cells, such as latent infection with uropathogenic *Escherichia coli* in the lower urinary tract [106] or *Mycobacterium tuberculosis* within macrophages [107, 108]. However, otherwise healthy, non-infected tissues have not been explored for the possible existence of microbiota.

Recent studies claiming the presence of unique tissue microbiota in tissues such as placental tissues [109] have been specifically been shown to be invalid. Among other studies claiming tissue microbiota at such sites as the lungs [110], liver [111], and the central nervous system [112], the only plausible results supporting the presence of resident tissue microbiota seem to be from a study of breast tissues [113]. In this study, Urbaniak et al. collected mammary tissues from 81 women with and without breast cancer and subjected them to 16S rRNA amplicon sequencing [113]. They found Proteobacteria phylum bacteria to be predominant in breast tissues. At the genus and family level, the most abundant taxa consisted of Enterobacteriaceae, *Bacillus*, *Acinetobacter, Pseudomonas, Staphylococcus, Propionibacterium*, and Comamonadaceae; all of these taxa are associated with either the gut or skin microbiomes [5]. Notably, this study additionally cultured eight species of bacteria: *Bacillus* sp., *Micrococcus luteus*, *Propionibacterium acnes*, *Propionibacterium granulosum*, *Staphylococcus* sp., *Staphylococcus saprophyticus*, *Streptococcus oralis*, and *Streptococcus agalactiae* from 43 out of the 81 subjects with amounts ranging from 75 to 2,000 CFU/gram of tissue [113]. The authors did not discuss the possible location in which these bacteria may be located within the breast

tissue. Breast tissue is anatomically unique in that it is surrounded by fascia and lined with epithelial cells making up lactation lobules (involved in lactation) that contain a direct link to the external body surface. Considering these features, the breast tissue may be one of the rare tissue sites that has some degree of bacterial colonization with microbiota possibly lining the epithelial cells of lobules, which directly communicate with the external environment rather than being completely excluded from the outside environment.

Considering Reagent Contamination in High-throughput Sequencing

There have been several studies reporting the presence of bacterial [114-120] fungal [121] and viral [122] contamination in extraction and sequencing reagents, which raises the question as to what degree microbiota derive from the tissue samples versus from contaminated reagents. These studies certainly indicate that the inclusion of negative blank controls is necessary to later computationally assess the degree of contamination contributed by reagents in DNA extraction and PCR amplification steps. Contamination in high-throughput sequencing studies is not unique to microbiome research, as studies have indicated various cross-contamination of genomes in the NCBI database, such as the *Neisseria gonorrhoeae* genome found in that of the domestic cow (*Bos Taurus*) [123] as well as human genome elements found in *Caenorhabditis elegans* (*C. elegans*), Xenopus, and Zebra fish [124]. Although low-biomass tissue samples in microbiome research are prone to contamination, raising questions about the very existence of a "tissue microbiome," the increasing number of studies identifying microbiota in tissues that were previously thought to be sterile makes the idea of such a tissue microbiome existing in the adipose compartment of the abdomen intriguing, given its proximity to the gut and

the possible mechanism of translocation. Our third area of study aimed to probe the existence of a visceral adipose tissue microbiome.

Chemical Perturbations on the microbiome

Both the composition, and enzymatic capacity of the gut microbiota are readily affected by various environmental factors, such as pesticides [125, 126], arsenic [37], triclosan [35, 36], and heavy-metals such as cadmium and lead in mice [127]. Regarding pesticides, glyphosate, an active component of the most widely used herbicide, has been shown to inhibit the growth of beneficial gut microbiota such as *Enterococcus facecalis* in both cattle and horse stool [125]. Whereas certain pathogenic gut bacteria of poultry such as *Salmonella enteritidia*, *Salmonella gallinarum* and *Closttridium perfringens* have been shown to be highly resistant to glyphosate [126]. Similarly, Breton et al. showed depletion of specific beneficial bacterial taxa such as Lacnospichaceae in cadmium/lead fed mice [127]. Fluoride is another abiotic trace element that has been widely added to drinking water and dental products; it has been shown to affect the growth of several bacterial taxa through various mono-culture-based studies (discussed below), but its potential effects on the gut and oral microbiota have not been explored.

Fluoride and Oral Bacteria

Fluoride is known to inhibit bacterial growth through inhibition of enzymes that are critical for bacterial metabolism such as enolase, which catalyzes the conversion of 2-phosphoglycerate to phosphoenolpyruvate (the last step of anaerobic glycolysis) and thus is critical for microbial energy harvesting and growth [128] [129]. The major mechanisms by which fluoride inhibits bacterial energy growth are direct binding of the fluorine ion to the active

sites of enolase [129] and ATPases [130] and disruption of the ion gradient across the bacterial cell membrane [128]. All of these mechanisms result in the reduction of adenosine triphosphate (ATP) synthesis [130]. Although a wide range of bacterial taxa are inhibited by this mechanism, the degree of resistance to fluoride's effects differs across taxa. For example, the enolases of *S. mutans and S. sanguis* are more susceptible (by 10-fold) to fluoride than those of *S. salivarius* and *Lactobacillus casei* in a monoculture system [130, 131].

Summary of Aims

This dissertation aims to contribute to the field's basic knowledge of the composition and factors influencing the distribution of microbiota within and beyond the GI tract and how environmental factor such as fluoride in drinking water and dental products influence oral and gut microbial composition and function. In Chapter 1, we comprehensively characterized the GI tract microbiota composition across five segments spanning both small and large intestine, including both mucosal and lumenal sites from each segment in rhesus macaques as a model for humans. In Chapter 2, we used 16S rRNA amplicon and metagenomic shotgun sequencing techniques to assess the effects of environmental factors such as fluoride on oral and gut microbial composition and function in mice. Finally, in Chapter 3, we characterized putative tissue microbiota composition in humans, macaques and mice.

This page intentionally left blank

CHAPTER 1

Biogeography of the Intestinal Mucosal and Lumenal Microbiome in the Rhesus Macaque

This Chapter is a reproduction of a published manuscript:
Yasuda, K., Oh, K., Ren, B., Tickle, T.L., Franzosa, E.A., Wachtman, L.M., Miller, A.D., Westmoreland, S.V., Mansfield, K.G., Vallender, E., Miller, G.M., Rowlett, J., Gevers, D., Huttenhower, C., Morgan, XC. Biogeography of the Intestinal Mucosal and Lumenal Microbiome in the Rhesus Macaque. *Cell Host and Microbe*. 2015;17(3):385- 91.

Author Contributions:
KY, KO, SVW, KGM, EV, GM, JR, DG, CH, and XCM designed the study. KY, SVW, KGM, DG processed samples and DNA sequenced. KY, KO, BR, TLT, EAF, CH, and XCM analyzed the data. KY, CH and XCM prepared the manuscript.

Abstract

The human gut microbiome is widely studied as stool, but the extent to which the stool microbiome reflects the composition of other intestinal sites is less well characterized. We investigated this relationship in the well-studied model organism *Macaca mulatta*, the rhesus macaque, by 16S sequencing stool and paired lumenal and mucosal samples from ten sites distal to the jejunum from 15 animals. Stool composition was highly correlated with the colonic lumen and mucosa (Spearman's r=0.98, 0.85), and moderately with the small intestine (r=0.53, 0.47). Facultative anaerobes (e.g. *Helicobacter*, *Treponema)* were mucosally enriched, while obligate anaerobes (e.g. *Firmicutes*) were lumenally enriched. The abundance of *Helicobacter*, *Faecalibacterium*, and *Lactobacillus* in stool was highly predictive of its abundance at most other sites in the gut. Our results precisely quantify the composition and biogeographic relationships between microbial communities in the macaque gut and support the use of stool for translational studies.

Introduction

Gut mucosal and lumenal microbial communities are distinct [1, 132, 133], and diseases such as colorectal cancer and inflammatory bowel disease induce site-specific epithelial inflammation at which the microbiota are disrupted relative to adjacent normal tissue [134-136]. Understanding the relationship between stool and the mucosal microbiome is thus of great interest, but large-scale human health-related studies typically focus on the stool microbiota due to technical limitations [5, 137-139].

Furthermore, human biopsy samples are near-universally collected after bowel preparation [140], which itself alters the mucosal community [141]; paired stool data is rarely available. Previous studies of human gut biogeography have included only samples from different individuals and / or timepoints [16, 133, 134, 142] or used a very small number of individuals [1, 132]. While the mucosa and lumenal

contents of mice are readily accessible for biogeographic studies, neither the pelleted, sparse nature of their colonic contents nor their native microbial composition are totally representative the human gut [143]. The captive rhesus macaque (*Macaca mulatta)*, which is widely used in biomedical research due to its genetic and physiological similarities to humans [70, 144-146], is an excellent model for detailed biogeographic study of the mucosal, lumenal, and stool microbiota. It further avoids confounding due to sample collection and manipulation methods (no colon preparation is required upon autopsy) or diet (synchronized meals).

In this study, we investigated i) the extent to which the stool microbiome reflects the composition of other intra-intestinal sites, ii) the biogeography of the composition of the rhesus macaque gut microbiome, and (iii) predictability of microbiota in the gut. Our results indicate that the stool microbiota community is a good proxy of the large intestinal (LI) lumen and mucosa and is surprisingly well-correlated with the small intestine (SI). The colonic mucosa was highly enriched in *Helicobacter*, which is flagellated and facultatively anaerobic. In contrast, obligate anaerobic Firmicutes were primarily localized to the intestinal lumen. This study thus provides the quantitative relationship between mucosal and lumenal microbial communities as assessed using stool.

Results

The macaque intestinal mucosa is dominated by non-pathogenic Pasteurellaceae and Helicobacteriaceae

Similarly to humans [1, 5, 137], the macaque intestine was colonized primarily by Bacteroidetes, Firmicutes, and Proteobacteria *(***Fig. 1-1A**). In contrast, the Actinobacteria and Verrucomicrobia were rare in macaques, and *Spirochaetes* and *Helicobacter* were much more abundant. To assess our data in the context of other human [5, 138] and macaque [70, 146] microbiome studies, we combined these datasets, calculated Bray-Curtis dissimilarity and weighted UniFrac distance, and performed principal

coordinate analysis (**Fig 1-S1D-E**; **Supplemental Methods**). Despite differences in sequencing technology, the three macaque studies were similar to one another, and more similar to the Malawian and Amerindian than to the US microbiomes.

We used both univariate [147] and multivariate analyses [133] to identify bacterial taxa significantly-enriched (FDR $q < 0.2$) in the mucosa or lumen; the multivariate analysis included location, sample type, weight, age, and primate center of origin as covariates (**Table 1-S1**). Relative to the mucosa, the stool and lumen were enriched for obligately anaerobic, short chain fatty acid-producing clades such as the Lachnospiraceae, Clostridiaceae, and Prevotellaceae [148]. In the mucosa, facultatively anaerobic clades were more abundant; these were mostly Proteobacteria, such as *Helicobacter* in the LI and *Pasteurella* in the SI (**Fig. 1-1A**). This likely reflects the higher host-derived oxygen content in the mucous layer compared to the lumen. Helicobacteraceae in particular was strongly associated with mucosa ($q < 10^{-21}$) and the ascending colon ($q=0.0011$). While *H. macacae* has been previously associated with chronic diarrhea and intestinal adenocarcinoma [149-151], our animals showed no evidence of tumorigenesis nor signs of excess inflammation upon routine histopathologic examination of the ileal, cecal, and colonic tissues.

All the animals in our study were housed at the New England Primate Research Center (NEPRC) for 2 years prior to sample collection, but 11 animals came from Oregon National Primate Center where they were housed outdoors. Research center was not associated with major systematic variation in microbial diversity, but was significantly associated with 23 OTUs (effect size -0.05 – 0.04; $q < 0.2$; **Table 1-S1; Fig. 1-S1A**). Most of these OTUs were Ruminococcaceae and Lachnospiraceae, which are primarily lumenally-enriched taxa. However, several mucosally-enriched taxa, including *Treponema*, *Desulfovibrio*, and *Corynebacterium*, were enriched in animals from one primate center, suggesting that their presence in the colonic mucosa may be highly influenced by early exposure.

The mucosal microbiota is most influenced by location, while the lumenal microbiota is most influenced by individual

The largest covariation within microbial community structure (as assessed by weighted UniFrac dissimilarity [152]) was explained by mucosal / lumenal sample origin (**Fig. 1-1B**). When mucosal and lumenal/stool samples were separated, the largest source of variation in mucosal samples corresponded to SI vs. LI sample origin (**Fig. 1-1C**), but no such pattern was observed for lumenal samples (**Fig. 1-1D**). As observed in previous human studies [1, 153], the stool microbiome showed high inter-individual variation, as did lumenal contents, which were themselves not substantially influenced by biogeographical location (**Fig. 1-S1B)**. The Bray-Curtis dissimilarity (based on species in common between sites) between stool and each of the other sites showed that stool was equally dissimilar to all mucosal sites regardless of anatomical proximity (**Fig. 1-S1C**); in contrast, lumenal dissimilarity increased with colonic distance. This suggests that despite the close anatomical proximity of distal mucosa and stool, lumenal flux of microbiota occurs more readily than transfer of microbiota between mucosa and lumen.

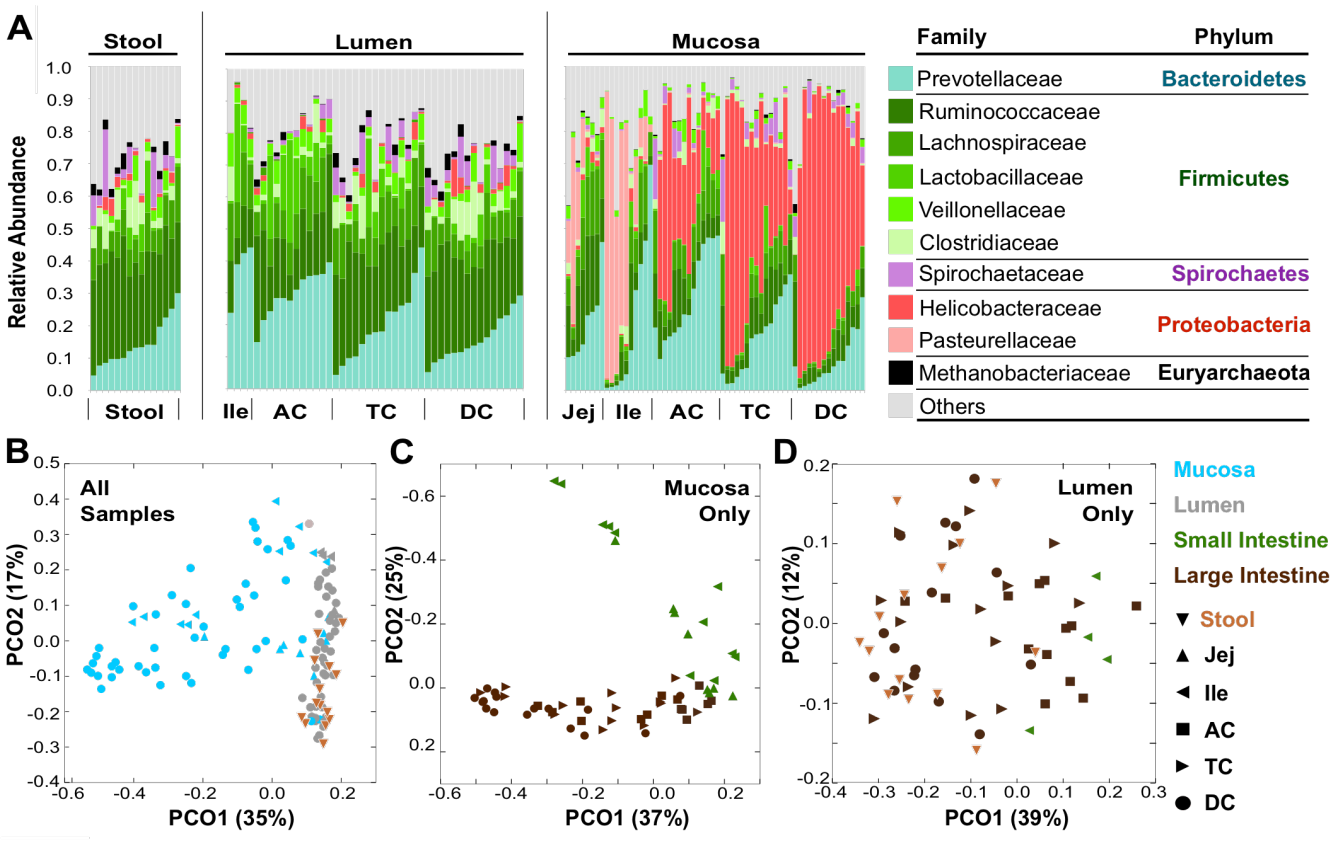


Figure 1-1. Biogeographic influences on macaque gut microbial composition A) Family-level relative abundance of intestinal microbiota in the stool (left), lumen (middle) and mucosa (right) of 15 healthy rhesus macaques. B) Principal coordinate analysis (PCoA) of all samples by weighted UniFrac distance. C) PCoA of mucosal-only samples. D) PCoA of lumen and stool-only samples. See also **Fig. 1-S1B**.

Stool microbial composition accurately reflects the colonic lumen and mucosa

We assessed the extent to which the mucosal and lumenal community of each individual was reflected in the stool by measuring the Spearman correlation between stool and the four major subdivisions of the distal gut (SI mucosa, SI lumen, LI mucosa, and LI lumen), thus accounting for both OTU rank order and the magnitude of relative abundances between the two sites being compared. Stool composition was highly correlated with the LI lumen (Spearman's $r=0.98$; $p<0.001$) and LI mucosa ($r=0.85$, $p<0.001$; **Fig. 1-2**). Stool composition was also surprisingly correlated with the SI mucosa ($r=0.465$, $p<0.001$) and lumen ($r=0.525$, $p<0.001$; **Fig. 1-2**). We examined these OTUs for a systematic taxonomic bias (**Fig. 1-S2**) and found that most mucosal OTUs that do not appear in stool are primarily Proteobacteria..

In the SI lumen and LI mucosa and lumen, over 97% of observed OTUs were also detected in stool, and stool-undetected OTUs had very low relative abundances ($<10^{-3}$) in the mucosa and lumen, and thus may have been detected with deeper sequencing of stool. In contrast, 10% of SI mucosal OTUs were stool-undetected despite relative abundance typically $>10^{-3}$; thus, increasing stool read depth may not necessarily improve the detection of these OTUs. Fusobacteria, β- and γ-proteobacteria are particularly likely to be stool-undetected (Fig 1-S2; Table 1-S2).

Nearly all (95%) OTUs detected within the LI mucosa lumen and in stool were detected in stool within two orders of magnitude ($10^{-1} – 10^{1}$) of their lumenal and mucosal relative abundances; this was only true for 50% of SI content and 66% of SI mucosal OTUs. Stool is therefore an excellent proxy for the LI lumen and mucosa, as it contains nearly all OTUs at preserved proportions.

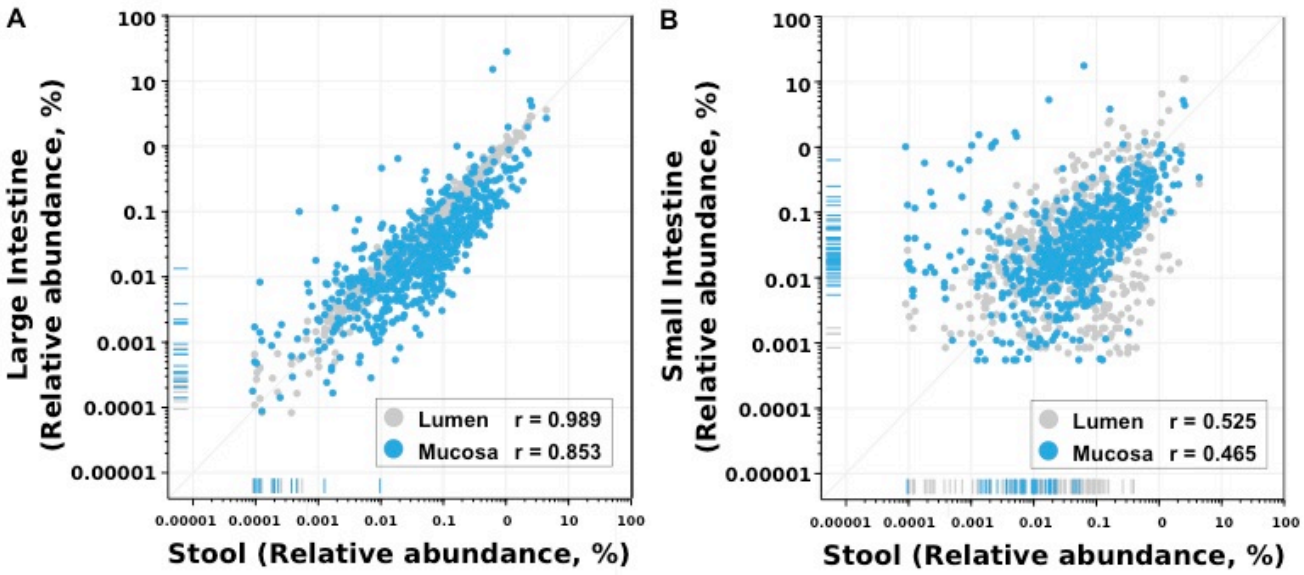


Figure 1-2: **Stool microbial composition mirrors that of the colonic lumen** Each dot corresponds to the average relative abundance of an OTU across 15 animals for each of 4 intestinal regions (SI mucosa and lumen, LI mucosa and lumen). To measure correlation, Spearman's r was calculated between stool and mean region OTU abundance. Marks on the x-axis (vertical lines) or y-axis (horizontal lines) margins represent OTUs with zero measured abundance at one site but non-zero abundance at the other. See also **Fig. 1-S2, Table 1-S2**.

Most OTUs are shared between adjacent sites, but each site has a small site-specific community

We found that ~40-70% of OTUs are typically shared between adjacent mucosal and lumenal sites (**Fig. 1-3**). It is unclear to what extent these overlapping taxa are persistent, metabolically active residents of the mucosa, rather than lumenal residents incidentally trapped on the mucosal surface (or vice versa). Although lumenal communities were generally more homogenous than those of the mucosa (**Fig. 1-2D**), 20-30% of OTUs were unique to each lumenal segment. As each mucosal sample contained a similar distribution of organisms within higher-order taxa, the variation we observed here at the genus or species level may be the result of colonization resistance by the more abundant members within similar functional groups. Whether the gut microbiota undergoes such nonrandom assembly remains unclear.

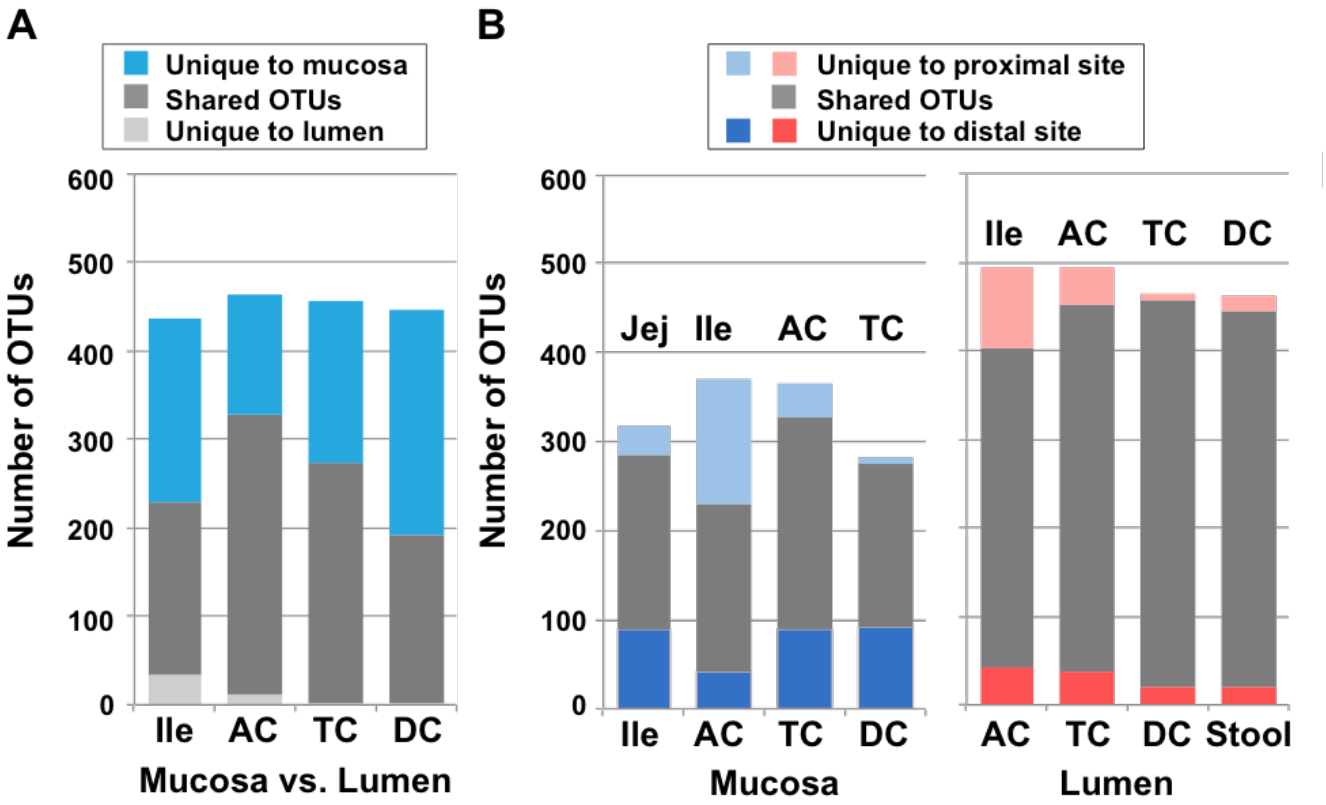


Figure 1-3: Microbial overlap between adjacent mucosal and lumenal sites / Differences in mucosal and lumenal community function are driven by *Helicobacter* A) Mean total, shared, and unique OTUs between the mucosa and lumen across all individuals at each paired site. See also Table S2. B) Mean total, shared, and unique OTUs between adjacent mucosal (left) and lumenal (right) sites, averaged across individuals. Most lumenal taxa are shared with the adjacent mucosa and lumenal sites, with a gradient of unique mucosal taxa occurring along the intestine.

Using logistic regression to distinguish mucosally and lumenally-enriched taxa and to predict bacterial flow

In order to understand microbial niches and potential bacterial flow within individuals, we used logistic regression to model the extent to which the abundance of a genus in mucosa and lumen (**Fig. 1-4A**) and at one site was predictive its abundance at an adjacent site (**Fig. 1-4B,** see Experimental Methods). Of the 56 genera identified in our cohort, 30% were mucosally-enriched, 40% were lumenally-enriched, and 30% showed no consistent enrichment. The proteobacteria comprised none of the lumenally-enriched taxa but one third of the mucosally-enriched taxa (6/17 genera). Conversely, nearly 70% of the lumenally-enriched taxa were Firmicutes. The mucosally-enriched genera were primarily gram-negative (13/17 genera) and frequently facultatively anaerobic (8/17 genera), while the lumenally-enriched taxa were primarily gram-positive (16/ 22 genera) and obligately anaerobic (19/22 genera). Most obligately anaerobic genera were not abundant in the mucosa; only *Treponema,* which is well-adapted to oxidative stress, showed a modest mucosal preference [154, 155]. This suggests that oxygen availability is a major, but not sole, determinant of mucosal composition [156].

Most mucosally-enriched genera identified here were not identified by univariate analysis because, with the exception of *Actinobacillus*, they were enriched only in either the SI or LI mucosa, but not both (e.g. *Klebsiella* in the LI, *Gemella* in the SI)(**Fig. 1-S3**); univariate analysis only detected mucosal enrichment consistent in both locations. The SI lumen was represented only by ileal samples, but most genera strongly enriched in the ileal lumen relative to the ileal mucosa (e.g. *Lactobacillus, Slackia*) were also strongly enriched at multiple locations along the LI lumen relative to the LI mucosa (**Fig. 1-4; Fig. 1-S3**). This lumenal community similarity may be partially explained by pH similarity between ileum (7.0-7.4) and colon (6.6-7.0) [157](**Fig. 1-S4C**).

Fig. 1-4B summarizes the relationships of β-values (regression slopes) between adjacent sites in the macaque gut. We observed four main patterns of microbial enrichment and potential flow: 1) SI mucosally-enriched taxa 2) LI mucosally-enriched taxa 3) SI and LI lumenally-enriched taxa and 4) clades following no consistent pattern. *Actinobacillus* and *Pasteurella* exemplify the pattern typically observed in SI mucosally-enriched clades (**Fig 1-S3**). They are most abundant in the ileal and jejunal mucosa, much less abundant in LI than SI, and more abundant in the LI mucosa than lumen. The differences in abundance between sites are very consistent, so the abundance at one site can be used to predict the abundance at another. Similarly, *Brachyspira* and *Helicobacter* are most abundant in the LI mucosa, and stool is highly predictive of their abundance elsewhere in the LI. *Lactobacillus*, *Ruminococcus*, and *Dialister* are enriched throughout the lumen, and predictably present in the mucosa at much lower abundance. Finally, several clades had predictable but atypical abundance patterns. For example, *Granulicatella* and *Enterococcus* were highly abundant in the SI mucosa (and nearly-absent in the SI lumen), and present at very low abundance in the LI lumen (but absent in the mucosa). *Pseudomonas* was only present in the SI mucosa, while *Klebsiella* was only present in the distal LI. In summary, using logistic regression modeling allowed us to group bacterial taxa that followed similar predictable ecological patterns across the intestine, and in some cases, we were able to predict where taxa may have originated within the intestine when observed in the stool.

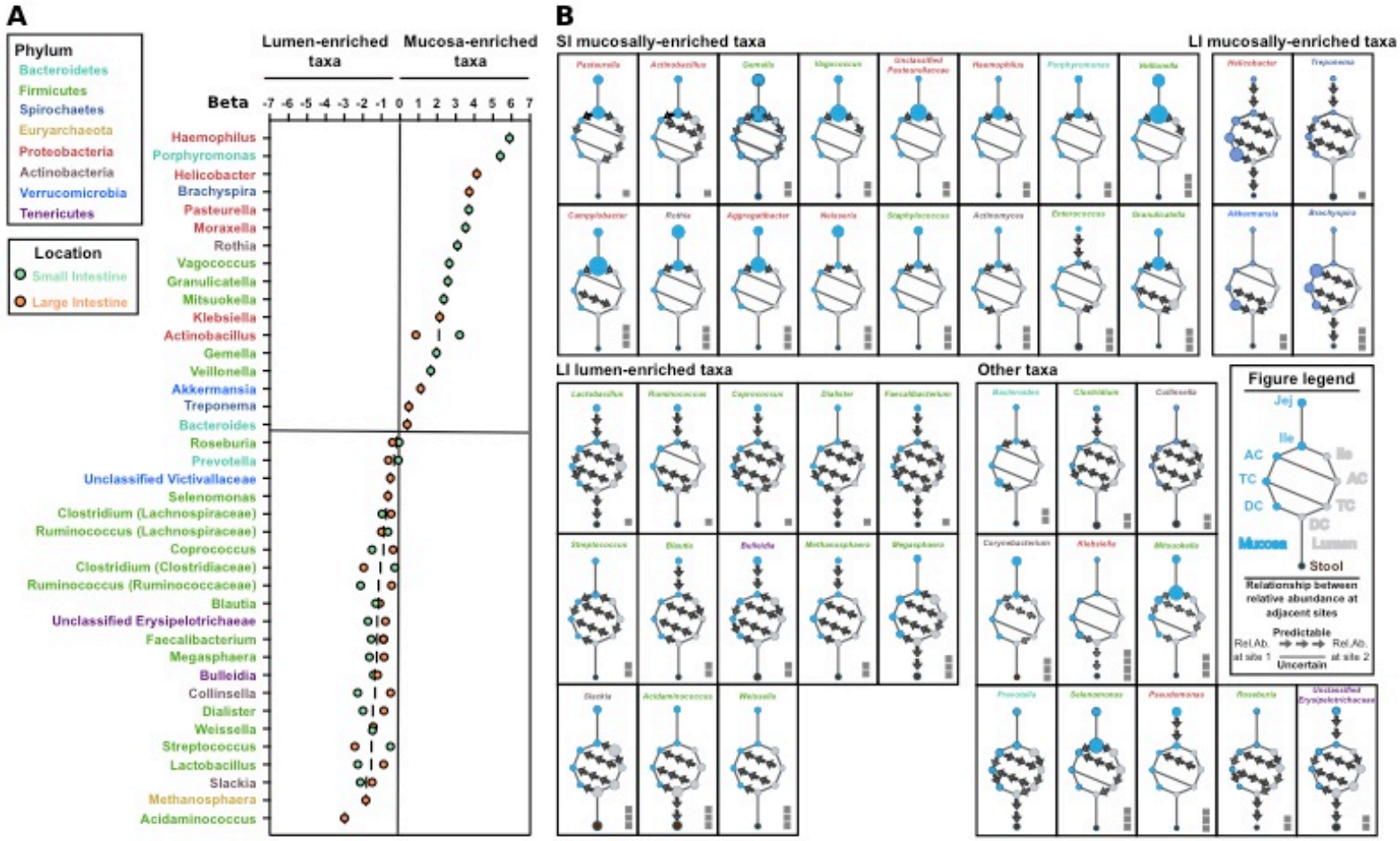

A
Phylum
Bacteroidetes
Firmicutes
Spirochaetes
Euryarchaeota
Proteobacteria
Actinobacteria
Verrucomicrobia
Tenericutes
Location
Small Intestine
Large Intestine
Lumen-enriched taxa
Mucosa-enriched taxa
Beta
-7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7
Haemophilus
Porphyromonas
Helicobacter
Brachyspira
Pasteurella
Moraxella
Rothia
Vagococcus
Granulicatella
Mitsuokella
Klebsiella
Actinobacillus
Gemella
Veillonella
Akkermansia
Treponema
Bacteroides
Roseburia
Prevotella
Unclassified Victivallaceae
Selenomonas
Clostridium (Lachnospiraceae)
Ruminococcus (Lachnospiraceae)
Coprococcus
Clostridium (Clostridiaceae)
Ruminococcus (Ruminococcaceae)
Blautia
Unclassified Erysipelotrichaeae
Faecalibacterium
Megasphaera
Bulleidia
Collinsella
Dialister
Weissella
Streptococcus
Lactobacillus
Slackia
Methanosphaera
Acidaminococcus
B
SI mucosally-enriched taxa
LI mucosally-enriched taxa
LI lumen-enriched taxa
Other taxa
Figure legend
Jej
Ile
AC
TC
DC
Mucosa
Lumen
Stool
Relationship between relative abundance at adjacent sites
Predictable
Rel.Ab. at site 1
Rel.Ab. at site 2
Uncertain

Figure 1-4: A logistic regression model of bacterial taxa site enrichment and flow through the macaque gut A) Average within small-intestine and within-large-intestine β values (regression slopes) for each genus in each biogeographic region. β corresponds to the magnitude of a difference in relative abundance between two sites (mucosa and lumen), and its consistency across 15 animals, and thus the enrichment of a taxa in mucosa or lumen. Only genera with at least one significant ($p < 0.05$ and $q < 0.05$) value for β are shown. B) Intra-intestinal microbial predictability between adjacent sites for each bacterial genus. Points on each clock-like diagram represent biogeographic sites, and point size corresponds to mean relative abundance across all animals for each genus. Adjacent sites with significant non-zero β (indicating that relative abundance at one site can predict that of the other) are connected with an arrow; arrows always start at the site with higher relative abundance. Solid lines indicate non-significant β, and arrows with significant β that point opposite of the physiological flow of intestinal contents inside the intestine (proximal to distal) are also replaced by a solid line. For visualization, the taxon relative abundances were adjusted by a factor of 10^1 to 10^3; the adjustment is indicated by the number of squares next to each diagram (e.g. *Pasteurella* was adjusted by 10^1 and has one square). See also **Fig 1-S2, Fig 1-S3**

Differences in mucosal and lumenal community function correspond to oxygen and nutrient availability

In order to understand the functional differences between communities at distinct biogeographic sites and their relation to community composition, we used PICRUSt [158] to infer community metabolic potential (Supplemental Methods), then used LEfSe [147] to identify functions that differed significantly between sites (see Methods; **Fig. 1-S4D; Table 1-S1**). PICRUSt is particularly useful for understanding mucosal microbial community function, as shotgun metagenomic sequencing of mucosal tissue samples yields high fractions (>90%) of host-derived nucleotides.

