ANTIBITOIC TREATMENT FOR TUBERCULOSIS INDUCES A PROFOUND DYSBIOSIS OF THE GUT MICROBIOME THAT PERSISTS LONG AFTER THERAPY IS COMPLETED

A Thesis

Presented to the Faculty of the Weill Cornell Graduate School

of Medical Sciences

in Partial Fulfillment of the Requirements for the Degree of

Masters of Science

by

Matthew F. Wipperman

May 2017

ABSTRACT

Mycobacterium tuberculosis, the cause of Tuberculosis (TB), infects one third of the world's population and causes substantial mortality worldwide. In its shortest format, treatment of drug sensitive TB requires six months of multidrug therapy with a mixture of broad spectrum and mycobacterial specific antibiotics, and treatment of multidrug resistant TB is much longer. The widespread use of this regimen worldwide makes this one the largest exposures of humans to antimicrobials, yet the effects of antimycobacterial agents on intestinal microbiome composition and long term stability are unknown. We compared the microbiome composition, assessed by both 16S rDNA and metagenomic DNA sequencing, of Haitian TB cases during antimycobacterial treatment and following cure by 6 months of TB therapy. TB treatment does not perturb overall diversity, but nonetheless dramatically depletes multiple immunologically significant commensal bacteria. The perturbation by TB therapy lasts at least 1.5 years after completion of treatment, indicating that the effects of TB treatment are long lasting and perhaps permanent. These results demonstrate that TB treatment has dramatic and durable effects on the intestinal microbiome and highlight unexpected extreme consequences of treatment for the world's most common infection on human ecology.

BIOGRAPHICAL SKETCH

NAME **Wipperman, Matthew Frederick**	POSITION TITLE Postdoctoral Researcher at Memorial Sloan Kettering Cancer Center
eRA COMMONS USER NAME MFWIPPERMAN	

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	MM/YY	FIELD OF STUDY
Franklin & Marshall College Lancaster, Pennsylvania	B.A.	2007 – 2011	Chemistry
Stony Brook University Stony Brook, New York	Ph.D.	2011 – 2015	Chemistry and Chemical Biology
Stony Brook Medicine Stony Brook, New York	Certificate	2013 – 2014	Biomedical science
Weill Cornell Medicine New York, New York	M.Sc.	2015 – 2017	Clinical and translational research
Memorial Sloan Kettering Cancer Center New York, New York	N/A	2015 – present	Postdoctoral research

A. Personal Statement

I am interested in using basic science to answer relevant questions in biomedical and translational scientific research. As an undergraduate I had the opportunity to work for four years under Dr. Phyllis Leber, who is a recognized expert in physical organic chemistry. My introduction to laboratory science during this time involved organic synthesis and physical chemistry techniques. I helped provide evidence that multicyclic vinylcyclobutanes rearrange to cyclohexenes via a diradical intermediate. Additionally, I helped characterize the properties of these diradical transition structures, helping to elucidate the mechanics of this class of pericyclic reactions. In graduate school, I worked under Dr. Nicole Sampson, a chemist and chemical biologist, which is where I learned techniques from classical biochemistry. I helped discover and characterize a new class of enzymes that make it possible for *M. tuberculosis* to metabolize steroids like cholesterol, synthesized chemical inhibitors that could target these enzymes, and helped to elucidate the steps of the cholesterol catabolic pathway. Additionally, I obtained a Certificate in

Biomedical Science at Stony Brook Medicine, and pursued a translational research project involving the development of pediatric biomarkers for TB disease using cholesterol metabolism as an indicator. Under the guidance of my graduate school co-mentor Dr. Sharon Nachman, I traveled to South Africa and went though the process of obtaining clinical blood samples for my research. This experience helped to form my current interests in using scientific tools to answer questions related to human health. In my current postdoctoral position under Dr. Michael Glickman, a distinguished physician-scientist, I am involved with a large translational research project studying immune and microbiomic predictors of innate resistance to Tuberculosis (TB) infection. I study how the microbiome may affect other aspects of immune function, and how the various stages of TB infection and TB treatment alter microbiomic composition in the short and the long term. Additionally, I am investigating Mycobacterial RecA and how its regulation via phosphorylation affects DNA repair. The interplay between basic science and the clinic interests me greatly. I enjoy working with a multidisciplinary network of individuals to answer medical questions using basic science and technology. I have a history as a highly motivated and self-sufficient researcher who communicates and works well both independently and with others.

B. Positions and Honors

Experience and Professional Memberships

2010 – Member, American Chemical Society

Honors

2012 Best Communicated Research (STEM) Award, Postdoctoral Research Symposium, Stony Brook University

2013 Scholarship, Scholars in the Biomedical Sciences at Stony Brook University

C. Contributions to Science

1. Tuberculosis continues to plague humanity as the number one killer of an infectious disease, and although much is known about what risk factors contribute to TB pathogenesis, virtually nothing is known about how the gut microbiome contributes to this risk. I integrated the research fields of TB and the microbiome using clinical samples from cross sectional and cohort studies to demonstrate microbiomic differences between different groups of people. I demonstrated how TB antibiotics affect the microbiome during and post treatment, and have correlated this to immunological health using RNAseq and immunophenotyping.

    - **Matthew F. Wipperman**, Dan Fitzgerald, Ying Taur, Sivaranjani Namasivayam, Alan Sher, James M. Bean, Vanni Bucci, Michael S.

Glickman. Antibiotic treatment for Tuberculosis induces a profound dysbiosis of the microbiome that persists long after completion of therapy. *Scientific Reviews, 2017,* ***submitted*****.**

- Sivaranjani Namasivayam, Mamoudou Maiga, Wuxing Yuan, Vishal Thovarai, Diego L. Costa, Lara R. Mittereder, **Matthew F. Wipperman**, Michael S. Glickman, Amiran Dzutsev, Giorgio, Trinchieri, Alan Sher. Longitudinal profiling reveals a persistent intestinal dysbiosis triggered by conventional antituberculosis therapy. *Microbiome.* **2017***,* 5:71. doi: 10.1186/s40168-017-0286-2.

2. The bacterial catabolism of steroids like cholesterol has been described for more than a century, however only recently has this process been implicated in disease pathogenesis, specifically of *M. tuberculosis*. I discovered and characterized a new class of acyl-CoA dehydrogenase enzymes using biochemical and bioinformatics techniques, and demonstrated their evolutionary distribution. I additionally synthesized an azasteroid inhibitor scaffold of these enzymes.

   - **Matthew F. Wipperman**, Meng Yang, Suzanne T. Thomas, and Nicole S. Sampson. Shrinking the FadE proteome of *Mycobacterium tuberculosis*: Insights into cholesterol metabolism through identification of an $\alpha_2\beta_2$ heterotetrameric acyl coenzyme A dehydrogenase family*. Journal of Bacteriology.* **2013.** *J. Bacteriol*. 195, 4331–4341. doi: 10.1128/JB.00502-13; PMCID: PMC3807453.

     - Commentary: Journal of Bacteriology: Voskuil, M. I. *Mycobacterium tuberculosis* cholesterol catabolism requires a new class of acyl coenzyme A dehydrogenase. 2013. *J. Bacteriol*. 195, 4319–4321. doi: 10.1128/JB.00867-13; PMCID: PMC3807469

     - Highlighted on the cover of the Journal of Bacteriology for the October, 2013 edition. http://jb.asm.org/content/195/19.cover-expansion

   - **Matthew F. Wipperman**, Nicole S. Sampson, and Suzanne T. Thomas. Pathogen roid rage: Cholesterol utilization by *Mycobacterium tuberculosis*. *Critical Reviews in Biochemistry and Molecular Biology*. July/August **2014**, 49 (4), 269-293. doi:10.3109/10409238.2014.895700; PMCID: PMC4255906

   - Meng Yang, Rui Lui, Kip E, Guja, **Matthew F. Wipperman**, Johnna R. St. Clair, Amber C. Bonds, Miguel Garcia-Diaz, Nicole S. Sampson. Unraveling Cholesterol Catabolism in *Mycobacterium tuberculosis*: ChsE4-ChsE5 $\alpha_2\beta_2$ Acyl-CoA Dehydrogenase Initiates

β-Oxidation of 3-Oxo-cholest-4-en-26-oyl CoA. *ACS Infectious Disease*. **2015**. doi**:** 10.1021/id500033m; PMCID: PMC4489319

3. Understanding the fundamental ways in which carbon-carbon bonds are made and broken has vast implications in our basic understanding of how organic reactions happen. I synthesized bicyclic and tricyclic vinylcyclobutanes, and studied [1,3] sigmatropic rearrangements of these molecules. I demonstrated the fundamental nature of diradical intermediates in this reaction process, and confirmed a theoretical prediction about the electron donating effects of the cyclopropylcarbinyl functionality.

    - Phyllis A. Leber, George R. Mann, III, William Hancock-Cerutti, **Matthew F. Wipperman**, Sylvia Zohrabian, Ryan M. Bell, and John E. Baldwin. *Thermal Isomerizations of cis,anti,cis-Tricyclo[7.4.0.0^{2,8}]tridec-10-ene. The Journal of Organic Chemistry.* **2012** 77 (7), 3468-3474. doi: 10.1021/jo300082a

    - Phyllis A. Leber, Anthony J. Nocket, **Matthew F. Wipperman**, Sylvia Zohrabian, Christopher Y. Bemis, and Munsha K. Sidhu. Experimental Evidence of a Cyclopropylbarbinyl Conjugative Electronic Effect. *Organic and Biomolecular Chemistry*. **2013**. 11, 5994–5997. doi: 10.1039/C3OB41365A

TABLE OF CONTENTS

LIST OF FUIGURES AND TABLES

LIST OF ABBREVIATIONS

Mtb	*Mycobacterium tuberculosis*
TB	active Tuberculosis disease
IFNγ	Interferon gamma
TNFα	Tumor necrosis factor α
IGRA	IFNγ release assay
HRZE	Isonasize (H), rifampin (R), pyrzanamide (Z), and ethambutol (E) combination drug therapy for TB
PAR%	population attributable risk percent
ILCs	innate lymphoid cells
Tregs	T regulatory cells
LTi	lymphoid tissue inducer cells
HMP	American Human Microbiome Project
SCFAs	Short chain fatty acids
NAD	Nicotinamide adenine dinucleotide
rDNA	16S ribosome-encoding deoxyribonucleic acid
BLASTn	Nucleotide Basic Local Alignment Search Tool
LeFSe	Linear discriminant analysis effect size
NMDS	Non-metric multidimensional scaling
DCA	Detrended correspondence analysis
OTU	Operational taxonomic unit
KEGG	Kyoto encyclopedia of genes and genomes

CHAPTER ONE - INTRODUCTION

Introduction to Tuberculosis disease

Each year, 3 – 4% of all deaths worldwide from any cause are attributable to infection with *Mycobacterium tuberculosis* (*Mtb*), the causative agent of Tuberculosis (TB) disease, which amounts to almost 5,000 TB-related deaths each day (1). This colossal disease burden necessitates a thorough understanding of both the pathogenic strategies *Mtb* uses to cause disease, as well as the host susceptibilities *Mtb* has evolved to exploit. Individuals can be uninfected, infected with latent *Mtb*, have active TB disease, or be cured through antibiotic therapy (Figure 1).

Many factors can influence the probability that some individuals transition from one of these stages to another, but most defined risk factors compromise immune function (2). For example, untreated HIV infection, which depletes CD4+ T cells, is associated with elevated risk of TB disease. Overall, immune status is also affected by age—the elderly and young infants are at a disproportionately high risk of *Mtb* infection and subsequent TB disease. Furthermore, genetic mutations in pathways involved in controlling *Mtb* infection and maintaining latency, like those of IFNγ and TNFα, also cause an increased risk of active TB disease (3). Despite these examples, known immune deficiencies are not sufficient to explain why the incidence of new active TB cases hovers over 10 million people each year, with a mortality rate between 1.5 – 2 million people (1). Furthermore, it is unknown why some individuals in TB endemic countries (where *Mtb* exposure is common) never become *Mtb* infected, why most latently infected individuals never progress to

active TB disease, and what may account for treatment failure, relapse, and re-infection. These epidemiological curiosities drive the questions proposed in this thesis.

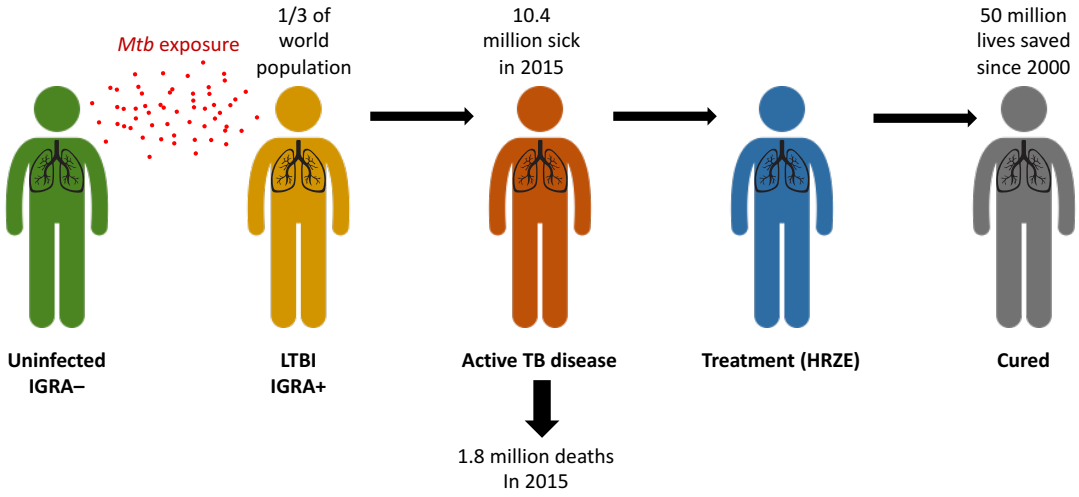


Figure 1: The stages of *Mtb* infection and TB disease. Individuals in TB endemic countries are exposed to *Mtb* at home or in the community throughout their lifetimes. Some individuals, despite this exposure, remain IGRA– while others develop latent TB disease (LTBI), and are IGRA+. Around 5 – 15% of individuals with LTBI will go on to have active TB disease, where they are able to infect others around them. Since TB is a curable disease, people can take treatment for a minimum of 6 months with four antibiotics (HRZE). Since the turn of the century, more than 50 million lives have been saved by antimycobacterial TB therapy.

Epidemiology of Tuberculosis

Mtb has co-evolved with humans for centuries, to the point where the only known natural reservoir is in *Homo sapiens*. Innumerable factors describe why some individuals get this disease while others do not, and despite all we know about TB, there are some elusive epidemiological questions that remain unsolved. First, individuals—even those who are exposed to *Mtb* throughout their lives—can remain uninfected, measured as being negative for the IFNγ release assay diagnostic test (IGRA⁻). Second, it is believed that nearly one third of the world's population is latently infected with live *Mtb* bacteria (clinically called LTBI), and are therefore IGRA+. More than 80% of these people will never get sick with active TB disease. Third, of the people who do get sick with active TB disease, depending on the susceptibility of the *Mtb* strain they are infected with to a standard course of antimycobacterial therapy, individuals with active disease can take treatment, thus, TB is a curable disease. These stages are depicted in Figure 1. Being cured from TB is achieved through a standard course of TB treatment, which consists of four antibiotics for two months, followed by two antibiotics for at least an additional four months. This makes antimycobacterial therapy for TB disease a full six months—one of the longest courses of antibiotic treatment administered to humans for any infectious disease.

It is unknown why nearly 20% of individuals in TB endemic countries remain IGRA– throughout their lifetimes, despite what is assumed to be persistent exposure. It is further unknown why 5 – 15% of individuals with LTBI will progress to active TB disease, whereas most people's immune systems are able to keep *Mtb* at bay for their entire lives. Finally, we do not know why

individuals who have had TB and have been cured through antibiotic treatment are more than four times more likely to become re-infected with *Mtb* again in their lifetimes.

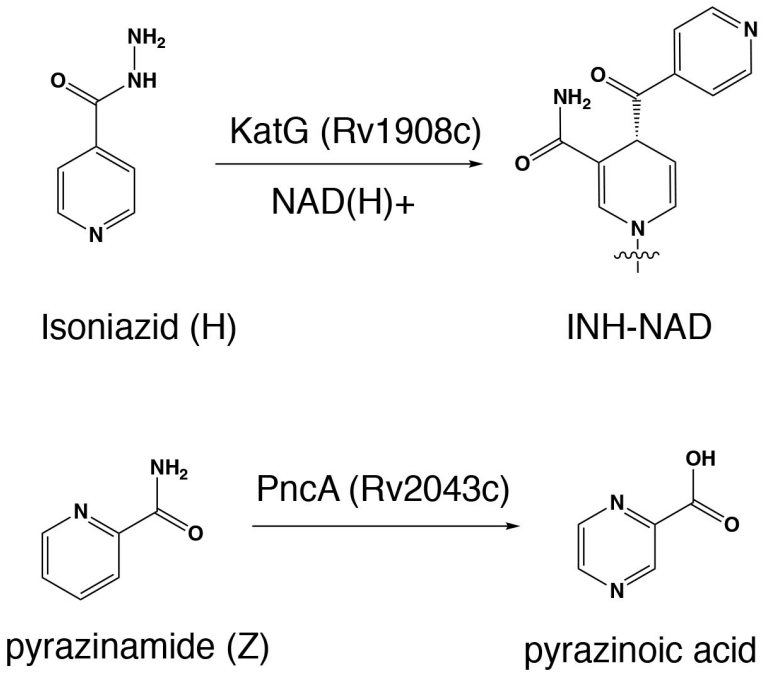


Figure 2
Mechanism of action and chemical structures of HRZE drugs and their described active products. Isoniazid (H) is a prodrug, which requires the mycobacterial protein KatG (Rv1908c) for ligation with NAD+, the product of which is the active drug that inhibits enoyl-acyl carrier protein reductase InhA, inhibiting fatty acid biosynthesis. Pyrazinamide (Z) is activated by the mycobacterial enzyme PncA to pyrazinoic acid. Ethambutol (E) targets arabinogalactan biosynthesis, and Rifampicin (R) inhibits *Mycobacterial* DNA-dependent RNA polymerase through binding to the β-subunit of the enzyme.

Calculating what epidemiological factors contribute to LTBI and TB disease incidence is by no means trivial. The goal is to estimate the percent of total incidence of disease in the population due to a given epidemiological factor (exposure), and to have the numbers from all contributing factors to add up to 100%. This is called the population attributable risk percent (PAR%), also known as the etiologic factor that describes what percent of the disease would be removed if that exposure was removed. For example, 12% of active TB disease incidence is attributable to HIV infection, thus if HIV infection was no longer a comorbidity with *Mtb* infection, there would be 12% fewer cases of active TB disease worldwide. Although these calculations are rough and fraught with assumptions, the question at hand is determining the full suite of etiological factors explain LTBI and active TB disease incidence.

Several of these factors link back to the sociological and historical understanding of TB. These include malnutrition and air pollution, each of which make up around one quarter of risk. Additionally, metabolic syndrome, alcohol use, and smoking, all increase risk for active TB disease significantly. Nevertheless, is it suspected that these factors are collectively unable to explain the full PAR% for why people get active TB, and even fewer factors explain why individuals are LTBI. This lack of understanding of the epidemiology of TB, in conjunction with the complete lack of understanding as to how the microbiome relates to LTBI and active TB disease, is the motivation for the study herein.

