Title
Investigating the Role of the Lung and Gut Microbiota in HIV-associated Bacterial Pneumonia

Author
Shenoy, Meera K.

Publication Date
2017

Peer reviewed|Thesis/dissertation

Investigating the Role of the Lung and Gut Microbiota in HIV-associated Bacterial Pneumonia

by

Meera K Shenoy

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biomedical Sciences

in the

GRADUATE DIVISION

[redacted]

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Susan Lynch for her unending support and amazing mentorship during my graduate career at UCSF. Sue's passion for scientific inquiry, willingness to think outside of the box, investment in mentoring, and infectious enthusiasm for microbiome research originally attracted me to her lab, but I did not realize until much later just how fortunate my choice of mentor truly was. She gave me the freedom to make my mistakes and the support to pick myself back up afterwards, allowing me to grow as a researcher, as a writer, as a presenter, and as a person. Whether I wanted to compete in a science communication competition or develop an experimental system foreign to the lab's expertise, Sue was always game. Sue is also the perfect role model: from her meticulous planning to her insightful questions and her perfect poise to her stellar heels (especially the heels), Sue always leads by example, and I can think of no better mentor for my graduate career.

One of Sue's greatest assets as a mentor has been her ability to recruit and maintain a diverse, talented, and lively research group. I would like to thank every member of the Lynch lab, past and present, for their willingness to learn, teach, laugh, and fire nerf guns with me during the joys and frustrations of my graduate career. Whether we were debating how best to clean up sequencing data (for two months straight without ever coming up with a "right answer") or lamenting the newest politic on-dit over coffee, I can think of no group of people, scientists or otherwise, who I would rather have spent the last four years with. Thank you for your wisdom, your expertise, your ridiculously bad jokes, and most importantly, your snarky commentary on every challenge that came our way. We definitely made the microbiome great again. While I

am incredibly grateful for every member of the Lynch lab, I need to give a special thank you to Professor (!) Emily Cope and my “work wife” Ariane Panzer. Emily’s patient explanations, constant encouragement, and indefatigable cheerfulness inspired me to keep trying no matter what; I knew long before she was offered her own lab that she was an amazing mentor and friend. Ariane and I bonded over books, tea, and most importantly, mustard (the color, not the food), and her enthusiasm was a large part of why I joined the lab. Ariane, research will not be the same without you; thank you for all the laughter you brought to lab life.

As much as I have loved my graduate research experience, I could never have survived my academic pursuits without the insights and distractions provided by my friends outside of lab. Thank you to each and every one of you. To my classmates, I have loved every pub trivia victory (and defeat…), escape room (loss), meat-alicious barbeque, Frisbee game, and board game shouting match. To the Singaporean crew, thank you for adopting me and forcing me to eat way too much food. To the Friday night dinner crew, you gave Jerusalem a whole new meaning. To my pre-graduate school friends, I truly appreciate your (genuine and otherwise) interest in my research and making me ditch academic life once in a while.

Of everyone I want to thank, my family is the hardest because I cannot possibly encompass everything they have done for me and mean to me. Thank you for being relentlessly, impossibly, and annoyingly optimistic about my graduate school experience. You always saw the silver lining and reminded me to spend time outside of lab and have fun. My family has always given me the support, motivation, and inspiration to do anything I have wanted, and I would never have dreamed of going to

graduate school without them. Thank you for believing in me even when I did not. To my brother, Ananth, you are the best older brother anyone could ask for. Thank you for the books (which I *borrowed*), your ridiculous banter which always cheers me up, and for inspiring me to try harder and be better. To my mother, Uma, thank you for your unconditional support, sensible advice in any situation, constant little kindnesses, and ability to listen to everything and anything with unending patience. You are the strongest, most amazing person I have ever met; thank you for everything. Finally, while he did not live to see it, thank you to my father for inspiring me to go to graduate school.

Firstly (because he could never be last), thank you to Yixuan for making each and every day of the last four years the best they could possibly be. Yixuan has been my constant source of joy, my motivation to be the best I can be (inside and outside of lab), and my voice of reason when all else fails. Yixuan is constantly patient, endlessly kind, shockingly brilliant, and patently ridiculous, and I could not have done this without him. The random facts, cooking experiments, weekend lab company (even when it meant clearing out a dead -80C freezer), endless dinners, Costco runs, symphony outings, surprise musicals, research chats, definitely not research chats, brunches, and so much more were the best part of my graduate school experience. Thank you for being my best friend, my confidant, my strength, and my sunshine.

CONTRIBUTIONS TO THE PRESENTED WORK

The work described in this dissertation was performed under the direct supervision and guidance of Dr. Susan V. Lynch, Ph.D. Additional guidance and insight were provided by thesis committee members Dr. Shomyseh Sanjabi, Ph.D., Dr. Jeffrey Cox, Ph.D., and Dr. Joseph M. McCune, M.D., Ph.D.

Contents in Chapter 1 are modified from the following *in press* publication:
M.K. Shenoy and S.V. Lynch. Role of the Lung Microbiome in HIV Pathogenesis. *Curr Opin HIV AIDS* (2018). M.K.S. and S.V.L. wrote the manuscript.

Contents in Chapter 2 are modified from the following publication:
M. K. Shenoy, S. Iwai, D. L. Lin, W. Worodria, I. Ayakaka, P. Byanyima, S. Kaswabuli, S. Fong, S. Stone, E. Chang, J. L. Davis, A. A. Faruqi, M. R. Segal, L. Huang, and S. V. Lynch. Immune Response and Mortality Risk Relate to Distinct Lung Microbiomes in Patients with HIV and Pneumonia. *Am J Respir Crit Care Med* 195, 104-114 (2017). L.H. and S.V.L. designed the study. W.W., I.A., P.B., S. K., S.F., S.S., E.C., J.L.D. and L.H., trained personnel in Uganda, enrolled patients, collected and shipped BAL and serum samples, and obtained demographic and clinical data. S.I. performed DNA extraction, amplification, and 16S rRNA sequencing. D.L.L. profiled host immune responses and prepared samples for metabolomics analysis. M.K.S processed microbiota data and analyzed clinical, microbiota, immune response, and metabolomic

data. A.A.F and M.R.S. provided statistical guidance. M.K.S. and S.V.L. wrote the manuscript. All authors reviewed and approved the final manuscript.

Contents in Chapter 3 are modified from the following manuscript in preparation: **M.K. Shenoy**, D.W. Fadrosh, D.L. Lin, W. Worodria, I. Ayakaka, P. Byanyima, S. Kaswabuli, M. McGing, M. Sommers, E. Chang, K. McCauley, L. Huang, and S.V. Lynch. Gut Microbiota is Related to Peripheral CD4 Counts, Lung Microbiota, and *In Vitro* Macrophage Dysfunction in HIV-Pneumonia Patients. L.H. and S.V.L. designed the study. W.W., I.A., P.B., S.K., E.C., M.M., M.S., and L.H. trained personnel in Uganda, enrolled patients, collected and shipped bronchoalveolar lavage and stool samples, and obtained demographic and clinical data. M.K.S. and D.W.F. performed microbiota processing. M.K.S. performed immune assays, microbiota profiling, and analyses. D.L.L. and M.K.S. designed immune assays. D.W.F. provided microbiota profiling guidance, and K.M. provided statistical guidance. M.K.S., D.W.F., and S.V.L. wrote the manuscript.

Investigating the Role of the Lung and Gut Microbiota in HIV-associated Bacterial Pneumonia

Meera K. Shenoy

ABSTRACT

In the era of highly active antiretroviral therapy (HAART), human immunodeficiency virus (HIV) infected patients are living longer, healthier lives, yet pulmonary infections still pose a common and frequently fatal co-morbidity in this population. Despite the frequency and severity of this disease, particularly within HIV and TB co-endemic regions of Africa, little is known about factors that determine patient endotypes, disease severity, and outcomes. Advanced HIV infection is known to compromise mucosal barrier defense, lead to macrophage dysfunction, and shift lung and gut microbiota composition. Hence we hypothesized that distinct pathogenic microbiota exist on these mucosal surfaces and relate to patient immune and disease endotypes in HIV infected patients with bacterial pneumonia. We initially profiled 16S rRNA in bronchoalveolar lavage (BAL) from a large cohort (n=182) of Ugandan HIV-pneumonia patients and identified three lower airway community states that repeated across the population and differentially related to microbiological, immunological, and clinical factors. Patients with the lowest mortality rate possessed *Gammaproteobacteria*-dominated communities, while patients with the highest mortality rate were colonized by *Prevotellaceae*-dominated communities, expressed more T-helper 2 (Th2) and Th17 cytokines within their lower airways, and were more frequently administered the

antibiotic ceftriaxone. Building on these observations, we examined bacterial and fungal microbiota composition in paired BAL and stool samples from an independent cohort (n=120) and validated these airway microbiota states. Additionally, we demonstrated that gut microbiota composition is related to circulating CD4 count, airway microbiota composition, and patient outcomes. Gut microbiota from patients with the lowest CD4 counts were enriched for microbes shared with patient airways (based on identical 16S rRNA sequences) and were depleted of traditional gut-associated microbes. Compared to patients with high CD4 counts, sterile microbial products from patients with low CD4 counts induced fewer activated and $CD206^{+}IL\text{-}10^{+}$ tissue repair macrophages and more $IL\text{-}1b^{+}$ pro-inflammatory macrophages *in vitro*. Overall, we demonstrate that there are a small number of distinct lower airway and gut microbiota within HIV infected patients with bacterial pneumonia that are differentially related to immune and patient outcomes. Stratifying patients based on lung and gut microbial communities may allow for more effective endotyping, leading to novel, tailored patient treatments.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

LIST OF TABLES

Chapter 1

Chapter 2

Chapter 3

CHAPTER 1

Introduction

Content in this chapter was modified from the following publication in press:

M.K. Shenoy and S.V. Lynch. Role of the Lung Microbiome in HIV Pathogenesis. *Curr Opin HIV AIDS* (2018).

The human body is extensively colonized by mixed-species communities of microbes (bacteria, fungi, and viruses), which play an important role in metabolism, immune response, and host health (*1*). These microbes primarily exist in close contact with each other both as biofilms that are adherent to host mucosal and dermal surfaces, and as free-swimming or planktonic cells (*2, 3*). Studies of the human microbiome, defined as the aggregate of mixed species microbial genomes and their interactions within a given body niche (for terminology see Table 1), have increased greatly over the last decade. This is primarily due to the advent of culture-independent approaches such as sequence-based 16S rRNA biomarker and shot-gun metagenomic approaches, as well as advances in mass spectrometry permitting detection and quantification of metabolites and proteins associated with human microbiomes. Unlike higher burden microbial sites, such as the gut, the lung was considered sterile prior to culture-independent methods. Due to the large surface area of the lungs, low microbial burden in healthy subjects, and difficulty in obtaining samples without sampling through the upper airways, study of the lung microbiome is challenging. Nonetheless, recent studies controlling for oral contamination have demonstrated that microbes can be detected in the lungs of both healthy subjects and patients with airway disease (*4, 5*).

While HIV infected individuals are living longer, healthier lives in the era of HAART, pulmonary disease still presents a common co-morbidity in this population (*6, 7*). Due to the increased risk of pulmonary disease and loss of mucosal barrier immunity in the HIV infected population, there is much interest in the role of the lung microbiome in HIV pathogenesis. Recent investigations have advanced the field by moving beyond descriptive microbiota (the profile of bacteria present in a given sample) studies towards

an integrated understanding of how distinct pathogenic lung microbiota and their associated metabolic products promote host immune dysfunction within HIV infected patient populations. This introduction summarizes these findings as well as discusses the potential role of the gut microbiome in HIV-associated pulmonary disease.

Table 1. Terminology used in lung microbiome research.

Table 1. Terminology used in lung microbiome research.

Term	Definition
16S rRNA gene	This gene is exclusive to, and ubiquitous within, the bacterial kingdom. It encodes multiple regions of sequence hyper-variability (V1-7) that permit identification of and phylogenetic classification of bacteria. Hence sequencing this gene or its hyper-variable regions is commonly used to assay bacterial microbiota composition.
Bronchoalveolar lavage (BAL)	Obtained via bronchoscopy, BAL fluid represents a wash of a small portion of the lower airways that allows for collection of fluid, metabolites, and microbes. It is one of the most common methods of sampling the lower airways.
BAL, acellular	Acellular BAL is generated by centrifuging BAL fluid at moderate speed and using the resulting supernatant but not the cell pellet for down-stream analyses.
BAL, whole	Whole BAL is generated by spinning BAL at high speed and using the cell pellet but not the supernatant for down-stream analyses.
Biofilm	Community of surface-attached microbes protected by an extracellular matrix. Biofilms are frequently comprised of multiple species and reside on both human mucosal and dermal surfaces.
Biomarker sequencing	A method of assessing microbial composition within a mixed-species community by sequencing a single conserved gene with high sequence variation, e.g., the16S rRNA gene for bacterial community analysis.
Burden	Total number of bacteria (or fungi, etc.) measured within a sample. Also referred to as biomass.
Dispersion	Measure of how variable microbial community composition is across subjects within a population, i.e. if a population shows high dispersion, then their microbiomes are very different from one another.
Dysbiosis	An imbalance within a microbial community i.e. differing from a healthy state.
Metabolome	The complete set of small-molecule chemicals found within a given site.
Microaspiration	Aspiration of nasopharyngeal or gastric fluid into the lower airways, allowing for microbes from other anatomic sites to take up residence in the lower airways.
Microbiome	The collection of microbial genomes, functions, and interactions within a given ecosystem.
Microbiota	A catalog of the collection of microbes within an ecosystem.
Mycobiome	The collection of fungal genomes found within an ecosystem.
Next-generation sequencing (NGS)	Sequencing approaches that permit faster, more economical, and high-throughput generation of nucleic acid reads.
Phylogenetic diversity	A composite metric taking into account the distribution, phylogenetic variability, and number of species within an ecosystem.
Quality filtering	The process of rigorously cleaning and filtering sequence data sets to only include high quality sequences.
Richness	The total number of species within an ecosystem.
Microbial selective pressures	Factors that shape the composition or activity of microbes.
Shotgun metagenomics	Sequencing and reassembly of multiple distinct genomes present within an ecosystem.
Systems biology	An interdisciplinary field focused on understanding human health and disease by reconstructing biological networks across cells, tissues, or whole body.
Virome	The collection of viral genomes found within an ecosystem.