The largest functional difference between the mucosal and lumenal difference was an upregulation in riboflavin biosynthesis. This is likely due to the fact that, while *Helicobacter* and gram-positive bacteria both have fused *ribAB* genes in their riboflavin biosynthesis operons, *Helicobacter* also have an additional copy of the *ribA* gene [159]. One of the main functions upregulated in the mucosa relative to the lumen was glutamate / aspartate transport. Notably, *H.pylori* preferentially consumes amino acids as an energy source [160], and its glutamate and aspartate transport and deamidase activity are essential for mouse colonization[161]. Glycolysis was correspondingly up-regulated in the lumen, where *Helicobacter* was not dominant. The SI was enriched for lipopolysaccharide biosynthesis, consistent with its high abundance of gram-negative Pasteurellaceae. It was also enriched in mannose-specific phosphotransferase system, which has also been characterized in *Pasteurella* [162]. Relative to the SI, the LI was enriched in oxidative phosphorylation and archaeal ribosome, the former potentially due to *Helicobacter* and the latter due to *Methanosphaera* (**Fig. 1-S4D**).

Discussion

In this study, we comprehensively examined the composition of the macaque gut microbiome at 10 different biogeographic locations within 15 individuals. The most similar previous study cross-sectionally compared the gut microbiota of healthy and sick (e.g. Simian Immunodeficiency Virus (SIV)-infected) macaques, although it also included several biogeographical locations drawn from distinct individuals [70]. In contrast, our study included only healthy individuals and comprehensively examined the microbiota of the same individual at many biogeographic sites at the same time. We quantified for the first time the degree to which microbial composition at one biogeographical location within the gut predicts that of another, particularly the extent to which stool samples reflect the mucosal microbiome. We found that between stool and colonic mucosa, both the conservation of taxa and their rank correlation were remarkably high ($r>0.85$), supporting the use of stool samples for translational studies of colonic mucosal inflammation in human subjects.

The human, macaque, and mouse gut microbiomes are fundamentally similar in containing Bacteroidetes, Firmicutes, and Proteobacteria [5, 70, 143]. In contrast, the macaque gut mucosa was most remarkable for its abundance of **ε**-proteobacteria, specifically *Helicobacter*, which reached up to 80% relative abundance in the LI mucosa of some animals. The macaque gut also included a substantial component of the Spirochaetes *Treponema* and *Brachyspira*, which comprised ~3% of the mucosal microbiota. A recent study of the gorilla, chimpanzee, and bonobo microbiomes found that they also contained *Brachyspira* and *Treponema*, but no *Helicobacter* was detected in their stool [163]. While Spirochaetes carriage is typically associated with intestinal pathology in humans, it can be asymptomatic and is more typically found in stool in residents of non-developed countries [164]. Recent studies have detected *Brachyspira spp.* and *Treponema spp* OTUs in the stool of Malawians, Amerindians, and children from rural Burkina Faso and Bangladesh slums, but not in comparison cohorts from the USA

[138, 165, 166]. The disappearance of these taxa in residents of developed countries may be associated with modern sanitation practices, including use of antibiotics and pesticides.

Integrating these results with our prior knowledge of intestinal ecology and microbial metabolism [5, 70, 132, 137, 138] refines our insights into the ecological dynamics of the gut microbiome. Relative to the inter-individual differences observed in human populations, inter-individual differences in this study were minimal, yet they were still a significant source of variation despite all animals being uniformly fed and housed for at least two years prior to sampling. Intestinal oxygen content appeared to determine the dominant taxa colonizing the mucosa and lumen, as the mucosa was colonized primarily by facilitative anaerobes, while the lumen was colonized mainly by obligate anaerobes. This in turn dictates the main patterns of community functionality, as lumenal bacteria (Prevotellaceae Lachnospiraceae, Ruminococcaceae) are as a result primarily carbohydrate fermenters [148]. There were correspondingly large differences between small and large-bowel mucosal communities, potentially corresponding to difference in pH, bile salt, and/or mucin composition [167, 168], although the difference between the lumenal microbiota of the SI and LI was much smaller.

The relationship between the stool and mucosal microbiota is highly relevant to human clinical studies, as disease may localize to specific locations while only stool remains readily accessible [134]**.** Patient biopsies are the current gold standard for study of human-associated mucosal communities, but invasiveness and expense limit the frequency with which they can be performed. At the same time, an increasing body of data [16, 133] underscores the importance of studying the microbiome longitudinally during the development of disease, a near-impossibility with mucosal biopsies. Our results quantify the extent to which the stool microbiota reflects the SI and LI mucosa and is highly correlated with the colonic lumen and mucosa. Thus, although biopsies are optimal for studying the SI mucosa, our results

in a primate model representative of the human gut microbiome show that stool is still surprisingly representative of the colonic mucosa.

Acknowledgements

We thank L. Waldron, E. Schwager, and J. Moon for assistance with statistical analyses; W. Garrett and the Garrett and Huttenhower research groups for comments on this manuscript; S. Macri, E. Curran, L. Kattenhorn, and B. Assaf for assistance with sample collection; and T. Poon for assistance with sequencing. This work was supported by the Robert F. Naka fellowship (KY), Base Grant (P51 RR000168-48) (NEPRC), NIH R01HG005969 (CH), NSF DBI-1053486 (CH), and ARO W911NF-11-1-0473 and W911NF-11-1-0429 (CH).

Experimental Procedures

Sample collection and sequencing

Stool and paired intestinal lumenal and mucosal samples from 10 segments distal to jejunum were collected from 15 clinically-healthy female rhesus macaques, and the V4 region of the16S rRNA gene was sequenced by Illumina MiSeq. All further details of animal husbandry, sample collection, preparation, sequencing, and bioinformatic processing are outlined in the Supplemental Data.

OTU overlap between sites

For each pair of adjacent sites, the number of OTUs observed in both adjacent sites was counted in each individual and subsequently averaged across all 15 animals. To minimize the influence of low prevalence OTUs and differences in sequencing depth, only OTUs with 15 reads per OTU in 15 animals were considered in this analysis.

Identification of microbial taxa enrichment sites and predictability by logistic regression

We built a logistic regression model for each taxa between each adjacent biogeographical site pair, as described in Supplemental Methods. The regression slope β between each pair of locations is calculated as the contrast between the coefficients of indicator variables corresponding to the locations. P-values of all β are adjusted for multiple hypothesis testing using the Benjamini-Hocheberg procedure (false discovery rate, FDR=0.05). Cytoscape was used to visualize **Figure 1-4B**.

Sequence accession numbers and availability

Sequences generated in this study are publicly available (NCBI BioProject ID number PRJNA259224)**.**

Supplemental Experimental Procedures

Animals and sample collection

All animals were housed at the NEPRC in accordance with all applicable regulations and in a facility accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International. Animals were maintained under an experimental protocol approved by Harvard Medical School's Standing Committee on Animals. Prior to sample collection, animals were housed and fed individually.

Intestinal lumen (ileum, ascending, transverse, and distal colon), mucosal scrapings (jejunum, ileum, ascending, transverse, descending colon), and stool samples were collected during autopsy from 15 clinically-healthy female rhesus macaques, ranging from 12 to 22-years old (See Table Cohort Metadata).

The entire intestinal tract was first removed from the body. Next, a 15-cm section from each biogeographical location was cross-sectionally transected, and then longitudinally transected on the anti-mesenteric side of the intestine to open the intestinal lumen (**Fig. 1-S4**). Lumenal samples were collected by advancing the lumenal contents into a cryotube (Nunc CryTubes, Sigma-Aldrich, St. Louis, MO) using a sterile spatula. Intestinal contents were removed from the lumen and rinsed with sterile saline to remove any visible contents without disturbing the intestinal mucosa. It was not possible to collect jejunal lumenal contents due to fasting of the animals prior to euthanasia. Intestinal mucosal samples were then collected by gently scraping the mucosal surface with a sterile glass slide (to avoid penetrating the basement membrane) and scraped samples were advanced to a cryotube. All intestinal samples were snap frozen in liquid nitrogen and stored at -80°C for further analysis. All histopathology of the intestinal tissues and major organs was normal.

16S rRNA sequencing and profiling

DNA from stool, mucosal, and lumenal samples was extracted using the MP BIO FASTDNA™ SPIN Kit for Soil (MP Bio, Santa Ana, CA) according to manufacturer's instructions. The amplification and sequencing of the V4 region by Illumina MiSeq were performed as described previously [138]. In brief, genomic DNA was subjected to 16S amplifications using primers designed incorporating the Illumina adapters and a sample barcode sequence, allowing directional sequencing covering variable region V4 (Primers: 515F [GTGCCAGCMGCCGCGGTAA] and 806R [GGACTACHVGGGTWTCTAAT]). PCR mixtures contained 10 µl of diluted template (1:50), 10 µl of HotMasterMix with the HotMaster Taq DNA Polymerase (5 Prime), and 5 µl of primer mix (2 µM of each primer). The cycling conditions consisted of an initial denaturation of 94°C for 3 min, followed by 30 cycles of denaturation at 94°C for 45 sec, annealing at 50 °C for 60 sec, extension at 72°C for 5 min, and a final extension at 72°C for 10 min.

Amplicons were quantified on the Caliper LabChipGX (PerkinElmer, Waltham, MA), pooled in equimolar concentrations, size selected (375-425 bp) on the Pippin Prep (Sage Sciences, Beverly, MA) to reduce non-specific amplification products from host DNA, and a final library size and quantification was performed on an Agilent Bioanalyzer 2100 DNA 1000 chips (Agilent Technologies, Santa Clara, CA). Sequencing was performed on the Illumina MiSeq platform (version 2) according to the manufacturer's specifications with addition of 5% PhiX, and yielded paired-end reads of 175 bp in length in each direction.

16S sequence bioinformatic processing

Overlapping paired-end reads were stitched together (approximately 97 bp overlap), size-selected to reduce non-specific amplification products from host DNA (225 - 275 bp), and further processed in a

data curation pipeline implemented for PICRUSt [158] in QIIME 1.6.0 as pick_closed_reference_otus.py [169]. In brief, this pipeline will (i) pick OTUs using a reference-based method and then (ii) constructs an OTU table. Taxonomy is assigned using the Greengenes (18 May 2012 version) predefined taxonomy map of reference sequence OTUs to taxonomy [170]. The resulting OTU tables are checked for mislabeling and contamination [171].

A mean sequence depth of 29,914/sample was obtained; samples with fewer than 3,000 filtered sequences and those Operational Taxonomic Units (OTUs) with less than 15 reads were excluded from downstream analysis. Further microbial community analysis such as beta diversity was calculated with QIIME 1.6.0 [169]. To test for statistically significant association between the microbiota and metadata including biogeographical locations, we used LEfSe (Segata et al., 2011) for univariate and MaAsLin (Multivariate Associations by Linear models) [133] for multivariate analyses (**Table 1-S1**). We used LEfSe to identify features (microbial taxa) that separate two classes (mucosa vs. lumen or small vs. large intestine) and quantify effect sizes (i.e. biological magnitude) of the association. We used MaAsLin to build a multivariate linear model combining fixed and random effects to identify associations between microbial communities with covariates including sample type (mucosa vs. lumen), locations (jejunum, ileum, ascending, transverse, and distal colon, and stool), age, body weight, and primate center origin). We controlled for individuals. For MaAsLin data, we used Benjamini-Hochberg false discovery rate corrections to accept no more than a 20% FDR.

In order to predict microbial functions from the microbial data, we used PICRUSt [158]. This algorithm estimates the functional potential of microbial communities given a marker gene survey and the set of currently-sequenced reference genomes with an accuracy of 80-90% on human gut communities. Although predicted metagenomes derived from PICRUSt provide informative functionalities of the microbial community, they are often specific (e.g. glycerol-3-phosphate dehydrogenase (NAD+)). We

thus used HUMAnN [172] to identify KEGG modules (version 56) based on the metagenome predicted from the 16S sequencing data using PICRUSt. KEGG module is a collection of manually-defined functional units and can be used to interpret biological functions of metagenomic data. The result of the univariate (LEfSe) and multivariate (MaAsLin) analyses are included in **Fig. 1-S4D** and **Table 1-S1**.

To assess the similarity of our data to previously-published macaque and human studies, microbiota data, either taxonomic or raw sequencing data were obtained from publically available sources (Handley et al., - RG-RAST: http://metagenomics.anl.gov/?page=MetagenomeSelect; Human Microbiome Project (HMP) – http://www.hmpdacc.org/reference_genomes/reference_genomes.php; Yatsunenko - https://gordonlab.wustl.edu/SuppData.html) or directly from the investigator (McKenna et al., 2008). Taxonomic tables were summarized to genus-level clades and merged. All studies except for Yatsunenko et al and the current study used different PCR amplification methods, sequencing platforms, and variable regions of the 16S rRNA gene (see below). The Bray-Curtis distance was used to assess the similarity between all five communities (Figure 1-S1D). Since Yatsunenko et al. and the current study used the same methods to amplify, sequence, and assign taxonomy, the weighted Unifrac distance, which measures the phylogenetic relatedness as well as the counts of each taxa, was used to assess similarity between the Yatsunenko dataset and the current study (Figure 1-S1E).

Study Name	Host Species	Sequence	Method	Regions	Sample Type
Human Microbiome Project	Human - US	454	16S	V1-V3, V3-V5	Stool
Yatsunenko *et al*	Human - US	Illumina	16S	V4	Stool
Yatsunenko *et al*	Human - Ameridian	Illumina	16S	V4	Stool
Yatsunenko *et al*	Human - Malawi	Illumina	16S	V4	Stool
McKenna *et al*	Macaque	454	16S	V1-3	Stool + Biogeography
Handley *et al*	Macaque	454	Shotgun metagenome		Stool
Yasuda *et al*	Macaque	Illumina	16S	V4	Stool + Biogeography

Identification of microbial taxa enrichment sites and predictability by logistic regression
For each OTU,

Where

, proportion of this OTU at this location;

, indicator variable for location;

, indicator variable for subject;

, reads corresponding to this OTU at this location for this subject;

, reads for all OTUs at this location for this subject

The *Circular layout* option included in Cytoscape [173] (Cytoscape version 3.0.1.) was used to visualize the predictability of microbial taxa between adjacent biogeographical sites for each taxa. The

direction of β (positive, negative, and none-significant) was used as the type of interaction, and attributes included relative abundance of each taxa at each location, and magnitude of β derived above. Although in some cases when abundances of distal sites are higher than proximal sites (i.e. abundances in stool are higher than distal colon lumen), in those cases, the negative βs suggested that this bacterial taxa may go from stool to distal colon, the fact that this is unlikely in reality considering the natural flow of intestinal contents. Therefore, when the direction of β (either positive, or negative) opposed the actual physiological flow (we assumed the actual physiological flow to be always proximal to distal amongst mucosa and lumen and interchangeable between mucosa and lumen), the errors were substituted with lines and combined with the non-significant group, which was also noted as a line.

This page intentionally left blank

CHAPTER 2

Fluoride depletes acidogenic taxa in oral but not gut microbial communities in mice

This Chapter is a reproduction of a published manuscript:
Yasuda, K., Hsu, T., Gallini, CA., Mclver, LJ., Schwager, E., Shi, A., DuLong, CR., Schwager, RN., Abu-Ali, GS., Franzosa, EA., Garrett, WS., Huttenhower, C., Morgan, XC. *mSystems*. 2017:2;4(e00047-17).

Author Contributions:
KY, WSG, CH, and XCM designed the study. KY, CAG, and CRD performed the experiments. KY, TH, ES, AS, RNS, GSA, LJM, EAF, CH and XCM analyzed the data. KY, TH, EAF, WSG, CH, and XCM wrote the manuscript.

Abstract

Fluoridation of drinking water and dental products prevents dental caries primarily by inhibiting energy harvest in oral cariogenic bacteria (such as *Streptococcus mutans* and *Streptococcus sanguinis*), thus leading to their depletion. However, the extent to which oral and gut microbial communities are affected by host fluoride exposure has been under-explored. In this study, we modeled human fluoride exposures to municipal water and dental products by treating mice with low or high levels of fluoride over a 12-week period. We then used 16S amplicon and shotgun metagenomic sequencing to assess fluoride's effects on oral and gut microbiome composition and function. In both the low- and high-fluoride groups, several operational taxonomic unit (OTUs) belonging to acidogenic bacterial genera (such as *Parabacteroides*, *Bacteroides*, and *Bilophila*) were depleted in the oral community. In addition, fluoride-associated changes in oral community composition resulted in depletion of gene families involved in central carbon metabolism and energy harvest (2-oxoglutarate ferredoxin oxidoreductase, succinate dehydrogenase, and the glyoxylate cycle). In contrast, fluoride treatment did not induce a significant shift in gut microbial community composition or function in our mouse model, possibly due to absorption in the upper gastrointestinal tract. Fluoride-associated perturbations thus appeared to have a selective effect on the composition of the oral but not gut microbial community in mice. Future studies will be necessary to understand possible implications of fluoride exposure for the human microbiome.

Importance

Fluoride has been added to drinking water and dental products since the 1950's. The beneficial effects of fluoride on oral health are due to its ability to inhibit the growth of bacteria that cause dental caries. Despite widespread human consumption of fluoride, there have only been two studies in humans that considered the effect of fluoride on human-associated microbial communities, which are increasingly understood to play important roles in health and disease. Notably, neither of these studies included a true cross-sectional control lacking fluoride exposure, as study subjects continued baseline fluoride treatment in their daily dental hygiene routines. To our knowledge, this work (in mice) is the first controlled study to assess the independent effects of fluoride exposure on the oral and gut microbial communities. Investigating how fluoride interacts with host-associated microbial communities in this controlled setting represents an effort towards understanding how common environmental exposures may potentially influence health.

Introduction

Since the 1940s, fluoridation of drinking water and dental products has been employed as a public health measure to prevent dental caries. In the United States, more than 60% of municipal water is fluoridated, and the majority of dental products contain fluoride [174-176]. Fluoridated compounds improve oral health by inhibiting bacterial growth through inhibition of the enzyme enolase, which catalyzes the conversion of 2-phosphoglycerate to phosphoenolpyruvate (the last step of anaerobic glycolysis) and thus is critical for microbial energy harvest and growth [128, 129]. Inhibition of individual oral bacteria such as *Streptococcus mutans* by fluoride has been well studied [177-179], but how fluoride affects the overall oral microbiome, or that of the gut, has been under-investigated.

The major mechanisms by which fluoride inhibits bacterial energy growth are direct binding of the fluorine ion to the active site of enolase [129] and ATPases [130], and disruption of the ion gradient across the bacterial cell membrane [180, 181]. All of these mechanisms result in the reduction of adenosine triphosphate (ATP) synthesis [130]. Although a wide range of bacterial taxa are inhibited by this mechanism, the degree of resistance differs across taxa. For example, the enolases of *S. mutans and S. sanguis* are more susceptible (by 10-fold) to fluoride than those of *S. salivarius* and *Lactobacillus casei* in monoculture system [130, 131]. This variation in fluoride resistance raises the question of how fluoride affects individual bacterial taxa within a complex microbial community. Two recent studies have begun extending this line of investigation to dental plaque microbial communities using high-throughput sequencing [182, 183]. Unfortunately, neither of these studies included a true control group to properly assess fluoride's effects on the oral microbiome (individuals in the studies' control groups retained access to commercial fluoride-containing dental products and fluoride in drinking water during the experimental period). One of these studies further considered a single dose of sodium fluoride mouthwash [182] and was therefore not designed to assess the effects of longer-term exposures. Thus,

no existing study has tested the effects of fluoride exposures at the levels found in municipal water and dental products on the oral and gut microbial communities.

We address these questions by assessing oral and stool microbiome structures and their functional potentials in mice given 1) non-fluoridated drinking water, 2) fluoridated drinking water, or 3) daily fluoride gavage in addition to fluoridated drinking water over 12 weeks. 16S rRNA gene amplicon and shotgun metagenomic sequencing revealed that fluoride exposure significantly perturbed oral but not gut microbiome composition in mice. Specifically, fluoride selectively depleted oral acidogenic bacteria, including *Bacteroides*, *Parabacteroides*, and *Bilophila*. In terms of the microbial functional profiles, fluoride exposure selectively depleted metabolic modules important for central carbon metabolism. Our results support that fluoride-associated perturbations have a selective effect on the composition of the oral microbial community in mice. Though limited by lack of human data, this study suggests that levels of fluoride currently added to drinking water and associated with routine use of dental products are unlikely to have significant effects on established gut microbial communities.

Results

Drivers of oral and gut microbial diversity in fluoride-treated mice

To elucidate the effects of chronic fluoride exposure on oral and gut microbial communities, wild type BALB/c mice (1 month of age) were randomized upon weaning to the following experimental groups: 1) unfluoridated (deionized) drinking water, 2) fluoridated drinking water (4 ppm), or 3) fluoridated drinking water (4 ppm) plus a daily gavage of fluoride similar to a dose ingested when swallowing dental fluoride products (2.25 micrograms of fluoride per day via gavage) for a period of 12 weeks. Oral samples were collected at 0 and 12 weeks, and stool samples were collected at 0, 4, 8, and 12 weeks. Since the collection of oral samples required animals to be sacrificed, 8 mice were taken for oral week-0 sample at random when the rest of the animals were allocated to the study groups. Samples were sequenced for the 16S rRNA gene V4 amplicon (referred to hereafter as 16S) and taxonomically profiled to measure the effects of fluoride on the microbial community (see **Methods**).

We first examined the major factors driving microbial diversity across our dataset by applying ADONIS on weighted UniFrac distance. Biogeographical site (oral vs. gut) explained the largest fraction of between-sample diversity (45%, p<0.001; **Fig. 2-1A**). This observation is consistent with the strong effect of biogeography on microbial community structure seen in other mammals [47, 71]. Downstream non-parametric comparison of oral and gut samples [147] revealed statistically significant enrichments of *Streptococcus spp.* and *Pastuerellaceae* species in the oral site, while the gut was enriched for *Bacteroides, Clostridiales*, *Lachnospiraceae*, and *Parabacteroides* (**Fig. 2-1E** and **Table 2-S1**). Notably, these clades are similarly enriched in human oral and gut sites, respectively [5, 71, 83].

Housing cage (**Fig. 2-S1A**) and treatment time point (**Fig. 2-1B** and **Fig. 2-S1B**) explained additional, significant fractions of between-sample diversity (29% and 4%, respectively). Specifically,

within-cage stool communities (**Fig. 2-S1 A** and **E**) were significantly more similar than between-cage communities

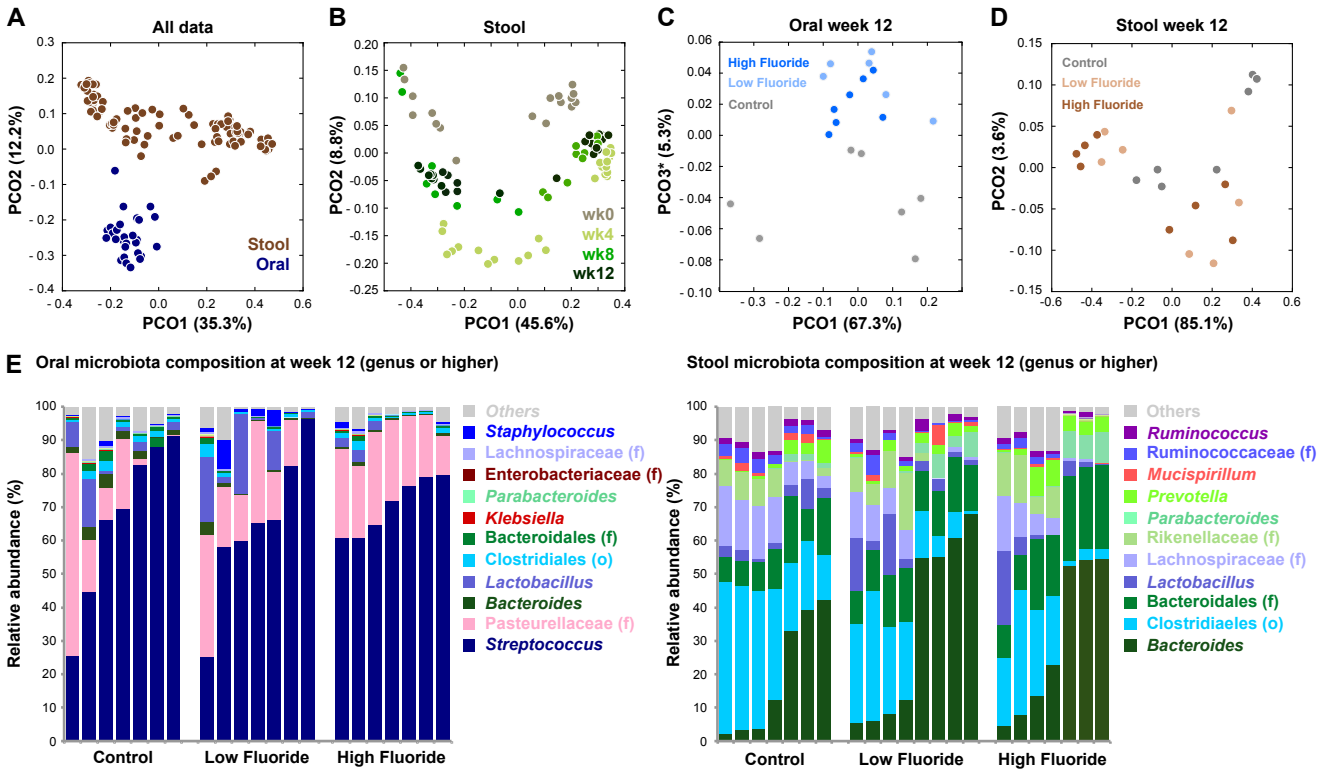


Figure 2-1: Drivers of oral and gut microbial diversity in fluoride-treated mice. (A) Principal coordinates analysis (PCoA) of all samples by weighted UniFrac distance. (B) PCoA of only stool samples by weighted UniFrac distance. (C) PCoA of oral samples at week 12 by weighted UniFrac distance. (D) PCoA of stool samples at week 12 by weighted UniFrac distance. (E) Genus-level relative abundance of the oral (left) and stool (right) microbiota. Individual columns represent each animal and are grouped by treatment (control, low, and high fluoride).

(p<0.001, Mann-Whitney U-test on weighted UniFrac distance), as observed in previous studies [184]. These cage effects corresponded to recognizable sample subgroups dominated by either 1) Clostridiales and Lachnospiraceae or 2) *Bacteroides* and *Parabacteroides*. Curiously, we did not find oral samples to be significantly more similar within-cage than between-cage (p=0.34), suggesting that cage effects may be less pronounced among the mouse oral microbiota. Based on these observations, we explicitly controlled for cage effects in downstream analyses of stool.

While overview ordination suggested a slight separation among week-12 oral samples associated with fluoride treatment (**Fig. 2-1C**), we did not observe a separation in stool (**Fig. 2-1D** and **Fig. 2-S1D**), and the component of between-sample diversity explained by fluoride was not statistically significant by ADONIS analysis (p=0.860). This suggests that fluoride treatment does not have a large effect on the global structure of the oral or gut microbiota in mice. However, it remained possible that fluoride treatment could selectively alter the abundance of individual microbial taxa and functions in these environments.

Fluoride selectively depletes oral acidogenic taxa in mice

To identify bacterial taxa selectively affected by fluoride in the oral site, we performed multivariate analyses using MaAsLin [15] with fluoride exposure coded as a as a categorical variable: high-fluoride vs. low-fluoride vs. no exposure (control). We considered only the week-12 oral samples (n=21) and isolated associations with FDR-corrected q-value<0.2 as statistically significant. In both the low- and high-fluoride groups, obligate anaerobes such as *Parabacteroides distasonis*, *Bacteroides uniformis*, and an unclassified species of *Bacteroides* were consistently depleted as compared to the control group (**Fig. 2-2A** and **Table 2-S2**). *Sutterella* and *Bilophila* were also depleted in fluoride-

treated animals, but only significantly so in the high-fluoride group (**Fig. 2-2B**). Such a pattern is consistent with

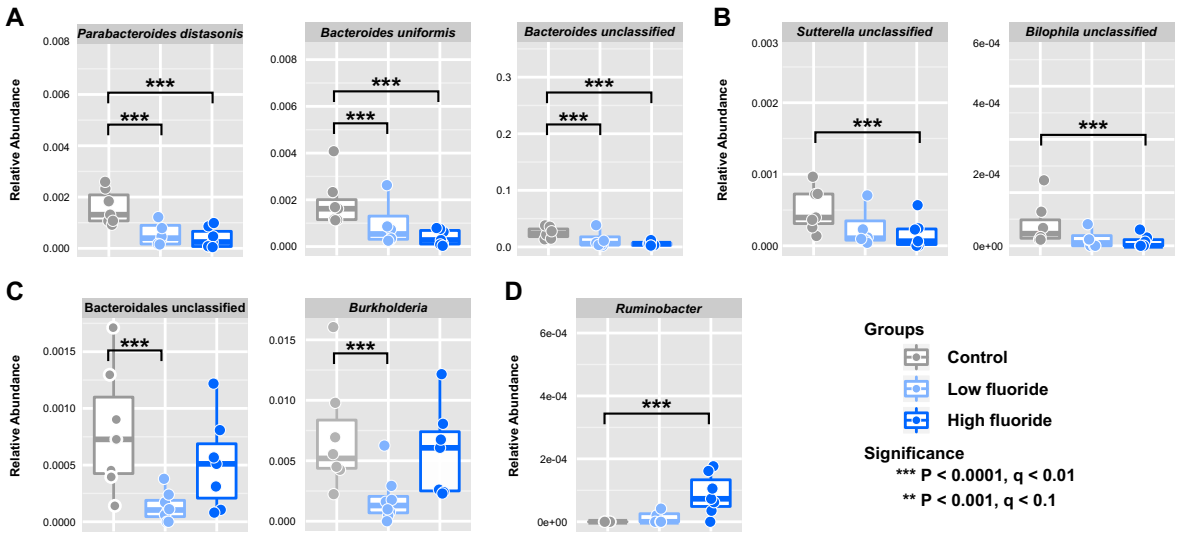


Figure 2-2: Fluoride selectively depletes acidogenic anaerobes in the oral microbiota. Multivariate linear model association results [15] showing bacterial OTU that are (A) consistently depleted in low- and high-fluoride treatment groups, (B) affected only in the high-fluoride group, or (C) the low-fluoride group as compared to controls. (D) A bacterial OTU enriched in the high-fluoride group.

previous *in vitro* studies of dosage-dependent inhibition of enolase in different microbial oral isolates [185]. Curiously, unclassified species of Bacteroidales and *Burkholderia* were significantly depleted only in the low-fluoride group compared to controls (**Fig. 2-2C**), possibly due to more extreme depletion of other taxa in the high-fluoride group. A similar mechanism would explain the expansion of another obligate anaerobe (an unclassified *Ruminobacter* species) in the high-fluoride group relative to the controls (**Fig. 2-2D**).

We performed a similar multivariate analysis to identify fluoride-sensitive taxa among stool samples. Relative to the model described above, we also considered treatment time point and animal cage as covariates (i.e. week-4, week-8, and week-12 samples were all considered). However, no species- or genus-level taxa showed significant (FDR q-value<0.2) associations with the high- or low-fluoride groups relative to controls. Results were similar when applying the model separately to samples stratified by treatment time point. While our inability to detect a significant treatment effect in stool could be a result of small sample size, an alternative explanation is that orally administered fluoride does not reach the gut in large-enough quantities to perturb the gut microbiota. Previous research has suggested that fluoride absorption occurs mainly in the stomach and upper small intestine [186]. We confirmed this result by directly measuring fluoride levels in the stool of treated vs. untreated mice. Even after twelve weeks of high-fluoride treatment, fluoride levels in treated stool were not significantly higher than baseline (control) stool fluoride levels (t-test, one-tailed p>0.05; see **Methods** and **Fig. 2-S2**). Hence, while the gut microbiota may be sensitive to fluoride, orally administered fluoride (even at high doses) is unlikely to expose this sensitivity.

Fluoride perturbs predicted oral microbial community functional potential

In addition to affecting individual microbial taxa, fluoride exposure may alter community-level function by enriching or depleting taxa that encode specific metabolic modules. Notably, this could be

true in the gut, even though individual taxa failed to show a significant effect there. To test this hypothesis in our data, we used PICRUSt [187] to infer community gene content from 16S amplicon sequencing data, followed by HUMAnN [172] to reconstruct functional modules. We then applied the same multivariate testing framework described above to identify modules whose relative abundance varied with fluoride treatment (focusing on modules with relative abundance >0.001% in at least 5 samples).

No functional modules varied significantly with fluoride treatment in the gut after adjusting for treatment time point and animal cage effects (all FDR-corrected *q*-values >0.2). However, in the oral samples, 19 of the 113 observed functional modules were differentially abundant amongst the treatment groups (**Table 2-S3).** Fluoride treatment was associated with depletions in the glyoxylate cycle (M00012), succinate dehydrogenase (M00149), and second-carbon oxidation (M00311) reflecting perturbed central carbon metabolism (**Fig. 2-S3A**). In addition, depletion in the mevalonate (MVA) pathway for isoprenoid biosynthesis (M00095) was also associated with fluoride treatment. 3-hydroxy-3-methylglutaryl-coenzyme A reductase, a key enzyme in the MVA pathway, has previously been shown to be fluoride-sensitive [188, 189]. Carriage of the MVA pathway is limited to gram-positive taxa such as *Streptococcus, Lactobacillus,* and *Staphylococcus* [190]. Individually, these taxa were not found to be significantly depleted in fluoride-treated communities via multivariate analysis (**Table 2-S2**). This analysis suggests that the changes reflected in differentially abundant oral functional modules may result from cumulative small changes in microbial community composition rather than large changes in specific taxa.