The Human Microbiome

Microbes and smaller organisms have co-evolved with more complex cellular life for the past 500 million years, thus it is not surprising that complex human immune systems are adapted to not only recognize and fight off these organisms when they are out of check, but to also take advantage of the host of biochemical and metabolic functions they possess. For every human cell, there is at least one other cell belonging to the microbes that inhabit our bodies. These microbes on nearly every exposed surface of the human body play an important role throughout the lifecycle of that organism they inhabit—collectively they are referred to as the *microbiota*. Most microbes that humans are exposed to are non-pathogenic, a relationship called commensalism. Many of these microbes actually benefit us, a relationship called mutualism. The microbiota help humans to digest food, they biosynthesize small molecules like serotonin that regulate our emotions, and teach the immune system how to respond to pathogens. The collection of the microorganisms, as well as the elaborate biochemical functionality in their genes is referred to as the *microbiome*. These indispensable functions have only been realized recently, since the mid 2000s. With the advent of next-generation DNA sequencing, culture-independent means to ask which organisms are present, and what they are doing, became a reality.

Barrier surfaces and the microbiome

Host-microbe interactions are complex, but one significant feature is that they generally involve the interaction of putative pathogens with mucosal barrier surfaces. This interaction establishes what sort of biochemical and immunological dialogue exists between at the onset of infection or disease, thus, understanding the immune status of mucosal surfaces before such an event is requisite to understanding these multifaceted host-pathogen interactions. Much of the work aimed at understanding these interactions involves the study of local microbiome composition at the site of contact.

Increasing evidence clearly demonstrates that these host-pathogen exchanges happen frequently and are immunologically important. For example, at the largest site of exposure of the immune system to the outside world—the gut—it has been shown that temporary infections can permanently alter the memory T cell repertoire (4). At baseline regulatory mechanisms exist to promote local immune homeostasis, especially involving the anti-inflammatory properties of intestinal macrophages and at least four unique subsets of intestinal dendritic cells (5). A population of innate CD11b+ CD103+ dendritic cells in the lamina proproia (LP) constantly surveil the luminal contents of the gut—they have mechanisms to transverse the tight junctions of the epithelium, sample antigens (including whole bacteria), and bring them back to the LP for processing and presentation to B and T cells. Additionally, CD11b+ Notch-2 dependent dendritic cells sense bacterial antigens. One outcome of this is TLR5 activation by the bacterial protein flagellin, resulting in IL-23 production. IL-23 binds to IL-23R receptors on innate lymphoid cells (ILCs), which result in the production of IL-22. Most epithelial cells express the

IL-22R receptor, and when activated, Il-22R causes antimicrobial peptide RegIIIβ and RegIIIγ production, mucous generation, and the promotion of overall intestinal barrier integrity. Additionally, CD4+ lymphoid tissue inducer (LTi) cells are rich sources of IL-22 in the gut (6). The importance of IL-22 in the context of intestinal homeostasis is highlighted by the collection of literature on the negative outcomes of *Citrobacter rodentium* infection in IL-22-/- mice, which models human enteropathogenic *Escherichia coli* infections (7, 8). This complex series of interactions are complex, and sometimes contradictory, given that in some contexts IL-22 can also promote inflammation (9). The cocktail of small molecules, chemokines, and cytokines, as well as the baseline host-commensal relationship are all factors to consider in host-commensal interactions.

Mucosal interactions with the adaptive arm of the immune system often rely on antigen-specific cues from the commensal microbiota. In the skin, mechanisms exist for establishing local immune homeostasis in conjunction with the microbiota. During early life, microbiota of the skin elicit antigen-specific T cells, and concomitantly establish tolerance by promoting the expansion of a population of activated T regulatory cells (Tregs) (10). If Tregs are prevented from entering neonatal skin, tolerance to commensals is attenuated. In the gut, bacterial antigens are presented to Notch-2-dependent dendritic cells (11) that then produce TGFβ. Tregs to inhibit effector responses, again promoting immune homeostasis in an antigen-dependent fashion (12). Collectively, immune tone is determined both locally and systemically through the interaction of the microbiota with the host. This immune tone is an increasingly important consideration in several contexts

and for a number of reasons, and may contribute to the pathogenesis of infectious disease.

The long-term response to a pathogen can be controlled spatially, as a function of which barrier surface that pathogen invades and what immune cell types respond to this invasion. For example, the dimorphic fungus and skin commensal organism *Candida albicans* can gain the ability to transvers the skin barrier surface and elicit an immune response. Skin dendritic cells called Langerhans cells use Dectin-1 to sense the yeast form of the organism during initial breaching of the skin barrier. These cells migrate to lymph nodes and produce IL-6, biasing naïve CD4+ cells to adopt a Th17 phenotype. Upon re-challenge of *C. albicans* on the skin, but not systemically, mice who have undergone this immune challenge are able to resist infection. A different population of CD103+ dermal dendritic cells that reside much deeper within the barrier surface use TLR2 to sense the filamentous (hyphae) form of *C. albicans*. Migration of these cells to the lymph nodes and production of IL-12 biases CD4+ naïve cells to the IFNγ producing Th1 phenotype. Upon re-challenge of *C. albicans* systemically, but not on the skin, mice who have undergone this immune challenge can resist infection. Thus, the same immune cell type located in two different locations and expressing two different receptors results in two different immune tones that can act either locally or at a distance.

The repertoire of immune cells at any barrier site and the receptors that they express are a function, in part, of the commensal microbiome. In the previous example, production of SCFAs by commensal anaerobic bacteria stimulate

colonization resistance and inhibit *C. albicans* growth. Additionally, SCFAs induce host immune response factors HIF-1α, a transcription factor involved in innate immunity, and the antimicrobial peptide LL-37 (13). For a pathogen, the establishment of a replicative niche depends on access to space, nutritional factors, and the ability to compete for these within the larger context of an ecological community that is the microbiome and the immune responses elicited by the host, primarily shaped by the local and distal microbiome.

The Microbiome and Tuberculosis

As discussed, one emerging biological factor that remains unstudied in the context of TB susceptibility and treatment is the intestinal microbiota. The organisms comprising the intestinal microbiota account for the largest exposure of the immune system to the environment, and in turn, the composition and metabolic activities of the intestinal bacterial community directly participate in the development and function of peripheral immunity (14).

Researchers recently described how SCFAs produced by members of the lung microbiota can induce peripheral immune phenotypes, leading to increased TB susceptibility in HIV+ individuals on antiretroviral therapy (15). In South Africa, where the TB epidemic is one of the most severe in the world, researchers recruited individuals in a longitudinal clinical study, and investigated both the composition of the lung microbiota, as well as the microbe-derived small molecules present in the periphery. Predictors of TB hazard include serum IFNγ IL-17A, butyrate, and propionate, which related to Treg expansion and likely attenuation of a productive pro-inflammatory immune response against *Mtb* upon initial infection or reactivation.

It is unknown how the composition of the gut microbiome affects TB, nor is it easy to predict. Even small changes in low abundance bacteria of the intestinal microbiota can have large effects on health outcomes. For example, the mucin-degrading species *Akkermansia muciniphila* comprises only 1 – 3% of the intestinal microbiota of some people, but its loss is associated with obesity, Type 2 diabetes, and inflammation in both mice and humans (16-18).

Similarly, babies with depletion of four microbiota constituents are at risk for asthma, and transfer of these four bacteria to germ free mice ameliorates Th2-mediated airway inflammation (19).

Although intestinal microbiome composition can be determined by many factors throughout a lifetime, there is an increasingly well-understood stability in its composition (20) in the absence of antibiotic perturbation. The human intestinal microbiome taxonomic composition is dominated by Firmicutes and Bacteroidetes, somewhat lower levels of Actinobacteria and Proteobacteria, and often low abundance but important phyla like Verrucomicrobia, Fusobacteria, and Euryarchaeota (21). Antibiotics can target any of these taxa, and have distinctive microbiome-altering effects both during and post treatment. For example, the fluoroquinolone ciprofloxacin has broad effects on intestinal microbiota, especially Enterobactericeae (22), and causes a marked decrease in taxonomic diversity (23). β-lactam antibiotics also cause detectable microbiomic shifts, resulting in altered carbohydrate processing and increased sugar metabolism (24). Antibiotic induced perturbations of the intestinal microbiome can also result in altered metabolite production by the microbiota. The third-generation cephalosporin cefoperazone, when administered to mice for 10 days, increased secondary bile acid, glucose, free fatty acids, and dipeptide production, while decreasing primary bile acids and sugar alcohols in the intestine (25). In people undergoing cancer treatment, treatment with metronidazole led to substantial derangement of the microbiota through its anti-anaerobic activity; in contrast, treatment with intravenous vancomycin had relatively little impact (26). Although the pre-treatment ecological state of the microbiome generally recovers after stopping antibiotic

treatment, there are noticeable effects that may persist for weeks, months, and even years after treatment is stopped (27). Although these are just a few examples, the current model is that antibiotic treatment can result in the establishment of an alternative state that can have systemic, often deleterious, consequences for immunity and disease susceptibility (21, 28).

Motivation for Study

Little is known about the effects of first-line TB antibiotics on the intestinal microbiome. In contrast to commonly used broad spectrum antimicrobials, most first-line antibiotics used to treat TB are narrow spectrum agents with *Mycobacteria*-specific targets. A standard course of TB therapy for drug sensitive *Mtb* consists of the administration of four drugs for two months, Isoniazid (H), Rifampin (R), Pyrazinamide (Z), Ethambutol (E), and then the continuation of HR for an additional four months. Isoniazid is a prodrug, which requires the mycobacterial catalase-peroxidase KatG (Rv1908c) for ligation with NAD+, the product of which is the active drug that inhibits enoyl-acyl carrier protein reductase InhA, thereby inhibiting mycolic acid biosynthesis (Figure 1) (29). Similarly, pyrazinamide is only used clinically for *Mtb* and is converted to the active pyrazinoic acid by the bacterial amidase PncA (Figure 1) (30). Ethambutol is thought to inhibit arabinosyltransferases, and like isoniazid, inhibit cell wall biosynthesis. Of the four standard TB antibiotics used in “short course” treatment (HRZE), only Rifampin, which inhibits bacterial RNA polymerase, is a broad-spectrum antimicrobial that is used for non-mycobacterial infections. The effects of this prolonged antibiotic regimen on the intestinal microbiota are unknown, and almost impossible to predict.

In this study, we characterize for the first time the intestinal microbiome of residents of Haiti and assess the immediate and long-term effects of TB treatment with HRZE antimycobacterial therapy on microbial diversity, taxonomic composition, and biochemical capacity. We demonstrate the substantial and unique disruptive effects of HRZE therapy on intestinal microbiome composition using both 16S and metagenomic DNA sequencing and demonstrate that long term gut microbiomic dysbiosis is a consequence of TB treatment.

CHAPTER 2 – RESULTS

Clinical characteristics of study groups from the TBRU study

We recruited four groups of individuals using a cross-sectional research study design. To characterize the intestinal microbiomes of individuals from the Haitian population, we recruited two groups of control individuals, 50 with no *Mtb* infection, and 25 latently infected by *Mtb* (LTBI), as defined by a positive Interferon Gamma Release Assay (IGRA) test. To determine the effect of HRZE antimycobacterial treatment on the intestinal microbiome, we recruited 19 volunteers currently on treatment with HRZE for drug sensitive Tuberculosis. In addition, to determine the duration of the microbiome perturbation of HRZE treatment, we recruited 14 previously treated cases who were cured of active TB. The clinical characteristics of the groups are given in Table 1. To appropriately control for age, we divided our LTBI group into two distinct control subgroups, designated LTBI (treatment control) and LTBI (cured control), since microbiome composition can vary significantly with age (31).

Group	Number of people	Average Age (range)	% female	Time on TB treatment	Time since TB treatment	Average number of 16S reads per patient (range)	Average number of OTUs per subject (range)	Shannon Diversity
Healthy control	50	33 (19 – 59)	62	N/A	N/A	35951 (690 – 116638)	230 (19 – 473)	3.412
LTBI (treatment control)	25	26 (17 – 32)	52	N/A	N/A	41038 (4713 – 118110)	229 (47 – 470)	3.341
LTBI (cured control)	17	24 (17 – 29)	50	N/A	N/A	48329 (8309 – 118110)	247 (119 – 453)	3.417
Treatment	19	20 (13 – 32)	54	3.2 months (13 – 217 days)	N/A	38489 (4360 – 140543)	150 (57 – 118)	3.218
Cured	14	23 (17 – 27)	36	6 months	417 days (34 – 1202 days)	21731 (5134 – 105583)	237 (142 – 388)	3.75
Human Microbiome Project (Americans)	319	N/A	57	N/A	N/A	15937 (114 – 53607)	58 (17 – 139)	2.347

Table 1: Patient populations analyzed in this study for 16S rDNA sequencing. Data are divided into study groups described in the text. The number of participants, participants average age, gender distribution, amount of time on HRZE treatment or time since treatment, average number of 16S reads and subsequent OTUs, and Shannon diversity index are shown if applicable.

Characterization of the healthy Haitian intestinal microbiome

Today, Haiti is the most financially impoverished country in the Western Hemisphere, with more than 80% of the population living below the poverty line. This cycle of poverty was recently amplified by a massive earthquake in 2010 that killed 300,000 people and displaced more than 1.5 million people. The socioeconomic reality of Haiti contributes to a TB epidemic that affects the entire population; the incidence of active TB in Haiti is 194 per 100,000 people (1).

Most microbiome studies in humans have been performed on relatively affluent populations in North America or Europe, with a relative paucity of large-scale studies in developing nations. Given the substantial effect of diet and other factors on microbiome composition, and the lack of data about microbiome composition in TB endemic countries, we initially characterized the healthy Haitian adult microbiome by enrolling 50 *Mtb*-uninfected, apparently healthy, Haitian individuals with no known recent *Mtb* exposure, and collected stool for 16S rDNA and metagenomic DNA sequencing. The mean age for volunteers in the healthy group was 33 years old (range 19 – 59 years), with 62% who were female. To our knowledge, this is the first characterization of the intestinal microbiome from Haitians.

Microbiome composition was first characterized through sequencing of the V4–V5 variable region of the 16S ribosomal DNA gene. Reads were filtered for size and quality and chimeras were removed to yield an average of 35,951 high-quality reads per patient (Table 1). Reads were clustered into operational taxonomic units (OTUs) at 97% identity, and taxonomic identification was

made using BLASTn against the refseq_rna database (32). We also analyzed both 16S (319 subjects) and metagenomic data (150 subjects) from healthy individuals recruited from the American Human Microbiome Project (HMP) (33, 34). Using metagenomic data, we found a statistically significant different mean Shannon diversity index between the HMP and Haitian populations (Wilcoxon rank sum test with continuity correction, p=0.0131), indicating that the Haitian microbiome is more diverse than the American counterpart (Figure 3). Non-metric multidimensional scaling (NMDS) principal component analysis on Bray Curtis distances clearly separated HMP samples from either *Mtb* uninfected Haitian samples or Haitian LTBI samples (Figure 4A).

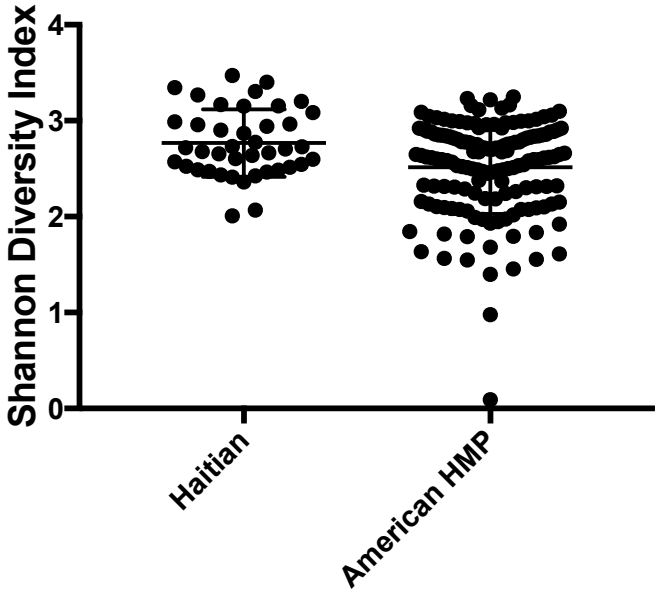


Figure 3: Shannon diversity index for the Haitian and Human Microbiome Project samples. Calculated using metagenomic data and the Vegan package in R (35). Unique species are acting as the input for each person. The Haitian vs HMP cohorts have statistically significant different Shannon diversity indices ($p=0.0127$).

To obtain species level resolution of the differences between the two populations, we performed taxonomic profiling of the metagenomes from the 150 healthy HMP individuals (see above) and from 43 healthy Haitian individuals (No TB: 18 + LTBI: 25) using the Metaphlan pipeline (36). The most abundant taxonomic genera in the healthy Haitian group come from *Prevotella, Faecalibacterium, Ruminococcus, Bacteroides, Oscillibacter, Blautia, Bifidobacterium, Agathobacter, Romboutsia, Gemmiger, Coprococcus, Holdemanella, Lactobacillus, Eubacterium, Anaerostipes, Sporobacter, Clostridium, Catenibacterium,* and *Akkermansia* (Figure 4). The most abundant genus is *Prevotella*, from the Phylum Bacteroidetes. The preponderance of *Prevotella* is similar to prior characterization of microbiome composition in other non-Western populations (37, 38).

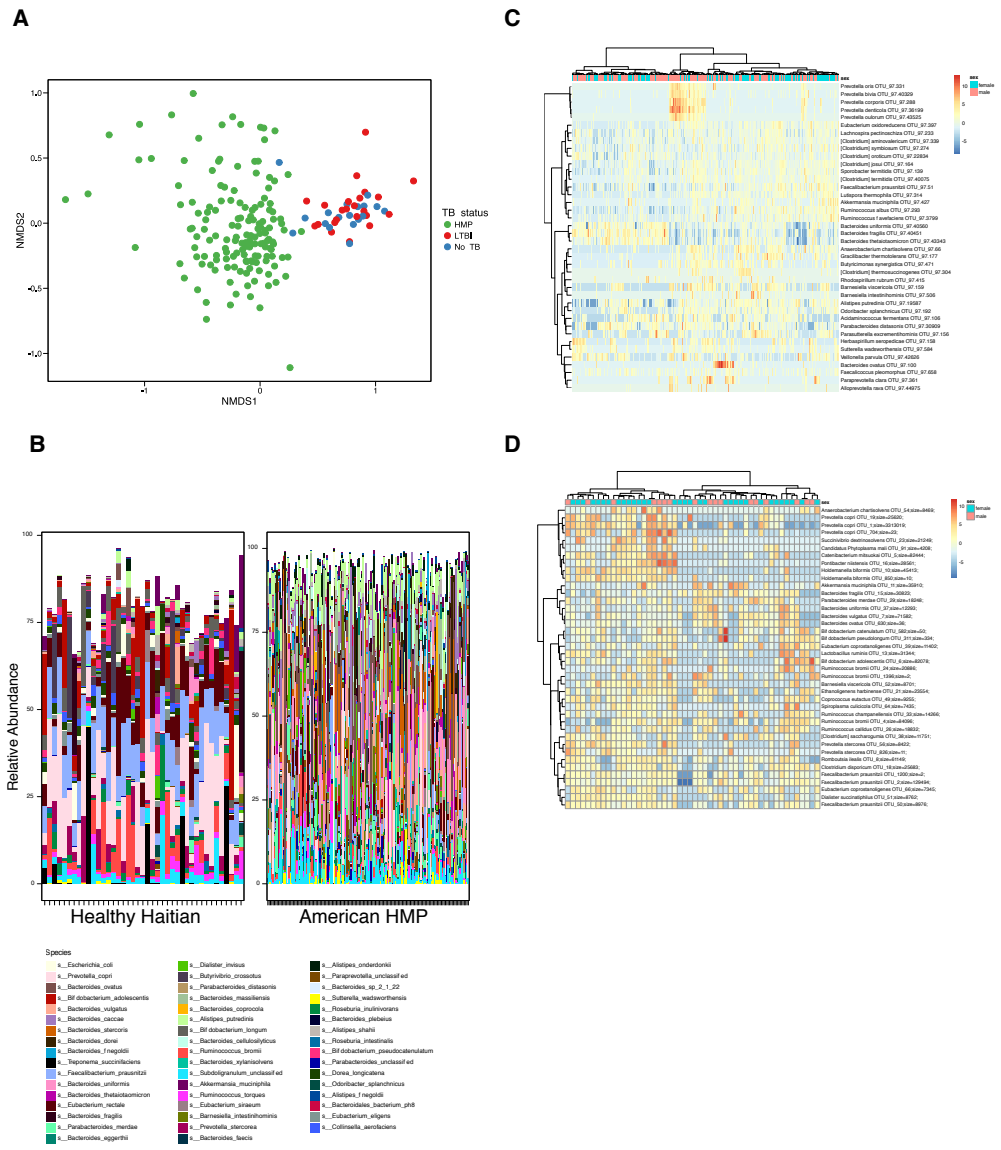


Figure 4: Characterization of the healthy Haitian microbiome in comparison to the Human Microbiome Project by 16S and metagenomic DNA sequencing. A. NMDS plot on Bray Curtis distances between healthy Haitian (both LTBI (red) and *Mtb* uninfected (blue) and HMP samples (green), based on metagenomic DNA sequencing, highlighting differences between the two populations and the lack of difference between LTBI and No TB subjects. **B.** The 40 most abundant taxa in both Haitian and HMP samples using metagenomic sequencing. **C.** Unsupervised hierarchical clustering of the 40 most abundant taxa of the 319 intestinal microbiome samples from the American Human Microbiome Project determined by 16S rDNA sequencing. **D.** Unsupervised hierarchical clustering the 40 most abundant taxa from 50 *Mtb* uninfected Haitian subjects using 16S rDNA sequencing.