HIV infected versus uninfected bacterial lung microbiota

Early studies by a multi-center consortium, the *Lung HIV Microbiome Project* (LHMP), sought to identify whether HIV infection resulted in changes to the lung bacterial microbiota. Their first study observed increased presence of *Tropheryma whipplei*, the causative agent of Whipple's disease, in HIV infected airways, which decreased following antiretroviral therapy (ART) (*8*). LHMP's follow-up study in 2015, however, was unable to validate this finding (*9*). Results may have differed due to the degree of patient immunodeficiency, study size, environmental factors such as geography, or microbiological variability across patients.

Animal models present a controlled way to address some of these confounding factors. Morris and colleagues developed a simian-HIV (SHIV)-infected cynomolgus macaque model for longitudinally studying the HIV-associated lung microbiome (*10*). Animal models allow repeated sampling over time, which is difficult in humans, and macaques are outbred, walk upright, and have a similar immune system to humans, making them a more accurate HIV infected lung microbiome model than mice. Using this model, the authors demonstrate relatively stable lower airway (BAL-derived) microbiota over time, with more variation observed between monkeys than between time points from the same monkey. While *Tropheryma* dominated a subset of lower airway samples, this finding was independent of SHIV infection and unique to one of the two macaque colonies studied, suggesting that environmental exposures, rather than SHIV infection, select for *Tropheryma* colonization. Thus *Tropheryma* is not necessarily a hallmark of SHIV infection, though it should be noted that this study was limited by small numbers and lack of a longitudinally-followed SHIV uninfected group.

While most lung microbiome studies examine relatively healthy (mean CD4>500 cells μl^{-1}) populations, a recent study of HIV infected patients with advanced disease (baseline CD4 count <262 cells μl^{-1}), longitudinally sampled bacterial composition of human acellular BAL (see Table 1) for three years following initiation of HAART (*11*). HIV uninfected subjects' lung microbiota significantly differed from that of HIV infected patients before and after HAART, suggesting that severe HIV-associated immune dysfunction leads to long-term changes in the lung microbiome. HIV infected, treatment-naïve individuals were enriched for *Streptococcus* and displayed increased population dispersion (or microbiota composition variability) compared to uninfected subjects. Interestingly, following one year of HAART, patients were enriched for *Prevotella* and *Veillonella* as well as *Streptococcus* and displayed decreased dispersion. *Prevotella*, *Veillonella*, and *Streptococcus* are considered part of a healthy oral microbiota (*12*), though their over-growth is associated with periodontal disease (*13*). Enrichment of these oral commensals within the lungs as well as increased community dispersion characterizes advanced HIV infection. Community dispersion lessens with HAART, implying that HAART promotes compositionally consistent lung microbiota but is unable to completely reverse microbial dysbiosis, at least within three years of treatment. This parallels the inability of HAART to fully restore the gut microbiota (*14*), and suggests HIV infection results in systemic microbial dysbiosis that HAART only partially reverses. Increased dispersion across the HIV infected population offers a possible explanation for the inability of studies with higher CD4 counts (*9, 10*) to identify consistent HIV infection signatures. The authors posit that HIV immunodeficiency permits many types of microbial dysbiosis, as illustrated by the large dispersion in HIV-associated lung

microbiota. Thus the emerging hypothesis is that multiple distinct pathogenic microbial communities exist and likely contribute to chronic inflammation, patient heterogeneity, and pulmonary disease progression across HIV infected populations.

HIV bacterial lung microbiota in pulmonary disease

Despite the efficacy of HAART in reducing viral load and increasing CD4 count, pulmonary disease is still a common co-morbidity within HIV infected patients. Bacterial pneumonia affects 5-30 percent of HIV infected patients compared to less than one percent of immunocompetent individuals, and results in lower pulmonary function and higher long-term mortality rates compared to CD4-matched controls (*15-17*). Prior studies demonstrated that the lungs of HIV infected pneumonia patients are enriched for a phylogenetically diverse array of bacteria, including *Firmicutes* and *Prevotellaceae*, compared to primarily *Proteobacteria*-enriched communities in HIV uninfected pneumonia patients (*18*). Additionally, HIV-pneumonia lower airway communities differed based on patient geographic location, with Ugandan HIV-pneumonia patients having significantly higher richness and diversity compared to San Franciscan patients (*19*). These differences in bacterial composition potentially suggest distinct mechanisms of microbial pathogenesis in HIV infected versus uninfected pneumonia as well as a role for geographic, ethnic, and environmental exposures in shaping lung communities.

In the largest HIV lung microbiota study to date (n=182 patients), distinct microbial community states (MCS) were identified in Ugandan, HIV infected pneumonia patients, the results of which are presented in Chapter 2 (*20*). Briefly, MCS were characterized by two repeating patterns of distinct airway (BAL) microbiota; the first was

dominated by *Pseudomonadaceae* co-associated with *Sphingomonadaceae*, and the second by a reciprocal gradient of *Prevotellaceae* to *Streptococcaceae* dominance with *Veillonellaceae* co-colonization. The first MCS colonized less than a quarter of patients and bore resemblance to HIV uninfected pneumonia lung communities. In contrast, the second colonization pattern encompassed over 100 patients who were frequently administered the antibiotic ceftriaxone, possessed significantly higher phylogenetic diversity and richness, and trended towards higher short- and long-term mortality. This MCS was characterized by the same oral commensals as aforementioned studies (*10, 11*), suggesting that lung colonization by oral commensals may precede and drive clinical presentation of pneumonia as well as explain the higher prevalence and severity of pneumonia in HIV infected patients. This study is the first to demonstrate that distinct microbial communities arise in the lower airways of HIV patients with pneumonia and are linked to patient outcomes.

Chronic obstructive pulmonary disease (COPD) has increased in prevalence in the HAART era, and lung microbial dysbiosis may contribute to COPD chronic inflammation (*21-23*). In the SHIV infected macaque model previously discussed, the authors induced COPD by co-housing SHIV infected macaques alongside SIV and *Pneumocystis* co-colonized macaques, a known model for inducing *Pneumocystis*-associated COPD (*24, 25*). Monkeys that eventually developed COPD were enriched for *Veillonella* prior to SHIV infection and displayed increasing abundances of *Prevotella* and *Fusobacterium* over the course of SHIV infection and following development of COPD (*10*). In light of *Prevotella* and *Veillonella* enrichment in advanced HIV infection, this correlation of their abundance with COPD development suggests that microbes

enriched in HIV infection may drive the increased prevalence of COPD in this population. The consistent enrichment of *Prevotella* and *Veillonella* across advanced HIV infection, pneumonia, and COPD argues that these microbes play a central role in HIV airway pathogenesis and may provide a new avenue for microbial species-targeted therapeutics to alleviate chronic pulmonary inflammation.

Lung mycobiome and virome in HIV infection

HIV infected individuals have increased susceptibility to airway infections with fungi such as *Pneumocystis* and viruses such as *Cytomegalovirus* in addition to bacterial infections (*26, 27*). Despite the frequency of these infections, there is little known about the lung mycobiome or virome in HIV infection. A recent study by Cui and colleagues sought to ameliorate this dearth, by studying fungal communities in HIV infected patients with and without COPD (*28*). The authors found that both HIV infection (relative to neither infection nor COPD) and COPD (relative to HIV infection without COPD) are associated with airway (BAL) enrichment of a number of fungi, including *Pneumocystis jiroveci* and *Ceriporia lacerata. Pneumocystis jiroveci* has previously been associated with COPD (*25, 29*) and both fungi are known causes of pneumonia (*16, 30*). Finding these fungi in the SHIV-COPD macaque model, without clinical manifestations of pneumonia, suggests that mycobiome dysbiosis and opportunistic fungal pathogens precede development of and predispose patients towards fungal pneumonia and COPD.

Even less is known about the lung virome than the mycobiome. To date, no peer-reviewed studies have examined the lung virome in HIV infection. A study of lung

transplant patients included three HIV infected, treatment-naïve individuals as a control and found that HIV infected and uninfected individuals shared low but detectable levels of *Anellovirus* and bacteriophages, whereas high levels were found in transplant patients (*31*). Additionally, a recent review included preliminary data examining ten healthy HIV infected individuals receiving long term ART and detected *Parvoviridae*, *Herpesviridae* (family of *Epstein-Barr virus* and *Cytomegalovirus*), *Flaviviridae,* and bacteriophages in BAL, though there was no uninfected comparison (*32*). Overall, the lung virome and mycobiome are areas that require far more study, particularly in the context of host immunity and the overall microbiome, before it will be possible to understand their contribution to HIV pathogenesis.

Metabolic and immune signatures associated with HIV lung microbiota

One of the main challenges in studying the lung microbiome is proving it plays a functional role in shaping human health and disease. More recently, studies of the gut microbiome have revealed that microbial-derived metabolites, e.g., short-chain fatty acids, play a key role in immune modulation (*33*). Cribbs and colleagues recently examined BAL bacterial community structure in relation to lower airway metabolic profiles (*34*). Using liquid chromatography-high-resolution mass spectrometry, the authors demonstrate that HIV infection and CD4 count are related to specific lower airway metabolic features, and a subset of these, including lineolate, glycerophospholipid, and fatty acid metabolism, are linked to members of *Caulobacteraceae, Staphylococcaceae, Nocardioidaceae,* and *Streptococcus*. It remains unclear whether these metabolites are indeed microbial derived or whether HIV

infection re-directs host metabolism, which in turn selects for these bacteria. However, this was the first study to indicate that the lower airway metabolome and microbiome of HIV infected patients are co-associated, and suggests a plausible mechanism for microbial pathogenesis in HIV infected patients.

A similar approach was applied to the aforementioned Ugandan HIV infected pneumonia patients, which will be discussed in Chapter 2. In this study, the predicted gene products of distinct airway MCS were enriched in paired serum, and significantly associated with lung immune response and mortality outcomes (*20*). Patients colonized by *Prevotellaceae*-dominated communities were enriched for products of branched chain amino acid (BCAA) metabolism within their serum, and possessed high levels of lower airway inflammation, including the T helper cell 17 (Th17) cytokine interleukin 17 (IL-17A). Products of the BCAA pathway can be bacterial derived (*35*) and positively regulate mechanistic target of rapamycin (mTOR) signaling (*36*), a key component of Th17 differentiation (*37*), suggesting that *Prevotellaceae*-dominated communities may induce Th17 inflammation through BCAA metabolism. Association between lung Th17 inflammation and *Prevotella* and its co-colonizer, *Veillonella,* is corroborated by multiple studies: the presence and abundance of these oral commensals within the lungs is associated with increased Th17 inflammation in healthy individuals (*38*) and can induce Th17 inflammation *in vitro* and in mouse models (*39*). Studying microbes in the context of the metabolome and immune response marks a profound transition in how we view the presence of microbes in the lungs. Decades of previous literature have focused on microbes within the lungs solely as isolated pathogens; however, newer data implicate

them as cooperative, metabolically active communities whose metabolic products influence immune function.

Segal and colleagues applied a similar systems biology approach to examining the microbial, metabolic, and immune correlates of tuberculosis (TB) risk in a cohort of HIV infected individuals in Cape Town, South Africa receiving ART (*40*). Individuals who proceeded to active pulmonary TB within 3 years were enriched for the microbial-derived short chain fatty acids (SCFA) propionate and butyrate and depleted for interferon gamma (IFNγ) and IL-17A within serum, relative to those who did not develop TB. *In vitro* addition of butyrate to peripheral blood mononuclear cells stimulated with *Mycobacterium tuberculosis* (MTB) antigen decreased IFNγ and IL-17A production, establishing an immune regulatory role for this microbial-derived metabolite. These patients were also colonized by *Prevotella* and *Veillonella*, emphasizing a role for these anaerobic oral commensals in HIV-associated pulmonary disease. While several studies have been unable to tie lung bacterial composition to circulating CD4 counts (*11, 20, 40*), lung CD4 counts were found to negatively correlate with *Prevotella* or *Veillonella* abundance. These recent studies provide evidence of lung microbial metabolites manipulating immune responses and disease outcomes, as well as emphasize the potential pathogenicity of oral commensal expansion in the lungs of HIV infected patients.

HIV infected bacterial gut microbiota and their potential role in HIV-associated pulmonary disease

A number of studies have demonstrated that HIV infected patients possess dysbiotic gut microbiota that are significantly and consistently distinct from the gut microbiota of HIV uninfected individuals. Initial analyses have disagreed in whether HIV infected microbiota displayed decreased richness and diversity (*41, 42*) relative to uninfected individuals or whether there is no difference or increased diversity (*43, 44*). However, taxonomic analyses of both stool samples and mucosal biopsies from across the United States have consistently revealed increased *Prevotella* and *Proteobacteria* abundance and decreased *Bacteroides* abundance as hallmarks of the HIV-associated gut microbiome (*14, 42-44*). In Western nations with high protein diets, such as the United States, the healthy gut microbiota is consistently characterized by a high ratio of *Bacteroides* to *Prevotella* abundance (*45*), which is shifted towards a high *Prevotella* to *Bacteroides* ratio in HIV infection. This gut bacterial dysbiosis was demonstrated by Vujkovic-Cvijin and colleagues to only partially and inconsistently be reversed by HAART treatment (*14*). Additionally, the authors found that an HIV infected long term non-progressor had gut microbiota similar to that of HIV uninfected individuals rather than HIV infected patients, indicating that gut microbiota dysbiosis is a hallmark of progressive HIV infection and cannot be fully reversed with ART, similar to airway microbiota dysbiosis.