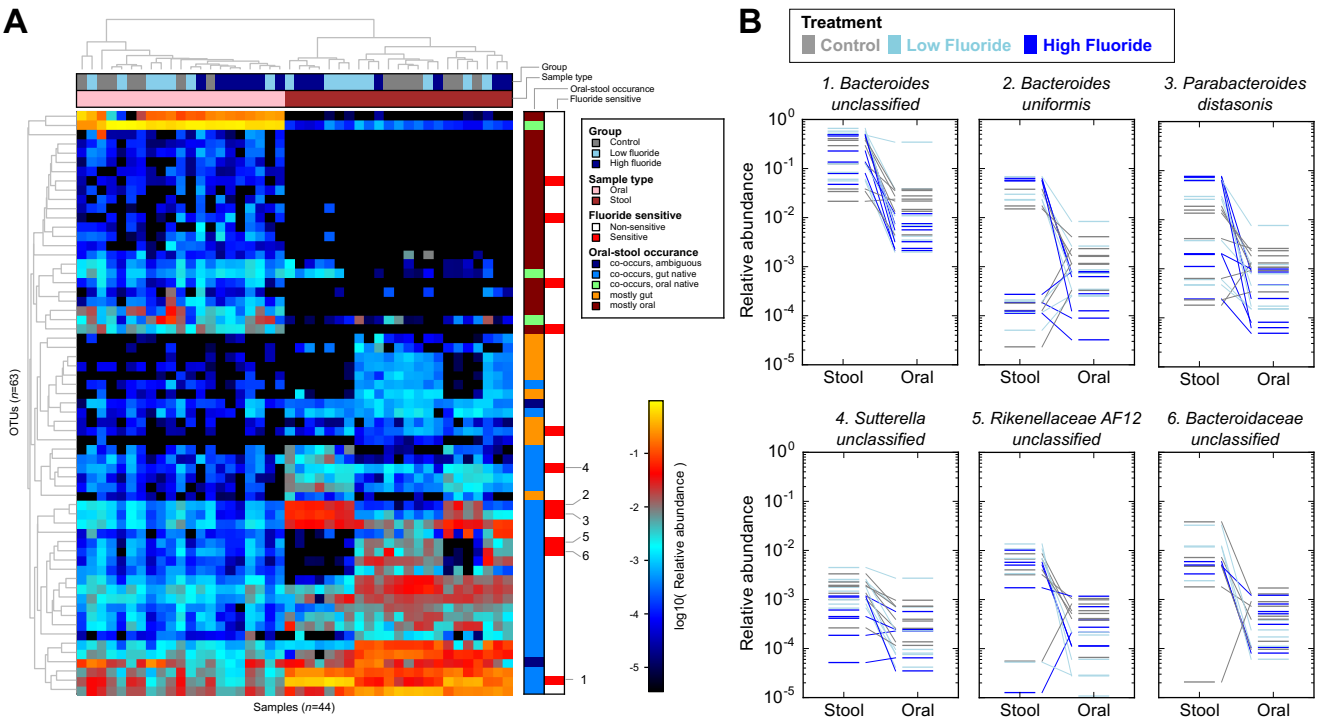


Figure 2-3. Fluoride affects stool-derived taxa found in the oral cavity. (A) 63 abundant OTUs (rows) across oral and stool samples (columns). Rows and columns are clustered by Bray-Curtis dissimilarity. OTUs are color-labeled according to their biogeographic occurrence/co-occurrence patterns (see main text for definitions). OTUs that were significantly depleted in fluoride-treated oral samples are highlighted in red, including a subset of orally-abundant, stool-derived OTUs. The six OTUs in this subset with the greatest treatment effects are highlighted in (B). Horizontal lines represent individual relative abundance measurements (colored by treatment group) and measurements from the same animal are connected.

The quality of the predicted functional profiles analyzed above can be estimated from the nearest sequenced taxon index (NSTI), which measures the closeness of a 16S-based profile to known reference genomes. NSTI values for our samples were low (i.e. close to reference) and mean NSTI scores for oral and gut samples were 0.047 and 0.148, respectively (**Table 2-S4** and **Fig. 2-S3C**). This range of NSTI values is 1) consistent with other non-human-mammal-associated samples and 2) suggests that predicted versus measured functional profiles for these samples should be in reasonably strong agreement [see Fig. 2-3 in [187]]. This agreement was also directly measured by subjecting a subset of samples to shotgun metagenomic sequencing and profiling (described in a subsequent section).

Fluoride affects stool-derived microbes in the mouse oral microbial community

We hypothesized above that the lack of fluoride treatment effect in the gut could be due to the low concentration of fluoride reaching that environment. In principle, gut microbes exposed to fluoride in the oral cavity (where concentrations are expected to be higher during treatment) might reveal additional sensitivity. While oral and gut microbial taxa are largely distinct in humans [71], mice are coprophagic, and hence have much greater potential for co-occurrence of oral and gut microbes (thus providing a basis to test this hypothesis).

We began by dividing mouse OTUs according to their biogeographic occurrence patterns (**Fig. 2-3A** and **Fig. 2-S4**). We focused on OTUs that were confidently detected among the week-12 oral or gut samples, defined as having relative abundance $>10^{-4}$ (0.01%) in at least five samples from at least one body site. OTUs that were confidently detected in only a single site were classified as "mostly oral" (n=21) and "mostly gut" (n=10). 32 additional OTUs were confidently detected at both sites. We further divided these OTUs into groups based on their likely point-of-origin. "Co-occurring, oral native" OTUs

(*n*=3) were defined to have mean oral abundance >2x larger than mean gut abundance, and may co-occur as a result of oral-gut transit. Conversely, 27 OTUs were classified as "co-occurring, gut native" due to >2x larger mean abundance in the gut (the two remaining co-occurring OTUs were classified as having "ambiguous" point of origin). Hence, oral sites are colonized by a relatively large number of gut bacterial taxa in mice, which is likely the result of direct or indirect ingestion of stool (with indirect including grooming of stool-contaminated body parts).

We next re-examined the behavior of the 27 orally-occurring, stool-derived taxa upon fluoride treatment based on the modeling results described above. While none of these taxa were significantly perturbed in the analysis of stool data, six were differentially abundant between week-12 control and low- or high-fluoride treatment oral samples. The most abundant of these include *Parabacteroides distasonis*, *Bacteroides uniformis*, and unclassified species of *Bacteroides*. (**Fig. 2-3B**). In fact, the taxa that were depleted among fluoride-treated oral samples were weakly enriched for taxa derived from the stool (Fisher's exact test, two-tailed p=0.034). This suggests that species native to the mouse stool microbiota are indeed sensitive to fluoride in concentrations typical of fluoridated water or dental products. However, these taxa are likely protected from fluoride exposure in stool due to the absorption of fluoride in the stomach and small intestine.

Targeted metagenomic sequencing supports 16S-based conclusions

In addition to the 16S-based sequencing profiles introduced above, we assayed subsets of oral (*n*=6) and stool (*n*=11) samples by shotgun metagenomic sequencing and profiled them with MetaPhlAn2 [191] (for microbial taxonomy) and HUMAnN2 (for gene families and pathways). While the shotgun-sequenced subset was too small for independent, well-powered statistical analysis, it proved

useful for supporting our 16S-based findings and for boosting taxonomic and functional resolution.

To facilitate shotgun-16S comparisons, all taxonomic features were collapsed to the family level. Of the 32 microbial families detected by reference-based shotgun metagenomic profiling, all were detected among the 16S profiles. 37 additional families were seen only in the 16S data, including several families that are uniquely enriched in mouse [Turicibacteraceae and Odoribacteraceae [11]; **Table 2-S5**]. This result underscores the utility of 16S-based taxonomic profiling in this study for detecting and quantifying clades that are underrepresented in isolate genome catalogs.

Conversely, reference-based shotgun metagenomic profiles were advantageous in providing increased taxonomic resolution in our dataset. For example, unclassified species of *Streptococcus* and Pasteurellaceae discovered through 16S sequencing were revealed in the shotgun sequencing data to be *Streptococcus parasanguinis* and *Haemophilus parainfluenza* (**Table 2-S5**). While we detected no major fungal community members by shotgun sequencing, several viral species were detected (fungi and viruses are notably invisible to 16S sequencing). Excluding trace viruses (i.e. present in one sample at <0.01% relative abundance), the commonly occurring viruses were mouse mammary tumor virus and murine osteosarcoma virus. Because these viruses belong to Retroviridae, their detection may be due to endogenous copies in the mouse genome that eluded host-read depletion during metagenomic quality control (e.g. due to absence or divergence from the mouse reference genome; see **Methods**).

In addition to the NSTI-based evaluation of our predicted functional profiles, we directly compared the predicted versus measured functional profiles for the subset of samples subjected to both 16S and shotgun metagenomic sequencing. Predicted versus measured gene family abundance [KEGG Orthogroups (KOs)] were reasonably concordant as measured by Spearman correlation, which ranged from 0.45 to 0.66 (**Fig. 2-S3**). Notably, these values are in line with the expected agreement between

predicted and measured KO abundance profiles inferred from NSTI scores [see Figure 2-3 in [187]], which lends further support to the accuracy of all predicted functional profiles considered in this study.

While most shotgun metagenomes were saturated with respect to their measured functional richness (see **Fig. 2-S3**), under-sampling of low-abundance KOs could in principle exaggerate the apparent disagreement between inferred and measured functional profiles. However, the Spearman correlations cited above should be reasonably robust to such events, meaning that any such exaggeration would be small. We suspect that disagreements between our 16S-inferred and metagenomic functional profiles are largely driven by the dependence of inferred profiles on a (complete) microbial reference-genome catalog, which may be missing ideal representatives for certain mouse-associated species (consistent with the low but non-zero NSTI scores described above).

Discussion

Our study surveyed the effect of chronic fluoride intake on oral and gut mouse microbial communities. Specifically, using a combination of 16S rRNA gene and shotgun metagenomic sequencing, we profiled changes in the taxonomic and functional composition of oral and gut communities following exposure to fluoride treatment. Our data revealed that fluoride exposures at levels commonly found in municipal water and dental products induced statistically significant changes in oral, but not stool, microbial community structure and function. In the mouse models used here, microbes in the oral community that are more typical of the gut microbiome (and likely derived from coprophagy) were also selectively depleted by fluoride treatment.

Due to its extremely widespread use in public health, it is important to understand how fluoride, even at low levels of exposure, might affect human-associated microbial communities during or after chronic use. Fluoride use in humans was first studied in the 1950s by assessing the toxicity of fluoride on different host organ systems [192, 193] and on cultures of select oral pathogens such as *Streptococcus mutans* [177-179]. More recently, two 16S sequencing studies in human populations have assessed the effect of fluoride on the dental plaque microbiome [182] (n=12) and orthodontic fixed appliances [183] (n=91). Both studies considered only the oral microbiome and, again, observed only minor, low-effect-size shifts in microbial composition longitudinally after fluoride exposure. Neither study included cross-sectional controls lacking fluoride exposure, nor were subjects at baseline free of chronic fluoride exposures from routine dental hygiene. To our knowledge, our work is thus the first study to assess the independent effects of fluoride on the oral and gut microbial communities.

Previous *in vitro* studies have shown that fluoride inhibits a wide variety of enzymes, including phosphatases, pyrophosphatases, esterases, and catalases [194, 195]. This inhibition is typically due to interactions with cationic metal cofactors. Among the best-characterized fluoride-sensitive enzymes are

those involved in glycolysis [e.g. enolase [196]] and the citric acid cycle [e.g. succinate dehydrogenase [197]]. However, the inputs into both of these processes (glucose and pyruvate) can be alternatively metabolized via the hexose monophosphate shunt or fermentation [198] if critical enzymes are inhibited by fluoride. Although enolase is the most well-characterized fluoride-sensitive enzyme, it was not depleted in our functional data due to its universal carriage in bacteria [199]. Our data show, however, that *in vivo* fluoride-associated modulation of *in vitro*-demonstrated fluoride-sensitive genes [such as succinate hydrogenase and glyoxylate reductase [200]] is detectable in metagenomic data. While 16S-based functional predictions cannot associate specific gene polymorphisms with fluoride sensitivity, future studies aimed at specific microbial molecular products or physiology may provide a higher resolution look at specific genes and activities affected by fluoride.

Our work, along with several previous, more open-ended studies in human populations [182, 183], suggest that physiological fluoride exposure levels have little effect on the established gut microbiome and even on the overall composition of the oral microbiome. An interesting open question, however, is the degree to which this conclusion might differ in developing microbial communities such as the infant gut or oral microbiota within the first few years of life [138, 201, 202]. While it is difficult to model early developing human microbial communities in mouse systems, we anticipate that future work in carefully designed models, human populations, or directly investigating microbial physiology will continue to characterize the effects of this daily environmental exposure on microbial community composition and function.

We have shown that exposures to fluoride levels found in municipal water and dental products altered oral, but not stool, microbial community structure and function in mice. Specifically, genera containing acidogenic bacteria such as *Parabacteroides*, *Bacteroides* and *Bilophila* were depleted in the mouth, and fluoride exposure was associated with depletion of genes involved in central carbon

metabolism and energy harvest. In contrast, fluoride treatment did not have a significant effect on gut community composition or function, which is consistent with administered fluoride not reaching the gut to an appreciable extent. While the specific responding taxa in humans and in mice are unlikely to be the same due to differences between human systems and mouse models, mechanisms of response and overall community structural ecological changes are likely to be shared [69]. We conclude that in our model, exposure to fluoride levels found in municipal water and dental products had selective effects on the composition of the oral microbial community, but are unlikely to have significant effects on established gut microbial communities.

Materials and Methods

Animal husbandry

Female BALB/c mice were weaned between post-natal days 18-21 and randomized into experimental cages (4 cages per treatment, n=7 per treatment) with an adjustment period of handling the animals for 1 week, each cage containing two mice. Mice were fed irradiated standard mouse chow (PicoLab Mouse Diet 20 [5058]; LabDiet, St. Louis, MO). All experiments were approved and conducted in accordance with Harvard Medical School Standing Committee on Animals and National Institutes of Health guidelines.

Sample collection and processing

Stool samples were collected at weeks 0, 4, 8, and 12. Oral samples were collected at weeks 0 and 12. Since the animals had to be euthanized to collect the oral samples (to avoid mouse skin contamination), our week-0 oral samples were collected from a separate set of mice (n=8) that were co-housed until randomization at week 0. The collected samples were stored at -80^{o}C before processing. DNA was extracted using the MP BIO FASTDNA™ SPIN Kit for Soil (MP Bio, Santa Ana, CA) according to manufacturer's instructions. For the 16S rRNA amplicon sequencing, the V4 region was amplified using the Earth Microbiome Project 16S sequencing protocol [169]. In brief, genomic DNA was subjected to 16S amplifications using primers designed incorporating the Illumina adapters and a sample barcode sequence, allowing directional sequencing covering variable region V4 (Primers: 515F [AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT ATG GTA ATT GTG TGC CAG CMG CCG CGG TAA] and 806R [GGACTACHVGGGTWTCTAAT]). PCR mixtures contained 2 µl of diluted template (5-50 ng/ul of DNA), 10 µl of HotMasterMix with the HotMaster Taq DNA Polymerase (5 Prime), and 0.5 µl of primer mix (10 µM of each primer). The cycling conditions consisted of an initial denaturation of 94°C for 3 min, followed by 32 cycles of denaturation at 94°C for 45 sec, annealing at

50 °C for 60 sec, extension at 72°C for 5 min, and a final extension at 72°C for 10 min. Amplicons were quantified using Qubit 2.0 fluorometer and Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen, Life technologies). Integrity of DNA was tested by gel electrophoresis (1% agarose gel). Quantified DNA was pooled in equimolar concentrations, size selected (375-425 bp) on the Pippin Prep (Sage Sciences, Beverly, MA) to reduce non-specific amplification products from host DNA. Sequencing was performed on the Illumina MiSeq platform (version 2) according to the manufacturer's specifications with addition of 15% PhiX, and yielded paired-end reads of 151 bp in length in each direction.

A subset of samples used for 16S sequencing was subjected to shotgun metagenomic sequencing. These comprise 6 oral (3 controls and 3 high fluoride group at the end of the study) and 11 stool samples (3 controls and 2-3 high fluoride group at the beginning and end of the study). Nextera libraries were prepared manually following the manufacturer's protocol (Nextera XT DNA Sample Prep Kit, Illumina Inc. San Diego, CA). Briefly, tagmentation of samples was performed using 1 ng of template, and PCR amplification was performed by a Bio-Rad T100 Thermocycler (Bio-Rad, Hercules, CA, USA) following manufacture's protocol. Agencourt AMPure PCR Purification System (A638801; Beckman Coulter, Brea, CA) was used to select for 300-500 bp fragments. The DNA libraries were validated with an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA) and quantified using Qubit 2.0 fluorometer (ThermoFisher Scientific, Waltham, MA). Equal volumes of normalized libraries were combined, diluted in hybridization buffer and heat denatured, according to Nextera XT protocol. Pair-end sequencing was performed using the NextSeq Mid 150 cycle (2 x 75 base pairs).

Bioinformatic analysis of 16S and metagenomic shotgun sequencing

Overlapping 16S paired-end reads were stitched together (approximately 97 bp overlap), and further processed in a data curation pipeline implemented in QIIME 1.8.0 as pick_closed_reference_otus.py [169]. In brief, this pipeline picks OTUs using a reference-based method

and then constructs an OTU table (Table S4). Taxonomy was assigned using the Greengenes (2013 version) predefined taxonomy map of reference sequence OTUs to taxonomy. A mean sequence depth of 134,046 reads per sample was obtained; samples with fewer than 50,000 filtered sequences were excluded from downstream analysis. Further microbial community analyses such as beta diversity calculation and analysis of similarities (ADONIS) between variables (i.e. treatment groups, time points, and cages) were performed using the weighted UniFrac distance measure with QIIME 1.8.0 [169]. Microbial functional modules were inferred from the 16S-based taxonomic profiles using PICRUSt [187] and functional modules were reconstructed using HUMAnN [172].

Shotgun metagenomic sequences were first adapter trimmed using cutadapt [203]. Mouse reads were removed using KneadData (http://huttenhower.sph.harvard.edu/kneaddata), which also trimmed low-quality base pairs (Phred score < 20) and filtered short reads (trimmed length <70% of original). Mouse reads were matched against the BALBc genome dated February 2017 (http://www.sanger.ac.uk/science/data/mouse-genomes-project). We performed taxonomic profiling using MetaPhlAn2 [191]. Species abundances (63 before filtering) were passed through a filter requiring each species to have at least 0.01% abundance in at least 10% of all samples, resulting in 36 species for analysis (**Table S5**). Functional profiles were generated using HUMAnN2 [172] (http://huttenhower.sph.harvard.edu/humann2). The UniRef90 database [204] was used for the translated search. UniRef90 abundances were collapsed to KEGG Orthology (KO) groups [205] for comparison with PICRUSt output by mapping through UniProt-derived annotations. KO rarefaction analysis (**Fig. S3E**) was carried out using the rarecurve function in R's vegan package (step set to 2,500) by providing per-sample KO and unmapped read counts as input.

To test for statistically significant microbial clades associated with fluoride treatment and metadata, we used the combination of LEfSe [147] for univariate and MaAsLin (Multivariate

Associations by Linear models) [15] for multivariate analyses. To find taxa enriched amongst oral and stool samples, we used LEfSe, where classes are set as biogeographical locations (oral and stool). To identify taxa and functional modules associated with fluoride treatment in each oral and stool samples, we used MaAsLin to build a multivariate linear model combining fixed and random effects on each sample type. For the oral samples, week-12 samples were used to identify taxa and functions that were perturbed by fluoride treatment group (control, low fluoride, or high fluoride). For the stool samples, fluoride's effects on taxa and functions were first tested on the combination of week-4, week-8, and week-12 samples, as these were the time windows that would allow us to observe potential treatment effects. In this model, animal age (week) and housing cage were included as covariates. We conducted a separate series of MaAsLin analyses stratifying the samples by week and including housing cage as the only covariate. Across linear models, we applied Benjamini-Hochberg multiple testing correction with a target false discovery rate (FDR) of 0.2.

Determining fluoride concentrations for use in mice

Mice in our study were treated with fluoride by inclusion in their drinking water (low- and high-fluoride groups) and through additional gavage (high-fluoride group). Drinking-water concentrations were designed to reflect human-equivalent exposures. Specifically, we prepared water with four parts per million (ppm) fluoride using sodium fluoride (solubility >99%, Product number S6776-100G, Sigma-aldrich). This is the highest FDA approved level of fluoride in municipal water in the US. Mice in the low- and high-fluoride groups drank from this solution daily throughout the experiment. This is equivalent to a 0.02 microgram/day exposure, based on an expected consumption of 5 ml of solution per day [206].

Mice in the high-fluoride group received an additional dose of 2.25 micrograms of fluoride per day. This dose was based on equivalent amounts of fluoride that might be consumed by ingestion of

fluoridated toothpastes in young children (1-3 years of age). Specifically, we assumed a 10-kg child consuming one quarter of 1 g of toothpaste per brushing session, twice per day, with a typical toothpaste fluoride concentration of 1,500 micrograms per gram. This equates to 750 micrograms per day for the 10 kg child, which is equivalent to 2.25 micrograms per day for a 30 g mouse (the expected average mass of our mice over the 12-week time course). This additional fluoride was given daily to mice in the high-fluoride group by gavage in 10 microliters of deionized water.

Measuring fluoride concentrations in the intestinal contents

We measured fluoride ion concentration in mouse stool using a fluoride ion electrode probe (Cole-Parmer Instrumental Company, Vernon Hills, IL). We calibrated the probe using an initial 1,000 ppm fluoride solution provided by the manufacturer serially diluted to 0.001 ppm in deionized water. To analyze a given stool sample, 100 mg of dry stool was dissolved in 10 ml of deionized water. Based on the calibration curve, we concluded that untreated mouse stool had a baseline fluoride ion concentration of 0.064 ± 0.032 μg / ml of dissolved stool (mean ± stdev, week-0 mice; **Fig. S1**). In comparison, stool fluoride concentrations after twelve weeks of low-fluoride treatment were not appreciably larger than untreated values at week 0 (0.036 ± 0.013 μg / ml of dissolved stool; *t*-test, one-tailed p=0.997). Fluoride concentrations in stool from the high-fluoride group were similarly unaffected (0.065 ± 0.023 μg / ml of dissolved stool; p=0.820).

To confirm that the probe was able to detect additional fluoride in stool, we mixed untreated dissolved stool with increasing concentrations of fluoride. This produced a trend similar to the standard calibration curve but flattened below 0.15 μg / ml of dissolved stool (consistent with the inability to resolve concentrations of fluoride below the baseline measurement for mouse stool). A follow-up experiment using stool subjected to bead beating (to potentially release additional intracellular fluoride) produced a similar trend.

Availability of data and materials

Data needed to evaluate the conclusions in the paper are present in the paper, Supplementary Materials, and the sequences generated in this study are publicly available (NCBI BioProject ID number PRJNA328099).

Acknowledgements

We thank the Huttenhower, Garrett and Chih-Hao Lee Labs for helpful discussions. We also thank Nicole Levesque and Jessica Lacy for copy-editing. This work was supported by Robert F. Naka Fellowship to KY, and grants from the NIH (R01HG005969), NSF (DBI-1053486) and ARO (W911NF-11-1-0473) to CH and NIH (R01CA154426, K08AI078942, R01GM09953), Burroughs Wellcome CAMS, and Danone Research Awards to WSG. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

This page intentionally left blank

CHAPTER 3

Adipose tissue microbiome

This chapter represents unpublished work.

Abstract

Gut microbial translocation to the visceral fat has been suggested as a source of tissue inflammation, but there has not yet been a study that comprehensively tests this hypothesis in humans and model systems. We used sterile techniques to collect intra-abdominal (mesenteric and omental) and peripheral (subcutaneous) adipose tissues and lymph-nodes, along with paired intestinal contents (small and large intestinal contents) from mice, rhesus macaques, and humans. Bacterial DNA in these samples was assessed by 16S rRNA gene sequencing. To address potential contamination in these low biomass samples, we included negative controls for extraction and PCR, as well as developing a novel bioinformatic protocol for contaminant sequence depletion using taxon prevalence and host-uniqueness measures. We identified a series of potential tissue-resident microbes in mouse and macaque in adipose tissues, lymph nodes, and liver including *Prevotella* and Helicobacter species, and *Allobaculum*, *Prevotella* and Clostridiales species in macaques and amice, respectively. These bore a striking resemblance to gut microbial profiles, particularly in the anatomically adjacent mesenteric fat. However, no non-contaminant reads were identified in human adipose tissues, suggesting that differences either in protocol (e.g. anesthesia versus sacrifice) or in biology may induce distinct tissue-resident microbial DNA in humans versus model systems. Further studies are needed to clarify the viability of tissue-resident microbes in animal models and the degree to which physiological or technical differences may drive the lack of adipose microbes in humans.

Introduction

Viable bacteria are often culturable [105] and their DNA fragments are detectable from sites of infection and from associated lymph nodes [17, 20]. Gut microbial translocation to the visceral fat has been suggested as a source of tissue inflammation; however, there has not been a study to systematically test this hypothesis. We hypothesized that there might also be a tissue microbiome associated with intra-abdominal adipose tissues, and a potential method of its establishment could be gut bacterial translocation.

Visceral adipose tissue microbiota are of particular interest for two reasons. First, adipose tissues are known to be closely involved in inflammation, as adipocytes themselves express microbial pattern recognition receptors such as TLR-4 and respond readily to infection by secreting antimicrobial peptides [38] and cytokines such as TNF-a, IL-6 and MCP-1 [207]. Secondly, rodent and human studies suggest that inflammation is more prominent in intra-abdominal fat compared to peripheral fat depots [40]. Given the close proximity of mesenteric-visceral fat to the gut where the gut microbiota are located, it is possible that there is a link between visceral fat inflammation and gut microbiota translocation.

While culture-based methods can provide a presence–absence readout and indicate whether bacteria are viable, culturing is low-throughput and many gut microbiota are unculturable or must be cultured on specialized media, making this method particularly unsuitable to asking questions about an unknown and potentially complex community of bacteria. Quantitative PCR-based methods can provide presence-absence readout and can be compared quantitatively, but this method also is low throughput and does not avoid the same contamination problems that are present in any DNA-based method. Metagenomic shotgun sequencing can provide strain-level taxonomic and community-level functional information; however, when dealing with tissues where the majority of the sequencing reads will come from host DNA, it is often cost prohibitive to obtain sufficient reads from microbial data for further

analysis. On the other hand, 16S rRNA gene amplicon based method can generate a large amount of bacterial taxonomic data at a relatively affordable cost, can detect live or dead bacterial DNA fragments, is semi-quantitative, and if given appropriate blank controls [115], one could possibly distinguish contaminant reads from real signals.

Since we do not know whether gut-to-tissue bacterial translocation occurs at a detectable rate with our 16S based method, we used high-fat diet fed mice to increase our chances of detecting gut bacterial translocation, as these mice have previously been shown to have increased gut permeability [208, 209]. Although mice gut microbiota resemble that of the human gut microbiota, the scarcity of colon contents (as colon contents are pelleted in mice and can only be found 1-2 pellets at a time) do not make mice the most ideal animal model to test gut bacterial translocation. On the other hand, rhesus macaques have similar colonic content and consistency as humans. We therefore also collected paired intra-abdominal adipose tissues and gut from rhesus macaques, as well as adipose tissues (omental and subcutaneous fat) from humans.

In this study, we investigated whether microbial DNA fragments are detectable in intra-abdominal adipose tissues, and if so whether it is possibly due to gut bacterial translocation. By taking into account the abundance, prevalence and host-uniqueness (for each taxon) data, we were able to distinguish microbial sequencing reads from those that are likely of contamination. Our data show that the majority of bacterial reads identified in human adipose tissues to be contaminants. There were, however, bacterial reads identified in fat tissues in macaques and mice resembling those of bacterial taxa from the gut. Although the macaque and mouse data could possibly support the notion of gut bacterial translocation, the discrepancy with the human data, and having significant proportion of sequencing reads in adipose tissues made up of contaminant reads, further studies are needed to clarify whether the bacterial reads commonly found in fat and gut of macaques indeed occurred *in vivo* or post-mortem.

Results

Comprehensive gut and tissue microbiota biogepgraphy sampling from humans, macaques and mice

We studied the presence of tissue microbiota and their possible mechanism as gut microbial translation in three mammalian systems including humans, macaques, and mice. We included obese non-diabetic (n = 25), and obese diabetic (n – 23) subjects from 18 – 65 years of age with a well-established diagnosis of type-two-diabetes with no concurrent infection or chronic disease. Omental adipose tissue and abdominal adipose tissue were collected during bariatric surgery. The macaques ranged from 12-18 years old, and weigh between 7.6 to 12.4 kg. During a routine autopsy, mesenteric, omental, and subcutaneous fat samples were collected from all macaques (n = 26), and additional mucosal and luminal gut samples from jejunum, ileum, proximal, transverse, and distal colon were collected from a subset of animals (n = 15). For the mice samples, mesenteric, epididymal, and subcutaneous fat samples (proximal to the right hind limb) and small and large intestinal and stool samples were collected from wild-type mice who were fed the normal chow and high-fat diet weighing on average 32g (WT: wild-type, n= 7) and 46g (HFD: high-fat diet, n = 7).

The samples were subjected for microbiome profiling using 16S rRNA gene sequencing on the illumina HiSeq platform (version 2) with 150 bp paired-end reads. After quality filtering and assembling overlapping paired-end reads, more than 21 million sequences were retained (mean of 29,121 sequences per sample), providing the comprehensive assessment of tissue and gut microbiota biogeography in three mammalian systems.

Distinguishing host-specific taxa from contaminants

Although all reagents and methods used to process tissue samples were considered sterile, we expected from previous studies that there may be some degree of contamination detectable in our samples, especially in low biomass samples such as adipose tissue. We therefore first assessed the degree of contamination and identified bacterial OTUs introduced during the sample processing steps.

Notably, there was significant overlap of bacterial oligotypes between samples and controls. Simply removing OTUs present in negative controls from our sample data may not be appropriate as this would result in the removal of the majority of OTUs in our dataset (**Figure 3-1**). For example, although the number of reads was low (22 reads), an OTU belonging to *Helicobacter macacae* was present in macaque samples as well as in a PCR negative control sample, where it is most likely that *H. macacae* reads from rhesus macaque sample(s) "spilled over" to the control samples.

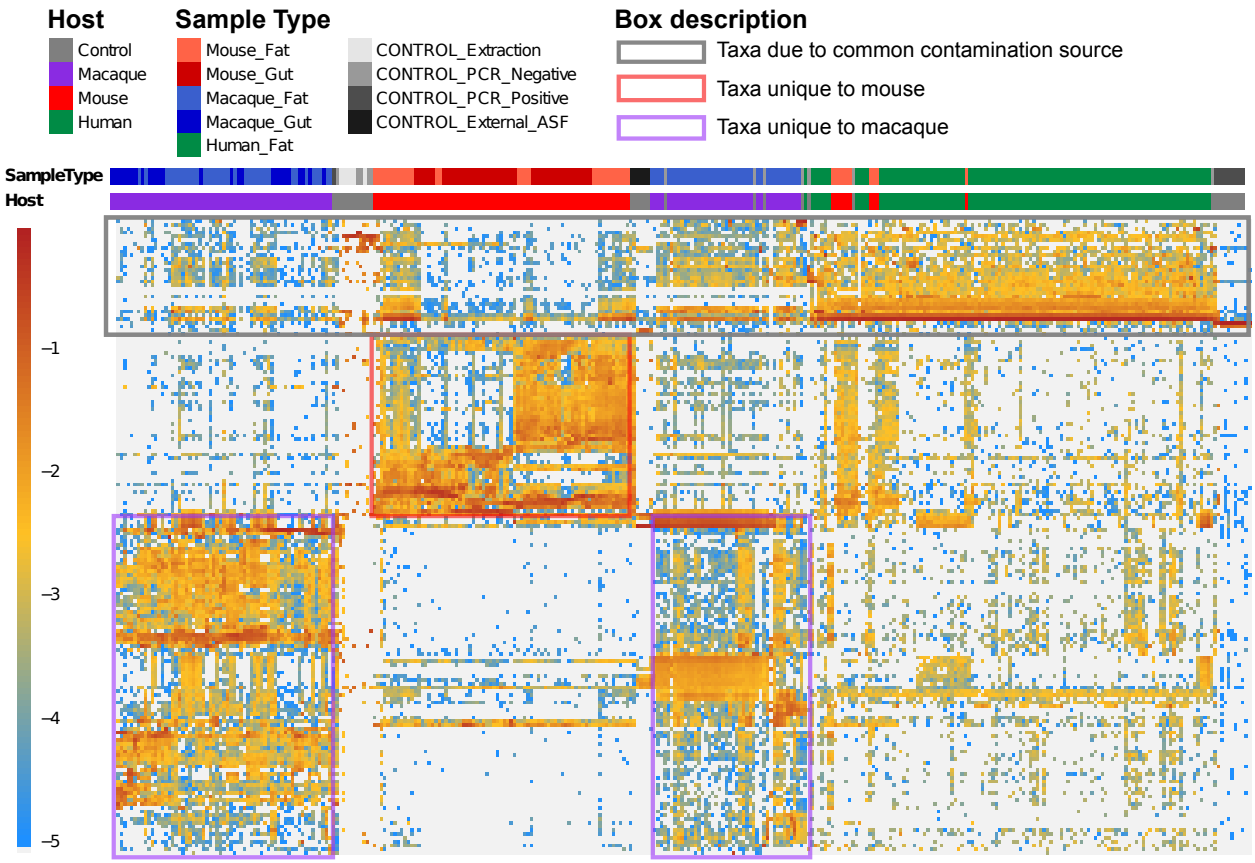


Figure 3-1: Bacterial relative abundance illustrates two types of contamination: common-source and cross contamination amongst samples derived from humans, mice and macaques A heatmap of bacterial OTUs (n = 162; >0.001% relative abundance) across all gut and adipose tissue samples were clustered by Bray-Curtis dissimilarity metric on samples (x-axis) and oligotypes (y-axis). Columns are colored (at the top) by the "Host" type (human, mouse, macaque, control), and "Sample type" (fat, and gut). Clusters of taxa are boxed to indicate host-specific taxa (grey = common source contamination; red = unique to mouse; purple = unique to macaque).

We next considered simply using the abundance or prevalence of each OTU to determine host-specificity. However, this was also inappropriate as there were multiple bacterial OTUs that were present across many samples (high prevalence) but were low in abundance. For example, *Bacteroidales 24.7_5957* was highly prevalent in normal-chow fed mice, but was also found in the majority of samples across different host types (high prevalence) (**Figure S3-2A**). Similarly, another *Bacteroidales 24.7_9408* oligotype which is highly prevalent and moderately abundant in macaques (as expected as it is unique to macaques) was also moderately prevalent in mice and human samples likely due to a spillover event (**Figure S3-2B**).

It has been shown that different host-species (e.g. macaque vs. mice) are colonized by distinct bacterial species/strains (e.g. *Helicobacter macaque* vs. *Helicobacter hepaticus*, respectively). Our data indeed supported this finding, where the overlap of the bacterial oligotypes between macaque and mouse is a rare event, occurring in 1 out of 341 oligotypes from high-biomass samples (stool of mice and macaques) (**Figure S3-3**). Therefore, we decided to make an assumption that since overlap of bacterial OTUs in high-biomass sample such as the gut is limited, an overlap of OTUs in low-biomass samples across different host-species would also be limited. Although there might be an overlap of host-specific bacterial OTUs in adipose tissue, we assumed the number of bacteria left out through this filtering scheme is much fewer than the number of true contaminants removed via this process.