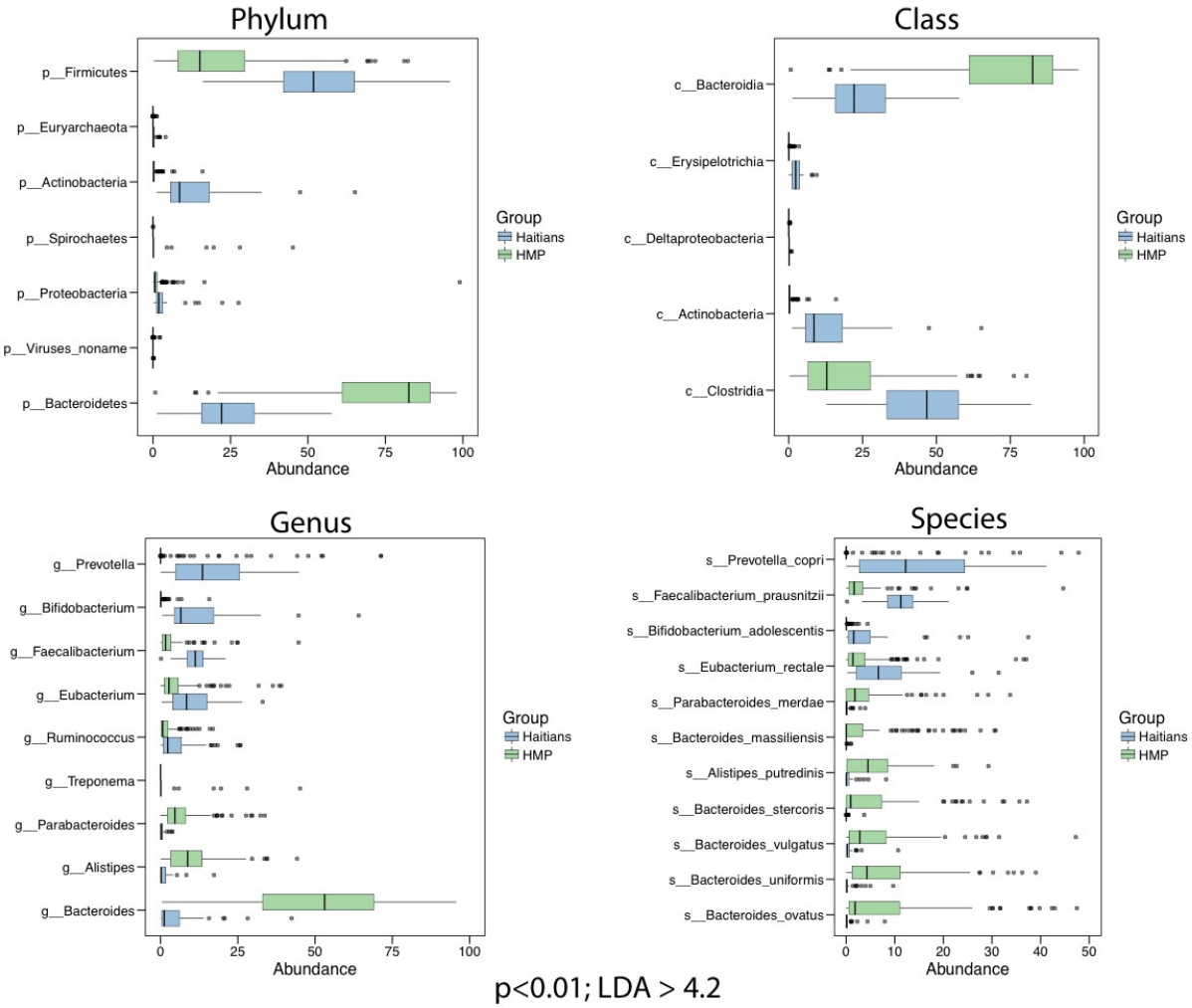


Figure 5: Differentially abundant taxa between Haitian and HMP samples. LEfSe was used to determine the major differences in taxonomy using Metaphlan taxonomic assignments on metagenomic data, and the relative abundance of the most significantly different taxa are plotted. Phylum, Class, Genus, and Species level comparisons are plotted between 150 HMP metagenomic samples with 42 healthy (No TB + LTBI) Haitian individuals. All biomarkers are derived from a Wilcoxon Rank Sum test $p < 0.01$ and have an linear discriminant analysis effect size score > 4.2. The median of the data in the box is indicated by a line, the whiskers show 1.5 times the value of the interquartile range of the box hinge, and outliers are shown separately as dots.

Hierarchical clustering of the top 40 most abundant OTUs for each group determined by 16S sequencing highlighted the heterogeneity of microbiome composition among healthy individuals within a population (Figure 4C and 4D). Comparison between Haitian and American data showed that despite being the dominant Phyla in both populations, *Firmicutes* were more abundant in the Haitian population (77.9% vs 64.8%), whereas *Bacteroidetes* were more abundant in the American population (21.4% vs 10.8%, Table S1, Figure 4). Two additional Phyla were found to have increased abundance in the Haitian population, *Actinobacteria*, the Phylum to which *Mtb* belongs, and *Euryarchaeota* (Archaeal Kingdom), whereas *Proteobacteria* and *Spirochetes* did not significantly differ between the two populations. Using LEfSe (39) to compare the most differentially abundant species between the two populations, we determined that on average Haitians have higher levels of *Bifidobacterium*, *Blautia, Catenibacterium, Collinsella, Coprococcus, Erysipelotrichaceae, Eubacterium, Faecalibacterium, Lactobacillus, Methanobrevibacter, Mitsuokella, Olsenella, Peptostreptococcaceae, Prevotella, Rothia*, and *Streptococcus*. By contrast, Americans have higher levels of *Adlercreutzia, Alistipes, Bacteroides, Burkholderiales, Holdemania, Parabacteroides*, and *Parasutterella* (Figure 4). We recognize that 16S data obtained from the HMP will differ with respect to DNA extraction, library preparation, and DNA sequencing. We downloaded the HMP dataset and performed OTU calling using our protocol and nomenclature (i.e., OTU names were called from the refseq_rna database), and at this level of comparison, we believe these OTUs accurately differ between the two populations.

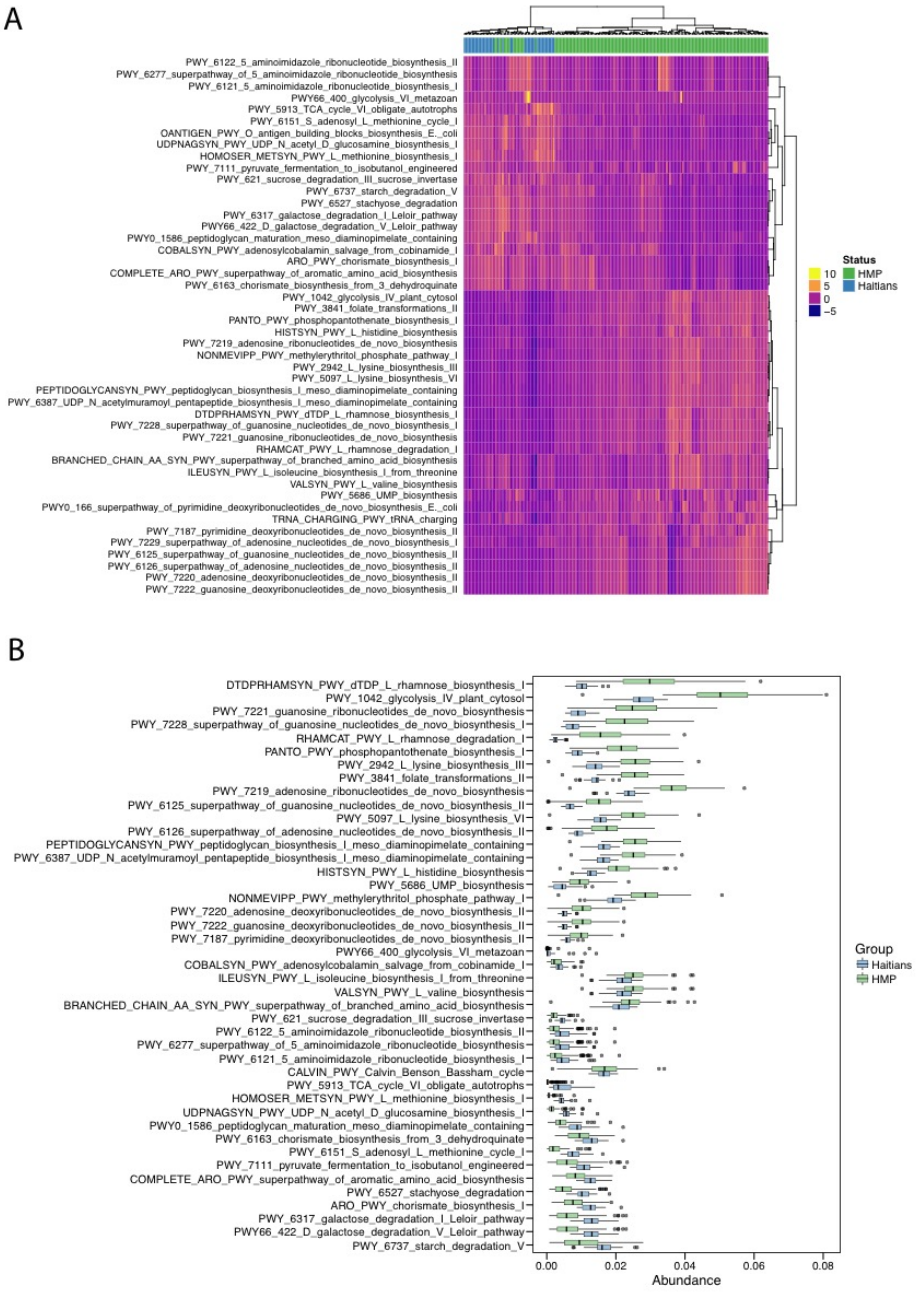


Figure 6: **Metacyc pathway abundance comparing the HMP dataset and Haitian healthy control people. A**. Unsuprivised hiericherical clustering of significantly altered Metacyc pathways. **B**. Relative abundnace of Metacyc pathways, significant by LeFSe ($p < 0.01$).

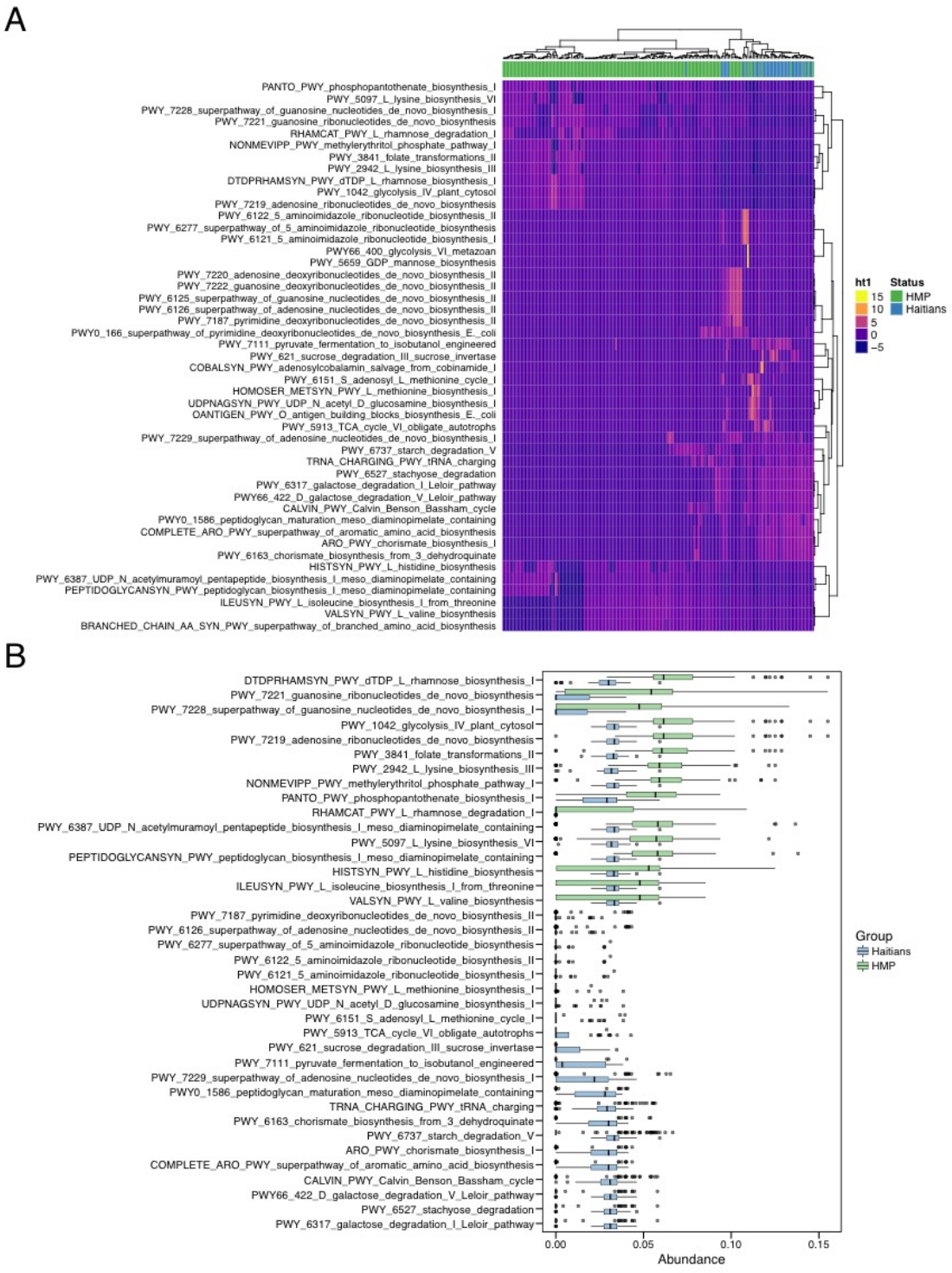


Figure 7: **Metacyc pathway coverage comparing the HMP dataset and Haitian healthy control people. A**. Unsuprivised hiericherical clustering of significantly altered Metacyc pathway coverage. **B**. Relative abundnace of Metacyc pathway coverage, significant by LeFSe ($p< 0.01$).

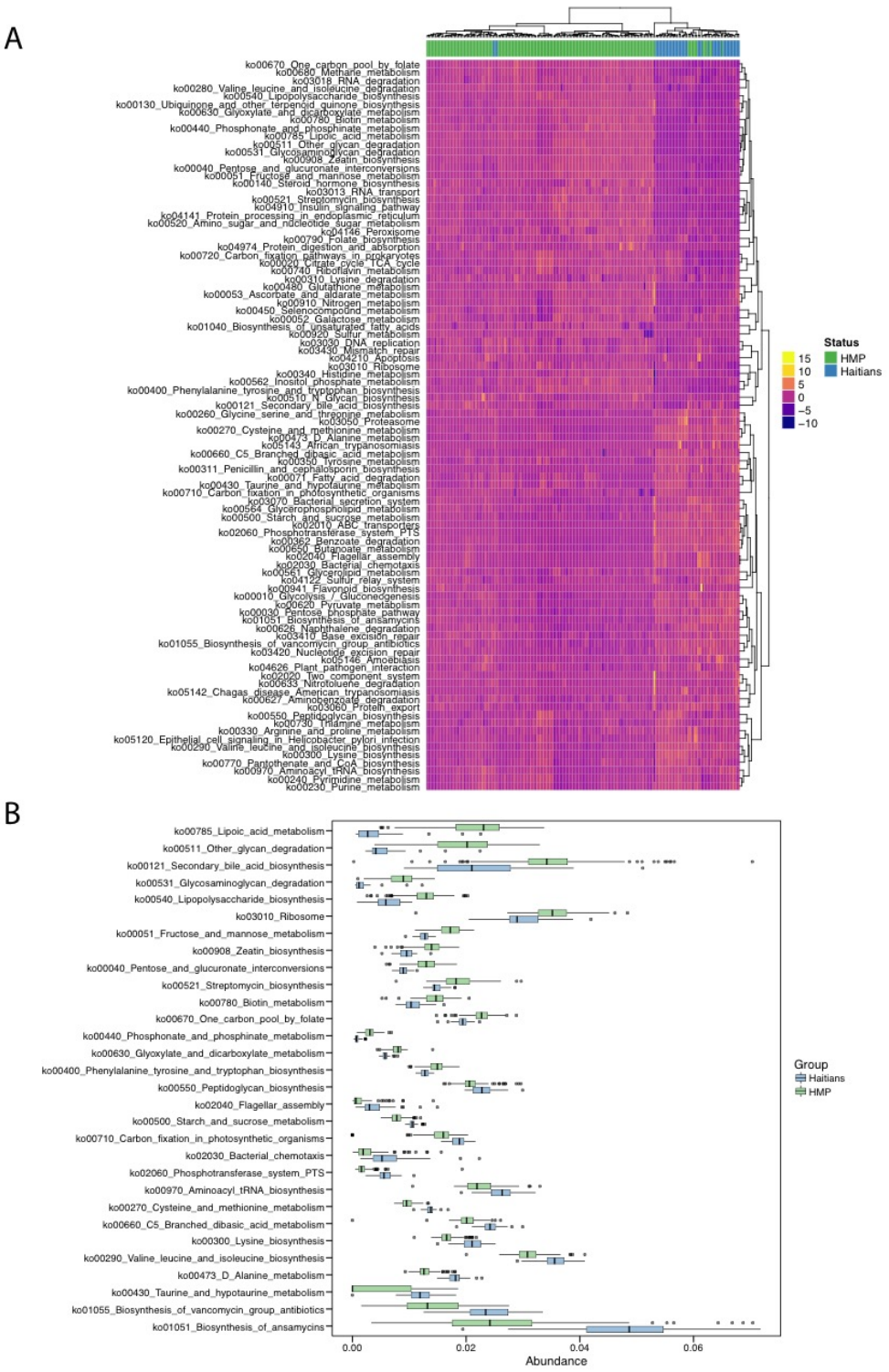


Figure 8: **KEGG pathways comparing the HMP dataset and Haitian healthy control people. A**. Unsuprivised hiericherical clustering of significantly altered KEGG pathways. **B**. Relative abundnace of KEGG pathways, significant by LeFSe ($p < 0.01$).