The aforementioned studies examined United States populations; however, a recent study examined the gut microbiota of rural Ugandan, HIV infected and uninfected individuals (*46*). The authors found that HIV infected patients in this rural, primarily vegetarian population possessed decreased richness and diversity, but were unable to detect a consistent HIV-associated gut microbiota that differed from HIV uninfected

individuals. Additionally, while the authors were unable to demonstrate a relationship between CD4 count and gut microbiota composition, they did detect shifts in specific taxa. Finally, soluble (s)CD14, a monocyte-derived cytokine and marker of microbial translocation, was increased in the circulation of HIV infected individuals with CD4 counts below 200 cells/μl and was moderately correlated (r=0.479) with viral load, indicating that the patients with the most severe HIV-associated disease possessed increased microbial translocation into the circulation. While this was the first investigation into the gut microbiota of HIV infected African populations, challenges with sequencing run effects and cohort size limited the conclusions of this study.

Investigations thus far have established that dysbiotic gut microbiota colonize HIV infected individuals, worsen with HIV disease progression, and are not successfully reversed with ART. However, these investigations have focused on HIV infected populations as a whole, without delving into HIV-associated diseases. HIV infected individuals are at increased risk of pulmonary diseases such as COPD and bacterial pneumonia even in the HAART era (*6, 21*). A recent investigation into the role of the gut microbiome in a mouse model of bacterial pneumonia demonstrated that systemic microbiome depletion prior to airway infection with *Streptococcus pneumonia* leads to more severe morbidity and mortality, which is reversed with gut microbial reconstitution by fecal microbial transplant (FMT). This finding in conjunction with the increased prevalence of pulmonary infection in HIV infected patients suggests a probable role for the gut microbiome in modulating and understanding pulmonary disease within this population. Since monocyte-derived macrophages are necessary for controlling the rampant inflammation that characterizes bacterial pneumonia patients with low CD4

counts and patients with advanced HIV infection display macrophage dysfunction, gut microbes and their products may prime airway and systemic immune responses through monocyte-derived macrophages. Findings presented in Chapter 3 are the first to link HIV-associated gut microbiota dysbiosis to bacterial pneumonia disease severity, airway microbe translocation, and macrophage dysfunction in HIV infected pneumonia patients, and establish a relationship between the gut microbiome and HIV-associated pulmonary disease.

SUMMARY OF THE FIELD SO FAR

The microbiome in HIV is still a field in its early stages, but recent studies, which expand beyond microbial description towards a multi-layered systems biology approach together with laboratory models, have provided novel insights and potential mechanisms of microbial pathogenesis in this patient population. Expansion of the oral commensals *Prevotella*, *Veillonella*, and *Streptococcus* within HIV infected lungs has been replicated across a number of studies, and implicated in metabolic and immune interactions (see Figure 1). Early work also suggests that fungal dysbiosis may precede and predispose HIV infected individuals to pulmonary diseases such as COPD, though these studies, like those of the lung virome, are nascent. While investigations into the gut microbiome in HIV infection has produced more consistent results, these studies also require further follow-up before we can truly understand how the gut microbiota contributes to and responds to the HIV pathogenesis.

While recent work has vastly expanded our understanding of the lung microbiome in HIV infection, standardizing methodology and directly proving microbial

colonization and functionality still remain challenges for the field. Since the lower airways in particular possess relatively low microbial burden and samples are difficult to collect, it is important to limit contamination and maximize representative nucleic acids recovery. Kit-based DNA extractions can introduce contamination, and can result in diminished nucleic acids recovery from more difficult-to-lyse microbes (*47, 48*). Additionally, the type of sample interrogated, e.g., whole versus acellular BAL, can also influence microbial detection (microbes are commonly adherent to airway mucosal cells). In addition to sample processing considerations, future studies will need to continue incorporating large cohorts, systems biology approaches, and mechanistic *in vitro* and animal models to ascertain how the lung microbiome drives HIV pathogenesis. Nonetheless, recent studies have set a firm foundation for the field. Numerous groups have shown that oral commensal colonization of the lower airways is associated with inflammation and metabolic reprogramming in HIV infected patients, suggesting a role for these microbes in HIV-associated airway pathogenesis. The studies described in this thesis contribute to and build on these findings. These studies collectively demonstrate that the gut-lung microbial axis is key to understanding HIV-associated pulmonary disease and that microbial community types and effects need to be considered when deciding how best to treat HIV infected pneumonia patients.

Figure 1.

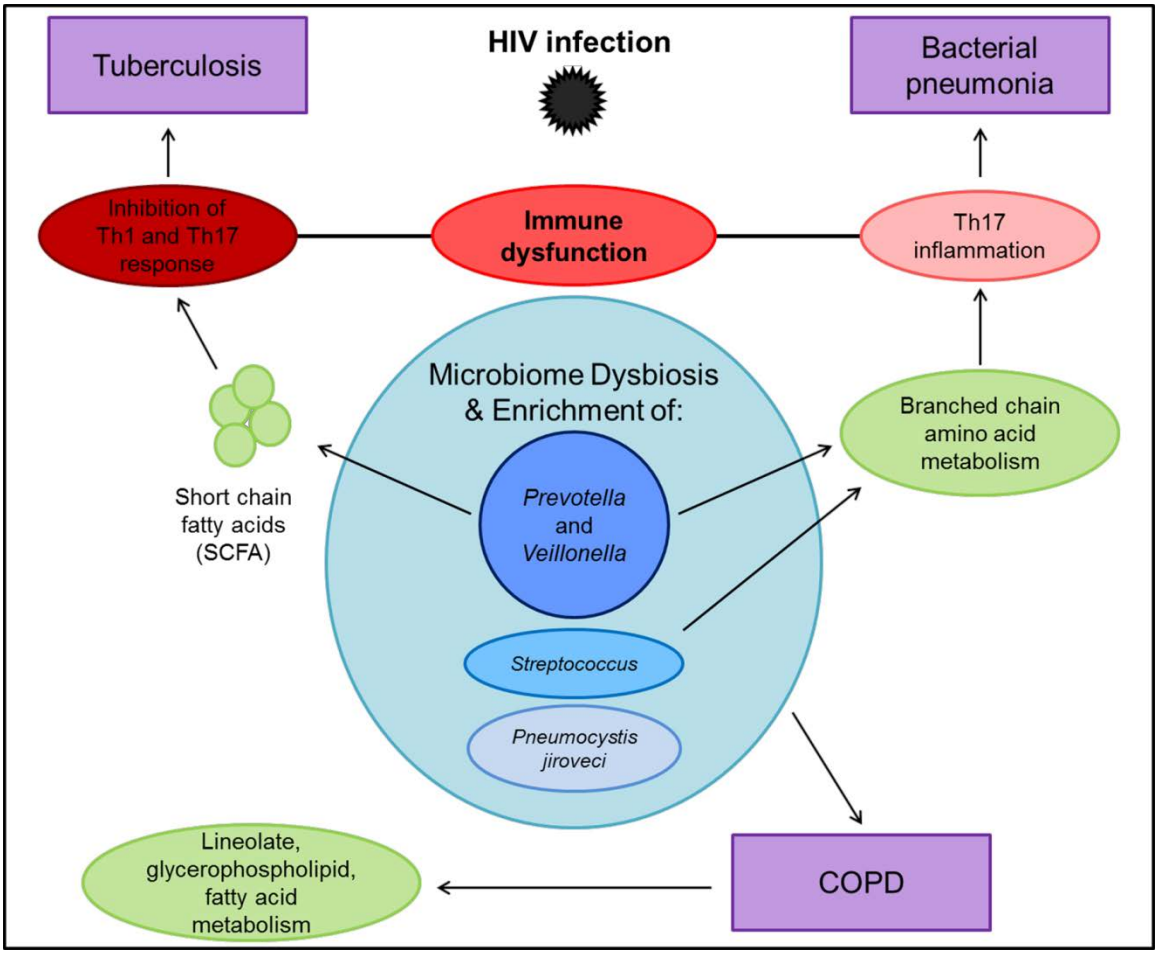


Figure 1. Consistent characteristics of lung microbiota in HIV infection, and their metabolic, immune, and pulmonary disease associations. Changes to the lung microbiota in HIV infection are replicated across multiple studies and are related to metabolic and immune functions. Enrichment of the oral commensal bacteria *Prevotella, Veillonella,* and *Streptococcus* as well as the fungi *Pneumocystis jiroveci* have been identified by numerous studies in the lower airways of HIV infected patients and are associated with pneumonia, COPD, and pulmonary tuberculosis (TB). Chronic immune dysfunction is characteristic of HIV infection and associated pulmonary diseases; Th17 inflammation is specifically related to bacterial pneumonia outcomes; reduced Th1 and

Th17 cytokine expression is related to increased risk of active TB. Microbial enrichments are depicted in blue, metabolism in green, immune response in red, and pulmonary disease in purple. Th1, T-helper cell 1. Th17, T-helper cell 17. COPD, chronic obstructive pulmonary disease.

CHAPTER 2

Immune Response and Mortality Risk Relate to Distinct Lung Microbiomes in HIV-Pneumonia Patients

Content in this chapter was modified from the following publication:

M. K. Shenoy, S. Iwai, D. L. Lin, W. Worodria, I. Ayakaka, P. Byanyima, S. Kaswabuli, S. Fong, S. Stone, E. Chang, J. L. Davis, A. A. Faruqi, M. R. Segal, L. Huang, and S. V. Lynch. Immune Response and Mortality Risk Relate to Distinct Lung Microbiomes in Patients with HIV and Pneumonia. *Am J Respir Crit Care Med* 195, 104-114 (2017).

ABSTRACT

The potential role of the airway microbiota in dictating immune responses and infection outcomes in HIV-associated pneumonia is largely unknown. Ugandan HIV infected pneumonia patient (n=182) BAL samples were obtained to assess lower airway bacterial community composition via amplification and sequencing of the V4 region of the 16S rRNA gene. Host immune response gene expression profiles were generated by QPCR using RNA extracted from BAL fluid. Liquid and gas chromatography mass spectrometry was used to profile serum metabolites. Based on airway microbiome composition, the majority of patients segregated into three distinct groups, each of which were predicted to encode metagenomes capable of producing metabolites characteristically enriched in paired serum samples from these patients. These three groups also exhibited differences in mortality; those with the highest rate had increased ceftriaxone administration and culturable *Aspergillus*, and demonstrated significantly increased induction of airway T-helper 2 responses. The group with the lowest mortality was characterized by increased expression of T cell immunoglobulin and mucin domain 3 (*TIM-3*), which down-regulates T-helper 1 pro-inflammatory responses and is associated with chronic viral infection. These data provide evidence that compositionally and structurally distinct lower airway microbiomes are associated with discrete local host immune responses, peripheral metabolic reprogramming, and different rates of mortality.

INTRODUCTION

Sub-Saharan Africa accounts for 71% of persons estimated to be living with HIV infection worldwide, with 63,000 acquired immunodeficiency syndrome-(AIDS)-related deaths per year in Uganda alone (*49*). Pulmonary infections pose a common and frequently fatal co-morbidity in HIV infected patients in Africa; two of the most prevalent are TB and bacterial pneumonia, incident in approximately 80% of this patient population (*50*). Overall, TB is the leading cause of death in HIV infected patients worldwide (*51, 52*). In HIV-TB co-endemic areas, bacterial pneumonia is a common cause of hospital admission, with mortality rates over 30% even with antiretroviral and antibiotic treatments (*17, 53*).

Even in the absence of acute respiratory infection, HIV infected patients exhibit a broader breadth of lower airway bacterial taxa compared to that detected in healthy subjects (*11*), indicating that HIV infection may present a risk factor for developing pulmonary infection. Despite high morbidity and mortality within this population, little is known about factors that influence heterogeneity in patient outcomes, and, specifically, whether variation in airway microbiota composition and immune response are related to patient survival. We demonstrated, in a non-HIV infected, antimicrobial-treated pneumonia cohort, that following antimicrobial treatment, a precipitous decline in airway microbiome diversity and domination of the community by a distinct respiratory pathogen, e.g., *Streptococcus pneumonia* or *Pseudomonas aeruginosa* is associated with increased 28-day mortality (*54*). In a study of 60 Ugandan HIV infected, antimicrobial-treated pneumonia patients, patients with reduced airway bacterial microbiota richness and diversity exhibited higher bacterial burden and increased

expression of pro-inflammatory *TNFα* and *MMP-9* (*19*), thus providing the first evidence that HIV-airway microbiota composition is related to immune response. These observations led us to hypothesize that in the context of HIV-associated immune dysfunction and antimicrobial administration, acute pneumonia patient subsets can be identified based on their lower airway bacterial composition. We further rationalized that these compositionally distinct airway microbiota function as discrete pathogenic units which induce characteristic airway immune responses and are associated with mortality. To address this hypothesis, we examined clinical and demographic factors related to the bacterial airway microbiome, as well as relationships between community composition, host immune response, and patient outcomes in a large cohort of Ugandan HIV infected pneumonia patients.

RESULTS

Lower airway microbiota composition is associated with demographic, clinical, and microbiological factors

Lower airway bacterial microbiota profiles of 190 Ugandan HIV infected patients with acute pneumonia were generated by 16S rRNA amplicon sequencing of whole BAL fluid (Fig. S1A: read depth). Overall, 182 samples with sufficient sequence reads and adequate bacterial community coverage were used for all microbiota analyses (Fig. S1B). A total of 6,915 operational taxonomical units (OTUs; >97% 16S rRNA V4-sequence similarity; range 124–869, median 335.5 taxa per sample) were identified, indicating robust bacterial presence.

Demographic and clinical data (Table S2) were used in PERMANOVA analysis (*55*). PERMANOVA allows for the identification of factors related to observed variation in bacterial beta-diversity (inter-sample bacterial compositional differences); we measured beta-diversity using a weighted UniFrac dissimilarity matrix, which considers phylogenetic relatedness and species abundance in distance calculations (*56*). Gender (R^2=0.021; p<0.017, Fig. S2A), consumption of alcohol ever (R^2=0.015, p<0.045, Fig. S2B), the presence of culturable *Aspergillus* in BAL (R^2=0.038, p<0.004; Fig. S2C), BAL or sputum culture positivity for *Mycobacterium* (R^2=0.027, p<0.021; Fig. S2D), and ceftriaxone administration within the last two weeks, or at the time of bronchoscopy (R^2=0.016, p<0.040 and R^2=0.061, p<0.001 respectively; Table 1; Fig. 1A) were significantly related to airway bacterial community composition. Seventy-day mortality trended strongly towards a relationship with airway microbiota composition [Canberra (beta-diversity distance based on taxa presence/absence); R^2=0.0061, p<0.053]. Mortality trended strongly towards significance using a Canberra but not a weighted UniFrac dissimilarity matrix, suggesting that presence (or absence) of particular taxa in airway communities is related to mortality, rather than relative abundance or phylogenetic relatedness of community members present.