Building upon this assumption, we therefore developed an algorithm (see more detail in the methods) to distinguish those bacterial OTUs that are likely to be present biologically ("real signal") from those that are likely to be introduced as contaminants by taking into account both prevalence and abundance. To do this, we first assigned a value (k-ratio) for each oligotype based on the ratio between the highest and second highest relative abundance values from each host-species. By using the ratio, we grouped 308 oligotypes into the following three categories: 1) "real" – likely biologically real signal (n =

151), 2) “common source” contamination – contaminant oligotype from common sources such as extraction buffer (n = 68), 3) “spillover” – oligotypes that can not be assigned to one host-species (e.g. oligotypes abundant in more than one host-species) (n = 68), or 4) “rare oligotypes” – rare oligotypes with prevalence of < 25% in samples (see methods for more detail) (n = 21) (**supplemental table 3-1**).

An abundance-based novel algorithm effectively identifies contaminate OTUs

Amongst two types of contamination events, we found that the majority (96% on average) of the contamination was derived from common source contamination and only a small portion (4%) derived from cross-contamination (**Figure 3-2A**). However, it is important to note that the way we defined common source contamination is prone to those OTUs falsely being called as “common source.” Some of the common source contaminants may actually be truly present in samples from two different host-species. Also, those truly biologically present OTUs that were difficult to be assigned unique hosts due to relative abundance difference not surpassing our threshold were also called as “common-source” contaminates.

After the removal of both types of contamination, the majority of human reads were removed and only one OTU (unclassified bacteria) remained in the human fat tissue composing less than 1% of the total abundance (**Figure 3-2A**). Expectedly, the gut contents (small and large intestinal contents) showed high proportion of the remaining OTUs (91% large intestinal contents; 79% small intestinal contents) compared to adipose tissue samples (46%), consistent with the fact that the gut contents having high microbial biomass *in vivo* (**Figure 3-2B**).

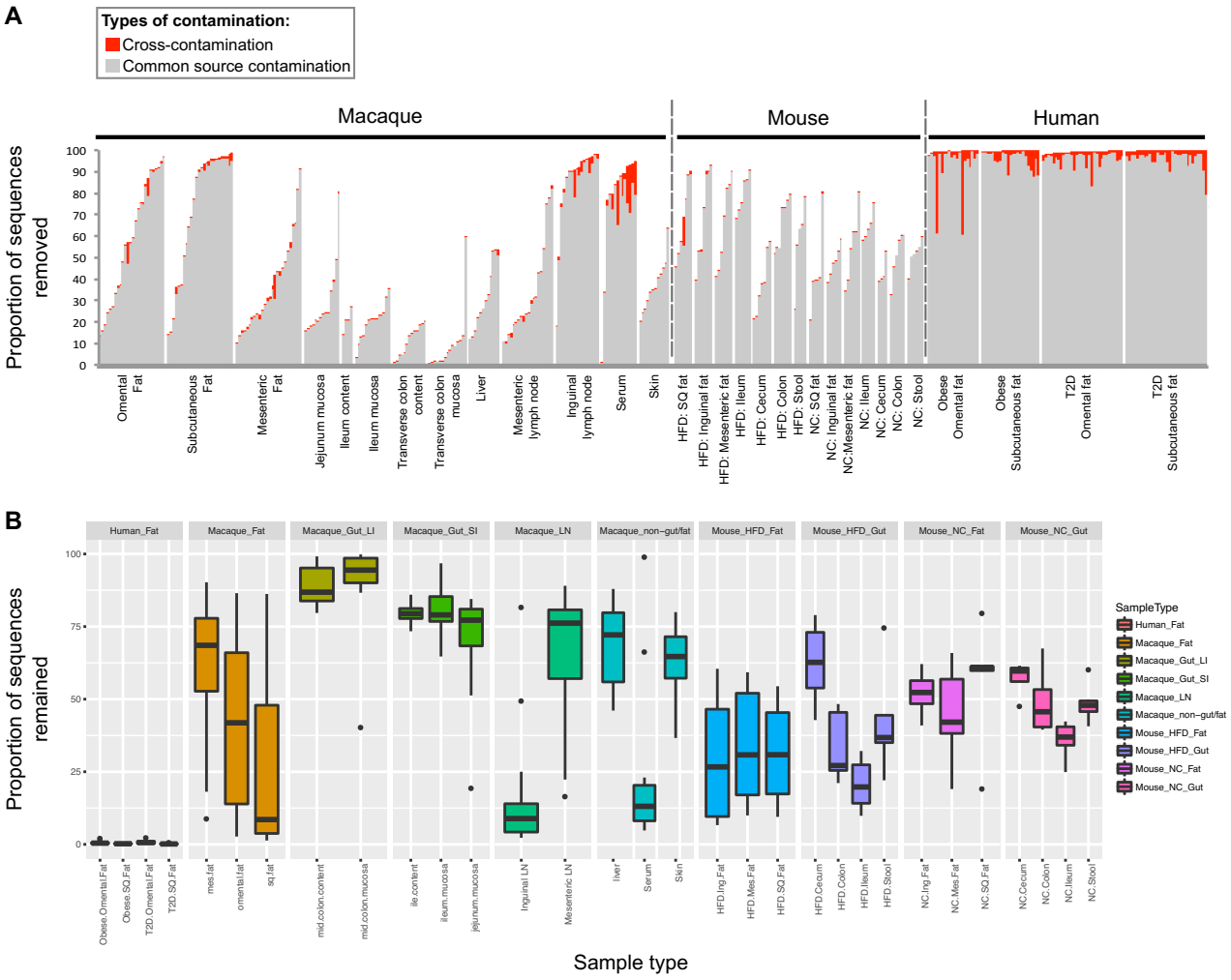


Figure 3-2: Proportion of contaminate-reads across different sample-types in macaque, mouse and humans A) Proportion of sequencing reads (y-axis) derived from common source contaminate (grey), cross-contamination (red), and likely present biologically (white). Each vertical bar represents a sample (x-axis). B) Human (black), macaque fat (brown) macaque gut large (dark green), and small intestine (light green), macaque non-gut/fat tissues (teal), mouse fat from high-fat-diet (HFD) (blue), mouse gut from HFD (purple), mouse fat from normal-chow (NC) (purple), mouse gut from NC (pink)

Adipose tissue microbiota resemble gut microbial profiles in macaques and mice

After the removal of contaminates bacterial OTUs, considerable number of bacterial OTUs remained in adipose tissues from macaques (128 OTUs) and mice (66 OTUs). The most abundant bacterial OTUs in the macaque adipose tissues were those that are also abundant in the gut including those OTUs belonging to *Prevotella*, *Helicobacter*, and Desulfovibrionaceae taxa (**Figure 3-3B**). Similarly in mice, bacterial OTUs associated with adipose tissues resembled that of the gut, including bacterial taxa belonging to *Prevotella*, *Bacteroides*, *Allobaculum*, *Oscillospira*, and *Succinivibrio*. There were no unique OTUs that are only associated with adipose tissues in macaques nor mice, and that all of the OTUs found in tissues were found in at least one of the GI segments.

Contrary to macaques and mice, our contamination removal steps resulted in the removal of all bacterial OTUs from human fat tissues (omental and abdominal adipose tissues). The most abundant (average relative abundance = 85%) contaminating bacteria genera in human adipose tissues belonged to genus *Ralstonia* (**Figure 3-3A**). This Ralstonia OTU was also found at moderate abundance in adipose tissues from macaques (6%) and mice (29%) supporting a common-source contamination. In addition, substantial portions of the macaque subcutaneous and omental fat consisted of three mice taxa; *Muscispillum schaedleri* (8%), *Turicibacter* (11%), and *Parabacteroides* (23%) (Figure 3-3A). Notably these three bacterial taxa were shared with another mouse gut microbiota in which the DNA libraries were prepared and samples were sequenced together on the same Illumina HiSeq lane [210].

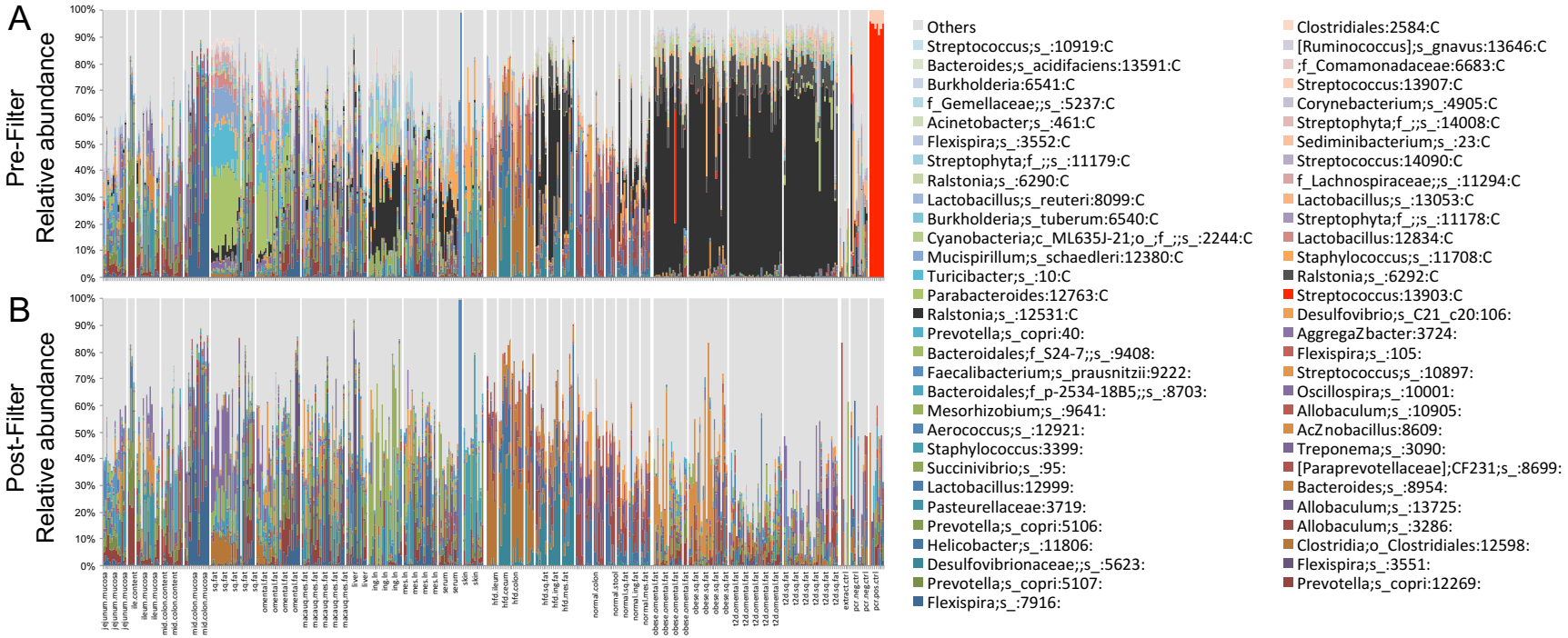


Figure 3-3: Proportion of reads remained after removing contamination reads for each sample type across humans, macaques and mice Oligotype relative abundance before (A) and after (B) contamination-reads removal.

contamination of environment (e.g. laboratory bench surfaces) and/or equipment (e.g. tubes, scalpels, and forceps).

The bacterial taxa identified in the current study mirror those found in other tissue microbiome studies. Urbaniak et al.,[113] found *Acinetobacter* and Comamonadaceae (Proteobacteria phylum) to be the predominant bacterial taxa in breast tissues, followed by Enterobacteriaceae, *Bacillus*, *Pseudomondas, Staphylococcus,* and *Propionibacterium*. The former two bacterial taxa were enriched in all of our samples as well as in the lung [110] and placenta studies [109]. It is worth mentioning that these taxa along with Ralstonia (the most abundant contaminant taxa in our human adipose samples) are also prevalent in extraction reagents [114, 115], raising a question of the origin of these bacteria.

It is important to note that unlike adipose tissues, the breast tissues are unique and might in fact host microbial colonization however. Breast tissue is anatomically unique in that it is surrounded by fascia and is composed of structural lobules lined with epithelial cells that have a direct link to the external body surface. Considering these features, the breast tissue may be one tissue that may actually have native bacterial colonization within the tissue, likely colonizing the lining epithelial cells making up the lobules rather than deeper within connective tissue. In fact the referenced study was able to culture eight species of bacteria including *Bacillus* sp., *Micrococcus luteus*, *Propionibacterium acnes*, *Propionibacterium granulosum*, *Staphylococcus* sp., *Staphylococcus saprophyticus*, *Streptococcus oralis*, and *Streptococcus agalactiae* from 43 out of the 81 subjects, with amounts ranging from 75 to 2,000 CFU/gram of tissue [113]. The exact location of bacterial colonization(s) within the breast tissue would be an intriguing topic to be explored.

Regardless of prior tissue microbiome results, it is clear that sequencing-based studies, particularly when dealing with low-biomass samples, are prone to an over-amplification of contaminant DNA [114-120]. This is not unique to bacterial DNA, but also fungal [121] and viral [122] DNA have

been detected in various laboratory reagents. Furthermore, the contamination in high-throughput sequencing studies also go beyond the microbiome research as well; cross-contamination has been found in whole genomes of many organisms whose genomes are deposited in the NCBI database. For example, components of *Neisseria gonorrhoeae* genome have been found in the domestic cow genome (*Bos Taurus*) [123], portions of human genomes have been found in deposited reference sequences for *Caenorhabditis elegans* (*C. elegans*), and Xenopus genomes have been found in Zebra fish genomes [124].

Our results indicate that the there are potential tissue microbiota in adipose tissues in macaques and mice models. Our study also indicates however the difficulty of detecting in vivo microbial microbial signals from low-biomass using 16S amplicon sequencing. Along with our current study-developing algorithm to distinguish contaminate OTUs from likely potentially real microbial OTUs, other have developed ways to handle contamination, an absence of relevant microbes is difficult to be proven as with any negative experimental results. This has been the case for other tissue microbiome studies [119], and can be addressed by the complementary technologies discussed above. In the case of a potential adipose tissue resident microbiome, the health consequences of local immune sensing, signaling, and inflammation would be sufficiently important to warrant further investigation into human subjects and a clearer comparison with differences in typical microbiome model systems.

Experimental Procedures

Animals and human subjects and sample collection

Macaques All macaque gut microbiome and tissues used in the current study came from the macaques used in the intestinal biogeography study described elsewhere [47]. Approximately 1 g of adipose tissues collected from different compartments including subcutaneous fat at the level of umbilicus; mesenteric fat adjacent to the cecal-colic junction and omental fat (center of the momentum) were resected and stored at -80^{o}C before processing. Similarly, a small sections of the liver from the quadrate lobe, and inguinal and mesenteric lymph nodes were resected prior to exposing the gastrointestinal tract content on the autopsy table to prevent samples and instruments from being cross contaminated by the GI bacteria.

Mice WT C57BL/6J female mice Beginning at 6 weeks of age was fed either a chow diet (Research Diet Inc, 10% kcal fat, 70% carbohydrate) or high-fat diet (HFD) (Research Diets Inc., 60% kcal fat, 20% carbohydrate) for 12 weeks. All mice had free access to food and water and were kept in temperature-controlled facilities at 23°C on a 12-hour light/dark cycle for 8 weeks. All protocols for animal use and euthanasia were approved by the Animal Care Committee of the Joslin Diabetes Center and are in accordance with National Institutes of Health (NIH) guidelines.

Humans Adipose tissues used in the current study were from a previously described study [211]. Briefly, adipose tissues were collected in pairs from omental and subcutaneous fat (percutaneous biopsies of abdominal subcutaneous adipose tissue) during bariatric surgery, and the collected tissues were placed in sterile falcon tubes, transported on ice and stored in -80^{o}C freezers until further processing. The study was approved by the Partners health the Joslin Diabetes Center Human Subject Committee approved the experimental protocol, and informed consent was obtained from all participants.

Sample processing and 16S rRNA gene amplicon sequencing

DNA was extracted using the MP BIO FASTDNA™ SPIN Kit for Soil (MP Bio, Santa Ana, CA) according to manufacturer's instructions. For the 16S rRNA amplicon sequencing, the V4 region was amplified using the Earth Microbiome Project 16S sequencing protocol [169]. In brief, genomic DNA was subjected to 16S amplifications using primers designed incorporating the Illumina adapters and a sample barcode sequence, allowing directional sequencing covering variable region V4. Negative blank control samples were included during the DNA extraction, and 16S PCR amplification In addition to extracting the DNA from tissues, negative controls were included before the extraction and PCR and well as a positive control of a DNA from known isolates such as *Streptococcus pneumonia* in order to access the degree of potential bacterial contamination introduced during these sample processing steps.

For the tissue samples, since we expected these samples to have low bacterial biomass, PCR mixtures contained 10 µl of diluted template (5-50 ng/ul of DNA) vs. 2 µl for all the gut-derived samples (these samples are considered high biomass samples), 10 µl of HotMasterMix with the HotMaster Taq DNA Polymerase (5 Prime), and 0.5 µl of primer mix (10 µM of each primer). The cycling conditions consisted of an initial denaturation of 94°C for 3 min, followed by 32 cycles of denaturation at 94°C for 45 sec, annealing at 50 °C for 60 sec, extension at 72°C for 5 min, and a final extension at 72°C for 10 min. Amplicons were quantified using Qubit 2.0 fluorometer and Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen, Life technologies). Integrity of DNA was tested by gel electrophoresis (1% agarose gel). Quantified DNA was pooled in equimolar concentrations, size selected (375-425 bp) on the Pippin Prep (Sage Sciences, Beverly, MA) to reduce non-specific amplification products from host DNA. Sequencing was performed on the Illumina HiSeq platform according to the manufacturer's specifications with addition of 10% PhiX, and yielded paired-end reads of 151 bp in length in each

direction, except the macaque gut samples, which were previously sequenced as described in chapter 2 of the dissertation.

Bioinformatics and statistical analyses

Overlapping 16S paired-end reads were stitched together (approximately 97 bp overlap) using the QIIME[45]. The stitched reads were processed through the MED pipeline described to derived oligoptypes (hereafter referred to as "OUT") and abundance table. Each representative oligotype sequences were then taxonomically assigned using the Greengenes (2013 version) predefined taxonomy map of reference sequence OTUs to taxonomy. We refer unique oligotypes as OTUs in this chapter.

Removal of potential contaminate reads

We developed a simple algorithm to distinguish those bacterial OTUs that are likely to be present biologically ("real signal") from those that are likely to be introduced as contaminate by assigning each oligotype into two categories; 1) common-source contaminants, or 2) cross-contaminate. By doing so, the leftover bacterial taxa should theoretically were present biologically within samples *in vivo*. The algorithm is as follows:

For each oligotype (n = 308):

Step 1: max (mean[#] (tissue types[❖])) per host species -> 6 (3 controls, 3 host) values per OTU

[❖]tissue types include different controls (PCR + and – and extraction control)

[#]mean of samples (animals/subjects)

Step 2: take top 2 values from #1 and calculate the ratio (k)

Step 3: assign each OTU into the following three categories based on the K-ratio:

Category 1: Tissue-specific OTU (REAL)

If $k_o > 3$, and X_h is from one of the host species = likely "tissue/host specific" -> keep.

Category 2: Contaminate OTU (CONTAMINATE):

If $k_o > 3$, and X_h is from one of the int/ext ctrls = likely contaminated from controls -> remove.

Category 3: Unknown OTU (COMMON SOURCE):

If $k_o < 3$, No tissue-specific, unknown= "contaminate" from a "common source" -> remove.

The following two exceptional rules were added to further retain "contaminate" oligotype that are to retain mice specific oligotypes that were removed Since two groups of mice fed different diet were treated as separate groups due to the fact that their microbiota composition differed dramatically, but it's likely that these two mice groups share significant number of bacterial oligotype, We also made an exception since

When both max & 2^{nd} max are from:

1. controls (i.e. *E.coli*), look for the highest Y^t from host.
2. the same host-species (i.e. mouse; *Helicobacter*), look for the next highest Y^t.

By doing so, the k_0 ratio allowed each OTU to be assigned to either 1) tissue-specific, 2) contaminate, and 3) spill-over OTUs.

This page intentionally left blank

CHAPTER 4

Conclusions and Future Directions

Conclusion

In this thesis, we first described our discovery that the stool microbiota is a fair proxy for the composition of the intra-intestinal mucosa and luminal gut microbiota, particularly those of the distal gut. We also found that intestinal oxygen content appears to be a strong factor in determining the composition of bacterial taxonomic communities in the mucosal-luminal environment. We then found that the level of fluoride typically added to drinking water and dental products may affect the oral microbiota. Finally, we discovered that higher bacterial DNA content can be detected in the visceral mesenteric fat and lymph nodes compared to their paired peripheral counterparts (subcutaneous fat and inguinal lymph nodes) in macaques, possibly indicating gut bacterial translocation from the gut to the visceral compartments. Exploring microbial composition through high-throughput sequencing and computational big data analytics enabled us to provide new insights within the few-centuries-old study of host-microbe interaction. There are, however, certain new questions that remain to be answered as a result of this work. These questions include; whether the gut-intra-intestinal microbial distribution holds true in humans, whether fluoride may affect the developing infant gut microbiome, and why there is a discrepancy between humans, macaques, and mice in terms of ability to detect microbial sequences in adipose tissues.

Limitations and Future Directions

As next steps, I am particularly interested in finding out which microbial taxa are mucosally enriched in the human gut. To do this, we could systematically collect biogeographic samples from healthy individuals who are organ donors or who are undergoing abdominal surgery. These approaches are also likely to be limited by the disease state of the host (e.g. recent death, or disease state necessitating the intestinal surgery such as trauma, inflammatory bowel disease, or obesity necessitating

gastric bypass). The samples are also likely to have extensive inter-subject variation depending on the diet, age, and medical history of the host. In addition, collection of samples from a sufficient number of subjects will likely take a significant amount of time and financial cost.

An interesting open question our study did not address in regards to the fluoride effects we observed in our mice is how our conclusions might differ in developing microbial communities such as the infant gut or oral microbiota within the first few years of life; as is true for the host body systems, the effects on the gut microbiota during this early stage have been shown to have lasting effects over an individual's life [138, 202, 212]. While it is difficult to model early developing human microbial communities in mouse systems, we anticipate that future work in carefully designed models, work in human populations, or work that directly investigates microbial physiology will continue to characterize the effects of this daily environmental exposure on microbial community composition and function.

With regard to our adipose microbiome investigations, further analysis should examine our hypothesis on the existence in intra-abdominal adipose tissues of resident microbial communities that originated from the gut. Such efforts could include using methods other than 16S rRNA sequencing, such as culturing of bacteria from adipose tissue homogenates, detection of bacteria *in situ* in tissues via imaging modalities such as *in situ* hybridization, and detection and quantification of bacterial taxa using fluorescence-activated cell sorting (FACS). Similarly to the 16S rRNA sequencing method, all of these methods require strategies to control contamination from non-host associated microbiota, which may involve use of microbe-free reagents and environment. As Salter et al. suggested in his paper, making a catalogue of known bacteria that are commonly present in laboratory reagents and the environment is urgently needed to permit possible subtraction of potential contaminants from the samples.

Long-term perspective

The longer-term scientific interest is to understand bacteria-host, virus-host, and virus-bacteria interactions within the context of the human microbiome. More specifically, I am interested in expanding the publically available viral genome reference by developing an efficient viral enrichment protocol to be applied to primary biological samples (e.g., stool, saliva). The timing is right to make significant progress in this line of work, as constant improvements are being made to generate longer fragments of genomes. For example, PacBio Sequel released in 2014 can sequence up to an average of 10 kb DNA fragments, in contrast with the 75-150 bp fragments by Illumina sequencing platforms. By having longer read sequences, we will no longer need to rely on the error-prone process of computational stitching of short reads. This stitching is particularly error-prone for those sequences that are without any reference genome. As the majority (80-90%) of the gut viromes currently do not have a reference genome, virome investigations will benefit greatly from longer fragment sequencing.

The virome in particular is interesting for understanding my clinical areas of interest of diabetes and obesity. There have been previous efforts towards identifying viral etiologies of type 2 diabetes mellitus (T2DM). Certain viruses in both mice and humans have been identified that appear to be causal of obesity and T2DM in animal models (e.g. adenovirus serotype 36); however, these studies were limited to specific individual viruses, and the broader context of the entire virome has remained unexplored. Next-generation sequencing will permit casting a wider net than the study of single viruses taken out of the context of the virome. Indeed, this broader context appears to be highly significant in other diseases such as inflammatory bowel disease (IBD) and antibiotic resistance states, where other studies have revealed a large number of previously unclassified viral sequences that are associated with these conditions. For example, these studies have revealed chronic viruses that may be considered commensals; the penetrance of overt disease in virus-infected persons is low in some cases; and viruses

may alter disease susceptibility via significant physiologic effects on the host independent of their role as pathogens. Given the advances in sequencing and the bioinformatic tools for its interpretation, it is timely to consider the role of the virome in obesity and T2DM.

Another fascinating area to study is the potential role that endogenous retroviruses may play in health and disease. It is estimated that 6-10% of the human genome is composed of human endogenous retroviruses (HERVs), which are identified in whole genome sequencing but often ignored as "junk" DNA. Although it had been believed that for the most part these HERVs are mostly silent and not involved in cellular functions, several recent studies have suggested that HERVs are expressed during different physiological conditions and are potentially an active part of both normal and abnormal physiology [213, 214]. For example, emerging evidence suggests a potential role of HERVs in triggering the onset of autoimmune diseases, such as rheumatoid arthritis and systemic lupus erythematosus [215]. Understanding HERVs may provide clues to some of the diseases that are yet to be fully understood.

The gut microbiota-phage interaction is another area that is under-studied. Despite the virome containing one of the most abundant and fastest-mutating genetic elements on the planet [216], the technical difficulties at all stages of virome characterization (e.g., viral isolation techniques, viral enrichment methods and viral reference genomes are limited) have severely limited our understanding of the viral components of the human gut microbiome. Similarly, these technical limitations apply to another important component of the gut microbiome: the gut fungal microbiome [217] [218]. To illustrate the limited amount of viral and fungal genome references, there are currently 105,374 bacterial genomes, 7,435 viral genomes, and 2,519fungal genomes with unique BioSample IDs deposited in the NCBI public database (accessed on 8/25/17). As with bacteria, viruses and fungi are also likely to play a

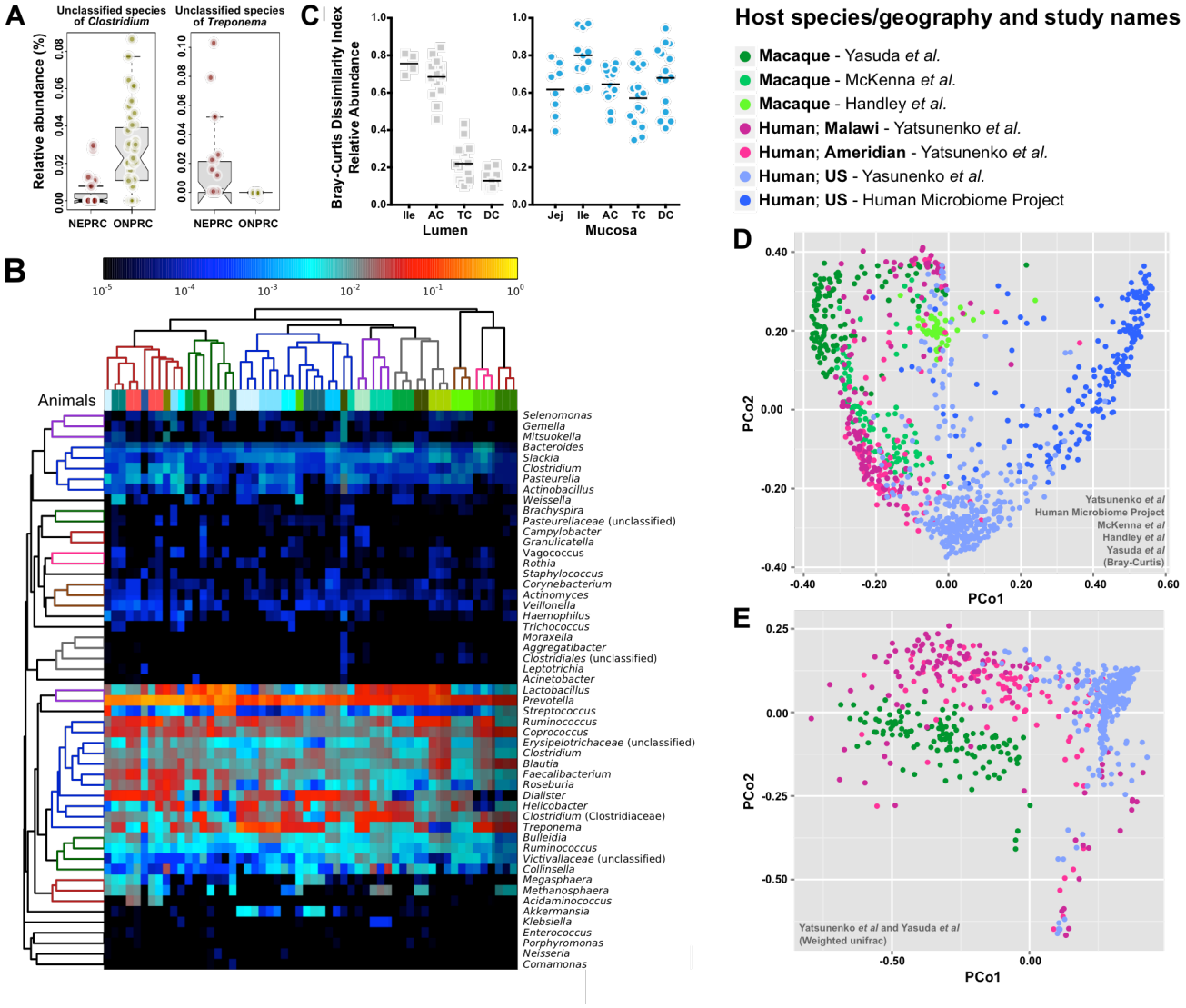
critical role in shaping host health, and understanding these microbiota will be essential to developing a full picture of all of the microbial elements at play in human health and disease.

Our discovery of the high correlation between stool and intra-colonic microbiota is remarkably relevant to the fields of both basic microbiological science and translational-clinical medicine, as the ecological variations within the colon may not greatly impact the gut microbiota composition during the colonic-stool transit. This knowledge of being able to use stool as a surrogate to understand the "gut microbiota" composition eliminates logistical hurdles associated with the collection of intra-gut samples and is extremely valuable, especially moving forward with the idea of developing stool-microbiota-based diagnostics. We now understand not only the microbial taxonomic spatial distribution within the gut but also that the oxygen content seems to be a strong factor in determining mucosal and lumenal microbiota composition. This knowledge is immensely relevant when designing small molecules to target specific microbial communities or probiotics that would effectively colonize specific areas of the gut. Our discovery of fluoride's ability to selectively deplete oral acidogenic bacteria raises questions as to how other chemicals that are also widely added to various products we use every day affect the microbiome, such as residuals from chlorine used to clean water or various chemicals included in pesticides applied to crops. Lastly, although our investigations detecting microbial fragments in intra-abdominal tissues that are of gut origin were inconsistent among humans, macaques and mice, if gut microbial translocation occurs, understanding its role in tissue inflammation and metabolism in the context of obesity and type 2 diabetes as well as other diseases that are associated with inflammation may be more deeply explored.