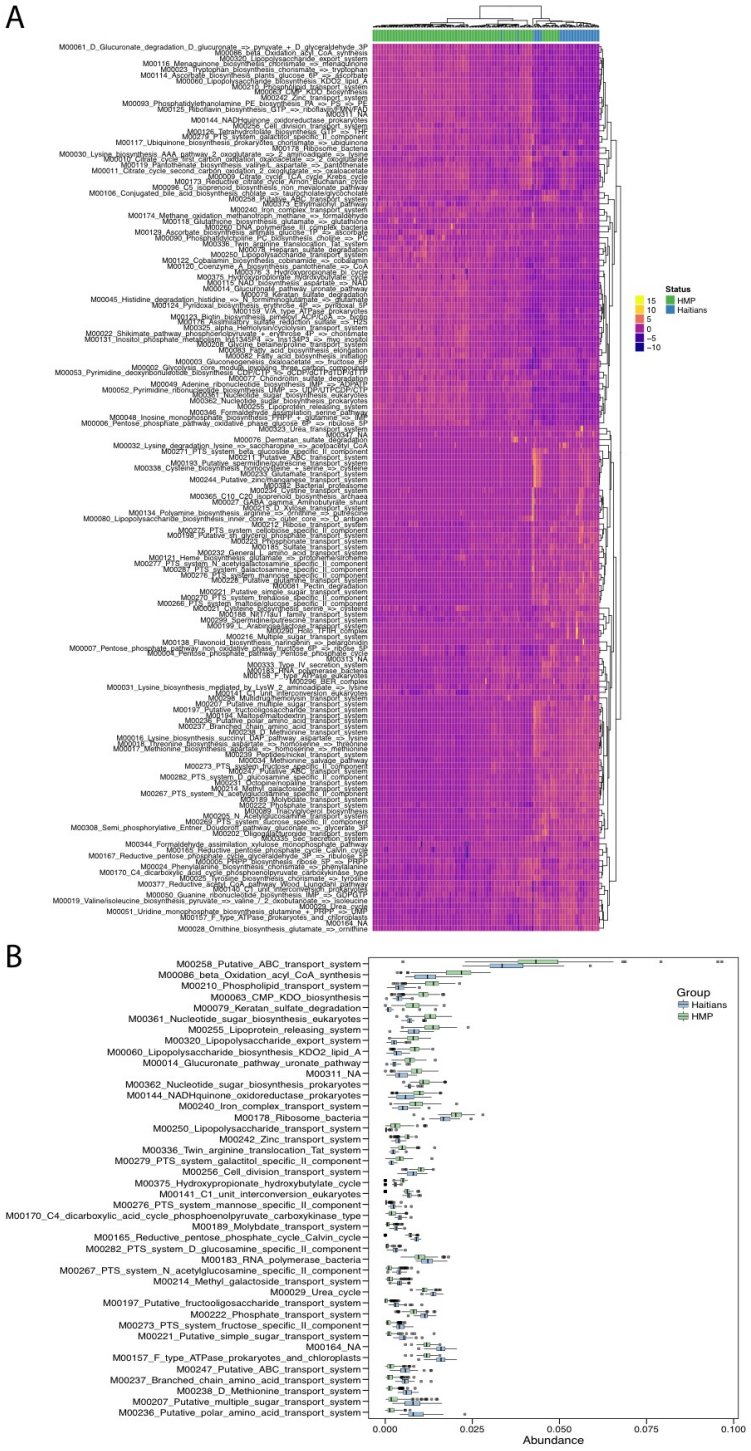


Figure 9: **Metacyc pathway coverage comparing the HMP dataset and Haitian healthy control people. A**. Unsuprivised hiericherical clustering of significantly altered Metacyc pathway coverage. **B**. Relative abundnace of Metacyc pathway coverage, significant by LeFSe ($p < 0.01$).

To determine whether the observed taxonomic compositional differences between Haitians and HMP samples are reflected in distinct functional differences (i.e., coding capacity) we mapped the abundances of Uniref/Metacyc pathway abundance and coverage, as well as KEGG pathways and modules using the HUMAnN2 software pipeline (40). Principal component analysis on pathway differences sharply distinguishes the Haitian and American groups, regardless of their *Mtb* infection status (not shown), so we further investigated these differences using differential statistics (see Methods). Comparison of the HMP and Haitian groups resulted in many pathway differences that were either enriched or reduced in each population, regardless of which statistical method was used to delineate group functional differences (Figures S4 – S7). Additionally, unsupervised hierarchical clustering of samples successfully delineates HMP samples from Haitian samples (Figures 6 – 9). These pathways broadly encompass multiple aspects of bacterial central metabolism and cofactor biosynthesis (see SI text for details), further emphasizing the significant differences between Haitian and American microbiome composition. Taken together, this analysis suggests that the two populations differ not only in terms of microbial composition but also in biochemical functionality.

In summary, we present the first taxonomic characterization of the intestinal microbiome from the Haitian population. The Haitian microbiome differs substantially from that of the United States samples included in the original Human Microbiome Project, and our data therefore augments a sparse literature on microbiome composition in people from both developing and TB endemic countries and regions. The most prominent differences in

microbiomic configuration are the abundance of *Prevotella*, *Bifidobacterium*, and *Ruminococcus* in the Haitian microbiome. The Haitian intestinal microbiome contains a high abundance of *Prevotella*, known to degrade xylan and other starches from dietary fiber, (41) and also observed in high abundance in other "non-Western" or African populations (31, 42, 43). Additionally, we found overlap in the differences in coding capacity between populations from developing and developed countries, which are similar to previous findings. One report in particular compares people recruited from the USA, native people from the Amazonas of Venezuela, and residents of rural Malawian communities (31). In the USA samples, compared to those of the people from Venezuela and Malawi, researchers found an increase in sugar (glycan) degradation, lipoic acid metabolism, and bile salt degradation. This result is strikingly similar to the findings presented here comparing HMP samples to healthy Haitians. Additionally, metagenomic studies on the Hadza hunter-gathers in Africa shows similarities with the Haitian data as well (42).

Pathway differences between HMP and Haitian microbiomes

In the HMP samples, several of the pathways that are differentially abundant generally relate to metabolism. For example, a variety of metabolic pathways including those for lipoic acid, glycan degradation, secondary bile acid synthesis, synthesis of LPS, and secondary polyketide synthesis are more abundant (Figures 6 – 9). Lipoic acid plays a role in lysine biosynthesis and metabolism (also enriched in HMP samples, PWY 2942), which is important for the regulation of energy production via the pyruvate dehydrogenase complex (44). Sugar biosynthesis and degradation, including that for glycosaminoglycan (ko00531) and fructose and mannose (ko00051),

rhamnose (DTDPRHAMSYN PWAY and RHAMCAT PWY), and galactose (PWY 6317 and PWY66 422) are significantly enriched in HMP samples, a finding recapitulated from the original HMP data, and possibly is a result of the American diet of the sampled individuals (45). Secondary bile acids, enriched in the HMP samples (ko00121), have been studied extensively and have both positive and negative health effects, including conferring resistance to colonization by *C. difficile* (46). Finally, secondary polyketide synthesis, specifically of antibiotics like streptomycin (ko00521), were more prevalent in the HMP samples. Synthesis of these molecules results in competition between susceptible and insensitive bacteria in the gut microbiome milieu.

Distinct functional pathways were enriched in the Haitian samples. Two pathways in particular, bacterial chemotaxis (ko02030) and flagellar assembly (ko02040) both related to mobility of bacteria and have been implicated in inflammation (47). Several pathways involved in amino acid biosynthesis including alanine (ko00473), valine, leucine, and isoleucine (ko00290), cysteine and methionine (ko00270), lysine (ko00300), methionine (HOMOSER METSYN PWY), aromatic amino acids (COMPLETE ARO PWY), and taurine and hypotaurine (ko00430). Finally, several enzymatic cofactors like S-adenosyl methionine utilization and recycling (PWY 6151), and chorismate biosynthesis (PWY 6163) are overrepresented in the Haitian samples (Figures 6 – 9).

Taken together, these results indicate that functional pathway differences between Haitian and HMP samples are significant, and recapitulate previous

findings where populations with non-Western diets and lifestyles were compared to those with them (31, 37, 42).

Antimycobacterial treatment alters intestinal microbiota taxonomic composition during treatment without affecting overall diversity

It is currently unknown if and how the standard regimen of antimycobacterial HRZE therapy affects the taxonomic composition of the intestinal microbiota, as none of these drugs have been studied, alone or in combination, for their effects in humans. We used a cross-sectional enrollment design to determine if and how the intestinal microbiota changes in response to HRZE therapy in people from Haiti. All cases were recruited from the Haitian community, had recently been diagnosed with microbiologically-confirmed *M. tuberculosis* infection, and have been on a combination regimen of HRZE antibiotics for at least two weeks (Treatment group, Table 1). We compared *Mtb* uninfected and LTBI controls with no known recent or historical *Mtb* exposure. Using either a DESeq or LeFSe analytical pipeline (see Methods), we were unable to detect any microbiomic differences between *Mtb* uninfected and LTBI individuals (data not shown), and quantitative Permanova analysis further confirm this finding (Table 2, see Methods). Thus, we conclude that LTBI has no detectable effect on intestinal microbiome composition.

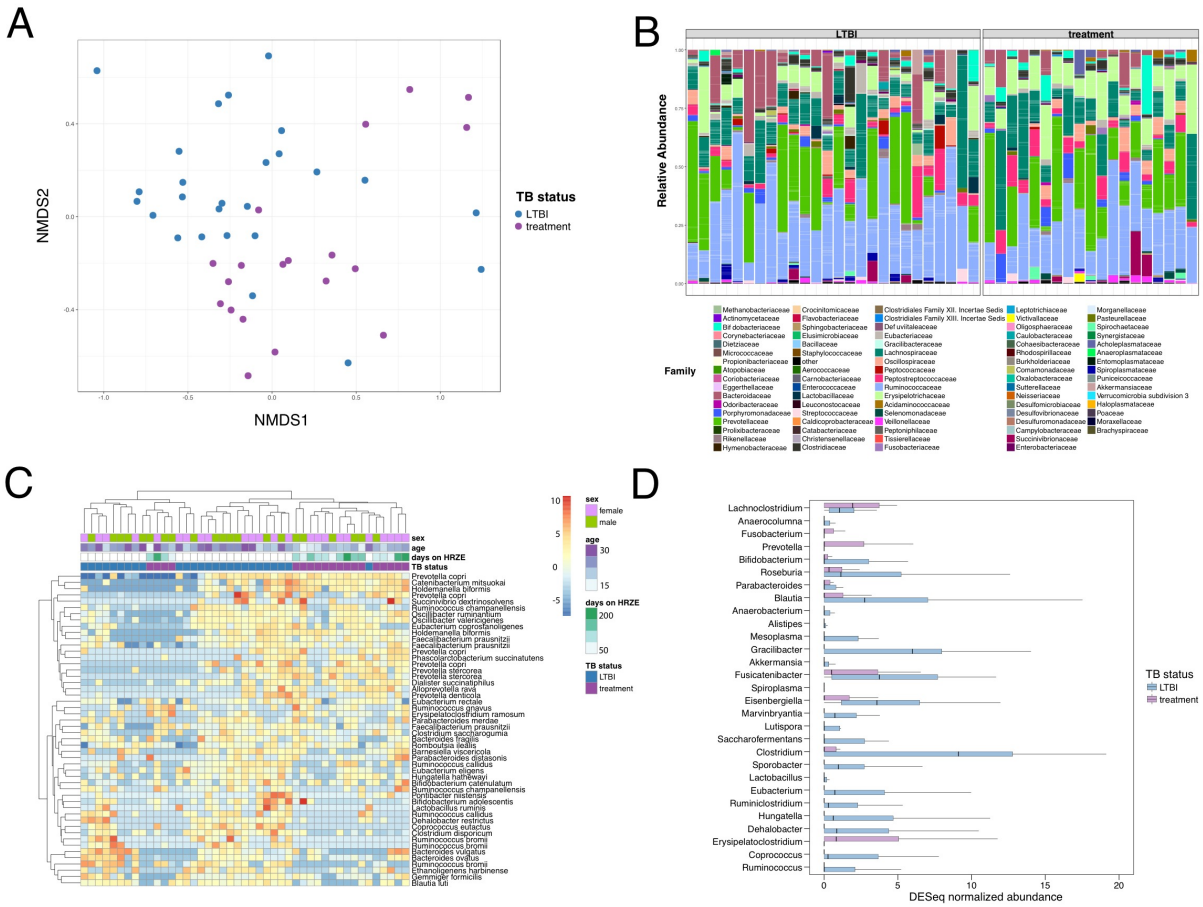


Figure 10: HRZE treatment perturbs the taxonomic structure of the microbiome. A. NMDS ordination of HRZE treated subjects (purple) or LTBI controls (blue) based on 16S rDNA sequencing **B.** Family taxonomic distribution of the intestinal microbiota from subjects with LTBI and subjects with TB on treatment. **C.** Heatmap of the top 50 most abundant taxa generated with DESeq2 showing unsupervised clustering of TB cases on treatment vs. LTBI controls. Age and sex are also shown but were not accounted for the in DESeq model. **D.** Taxonomic abundance profiling comparing treatment vs LTBI participants using LeFSe to determine differentially abundant Genera. Box and whisker plots of differentially abundant genera are shown based on the DESeq normalized data. Plots show the first and third quartiles of the abundance data, the line represents the median, and the whiskers show 1.5 times the value of the interquartile range. Outliers are shown as dots.

In comparing individuals on HRZE treatment with LTBI controls, we were surprised to find that overall microbiomic diversity of volunteers treated for active TB with HRZE did not differ from controls, as measured by the Shannon diversity index, despite being on therapy for an average of 3.2 months (Figure 11B, Table 1). This lack of effect on diversity is in stark contrast to the dramatic and rapid loss of diversity seen with broadly acting antimicrobials (48), and is consistent with the narrower spectrum of antimycobacterial agents used for treatment of TB. We additionally compared the Treatment group to the *Mtb* uninfected controls, and found similar differences (data not shown). Nevertheless, there was a highly significant loss of specific taxa of bacteria with antimycobacterial treatment (i.e., the number of unique OTUs). The number of observed OTUs was significantly lower in the treatment group compared to the *Mtb* uninfected (two-samples t-test; $p = 0.0074$) or LTBI-treatment (two-samples t-test; $p = 0.0122$) control groups (Figure 11C).

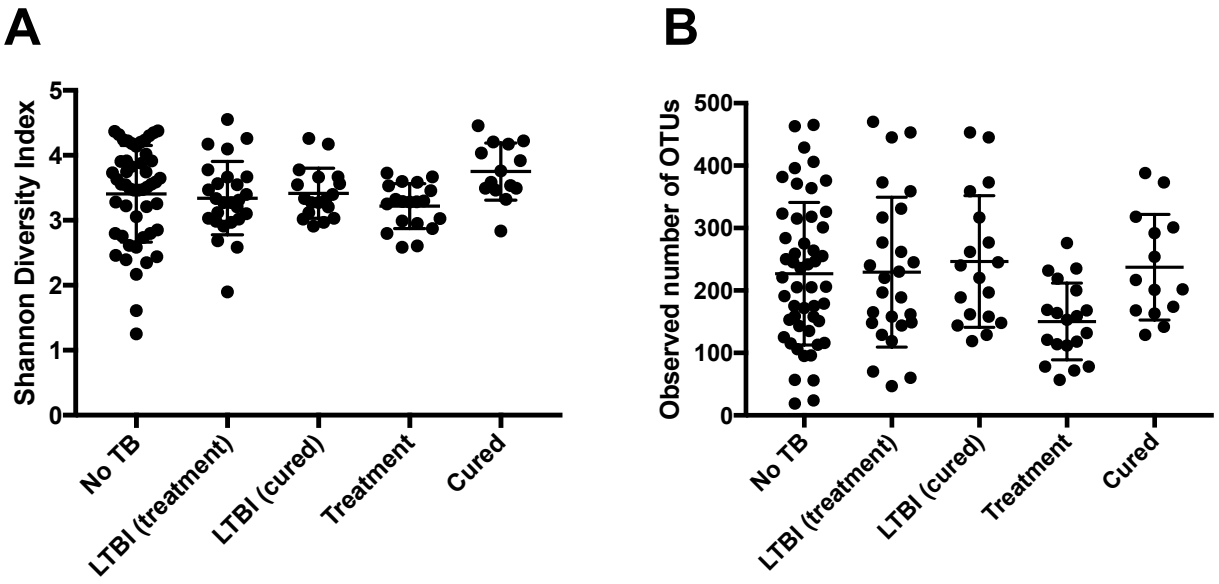


Figure 11: Taxonomic richness and diversity in Haitian cohorts used in this study. A. Shannon diversity index measured for all groups used in this study, based on 16S rDNA sequencing data. The LTBI (treatment) group indicates subjects who are the age-matched controls for the treatment group, and the LTBI (cured) group indicates the age-matched controls for the cured group. **B.** Raw number of observed OTUs clustered at 97% similarity for the clinical groups in this study.

Despite the lack of an effect on overall diversity, closer examination of specific microbiomic changes associated with treatment revealed substantial changes in the treated group compared to either the *Mtb* uninfected control group or LTBI controls. Principal coordinate analysis on differential taxonomic diversity calculated either using 16S (Figure 10A, HRZE vs LTBI) or metagenomic DNA sequencing (Figure 12A, HRZE vs healthy) clearly demarcates people who are on treatment from those who are not, indicating substantial antimycobacterial-induced perturbations. A quantitative test of what variables account for variance between two groups (Permanova) confirms differences between the treatment and healthy groups ($p=0.023$, Table 2, see Methods). We were unable to find differences based on sex of the individual ($p=0.44$, Table 2), which we accounted for due to the unbalanced gender distribution of the groups, and in all analyses performed, treatment contributed the largest to the observed variance (Table 2). Unsupervised hierarchical clustering of the 40 most abundant OTUs revealed similar distinctions between treated and control subjects (Figure 10C), and revealed that treated people clustered into two major groups. The predominant cluster in the treated group is represented by an overabundance of *Prevotella copri* and *P. denticola* (as representative organisms), and the other by its depletion (Figure 10C).

Treatment vs LTBI			
Unifrac	**F.Model**	**R^2**	**Pr(>F)**
TB status	3.6433	0.07804	0.001
age	1.0898	0.02334	0.297
sex	0.9546	0.02045	0.446
Bray	**F.Model**	**R^2**	**Pr(>F)**
TB status	1.93533	0.04266	0.019
age	1.43693	0.03167	0.086
sex	0.99648	0.02196	0.414
Jaccard	**F.Model**	**R^2**	**Pr(>F)**
TB status	1.63854	0.03652	0.007
age	1.24754	0.0278	0.098
sex	0.98374	0.02192	0.457
No TB vs LTBI (30 in each group, Haitian			
Unifrac	**F.Model**	**R^2**	**Pr(>F)**
TB status	0.64947	0.00961	0.924
age	1.29284	0.01913	0.164
sex	0.63887	0.00945	0.938
Bray	**F.Model**	**R^2**	**Pr(>F)**
TB status	0.75282	0.01099	0.823
age	1.69801	0.02478	0.035
sex	1.0744	0.01568	0.296
Jaccard	**F.Model**	**R^2**	**Pr(>F)**
TB status	0.8399	0.01229	0.837
age	1.4459	0.02116	0.029
sex	1.0342	0.01514	0.337
Cured vs LTBI			
Unifrac	**F.Model**	**R^2**	**Pr(>F)**
TB status	1.1042	0.03532	0.272
age	1.1169	0.03572	0.277
sex	1.0464	0.03347	0.359
Bray	**F.Model**	**R^2**	**Pr(>F)**
TB status	1.89019	0.0587	0.009
age	1.46281	0.04543	0.042
sex	0.84962	0.02638	0.671
Jaccard	**F.Model**	**R^2**	**Pr(>F)**
TB status	1.5714	0.04957	0.007
age	1.2373	0.03903	0.085
sex	0.89447	0.02821	0.725

Table 2: PERMANOVA test results for the treatment vs LTBI and cured vs LTBI comparisons, as well as the comparison between No TB and LTBI cohorts as a control, using three different distance metrics (Bray Curtis distances, Jaccard distances, and unweighted Unifrac distances). This demonstrates what factors (age, sex, TB status) contribute to the observed variation between the cohorts.