Since microbes engage in inter-species cell-cell communication that dictates abundance and behavior of other microorganisms in their environment (*57, 58*), we rationalized that inter-species interactions also occur in complex multi-species bacterial microbiota, resulting in deterministic community structures. Indeed, Shannon's diversity index (which considers abundance and richness, PERMANOVA: R^2=0.17, p<0.001; Fig. 1B), Faith's Phylogenetic diversity (phylogenetic variation, R^2=0.09, p<0.001), Chao1

index (species richness estimator, R^2=0.08, p<0.001), and observed species richness (total species, R^2=0.08, p<0.001), were all significantly associated with airway bacterial beta-diversity. These alpha-diversity indices (measurements of variation within samples) explained a greater degree of microbial community variability (8-17%) than clinical or demographic features (reflected in the strength of PCoA groupings in Figs. 1 and S1), suggesting that microbiological influences appear to play a larger role in defining airway taxonomic content than clinical-demographic features.

Table 1. Clinical, demographic, and microbiological features are significantly associated with airway bacterial beta-diversity in HIV infected patients with pneumonia.

Variable Type	Variable Name	Sample (n)	Yes | No	PERMANOVA	
				R^2	p-value
Clinical and demographic	70-Day mortality[a]	182	143 | 39 (Live | Dead)	0.006	0.053
	Alcohol ever consumed	182	113 | 69	0.015	0.045
	Ceftriaxone at bronchoscopy	174	54 | 120	0.061	0.001
	Ceftriaxone within last 2 weeks	178	137 | 41	0.016	0.040
	Culture identified *Aspergillus*	157	15 | 142	0.038	0.004
	Gender	182	110 | 72 (F | M)	0.021	0.017
	TB positive by culture	182	40 | 1 | 141 (Positive | Scanty | Negative)	0.027	0.021
Variable Type	**Variable Name**	**Sample (n)**	**Range [min-max (med)]**	**PERMANOVA**	
				R^2	**p-value**
Microbiological	Chao1	182	170-1326 (484.1)	0.080	0.001
	Faith's Phylogenetic Diversity	182	8.792-45.83 (21.97)	0.091	0.001
	Observed species	182	39-865 (340.5)	0.076	0.001
	Shannon diversity	182	0.642-6.427 (4.011)	0.169	0.001
	Simpson diversity	182	0.112-0.977 (0.867)	0.154	0.001

[a]PERMANOVA value calculated using a Canberra distance matrix.

Figure 1.

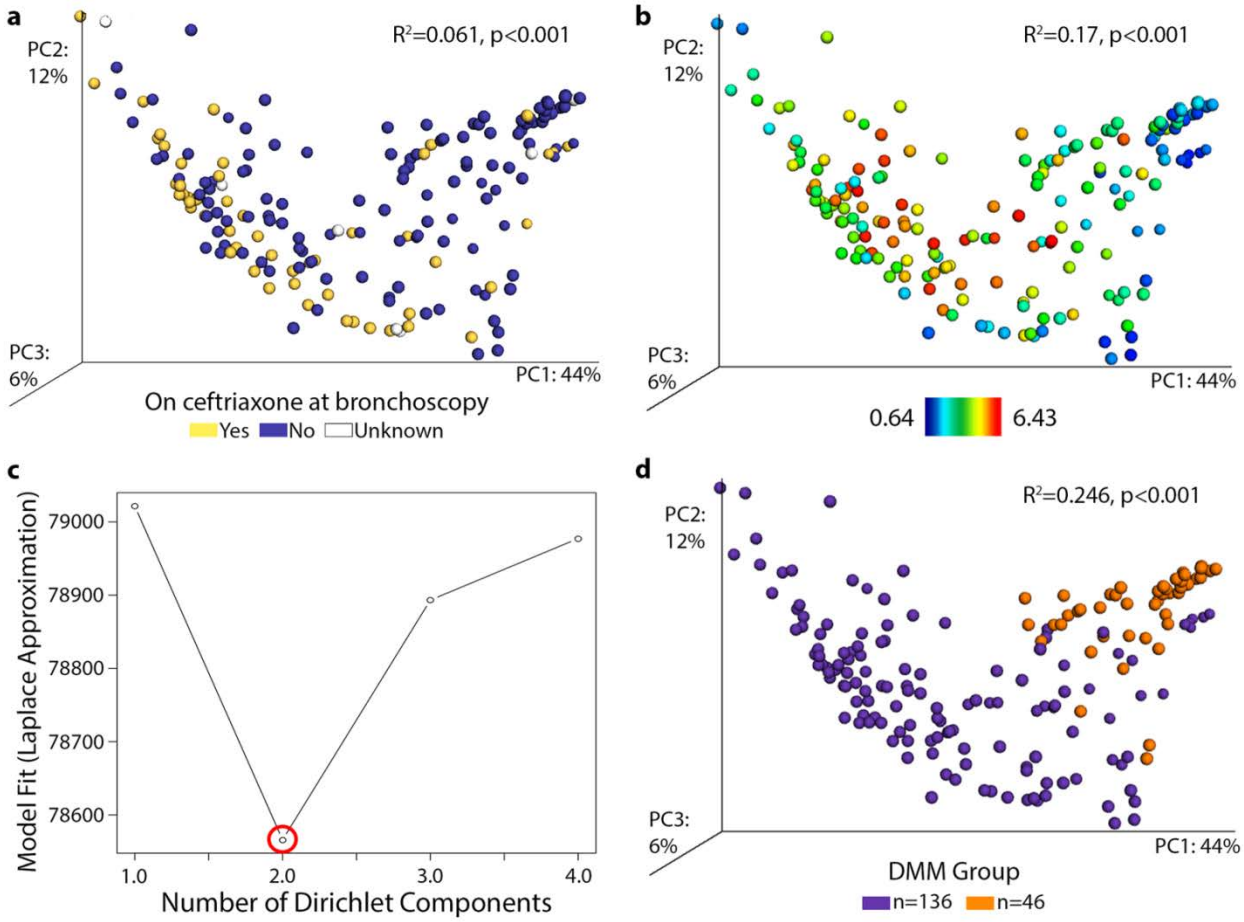


Figure 1. Antibiotic administration, alpha-diversity, and probabilistic modeling differentiate bacterial community types within the lower airways of HIV-pneumonia patients. PCoA of n=182 lower airway BAL bacterial community profiles of Ugandan HIV-pneumonia patients illustrates that **A.** Ceftriaxone (in yellow versus no ceftriaxone in purple), a third generation cephalosporin, administered at time of bronchoscopy is significantly associated with community composition (PERMANOVA, R^2=0.061, p<0.001), as is **B.** Shannon diversity (PERMANOVA, R^2=0.17, p<0.001, scaled from red high to blue low). **C.** Based on Laplace approximation, where a lower value indicates a better model fit, Dirichlet Multinomial Mixtures identified two

compositionally distinct bacterial microbiota (n=136 and n=46) in the lower airways of HIV-pneumonia patients. **D.** PCoA illustrates that DMM-defined lower airway bacterial communities are compositionally distinct (PERMANOVA, R^2=0.246, p<0.001).

HIV-pneumonia patients stratify into two groups based on bacterial community composition

Dirichlet Multinomial Mixtures [DMM (*59*)] modeling examines taxa frequencies and determines how many "meta-communities" or microbial community states (MCS), exist within a dataset. Application of DMM to our cohort identified two significantly distinct MCS (n=46 and n=136, R^2=0.246, p<0.001) using a Laplace Approximation (Fig. 1C and D), which evaluates model fit [lowest Laplace value corresponds to the number of meta-communities that best-fit the model (Fig. 1C)]. Specific co-occurring bacterial families were characteristically enriched in these two groups; MCS1 microbiota were characteristically dominated by *Pseudomonadaceae,* which typically co-occurred with *Sphingomonadaceae* and *Prevotellaceae*. The second, larger group, exhibited a reciprocal gradient of *Streptococcaceae* or *Prevotellaceae* domination, which we designated MCS2A and MCS2B, respectively. *Streptococcaceae*-dominated MCS2A communities co-associated with *Prevotellaceae* and *Veillonellaceae*, and *Prevotellaceae*-dominated MCS2B assemblages with *Veillonellaceae* and *Streptococcaceae* (Fig. 2A). These distinct microbial states exhibited significant differences in diversity, with MCS1 exhibiting the lowest mean diversity compared to MCS2A or MCS2B communities (Faith's Phylogenetic diversity; one-way ANOVA, p<0.001, Fig. 2B).

Using dominant family to classify samples, PCoA-ordination of weighted UniFrac distance matrices confirmed a strong and significant relationship between MCS class and bacterial beta-diversity (PERMANOVA, R^2=0.670, p<0.001, Fig. 2C), corroborating the existence of compositionally distinct microbial states. Removal of the dominant

family reads and re-application of DMM to the remaining data yielded the same two groups (n=46, n=136), indicating that dominant family is not the sole defining feature of these airway microbiota.

Figure 2.

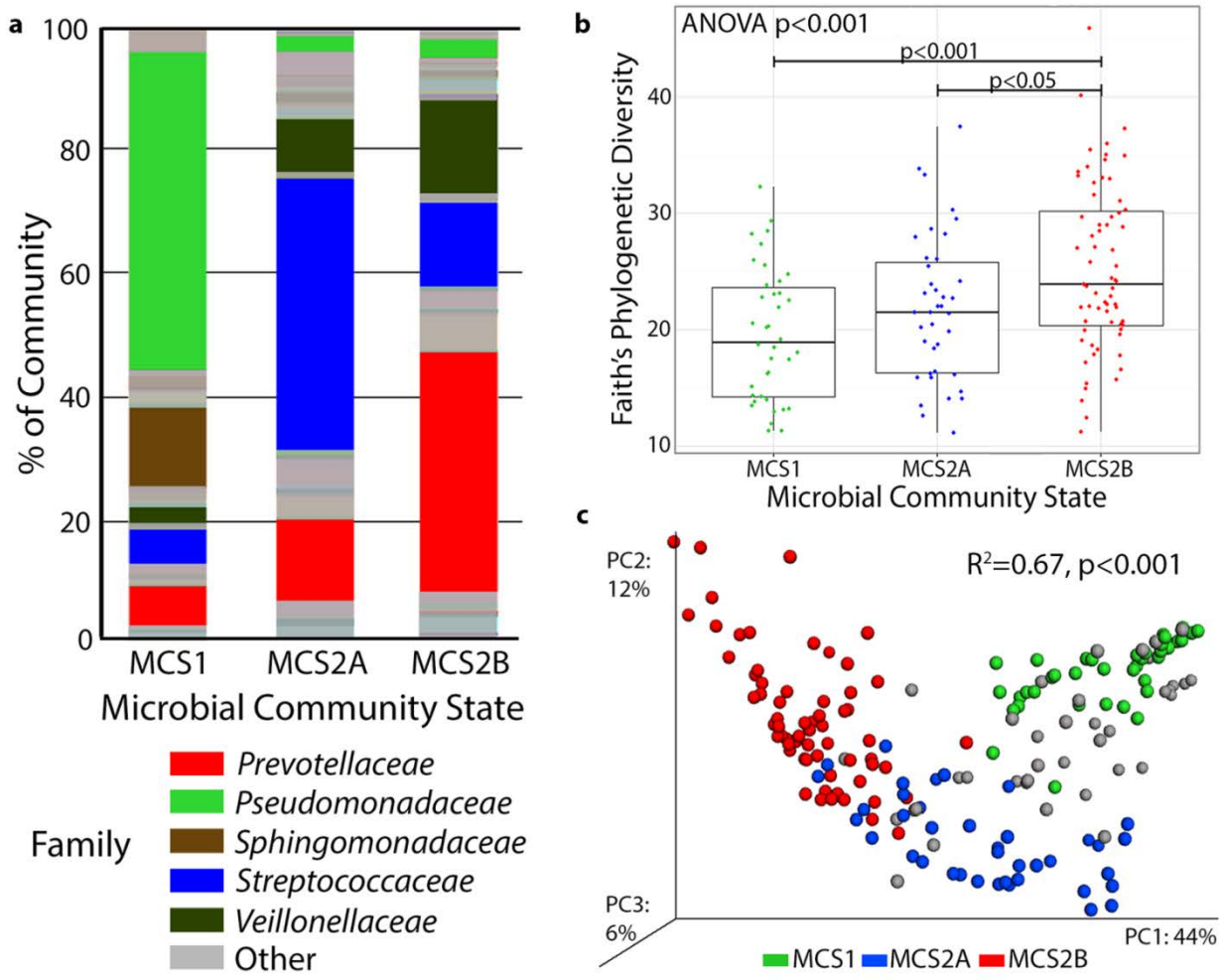


Figure 2. Two compositionally distinct lower airway microbial states exist in HIV infected pneumonia patients. A. Mean community composition of each state at the family level. **B**. Lower airway phylogenetic diversity differs significantly across microbial states (one-way ANOVA, p<0.001). **C**. PCoA plot illustrating weighted UniFrac distances permits visualization of MCS1 (green) and the sister states MCS2A (blue) and MCS2B (red), which collectively explain a significant proportion of bacterial community variation (PERMANOVA, R^2=0.67, p<0.001) within the lower airways of this patient

population. Patient lower airway communities that do not fit one of these three mean community compositions are depicted in gray.