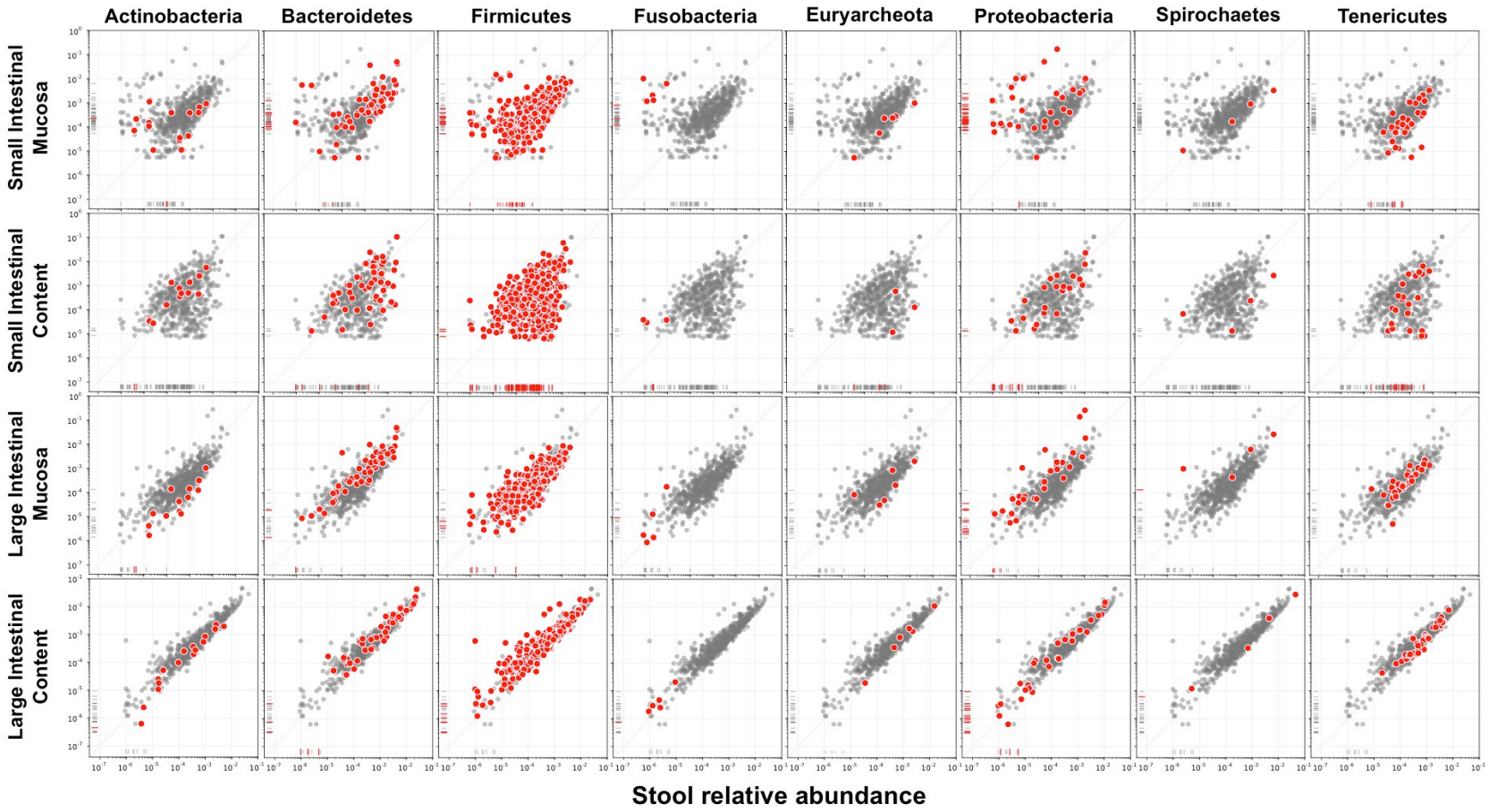
This page intentionally left blank

APPENDIX 1

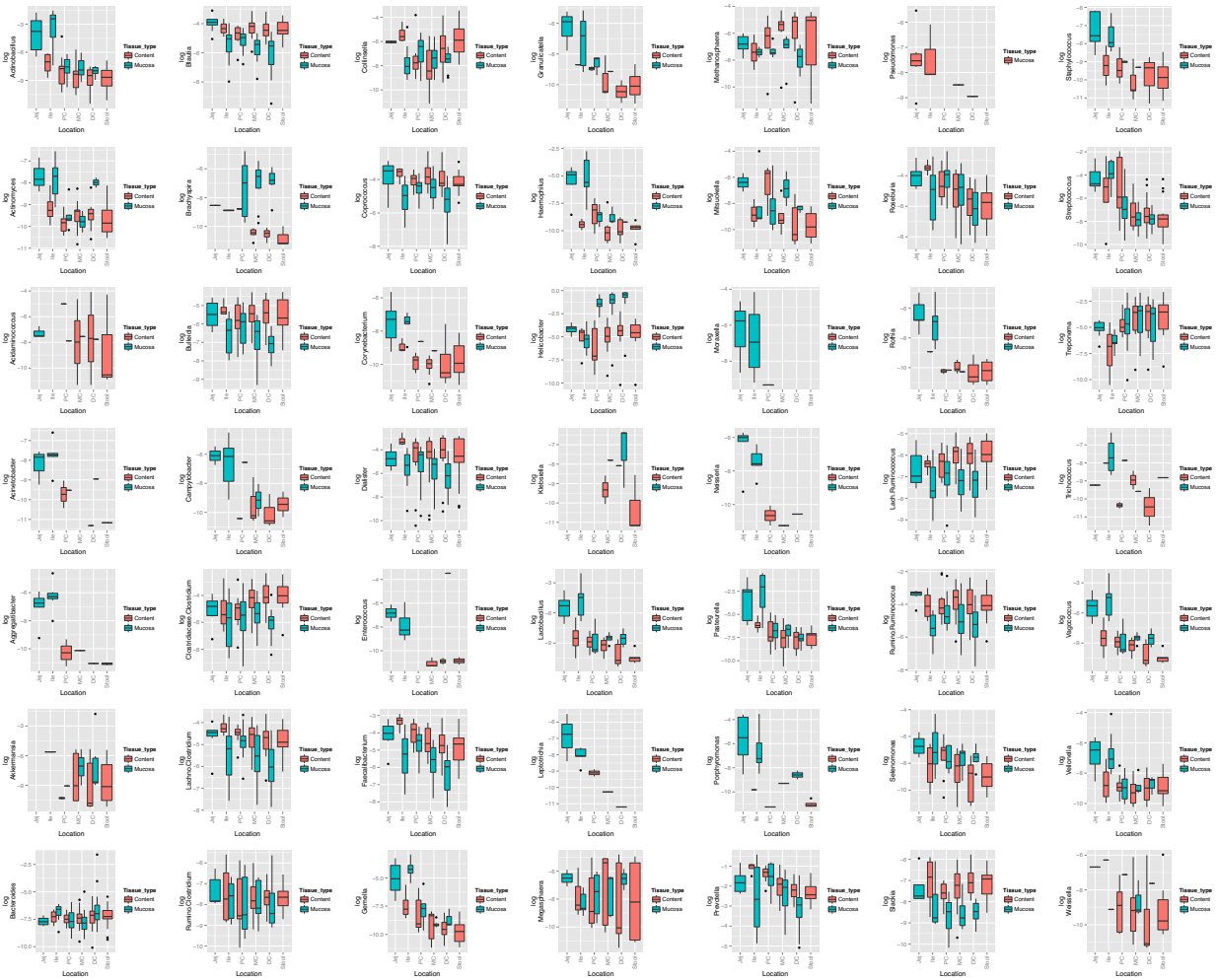
Supplementary Materials


A
Unclassified species of Clostridium
Unclassified species of Treponema
Relative abundance (%)
NEPRC ONPRC
NEPRC ONPRC
C
Bray-Curtis Dissimilarity Index
Relative Abundance
Ile AC TC DC
Lumen
Jej Ile AC TC DC
Mucosa
B
Animals
Selenomonas
Gemella
Mitsuokella
Bacteroides
Slackia
Clostridium
Pasteurella
Actinobacillus
Weissella
Brachyspira
Pasteurellaceae (unclassified)
Campylobacter
Granulicatella
Vagococcus
Rothia
Staphylococcus
Corynebacterium
Actinomyces
Veillonella
Haemophilus
Trichococcus
Moraxella
Aggregatibacter
Clostridiales (unclassified)
Leptotrichia
Acinetobacter
Lactobacillus
Prevotella
Streptococcus
Ruminococcus
Coprococcus
Erysipelotrichaceae (unclassified)
Clostridium
Blautia
Faecalibacterium
Roseburia
Dialister
Helicobacter
Clostridium (Clostridiaceae)
Treponema
Bulleidia
Ruminococcus
Victivallaceae (unclassified)
Collinsella
Megasphaera
Methanosphaera
Acidaminococcus
Akkermansia
Klebsiella
Enterococcus
Porphyromonas
Neisseria
Comamonas
Host species/geography and study names
Macaque - Yasuda et al.
Macaque - McKenna et al.
Macaque - Handley et al.
Human; Malawi - Yatsunenko et al.
Human; Ameridian - Yatsunenko et al.
Human; US - Yasunenko et al.
Human; US - Human Microbiome Project
D
PCo2
PCo1
Yatsunenko et al
Human Microbiome Project
McKenna et al
Handley et al
Yasuda et al
(Bray-Curtis)
E
PCo2
PCo1
Yatsunenko et al and Yasuda et al
(Weighted unifrac)

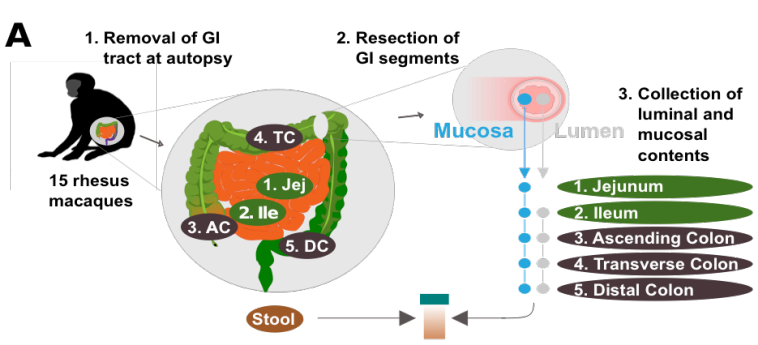
Supplemental Figure 1-1: Influences on gut microbial composition and relating macaque and human microbiota A) A multivariate analysis identified twenty-three taxa that were differentially abundant between the source primate centers from which our cohort originated. Two examples are shown here: an unclassified species of *Treponema,* and an unclassified species of *Clostridium.* Whiskers on boxplot correspond to 1.5 interquartile ranges of the data. The complete list of differentially expressed taxa is available in **Table S1**. B) The Bray-Curtis distance between each sample and the stool sample of the same macaque is plotted for lumenal (left, gray) and mucosal (right, light blue) samples. Samples are stratified by intestinal region. C) To address the influence of host on microbial diversity, all colonic lumen and stool samples are hierarchically clustered based on Bray-Curtis dissimilarity index. Top bar indicates individual animal. D and E) The similarity of microbial communities described in this study, two other macaque studies (McKenna et al., 2008; Handley et al., 2012) and two human studies (Human Microbiome Project, 2012; Yatsunenko et al., 2012) was assessed by calculating the Bray-Curtis dissimilarity and weighted Unifrac distances, then performing principal coordinate analysis. D) Community distance was measured by Bray-Curtis dissimilarity. This plot includes all five studies. E) Community distance was measured by weighted Unifrac distance. This plot only includes the Yatsunenko et al and Yasuda et al datasets.

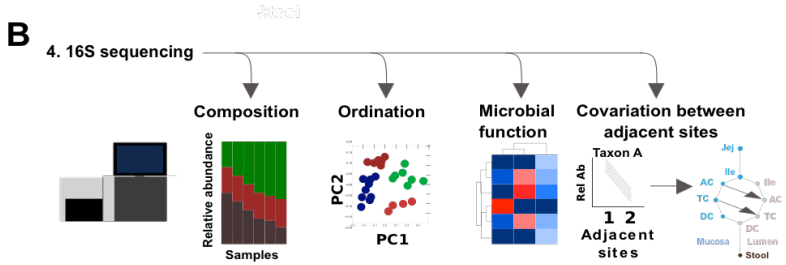


Supplemental Figure 1-2: A phylum-level view of mucosal taxa underrepresented in stool Each dot corresponds to the average relative abundance of an OTU across 15 animals in each intestinal region (SI mucosa and content, LI mucosa and content). Clades of interest are highlighted in red. Marks on the x-axis (vertical lines) or y-axis (horizontal lines) margins represent OTUs with zero measured abundance at one site but non-zero abundance at the other.



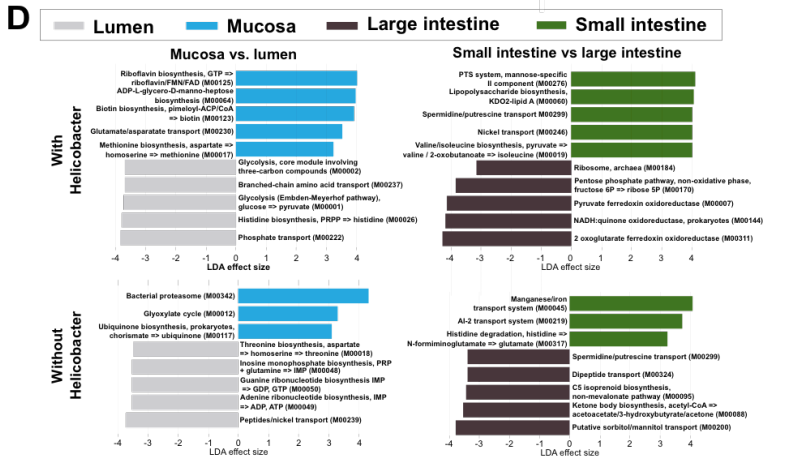
Supplemental Figure 1-3: Biogeographic distribution of taxa throughout the gut For each genus, distribution of its mucosal, lumenal, and stool abundance is shown across the population of 15 macaques, stratified by geography (jejunum, ileum, proximal colon, mid colon, distal colon, stool) and mucosa / lumena. All mucosal samples are shown in blue (except *Pseudomonas*); all lumenal/stool samples are shown in red, and all y axes are log relative abundance. Jejunum and stool samples appear twice as wide as the others because there is not a corresponding paired sample – jejunum is mucosal-only, and stool is stool-only. Whiskers on boxplot correspond to 1.5 interquartile ranges of the data.



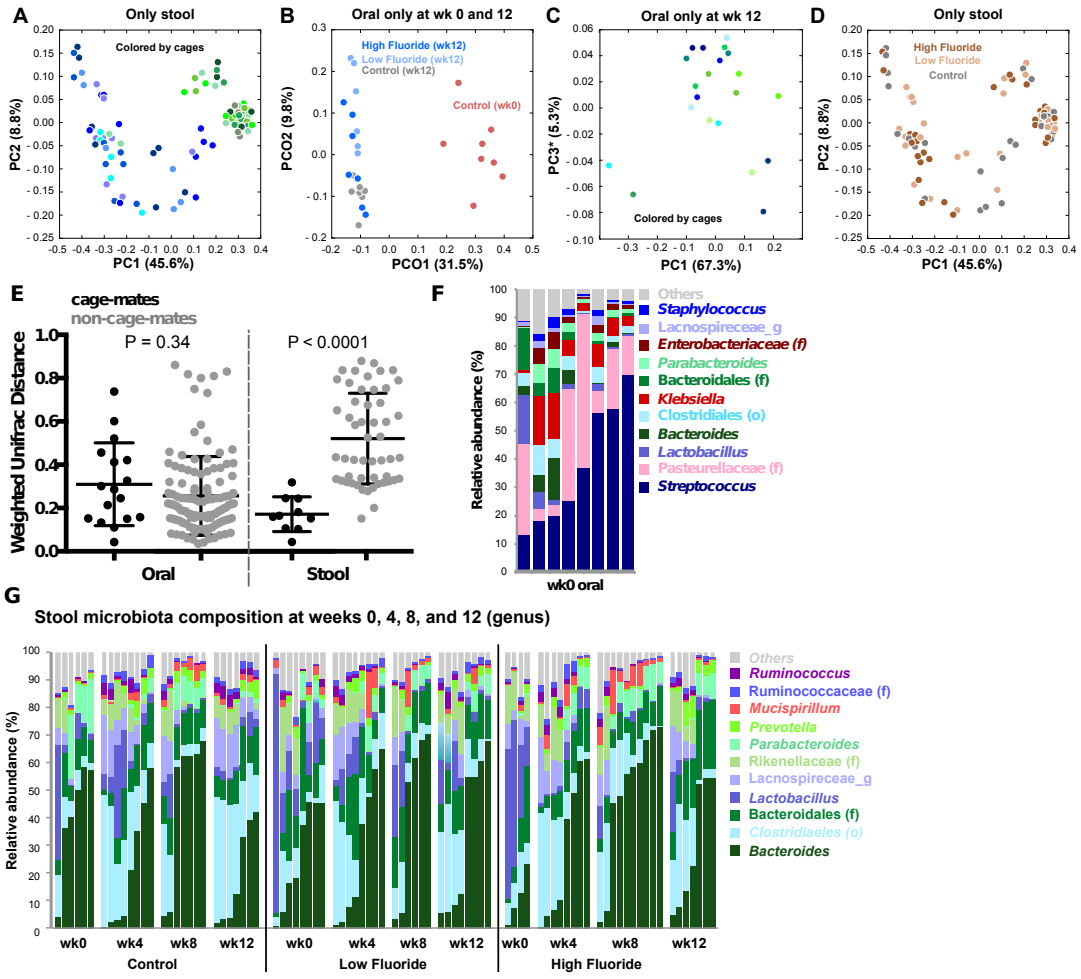


C

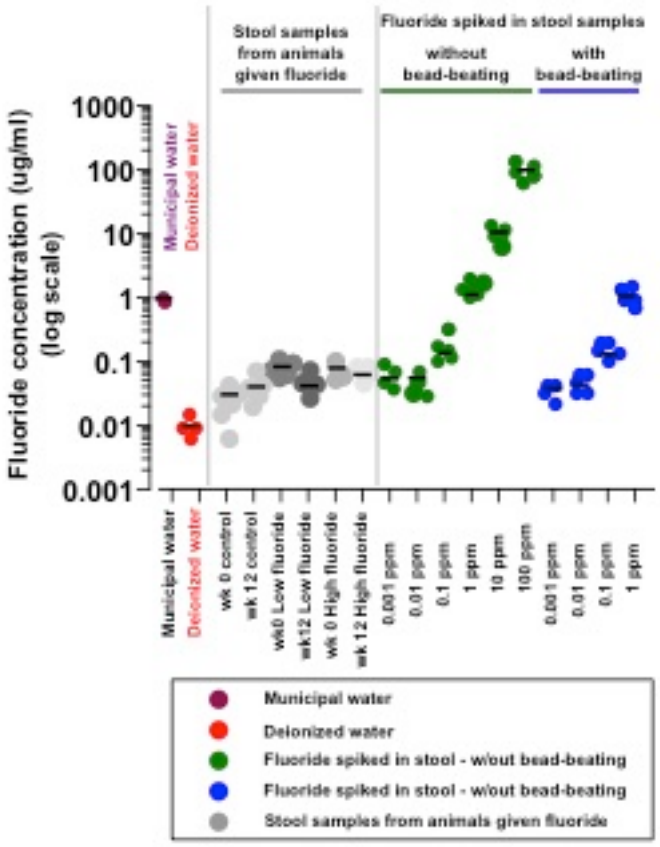
	Microbial mass (cell/ml)		pH		Transit time (h)		Intestinal lengths* (m (%))	
Stomach	$10^{2\text{-}3}$	N/A	1.5-3.0	1.5-3.0	0.5-1.0	-	0.25 - 0.28 (0.04)	0.10 - 0.18 (0.06 - 0.10)
Duodenum	$10^{3\text{-}4}$	N/A	6.0-6.1	5.6-5.8	1.0	Ranges: Low 0.5-1.5 High 4.0-6.0	0.25 - 0.38 (0.04 - 0.06)	0.10 - 0.15 (0.05 - 0.09)
Jejunum	$10^{4\text{-}5}$	N/A	6.1-7.2	5.8-7.0	1.0		2.4 - 2.6 (0.37 - 0.40)	0.7 - 1.2 (0.40 - 0.69)
Ileum	10^{8}	N/A	7.0-7.8	7.0-7.4	1.0		2.0 - 4.0 (0.31 - 0.62)	0.6 - 1.2 (0.34 - 0.69)
Cecum	N/A	N/A	N/A	6.6-7.0	N/A	2.4-5.0	N/A	0.1 - 0.14 (0.06 - 0.08)
Colon	10^{11}	N/A	6.5-7.0	6.6-7.0	3.0-8.0		1.2 - 1.8 (0.19 - 0.28)	0.3 - 0.6 (0.17 - 0.34)



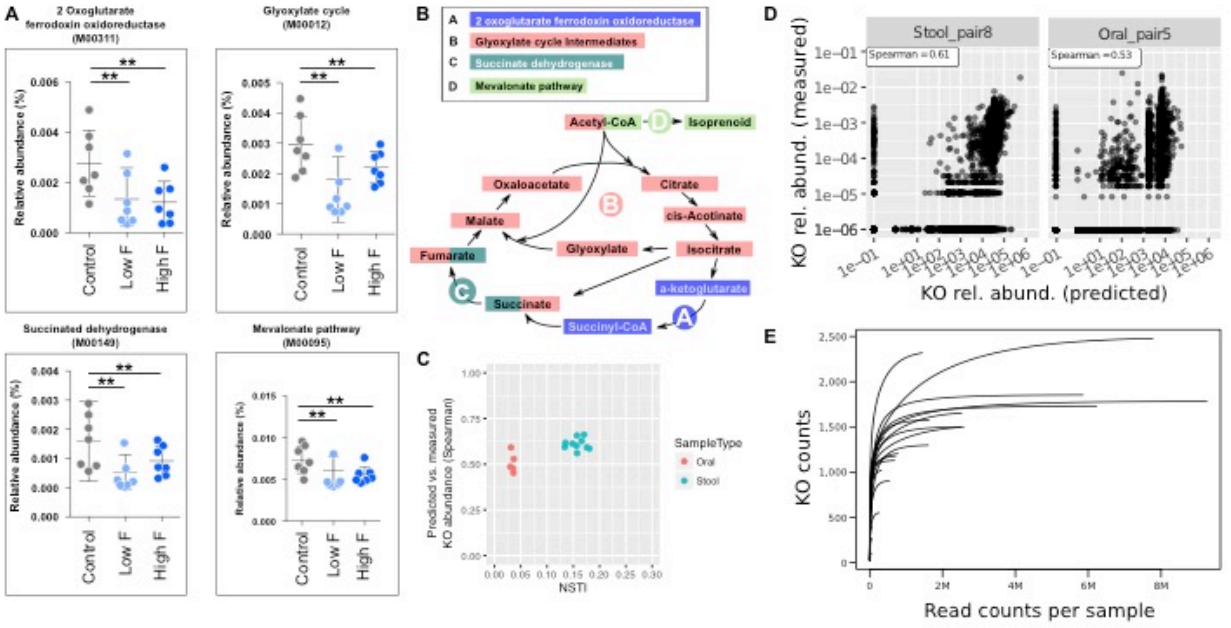
Supplemental figure 1-4: Experimental Procedures. Study design and survey of primate gut microbial biogeography and microbial functional potentials with and without *Helicobacter* A) Paired intestinal mucosal and lumenal contents were collected from the ileum, ascending, transverse, and descending colon of 15 clinically-healthy rhesus macaques, in addition to stool and a sample of jejunal mucosa. The microbiome of the samples was profiled by sequencing the V4 region of the 16S rRNA gene. B) After sequencing, community structure, function, and covariation with biogeography were characterized by ordination [169], univariate [219] and multivariate [133] association testing, metagenomic inference [158], and logistic regression. C) Comparison of the gastrointestinal tracts of humans and rhesus macaques. In contrast to macaques, humans lack a prominent cecum. The total length of the GI tract is 6-7 m for an adult human and 1.5-2m for an adult rhesus macaque. Comparison of intestinal microbial mass [220, 221], pH [220, 222] and transit time [222, 223]. Percent of intestinal length is normalized to an intestinal length of 6.5 m for humans and 1.75 m for a macaque, for comparison purposes. D) PICRUSt [158] was used to infer community function, and LEFSe [147] was used to determine which functions were most differential between the mucosa and lumen and LI and SI. Due to the high abundance of *Helicobacter*, this analysis was repeated with *Helicobacter* removed. The ten largest LDA effects are shown here. The top and bottom two panels are derived from 16S data including *Helicobacte*r, and excluding *Helicobacter* OTUs, respectively.



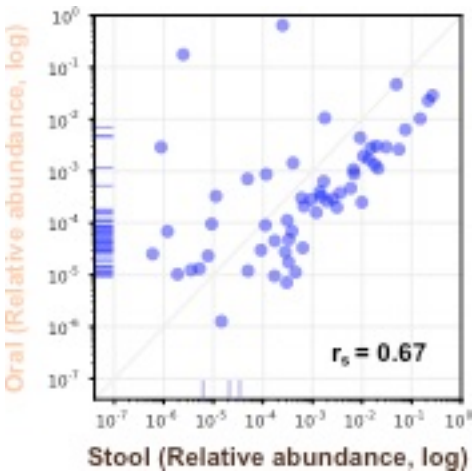
Supplemental figure 2-1: Drivers of oral and gut microbial diversity in fluoride-treated mice (A) Principal coordinates analysis (PCoA) of stool samples from all time-points by weighted UniFrac distance colored by cages. (B) PCoA of week-12 oral samples by weighted UniFrac distance. (C) PCoA of week-12 stool by weighted UniFrac distance colored by cages. (D) PCoA of stool samples from all time-points by weighted UniFrac distance colored by fluoride treatment-groups. (E) Microbial community dissimilarity of oral (left panel) and stool (right panel) sites measured by weighted UniFrac distance amongst mice within cage-mates (intra-cage) and non-cage-mates (inter-cage). (F) Genus-level or higher relative abundance of the oral microbial compositions at week zero. (G) Genus level or higher stool microbiota composition stratified by fluoride groups and time.



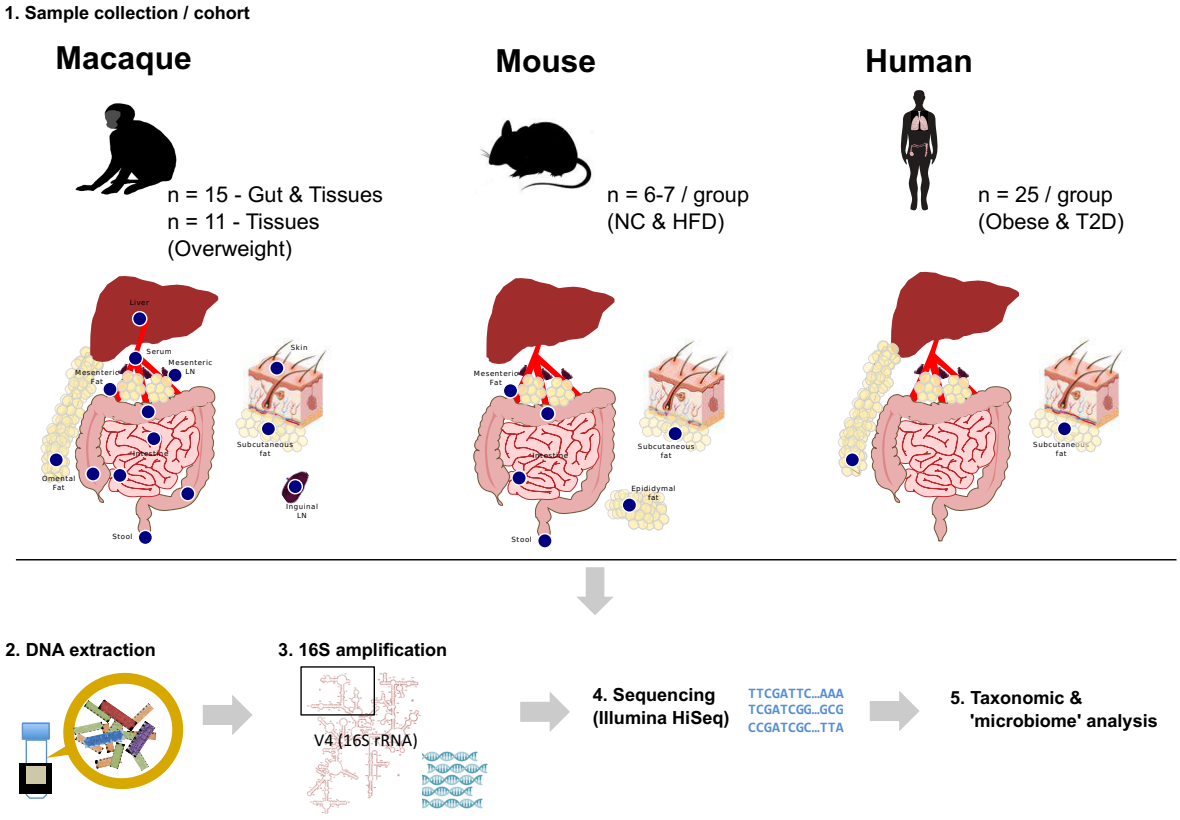
Supplemental figure 2-2: Measuring fluoride concentrations Using a pre-calibrated fluoride ion probe, we measured fluoride concentrations in 1) deionized water, 2) municipal water sampled at the Harvard T. H. Chan School of Public Health (Boston, MA), and 3) mouse stool (100 mg in 10 ml of deionized water) from each treatment group at week 0 and week 12. Post-treatment stool fluoride concentrations were not significantly higher than baseline levels (*t*-test, one-tailed $p>0.05$ in all comparisons). Separate calibration series confirmed that the probe was able to detect the presence of added fluoride in the presence of stool with (blue) and without (green) bead beating (applied to potentially release additional intracellular fluoride). Added fluoride could not be distinguished below baseline stool concentrations (~0.015 ug/ul of dissolved stool).



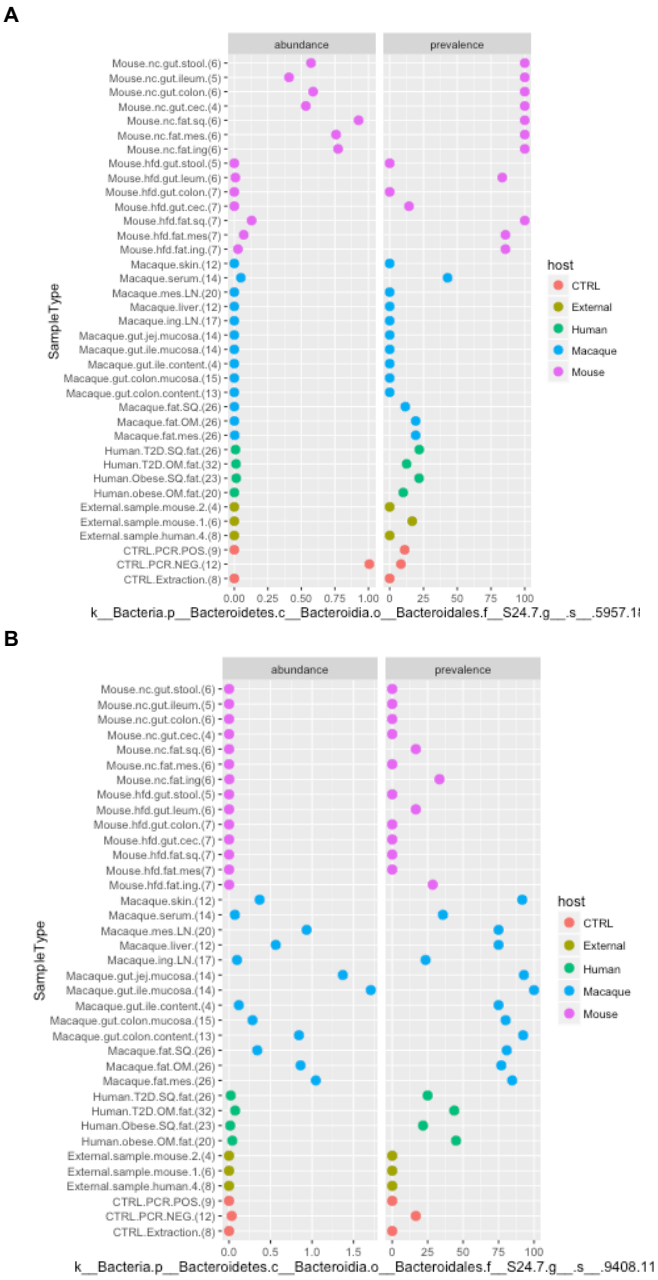
Supplemental figure 2-3: Bacterial functional modules and genes associated with energy production are affected by fluoride in the oral microbial community (**A**) Statistically significant functional modules that are consistently depleted in low- and high-fluoride treatment groups compared to controls. Gene families were inferred from 16S-based taxonomic profiles using PiCRUSt and functional modules were reconstructed using HUMAnN [172]. (**B**) Summary of possible fluoride inhibitory mechanism on microbial carbohydrate metabolism. (**C**) PICRUSt accuracy relative to NSTI scores colored by site (oral and stool). (**D**) Scatterplots illustrate correlation between KO relative abundance measured (shotgun-HUMAnN-imputed; y-axis) and predicted (PICRUSt-16S-imputed; x-axis). Median stool (left) and oral (right) samples are shown. (**E**) Rarefaction curve illustrating the number of KEGG Orthogroups (KOs; y-axis) identified as a function of the number of reads in each sample (x-axis). 89% of samples were saturating with respect to KO richness, defined here as a <10% increase in detected KOs after doubling sequencing depth.



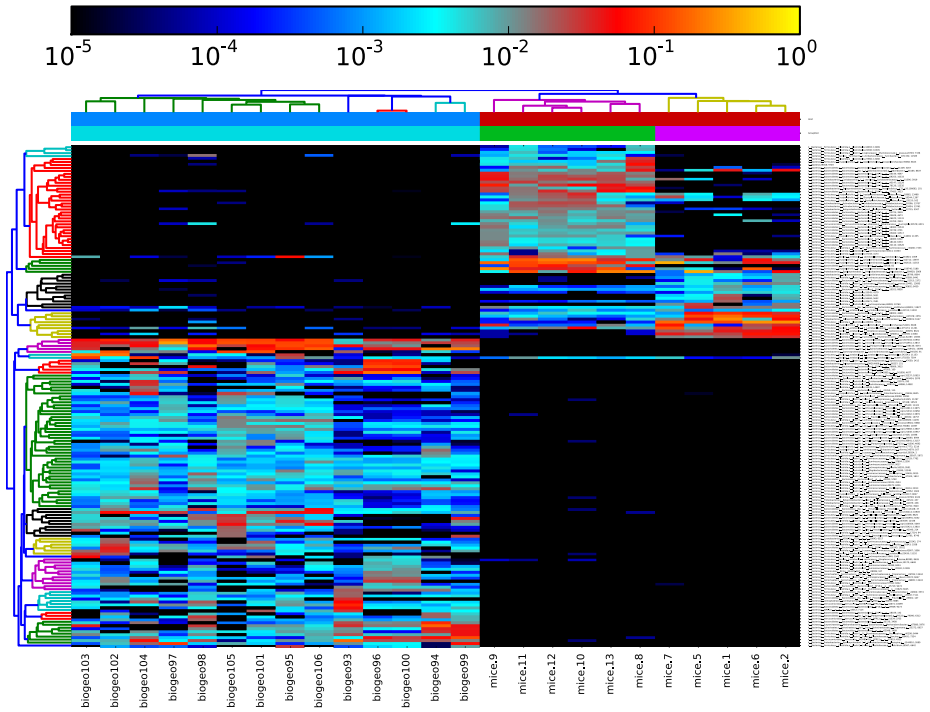
Supplemental figure 2-4: Correlation of bacterial OTUs found in both oral site and stool. Each dot corresponds to the mean relative abundance of an OTU across seven mice for oral (*y*-axis) and stool (*x*-axis) samples. Marks on the *x*-axis (vertical lines) or *y*-axis (horizontal lines) margins represent OTUs with zero measured abundance at one site but non-zero abundance at the other. r_s indicates Spearman correlation.



Supplemental figure 3-1: Study design: cohorts, tissue and gut content sample locations and processing scheme. The adipose tissues samples from mesenteric, omental and subcutaneous depots, mesenteric and inguinal lymph nodes, liver, serum, and skin swabs and intestinal lumenal and mucosal samples from the jejunum, ileum, ascending, transverse, and descending colon, and stool were collected from macaques; adipose tissues from mesenteric, subcutaneous, and epididymal fat depots, and intestinal contents from jejunum/ileum, colon and stool from mouse; omental and subcutaneous fat depots from human. DNA from all samples was extracted, 16S rRNA genes were amplified, multiplexed, and sequenced on Illumina HiSeq and taxonomic composition was analyzed.



Supplemental figure 3-2: Number of sequencing reads per sample across sample types. Each dot represents a sample from different tissue types (rows – y-axis) by different host-species and controls (color) showing sequencing depth (*x*-axis). Red-vertical line is drawn at 10,000 reads.



Supplemental figure 3-3: Unique gut microbiota oligotypes colonize different mammalian hosts (macaque and mouse) There were 103 and 66 unique oligotypes found macaque and mouse gut, respectively. There was one oligotype (*Lactobacillus reuti*) that was shared by two host species at this threshold. Samples and oligoptypes were clustered using bray-curtis distance. Oligotypes were filtered at 0.1% abundance and 50% prevalence for each host species.

Supplemental table 1-1: Bacterial taxa and functions significantly enriched in the mucosa or lumen, a location, or a primate center of origin.

See Cell Host and Microbe Link:
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4369771/bin/NIHMS668212-supplement-3.xls

Supplemental table 1-2: Bacterial OTUs identified in 4 major regions of the intestine but not identified in stool.