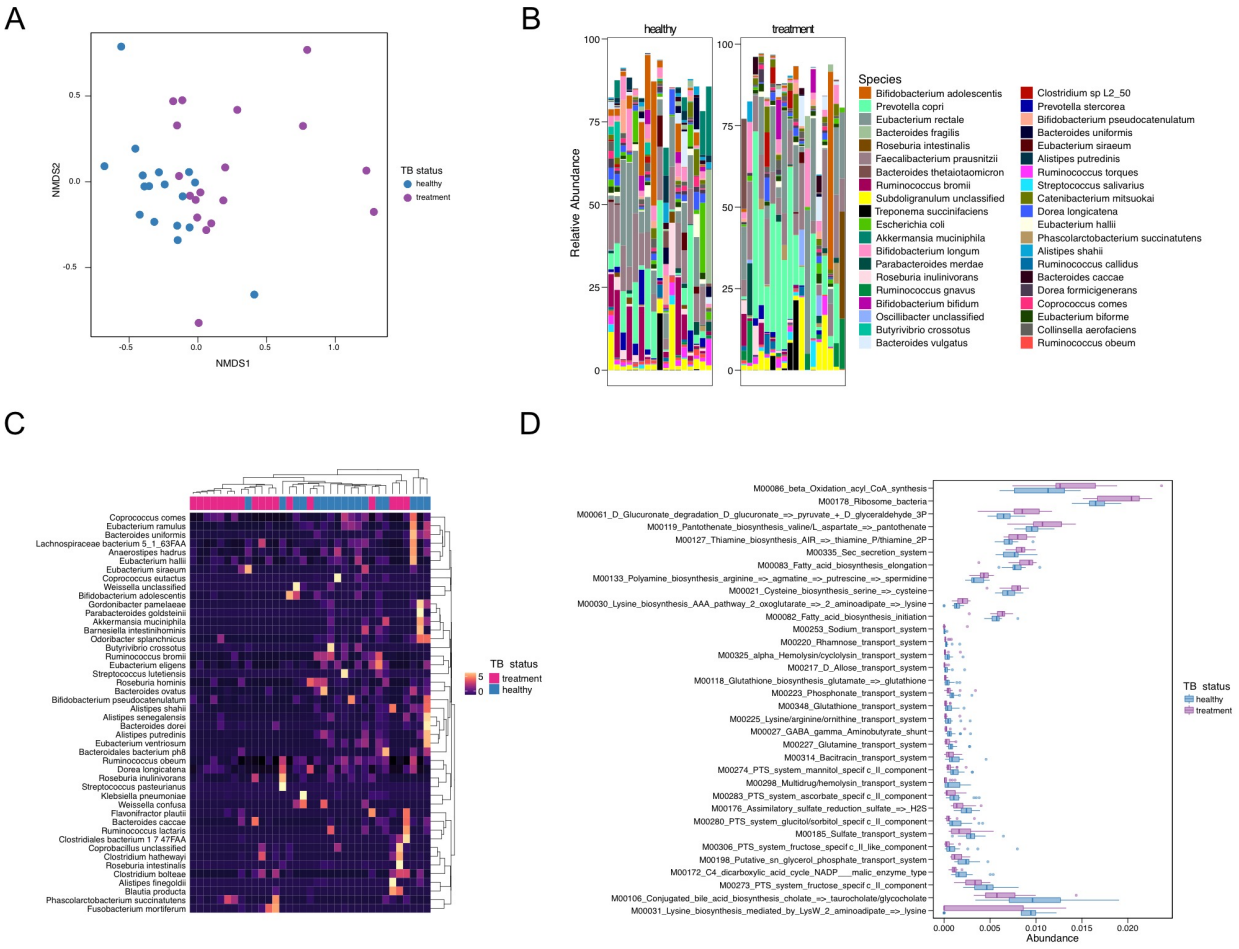


Figure 12: Taxonomic and biochemical microbiomic perturbation induced by HRZE. A. NMDS ordination plot on metagenomic taxonomy data demonstrating microbiomic differences between healthy individuals and subjects on HRZE treatment. For this comparison, the healthy group consists of LTBI and *Mtb* uninfected subjects. **B.** Comparative abundance plots between healthy Haitian individuals and cases on HRZE treatment showing the most abundant species. **C**. Unsupervised hiericheral clustering of significantly altered taxa from species-level metagenomic data. **D.** Abundance of significantly different KEGG modules between healthy volunteers and cases on treatment.

To determine which taxa are significantly affected by HRZE therapy, we used both the LEfSe and DESeq2 pipelines with 16S rDNA sequencing data (see Methods) and observed dramatic changes at the Genus level. Subjects taking HRZE have, on average, an enrichment of Erysipelatoclostridium (8.8% in the Treatment group vs. 3.4% in LTBI controls), Fusobacterium (0.08% in the Treatment group vs. 0.54% in LTBI controls) and Prevotella (7.11% in the Treatment group vs. 3.76% in LTBI controls). HRZE treatment resulted in a 10-fold reduction in *Blautia*, a more than a 200-fold reduction in *Lactobacillus* and *Coprococcus*, and a 675-fold decrease in *Ruminococcus* compared to the LTBI group (Figure 10D). In the Phylum Actinobacteria, there was a nearly 20-fold depletion of *Bifidobacterium* (Figure 10D).

Although 16S rDNA based taxonomic analysis provides useful information about the relative abundances of microbiome constituents, and clearly HRZE treatment induces a profound alteration in the compositional structure of the gut microbiome, this technique cannot directly interrogate the coding capacity of the microbiota, as the gene content of taxonomically identical OTUs can differ substantially. To ask whether HRZE treatment altered metabolic coding capacity of the microbiota, we performed metagenomic sequencing and analyzed the coding capacity of HRZE treated subjects vs a control group consisting of LTBI and TB uninfected subjects (designated healthy in Figure 12). Taxonomic analysis by LEfSe revealed similar depletion of *Ruminococcus*, *Eubacterium*, and other species as was observed with 16S profiling, and differentially abundant species by LeFSe clustered well in a heatmap analysis (Figure 12C). Using the HUMANn2 software pipeline (40) we observed both enrichment and depletion of biochemical pathways within

the microbiome of treated volunteers, with the most dramatic changes including overabundance of fatty acid oxidation and vitamin biosynthesis and depletion of conjugated bile acid biosynthesis with HRZE treatment (Figure 12D, Figure S10-S12). These results further confirm that HRZE treatment broadly perturbs the microbiome, both on the taxonomic and functional levels.

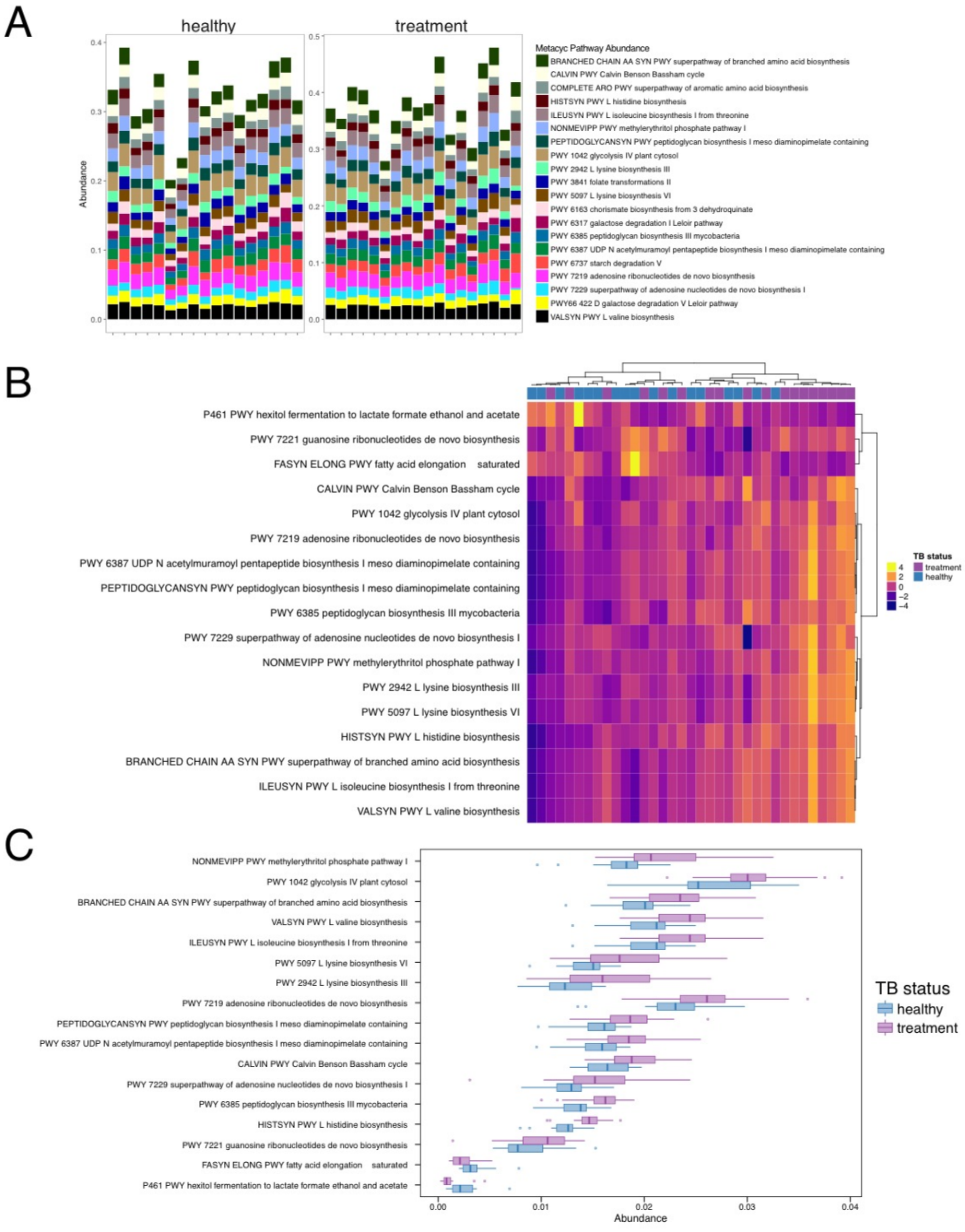


Figure 13: Metacyc pathway abundance calculated against the UniRef50 gene database comparing healthy (uninfected and LTBI) to treated people. A. Top 20 most abundant pathways in the healthy and treated groups. **B**. Unsuprivised hiericherical clustering of significantly altered Metacyc pathways. **C**. Relative abundnace of Metacyc pathways, significant by LeFSe ($p < 0.05$), between healthy and treatment groups using Metacyc pathway coverage.

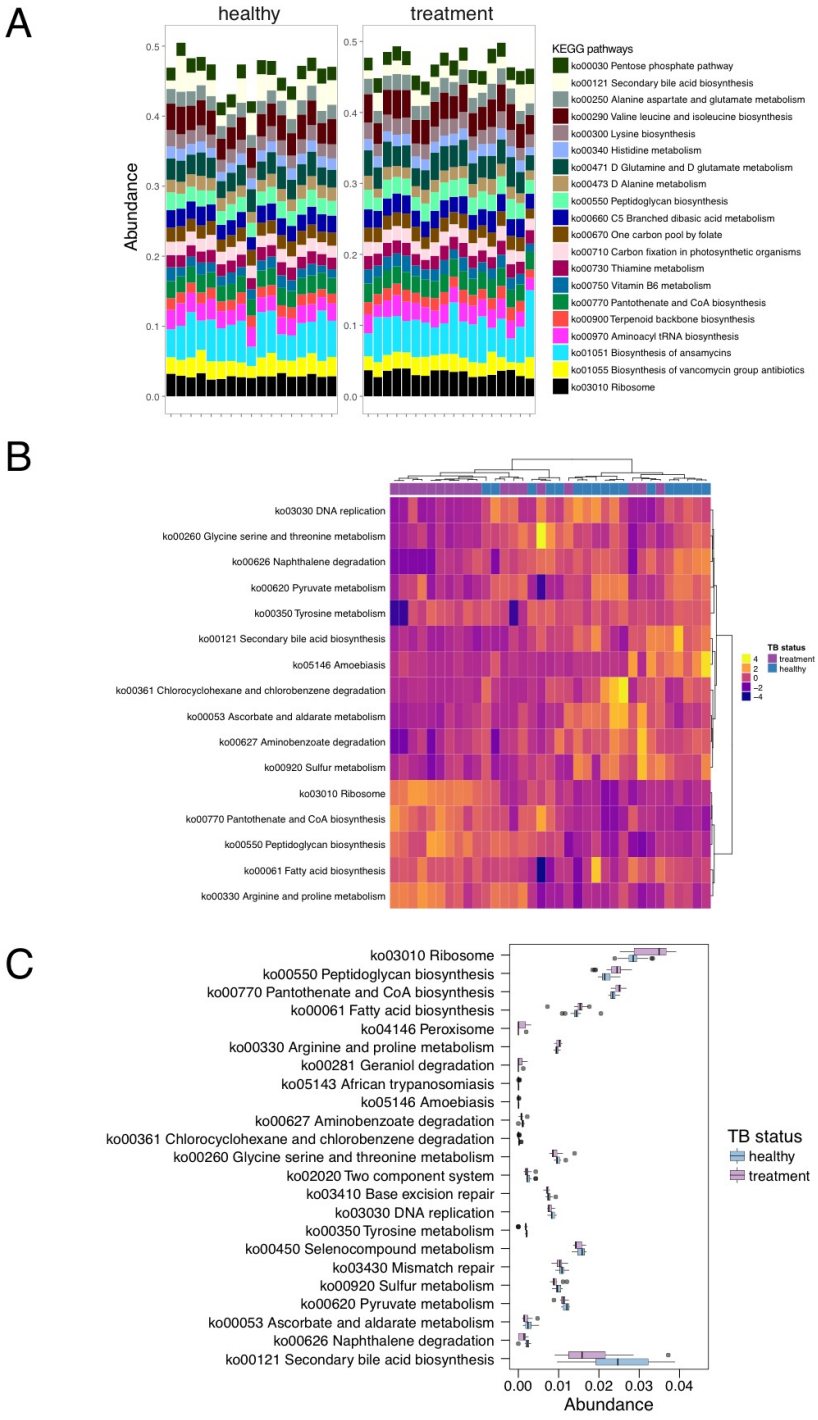


Figure 14: KEGG pathways comparing healthy (uninfected and LTBI) to treated people. **A**. Top 20 most abundant pathways in the healthy and treated groups. **B**. Unsuprivised hiericherical clustering of significantly altered KEGG pathways. **C**. Relative abundnace of KEGG pathways, significant by LeFSe ($p < 0.05$), between healthy and treatment groups using KEGG pathways.

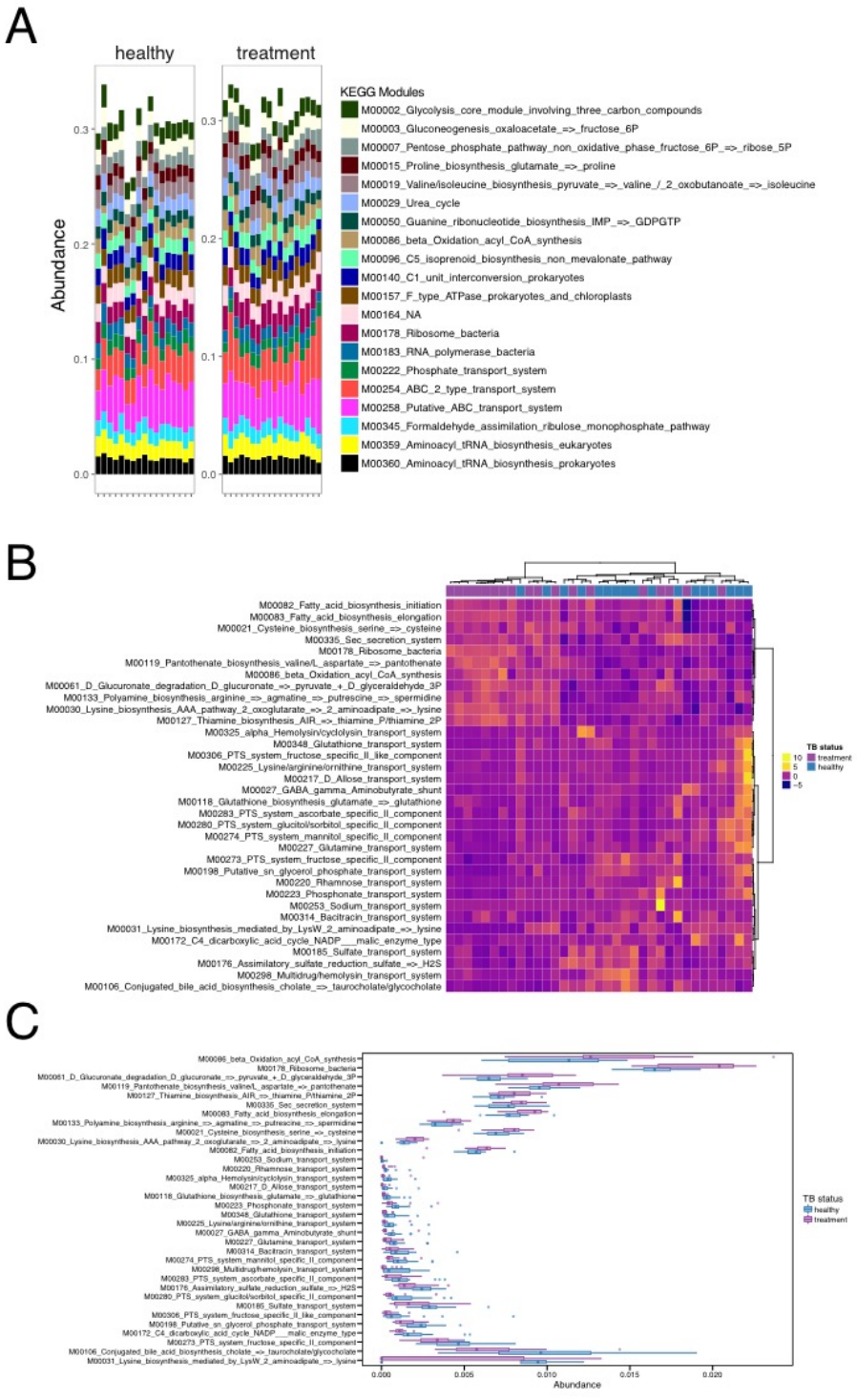


Figure 15: KEGG modules comparing healthy (uninfected and LTBI) to treated people. **A**. Top 20 most abundant pathways in the healthy and treated groups. **B**. Unsuprivised hiericherical clustering of significantly altered KEGG modules in patients on HRZE treatmnet. **C**. Relative abundnace of KEGG modules between healthy and treatment groups.

TB treatment is associated with a lasting intestinal microbiome dysbiosis

Compelling evidence suggests that alterations in intestinal microbiome composition through antibiotic perturbation can result in the development of novel ecological states with preliminarily characterized, but poorly defined, health outcomes (49). The data presented above clearly indicates that HRZE treatment induces a detectable microbiomic perturbation, which could be long lasting given the prolonged duration of antibiotic exposure during TB treatment. To determine whether the dysbiosis induced by antimycobacterials persists after discontinuation of therapy, we recruited a group of cured TB cases who had completed 6-month HRZE therapy for active Tuberculosis and compared their microbiome composition to age matched LTBI subjects as controls. The average time since completion of treatment in the cured group was 1.2 years (mean 417 days, range 34-1202 days, Table 1). We found that taxonomic alpha diversity in the cured subjects remained at levels comparable with those in the LTBI control groups (Figure 11B). Indeed, using a Mann-Whitney unpaired t-test, there is a slight significant increase in Shannon diversity for the cured cases ($p=0.0336$, Figure 11B). However, the intestinal microbiomes of cured TB cases were distinguishable from LTBI controls when examined by detrended correspondence principal component analysis (Figure 16A), indicating that HRZE therapy has long-lasting effects on microbiome composition (Figure 16B).