Lower airway states are related to clinical and microbiological factors

We next asked whether the specific factors that explained the variation in bacterial beta-diversity were differentially associated with microbial state (Table S2). Neither gender nor alcohol consumption significantly differed across groups; however, MCS1 communities had significantly higher *Mycobacterium* detection (Chi-squared, p=0.006; Fig. 3A), while MCS2B communities exhibited increased culturable *Aspergillus* (Chi-squared, p=0.07; Fig. 3B). In parallel, the MCS2B group had significantly increased ceftriaxone administered (n=29/65), whereas MCS1 patients were almost never treated with ceftriaxone (n=1/36, Chi-squared, p<0.0001; Fig. 3C). Since ceftriaxone administration may be reflective of infection severity, we compared variables associated with disease severity upon study enrollment (e.g., fever, sputum production, chest pain, etc.) between ceftriaxone-treated and untreated patients; however, no statistically significant differences were observed.

Mortality was tracked from enrollment through 70-days post-bronchoscopy. MCS2B patients exhibited the most deaths at 1-week post-enrollment (n=5/67) while MCS1 patients all survived (p=0.08; Fig. 3D). At 70-days, MCS2B patients still had the highest mortality (22%), followed by MCS2A (16%) and MCS1 patients (13%), though this trend did not reach significance (log-rank test, p>0.05, Fig. S3).

Figure 3.

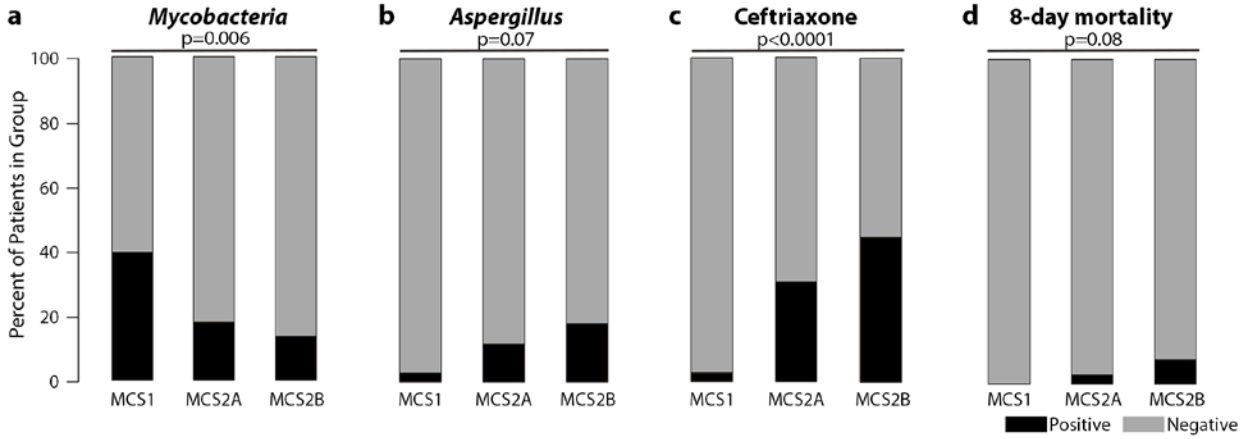


Figure 3. Culture positivity for *Mycobacterium* or *Aspergillus*, as well as antibiotic administration and mortality differ between MCS. A. *Mycobacterium* (Chi-squared, p=0.006) or **B.** *Aspergillus* (p=0.07) culture positivity, **C.** ceftriaxone administration at bronchoscopy (p<0.0001), and **D.** mortality after 1 week of enrollment (p=0.08) differ amongst microbial states (positive in black, negative in gray).

Airway microbial states are predicted to encode functionally distinct metagenomes

We next predicted the metagenomic content of each microbial state using the Phylogenetic Investigation of Communities by Reconstruction of Unobserved States [PICRUSt (*60*)] package. Each microbial state was predicted to encode significantly distinct metagenomes and enriched for a characteristic set of gene pathways (PERMANOVA; R^2=0.10, p<0.001; Fig. 4A). A total of 238 Kyoto Encyclopedia of Genes and Genomes [KEGG (*61, 62*)] pathways differed significantly between the three groups (Kruskal-Wallis, 329 pathways tested, q<0.05, Table S3). Despite decreased community diversity, MCS1 communities were predicted to encode a greater range of functional pathways compared to the other groups (Kruskal-Wallis, pairwise, p<0.001, q<0.05). This group was predicted to be significantly enriched for a broad range of pathways involved in β-lactam, linoleic and arachidonic acid, and tryptophan metabolism; the majority of which (69%) were encoded by *Pseudomonadaceae* in these communities. MCS2A bacterial communities were enriched for pathways involved in biosynthesis of flavonols and ion channels, while MCS2B bacterial communities encoded glycan metabolism and glycosphingolipid biosynthesis pathways and lacked type II polyketide biosynthesis. Few pathways were predicted to be significantly enriched in MCS2B, indicating that associated increased mortality risk may be either driven by differential expression of pathways shared across compositionally distinct communities, or that non-bacterial species such as *Aspergillus*, detected with greater frequency in this microbial state, contribute substantially to their associated pathogenesis.

Microbial community types induce distinct and characteristic lower airway immune responses

RNA extracted from a subset of compositionally representative BAL samples (n=10/MCS) was used to analyze expression of a diverse panel of immune markers, chosen for their known associations with HIV, chronic bacterial infections, or airway inflammatory responses. mRNA expression levels (*GAPDH*-normalized) were used to generate a multivariate profile of host immune response. PERMANOVA analysis indicated that airway immune response was significantly related to the MCS present (PERMANOVA; R^2=0.168, p<0.005). Specific immune responses were significantly enriched in particular MCS (one-way ANOVA, p<0.05; Fig. 4B). For example, MCS1 patients exhibited significantly higher expression of T cell immunoglobulin and mucin domain 3 (*TIM-3*), a glycoprotein expressed by T and innate cells that down-regulates T-helper 1 activity and pro-inflammatory responses (*63*), and plays a key role in the T cell dysfunction that occurs during chronic viral infection (*64, 65*).

By contrast, MCS2A demonstrated the lowest *TIM-3* expression and significantly increased expression of protein-arginine deiminase type-4 (*PADI4*), which converts arginine to citrulline, an α-amino acid post-translationally incorporated into histones, filaggrin and proteins involved in myelination (*66*). Additionally, this MCS trended toward significantly higher levels of interleukin-10 (*IL-10*; anti-inflammatory cytokine) and programmed cell death protein 1 (*PD-1*; T cell negative regulator and exhaustion marker), and lower levels of forkhead box P3 (*FOXP3*; master regulator of regulatory T cells) expression. MCS2B subjects displayed increased interferon-alpha (*IFNα*), which

characteristically protects against infection in immunocompetent subjects, but is associated with rapid disease progression in HIV infection (*67*), *IL-13* (mediator of T-helper 2 cell function), occludin/ELL Domain-Containing Protein 1 (*OCEL1*; maintains and regulates tight junctions), and protein tyrosine phosphatase receptor type C (*PTPRC*; expressed on micro-vesicles produced by HIV infected cells) were also significantly increased. These patients also trended towards increased expression of *IL-5* (mediator of T-helper 2 cell function which stimulates B cell growth), *MUC5AC* (the primary airway mucin), *PD-1*, and *IL-33* (pro-inflammatory cytokine which induces Th2 responses), indicating a significant T-helper 2 skew.

Airway microbial states are associated with distinct circulating metabolites

Paired serum from patients for whom BAL immune profiles were generated (n=30), were examined using Liquid and Gas Chromatography-Mass Spectrometry to determine whether distinct systemic metabolic profiles were associated with airway MCS. A total of 60 metabolites differed significantly between all three groups (Kruskal-Wallis, $p<0.05$; Fig. 4C and Table S4). As the *in silico* metagenomic analysis predicted, MCS1 patients were characterized by significant enrichment of xanthurenate (a tryptophan metabolite) and arachidonic acid metabolites, including the eicosanoids leukotriene B4, a pro-inflammatory lipid-mediator, and 15-HETE (15-hydroxyeicosatetraenoic acid), which induces pulmonary vasoconstriction and edema (*68*). In addition, MCS1 patients were significantly enriched for multiple products of primary and secondary bile acid metabolism (e.g., chenodeoxycholate, glycodeoxycholate, taurochenodeoxycholate, and ursodeoxycholate), several of which

have been shown to upregulate inflammatory responses along with LPS (*69*), indicating that the activities of the gastrointestinal microbiome may also contribute to the tone of host inflammation in MCS1.

As predicted, MCS2B patients were characterized by significantly reduced relative levels of circulating metabolites compared to MCS1 patients. However, significant increases in amino acid metabolites 3-methyl-2-oxobutyrate and 4-methyl-2-oxopentanoate (both involved in valine and leucine metabolism), monoacylglycerols associated with lipid metabolism (1-dihomo-linolenylglycerol and 1-myristoylglycerol), and inosine (purine metabolism) were significantly enriched. MCS2A patients exhibited a mixture of metabolites identified in the other MCS but at lower concentrations, with a unique increase in lysolipid and pyrimidine metabolism and a decrease in monoacylglycerols. Thus, products of several of the biosynthetic or metabolic pathways predicted to discriminate between patients with specific MCS were significantly and differentially enriched in their circulation.

To verify that specific MCS, their predicted metagenomes, local airway immune responses, and serum metabolites were inter-related, we applied Procrustes (*70, 71*) and Mantel (*72*) analyses. Both confirmed a strong and significant correlation between each of these data matrices [correlation between bacterial community composition and **1.** PICRUSt metagenomic prediction (Procrustes: r^2=0.513, p<0.001; Mantel: r^2=0.674, p<0.001), or **2.** Airway immune expression (Procrustes: r^2=0.147, p=0.031; Mantel: r^2=0.122, p=0.067), or **3.** Serum metabolites (Procrustes: r^2=0.414, p<0.001; Mantel r^2=0.286, p<0.001)]. This result indicates that our HIV infected pneumonia patients who

possess distinct airway MCS exhibit corresponding features of immune dysfunction and a characteristic peripheral metabolome.

Figure 4.