Large Intestinal Content		
Taxa	**Stool_(RelAb)**	**LI_Mucosa_(RelAb)**
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pasteurellales|f__Pasteurellaceae|g__unclassified|s__unclassified|279270	0	8.29E-06
k__Bacteria|p__Cyanobacteria|c__Chloroplast|o__Streptophyta|f__unclassified|g__unclassified|s__unclassified|143720	0	2.91E-06
k__Bacteria|p__Proteobacteria|c__Epsilonproteobacteria|o__Campylobacterales|f__Campylobacteraceae|g__Campylobacter|s__Campylobacterrectus|469893	0	2.61E-06
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Streptococcaceae|g__Streptococcus|s__unclassified|305963	0	2.57E-06
k__Bacteria|p__Spirochaetes|c__Brachyspirae|o__Brachyspirales|f__Brachyspiraceae|g__Brachyspira|s__unclassified|84927	0	2.18E-06
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pasteurellales|f__Pasteurellaceae|g__Haemophilus|s__unclassified	0	1.92E-06
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Neisseriales|f__Neisseriaceae|g__Neisseria|s__unclassified|147801	0	1.73E-06
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Neisseriales|f__Neisseriaceae|g__Eikenella|s__Eikenellacorrodens|574200	0	1.37E-06
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Actinomycetales|f__Micrococcaceae|g__Rothia|s__Rothiaaeria|33563	0	1.24E-06
k__Bacteria|p__Cyanobacteria|c__Chloroplast|o__Streptophyta|f__unclassified|g__unclassified|s__unclassified|84627	0	9.45E-07

Large Intestinal Mucosa		
Taxa	**Stool_(RelAb)**	**LI_Mucosa_(RelAb)**
k__Bacteria|p__Spirochaetes|c__Brachyspirae|o__Brachyspirales|f__Brachyspiraceae|g__Brachyspira|s__unclassified|84927	0	0.000134341
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__unclassified|g__Aquabacterium|s__unclassified|560907	0	3.85E-05
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Sphingomonadales|f__Sphingomonadaceae	0	2.23E-05
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Hylemonella|s__unclassified|569527	0	2.02E-05
k__Bacteria|p__Bacteroidetes|c__Flavobacteria|o__Flavobacteriales|f__Flavobacteriaceae|g__unclassified|s__unclassified|388958	0	1.91E-05
k__Bacteria|p__Fusobacteria|c__Fusobacteria|o__Fusobacteriales|f__Fusobacteriaceae|g__Streptobacillus|s__unclassified|449686	0	9.41E-06
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Pseudomonadaceae|g__Pseudomonas|s__unclassified|271906	0	7.66E-06
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Streptococcaceae|g__Streptococcus|s__unclassified|305963	0	6.54E-06
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamonadaceae|g__Curvibacter|s__unclassified	0	6.51E-06
k__Bacteria|p__Firmicutes|c__Clostridia|o__Clostridiales|f__Veillonellaceae|g__Veillonella|s__unclassified|428006	0	4.34E-06

k__Bacteria|p__Cyanobacteria|c__Chloroplast|o__Streptophyta|f__unclassified|g__unclassified|s__unclassified|143720	0	3.52E-06
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Carnobacteriaceae|g__Carnobacterium|s__Carnobacteriumviridans|140940	0	3.45E-06
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae|g__Ralstonia|s__unclassified|534464	0	3.27E-06
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhodospirillales	0	2.71E-06
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Neisseriales|f__Neisseriaceae|g__Eikenella|s__Eikenellacorrodens|574200	0	2.47E-06
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__unclassified|g__Mitsuaria|s__unclassified|7148	0	2.06E-06
k__Bacteria|p__Bacteroidetes|c__Bacteroidia|o__Bacteroidales|f__Prevotellaceae|g__Prevotella|s__Prevotellahisticola|423307	0	1.42E-06

Small Intestinal Mucosa		
Taxa	**Stool_(RelAb)**	**SI_Content_(RelAb)**
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Streptococcaceae|g__Streptococcus|s__unclassified|305963	0	1.38E-05
k__Bacteria|p__Firmicutes|c__Clostridia|o__Clostridiales|f__Lachnospiraceae|g__unclassified|s__unclassified|537098	0	8.45E-06
k__Bacteria|p__Firmicutes|c__Clostridia|o__Clostridiales|f__Veillonellaceae|g__Veillonella|s__unclassified|428006	0	1.69E-05
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pasteurellales|f__Pasteurellaceae|g__unclassified|s__unclassified|279270	0	1.38E-05

Small Intestinal Mucosa		
Taxa	**Stool_(RelAb)**	**SI_Mucosa_(RelAb)**
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Actinomycetales|f__Micrococcaceae|g__Rothia|s__Rothiaaeria|33563	0	2.29E-04
k__Bacteria|p__Bacteroidetes|c__Bacteroidia|o__Bacteroidales|f__Porphyromonadaceae|g__Porphyromonas|s__unclassified|256893	0	1.28E-03
k__Bacteria|p__Bacteroidetes|c__Bacteroidia|o__Bacteroidales|f__Prevotellaceae|g__Prevotella|s__Prevotellafalsenii|288721	0	1.66E-04
k__Bacteria|p__Bacteroidetes|c__Bacteroidia|o__Bacteroidales|f__Prevotellaceae|g__Prevotella|s__Prevotellahisticola|423307	0	8.91E-05
k__Bacteria|p__Bacteroidetes|c__Bacteroidia|o__Bacteroidales|f__Prevotellaceae|g__Prevotella|s__unclassified|149558	0	4.05E-04
k__Bacteria|p__Bacteroidetes|c__Bacteroidia|o__Bacteroidales|f__Prevotellaceae|g__Prevotella|s__unclassified|2181	0	1.07E-04
k__Bacteria|p__Bacteroidetes|c__Bacteroidia|o__Bacteroidales|f__Prevotellaceae|g__Prevotella|s__unclassified|251453	0	1.34E-04
k__Bacteria|p__Bacteroidetes|c__Bacteroidia|o__Bacteroidales|f__Prevotellaceae|g__Prevotella|s__unclassified|469379	0	4.08E-04
k__Bacteria|p__Bacteroidetes|c__Flavobacteria|o__Flavobacteriales|f__Flavobacteriaceae|g__unclassified|s__unclassified|286716	0	3.23E-04

k__Bacteria|p__Bacteroidetes|c__Flavobacteria|o__Flavobacteriales|f__Flavobacteriacea e|g__unclassified|s__unclassified|388958	0	3.89E-04
k__Bacteria|p__Cyanobacteria|c__Chloroplast|o__Streptophyta|f__unclassified|g__uncla ssified|s__unclassified|143720	0	6.33E-03
k__Bacteria|p__Cyanobacteria|c__Chloroplast|o__Streptophyta|f__unclassified|g__uncla ssified|s__unclassified|84627	0	5.58E-04
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Carnobacteriaceae|g__Carno bacterium|s__Carnobacteriumviridans|140940	0	1.76E-04
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Carnobacteriaceae|g__unclas sified|s__unclassified|378347	0	1.30E-04
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Enterococcaceae|g__Enteroc occus|s__Enterococcussulfureus|53470	0	2.84E-04
k__Bacteria|p__Firmicutes|c__Bacilli|o__Lactobacillales|f__Streptococcaceae|g__Strepto coccus|s__unclassified|305963	0	2.36E-04
k__Bacteria|p__Firmicutes|c__Clostridia|o__Clostridiales|f__Lachnospiraceae|g__Catonel la|s__unclassified|48583	0	1.63E-04
k__Bacteria|p__Firmicutes|c__Clostridia|o__Clostridiales|f__Lachnospiraceae|g__unclass ified|s__unclassified|537098	0	4.07E-04
k__Bacteria|p__Firmicutes|c__Clostridia|o__Clostridiales|f__Peptococcaceae|g__Peptoco ccus|s__unclassified|249899	0	5.46E-05
k__Bacteria|p__Firmicutes|c__Clostridia|o__Clostridiales|f__Peptostreptococcaceae|g__ Peptostreptococcus|s__unclassified|527485	0	1.46E-04
k__Bacteria|p__Firmicutes|c__Clostridia|o__Clostridiales|f__Veillonellaceae|g__Selenom onas|s__unclassified|59529	0	2.66E-04
k__Bacteria|p__Firmicutes|c__Clostridia|o__Clostridiales|f__Veillonellaceae|g__Veillonell a|s__unclassified|428006	0	6.09E-04
k__Bacteria|p__Fusobacteria|c__Fusobacteria|o__Fusobacteriales|f__Fusobacteriaceae|g __Leptotrichia|s__Leptotrichiabuccalis|535068	0	1.16E-04
k__Bacteria|p__Fusobacteria|c__Fusobacteria|o__Fusobacteriales|f__Fusobacteriaceae|g __Leptotrichia|s__unclassified|324532	0	2.21E-04
k__Bacteria|p__Fusobacteria|c__Fusobacteria|o__Fusobacteriales|f__Fusobacteriaceae|g __Streptobacillus|s__unclassified|449686	0	7.83E-04
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Rhodospirillales	0	7.84E-05
k__Bacteria|p__Proteobacteria|c__Alphaproteobacteria|o__Sphingomonadales|f__Sphin gomonadaceae	0	2.12E-04
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholde riaceae|g__Ralstonia|s__unclassified|534464	0	4.11E-04
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamon adaceae|g__Acidovorax|s__unclassified|141709	0	2.79E-04
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamon adaceae|g__Curvibacter|s__unclassified	0	1.72E-04
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Comamon adaceae|g__Hylemonella|s__unclassified|569527	0	7.36E-04
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__unclassifie d|g__Aquabacterium|s__unclassified|560907	0	2.78E-04
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__unclassifie d|g__Mitsuaria|s__unclassified|7148	0	1.88E-04
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Neisseriales|f__Neisseriaceae |g__Eikenella|s__Eikenellacorrodens|574200	0	1.49E-03
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Neisseriales|f__Neisseriaceae |g__Neisseria|s__Neisserialactamica|548102	0	1.59E-04

k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Neisseriales|f__Neisseriaceae|g__Neisseria|s__unclassified|147801	0	5.62E-04
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Neisseriales|f__Neisseriaceae|g__Simonsiella|s__unclassified|447914	0	7.47E-05
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Rhodocyclales|f__Rhodocyclaceae|g__unclassified|s__unclassified|249511	0	1.74E-04
k__Bacteria|p__Proteobacteria|c__Epsilonproteobacteria|o__Campylobacterales|f__Campylobacteraceae|g__Campylobacter|s__Campylobacterconcisus|380913	0	2.02E-04
k__Bacteria|p__Proteobacteria|c__Epsilonproteobacteria|o__Campylobacterales|f__Campylobacteraceae|g__Campylobacter|s__Campylobacterrectus|469893	0	9.00E-04
k__Bacteria|p__Proteobacteria|c__Epsilonproteobacteria|o__Campylobacterales|f__Campylobacteraceae|g__Campylobacter|s__unclassified|200309	0	5.38E-04
k__Bacteria|p__Proteobacteria|c__Epsilonproteobacteria|o__Campylobacterales|f__Helicobacteraceae|g__Helicobacter|s__Helicobactersuis|550809	0	5.89E-04
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Enterobacteriales|f__Enterobacteriaceae|g__unclassified|s__unclassified|464068	0	1.00E-04
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pasteurellales|f__Pasteurellaceae|g__Aggregatibacter|s__Aggregatibacteraphrophilus|269356	0	1.35E-04
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pasteurellales|f__Pasteurellaceae|g__Haemophilus|s__unclassified	0	1.89E-04
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pasteurellales|f__Pasteurellaceae|g__unclassified|s__BisgaardTaxon10|140697	0	3.59E-04
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pasteurellales|f__Pasteurellaceae|g__unclassified|s__unclassified|109416	0	2.38E-04
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pasteurellales|f__Pasteurellaceae|g__unclassified|s__unclassified|279270	0	2.52E-03
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Moraxellaceae|g__Acinetobacter|s__unclassified|350209	0	1.44E-04
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Moraxellaceae|g__Moraxella|s__unclassified|246528	0	1.72E-03
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Pseudomonadales|f__Pseudomonadaceae|g__Pseudomonas|s__unclassified|271906	0	1.76E-04

Supplemental table 2-1: Differentially abundant bacterial genera by oral-stool biogeography in mice inferred from univariate analysis.

Taxon	Enriched site	LDA	P-value
k__Bacteria| p__Proteobacteria| c__Gammaproteobacteria| o__Pasteurellales| f__Pasteurellaceae_g__	Oral	4.98	3.909E-18
k__Bacteria| p__Firmicutes| c__Bacilli| o__Lactobacillales| f__Streptococcaceae| g__Streptococcus	Oral	5.46	1.784E-14
k__Bacteria| p__Bacteroidetes| c__Bacteroidia| o__Bacteroidales| f__S24_7_g__	Stool	4.74	1.051E-13
k__Bacteria| p__Bacteroidetes| c__Bacteroidia| o__Bacteroidales| f__Bacteroidaceae| g__Bacteroides	Stool	5.16	1.211E-11
k__Bacteria| p__Firmicutes| c__Clostridia| o__Clostridiales| f___g__	Stool	4.83	1.133E-10
k__Bacteria| p__Firmicutes| c__Clostridia| o__Clostridiales| f__Lachnospiraceae| g__	Stool	4.40	5.437E-04
k__Bacteria| p__Bacteroidetes| c__Bacteroidia| o__Bacteroidales| f__Porphyromonadaceae| g__Parabacteroides	Stool	4.20	7.669E-04
k__Bacteria| p__Bacteroidetes| c__Bacteroidia| o__Bacteroidales| f__Rikenellaceae| g__	Stool	4.41	2.570E-03
k__Bacteria| p__Firmicutes| c__Bacilli| o__Lactobacillales| f__Lactobacillaceae| g__Lactobacillus	Stool	4.33	3.708E-02

Supplemental table 2-2: Fluoride treatment effects on oral bacterial OTUs at week 12.

Taxon depleted in both low and high fluoride groups	Value	Coefficient	P.value	Q.value
k__Bacteria|p__Bacteroidetes|c__Bacteroidia|o__Bacteroidales|f__Porphyromonadaceae|g__Parabacteroides|s__distasonis	Depleted in High Fluoride	-2.2E-02	2.2E-04	3.1E-02
k__Bacteria|p__Bacteroidetes|c__Bacteroidia|o__Bacteroidales|f__Porphyromonadaceae|g__Parabacteroides|s__distasonis	Depeleted in Low Fluoride	-1.8E-02	8.1E-04	3.9E-02
k__Bacteria|p__Bacteroidetes|c__Bacteroidia|o__Bacteroidales|f__Bacteroidaceae|g__Bacteroides|s__uniformis	Depleted in High Fluoride	-2.4E-02	4.4E-04	3.2E-02
k__Bacteria|p__Bacteroidetes|c__Bacteroidia|o__Bacteroidales|f__Bacteroidaceae|g__Bacteroides|s__uniformis	Depeleted in Low Fluoride	-1.7E-02	7.9E-03	1.5E-01
k__Bacteria|p__Bacteroidetes|c__Bacteroidia|o__Bacteroidales|f__Bacteroidaceae|g__Bacteroides|s__	Depleted in High Fluoride	-8.4E-02	3.6E-04	3.2E-02
k__Bacteria|p__Bacteroidetes|c__Bacteroidia|o__Bacteroidales|f__Bacteroidaceae|g__Bacteroides|s__	Depeleted in Low Fluoride	-6.6E-02	2.5E-03	8.0E-02

Taxon depleted in high fluoride	Value	Coefficient	P.value	Q.value
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Alcaligenaceae|g__Sutterella|s__	Depleted in High Fluoride	-1.2E-02	7.4E-03	1.5E-01
k__Bacteria|p__Proteobacteria|c__Deltaproteobacteria|o__Desulfovibrionales|f__Desulfovibrionaceae|g__Bilophila|s__	Depleted in High Fluoride	-5.7E-03	9.0E-03	1.6E-01

Taxon depleted in low fluoride	Value	Coefficient	P.value	Q.value
k__Bacteria|p__Bacteroidetes|c__Bacteroidia|o__Bacteroidales|f__|g__|s__	Depeleted in Low Fluoride	-1.7E-02	1.2E-03	4.5E-02
k__Bacteria|p__Proteobacteria|c__Betaproteobacteria|o__Burkholderiales|f__Burkholderiaceae|g__Burkholderia|s__	Depeleted in Low Fluoride	-4.3E-02	3.0E-03	8.0E-02
k__Bacteria|p__Bacteroidetes|c__Bacteroidia|o__Bacteroidales|f__Rikenellaceae|g__AF12|s__	Depeleted in Low Fluoride	-1.4E-02	3.0E-03	8.0E-02
k__Bacteria|p__Actinobacteria|c__Actinobacteria|o__Actinomycetales|f__Corynebacteriaceae|g__Corynebacterium|s__mastitidis	Depeleted in Low Fluoride	-5.8E-02	6.3E-03	1.5E-01
k__Bacteria|p__Proteobacteria|c__Deltaproteobacteria|o__|f__|g__|s__	Depeleted in Low Fluoride	-1.8E-02	6.5E-03	1.5E-01

Taxon enriched in high fluoride	Value	Coefficient	P.value	Q.value
k__Bacteria|p__Proteobacteria|c__Gammaproteobacteria|o__Aeromonadales|f__Succinivibrionaceae|g__Ruminobacter|s__	Enriched in High Fluoride	8.4E-03	1.6E-04	3.1E-02

Supplemental table 2-3: Fluoride treatment effects on oral bacterial functional modules at week 12.

Depleted functional modules in both low and high fluoride groups	Value	Coefficient	P.value	Q.value
M00311_2 Oxoglutarate ferredoxin oxidoreductase	Depeleted in Low Fluoride	-2.42E-02	6.33E-03	1.91E-01
M00311_2 Oxoglutarate ferredoxin oxidoreductase	Depleted in High Fluoride	-1.84E-02	6.14E-03	1.91E-01
M00012_Glyoxylate cycle	Depeleted in Low Fluoride	-2.45E-02	6.55E-03	1.91E-01
M00012_Glyoxylate cycle	Depleted in High Fluoride	-1.64E-02	6.90E-03	1.91E-01
M00149_Succinate dehydragenase	Depeleted in Low Fluoride	-2.42E-02	7.43E-03	1.91E-01
M00149_Succinate dehydragenase	Depleted in High Fluoride	-1.19E-02	7.69E-03	1.91E-01
M00095 Mevalonate Pathway	Depeleted in Low Fluoride	-1.57E-02	7.95E-03	1.91E-01
M00095 Mevalonate Pathway	Depleted in High Fluoride	-9.25E-03	7.95E-03	1.91E-01
M00348_Glutathione transport system	Depeleted in Low Fluoride	-3.42E-02	7.95E-03	1.97E-01
M00348_Glutathione transport system	Depleted in High Fluoride	-3.93E-02	8.26E-03	1.97E-01
M00025_Tyrosine biosynthesis (chorismate => tyrosine)	Depeleted in Low Fluoride	-5.91E-02	8.16E-03	1.97E-01
M00025_Tyrosine biosynthesis (chorismate => tyrosine)	Depleted in High Fluoride	-1.91E-02	7.90E-03	1.97E-01

Depleted or enriched functional modules in either low or high fluoride groups	Value	Coefficient	P.value	Q.value
M00034_Methionine salvage pathway	Enriched in Low Fluoride	4.23E-02	6.04E-03	1.91E-01
M00126_Tetrahydrofolate biosynthesis	Depleted in High Fluoride	-3.93E-02	6.52E-03	1.91E-01
M00136_GABA biosynthesis (putrescine => GABA)	Enriched in High Fluoride	5.01E-02	6.99E-03	1.91E-01
M00202_Oligogalacturonide transport system	Enriched in High Fluoride	2.11E-02	7.11E-03	1.91E-01
M00220_Rhamnose transport system	Depleted in High Fluoride	-9.12E-03	7.00E-03	1.91E-01
M00225_Lysine/arginine/ornithine transport system	Depleted in Low Fluoride	-1.73E-02	7.66E-03	1.91E-01
M00278_PT System (sorbose-specific component)	Enriched in High Fluoride	4.30E-02	7.94E-03	1.91E-01
M00302_1-Aminoethylphosphonate transport system	Depleted in Low Fluoride	-2.99E-02	7.03E-03	1.91E-01
M00198_sn-Glycerol3-phosphate transport	Depleted in High Fluoride	-1.93E-02	8.10E-03	1.97E-01
M00349_Microcin C Transport system	Enriched in Low Fluoride	3.13E-02	8.65E-03	1.97E-01
M00300_Putrescine transport	Depleted in High Fluoride	-4.23E-01	7.33E-03	1.97E-01
M00226_Histidine transport	Depleted in High Fluoride	-3.00E-02	8.21E-03	1.97E-01
M00230_Glutamate/aspartate transport	Depleted in High Fluoride	-3.11E-02	8.01E-03	1.97E-01

Supplemental table 2-4: Genus-level taxonomic profiles of all samples from 16S-sequencing and their associated metadata.

Data available from mSystems:

http://msystems.asm.org/content/2/4/e00047-17#DC7

Supplemental table 2-5: Taxonomic profiles (MetaPhlAn2) for the subset of samples analyzed with shotgun sequencing and their associated metadata

Data available from mSystems:

http://msystems.asm.org/content/msys/2/4/e00047-17/DC9/embed/inline-supplementary-material-9.xls?download=true

Supplemental table 3-1: Bacterial taxa (oligotype) assigned to different categories based on the k-ratio.

k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus:13907:109874	control.lo	REAL
k__Bacteria:15:30117	human.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__melaninogenica:59:36535	human.lo	REAL
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Bacillaceae;g__Bacillus;s__cereus:3404:24199	human.lo	REAL
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Aerococcaceae;g__Alloiococcus;s__:4156:33038	human.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Veillonella;s__dispar:7065:43406	human.lo	REAL
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Micrococcaceae;g__Rothia;s__mucilaginosa:12726:20314	human.lo	REAL
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Micrococcaceae;g__Rothia;s__mucilaginosa:12727:51407	human.lo	REAL
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus:14095:21909	human.lo	REAL
k__Bacteria;p__Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Helicobacteraceae;g__Flexispira;s__:4470:31041	macaque.hi	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium;s__prausnitzii:2:39763	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__copri:40:113182	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Moraxella;s__:78:50913	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__RF16;g__;s__:79:21467	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Helicobacteraceae;g__Flexispira;s__:105:116104	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Tenericutes;c__RF3;o__ML615J-28;f__;g__;s__:114:28112	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__:156:25760	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__GMD14H09;f__;g__;s__:161:61591	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__;s__:179:48246	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Ruminococcus;s__:184:34566	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__:218:41852	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Lentisphaerae;c__[Lentisphaeria];o__Z20;f__R4-45B;g__;s__:262:30594	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__;s__:646:41761	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae;g__Sutterella;s__:995:24745	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae;g__Sutterella;s__:996:22763	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Christensenellaceae;g__;s__:1263:22070	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae;g__Sutterella;s__:1266:20160	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Coprococcus;s__:1428:47645	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__[Ruminococcus];s__gnavus:2033:30934	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__RF32;f__;g__;s__:2247:18826	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae:2394:2	macaque.hi;macaque.lo	REAL

4853		
k__Bacteria;p__Verrucomicrobia;c__Verruco-5;o__WCHB1-41;f__RFP12;g__;s__:2791:67732	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Verrucomicrobia;c__Verruco-5;o__WCHB1-41;f__RFP12;g__;s__:2924:21336	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae;g__Treponema;s__:3090:218169	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae:3092:39163	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Verrucomicrobia;c__Verruco-5;o__WCHB1-41;f__RFP12;g__;s__:3096:44372	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Helicobacteraceae;g__Flexispira;s__:3551:548522	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Helicobacteraceae;g__Flexispira;s__:3555:26074	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae:3719:307075	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Aggregatibacter:3720:31322	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Aggregatibacter;s__:3721:56092	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Aggregatibacter:3724:117171	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Aggregatibacter:3850:32975	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Aggregatibacter:3851:33887	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae;g__Treponema;s__:3975:30294	macaque.hi;macaque.lo	REAL
k__Archaea;p__Euryarchaeota;c__Methanobacteria;o__Methanobacteriales;f__Methanobacteriaceae;g__Methanobrevibacter;s__:3976:67797	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__stercorea:4343:59566	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__stercorea:4344:35601	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__:4348:36041	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:4650:76492	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Phascolarctobacterium;s__:4657:31771	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae;g__Treponema;s__:4898:55108	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__copri:5106:329654	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__copri:5110:32245	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__copri:5114:20814	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Peptococcaceae;g__Peptococcus;s__:5239:32458	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__:5440:69790	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__copri:5445:40697	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae:5795:23020	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__YRC22;s__:5945:78999	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Ruminococcus;s__bromii:6546:36913	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Megasphaera;s__:7049:52024	macaque.hi;macaque.lo	REAL

k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__;s__:7263:35713	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__;s__:7264:62124	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__Sarcina;s__:7429:89142	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Helicobacteraceae;g__Flexispira;s__:7916:1547448	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Helicobacteraceae;g__Flexispira;s__:7917:23180	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Helicobacteraceae;g__Flexispira;s__:7921:27343	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus;s__mucosae:8192:31939	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus:8202:26026	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales:8324:45423	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Ruminococcus;s__flavefaciens:8467:42859	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Oribacterium;s__:8474:27781	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__CF231;s__:8699:237578	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__p-2534-18B5;g__;s__:8703:140668	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__CF231;s__:8721:40852	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Dialister;s__:8723:29349	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__copri:9092:100259	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__copri:9095:60864	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium;s__prausnitzii:9222:136472	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Faecalibacterium;s__prausnitzii:9227:45207	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Anaerovibrio;s__:9410:97476	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Oscillospira;s__:9648:28090	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Oscillospira;s__:9875:48037	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Oscillospira;s__:9877:34354	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae:9879:21746	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae:10000:57531	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae:10004:40260	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Oscillospira;s__:10025:40302	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__:10202:22700	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Catenibacterium;s__:10684:24871	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus;s__:10897:140937	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus;s__:10923:39347	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Fusobacteria;c__Fusobacteriia;o__Fusobacteriales;f__Fusobacteriaceae;g__Fusobacterium;s__:10924:44750	macaque.hi;macaque.lo	REAL

k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Roseburia:11393:44336	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__stercorea:12017:36411	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__stercorea:12018:57849	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae;g__Treponema;s__:12151:49859	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Spirochaetes;c__Spirochaetes;o__Spirochaetales;f__Spirochaetaceae;g__Treponema;s__:12152:26769	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__copri:12205:29397	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__copri:12208:55475	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__copri:12276:24438	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__;s__:12485:20315	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__;s__:12486:37548	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella];s__:13105:95827	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella];s__:13108:28565	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella];s__:13207:60373	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella];s__:13208:98869	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__SMB53;s__:13267:62708	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__[Mogibacteriaceae];g__;s__:13271:22003	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Veillonella;s__:13360:26796	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Veillonellaceae;g__Dialister;s__:13361:104375	macaque.hi;macaque.lo	REAL
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Corynebacteriaceae;g__Corynebacterium;s__:849:20551	macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Staphylococcaceae;g__Staphylococcus;s__:3401:43275	macaque.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__YRC22;s__:5944:24610	macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria:6155:39414	macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__[Tissierellaceae];g__Peptoniphilus;s__:7070:41444	macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__[Tissierellaceae];g__Anaerococcus;s__:8710:30145	macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Phyllobacteriaceae;g__Mesorhizobium;s__:9641:147459	macaque.lo	REAL
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Corynebacteriaceae;g__Corynebacterium;s__:11520:45961	macaque.lo	REAL
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Corynebacteriaceae;g__Corynebacterium;s__:11522:63380	macaque.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__[Ruminococcus];s__gnavus:13645:40498	macaque.lo	REAL
k__Bacteria;p__Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Helicobacteraceae;g__Helicobacter:4469:220442	mouse.hi	REAL
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus;s__reuteri:8102:43598	mouse.hi	REAL
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus;s__reuteri:8104:44732	mouse.hi	REAL
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae:11580:34369	mouse.hi	REAL

k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus;s__:13050:33373	mouse.hi	REAL
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus;s__:13058:26742	mouse.hi	REAL
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rickettsiales;f__;g__;s__:47:19589	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__;g__;s__:1618:24683	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__;s__:1811:37538	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales:2044:37812	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Allobaculum;s__:3289:20487	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Desulfovibrionales;f__Desulfovibrionaceae;g__Desulfovibrio;s__:4906:21049	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae:5443:51028	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:5957:18329	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Odoribacteraceae];g__Odoribacter;s__:7063:63235	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:8706:94399	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides;s__:8954:261472	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:9413:23460	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:9415:23886	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Oscillospira;s__:10022:27176	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae:10681:159057	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae:10682:40886	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Tenericutes;c__Mollicutes;o__Mycoplasmatales;f__Mycoplasmataceae;g__;s__:10904:44954	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__[Ruminococcus];s__gnavus:11449:60413	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__[Ruminococcus];s__gnavus:11451:20646	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae;g__Sutterella;s__:12537:46194	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Alcaligenaceae;g__Sutterella;s__:12540:177920	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Allobaculum;s__:13726:19702	mouse.hi;mouse.lo	REAL
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales:1615:138897	mouse.lo	REAL

k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Corynebacteriaceae;g__Corynebacterium;s__:1413:66362	control.lo;human.lo;macaque.hi;macaque.lo	COMMON
k__Bacteria;p__Firmicutes;c__Bacilli;o__Gemellales;f__Gemellaceae;g__;s__:5237:156322	control.lo;human.lo;macaque.hi;macaque.lo	COMMON
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus;s__:10899:193187	control.lo;human.lo;macaque.hi;macaque.lo	COMMON
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus:13837:25967	control.lo;human.lo;macaque.hi;macaque.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae;g__Escherichia;s__coli:2927:137277	control.lo;human.lo;macaque.hi;macaque.lo;mouse.hi;mouse.lo	COMMON
k__Bacteria;p__Deferribacteres;c__Deferribacteres;o__Deferribacterales;f__Deferribacteraceae;g__Mucispirillum;s__schaedleri:12380:598698	control.lo;human.lo;macaque.hi;macaque.lo;mouse.hi;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Ralstonia;s__:12531:7569131	control.lo;human.lo;macaque.hi;macaque.lo;mouse.hi;mouse.lo	COMMON

k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus:12834:717468	control.lo;human.lo;macaque.hi;macaque.lo;mouse.hi;mouse.lo	COMMON
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus;s__:13053:1703758	control.lo;human.lo;macaque.hi;macaque.lo;mouse.hi;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Neisseriales;f__Neisseriaceae;g__;s__:135:54845	control.lo;human.lo;macaque.hi;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Bacteroidetes;c__Flavobacteriia;o__Flavobacteriales;f__[Weeksellaceae];g__Cloacibacterium;s__:843:68669	control.lo;human.lo;macaque.hi;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Ralstonia;s__:6292:1137405	control.lo;human.lo;macaque.hi;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Haemophilus:8606:67212	control.lo;human.lo;macaque.hi;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Staphylococcaceae;g__Staphylococcus;s__:11708:815689	control.lo;human.lo;macaque.hi;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus:14090:365889	control.lo;human.lo;macaque.hi;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Bradyrhizobiaceae:33:45774	control.lo;human.lo;macaque.hi;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Enterobacteriales;f__Enterobacteriaceae:11579:46195	control.lo;human.lo;macaque.lo;mouse.hi	COMMON
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus:13903:1751618	control.lo;human.lo;macaque.lo;mouse.hi	COMMON
k__Bacteria;p__Firmicutes;c__Bacilli;o__Turicibacterales;f__Turicibacteraceae;g__Turicibacter;s__:10:772706	control.lo;human.lo;macaque.lo;mouse.hi;mouse.lo	COMMON
k__Bacteria;p__Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Allobaculum;s__:3286:500051	control.lo;human.lo;macaque.lo;mouse.hi;mouse.lo	COMMON
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:9404:74476	control.lo;human.lo;macaque.lo;mouse.hi;mouse.lo	COMMON
k__Bacteria;p__Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Allobaculum;s__:10905:168768	control.lo;human.lo;macaque.lo;mouse.hi;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Helicobacteraceae;g__Helicobacter;s__:11806:414660	control.lo;human.lo;macaque.lo;mouse.hi;mouse.lo	COMMON
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides;s__acidifaciens:13591:108440	control.lo;human.lo;macaque.lo;mouse.hi;mouse.lo	COMMON
k__Bacteria;p__Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Allobaculum;s__:13725:345926	control.lo;human.lo;macaque.lo;mouse.hi;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Sphingomonadales;f__Sphingomonadaceae;g__Novosphingobium;s__:94:21638	control.lo;human.lo;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae:6162:34882	control.lo;human.lo;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__Burkholderia;s__tuberum:6540:244550	control.lo;human.lo;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae:6682:37035	control.lo;human.lo;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__Streptophyta;f__;g__;s__:11179:134884	control.lo;human.lo;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Ralstonia;s__:12532:29617	control.lo;human.lo;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Acinetobacter;s__:461:117467	human.lo;macaque.hi;macaque.lo	COMMON
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Corynebacteriaceae;g__Corynebacterium;s__:4900:48321	human.lo;macaque.hi;macaque.lo	COMMON
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__copri:5107:604167	human.lo;macaque.hi;macaque.lo	COMMON
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales:8326:37369	human.lo;macaque.hi;macaque.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Actinobacillus:8609:212188	human.lo;macaque.hi;macaque.lo	COMMON
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus;s__luteciae:10685:88857	human.lo;macaque.hi;macaque.lo	COMMON
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus;s__:13832:162303	human.lo;macaque.hi;macaque.lo	COMMON
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus;s__reuteri:8099:344364	human.lo;macaque.hi;macaque.lo;mouse.hi;mouse.lo	COMMON

k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus;s__:10679:91124	human.lo;macaque.hi;macaque.lo;mouse.hi;mouse.lo	COMMON
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus;s__:13960:37393	human.lo;macaque.hi;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Corynebacteriaceae;g__Corynebacterium;s__:4905:115727	human.lo;macaque.hi;mouse.hi	COMMON
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Pseudomonadaceae;g__Pseudomonas;s__fragi:102:24182	human.lo;macaque.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Acinetobacter:647:52274	human.lo;macaque.lo	COMMON
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Corynebacteriaceae;g__Corynebacterium;s__:1416:20020	human.lo;macaque.lo	COMMON
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Actinomycetaceae;g__Actinomyces;s__:2790:37509	human.lo;macaque.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Acinetobacter;s__lwoffii:3974:89871	human.lo;macaque.lo	COMMON
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus;s__:10919:78305	human.lo;macaque.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Acinetobacter;s__johnsonii:11137:24351	human.lo;macaque.lo	COMMON
k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__Streptophyta;f__;g__;s__:11178:406141	human.lo;macaque.lo	COMMON
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__;s__:11298:55012	human.lo;macaque.lo	COMMON
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Staphylococcaceae;g__Staphylococcus:11711:23793	human.lo;macaque.lo	COMMON
k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__Streptophyta;f__;g__;s__:14008:125962	human.lo;macaque.lo	COMMON
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__[Ruminococcus];s__gnavus:7432:20751	human.lo;macaque.lo;mouse.hi;mouse.lo	COMMON
k__Bacteria;p__Bacteroidetes;c__[Saprospirae];o__[Saprospirales];f__Chitinophagaceae;g__Sediminibacterium;s__:23:129801	human.lo;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Ralstonia;s__:183:22974	human.lo;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Acidobacteria;c__DA052;o__Ellin6513;f__;g__;s__:258:25097	human.lo;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Xanthomonadales;f__Xanthomonadaceae;g__Dyella;s__:644:30701	human.lo;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Cyanobacteria;c__ML635J-21;o__;f__;g__;s__:2244:432652	human.lo;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__Burkholderia:6541:101635	human.lo;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae:6683:91810	human.lo;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Burkholderiaceae;g__Salinispora;s__:6687:56926	human.lo;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Chlamydiae;c__Chlamydiia;o__Chlamydiales;f__Rhabdochlamydiaceae;g__Candidatus Rhabdochlamydia;s__:10353:34893	human.lo;macaque.lo;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Ralstonia;s__:110:23050	human.lo;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Oxalobacteraceae;g__Ralstonia;s__:6290:161902	human.lo;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pseudomonadales;f__Moraxellaceae;g__Acinetobacter:11138:28337	human.lo;mouse.lo	COMMON
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus;s__reuteri:8100:135371	macaque.hi;macaque.lo;mouse.hi;mouse.lo	COMMON
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Aggregatibacter:11937:1339456	macaque.hi;macaque.lo;mouse.hi;mouse.lo	COMMON