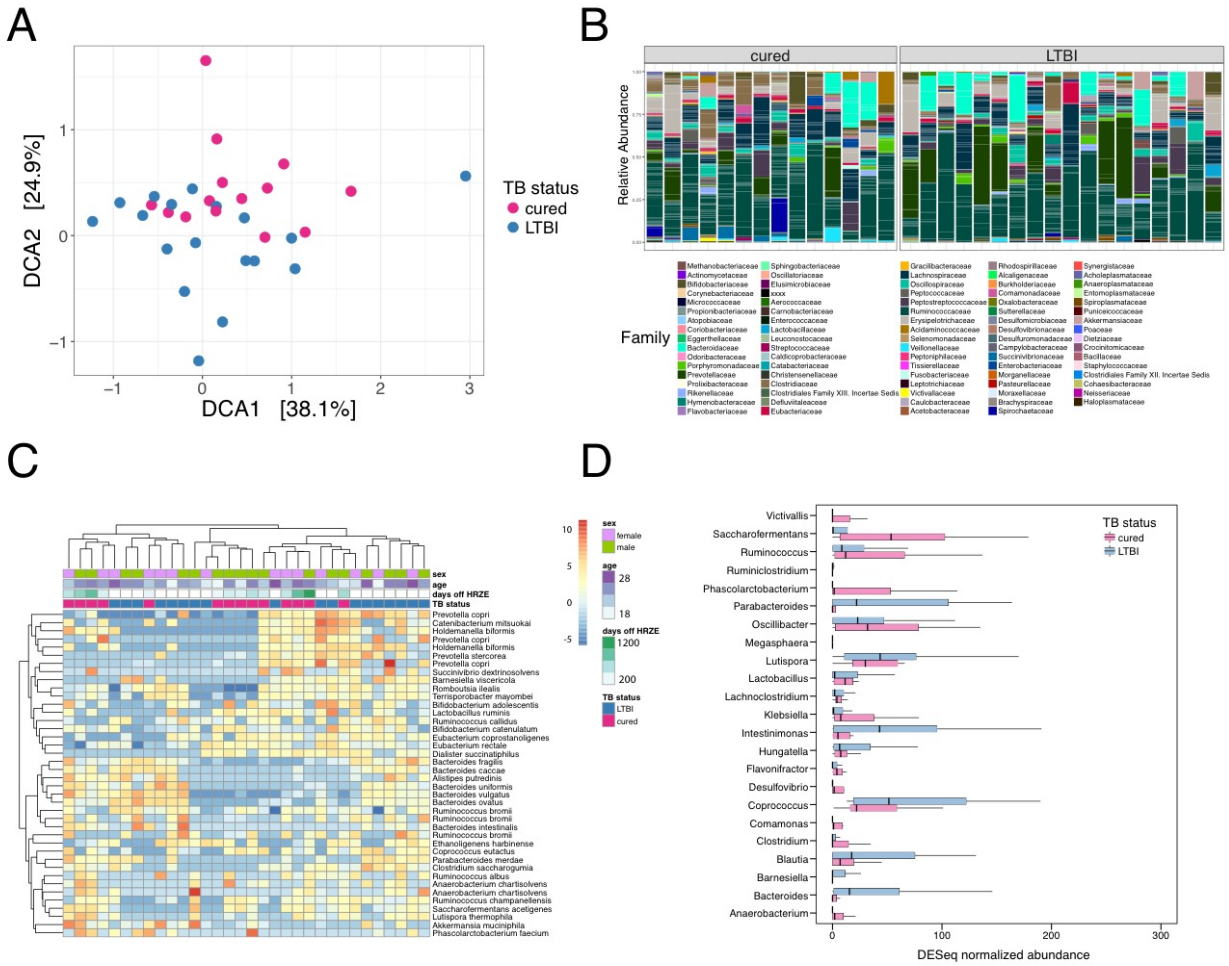


Figure 16: TB treatment induces a lasting alteration in microbiome structure.
A. DCA ordination plot of cured cases compared to LTBI controls **B.** Family level taxonomic distribution of the intestinal microbiota from subjects with LTBI or who are cured. **C.** Heatmap of the 40 most abundant taxa generated with DESeq2 showing unsupervised clustering of cured vs. LTBI subjects. Age and sex are also shown but were not accounted for the in DESeq model. The number of days that each patient has been off treatment is also shown. **D.** Taxonomic abundance profiling comparing cured vs LTBI subjects. Taxa are significant from LeFSe ($p < 0.05$ and LDA cutoff > 3.0).

To quantitatively asses the principal coordinate analysis result, a Permanova test was used to confirm differences between the two populations (p=0.046). Again, given the uneven distribution of sex in our cohorts, we performed a Permanova test on sex, but are unable to detect differences between the populations using this measure (p=0.98). Unsupervised hierarchical clustering of the 40 most abundant OTUs revealed that in general cured individuals clustered away from LTBI controls (Figure 16C). We identified some degree of heterogeneity in cured individual's microbiomes, since there is not perfect clustering, but this in no way rules out the presence of altered taxonomic or functional microbiomic states. We asked whether this heterogeneity in the cured people was correlated with the duration of time since treatment and found that both distantly and recently cured people clustered well together, indicating that the persistent dysbiosis observed is not simply a marker for time since cure (Figure 16C).

We used LeFSe and DESeq to analyze differences between cured individuals and LTBI age matched controls. Cured people were depleted in the Bacteroidetes Phylum of *Bacteroides*, *Barnesiella*, and *Blautia* (Figure 17D). Using metagenomic DNA sequencing and LeFSe, we found that *Enterobacter cloacae*, *Phascolarctobacterium succinatutens*, *Methanobrevibacter smithii*, *Bilophila,* and *Parabacteroides* are biomarkers of cured individuals. Although DCA performed with species level abundances on metagenomic data failed to cluster healthy and treated cases (Figure 17A), and community structure was grossly similar (Figure 17B), pathway abundance analysis revealed that cured cases demonstrated altered coding capacity compared to controls. (Figure 17D and Figures S13 – S15). As with the previous comparison of healthy

individuals to cases on HRZE treatment, the perturbed pathways represent diverse microbial functions including sugar biosynthesis, protein secretion, and central metabolism. We conclude that TB treatment results in long term, and perhaps permanent, taxonomic, metagenomic, and biochemical consequences via perturbation of the intestinal microbiome.

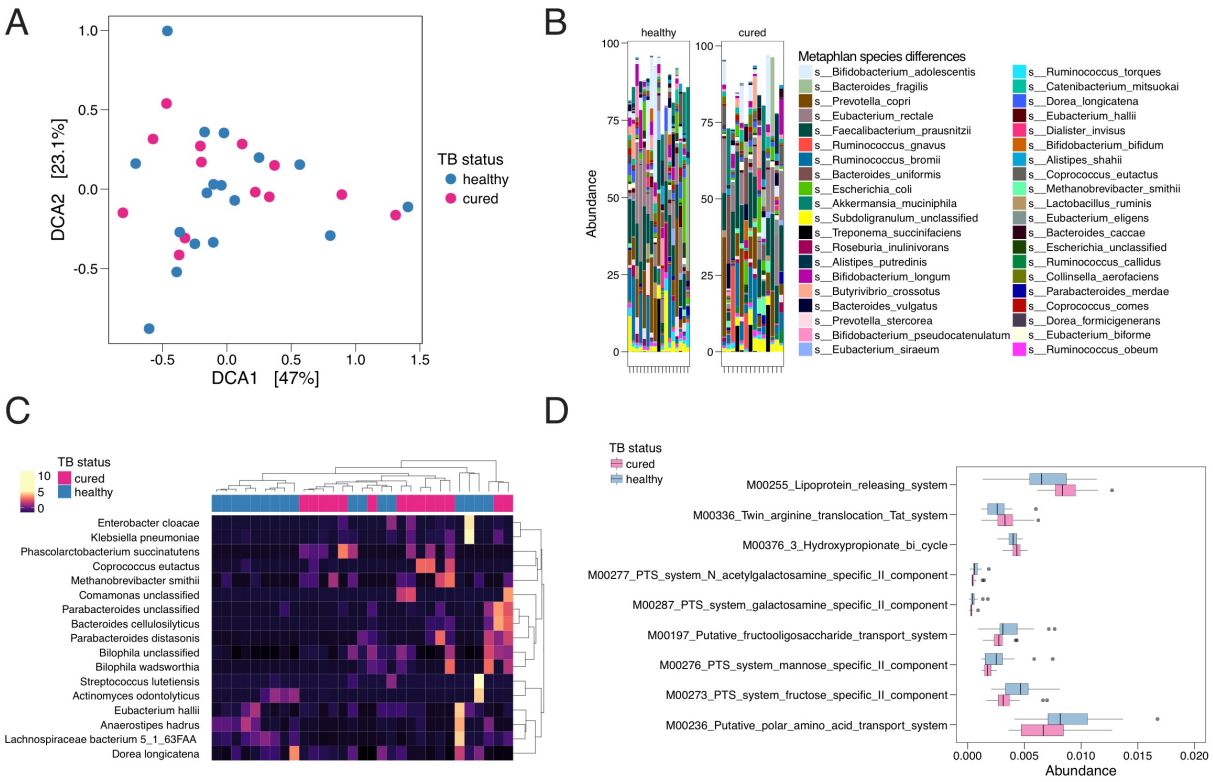


Figure 17: TB treatment induces a lasting alteration in microbiome structure and function. A. DCA ordination plot on metagenomic taxonomy data in healthy (*Mtb*-uninfected and LTBI community controls) and cured individuals. **B.** Comparative abundance plots between healthy Haitian individuals and cured subjects showing the top 40 most abundant species between the two groups. **C.** Unsupervised hierarchical clustering of significantly altered taxa. **D**. Abundance of significantly different KEGG modules between healthy and cured subjects.

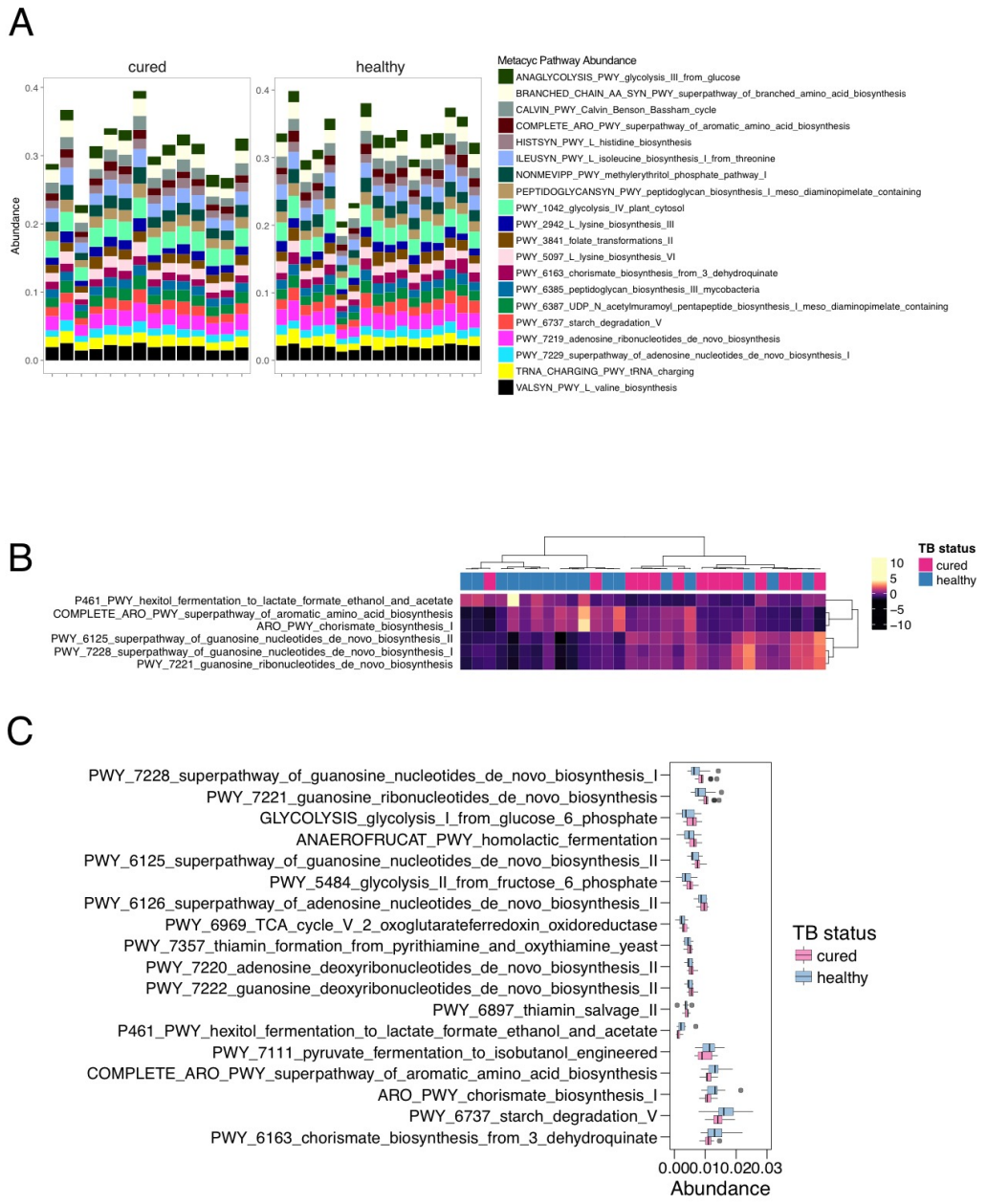


Figure 18: Metacyc pathway abundance calculated against the UniRef50 gene database comparing healthy to cured people. A. Top 20 most abundant pathways in the healthy and cured groups. **B**. Unsuprivised hiericherical clustering of significantly altered Metacyc pathways. **C**. Relative abundnace of Metacyc pathways, significant by LeFSe ($p < 0.05$), between healthy and cured groups.

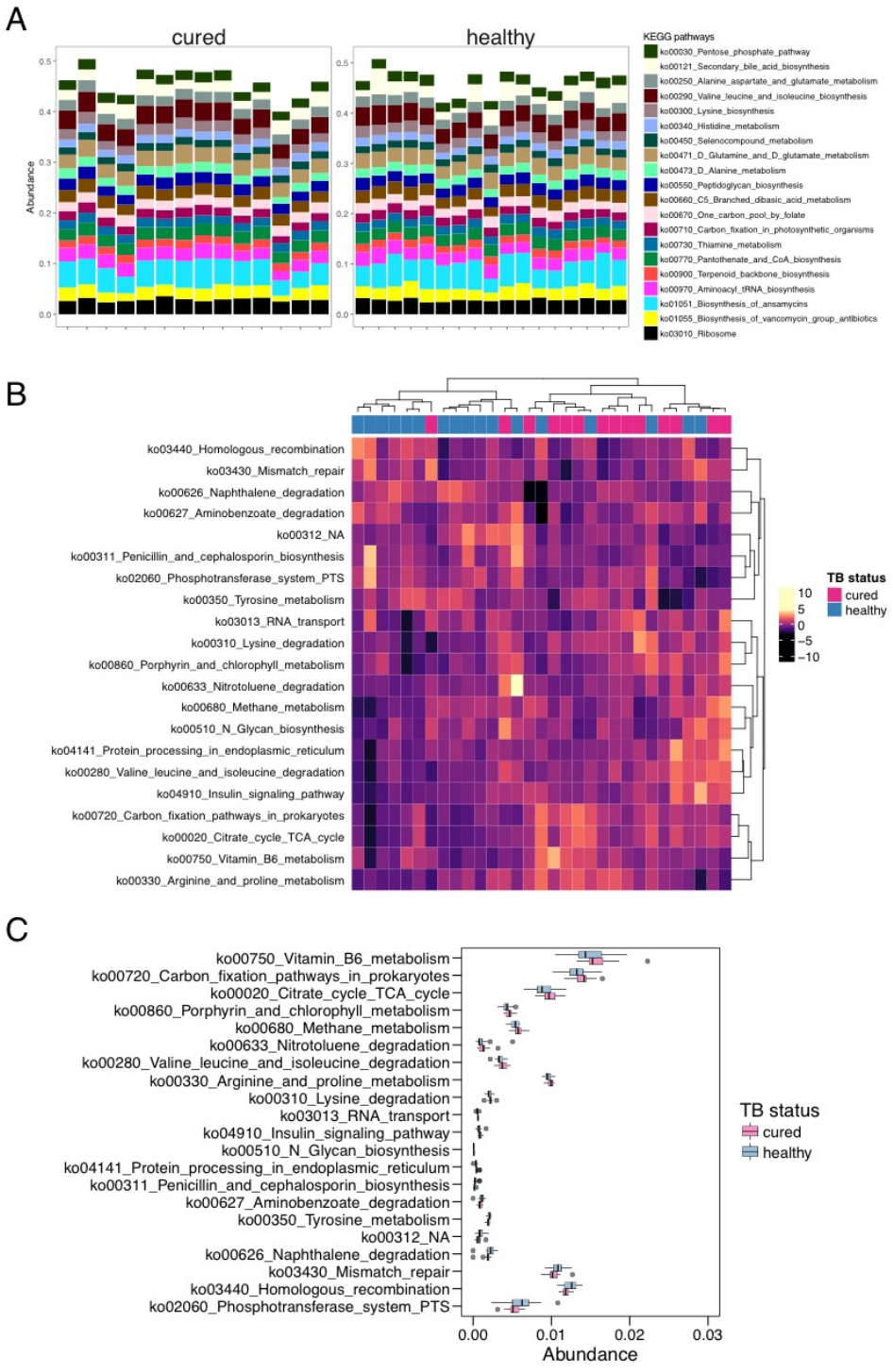


Figure 19: KEGG pathways comparing healthy (uninfected and LTBI) to cured people. **A.** KEGG pathways comparing healthy to cured people. **A**. Top 20 most abundant KEGG pathways in the healthy and cured groups. **B**. Unsuprivised hiericherical clustering of significantly altered KEGG pathways. **C**. Relative abundnace of KEGG pathways, significant by LeFSe ($p < 0.1$), between healthy and cured groups.