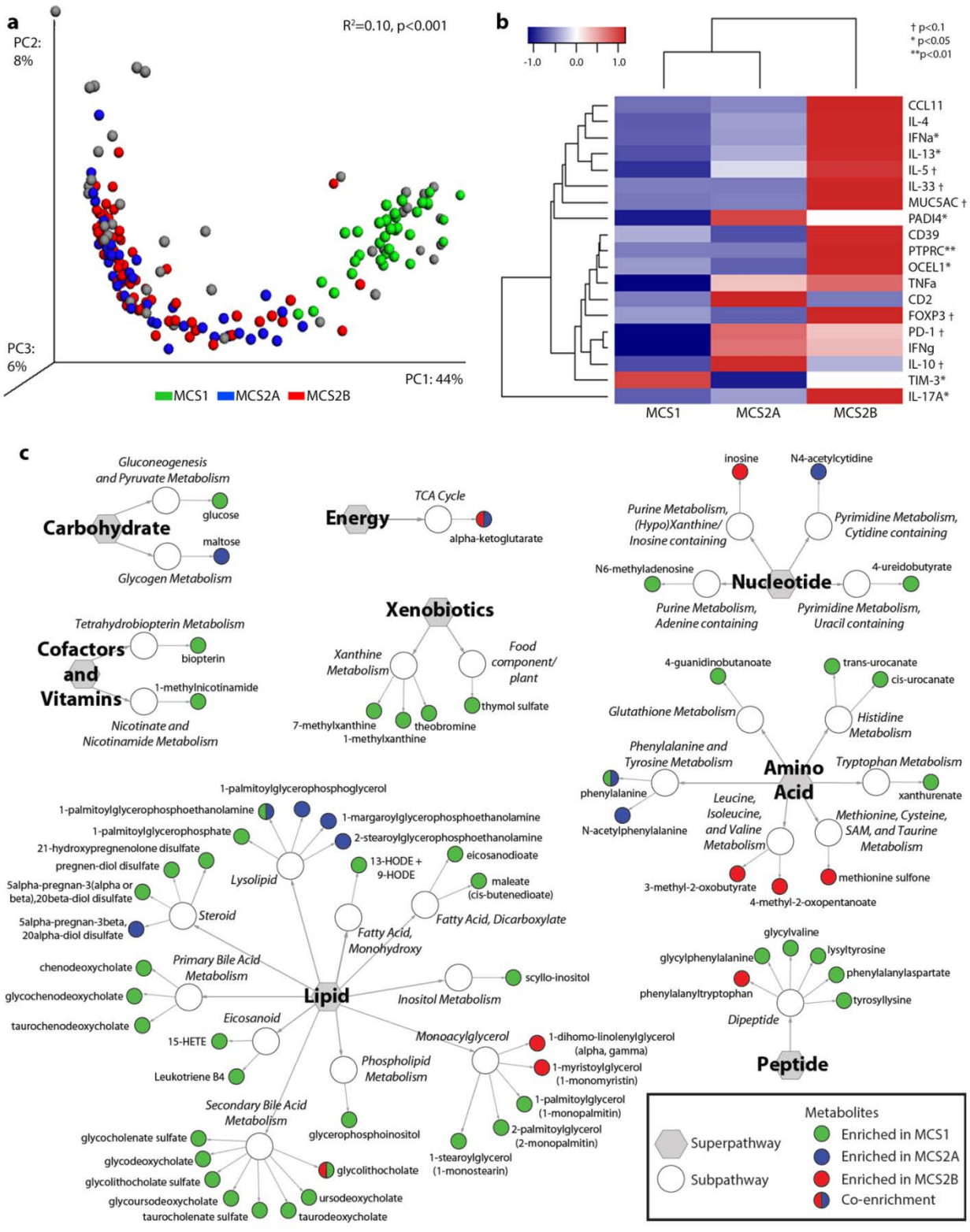

a
PC2: 8%
PC3: 6%
PC1: 44%
R²=0.10, p<0.001
MCS1 MCS2A MCS2B
b
-1.0 0.0 1.0
† p<0.1
* p<0.05
**p<0.01
CCL11
IL-4
IFNa*
IL-13*
IL-5 †
IL-33 †
MUC5AC †
PADI4*
CD39
PTPRC**
OCEL1*
TNFa
CD2
FOXP3 †
PD-1 †
IFNg
IL-10 †
TIM-3*
IL-17A*
MCS1 MCS2A MCS2B
c
Carbohydrate
Gluconeogenesis and Pyruvate Metabolism
glucose
maltose
Glycogen Metabolism
Energy
TCA Cycle
alpha-ketoglutarate
Cofactors and Vitamins
Tetrahydrobiopterin Metabolism
biopterin
1-methylnicotinamide
Nicotinate and Nicotinamide Metabolism
Xenobiotics
Xanthine Metabolism
Food component/ plant
thymol sulfate
7-methylxanthine
theobromine
1-methylxanthine
Nucleotide
inosine
N4-acetylcytidine
Purine Metabolism, (Hypo)Xanthine/ Inosine containing
Pyrimidine Metabolism, Cytidine containing
N6-methyladenosine
4-ureidobutyrate
Purine Metabolism, Adenine containing
Pyrimidine Metabolism, Uracil containing
Amino Acid
4-guanidinobutanoate
trans-urocanate
cis-urocanate
Glutathione Metabolism
Histidine Metabolism
Phenylalanine and Tyrosine Metabolism
Tryptophan Metabolism
phenylalanine
xanthurenate
N-acetylphenylalanine
Leucine, Isoleucine, and Valine Metabolism
Methionine, Cysteine, SAM, and Taurine Metabolism
3-methyl-2-oxobutyrate
methionine sulfone
4-methyl-2-oxopentanoate
Peptide
glycylvaline
glycylphenylalanine
lysyltyrosine
phenylalanylaspartate
phenylalanyltryptophan
tyrosyllysine
Dipeptide
Lipid
1-palmitoylglycerophosphoglycerol
1-palmitoylglycerophosphoethanolamine
1-margaroylglycerophosphoethanolamine
1-palmitoylglycerophosphate
2-stearoylglycerophosphoethanolamine
21-hydroxypregnenolone disulfate
eicosanodioate
pregnen-diol disulfate
13-HODE + 9-HODE
maleate (cis-butenedioate)
5alpha-pregnan-3(alpha or beta),20beta-diol disulfate
Lysolipid
Steroid
Fatty Acid, Dicarboxylate
5alpha-pregnan-3beta, 20alpha-diol disulfate
Fatty Acid, Monohydroxy
chenodeoxycholate
Primary Bile Acid Metabolism
scyllo-inositol
glycochenodeoxycholate
Inositol Metabolism
taurochenodeoxycholate
Eicosanoid
15-HETE
Monoacylglycerol
1-dihomo-linolenylglycerol (alpha, gamma)
Leukotriene B4
Phospholipid Metabolism
1-myristoylglycerol (1-monomyristin)
1-palmitoylglycerol (1-monopalmitin)
Secondary Bile Acid Metabolism
glycerophosphoinositol
2-palmitoylglycerol (2-monopalmitin)
glycocholenate sulfate
1-stearoylglycerol (1-monostearin)
glycodeoxycholate
glycolithocholate
glycolithocholate sulfate
glycoursodeoxycholate
ursodeoxycholate
taurocholenate sulfate
taurodeoxycholate
Superpathway
Subpathway
Metabolites
Enriched in MCS1
Enriched in MCS2A
Enriched in MCS2B
Co-enrichment

Fig. 4. Airway microbial states are predicted to encode distinct metagenomes, and each is shown to induce different lower airway immunological responses and is associated with significantly different serum metabolomes. A. PCoA of *in silico* metagenomic predictions (generated using PICRUSt), indicates that the variation in predicted metagenomic content is significantly explained by MCS designation (PERMANOVA, R^2=0.10, p<0.001). **B.** qRT-PCR array assessing immune-associated gene expression within the lower airways of patients is significantly associated with MCS (PERMANOVA, R^2=0.168, p<0.005). PERMANOVA explains how well specific factors such as MCS explain multivariate data such as cytokine expression (not confined to PCoA or microbiota). Pathway significance: †p<0.1 *p<0.05 **p<0.01. **C.** Comparative LC/MS metabolomic analysis of paired patient serum identified 60 metabolites that differed significantly between all three groups (Kruskal-Wallis; p<0.05). Color code: MCS1 (green), MCS2A (blue), MCS2B (red), and communities not fitting into a MCS mean community composition (gray).

DISCUSSION

Factors that influence pneumonia outcomes in HIV infected patients are poorly defined and we hypothesized that the airway microbiome may influence these outcomes. In this study, three distinct microbial states were identified that exhibited significant differences in alpha diversity, culturable *Aspergillus* or *Mycobacterium,* ceftriaxone administration, immune responses, and metabolic signatures, and trended towards differences in mortality. Recent work by Cribbs and colleagues demonstrated HIV infected patients are enriched for pneumonia-associated bacteria, including *Streptococcus*, even in the absence of airway infection, and exhibit a distinct metabolic microenvironment compared to healthy subjects (*34*). Together with this study, our work suggests that specific lower airway microbial states may lead to functionally relevant metabolic shifts that relate to distinct pathways of disease pathogenesis in HIV infected individuals.

Segal and colleagues have shown that healthy individuals who have an enrichment of oral taxa, including *Prevotella*, within their lower airways, exhibit increased inflammatory cytokines and Th17 cells (*38*). This corroborates our findings that *Prevotellaceae*-dominated airway microbiota promote inflammation within the lower airways, including IL-17A expression. Recent studies have demonstrated that the composition of the airway microbiota influences susceptibility to *Aspergillus* infection (*73*) and that HIV-associated airway disease is related to fungal community alterations, including *Aspergillus* enrichment (*28*). Our data support these findings and suggests that *Aspergillus* prospers in a *Prevotellaceae*-dominated microbiota in the context of a Th2-skewed airway immune response. Several recent studies have confirmed the

capacity of multiple *Aspergillus* species to induce Th2 responses (*74*), particularly in early-stage airway infection (*75*), suggesting that this species may not simply co-colonize MCS2B airways, but may actively define immunological responses characteristic of this patient subgroup. Patients with this airway microbiota were more likely to have been administered ceftriaxone and exhibited the highest mortality rates; one possible conclusion from these observations is that ceftriaxone administration selectively enriches for an MCS2B microbial community, and that their inter-kingdom microbial activities elicit a host immune response that increases mortality risk. However, the paucity of pre-antibiotic bronchoscopic samples, which are both ethically and logistically difficult to obtain, precludes definitive conclusions on whether ceftriaxone administration is responsible for the presence of this more severe MCS, or whether MCS2B assemblages pre-existed in these patients' airways prior to hospitalization.

MCS1 patients, who exhibited the lowest levels of profiled immune activation markers, were predicted to be enriched for pathways involved in linoleic and arachidonic acid metabolism. Leukotriene B4, a product of arachidonic acid metabolism typically produced by leukocytes in response to inflammatory mediators, was detected in significantly increased concentrations in these patients' serum. While circulating leukotriene B4 in MCS1 patients is likely produced by leukocytes, our data suggest that microbial metabolism of arachidonic acid may contribute to their circulating leukotriene B4 and that microbial-derived lipid inflammation may underlie their immune dysfunction. *Mycobacterium* was more prevalent in MCS1 patients, who also exhibited a significant increase in *TIM-3* expression. This is consistent with the findings of Behar and

colleagues who demonstrated *in vivo* surface expression of Tim-3 on macrophages infected with *M. tuberculosis* (*76*).

MCS2A appears to represent an intermediate microbial state, between the MCS2B and MCS1 groups in terms of clinical associations, alpha-diversity, composition, metabolites, and immune expression. This raises the possibility that airway microbiota may be dynamic and transition through distinct microbiological states, particularly under antimicrobial selective pressure; however, large longitudinal studies are necessary to address this possibility. Nonetheless, patients with MCS2A were uniquely characterized by increased lower airways *PADI4* expression. Extracellular bronchial PADI4 has been shown to citrullinate the innate immune defensin human cathelicidin LL-37/human cationic antimicrobial protein-18, rendering it less efficient at neutralizing lipopolysaccharide. PADI4 is detected in the airways of patients with chronic obstructive pulmonary disease, who also exhibit impaired antibacterial response against *Streptococcus* (*77*), indicating that MCS2A patients, who exhibit expansion of *Streptococcaceae* and induction of PADI4, may also have diminished capacity to respond to the dominant bacterial family present in their airways.

Although lower airway colonization is considered uniformly detrimental to patients, we show that specific, repeated airway microbiome states, discriminated upon the basis of microbial composition, function, host immune response, and clinical outcomes, exist in HIV infected pneumonia patient subsets. Though the majority of patients fall into the three microbial states described, we recognize that not all patients belong to these groupings, which is not surprising in light of recent work by Twigg and colleagues showing far greater variation across lower airway communities in advanced-

HIV patients than healthy individuals (*11*). Larger cohorts of patients are necessary to sufficiently power studies examining other rarer microbial states and their immunological and clinical implications. This cohort provides insight into HIV infected Ugandan pneumonia patients; however, these results may have limited applicability to patients in Western countries due to differences in patient demographics, laboratory testing and antibiotic availability, and high HIV-TB co-prevalence in Uganda. Furthermore, while BAL allows identification of general microbiota patterns within the lower airways, examining spatial-specific microbiota and their interactions with the host requires lung biopsies or brushings, which are beyond the scope of this study. While this study did not use paired oral-BAL samples to control for oral contamination, we have previously shown that oral and lower airway microbiota within HIV infected pneumonia patients display niche specificity (*18*). Our data trend towards a significant relationship between mortality and airway MCS. We calculated that to achieve an eighty percent likelihood of detecting a significant difference in mortality (power=0.8), 100 patients per MCS would have to be studied, underscoring the need and utility of larger cohorts. Despite these limitations, this study identified several factors that shape microbial community composition in the lower airways of HIV infected pneumonia patients. Moreover, it identifies distinct bacterial microbiota states that repeat over large numbers of patients and builds an argument that pneumonia patient heterogeneity, with respect to both immunological and clinical outcomes, may be related to compositional and functional differences in airway microbiomes.

MATERIALS AND METHODS

Subjects and sample collection

We enrolled HIV infected subjects admitted to Mulago Hospital in Kampala, Uganda for acute pneumonia from October 2009-December 2011 as part of the Lung Microbiome in Cohorts of HIV Infected Persons (Lung MicroCHIP) Study. Patients underwent two sputum acid fast bacilli (AFB) smear examinations to diagnose pulmonary TB. AFB smear-negative patients underwent bronchoscopy with BAL for clinical diagnosis, with 10mL set aside for microbiome analysis (*19*). Bronchoscopy was performed a median of 3 days after hospital admission (interquartile range 1-4 days). Bronchoscopy with BAL was performed, and BAL was stored 1:2 in RNAlater (Thermo Fisher Scientific). Notably, during the bronchoscopy, suction is not started until the bronchoscope has been wedged, limiting oral contamination. Venipuncture for blood specimens was performed at enrollment, hospital day #1, and collected into serum separator tubes. MCS groups were rigorously checked across dates of collection, shipping groups, and sequencing batches to ensure that microbial community was not a result of collection, shipping, or sample processing methods. Each of the three MCS described were observed throughout the timeline of sample collection and processing. Over 98% of subjects had received antibiotics prior to bronchoscopy. Clinical data were collected and diagnoses were assigned using standardized forms as previously described (*19*). Study endpoint was mortality follow-up at 70-days post bronchoscopy.

Ethics statement

The Makerere University School of Medicine Research Ethics Committee, the Mulago Hospital Research and Ethics Committee, the Uganda National Council for Science and Technology, and the University of California San Francisco Committee on Human Research approved the protocol. Subjects provided written, informed consent.

BAL processing and extraction

BAL was stored 1:2 vol:vol in RNAlater for stable storage and shipping from Uganda to San Francisco. RNA and DNA were extracted in parallel from whole BAL. Thawed BAL samples were centrifuged at high speed, the supernatant was removed and pellets were resuspended in sterile PBS, and centrifuged again at high speed, prior to extraction. Total DNA and RNA were extracted from whole BAL in parallel using an AllPrep DNA/RNA extraction kit (Qiagen) according to manufacturer's instructions.

DNA and RNA extraction

Total DNA and RNA were extracted from whole BAL in parallel using an AllPrep DNA/RNA extraction kit [Qiagen, (*18*)]; RNA quality and purity were assessed as previously described (*19*).

16S rRNA gene amplification and sequencing

The V4-region of bacterial 16S rRNA gene was amplified using primers with barcodes for multiplex sequencing (Table S1). PCRs were performed in triplicate with parallel no-template control reactions in which no amplification product was observed. Triplicate PCR product was pooled and visualized on 2% agarose gel. If there were no

visible background bands >150bp, Agencourt AMPure XP System (Beckman Coulter) was used for PCR product purification, otherwise the product was purified using QIAquick Gel Extraction Kit (Qiagen). Quality of purified amplicon was confirmed using a Bioanalyzer and the Agilent DNA 1000 Kit (Agilent Technologies). Purified products were quantified using Qubit dsDNA HS Assay Kit (Life Technologies) and pooled equimolar for multiplex sequencing. Sequencing was performed using a MiSeq platform and MiSeq Control Software v2.2.0 (Illumina).

A mock community composed of equal genomic concentration (2 ng per reaction) of *Escherichia coli* ATCC25922, *Pseudomonas aeruginosa* ATCC27853, *Corynebacterium tuberculostearicum* ATCC35692, *Lactobacillus sakei* ATCC15521, and *L. rhamnosus* ATCC53103 was used to monitor and standardize data between runs.

Microbiome data processing

251 bp paired-end reads were assembled using FLASh [Fast Length Adjustment of SHort reads(*78*)] with parameters: -r 251 -f 300 -s 30 -m 15. Assembled sequences were quality trimmed using QIIME software (*71*) with default settings except Phred quality threshold -q 30. Chimeras were removed using ChimeraSlayer (*79*). Each sample was rarefied 100 times to 100,000 reads in the R environment (*80*); the centroid of each sample distribution was subsequently used for analysis (n=182). Greengenes database May 2013 (*81*) was used to classify taxa; singleton OTUs were removed.