k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Aeromonadales;f__Succinivibrionaceae;g__Succinivibrio;s__:95:241741	control.lo;human.lo;macaque.hi;macaque.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales:1616:27218	control.lo;human.lo;macaque.lo;mouse.hi	SP

k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales:2584:70419	control.lo;human.lo;macaque.lo;mouse.hi	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__;s__:11294:172720	control.lo;human.lo;macaque.lo;mouse.hi	SP
k__Bacteria;p__Deferribacteres;c__Deferribacteres;o__Deferribacterales;f__Deferribacteraceae;g__Mucispirillum;s__schaedleri:12384:56730	control.lo;human.lo;macaque.lo;mouse.hi	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:4653:122914	control.lo;human.lo;macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Oscillospira;s__:10001:149849	control.lo;human.lo;macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria:10205:34081	control.lo;human.lo;macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales:12598:456227	control.lo;human.lo;macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Parabacteroides:12763:1654741	control.lo;human.lo;macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Bacilli;o__Bacillales;f__Staphylococcaceae;g__Staphylococcus:3399:217612	control.lo;macaque.hi;macaque.lo	SP
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus:12999:243472	control.lo;macaque.hi;macaque.lo	SP
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus:12836:56061	control.lo;macaque.lo;mouse.hi	SP
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus:8187:62731	control.lo;macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales:11242:38821	control.lo;macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Clostridium;s__citroniae:2591:80504	control.lo;mouse.hi	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Rikenellaceae;g__;s__:181:31645	control.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Desulfovibrionales;f__Desulfovibrionaceae;g__;s__:5623:490281	control.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides;s__:9223:46276	control.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Oscillospira;s__:10356:21451	control.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales:10545:59835	control.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Neisseriales;f__Neisseriaceae;g__Neisseria;s__:139:23811	human.lo;macaque.hi	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Porphyromonas;s__:263:37182	human.lo;macaque.hi;macaque.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Blautia;s__:4153:92769	human.lo;macaque.hi;macaque.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Blautia;s__obeum:6863:45419	human.lo;macaque.hi;macaque.lo	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__[Paraprevotellaceae];g__[Prevotella];s__:8717:46228	human.lo;macaque.hi;macaque.lo	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:9408:115568	human.lo;macaque.hi;macaque.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__Roseburia;s__faecis:11391:85872	human.lo;macaque.hi;macaque.lo	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__copri:12269:750861	human.lo;macaque.hi;macaque.lo	SP
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus;s__:13958:30976	human.lo;macaque.hi;macaque.lo	SP
k__Bacteria;p__Proteobacteria;c__Betaproteobacteria;o__Burkholderiales;f__Comamonadaceae:6153:50910	human.lo;macaque.hi;macaque.lo;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__;g__;s__:68:28261	human.lo;macaque.lo;mouse.hi	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Parabacteroides:7774:54879	human.lo;macaque.lo;mouse.hi	SP
k__Bacteria;p__Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Helicobacteraceae;g__Helicobacter:11808:30100	human.lo;macaque.lo;mouse.hi	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__[Ruminococcu	human.lo;macaque.lo;mouse.hi	SP

s];s__gnavus:13646:72389		
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:8726:53419	human.lo;macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus;s__:6408:106835	macaque.hi;macaque.lo;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Aerococcaceae;g__Aerococcus;s__:12921:202954	macaque.hi;macaque.lo;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus:143:27747	macaque.lo;mouse.hi	SP
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Actinomycetales;f__Corynebacteriaceae;g__Corynebacterium;s__:1419:25629	macaque.lo;mouse.hi	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Porphyromonadaceae;g__Parabacteroides:12767:25847	macaque.lo;mouse.hi	SP
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus:12997:19545	macaque.lo;mouse.hi	SP
k__Bacteria;p__Verrucomicrobia;c__Verrucomicrobiae;o__Verrucomicrobiales;f__Verrucomicrobiaceae;g__Akkermansia;s__muciniphila:50:92046	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__;g__;s__:87:42539	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Proteobacteria;c__Deltaproteobacteria;o__Desulfovibrionales;f__Desulfovibrionaceae;g__Desulfovibrio;s__C21_c20:106:118146	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Erysipelotrichi;o__Erysipelotrichales;f__Erysipelotrichaceae;g__Allobaculum;s__:171:29825	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales:1815:24652	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Lachnospiraceae;g__;s__:2034:18583	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__Adlercreutzia;s__:4652:42870	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Actinobacteria;c__Coriobacteriia;o__Coriobacteriales;f__Coriobacteriaceae;g__;s__:4654:142355	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:4661:26143	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:4662:36285	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:5954:19679	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus;s__:6407:112944	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Ruminococcus;s__:7431:31051	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:8697:31010	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:8698:27119	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:9403:28648	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Ruminococcaceae;g__Oscillospira;s__:10008:33385	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales:10544:62496	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__;g__;s__:11243:17889	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__Bifidobacterium;s__:11633:20006	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Actinobacteria;c__Actinobacteria;o__Bifidobacteriales;f__Bifidobacteriaceae;g__Bifidobacterium;s__:11635:73370	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:12078:53672	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:12080:31911	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae;g__SMB53;s__:12599:108429	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:13440:19902	macaque.lo;mouse.hi;mouse.lo	SP
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__S24-7;g__;s__:13443:80323	macaque.lo;mouse.hi;mouse.lo	SP

k__Bacteria:34:54126	--	RARE

k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rhizobiales;f__Methylobacteriaceae;g__Methylobacterium;s__:101:20505	--	RARE
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae:1099:197829	--	RARE
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae:1100:20515	--	RARE
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella:1191:34351	--	RARE
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Prevotellaceae;g__Prevotella;s__:1192:21086	--	RARE
k__Bacteria;p__Proteobacteria;c__Epsilonproteobacteria;o__Campylobacterales;f__Helicobacteraceae;g__Flexispira;s__:3552:126344	--	RARE
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Aggregatibacter:3717:618218	--	RARE
k__Bacteria;p__Firmicutes;c__Clostridia;o__Clostridiales;f__Clostridiaceae:6860:852753	--	RARE
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Enterococcaceae;g__Enterococcus;s__:8191:302853	--	RARE
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus:10683:19335	--	RARE
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Streptococcus;s__:10892:735909	--	RARE
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Streptococcaceae;g__Lactococcus;s__garvieae:10902:55881	--	RARE
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rickettsiales;f__mitochondria;g__Citrullus;s__lanatus:11852:31926	--	RARE
k__Bacteria;p__Proteobacteria;c__Alphaproteobacteria;o__Rickettsiales;f__mitochondria;g__Lupinus;s__luteus:11853:54339	--	RARE
k__Bacteria;p__Proteobacteria;c__Gammaproteobacteria;o__Pasteurellales;f__Pasteurellaceae;g__Aggregatibacter:11946:20986	--	RARE
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Aerococcaceae;g__Aerococcus;s__:12922:33219	--	RARE
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus;s__:13051:37386	--	RARE
k__Bacteria;p__Firmicutes;c__Bacilli;o__Lactobacillales;f__Lactobacillaceae;g__Lactobacillus;s__:13052:31906	--	RARE
k__Bacteria;p__Bacteroidetes;c__Bacteroidia;o__Bacteroidales;f__Bacteroidaceae;g__Bacteroides;s__fragilis:13592:44240	--	RARE
k__Bacteria;p__Cyanobacteria;c__Chloroplast;o__Streptophyta;f__;g__;s__:14007:22015	--	RARE

References

1. Eckburg, P.B., E.M. Bik, C.N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S.R. Gill, K.E. Nelson, and D.A. Relman, ***Diversity of the human intestinal microbial flora.*** Science, 2005. **308**(5728): p. 1635-8.

2. Backhed, F., R.E. Ley, J.L. Sonnenburg, D.A. Peterson, and J.I. Gordon, ***Host-bacterial mutualism in the human intestine.*** Science, 2005. **307**(5717): p. 1915-20.

3. Rooks, M.G. and W.S. Garrett, ***Gut microbiota, metabolites and host immunity.*** Nat Rev Immunol, 2016. **16**(6): p. 341-52.

4. Ipci, K., N. Altintoprak, N.B. Muluk, M. Senturk, and C. Cingi, ***The possible mechanisms of the human microbiome in allergic diseases.*** Eur Arch Otorhinolaryngol, 2017. **274**(2): p. 617-626.

5. Human Microbiome Project, C., ***Structure, function and diversity of the healthy human microbiome.*** Nature, 2012. **486**(7402): p. 207-14.

6. Lloyd-Price, J., G. Abu-Ali, and C. Huttenhower, ***The healthy human microbiome.*** Genome Med, 2016. **8**(1): p. 51.

7. Turnbaugh, P.J., R.E. Ley, M.A. Mahowald, V. Magrini, E.R. Mardis, and J.I. Gordon, ***An obesity-associated gut microbiome with increased capacity for energy harvest.*** Nature, 2006. **444**(7122): p. 1027-31.

8. Kau, A.L., P.P. Ahern, N.W. Griffin, A.L. Goodman, and J.I. Gordon, ***Human nutrition, the gut microbiome and the immune system.*** Nature, 2011. **474**(7351): p. 327-36.

9. Smith, P.M., M.R. Howitt, N. Panikov, M. Michaud, C.A. Gallini, Y.M. Bohlooly, J.N. Glickman, and W.S. Garrett, ***The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis.*** Science, 2013. **341**(6145): p. 569-73.

10. Mazmanian, S.K., J.L. Round, and D.L. Kasper, ***A microbial symbiosis factor prevents intestinal inflammatory disease.*** Nature, 2008. **453**(7195): p. 620-5.

11. Chung, H., S.J. Pamp, J.A. Hill, N.K. Surana, S.M. Edelman, E.B. Troy, N.C. Reading, E.J. Villablanca, S. Wang, J.R. Mora, et al., ***Gut immune maturation depends on colonization with a host-specific microbiota.*** Cell, 2012. **149**(7): p. 1578-93.

12. Ivanov, II, K. Atarashi, N. Manel, E.L. Brodie, T. Shima, U. Karaoz, D. Wei, K.C. Goldfarb, C.A. Santee, S.V. Lynch, et al., ***Induction of intestinal Th17 cells by segmented filamentous bacteria.*** Cell, 2009. **139**(3): p. 485-98.

13. Theriot, C.M., M.J. Koenigsknecht, P.E. Carlson, Jr., G.E. Hatton, A.M. Nelson, B. Li, G.B. Huffnagle, Z.L. J, and V.B. Young, ***Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to Clostridium difficile infection.*** Nat Commun, 2014. **5**: p. 3114.

14. van Nood, E., A. Vrieze, M. Nieuwdorp, S. Fuentes, E.G. Zoetendal, W.M. de Vos, C.E. Visser, E.J. Kuijper, J.F. Bartelsman, J.G. Tijssen, et al., ***Duodenal infusion of donor feces for recurrent Clostridium difficile.*** N Engl J Med, 2013. **368**(5): p. 407-15.

15. Morgan, X.C., T.L. Tickle, H. Sokol, D. Gevers, K.L. Devaney, D.V. Ward, J.A. Reyes, S.A. Shah, N. LeLeiko, S.B. Snapper, et al., ***Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment.*** Genome Biol, 2012. **13**(9): p. R79.

16. Gevers, D., S. Kugathasan, L.A. Denson, Y. Vazquez-Baeza, W. Van Treuren, B. Ren, E. Schwager, D. Knights, S.J. Song, M. Yassour, et al., ***The treatment-naive microbiome in new-onset Crohn's disease.*** Cell Host Microbe, 2014. **15**(3): p. 382-92.

17. Rooks, M.G., P. Veiga, L.H. Wardwell-Scott, T. Tickle, N. Segata, M. Michaud, C.A. Gallini, C. Beal, J.E. van Hylckama-Vlieg, S.A. Ballal, et al., ***Gut microbiome composition and function in experimental colitis during active disease and treatment-induced remission.*** ISME J, 2014. **8**(7): p. 1403-17.

18. Parsonnet, J., G.D. Friedman, D.P. Vandersteen, Y. Chang, J.H. Vogelman, N. Orentreich, and R.K. Sibley, ***Helicobacter pylori infection and the risk of gastric carcinoma.*** N Engl J Med, 1991. **325**(16): p. 1127-31.

19. Kostic, A.D., E. Chun, L. Robertson, J.N. Glickman, C.A. Gallini, M. Michaud, T.E. Clancy, D.C. Chung, P. Lochhead, G.L. Hold, et al., ***Fusobacterium nucleatum potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment.*** Cell Host Microbe, 2013. **14**(2): p. 207-15.

20. Kostic, A.D., D. Gevers, C.S. Pedamallu, M. Michaud, F. Duke, A.M. Earl, A.I. Ojesina, J. Jung, A.J. Bass, J. Tabernero, et al., ***Genomic analysis identifies association of Fusobacterium with colorectal carcinoma.*** Genome Res, 2012. **22**(2): p. 292-8.

21. Castellarin, M., R.L. Warren, J.D. Freeman, L. Dreolini, M. Krzywinski, J. Strauss, R. Barnes, P. Watson, E. Allen-Vercoe, R.A. Moore, et al., ***Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma.*** Genome Res, 2012. **22**(2): p. 299-306.

22. Vatanen, T., A.D. Kostic, E. d'Hennezel, H. Siljander, E.A. Franzosa, M. Yassour, R. Kolde, H. Vlamakis, T.D. Arthur, A.M. Hamalainen, et al., ***Variation in Microbiome LPS Immunogenicity Contributes to Autoimmunity in Humans.*** Cell, 2016. **165**(4): p. 842-53.

23. Scher, J.U., A. Sczesnak, R.S. Longman, N. Segata, C. Ubeda, C. Bielski, T. Rostron, V. Cerundolo, E.G. Pamer, S.B. Abramson, et al., ***Expansion of intestinal Prevotella copri correlates with enhanced susceptibility to arthritis.*** eLife, 2013. **2**: p. e01202.

24. Wu, S., J. Yi, Y.G. Zhang, J. Zhou, and J. Sun, ***Leaky intestine and impaired microbiome in an amyotrophic lateral sclerosis mouse model.*** Physiol Rep, 2015. **3**(4).

25. Bonaz, B.L. and C.N. Bernstein, ***Brain-gut interactions in inflammatory bowel disease.*** Gastroenterology, 2013. **144**(1): p. 36-49.

26. Foster, J.A. and K.A. McVey Neufeld, ***Gut-brain axis: how the microbiome influences anxiety and depression.*** Trends Neurosci, 2013. **36**(5): p. 305-12.

27. Dethlefsen, L. and D.A. Relman, ***Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation.*** Proc Natl Acad Sci U S A, 2011. **108 Suppl 1**: p. 4554-61.

28. Buffie, C.G., I. Jarchum, M. Equinda, L. Lipuma, A. Gobourne, A. Viale, C. Ubeda, J. Xavier, and E.G. Pamer, ***Profound alterations of intestinal microbiota following a single dose of clindamycin results in sustained susceptibility to Clostridium difficile-induced colitis.*** Infect Immun, 2012. **80**(1): p. 62-73.

29. Yassour, M., T. Vatanen, H. Siljander, A.M. Hamalainen, T. Harkonen, S.J. Ryhanen, E.A. Franzosa, H. Vlamakis, C. Huttenhower, D. Gevers, et al., ***Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability.*** Sci Transl Med, 2016. **8**(343): p. 343ra81.

30. Everard, A., V. Lazarevic, M. Derrien, M. Girard, G.G. Muccioli, A.M. Neyrinck, S. Possemiers, A. Van Holle, P. Francois, W.M. de Vos, et al., ***Responses of gut microbiota and glucose and lipid metabolism to prebiotics in genetic obese and diet-induced leptin-resistant mice.*** Diabetes, 2011. **60**(11): p. 2775-86.

31. Everard, A., V. Lazarevic, N. Gaia, M. Johansson, M. Stahlman, F. Backhed, N.M. Delzenne, J. Schrenzel, P. Francois, and P.D. Cani, ***Microbiome of prebiotic-treated mice reveals novel targets involved in host response during obesity.*** ISME J, 2014. **8**(10): p. 2116-30.

32. Sonnenburg, E.D., S.A. Smits, M. Tikhonov, S.K. Higginbottom, N.S. Wingreen, and J.L. Sonnenburg, ***Diet-induced extinctions in the gut microbiota compound over generations.*** Nature, 2016. **529**(7585): p. 212-5.

33. David, L.A., C.F. Maurice, R.N. Carmody, D.B. Gootenberg, J.E. Button, B.E. Wolfe, A.V. Ling, A.S. Devlin, Y. Varma, M.A. Fischbach, et al., ***Diet rapidly and reproducibly alters the human gut microbiome.*** Nature, 2014. **505**(7484): p. 559-63.

34. Turnbaugh, P.J., V.K. Ridaura, J.J. Faith, F.E. Rey, R. Knight, and J.I. Gordon, ***The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice.*** Sci Transl Med, 2009. **1**(6): p. 6ra14.

35. Poole, A.C., L. Pischel, C. Ley, G. Suh, J.K. Goodrich, T.D. Haggerty, R.E. Ley, and J. Parsonnet, ***Crossover Control Study of the Effect of Personal Care Products Containing Triclosan on the Microbiome.*** mSphere, 2016. **1**(3).

36. Hartmann, E.M., R. Hickey, T. Hsu, C.M. Betancourt Roman, J. Chen, R. Schwager, J. Kline, G.Z. Brown, R.U. Halden, C. Huttenhower, et al., ***Antimicrobial Chemicals Are Associated with Elevated Antibiotic Resistance Genes in the Indoor Dust Microbiome.*** Environ Sci Technol, 2016. **50**(18): p. 9807-15.

37. Lu, K., R. Mahbub, P.H. Cable, H. Ru, N.M. Parry, W.M. Bodnar, J.S. Wishnok, M. Styblo, J.A. Swenberg, J.G. Fox, et al., ***Gut microbiome phenotypes driven by host genetics affect arsenic metabolism.*** Chem Res Toxicol, 2014. **27**(2): p. 172-4.

38. Zhang, L.J., C.F. Guerrero-Juarez, T. Hata, S.P. Bapat, R. Ramos, M.V. Plikus, and R.L. Gallo, *Innate immunity.* ***Dermal adipocytes protect against invasive Staphylococcus aureus skin infection.*** Science, 2015. **347**(6217): p. 67-71.

39. Nakatsuji, T., H.I. Chiang, S.B. Jiang, H. Nagarajan, K. Zengler, and R.L. Gallo, ***The microbiome extends to subepidermal compartments of normal skin.*** Nat Commun, 2013. **4**: p. 1431.

40. Zulian, A., R. Cancello, C. Ruocco, D. Gentilini, A.M. Di Blasio, P. Danelli, G. Micheletto, E. Cesana, and C. Invitti, ***Differences in visceral fat and fat bacterial colonization between ulcerative colitis and Crohn's disease. An in vivo and in vitro study.*** PLoS One, 2013. **8**(10): p. e78495.

41. Fallingborg, J., ***Intraluminal pH of the human gastrointestinal tract.*** Dan Med Bull, 1999. **46**(3): p. 183-96.

42. Sayin, S.I., A. Wahlstrom, J. Felin, S. Jantti, H.U. Marschall, K. Bamberg, B. Angelin, T. Hyotylainen, M. Oresic, and F. Backhed, ***Gut microbiota regulates bile acid metabolism by reducing the levels of tauro-beta-muricholic acid, a naturally occurring FXR antagonist.*** Cell Metab, 2013. **17**(2): p. 225-35.

43. Devkota, S., Y. Wang, M.W. Musch, V. Leone, H. Fehlner-Peach, A. Nadimpalli, D.A. Antonopoulos, B. Jabri, and E.B. Chang, ***Dietary-fat-induced taurocholic acid promotes pathobiont expansion and colitis in Il10-/- mice.*** Nature, 2012. **487**(7405): p. 104-8.

44. Goodrich, J.K., E.R. Davenport, M. Beaumont, M.A. Jackson, R. Knight, C. Ober, T.D. Spector, J.T. Bell, A.G. Clark, and R.E. Ley, ***Genetic Determinants of the Gut Microbiome in UK Twins.*** Cell Host Microbe, 2016. **19**(5): p. 731-43.

45. Donaldson, G.P., S.M. Lee, and S.K. Mazmanian, ***Gut biogeography of the bacterial microbiota.*** Nat Rev Microbiol, 2016. **14**(1): p. 20-32.

46. Albenberg, L., T.V. Esipova, C.P. Judge, K. Bittinger, J. Chen, A. Laughlin, S. Grunberg, R.N. Baldassano, J.D. Lewis, H. Li, et al., ***Correlation Between Intraluminal Oxygen Gradient and Radial Partitioning of Intestinal Microbiota.*** Gastroenterology, 2014.

47. Yasuda, K., K. Oh, B. Ren, T.L. Tickle, E.A. Franzosa, L.M. Wachtman, A.D. Miller, S.V. Westmoreland, K.G. Mansfield, E.J. Vallender, et al., ***Biogeography of the intestinal mucosal and lumenal microbiome in the rhesus macaque.*** Cell Host Microbe, 2015. **17**(3): p. 385-91.

48. Round, J.L. and S.K. Mazmanian, ***The gut microbiota shapes intestinal immune responses during health and disease.*** Nat Rev Immunol, 2009. **9**(5): p. 313-23.

49. Graessler, J., Y. Qin, H. Zhong, J. Zhang, J. Licinio, M.L. Wong, A. Xu, T. Chavakis, A.B. Bornstein, M. Ehrhart-Bornstein, et al., ***Metagenomic sequencing of the human gut microbiome before and after bariatric surgery in obese patients with type 2 diabetes: correlation with inflammatory and metabolic parameters.*** Pharmacogenomics J, 2013. **13**(6): p. 514-22.

50. Li, J.V., H. Ashrafian, M. Bueter, J. Kinross, C. Sands, C.W. le Roux, S.R. Bloom, A. Darzi, T. Athanasiou, J.R. Marchesi, et al., ***Metabolic surgery profoundly influences gut microbial-host metabolic cross-talk.*** Gut, 2011. **60**(9): p. 1214-23.

51. Furet, J.P., L.C. Kong, J. Tap, C. Poitou, A. Basdevant, J.L. Bouillot, D. Mariat, G. Corthier, J. Dore, C. Henegar, et al., ***Differential adaptation of human gut microbiota to bariatric surgery-induced weight loss: links with metabolic and low-grade inflammation markers.*** Diabetes, 2010. **59**(12): p. 3049-57.

52. Tremaroli, V., F. Karlsson, M. Werling, M. Stahlman, P. Kovatcheva-Datchary, T. Olbers, L. Fandriks, C.W. le Roux, J. Nielsen, and F. Backhed, ***Roux-en-Y Gastric Bypass and Vertical Banded Gastroplasty Induce Long-Term Changes on the Human Gut Microbiome Contributing to Fat Mass Regulation.*** Cell Metab, 2015. **22**(2): p. 228-38.

53. Everard, A., C. Belzer, L. Geurts, J.P. Ouwerkerk, C. Druart, L.B. Bindels, Y. Guiot, M. Derrien, G.G. Muccioli, N.M. Delzenne, et al., ***Cross-talk between Akkermansia muciniphila and intestinal epithelium controls diet-induced obesity.*** Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(22): p. 9066-71.

54. Li, H., J.P. Limenitakis, T. Fuhrer, M.B. Geuking, M.A. Lawson, M. Wyss, S. Brugiroux, I. Keller, J.A. Macpherson, S. Rupp, et al., ***The outer mucus layer hosts a distinct intestinal microbial niche.*** Nat Commun, 2015. **6**: p. 8292.

55. Johansson, M.E., H. Sjovall, and G.C. Hansson, ***The gastrointestinal mucus system in health and disease.*** Nat Rev Gastroenterol Hepatol, 2013. **10**(6): p. 352-61.

56. Sonnenburg, J.L., J. Xu, D.D. Leip, C.H. Chen, B.P. Westover, J. Weatherford, J.D. Buhler, and J.I. Gordon, ***Glycan foraging in vivo by an intestine-adapted bacterial symbiont.*** Science, 2005. **307**(5717): p. 1955-9.

57. Kashyap, P.C., A. Marcobal, L.K. Ursell, S.A. Smits, E.D. Sonnenburg, E.K. Costello, S.K. Higginbottom, S.E. Domino, S.P. Holmes, D.A. Relman, et al., ***Genetically dictated change in host mucus carbohydrate landscape exerts a diet-dependent effect on the gut microbiota.*** Proc Natl Acad Sci U S A, 2013. **110**(42): p. 17059-64.

58. Marcobal, A., M. Barboza, E.D. Sonnenburg, N. Pudlo, E.C. Martens, P. Desai, C.B. Lebrilla, B.C. Weimer, D.A. Mills, J.B. German, et al., ***Bacteroides in the infant gut consume milk oligosaccharides via mucus-utilization pathways.*** Cell Host Microbe, 2011. **10**(5): p. 507-14.

59. Wexler, A.G., Y. Bao, J.C. Whitney, L.M. Bobay, J.B. Xavier, W.B. Schofield, N.A. Barry, A.B. Russell, B.Q. Tran, Y.A. Goo, et al., ***Human symbionts inject and neutralize antibacterial toxins to persist in the gut.*** Proc Natl Acad Sci U S A, 2016. **113**(13): p. 3639-44.

60. Russell, A.B., A.G. Wexler, B.N. Harding, J.C. Whitney, A.J. Bohn, Y.A. Goo, B.Q. Tran, N.A. Barry, H. Zheng, S.B. Peterson, et al., ***A type VI secretion-related pathway in Bacteroidetes mediates interbacterial antagonism.*** Cell Host Microbe, 2014. **16**(2): p. 227-36.

61. Cassat, J.E. and E.P. Skaar, ***Iron in infection and immunity.*** Cell Host Microbe, 2013. **13**(5): p. 509-19.

62. Gao, B., H. Vorwerk, C. Huber, M. Lara-Tejero, J. Mohr, A.L. Goodman, W. Eisenreich, J.E. Galan, and D. Hofreuter, ***Metabolic and fitness determinants for in vitro growth and intestinal colonization of the bacterial pathogen Campylobacter jejuni.*** PLoS Biol, 2017. **15**(5): p. e2001390.

63. Thiemann, S., N. Smit, U. Roy, T.R. Lesker, E.J.C. Galvez, J. Helmecke, M. Basic, A. Bleich, A.L. Goodman, U. Kalinke, et al., ***Enhancement of IFNgamma Production by Distinct Commensals Ameliorates Salmonella-Induced Disease.*** Cell Host Microbe, 2017. **21**(6): p. 682-694 e5.

64. Cullen, T.W., W.B. Schofield, N.A. Barry, E.E. Putnam, E.A. Rundell, M.S. Trent, P.H. Degnan, C.J. Booth, H. Yu, and A.L. Goodman, ***Gut microbiota. Antimicrobial peptide resistance mediates resilience of prominent gut commensals during inflammation.*** Science, 2015. **347**(6218): p. 170-5.

65. Franzosa, E.A., K. Huang, J.F. Meadow, D. Gevers, K.P. Lemon, B.J. Bohannan, and C. Huttenhower, ***Identifying personal microbiomes using metagenomic codes.*** Proc Natl Acad Sci U S A, 2015. **112**(22): p. E2930-8.

66. Dethlefsen, L., S. Huse, M.L. Sogin, and D.A. Relman, ***The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing***. PLoS Biol, 2008. **6**(11): p. e280.

67. Kumar, R., N. Yi, D. Zhi, P. Eipers, K.T. Goldsmith, P. Dixon, D.K. Crossman, M.R. Crowley, E.J. Lefkowitz, J.M. Rodriguez, et al., ***Identification of donor microbe species that colonize and persist long term in the recipient after fecal transplant for recurrent Clostridium difficile.*** NPJ Biofilms Microbiomes, 2017. **3**: p. 12.

68. Yasuda, K., T. Hsu, C.A. Gallini, L.J. McLver, E. Schwager, A. Shi, C.R. DuLong, R.N. Schwager, G.S. Abu-Ali, E.A. Franzosa, et al., ***Fluoride Depletes Acidogenic Taxa in Oral but Not Gut Microbial Communities in Mice.*** mSystems, 2017. **2**(4).

69. Kostic, A.D., M.R. Howitt, and W.S. Garrett, ***Exploring host-microbiota interactions in animal models and humans.*** Genes Dev, 2013. **27**(7): p. 701-18.

70. McKenna, P., C. Hoffmann, N. Minkah, P.P. Aye, A. Lackner, Z. Liu, C.A. Lozupone, M. Hamady, R. Knight, and F.D. Bushman, ***The macaque gut microbiome in health, lentiviral infection, and chronic enterocolitis.*** PLoS Pathog, 2008. **4**(2): p. e20.

71. Franzosa, E.A., X.C. Morgan, N. Segata, L. Waldron, J. Reyes, A.M. Earl, G. Giannoukos, M.R. Boylan, D. Ciulla, D. Gevers, et al., ***Relating the metatranscriptome and metagenome of the human gut.*** Proc Natl Acad Sci U S A, 2014. **111**(22): p. E2329-38.

72. Griffin, S.O., J.A. Jones, D. Brunson, P.M. Griffin, and W.D. Bailey, ***Burden of oral disease among older adults and implications for public health priorities.*** Am J Public Health, 2012. **102**(3): p. 411-8.

73. Dewhirst, F.E., T. Chen, J. Izard, B.J. Paster, A.C. Tanner, W.H. Yu, A. Lakshmanan, and W.G. Wade, ***The human oral microbiome***. J Bacteriol, 2010. **192**(19): p. 5002-17.

74. Wade, W.G., ***The oral microbiome in health and disease.*** Pharmacol Res, 2013. **69**(1): p. 137-43.

75. Kroes, I., P.W. Lepp, and D.A. Relman, ***Bacterial diversity within the human subgingival crevice.*** Proc Natl Acad Sci U S A, 1999. **96**(25): p. 14547-52.

76. Paster, B.J., S.K. Boches, J.L. Galvin, R.E. Ericson, C.N. Lau, V.A. Levanos, A. Sahasrabudhe, and F.E. Dewhirst, ***Bacterial diversity in human subgingival plaque.*** J Bacteriol, 2001. **183**(12): p. 3770-83.

77. Avila, M., D.M. Ojcius, and O. Yilmaz, ***The oral microbiota: living with a permanent guest.*** DNA Cell Biol, 2009. **28**(8): p. 405-11.

78. Mark Welch, J.L., B.J. Rossetti, C.W. Rieken, F.E. Dewhirst, and G.G. Borisy, ***Biogeography of a human oral microbiome at the micron scale.*** Proc Natl Acad Sci U S A, 2016. **113**(6): p. E791-800.

79. Donlan, R.M. and J.W. Costerton, ***Biofilms: survival mechanisms of clinically relevant microorganisms.*** Clin Microbiol Rev, 2002. **15**(2): p. 167-93.

80. Struzycka, I., ***The oral microbiome in dental caries.*** Pol J Microbiol, 2014. **63**(2): p. 127-35.

81. Nyvad, B., W. Crielaard, A. Mira, N. Takahashi, and D. Beighton, ***Dental caries from a molecular microbiological perspective.*** Caries Res, 2013. **47**(2): p. 89-102.

82. Docktor, M.J., B.J. Paster, S. Abramowicz, J. Ingram, Y.E. Wang, M. Correll, H. Jiang, S.L. Cotton, A.S. Kokaras, and A. Bousvaros, ***Alterations in diversity of the oral microbiome in pediatric inflammatory bowel disease.*** Inflamm Bowel Dis, 2012. **18**(5): p. 935-42.

83. Rautava, J., L.J. Pinnell, L. Vong, N. Akseer, A. Assa, and P.M. Sherman, ***Oral microbiome composition changes in mouse models of colitis.*** J Gastroenterol Hepatol, 2015. **30**(3): p. 521-7.

84. Said, H.S., W. Suda, S. Nakagome, H. Chinen, K. Oshima, S. Kim, R. Kimura, A. Iraha, H. Ishida, J. Fujita, et al., ***Dysbiosis of salivary microbiota in inflammatory bowel disease and its association with oral immunological biomarkers.*** DNA Res, 2014. **21**(1): p. 15-25.

85. Michaud, D.S. and J. Izard, ***Microbiota, oral microbiome, and pancreatic cancer.*** Cancer J, 2014. **20**(3): p. 203-6.

86. Slocum, C., C. Kramer, and C.A. Genco, ***Immune dysregulation mediated by the oral microbiome: potential link to chronic inflammation and atherosclerosis.*** J Intern Med, 2016. **280**(1): p. 114-28.

87. Fernandez-Feo, M., G. Wei, G. Blumenkranz, F.E. Dewhirst, D. Schuppan, F.G. Oppenheim, and E.J. Helmerhorst, ***The cultivable human oral gluten-degrading microbiome and its potential implications in coeliac disease and gluten sensitivity.*** Clin Microbiol Infect, 2013. **19**(9): p. E386-94.

88. Lionetti, E., S. Castellaneta, R. Francavilla, A. Pulvirenti, E. Tonutti, S. Amarri, M. Barbato, C. Barbera, G. Barera, A. Bellantoni, et al., ***Introduction of gluten, HLA status, and the risk of celiac disease in children.*** N Engl J Med, 2014. **371**(14): p. 1295-303.

89. Bik, E.M., P.B. Eckburg, S.R. Gill, K.E. Nelson, E.A. Purdom, F. Francois, G. Perez-Perez, M.J. Blaser, and D.A. Relman, ***Molecular analysis of the bacterial microbiota in the human stomach.*** Proc Natl Acad Sci U S A, 2006. **103**(3): p. 732-7.

90. Beasley, D.E., A.M. Koltz, J.E. Lambert, N. Fierer, and R.R. Dunn, ***The Evolution of Stomach Acidity and Its Relevance to the Human Microbiome.*** PLoS One, 2015. **10**(7): p. e0134116.

91. Franklin, M.A. and S.C. Skoryna, ***Studies on natural gastric flora: survival of bacteria in fasting human subjects.*** Can Med Assoc J, 1971. **105**(4): p. 380-6.

92. Franklin, M.A. and S.C. Skoryna, ***Studies on natural gastric flora. I. Bacterial flora of fasting human subjects.*** Can Med Assoc J, 1966. **95**(26): p. 1349-55.

93. Wen, Y., E.A. Marcus, U. Matrubutham, M.A. Gleeson, D.R. Scott, and G. Sachs, ***Acid-adaptive genes of Helicobacter pylori.*** Infect Immun, 2003. **71**(10): p. 5921-39.

94. Marshall, B.J. and J.R. Warren, ***Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration.*** Lancet, 1984. **1**(8390): p. 1311-5.