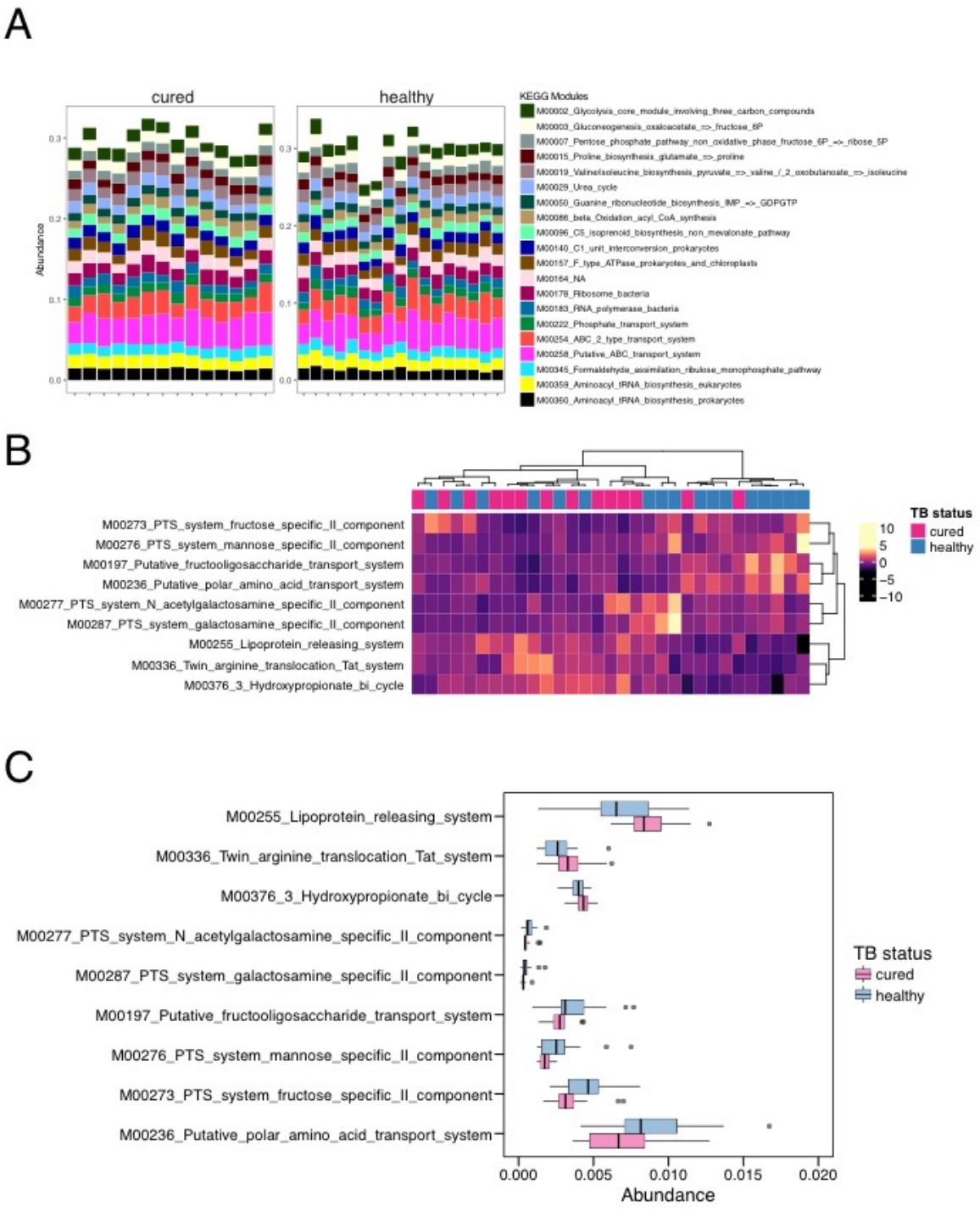


Figure 20: KEGG modules comparing healthy (uninfected and LTBI) to cured people. A. KEGG modules comparing healthy to cured people. **A**. Top 20 most abundant KEGG modules in the healthy and cured groups. **B**. Unsuprivised hiericherical clustering of significantly altered KEGG modules. **C**. Relative abundnace of KEGG pathways, significant by LeFSe ($p < 0.1$), between healthy and cured groups.

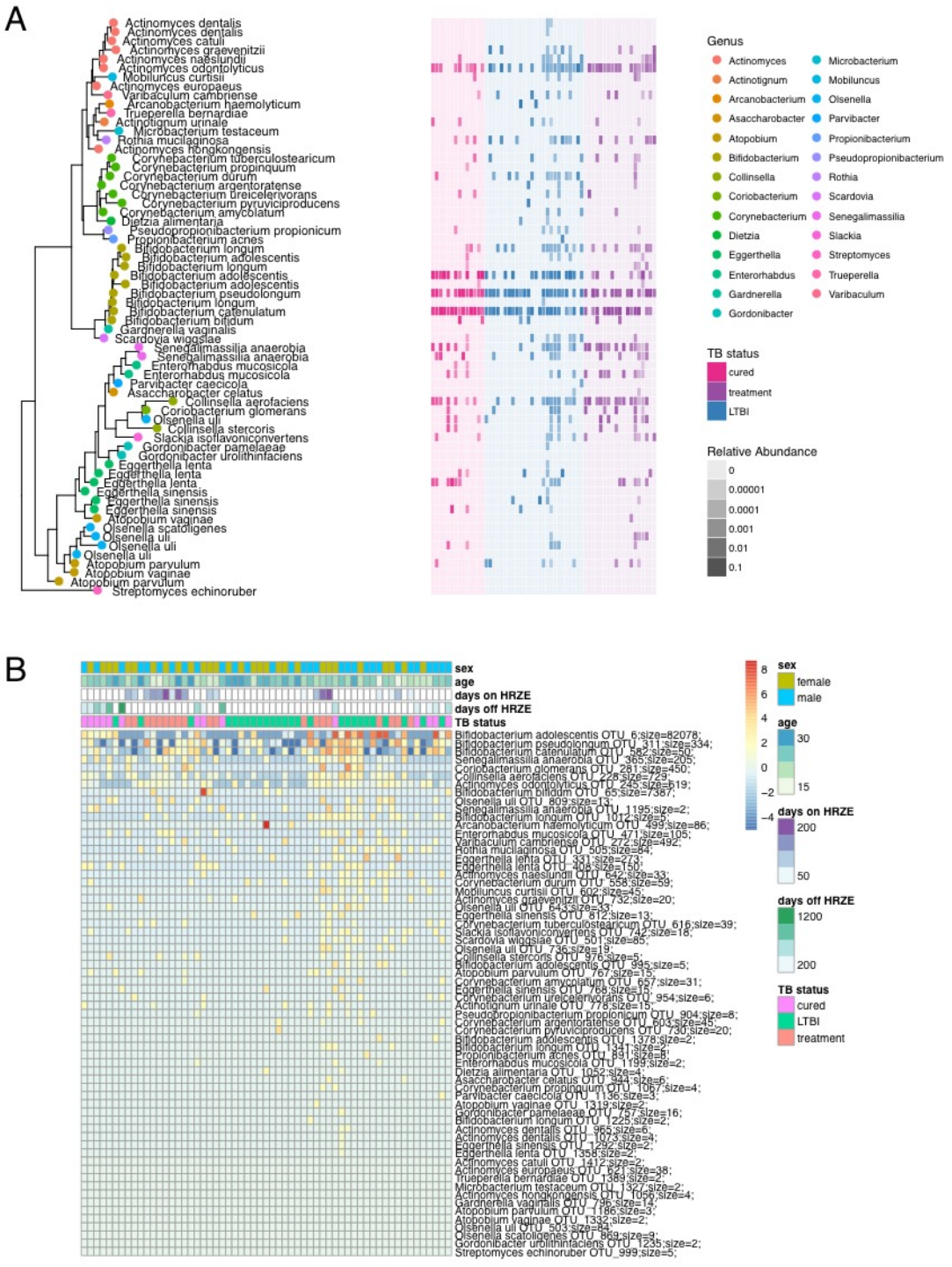


Figure 21: Profile of Actinobacteria in the Haitian microbiome. A. Scaled relative abundnace of all OTUs from the phylum Actinobacteria from LTBI, treatment, and cured individuals shown on a phylogenetic tree layout. **B**. DESeq variance stabilized transformed relative abundances of Actinobacterial OTUs sampled from the intestinal microbiomes of LTBI, treatment, and cured individuals.

CHAPTER 3 – DISCUSSION

We present the first characterization of the short and long term effects of standard HRZE TB antibiotic treatment on the intestinal microbiome. Antibiotics are recognized to perturb the composition of the intestinal microbiome, and their use has been associated with potentially deleterious consequences (21). This perturbation is best documented for broad spectrum antibacterial agents which are active against swaths of bacterial microbiome constituents. As such, broad spectrum antimicrobials like the fluoroquinolone ciprofloxacin (27) may cause rapid loss of overall diversity and disruption of the microbiome's ability to resist pathogenic colonization, which can predispose to disease such as *Enterococcus* domination and *C. difficile* infection (50). However, treatment of TB employs antimicrobial agents with narrower spectrums of activity. Although Rifampin is used for non-mycobacterial infections, Isoniazid, Pyrazinamide, and Ethambutol are only prescribed for TB and are activated by and/or target mycobacterial proteins not widely distributed throughout the bacterial Kingdom. Thus, we think it is important to understand the potential consequences of HRZE treatment on microbiome composition and function in the short and the long term. Since the turn on the millennium, 49 million people have been treated with this or similar regimens, comprising almost 10 billion doses each drug, making TB treatment the most widely administered antibiotic regimen in the world.

Our data indicates that the narrow spectrum of the TB treatment regimen is reflected in the preserved overall diversity in HRZE treated cases. However, this relatively gross measure of perturbation fails to capture the profound effects of HRZE on specific components of the microbiome. Most dramatic is

the depletion of multiple species of *Ruminococcus*, *Eubacterium*, *Lactobacillus*, and *Bacteroides* along with a simultaneous increase of *Erysipeloclostridium* and *Prevotella*. The consequences of this HRZE-induced taxonomic perturbation are unknown at present, but several of these bacteria have been associated with immune-inducing phenotypes relevant to TB immunity. *Bacterioides* (depleted in treated and cured subjects) polysaccharide can modulate host inflammatory responses in mice (51). *Ruminococcus* and *Coprococcus* are two of the most dramatically depleted phyla in HRZE treated patients, and these organisms modulate peripheral cytokine production, including IL-1, and IFNγ (52). Similarly, *Bifidobacterium*, which we find depleted in HRZE treated cases, can induce a Th17 immune response in mice (53). Taken together, these findings suggest that the HRZE-induced perturbation of the microbiome may have significant effects on peripheral immune tone. These potential effects on immunity, coupled with the well-documented variability in treatment response to TB, may suggest that variability in microbiotic perturbation and peripheral immunity could affect the efficacy of TB treatment. The data presented here will now allow for testing of this hypothesis using prospectively collected cohorts of TB cases beginning treatment, with the aim to correlate their microbiomic disruption with microbiologic and immunologic markers of treatment success.

Our findings are also corroborated by a recent study that examined the effect of TB treatment in mice (54). During HRZ treatment in mice, a decrease in species richness is observed, similar to the significant decrease in the number of OTUs during HRZE treatment in humans. In mice, RIF is the major driver of taxonomic alteration in the intestinal microbiome, but interestingly,

combination therapy gives rise to alterations not found for monotherapy of any single antibiotic. Additionally, in both mice and humans there is a significant decrease in the number of Clostridia during treatment, including the genera *Bacteroides*, *Blautia*, *Clostridium*, and *Roseburia*.

The other prominent finding from our study is the long-lasting duration of the microbiotic disruption induced by HRZE. Our cured group had completed treatment on average 1.2 years earlier, yet their microbiomes were still detectably different from age matched control subjects. This finding suggests that the duration of 6 months of HRZE therapy has long lasting effects on the community structure of the microbiome. Furthermore, based on data from Namasivayam et al., in both mice and humans there is a persistent microbiomic dysbiosis after the completion of HRZ(E) treatment (54). Although alpha diversity recovers, taxonomic profiling in mice and humans, and functional pathway profiling in humans, suggests that 6-month administration of HRZ(E) treatment causes persistent changes. Our study raises the possibility that TB therapy permanently changes the ecology of the human microbiome.

As is the case for the acute perturbation by HRZE, the consequences of these durable changes in cured cases will require further study. It is possible that the persistent microbiomic perturbation could predispose to enteric pathogens due to loss of colonization resistance. It is also possible that cured individuals could be more susceptible to systemic infection due to effects of microbiotic alteration and disruption on peripheral immunity. Multiple epidemiological studies have indicated that people cured of TB are at higher risk of a second

case of TB due to reinfection (55, 56). Although multiple environmental and genetic factors likely contribute to this risk, including HIV infection, the findings in this study raise the possibility that the persistent microbiomic disruption that accompanies curative TB treatment could contribute to post-treatment susceptibility to reinfection, perhaps not just with *Mtb*, but also with other diseases associated with an altered immune state.

In summary, we have shown that TB treatment with HRZE in humans perturbs the intestinal microbiome in distinct and long lasting ways. Specific genera of bacteria are depleted during treatment and functional profiling demonstrates altered functional pathway composition. These changes, in terms of both taxonomic and metagenomic function, are protracted for more than one year after the completion of therapy for TB disease, and are possibly permanent.

CHAPTER 4 – METHODS

DNA extraction from stool

Stool specimens were collected and stored for less than 24 hours at 4 °C, aliquoted (~2 mL each), frozen at -80 °C, and shipped to MSKCC. About 500 mg of stool from frozen samples was suspended in 500 µl of extraction buffer (200 mM Tris-HCl, pH=8.0; 200 mM NaCl; 20 mM EDTA), 210 µl of 20% SDS, 500 µl of phenol/chloroform/isoamyl alcohol (25:24:1), and 500 µl of 0.1-mm-diameter zirconia/silica beads (BioSpec Products). Samples were lysed via mechanical disruption with a bead beater (BioSpec Products) for two minutes, followed by two extractions with phenol/chloroform/isoamyl alcohol (25:24:1). DNA was precipitated with ethanol and sodium acetate at -80 °C for 1 hour, re-suspended in 200 µl of nuclease-free water, and further purified with the QIAamp DNA Mini Kit (Qiagen) according to the manufacturer's protocols, including Protein removal by Proteinase K treatment. DNA was eluted in 200 µl of nuclease-free water and sorted at -20 °C.

16S rDNA sequencing

Primers used to amplify rDNA were: 563F (59-nnnnnnnn-NNNNNNNNNNNN-AYTGGGYDTAAAGN G-39) and 926R (59-nnnnnnnn-NNNNNNNNNNNN-CCGTCAATTYHTTTR AGT-39). Each reaction contained 50 ng of purified DNA, 0.2 mM dNTPs, 1.5 µM $MgCl_2$, 1.25 U Platinum TaqDNA polymerase, 2.5 µl of 10 × PCR buffer and 0.2 µM of each primer. A unique 12-base Golay barcode (Ns) preceded the primers for sample identification after pooling amplicons. One to eight additional nucleotides were added before the barcode to offset the sequencing of the primers. Cycling conditions were the following: 94 °C for 3 min, followed by 27 cycles of 94 °C for 50 s, 51 °C for 30 s and

72 °C for 1 min, where the final elongation step was performed at 72 °C for 5 min. Replicate PCRs were combined and were subsequently purified using the Qiaquick PCR Purification Kit (Qiagen) and Qiagen MinElute PCR Purification Kit. PCR products were quantified and pooled at equimolar amounts before Illumina barcodes and adaptors were ligated on using the Illumina TruSeq Sample Preparation procedure. The completed library was sequenced on an Ilumina Miseq platform per the Illumina recommended protocol.

Bioinformatics Analysis

For 16S MiSeq sequencing, paired-end reads were joined, demultiplexed, filtered for quality using maximum expected error (Emax=1), and dereplicated. Sequences were grouped into operational taxonomic units (OTUs) of 97% distance-based similarity using UPARSE (57). Potentially chimeric sequences were removed using both de novo and reference-based methods (where the Gold database was used for the latter) (58). Taxonomic assignments were made using BLASTN (59) against the NCBI refseq_rna database with custom scripts (32). Our approach allows for the identification of the top 30 taxa associated with a particular OTU, thus the taxonomic nomenclature that we use for 16S is versatile (data not shown). A biological observation matrix (biom) (60) file, a taxonomy file, reference sequence file, and tree file were constructed using QIIME commands. These files were imported into R (61) and merged with a metadata file into a single Phyloseq object (62). Phyloseq was used for all downstream analysis of 16S taxonomic data, and plots were made with the ggplot2 package (63).

Shotgun Metagenomic Sequencing

Between 150 and 200 ng of DNA isolated from stool (vide supra) was sheared acoustically. Hiseq sequencing libraries were prepared using the KAPA Hyper Prep Kit (Roche). PCR amplification of the libraries was carried out for 6 cycles. Samples were run on a Hiseq 4000 in a 125bp/125bp paired end run, using the TruSeq SBS Kit v3 (Illumina). There were an average number of read pairs per sample of around 11 million.

Shotgun metagenomics pair-ends reads for corresponding to stool samples from 150 individuals (64), were downloaded from the Human Microbiome Project Portal at http://www.hmpdacc.org/HMASM/.

For the analysis of shotgun metagenomic reads (i.e. those generated in-house, characterizing the Haitian group as well as those obtained from the HMP portal), sequences were first trimmed and removed of host contamination using Trimmomatic (65) and Bowtie2 (66). Host-decontaminated reads were then profiled for microbial species abundances using Metaphlan2 (67), and for abundance of Uniref gene and KEGG orthologues, and functional pathways (Metacyc pathways, KEGG pathways, and KEGG modules) using the software pipeline HUMAnN2 (40) and in-house written scripts (available upon request). Normalized taxonomic, gene, and pathway abundances were then used for downstream statistical analysis in R. All intestinal microbiome samples were sequenced using 16S rDNA sequencing, however, only a subset of controls were sequenced using metagenomics. Due to sample size limitations, for the metagenomic DNA sequencing comparisons, we combined both uninfected and LTBI individuals

into a healthy control group which was used as the comparator for metagenomic analyses.

Statistical Analysis

The ability to detect differentially abundant OTUs between groups of people is critical for comparison between groups, and various methods exist and have been validated for this sort of analysis. For 16S rDNA sequencing, we employed the tools available within the Phyloseq package to manipulate the data and metadata for downstream analysis. Raw counts with taxonomy and metadata were piped into the DESeq2 package for differential abundance analysis using the negative binomial distribution assumption with zero inflation (68). This method assumes that for many OTUs, the variance in abundance (i.e., read count) between samples or groups exceeds the mean read count (often zero). When this is true, the DESeq method can be used to transform the data so that between sample or between group differences may be compared more accurately. Homoscedastic abundance data was used to generate heatmaps in Figures 3C and 3D, by applying a variance stabilizing transformation from fitted dispersion-means to transform the count data. We additionally employed the microbiome-friendly linear discriminant analysis, effect size (LEfSe) tool (39) to detect statistically significant differences between clinical groups. This technique first employs the non-parametric Kruskal-Wallis (KW) sum-rank test between different groups of people (i.e., HMP, healthy [comprised of *Mtb* uninfected and LTBI], on HRZE treatment, or cured), followed by linear discriminant analysis to estimate the size of the effect (i.e., the degree of significant differential abundance between a particular OTU, taxa, gene, or pathway between groups). We attempted to

employ both the DESeq2 and LEfSe methods, and try to emphasize where there is overlap. All figures in the paper that are related to 16S sequence analysis are plotted using the normalized and transformed abundances from the DESeq2 package. For the statistical analysis of the results from shotgun metagenomics reads, data were imported into R and converted to Phyloseq objects with custom scripts. Custom code implementing non-parametric tests (Wilcoxon-signed rank) with FDR correction (Benjamini & Hochberg method) as well as LEfSe (39) were used to test for differential abundances for taxa, and functional pathways. For the Haitian-HMP LEfSe comparison significance for the Kruskall-Wallis test (no subclasses) was assessed at the p-value threshold of 0.01. For the LTBI-Treatment and LTBI-Cured comparisons p-value threshold was kept at 0.05 for both the initial Kruskall-Wallis test and the subsequent sex-matched subclasses Wilcoxon-signed rank tests. Shannon diversity for the Haitian-HMP comparison was calculated using the Vegan package in R based on species number and abundance (35). We additionally employed the Permanova and Betadisper tests using the adonis function in the Vegan package in R. Adonis partitions a distance matrix of OTU count data and runs an analysis of variance between groups of samples. Betadisper further supports this conclusion by determining if the variance between the two groups is similarly distributed. All box-and-whisker plots were generated with the ggplot2 (63) function geom_boxplot, which shows the first and third quartiles of the dataset and the median of the data in the box, the whiskers show 1.5 times the value of the interquartile range of the box hinge, and outliers are shown as dots. All other plots were made using Prism 7.