Immune gene expression

Total RNA (0.5 ug) was reverse transcribed using the RT First Strand Kit, and cDNA gene expression was assessed on a custom RT Profiler PCR Array (both Qiagen). Real time PCR and detection was performed on QuantStudio 6 Flex (Thermo Fisher Scientific) by two-step cycling: 95°C for 10 minutes, then 40 cycles: 15 seconds at 95°C, 1 minute at 60°C. The custom PCR array included *IFNγ*, *IFNα*, *TNFα*, *MUC5AC*, *IL-17A*, *IL-4*, *IL-5*, *IL-13*, *IL-33*, *OCEL1*, *CCL11*, *PADI4*, *IL-10*, *FOXP3*, *PD-1*, *TIM-3*, *CD45RO*, *CD2*, *CD39*, and *GAPDH*; the last was used with the delta-delta CT method to normalize host immune gene expression (*82*).

Metabolic profiling

Metabolic profiles were generated from 100 uL of patient serum (n=30) by Ultrahigh Performance Liquid and Gas Chromatography-Tandem Mass Spectrometry (UPLC-MS/MS, GC-MS) at Metabolon according to a standard protocol.

Microbial and statistical analyses

Microbial analyses were performed using QIIME software (*71*). Results were visualized using Emperor (*83*). Metagenomic predictions were generated using PICRUSt (*60*). Procrustes ("transform_coordinate_matrices.py" script, "-r 1000") and Mantel tests ("compare_distance_matrices.py" script) were performed in QIIME using Bray Curtis dissimilarity [compositional dissimilarity based on taxon relative abundance (*84*)]. Statistical analyses (e.g., one-way ANOVA, Kruskal-Wallis, etc.) were performed in the R-environment. Dirichlet Multinomial Mixtures and log-rank test were performed using the *DirichletMultinomial* (*59*) and *survival* packages, respectively. Permutational

Multivariate Analysis of Variance [PERMANOVA (*55*), *vegan* v.2.3.0, 1000 permutations] and PCoA were performed using weighted UniFrac (*56, 85*) and Canberra dissimilarity measurements. PERMANOVA independently considers each factor (e.g., age, gender) against bacterial community beta-diversity variance, permuting data independently, and thus does not require false discovery correction. The resulting R^2 provides the proportion of variation explained, e.g., a factor that has a R^2=0.021 explains 2% of the variation in community composition.

Data and materials availability

Raw sequencing data available via the SRA database under SRP077299.

SUPPLEMENTARY FIGURES AND TABLES

Figure S1.

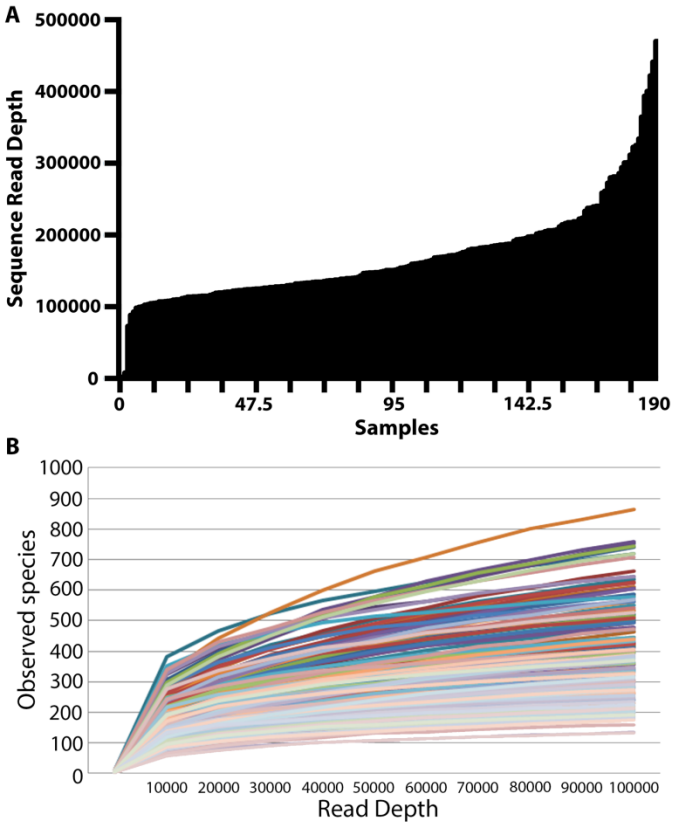


Figure S1. Sequence read depth and observed species distributions. A. Distribution of sequence read depths across all 190 sequenced BAL samples. Read-depth sub-sampling cut-off was set to 100,000 sequences per sample, resulting in the exclusion of eight samples (with insufficient read depth) from further analyses. **B.** Rarefaction curves showing total observed species per sample across multiple read depths, indicates that 100,000 reads results in good community coverage, demonstrated by curves approaching a plateau.

Figure S2.

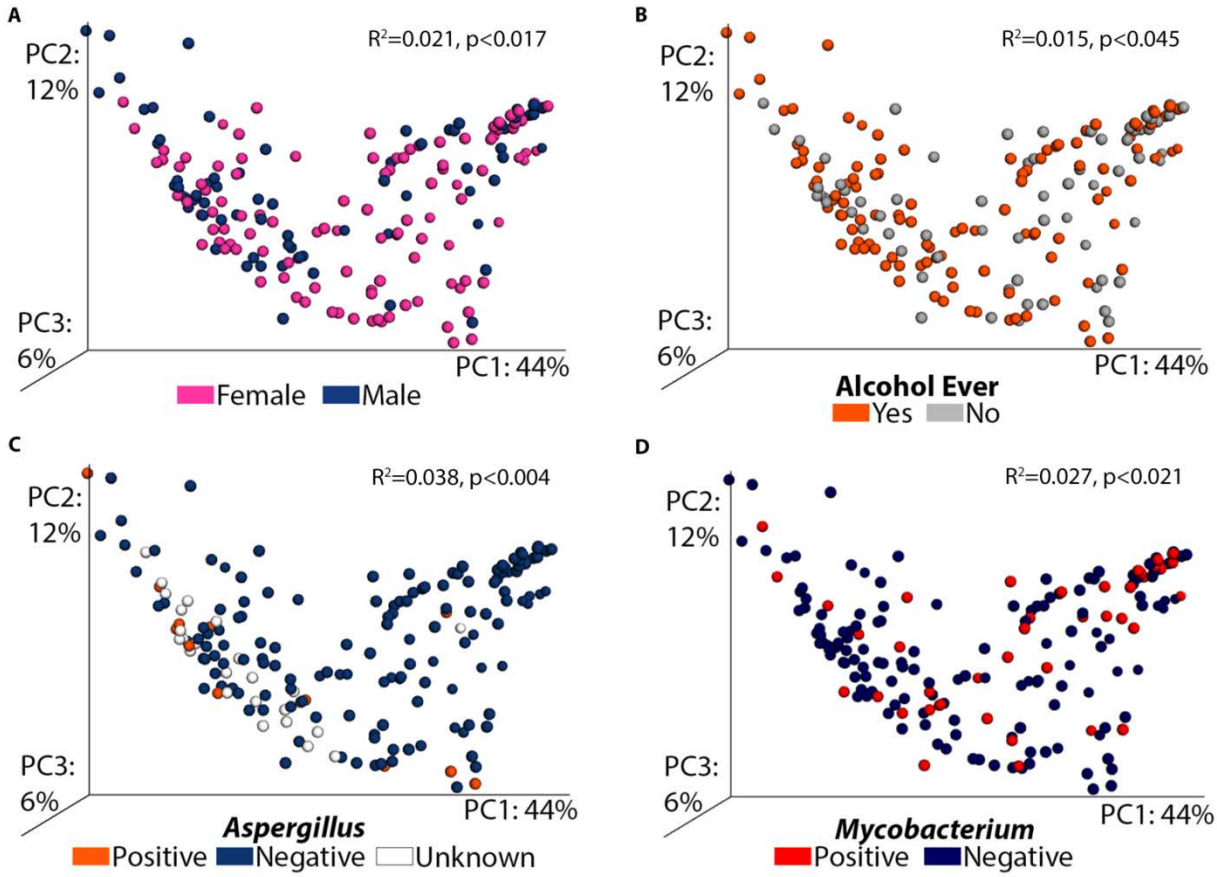


Figure S2. Clinical and demographic factors are significantly associated with community composition. PCoAs comparing community composition of the lower airways to **A.** gender (PERMANOVA, R^2=0.021, p<0.017), **B.** alcohol ever consumed (PERMANOVA, R^2=0.015, p<0.045), **C.** *Aspergillus* positive culture (PERMANOVA, R^2=0.038, p<0.004), and **D.** *Mycobacterium* positive culture (PERMANOVA, R^2=0.027, p<0.021).

Figure S3.

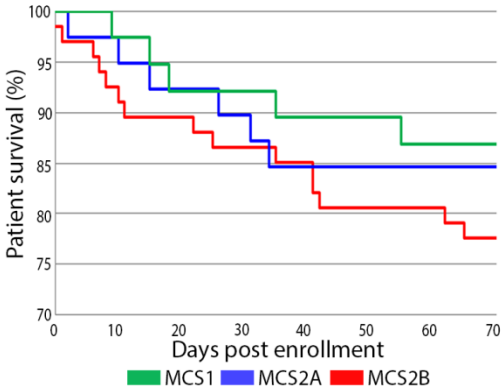


Figure S3. Microbial community states are associated with mortality outcomes.

Kaplan-Meier curves comparing mortality among microbial states from enrollment for 70 days ($p>0.05$).

Table S1. Primer sequences

Primer Name	Primer Sequence
515F_w_adapter	5'-AATGATACGGCGACCACCGAGATCTACAC-TATGGTAATT-GTGTGCCAGCMGCCGCGGTAA -3'
806R_w_adapter	5'-CAAGCAGAAGACGGCATACGAGAT-XXXXXXXXXXXX-AGTCAGTCAG-CCGGACTACHVGGGTWTCTAAT-3'[a]

[a] 12 "X" represents 12-bp Golay barcode

Table S2. Summary of demographic variables.

Variable Name	Samples (n)	Yes | No Or Range [min-max (med)]	Statistical Test across MCS	MCS1	MCS2A	MCS2B	p-value	Sig.
Pneumocystis positive (diagnosis by BAL microscopy)	182	6 | 176 (y|n)	χ^2 test	1 | 37	3 | 36	2 | 65	0.43	
Mycobacteria culture positive	182	41 | 141 (y|n)	χ^2 test	23 | 15	32 | 7	58 | 9	0.006	**
Aspergillus culture positive	157	15 | 142 (y|n)	χ^2 test	1 | 37	4 | 31	9 | 41	0.07	†
Pulmonary *Kaposi's Sarcoma*	182	14 | 168 (y|n)	χ^2 test	2 | 36	2 | 37	7 | 60	0.49	
Age	182	19.1 - 62.6 (34.1)	One-way ANOVA	36.5	35.1	35.6	0.84	
Gender	182	110 | 72 (F|M)	χ^2 test	23 | 15	28 | 11	36 | 31	0.19	
Clinical score at admission	133	1 | 4 | 59 | 39 | 28 | 2 (Normal |Unaffected| Ambulatory| <50% in bed| >50% in bed| Bed bound	χ^2 test	0 | 1 | 11 | 4 | 5 | 0	0 | 2 | 16 | 5 | 6 | 0	1 | 1 | 25 | 14 | 8 | 2	0.83	
Temperature at admission (Celsius)	182	33.7 - 39.9 (36.7)	One-way ANOVA	36.7	37.1	36.7	0.25	
Heart rate at admission	182	48 - 165 (101.5)	One-way ANOVA	101.1	107.6	97.2	0.06	†
Respiratory rate at admission (breaths/ minute)	182	18 - 64 (30)	One-way ANOVA	30.5	32.9	30.1	0.19	
Oxygen saturation at admission	182	76 - 99 (96)	One-way ANOVA	94.3	94.6	94.7	0.90	
Reported fevers, chills, night sweats	182	159 | 23 (y|n)	χ^2 test	31 | 7	38 | 1	58 | 9	0.08	†
How long fevers (weeks)	159	1 - 24 (4)	One-way ANOVA	5.0	4.4	5.0	0.8	
Weight loss at admission	182	170 | 12 (y|n)	χ^2 test	36 | 2	39 | 0	58 | 9	0.03	*
Amount of weight loss	170	38 | 132 (< 5kg | >= 5kg)	χ^2 test	11 | 25	9 | 30	12 | 46	0.55	
How long cough (weeks)	182	1 - 24 (4)	One-way ANOVA	6.2	7.0	6.4	0.76	
Coughing sputum at admission	182	178 | 4 (y|n)	χ^2 test	38 | 0	39 | 0	64 | 3	0.17	

Variable Name	Samples (n)	Yes | No Or Range [min-max (med)]	Statistical Test across MCS	MCS1	MCS2A	MCS2B	p-value	Sig.
How long coughing sputum (weeks)	178	1 - 24 (3)	One-way ANOVA	4.3	5.6	4.5	0.33	
Color of sputum at admission	178	93 | 85 (White or clear | Discolored)	χ^2 test	20 | 18	19 | 20	34 | 30	0.90	
Blood in sputum at admission	178	35 | 143 (y|n)	χ^2 test	7 | 31	10 | 29	13 | 51	0.72	
Dyspnea at admission	182	82 | 100 (y|n)	χ^2 test	22 | 16	19 | 20	29 | 38	0.35	
How long dyspnea (weeks)	82	1 - 24 (4)	One-way ANOVA	4.7	4.7	4.2	0.9	
Severity of dyspnea	82	42 | 40 (Only exercise | At rest)	χ^2 test	10 | 12	10 | 9	15 | 14	0.87	
Chest pain at admission	182	116 | 66 (y|n)	χ^2 test	24 | 14	28 | 11	42 | 25	0.61	
How long chest pain (weeks)	116	1 - 24 (2.5)	One-way ANOVA	4.1	3.7	5.1	0.39	
Wheeze at admission	182	39 | 143 (y|n)	χ^2 test	10 | 28	10 | 29	13 | 54	0.64	
How long wheeze (weeks)	39	1 - 16 (3)	One-way ANOVA	3.7	4.9	3.7	0.67	
On *Pneumocystis* prophylaxis	121	94 | 27 (y|n)	χ^2 test	18 | 9	22 | 3	30 | 10	0.19	
On antiretrovirals	121	39 | 82 (y|n)	χ^2 test	8 | 19	8 | 17	13 | 27	0.97	
CD4 count	179	1 - 697 (71)	One-way ANOVA	179	119	134	0.19	
Previous diagnosis *TB*	182	14 | 168 (y|n)	χ^2 test	1 | 37	5 | 34	4 | 63	0.19	
Smoked >99 cigarettes in lifetime	182	41 | 141 (y|n)	χ^2 test	10 | 28	4 | 35	19 | 48	0.09	†
Alcohol ever	182	113 | 69 (y|n)	χ^2 test	20 | 18	27 | 12	44 | 23	0.27	
Last known vital status	182	143 | 39 (live|dead)	χ^2 test	32 | 6	32 | 7	51 | 16	0.56	
Vital status 30 days post bronchoscopy	182	163 | 19 (live|dead)	χ^2 test	35 | 3	35 | 4	58 | 9	0.67	
Vital status 70 days post bronchoscopy	182	148 | 34 (live|dead)	χ^2 test	33 | 5	33 | 6	52 | 15	0.44	
Reported sulfamethoxazole-trimethoprim within last 2 weeks	178	137 | 41 (y|n)	χ^2 test	30 | 7	30 | 6	48 | 19	0..32	
Reported penicillin within last 2 weeks	178	83 | 95 (y|n)	χ^2 test	17 | 20	16 | 20	31 | 36	0.98	