95. Peterson, W.L., ***Helicobacter pylori and peptic ulcer disease.*** N Engl J Med, 1991. **324**(15): p. 1043-8.

96. Kusters, J.G., A.H. van Vliet, and E.J. Kuipers, ***Pathogenesis of Helicobacter pylori infection.*** Clin Microbiol Rev, 2006. **19**(3): p. 449-90.

97. Wroblewski, L.E., R.M. Peek, Jr., and K.T. Wilson, ***Helicobacter pylori and gastric cancer: factors that modulate disease risk.*** Clin Microbiol Rev, 2010. **23**(4): p. 713-39.

98. Blaser, M.J., ***Helicobacter pylori and gastric diseases.*** BMJ, 1998. **316**(7143): p. 1507-10.

99. Conway, P.L., S.L. Gorbach, and B.R. Goldin, ***Survival of lactic acid bacteria in the human stomach and adhesion to intestinal cells.*** J Dairy Sci, 1987. **70**(1): p. 1-12.

100. Stearns, J.C., M.D. Lynch, D.B. Senadheera, H.C. Tenenbaum, M.B. Goldberg, D.G. Cvitkovitch, K. Croitoru, G. Moreno-Hagelsieb, and J.D. Neufeld, ***Bacterial biogeography of the human digestive tract.*** Sci Rep, 2011. **1**: p. 170.

101. Hartman, A.L., D.M. Lough, D.K. Barupal, O. Fiehn, T. Fishbein, M. Zasloff, and J.A. Eisen, ***Human gut microbiome adopts an alternative state following small bowel transplantation.*** Proc Natl Acad Sci U S A, 2009. **106**(40): p. 17187-92.

102. Qin, J., R. Li, J. Raes, M. Arumugam, K.S. Burgdorf, C. Manichanh, T. Nielsen, N. Pons, F. Levenez, T. Yamada, et al., ***A human gut microbial gene catalogue established by metagenomic sequencing.*** Nature, 2010. **464**(7285): p. 59-65.

103. Koren, O., D. Knights, A. Gonzalez, L. Waldron, N. Segata, R. Knight, C. Huttenhower, and R.E. Ley, ***A guide to enterotypes across the human body: meta-analysis of microbial community structures in human microbiome datasets.*** PLoS Comput Biol, 2013. **9**(1): p. e1002863.

104. Knights, D., T.L. Ward, C.E. McKinlay, H. Miller, A. Gonzalez, D. McDonald, and R. Knight, ***Rethinking "enterotypes".*** Cell Host Microbe, 2014. **16**(4): p. 433-7.

105. Bowler, P.G., B.I. Duerden, and D.G. Armstrong, ***Wound microbiology and associated approaches to wound management.*** Clin Microbiol Rev, 2001. **14**(2): p. 244-69.

106. Norman, J.M., S.A. Handley, M.T. Baldridge, L. Droit, C.Y. Liu, B.C. Keller, A. Kambal, C.L. Monaco, G. Zhao, P. Fleshner, et al., ***Disease-specific alterations in the enteric virome in inflammatory bowel disease.*** Cell, 2015. **160**(3): p. 447-60.

107. Fortune, S.M., A. Solache, A. Jaeger, P.J. Hill, J.T. Belisle, B.R. Bloom, E.J. Rubin, and J.D. Ernst, ***Mycobacterium tuberculosis inhibits macrophage responses to IFN-gamma through myeloid differentiation factor 88-dependent and -independent mechanisms.*** J Immunol, 2004. **172**(10): p. 6272-80.

108. Rengarajan, J., B.R. Bloom, and E.J. Rubin, ***Genome-wide requirements for Mycobacterium tuberculosis adaptation and survival in macrophages.*** Proc Natl Acad Sci U S A, 2005. **102**(23): p. 8327-32.

109. Aagaard, K., J. Ma, K.M. Antony, R. Ganu, J. Petrosino, and J. Versalovic, ***The placenta harbors a unique microbiome.*** Sci Transl Med, 2014. **6**(237): p. 237ra65.

110. Sze, M.A., P.A. Dimitriu, S. Hayashi, W.M. Elliott, J.E. McDonough, J.V. Gosselink, J. Cooper, D.D. Sin, W.W. Mohn, and J.C. Hogg, ***The lung tissue microbiome in chronic obstructive pulmonary disease.*** Am J Respir Crit Care Med, 2012. **185**(10): p. 1073-80.

111. Chng, K.R., S.H. Chan, A.H. Ng, C. Li, A. Jusakul, D. Bertrand, A. Wilm, S.P. Choo, D.M. Tan, K.H. Lim, et al., ***Tissue Microbiome Profiling Identifies an Enrichment of Specific Enteric Bacteria in Opisthorchis viverrini Associated Cholangiocarcinoma.*** EBioMedicine, 2016. **8**: p. 195-202.

112. Kumar, D.K., S.H. Choi, K.J. Washicosky, W.A. Eimer, S. Tucker, J. Ghofrani, A. Lefkowitz, G. McColl, L.E. Goldstein, R.E. Tanzi, et al., ***Amyloid-beta peptide protects against microbial infection in mouse and worm models of Alzheimer's disease.*** Sci Transl Med, 2016. **8**(340): p. 340ra72.

113. Urbaniak, C., J. Cummins, M. Brackstone, J.M. Macklaim, G.B. Gloor, C.K. Baban, L. Scott, D.M. O'Hanlon, J.P. Burton, K.P. Francis, et al., ***Microbiota of human breast tissue.*** Appl Environ Microbiol, 2014. **80**(10): p. 3007-14.

114. Glassing, A., S.E. Dowd, S. Galandiuk, B. Davis, and R.J. Chiodini, ***Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples.*** Gut Pathog, 2016. **8**: p. 24.

115. Salter, S.J., M.J. Cox, E.M. Turek, S.T. Calus, W.O. Cookson, M.F. Moffatt, P. Turner, J. Parkhill, N.J. Loman, and A.W. Walker, ***Reagent and laboratory contamination can critically impact sequence-based microbiome analyses.*** BMC Biol, 2014. **12**: p. 87.

116. Batjer, J.D., L.J. Williamson, L. Polissar, and W.B. Hamlin, ***Effects of bacterial contamination of reagent water on selected laboratory tests.*** Am J Clin Pathol, 1979. **71**(3): p. 319-25.

117. Laurence, M., C. Hatzis, and D.E. Brash, ***Common contaminants in next-generation sequencing that hinder discovery of low-abundance microbes.*** PLoS One, 2014. **9**(5): p. e97876.

118. Weiss, S., A. Amir, E.R. Hyde, J.L. Metcalf, S.J. Song, and R. Knight, ***Tracking down the sources of experimental contamination in microbiome studies.*** Genome Biol, 2014. **15**(12): p. 564.

119. Lauder, A.P., A.M. Roche, S. Sherrill-Mix, A. Bailey, A.L. Laughlin, K. Bittinger, R. Leite, M.A. Elovitz, S. Parry, and F.D. Bushman, ***Comparison of placenta samples with contamination controls does not provide evidence for a distinct placenta microbiota.*** Microbiome, 2016. **4**(1): p. 29.

120. Adams, R.I., A.C. Bateman, H.M. Bik, and J.F. Meadow, ***Microbiota of the indoor environment: a meta-analysis.*** Microbiome, 2015. **3**: p. 49.

121. Czurda, S., S. Smelik, S. Preuner-Stix, F. Nogueira, and T. Lion, ***Occurrence of Fungal DNA Contamination in PCR Reagents: Approaches to Control and Decontamination.*** J Clin Microbiol, 2016. **54**(1): p. 148-52.

122. Moustafa, A., C. Xie, E. Kirkness, W. Biggs, E. Wong, Y. Turpaz, K. Bloom, E. Delwart, K.E. Nelson, J.C. Venter, et al., ***The blood DNA virome in 8,000 humans.*** PLoS Pathog, 2017. **13**(3): p. e1006292.

123. Merchant, S., D.E. Wood, and S.L. Salzberg, ***Unexpected cross-species contamination in genome sequencing projects.*** PeerJ, 2014. **2**: p. e675.

124. Longo, M.S., M.J. O'Neill, and R.J. O'Neill, ***Abundant human DNA contamination identified in non-primate genome databases.*** PLoS One, 2011. **6**(2): p. e16410.

125. Kruger, M., A.A. Shehata, W. Schrodl, and A. Rodloff, ***Glyphosate suppresses the antagonistic effect of Enterococcus spp. on Clostridium botulinum.*** Anaerobe, 2013. **20**: p. 74-8.

126. Shehata, A.A., W. Schrodl, A.A. Aldin, H.M. Hafez, and M. Kruger, ***The effect of glyphosate on potential pathogens and beneficial members of poultry microbiota in vitro.*** Curr Microbiol, 2013. **66**(4): p. 350-8.

127. Breton, J., S. Massart, P. Vandamme, E. De Brandt, B. Pot, and B. Foligne, ***Ecotoxicology inside the gut: impact of heavy metals on the mouse microbiome.*** BMC Pharmacol Toxicol, 2013. **14**: p. 62.

128. Marquis, R.E., ***Antimicrobial actions of fluoride for oral bacteria.*** Can J Microbiol, 1995. **41**(11): p. 955-64.

129. Qin, J., G. Chai, J.M. Brewer, L.L. Lovelace, and L. Lebioda, ***Fluoride inhibition of enolase: crystal structure and thermodynamics***. Biochemistry, 2006. **45**(3): p. 793-800.

130. Sutton, S.V., G.R. Bender, and R.E. Marquis, ***Fluoride inhibition of proton-translocating ATPases of oral bacteria.*** Infect Immun, 1987. **55**(11): p. 2597-603.

131. Guha-Chowdhury, N., A.G. Clark, and C.H. Sissons, ***Inhibition of purified enolases from oral bacteria by fluoride.*** Oral Microbiol Immunol, 1997. **12**(2): p. 91-7.

132. Stearns, J.C., M.D. Lynch, D.B. Senadheera, H.C. Tenenbaum, M.B. Goldberg, D.G. Cvitkovitch, K. Croitoru, G. Moreno-Hagelsieb, and J.D. Neufeld, ***Bacterial biogeography of the human digestive tract.*** Scientific reports, 2011. **1**: p. 170.

133. Morgan, X.C., T.L. Tickle, H. Sokol, D. Gevers, K.L. Devaney, D.V. Ward, J.A. Reyes, S.A. Shah, N. LeLeiko, S.B. Snapper, et al., ***Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment.*** Genome biology, 2012. **13**(9): p. R79.

134. Darfeuille-Michaud, A., J. Boudeau, P. Bulois, C. Neut, A.L. Glasser, N. Barnich, M.A. Bringer, A. Swidsinski, L. Beaugerie, and J.F. Colombel, ***High prevalence of adherent-invasive Escherichia coli associated with ileal mucosa in Crohn's disease.*** Gastroenterology, 2004. **127**(2): p. 412-21.

135. Kostic, A.D., E. Chun, L. Robertson, J.N. Glickman, C.A. Gallini, M. Michaud, T.E. Clancy, D.C. Chung, P. Lochhead, G.L. Hold, et al., ***Fusobacterium nucleatum potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment.*** Cell host & microbe, 2013. **14**(2): p. 207-15.

136. Sobhani, I., J. Tap, F. Roudot-Thoraval, J.P. Roperch, S. Letulle, P. Langella, G. Corthier, J. Tran Van Nhieu, and J.P. Furet, ***Microbial dysbiosis in colorectal cancer (CRC) patients.*** PLoS One, 2011. **6**(1): p. e16393.

137. Qin, J., R. Li, J. Raes, M. Arumugam, K.S. Burgdorf, C. Manichanh, T. Nielsen, N. Pons, F. Levenez, T. Yamada, et al., ***A human gut microbial gene catalogue established by metagenomic sequencing.*** Nature, 2010. **464**(7285): p. 59-65.

138. Yatsunenko, T., F.E. Rey, M.J. Manary, I. Trehan, M.G. Dominguez-Bello, M. Contreras, M. Magris, G. Hidalgo, R.N. Baldassano, A.P. Anokhin, et al., ***Human gut microbiome viewed across age and geography.*** Nature, 2012. **486**(7402): p. 222-7.

139. Qin, J., Y. Li, Z. Cai, S. Li, J. Zhu, F. Zhang, S. Liang, W. Zhang, Y. Guan, D. Shen, et al., ***A metagenome-wide association study of gut microbiota in type 2 diabetes.*** Nature, 2012. **490**(7418): p. 55-60.

140. Whitlock, E.P., J.S. Lin, E. Liles, T.L. Beil, and R. Fu, ***Screening for colorectal cancer: a targeted, updated systematic review for the U.S. Preventive Services Task Force.*** Ann Intern Med, 2008. **149**(9): p. 638-58.

141. Ahmed, S., G.T. Macfarlane, A. Fite, A.J. McBain, P. Gilbert, and S. Macfarlane, ***Mucosa-associated bacterial diversity in relation to human terminal ileum and colonic biopsy samples.*** Appl Environ Microbiol, 2007. **73**(22): p. 7435-42.

142. Huse, S.M., V.B. Young, H.G. Morrison, D.A. Antonopoulos, J. Kwon, S. Dalal, R. Arrieta, N.A. Hubert, L. Shen, J.H. Vineis, et al., ***Comparison of brush and biopsy sampling methods of the ileal pouch for assessment of mucosa-associated microbiota of human subjects.*** Microbiome, 2014. **2**(1): p. 5.

143. Kostic, A.D., M.R. Howitt, and W.S. Garrett, ***Exploring host-microbiota interactions in animal models and humans.*** Genes & development, 2013. **27**(7): p. 701-18.

144. Vallender, E.J. and G.M. Miller, ***Nonhuman primate models in the genomic era: a paradigm shift.*** ILAR J, 2013. **54**(2): p. 154-65.

145. Bauer, S.A., T.P. Arndt, K.E. Leslie, D.L. Pearl, and P.V. Turner, ***Obesity in rhesus and cynomolgus macaques: a comparative review of the condition and its implications for research.*** Comp Med, 2011. **61**(6): p. 514-26.

146. Handley, S.A., L.B. Thackray, G. Zhao, R. Presti, A.D. Miller, L. Droit, P. Abbink, L.F. Maxfield, A. Kambal, E. Duan, et al., ***Pathogenic simian immunodeficiency virus infection is associated with expansion of the enteric virome.*** Cell, 2012. **151**(2): p. 253-66.

147. Segata, N., J. Izard, L. Waldron, D. Gevers, L. Miropolsky, W.S. Garrett, and C. Huttenhower, ***Metagenomic biomarker discovery and explanation.*** Genome Biol, 2011. **12**(6): p. R60.

148. Duncan, S.H., A. Belenguer, G. Holtrop, A.M. Johnstone, H.J. Flint, and G.E. Lobley, ***Reduced dietary intake of carbohydrates by obese subjects results in decreased concentrations of butyrate and butyrate-producing bacteria in feces.*** Applied and environmental microbiology, 2007. **73**(4): p. 1073-8.

149. Fox, J.G., S.R. Boutin, L.K. Handt, N.S. Taylor, S. Xu, B. Rickman, R.P. Marini, F.E. Dewhirst, B.J. Paster, S. Motzel, et al., ***Isolation and characterization of a novel helicobacter species, "Helicobacter macacae," from rhesus monkeys with and without chronic idiopathic colitis***. J Clin Microbiol, 2007. **45**(12): p. 4061-3.

150. Marini, R.P., S. Muthupalani, Z. Shen, E.M. Buckley, C. Alvarado, N.S. Taylor, F.E. Dewhirst, M.T. Whary, M.M. Patterson, and J.G. Fox, ***Persistent infection of rhesus monkeys with 'Helicobacter macacae' and its isolation from an animal with intestinal adenocarcinoma.*** J Med Microbiol, 2010. **59**(Pt 8): p. 961-9.

151. Lertpiriyapong, K., L. Handt, Y. Feng, T.W. Mitchell, K.E. Lodge, Z. Shen, F.E. Dewhirst, S. Muthupalani, and J.G. Fox, ***Pathogenic properties of enterohepatic Helicobacter spp. isolated from rhesus macaques with intestinal adenocarcinoma.*** **Journal of medical microbiology,** 2014. **63**(Pt 7): p. 1004-16.

152. Lozupone, C. and R. Knight, ***UniFrac: a new phylogenetic method for comparing microbial communities.*** Applied and environmental microbiology, 2005. **71**(12): p. 8228-35.

153. Fox, J.G., C.C. Chien, F.E. Dewhirst, B.J. Paster, Z. Shen, P.L. Melito, D.L. Woodward, and F.G. Rodgers, ***Helicobacter canadensis sp. nov. isolated from humans with diarrhea as an example of an emerging pathogen.*** J Clin Microbiol, 2000. **38**(7): p. 2546-9.

154. Jovanovic, T., C. Ascenso, K.R. Hazlett, R. Sikkink, C. Krebs, R. Litwiller, L.M. Benson, I. Moura, J.J. Moura, J.D. Radolf, et al., ***Neelaredoxin, an iron-binding protein from the syphilis spirochete, Treponema pallidum, is a superoxide reductase.*** The Journal of biological chemistry, 2000. **275**(37): p. 28439-48.

155. Giacani, L., O. Denisenko, M. Tompa, and A. Centurion-Lara, ***Identification of the Treponema pallidum subsp. pallidum TP0092 (RpoE) regulon and its implications for pathogen persistence in the host and syphilis pathogenesis.*** Journal of Bacteriology, 2013. **195**(4): p. 896-907.

156. Albenberg, L., T.V. Esipova, C.P. Judge, K. Bittinger, J. Chen, A. Laughlin, S. Grunberg, R.N. Baldassano, J.D. Lewis, H. Li, et al., ***Correlation Between Intraluminal Oxygen Gradient and Radial Partitioning of Intestinal Microbiota in Human Beings and Mice.*** Gastroenterology, 2014.

157. Mercier, G.T., P.N. Nehete, M.F. Passeri, B.N. Nehete, E.A. Weaver, N.S. Templeton, K. Schluns, S.S. Buchl, K.J. Sastry, and M.A. Barry, ***Oral immunization of rhesus macaques with adenoviral HIV vaccines using enteric-coated capsules.*** Vaccine, 2007. **25**(52): p. 8687-701.

158. Langille, M.G., J. Zaneveld, J.G. Caporaso, D. McDonald, D. Knights, J.A. Reyes, J.C. Clemente, D.E. Burkepile, R.L. Vega Thurber, R. Knight, et al., ***Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences.*** Nature biotechnology, 2013. **31**(9): p. 814-21.

159. Fassbinder, F., M. Kist, and S. Bereswill, ***Structural and functional analysis of the riboflavin synthesis genes encoding GTP cyclohydrolase II (ribA), DHBP synthase (ribBA), riboflavin synthase (ribC), and riboflavin deaminase/reductase (ribD) from Helicobacter pylori strain P1.*** FEMS microbiology letters, 2000. **191**(2): p. 191-7.

160. Stark, R.M., M.S. Suleiman, I.J. Hassan, J. Greenman, and M.R. Millar, ***Amino acid utilisation and deamination of glutamine and asparagine by Helicobacter pylori.*** Journal of medical microbiology, 1997. **46**(9): p. 793-800.

161. Leduc, D., J. Gallaud, K. Stingl, and H. de Reuse, ***Coupled amino acid deamidase-transport systems essential for Helicobacter pylori colonization.*** Infection and immunity, 2010. **78**(6): p. 2782-92.

162. Binet, M.R. and O.M. Bouvet, ***Transport of glucose by a phosphoenolpyruvate:mannose phosphotransferase system in Pasteurella multocida.*** Research in microbiology, 1998. **149**(2): p. 83-94.

163. Moeller, A.H., M. Peeters, J.B. Ndjango, Y. Li, B.H. Hahn, and H. Ochman, ***Sympatric chimpanzees and gorillas harbor convergent gut microbial communities.*** Genome research, 2013. **23**(10): p. 1715-20.

164. Tsinganou, E. and J.O. Gebbers, ***Human intestinal spirochetosis--a review.*** German medical science : GMS e-journal, 2010. **8**: p. Doc01.

165. De Filippo, C., D. Cavalieri, M. Di Paola, M. Ramazzotti, J.B. Poullet, S. Massart, S. Collini, G. Pieraccini, and P. Lionetti, ***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, 2010. **107**(33): p. 14691-6.

166. Lin, A., E.M. Bik, E.K. Costello, L. Dethlefsen, R. Haque, D.A. Relman, and U. Singh, ***Distinct distal gut microbiome diversity and composition in healthy children from Bangladesh and the United States.*** PLoS One, 2013. **8**(1): p. e53838.

167. Walter, J. and R. Ley, ***The human gut microbiome: ecology and recent evolutionary changes.*** Annual review of microbiology, 2011. **65**: p. 411-29.

168. Croix, J.A., F. Carbonero, G.M. Nava, M. Russell, E. Greenberg, and H.R. Gaskins, ***On the relationship between sialomucin and sulfomucin expression and hydrogenotrophic microbes in the human colonic mucosa.*** PLoS One, 2011. **6**(9): p. e24447.

169. Caporaso, J.G., J. Kuczynski, J. Stombaugh, K. Bittinger, F.D. Bushman, E.K. Costello, N. Fierer, A.G. Pena, J.K. Goodrich, J.I. Gordon, et al., ***QIIME allows analysis of high-throughput community sequencing data.*** Nat Methods, 2010. **7**(5): p. 335-6.

170. McDonald, D., J.C. Clemente, J. Kuczynski, J.R. Rideout, J. Stombaugh, D. Wendel, A. Wilke, S. Huse, J. Hufnagle, F. Meyer, et al., ***The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome.*** Gigascience, 2012. **1**(1): p. 7.

171. Knights, D., J. Kuczynski, O. Koren, R.E. Ley, D. Field, R. Knight, T.Z. DeSantis, and S.T. Kelley, ***Supervised classification of microbiota mitigates mislabeling errors.*** ISME J, 2011. **5**(4): p. 570-3.

172. Abubucker, S., N. Segata, J. Goll, A.M. Schubert, J. Izard, B.L. Cantarel, B. Rodriguez-Mueller, J. Zucker, M. Thiagarajan, B. Henrissat, et al., ***Metabolic reconstruction for metagenomic data and its application to the human microbiome.*** PLoS Comput Biol, 2012. **8**(6): p. e1002358.

173. Cline, M.S., M. Smoot, E. Cerami, A. Kuchinsky, N. Landys, C. Workman, R. Christmas, I. Avila-Campilo, M. Creech, B. Gross, et al., ***Integration of biological networks and gene expression data using Cytoscape.*** Nat Protoc, 2007. **2**(10): p. 2366-82.

174. Griffin, S.O., E. Regnier, P.M. Griffin, and V. Huntley, ***Effectiveness of fluoride in preventing caries in adults.*** J Dent Res, 2007. **86**(5): p. 410-5.

175. Newbrun, E., ***Effectiveness of water fluoridation.*** J Public Health Dent, 1989. **49**(5 Spec No): p. 279-89.

176. Armfield, J.M., ***Community effectiveness of public water fluoridation in reducing children's dental disease.*** Public Health Rep, 2010. **125**(5): p. 655-64.

177. Kanapka, J.A. and I.R. Hamilton, ***Fluoride inhibition of enolase activity in vivo and its relationship to the inhibition of glucose-6-P formation in Streptococcus salivarius.*** Arch Biochem Biophys, 1971. **146**(1): p. 167-74.

178. van der Hoeven, J.S. and H.C. Franken, ***Effect of fluoride on growth and acid production by Streptococcus mutans in dental plaque.*** Infect Immun, 1984. **45**(2): p. 356-9.

179. van Loveren, C., J.F. Buijs, N. Kippuw, and J.M. ten Cate, ***Plaque composition, fluoride tolerance and acid production of mutans streptococci before and after the suspension of the use of fluoride toothpastes.*** Caries Res, 1995. **29**(6): p. 442-8.

180. Marquis, R.E., S.A. Clock, and M. Mota-Meira, ***Fluoride and organic weak acids as modulators of microbial physiology.*** FEMS Microbiol Rev, 2003. **26**(5): p. 493-510.

181. Marquis, R.E., ***Diminished acid tolerance of plaque bacteria caused by fluoride.*** J Dent Res, 1990. **69 Spec No**: p. 672-5; discussion 682-3.

182. Reilly, C., M. Goettl, M. Steinmetz, J. Nikrad, and R.S. Jones, ***Short-term effects of povidone iodine and sodium fluoride therapy on plaque levels and microbiome diversity.*** Oral Dis, 2016. **22**(2): p. 155-61.

183. Koopman, J.E., N.C. van der Kaaij, M.J. Buijs, Y. Elyassi, M.H. van der Veen, W. Crielaard, J.M. Ten Cate, and E. Zaura, ***The Effect of Fixed Orthodontic Appliances and Fluoride Mouthwash on the Oral Microbiome of Adolescents - A Randomized Controlled Clinical Trial.*** PLoS One, 2015. **10**(9): p. e0137318.

184. Hildebrand, F., T.L. Nguyen, B. Brinkman, R.G. Yunta, B. Cauwe, P. Vandenabeele, A. Liston, and J. Raes, ***Inflammation-associated enterotypes, host genotype, cage and inter-individual effects drive gut microbiota variation in common laboratory mice.*** Genome Biol, 2013. **14**(1): p. R4.

185. Bowden, G.H., ***Effects of fluoride on the microbial ecology of dental plaque.*** J Dent Res, 1990. **69 Spec No**: p. 653-9; discussion 682-3.

186. Maguire, A., F.V. Zohouri, J.C. Mathers, I.N. Steen, P.N. Hindmarch, and P.J. Moynihan, ***Bioavailability of fluoride in drinking water: a human experimental study.*** J Dent Res, 2005. **84**(11): p. 989-93.

187. Langille, M.G., J. Zaneveld, J.G. Caporaso, D. McDonald, D. Knights, J.A. Reyes, J.C. Clemente, D.E. Burkepile, R.L. Vega Thurber, R. Knight, et al., ***Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences.*** Nat Biotechnol, 2013. **31**(9): p. 814-21.

188. Feingold, K.R., M.H. Wiley, A.H. Moser, S.R. Lear, and M.D. Siperstein, ***Activation of HMG-CoA reductase by microsomal phosphatase.*** J Lipid Res, 1983. **24**(3): p. 290-6.

189. Beg, Z.H., J.A. Stonik, and H.B. Brewer, Jr., ***In vivo modulation of rat liver 3-hydroxy-3-methylglutaryl-coenzyme A reductase, reductase kinase, and reductase kinase kinase by mevalonolactone.*** Proc Natl Acad Sci U S A, 1984. **81**(23): p. 7293-7.

190. Potekaev, N.S., V.V. Pokrovskii, B.I. Zudin, E. Karetkina, and E.S. Gorbacheva, ***[Secondary relapsing syphilis in a patient with an infection caused by the human immunodeficiency virus].*** Vestn Dermatol Venerol, 1988(12): p. 55-9.

191. Truong, D.T., E.A. Franzosa, T.L. Tickle, M. Scholz, G. Weingart, E. Pasolli, A. Tett, C. Huttenhower, and N. Segata, ***MetaPhlAn2 for enhanced metagenomic taxonomic profiling.*** Nat Methods, 2015. **12**(10): p. 902-3.

192. Galletti, P.M. and G. Joyet, ***Effect of fluorine on thyroidal iodine metabolism in hyperthyroidism.*** J Clin Endocrinol Metab, 1958. **18**(10): p. 1102-10.

193. Schlesinger, E.R., D.E. Overton, H.C. Chase, and K.T. Cantwell, ***Newburgh-Kingston caries-fluorine study. XIII. Pediatric findings after ten years.*** J Am Dent Assoc, 1956. **52**(3): p. 296-306.

194. Gordon, J.A., ***Use of vanadate as protein-phosphotyrosine phosphatase inhibitor.*** Methods Enzymol, 1991. **201**: p. 477-82.

195. Wiseman, A., ***Effect of Inorganic Fluoride on Enzymes*****, in** ***Pharmacology of Fluorides*****,** F.A. Smith, Editor. 1970, Springer Berlin Heidelberg: Verlag Berlin. p. 48-97.

196. Warburg, O. and W. Christian, ***Isolation and crystallization of enolase.*** Biochemische Zeitschrift, 1942. **310**: p. 384-421.

197. Slater, E.C. and W.D. Borner, Jr., ***The effect of fluoride on the succinic oxidase system.*** Biochem J, 1952. **52**(2): p. 185-96.

198. Jurtshuk, P., Jr., ***Bacterial Metabolism*****, in** ***Medical Microbiology*****,** S. Baron, Editor. 1996: Galveston (TX).

199. Rautava, S., R. Luoto, S. Salminen, and E. Isolauri, ***Microbial contact during pregnancy, intestinal colonization and human disease.*** Nat Rev Gastroenterol Hepatol, 2012. **9**(10): p. 565-76.

200. Schomburg, D. and I. Schomburg, ***Class 1 Oxidoreductases VIII: EC 1.5*, in *Springer Handbook Of Enzymes*,** A. Chang, Editor. 2008, Springer Science & Business Media. p. 1-401.

201. Lim, E.S., Y. Zhou, G. Zhao, I.K. Bauer, L. Droit, I.M. Ndao, B.B. Warner, P.I. Tarr, D. Wang, and L.R. Holtz, ***Early life dynamics of the human gut virome and bacterial microbiome in infants***. Nat Med, 2015. **21**(10): p. 1228-34.

202. Vatanen, T., A.D. Kostic, E. d'Hennezel, H. Siljander, E.A. Franzosa, M. Yassour, R. Kolde, H. Vlamakis, T.D. Arthur, A.M. Hamalainen, et al., ***Variation in Microbiome LPS Immunogenicity Contributes to Autoimmunity in Humans.*** Cell, 2016. **165**(6): p. 1551.

203. Bolger, A.M., M. Lohse, and B. Usadel, ***Trimmomatic: a flexible trimmer for Illumina sequence data.*** Bioinformatics, 2014. **30**(15): p. 2114-20.

204. Suzek, B.E., H. Huang, P. McGarvey, R. Mazumder, and C.H. Wu, ***UniRef: comprehensive and non-redundant UniProt reference clusters.*** Bioinformatics, 2007. **23**(10): p. 1282-8.

205. Kanehisa, M., S. Goto, Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe, ***Data, information, knowledge and principle: back to metabolism in KEGG.*** Nucleic Acids Res, 2014. **42**(Database issue): p. D199-205.

206. Permutt, M.A., K.C. Chiu, and Y. Tanizawa, ***Glucokinase and NIDDM. A candidate gene that paid off.*** Diabetes, 1992. **41**(11): p. 1367-72.

207. Hanses, F., A. Kopp, M. Bala, C. Buechler, W. Falk, B. Salzberger, and A. Schaffler, ***Intracellular survival of Staphylococcus aureus in adipocyte-like differentiated 3T3-L1 cells is glucose dependent and alters cytokine, chemokine, and adipokine secretion***. Endocrinology, 2011. **152**(11): p. 4148-57.

208. Campbell, S.C., P.J. Wisniewski, M. Noji, L.R. McGuinness, M.M. Haggblom, S.A. Lightfoot, L.B. Joseph, and L.J. Kerkhof, ***The Effect of Diet and Exercise on Intestinal Integrity and Microbial Diversity in Mice.*** PLoS One, 2016. **11**(3): p. e0150502.

209. C, Y.L., ***The Effect of High-Fat Diet-Induced Pathophysiological Changes in the Gut on Obesity: What Should be the Ideal Treatment?*** Clin Transl Gastroenterol, 2013. **4**: p. e39.

210. Rooks, M.G., P. Veiga, A.Z. Reeves, S. Lavoie, K. Yasuda, Y. Asano, K. Yoshihara, M. Michaud, L. Wardwell-Scott, C.A. Gallini, et al., ***QseC inhibition as an antivirulence approach for colitis-associated bacteria.*** Proc Natl Acad Sci U S A, 2017. **114**(1): p. 142-147.

211. Goldfine, A.B., S. Crunkhorn, M. Costello, H. Gami, E.J. Landaker, M. Niinobe, K. Yoshikawa, D. Lo, A. Warren, J. Jimenez-Chillaron, et al., ***Necdin and E2F4 are modulated by rosiglitazone therapy in diabetic human adipose and muscle tissue.*** Diabetes, 2006. **55**(3): p. 640-50.

212. Kostic, A.D., D. Gevers, H. Siljander, T. Vatanen, T. Hyotylainen, A.M. Hamalainen, A. Peet, V. Tillmann, P. Poho, I. Mattila, et al., ***The dynamics of the human infant gut microbiome in development and in progression toward type 1 diabetes.*** Cell Host Microbe, 2015. **17**(2): p. 260-73.

213. Serafino, A., E. Balestrieri, P. Pierimarchi, C. Matteucci, G. Moroni, E. Oricchio, G. Rasi, A. Mastino, C. Spadafora, E. Garaci, et al., ***The activation of human endogenous retrovirus K (HERV-K) is implicated in melanoma cell malignant transformation.*** Exp Cell Res, 2009. **315**(5): p. 849-62.

214. Zhou, F., M. Li, Y. Wei, K. Lin, Y. Lu, J. Shen, G.L. Johanning, and F. Wang-Johanning, ***Activation of HERV-K Env protein is essential for tumorigenesis and metastasis of breast cancer cells.*** Oncotarget, 2016. **7**(51): p. 84093-84117.

215. Tugnet, N., P. Rylance, D. Roden, M. Trela, and P. Nelson, ***Human Endogenous Retroviruses (HERVs) and Autoimmune Rheumatic Disease: Is There a Link?*** Open Rheumatol J, 2013. **7**: p. 13-21.

216. Virgin, H.W., ***The virome in mammalian physiology and disease.*** Cell, 2014. **157**(1): p. 142-50.

217. Hoffmann, C., S. Dollive, S. Grunberg, J. Chen, H. Li, G.D. Wu, J.D. Lewis, and F.D. Bushman, ***Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents.*** PLoS One, 2013. **8**(6): p. e66019.

218. Huffnagle, G.B. and M.C. Noverr, ***The emerging world of the fungal microbiome.*** Trends Microbiol, 2013. **21**(7): p. 334-41.

219. Segata, N., J. Izard, L. Waldron, D. Gevers, L. Miropolsky, W.S. Garrett, and C. Huttenhower, ***Metagenomic biomarker discovery and explanation.*** Genome biology, 2011. **12**(6): p. R60.

220. Walter, J. and R. Ley, ***The human gut microbiome: ecology and recent evolutionary changes.*** Annu Rev Microbiol, 2011. **65**: p. 411-29.

221. Solnick, J.V., J. Fong, L.M. Hansen, K. Chang, D.R. Canfield, and J. Parsonnet, ***Acquisition of Helicobacter pylori infection in rhesus macaques is most consistent with oral-oral transmission.*** J Clin Microbiol, 2006. **44**(10): p. 3799-803.

222. Mercier, G.T., P.N. Nehete, M.F. Passeri, B.N. Nehete, E.A. Weaver, N.S. Templeton, K. Schluns, S.S. Buchl, K.J. Sastry, and M.A. Barry, ***Oral immunization of rhesus macaques with adenoviral HIV vaccines using enteric-coated capsules.*** Vaccine. Vol. 25. 2007. 8687-701.

223. Dubois, A., B.H. Natelson, P. van Eerdewegh, and J.D. Gardner, ***Gastric emptying and secretion in the rhesus monkey.*** Am J Physiol, 1977. **232**(2): p. E186-92.