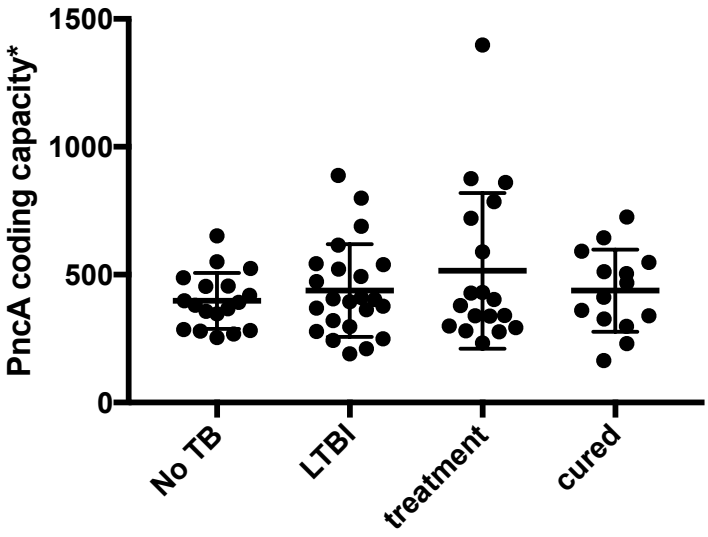
PncA Coding Capacity Analysis

To investigate mechanistically how HRZE therapy may be altering the microbiome, we reasoned that coding capacity for the gene that is known to activate pyrazinamide, PncA, may indicate the degree of perturbation one's microbiome would experience upon HRZE exposure. To search for PncA orthologs, we used the Ortholuge Database (http://www.pathogenomics.sfu.ca/ortholugedb/), an online resource that is able to provide a list of orthologous genes for a given input sequence for both Bacteria and Archaea (18). The advantage of using a database like this is that it utilizes reciprocal BLAST, which ensures that all of the genes in the output are true orthologs of the input gene. Reciprocal BLAST first uses standard BLASTp to obtain protein homologs, and then takes the top hit of the BLASTp output and re-performs BLAST against the reference genome. Only if the top hit of this second step is the input sequence does the program consider the two sequences orthologous—a standard generally accepted by biologists (19, 20). The Ortholuge Database output is a list of protein-coding gene sequences of all true orthologs of PncA. We then obtained the nucleotide sequences using the NIH command line E-Utilities and E-Direct programs (http://www.ncbi.nlm.nih.gov/books/NBK179288/), designed for high-throughput queries (21). Using a custom BASH script to automatically take the ortholog output from the reciprocal BLAST as input into the esearch and efetch commands in EDirect, we obtained a fasta file of all protein coding true orthologs of PncA.

Next, we used shortBRED to find a unique set of PncA markers to which metagenomic reads could be mapped. This software relies on k-mer based

identification of protein families from the input PncA sequence. We used shortbred_identify.py to cluster the PncA orthologs from all Bacteria and Archaea into families and identify k-mers, then used shortbred_quantiy.py to map metagenomic reads from the Haitian metagenomic samples to the PncA k-mer database.

The output was quantified using Prism 7 and is shown in Figure 22. We were unable to detect any statistically significant difference in the coding capacity of the microbiomes of relevant cohorts in our study. We thus conclude that perturbations induced by pyrazinamide may not be caused directly by the canonical active metabolite, pyrazinoic acid (Figure 1).



*PncA coding capacity = reads per kilobase per million mapped reads (RPKM)

Figure 22: PncA coding capacity for each person calculated from metagenomic data and using the ShortBRED program (69). There is no statistically significant difference between any cohort's ability to code for PncA.

Patient Recruitment and Protection of Human Subjects

Subjects were enrolled through the Tri-Intuitional Tuberculosis Research Unit (TBRU) in conjunction with the GHESKIO Centers, in Port-au-Prince, Haiti, where all participants provided written, informed consent. All TBRU protocols and consent forms for samples collected at GHESKIO were approved by Institutional Review Boards at the GHESKIO and Weill Cornell Medicine (see Study Approval). A dedicated clinical field team at the GHESKIO Centers in Port-au-Prince, Haiti recruited research volunteers as part of the NIH U19-funded Tuberculosis Research Unit (AI111143). Patient *Mtb*-infection status is determined using IFNγ release assay (IGRA) status, and active TB disease is determined using standard clinical assessments. All cases with active pulmonary TB receive periodic follow-up appointments while on treatment, and anyone with known contact with an active TB patient receives a six-month follow-up and is re-screened for IGRA status. All patient samples were de-identified on site using a barcode system, before they were shipped to NYC for analysis. Human DNA was decontaminated from metagenomic shotgun sequencing data before analysis and publication, consistent with the removal of all biometric identifiers according the Health Insurance Portability and Accountability Act (70). All clinical metadata was collected on site and managed through the REDCap data management system (71).

Study Approval

All volunteers provided written informed consent to participate in this study. All protocols and consent forms have been approved by the GHESKIO and Weill Cornell Medicine institutional review boards.

REFERENCES

1. . World Health Organization; 2016.
2. Nathan C. What can immunology contribute to the control of the world's leading cause of death from bacterial infection? *Immunol Rev.* 2015;264(1):2-5.
3. Abel L, El-Baghdadi J, Bousfiha AA, Casanova JL, and Schurr E. Human genetics of tuberculosis: a long and winding road. *Philos Trans R Soc Lond B Biol Sci.* 2014;369(1645):20130428.
4. Hand TW, Dos Santos LM, Bouladoux N, Molloy MJ, Pagan AJ, Pepper M, Maynard CL, Elson CO, 3rd, and Belkaid Y. Acute gastrointestinal infection induces long-lived microbiota-specific T cell responses. *Science.* 2012;337(6101):1553-6.
5. Cerovic V, Bain CC, Mowat AM, and Milling SW. Intestinal macrophages and dendritic cells: what's the difference? *Trends Immunol.* 2014;35(6):270-7.
6. Sonnenberg GF, Monticelli LA, Elloso MM, Fouser LA, and Artis D. CD4(+) lymphoid tissue-inducer cells promote innate immunity in the gut. *Immunity.* 2011;34(1):122-34.
7. Zheng Y, Valdez PA, Danilenko DM, Hu Y, Sa SM, Gong Q, Abbas AR, Modrusan Z, Ghilardi N, de Sauvage FJ, et al. Interleukin-22 mediates early host defense against attaching and effacing bacterial pathogens. *Nat Med.* 2008;14(3):282-9.
8. Zheng Y, Danilenko DM, Valdez P, Kasman I, Eastham-Anderson J, Wu J, and Ouyang W. Interleukin-22, a T(H)17 cytokine, mediates IL-23-induced dermal inflammation and acanthosis. *Nature.* 2007;445(7128):648-51.
9. Song C, Lee JS, Gilfillan S, Robinette ML, Newberry RD, Stappenbeck TS, Mack M, Cella M, and Colonna M. Unique and redundant functions of NKp46+ ILC3s in models of intestinal inflammation. *J Exp Med.* 2015;212(11):1869-82.
10. Scharschmidt TC, Vasquez KS, Truong HA, Gearty SV, Pauli ML, Nosbaum A, Gratz IK, Otto M, Moon JJ, Liese J, et al. A Wave of Regulatory T Cells into Neonatal Skin Mediates Tolerance to Commensal Microbes. *Immunity.* 2015;43(5):1011-21.
11. Satpathy AT, Briseno CG, Lee JS, Ng D, Manieri NA, Kc W, Wu X, Thomas SR, Lee WL, Turkoz M, et al. Notch2-dependent classical dendritic cells orchestrate intestinal immunity to attaching-and-effacing bacterial pathogens. *Nat Immunol.* 2013;14(9):937-48.
12. Nutsch K, Chai JN, Ai TL, Russler-Germain E, Feehley T, Nagler CR, and Hsieh CS. Rapid and Efficient Generation of Regulatory T Cells to Commensal Antigens in the Periphery. *Cell Rep.* 2016;17(1):206-20.
13. Fan D, Coughlin LA, Neubauer MM, Kim J, Kim MS, Zhan X, Simms-Waldrip TR, Xie Y, Hooper LV, and Koh AY. Activation of HIF-1alpha and LL-37 by commensal bacteria inhibits Candida albicans colonization. *Nat Med.* 2015;21(7):808-14.
14. Honda K, and Littman DR. The microbiota in adaptive immune homeostasis and disease. *Nature.* 2016;535(7610):75-84.
15. Segal LN, Clemente JC, Li Y, Ruan C, Cao J, Danckers M, Morris A, Tapyrik S, Wu BG, Diaz P, et al. Anaerobic Bacterial Fermentation Products Increase Tuberculosis Risk in Antiretroviral-Drug-Treated HIV Patients. *Cell Host Microbe.* 2017;21(4):530-7 e4.

16. Everard A, Belzer C, Geurts L, Ouwerkerk JP, Druart C, Bindels LB, Guiot Y, Derrien M, Muccioli GG, Delzenne NM, et al. Cross-talk between Akkermansia muciniphila and intestinal epithelium controls diet-induced obesity. *Proc Natl Acad Sci U S A.* 2013;110(22):9066-71.
17. Shin NR, Lee JC, Lee HY, Kim MS, Whon TW, Lee MS, and Bae JW. An increase in the Akkermansia spp. population induced by metformin treatment improves glucose homeostasis in diet-induced obese mice. *Gut.* 2014;63(5):727-35.
18. Anhe FF, Roy D, Pilon G, Dudonne S, Matamoros S, Varin TV, Garofalo C, Moine Q, Desjardins Y, Levy E, et al. A polyphenol-rich cranberry extract protects from diet-induced obesity, insulin resistance and intestinal inflammation in association with increased Akkermansia spp. population in the gut microbiota of mice. *Gut.* 2015;64(6):872-83.
19. Arrieta MC, Stiemsma LT, Dimitriu PA, Thorson L, Russell S, Yurist-Doutsch S, Kuzeljevic B, Gold MJ, Britton HM, Lefebvre DL, et al. Early infancy microbial and metabolic alterations affect risk of childhood asthma. *Sci Transl Med.* 2015;7(307):307ra152.
20. Moya A, and Ferrer M. Functional Redundancy-Induced Stability of Gut Microbiota Subjected to Disturbance. *Trends Microbiol.* 2016;24(5):402-13.
21. Becattini S, Taur Y, and Pamer EG. Antibiotic-Induced Changes in the Intestinal Microbiota and Disease. *Trends Mol Med.* 2016;22(6):458-78.
22. Brismar B, Edlund C, Malmborg AS, and Nord CE. Ciprofloxacin concentrations and impact of the colon microflora in patients undergoing colorectal surgery. *Antimicrob Agents Chemother.* 1990;34(3):481-3.
23. Dethlefsen L, Huse S, Sogin ML, and Relman DA. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol.* 2008;6(11):e280.
24. Hernandez E, Bargiela R, Diez MS, Friedrichs A, Perez-Cobas AE, Gosalbes MJ, Knecht H, Martinez-Martinez M, Seifert J, von Bergen M, et al. Functional consequences of microbial shifts in the human gastrointestinal tract linked to antibiotic treatment and obesity. *Gut Microbes.* 2013;4(4):306-15.
25. Theriot CM, Koenigsknecht MJ, Carlson PE, Jr., Hatton GE, Nelson AM, Li B, Huffnagle GB, J ZL, and Young VB. Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to Clostridium difficile infection. *Nat Commun.* 2014;5(3114.
26. Taur Y, Xavier JB, Lipuma L, Ubeda C, Goldberg J, Gobourne A, Lee YJ, Dubin KA, Socci ND, Viale A, et al. Intestinal domination and the risk of bacteremia in patients undergoing allogeneic hematopoietic stem cell transplantation. *Clin Infect Dis.* 2012;55(7):905-14.
27. Dethlefsen L, and Relman DA. 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(4554-61.
28. Levy M, Kolodziejczyk AA, Thaiss CA, and Elinav E. Dysbiosis and the immune system. *Nat Rev Immunol.* 2017;17(4):219-32.
29. Rawat R, Whitty A, and Tonge PJ. The isoniazid-NAD adduct is a slow, tight-binding inhibitor of InhA, the Mycobacterium tuberculosis enoyl reductase: adduct affinity and drug resistance. *Proc Natl Acad Sci U S A.* 2003;100(24):13881-6.

30. Scorpio A, and Zhang Y. Mutations in pncA, a gene encoding pyrazinamidase/nicotinamidase, cause resistance to the antituberculous drug pyrazinamide in tubercle bacillus. *Nat Med.* 1996;2(6):662-7.
31. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP, et al. Human gut microbiome viewed across age and geography. *Nature.* 2012;486(7402):222-7.
32. Geer LY, Marchler-Bauer A, Geer RC, Han L, He J, He S, Liu C, Shi W, and Bryant SH. The NCBI BioSystems database. *Nucleic Acids Res.* 2010;38(Database issue):D492-6.
33. Human Microbiome Project C. A framework for human microbiome research. *Nature.* 2012;486(7402):215-21.
34. Group NHW, Peterson J, Garges S, Giovanni M, McInnes P, Wang L, Schloss JA, Bonazzi V, McEwen JE, Wetterstrand KA, et al. The NIH Human Microbiome Project. *Genome Res.* 2009;19(12):2317-23.
35. Dixon P. VEGAN, a package of R functions for community ecology. *J Veg Sci.* 2003;14(6):927-30.
36. Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, and Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat Methods.* 2012;9(8):811-4.
37. Schnorr SL, Candela M, Rampelli S, Centanni M, Consolandi C, Basaglia G, Turroni S, Biagi E, Peano C, Severgnini M, et al. Gut microbiome of the Hadza hunter-gatherers. *Nat Commun.* 2014;5(3654.
38. Dehingia M, Devi KT, Talukdar NC, Talukdar R, Reddy N, Mande SS, Deka M, and Khan MR. Gut bacterial diversity of the tribes of India and comparison with the worldwide data. *Sci Rep.* 2015;5(18563.
39. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, and Huttenhower C. Metagenomic biomarker discovery and explanation. *Genome Biol.* 2011;12(6):R60.
40. Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, Rodriguez-Mueller B, Zucker J, Thiagarajan M, Henrissat B, et al. Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput Biol.* 2012;8(6):e1002358.
41. Flint HJ, Bayer EA, Rincon MT, Lamed R, and White BA. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nat Rev Microbiol.* 2008;6(2):121-31.
42. Rampelli S, Schnorr SL, Consolandi C, Turroni S, Severgnini M, Peano C, Brigidi P, Crittenden AN, Henry AG, and Candela M. Metagenome Sequencing of the Hadza Hunter-Gatherer Gut Microbiota. *Curr Biol.* 2015;25(13):1682-93.
43. Turroni S, Fiori J, Rampelli S, Schnorr SL, Consolandi C, Barone M, Biagi E, Fanelli F, Mezzullo M, Crittenden AN, et al. Fecal metabolome of the Hadza hunter-gatherers: a host-microbiome integrative view. *Sci Rep.* 2016;6(32826.
44. Goraca A, Huk-Kolega H, Piechota A, Kleniewska P, Ciejka E, and Skibska B. Lipoic acid - biological activity and therapeutic potential. *Pharmacol Rep.* 2011;63(4):849-58.
45. Lacombe A, Li RW, Klimis-Zacas D, Kristo AS, Tadepalli S, Krauss E, Young R, and Wu VC. Lowbush wild blueberries have the potential to modify gut microbiota and xenobiotic metabolism in the rat colon. *PLoS One.* 2013;8(6):e67497.

46. Buffie CG, Bucci V, Stein RR, McKenney PT, Ling L, Gobourne A, No D, Liu H, Kinnebrew M, Viale A, et al. Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile. *Nature.* 2015;517(7533):205-8.
47. Anahtar MN, Byrne EH, Doherty KE, Bowman BA, Yamamoto HS, Soumillon M, Padavattan N, Ismail N, Moodley A, Sabatini ME, et al. Cervicovaginal bacteria are a major modulator of host inflammatory responses in the female genital tract. *Immunity.* 2015;42(5):965-76.
48. Zaborin A, Smith D, Garfield K, Quensen J, Shakhsheer B, Kade M, Tirrell M, Tiedje J, Gilbert JA, Zaborina O, et al. Membership and behavior of ultra-low-diversity pathogen communities present in the gut of humans during prolonged critical illness. *MBio.* 2014;5(5):e01361-14.
49. Jernberg C, Lofmark S, Edlund C, and Jansson JK. Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. *ISME J.* 2007;1(1):56-66.
50. Taur Y, and Pamer EG. Harnessing microbiota to kill a pathogen: Fixing the microbiota to treat Clostridium difficile infections. *Nat Med.* 2014;20(3):246-7.
51. Shen Y, Giardino Torchia ML, Lawson GW, Karp CL, Ashwell JD, and Mazmanian SK. Outer membrane vesicles of a human commensal mediate immune regulation and disease protection. *Cell Host Microbe.* 2012;12(4):509-20.
52. Schirmer M, Smeekens SP, Vlamakis H, Jaeger M, Oosting M, Franzosa EA, Horst RT, Jansen T, Jacobs L, Bonder MJ, et al. Linking the Human Gut Microbiome to Inflammatory Cytokine Production Capacity. *Cell.* 2016;167(7):1897.
53. Tan TG, Sefik E, Geva-Zatorsky N, Kua L, Naskar D, Teng F, Pasman L, Ortiz-Lopez A, Jupp R, Wu HJ, et al. Identifying species of symbiont bacteria from the human gut that, alone, can induce intestinal Th17 cells in mice. *Proc Natl Acad Sci U S A.* 2016;113(50):E8141-E50.
54. Namasivayam S, Maiga M, Yuan W, Thovarai V, Costa DL, Mittereder LR, Wipperman MF, Glickman MS, Dzutsev A, Trinchieri G, et al. Longitudinal profiling reveals a persistent intestinal dysbiosis triggered by conventional anti-tuberculosis therapy. *Microbiome.* 2017;5(1):71.
55. Verver S, Warren RM, Beyers N, Richardson M, van der Spuy GD, Borgdorff MW, Enarson DA, Behr MA, and van Helden PD. Rate of reinfection tuberculosis after successful treatment is higher than rate of new tuberculosis. *Am J Respir Crit Care Med.* 2005;171(12):1430-5.
56. Glynn JR, Murray J, Bester A, Nelson G, Shearer S, and Sonnenberg P. High rates of recurrence in HIV-infected and HIV-uninfected patients with tuberculosis. *J Infect Dis.* 2010;201(5):704-11.
57. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods.* 2013;10(10):996-8.
58. Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G, Ciulla D, Tabbaa D, Highlander SK, Sodergren E, et al. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Res.* 2011;21(3):494-504.
59. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, and Madden TL. BLAST+: architecture and applications. *BMC Bioinformatics.* 2009;10(421.

60. McDonald D, Clemente JC, Kuczynski J, Rideout JR, Stombaugh J, Wendel D, Wilke A, Huse S, Hufnagle J, Meyer F, et al. The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome. *Gigascience.* 2012;1(1):7.
61. . *R Core Team*. Vienna, Austria: R Foundation for Statistical Computing; 2016.
62. McMurdie PJ, and Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One.* 2013;8(4):e61217.
63. Wickham H. *ggplot2: Elegant Graphics for Data Analysis.* Springer Publishing Company, Incorporated; 2009.
64. Human Microbiome Project C. Structure, function and diversity of the healthy human microbiome. *Nature.* 2012;486(7402):207-14.
65. Bolger AM, Lohse M, and Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114-20.
66. Langmead B, and Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357-9.
67. Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, Tett A, Huttenhower C, and Segata N. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat Methods.* 2015;12(10):902-3.
68. Love MI, Huber W, and Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.
69. Kaminski J, Gibson MK, Franzosa EA, Segata N, Dantas G, and Huttenhower C. High-Specificity Targeted Functional Profiling in Microbial Communities with ShortBRED. *PLoS Comput Biol.* 2015;11(12):e1004557.
70. United S. Health Insurance Portability and Accountability Act of 1996. Public Law 104-191. *US Statut Large.* 1996;110(1936-2103.
71. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, and Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42(2):377-81.