Variable Name	Samples (n)	Yes | No Or Range [min-max (med)]	Statistical Test across MCS	MCS1	MCS2A	MCS2B	p-value	Sig.
Reported ceftriaxone within last 2 weeks	178	137 | 41 (y|n)	χ^2 test	32 | 6	25 | 11	54 | 13	0.26	
Reported quinolone within last 2 weeks	172	13 | 159 (y|n)	χ^2 test	3 | 34	3 | 30	7 | 59	0.91	
Reported macrolide within last 2 weeks	178	50 | 128 (y|n)	χ^2 test	12 | 25	11 | 25	16 | 51	0.59	
Sulfamethoxazole-trimethoprim at bronchoscopy	177	103 | 74 (y|n)	χ^2 test	19 | 18	25 | 11	33 | 34	0.13	
Penicillin at bronchoscopy	177	46 | 131 (y|n)	χ^2 test	8 | 29	9 | 27	18 | 49	0.84	
Ceftriaxone at bronchoscopy	174	54 | 120 (y|n)	χ^2 test	1 | 35	11 | 25	29 | 36	$6.3x10^{-5}$	***
Quinolone at bronchoscopy	170	7 | 163 (y|n)	χ^2 test	3 | 34	2 | 31	2 | 63	0.53	
Macrolide at bronchoscopy	174	27 | 147 (y|n)	χ^2 test	4 | 33	5 | 31	11 | 54	0.70	

Significance: † p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Table S3. PICRUSt predicted pathways enriched in each microbial state.

Enriched Group	Pathway	Sub Pathway	Super Pathway	p-value	Bonferroni p-value
MCS1	Apoptosis	Cell Growth and Death	Cellular Processes	9.30E-22	3.00E-19
	Cell cycle - *Caulobacter*	Cell Growth and Death	Cellular Processes	1.60E-15	5.30E-13
	Meiosis - yeast	Cell Growth and Death	Cellular Processes	2.80E-11	9.30E-09
	p53 signaling pathway	Cell Growth and Death	Cellular Processes	3.10E-18	1.00E-15
	Bacterial chemotaxis	Cell Motility	Cellular Processes	5.50E-19	1.80E-16
	Bacterial motility proteins	Cell Motility	Cellular Processes	9.90E-19	3.20E-16
	Cytoskeleton proteins	Cell Motility	Cellular Processes	3.30E-19	1.10E-16
	Flagellar assembly	Cell Motility	Cellular Processes	1.20E-18	3.80E-16
	Endocytosis	Transport and Catabolism	Cellular Processes	3.80E-09	1.30E-06
	Peroxisome	Transport and Catabolism	Cellular Processes	7.70E-18	2.50E-15
	ABC transporters	Membrane Transport	Environmental Information Processing	4.30E-20	1.40E-17
	Bacterial secretion system	Membrane Transport	Environmental Information Processing	5.20E-19	1.70E-16
	Secretion system	Membrane Transport	Environmental Information Processing	7.70E-21	2.50E-18
	Transporters	Membrane Transport	Environmental Information Processing	1.10E-21	3.60E-19
	MAPK signaling pathway – yeast	Signal Transduction	Environmental Information Processing	6.30E-18	2.10E-15
	Phosphatidylinositol signaling system	Signal Transduction	Environmental Information Processing	1.10E-21	3.50E-19
	Two-component system	Signal Transduction	Environmental Information Processing	2.30E-20	7.40E-18
	Bacterial toxins	Signaling Molecules and Interaction	Environmental Information Processing	7.30E-12	2.40E-09
	Cellular antigens	Signaling Molecules and Interaction	Environmental Information Processing	3.10E-22	1.00E-19
	Chaperones and folding catalysts	Folding, Sorting and Degradation	Genetic Information Processing	1.40E-19	4.70E-17

Enriched Group	Pathway	Sub Pathway	Super Pathway	p-value	Bonferroni p-value
	Proteasome	Folding, Sorting and Degradation	Genetic Information Processing	1.60E-16	5.30E-14
	Protein export	Folding, Sorting and Degradation	Genetic Information Processing	4.00E-06	1.30E-03
	Protein processing in endoplasmic reticulum	Folding, Sorting and Degradation	Genetic Information Processing	1.50E-13	4.90E-11
	RNA degradation	Folding, Sorting and Degradation	Genetic Information Processing	9.90E-18	3.20E-15
	Sulfur relay system	Folding, Sorting and Degradation	Genetic Information Processing	4.20E-19	1.40E-16
	Ubiquitin system	Folding, Sorting and Degradation	Genetic Information Processing	7.80E-15	2.50E-12
	Base excision repair	Replication and Repair	Genetic Information Processing	1.40E-17	4.50E-15
	Chromosome	Replication and Repair	Genetic Information Processing	8.00E-18	2.60E-15
	DNA repair and recombination proteins	Replication and Repair	Genetic Information Processing	7.00E-15	2.30E-12
	DNA replication	Replication and Repair	Genetic Information Processing	1.80E-05	6.00E-03
	DNA replication proteins	Replication and Repair	Genetic Information Processing	1.50E-06	5.00E-04
	Mismatch repair	Replication and Repair	Genetic Information Processing	1.60E-05	5.40E-03
	Non-homologous end-joining	Replication and Repair	Genetic Information Processing	1.90E-18	6.20E-16
	Nucleotide excision repair	Replication and Repair	Genetic Information Processing	7.60E-09	2.50E-06
	Transcription factors	Transcription	Genetic Information Processing	4.20E-22	1.40E-19
	Transcription machinery	Transcription	Genetic Information Processing	8.90E-20	2.90E-17
	Ribosome biogenesis	Translation	Genetic Information Processing	1.50E-18	4.80E-16
	Ribosome biogenesis in eukaryotes	Translation	Genetic Information Processing	1.40E-17	4.50E-15
	Bladder cancer	Cancers	Human Diseases	6.70E-13	2.20E-10
	Colorectal cancer	Cancers	Human Diseases	3.10E-18	1.00E-15
	Pathways in cancer	Cancers	Human Diseases	1.40E-19	4.50E-17
	Prostate cancer	Cancers	Human Diseases	2.50E-15	8.00E-13

Enriched Group	Pathway	Sub Pathway	Super Pathway	p-value	Bonferroni p-value
	Renal cell carcinoma	Cancers	Human Diseases	1.10E-18	3.70E-16
	Small cell lung cancer	Cancers	Human Diseases	3.00E-18	9.80E-16
	Hypertrophic cardiomyopathy (HCM)	Cardiovascular Diseases	Human Diseases	3.90E-16	1.30E-13
	Viral myocarditis	Cardiovascular Diseases	Human Diseases	3.10E-18	1.00E-15
	Primary immunodeficiency	Immune System Diseases	Human Diseases	3.30E-13	1.10E-10
	Systemic lupus erythematosus	Immune System Diseases	Human Diseases	2.30E-17	7.40E-15
	African trypanosomiasis	Infectious Diseases	Human Diseases	2.50E-18	8.30E-16
	Chagas disease (American trypanosomiasis)	Infectious Diseases	Human Diseases	6.10E-18	2.00E-15
	Epithelial cell signaling in *Helicobacter pylori* infection	Infectious Diseases	Human Diseases	1.10E-12	3.60E-10
	Influenza A	Infectious Diseases	Human Diseases	3.10E-18	1.00E-15
	Pertussis	Infectious Diseases	Human Diseases	2.80E-18	9.20E-16
	Toxoplasmosis	Infectious Diseases	Human Diseases	3.10E-18	1.00E-15
	Tuberculosis	Infectious Diseases	Human Diseases	4.20E-15	1.40E-12
	Vibrio cholerae infection	Infectious Diseases	Human Diseases	1.50E-13	4.90E-11
	Vibrio cholerae pathogenic cycle	Infectious Diseases	Human Diseases	4.00E-19	1.30E-16
	Type I diabetes mellitus	Metabolic Diseases	Human Diseases	7.50E-05	2.50E-02
	Type II diabetes mellitus	Metabolic Diseases	Human Diseases	5.70E-18	1.90E-15
	Alzheimer's disease	Neurodegenerative Diseases	Human Diseases	1.10E-18	3.60E-16
	Amyotrophic lateral sclerosis (ALS)	Neurodegenerative Diseases	Human Diseases	2.00E-18	6.50E-16
	Huntington's disease	Neurodegenerative Diseases	Human Diseases	6.50E-21	2.10E-18
	Parkinson's disease	Neurodegenerative Diseases	Human Diseases	3.80E-19	1.30E-16

Enriched Group	Pathway	Sub Pathway	Super Pathway	p-value	Bonferroni p-value
	Prion diseases	Neurodegenerative Diseases	Human Diseases	4.20E-16	1.40E-13
	Alanine, aspartate, and glutamate metabolism	Amino Acid Metabolism	Metabolism	1.30E-18	4.30E-16
	Amino acid related enzymes	Amino Acid Metabolism	Metabolism	1.10E-16	3.80E-14
	Arginine and proline metabolism	Amino Acid Metabolism	Metabolism	7.50E-19	2.50E-16
	Cysteine and methionine metabolism	Amino Acid Metabolism	Metabolism	2.10E-17	6.80E-15
	Glycine, serine, and threonine metabolism	Amino Acid Metabolism	Metabolism	1.50E-18	5.00E-16
	Histidine metabolism	Amino Acid Metabolism	Metabolism	9.00E-18	2.90E-15
	Lysine biosynthesis	Amino Acid Metabolism	Metabolism	1.70E-17	5.60E-15
	Lysine degradation	Amino Acid Metabolism	Metabolism	2.00E-18	6.60E-16
	Phenylalanine metabolism	Amino Acid Metabolism	Metabolism	1.30E-18	4.30E-16
	Phenylalanine, tyrosine, and tryptophan biosynthesis	Amino Acid Metabolism	Metabolism	2.70E-17	8.90E-15
	Tryptophan metabolism	Amino Acid Metabolism	Metabolism	1.30E-18	4.40E-16
	Tyrosine metabolism	Amino Acid Metabolism	Metabolism	2.80E-19	9.10E-17
	Valine, leucine, and isoleucine biosynthesis	Amino Acid Metabolism	Metabolism	6.00E-21	2.00E-18
	Valine, leucine, and isoleucine degradation	Amino Acid Metabolism	Metabolism	4.20E-18	1.40E-15
	beta-Lactam resistance	Biosynthesis of Other Secondary Metabolites	Metabolism	5.60E-15	1.80E-12
	Betalain biosynthesis	Biosynthesis of Other Secondary Metabolites	Metabolism	1.10E-07	3.50E-05
	Caffeine metabolism	Biosynthesis of Other Secondary Metabolites	Metabolism	1.10E-09	3.70E-07
	Flavonoid biosynthesis	Biosynthesis of Other Secondary Metabolites	Metabolism	1.20E-13	3.90E-11
	Isoflavonoid biosynthesis	Biosynthesis of Other Secondary Metabolites	Metabolism	5.30E-05	1.80E-02
	Isoquinoline alkaloid biosynthesis	Biosynthesis of Other Secondary Metabolites	Metabolism	6.40E-21	2.10E-18

Enriched Group	Pathway	Sub Pathway	Super Pathway	p-value	Bonferroni p-value
	Novobiocin biosynthesis	Biosynthesis of Other Secondary Metabolites	Metabolism	5.60E-20	1.80E-17
	Penicillin and cephalosporin biosynthesis	Biosynthesis of Other Secondary Metabolites	Metabolism	1.10E-17	3.70E-15
	Phenylpropanoid biosynthesis	Biosynthesis of Other Secondary Metabolites	Metabolism	7.70E-17	2.50E-14
	Stilbenoid, diarylheptanoid, and gingerol biosynthesis	Biosynthesis of Other Secondary Metabolites	Metabolism	9.30E-15	3.10E-12
	Streptomycin biosynthesis	Biosynthesis of Other Secondary Metabolites	Metabolism	4.20E-15	1.40E-12
	Tropane, piperidine, and pyridine alkaloid biosynthesis	Biosynthesis of Other Secondary Metabolites	Metabolism	1.30E-20	4.20E-18
	Amino sugar and nucleotide sugar metabolism	Carbohydrate Metabolism	Metabolism	8.10E-12	2.70E-09
	Ascorbate and aldarate metabolism	Carbohydrate Metabolism	Metabolism	2.90E-17	9.60E-15
	Butanoate metabolism	Carbohydrate Metabolism	Metabolism	2.70E-19	8.80E-17
	C5-Branched dibasic acid metabolism	Carbohydrate Metabolism	Metabolism	3.60E-19	1.20E-16
	Citrate cycle (TCA cycle)	Carbohydrate Metabolism	Metabolism	9.00E-19	2.90E-16
	Fructose and mannose metabolism	Carbohydrate Metabolism	Metabolism	9.20E-14	3.00E-11
	Glycolysis / Gluconeogenesis	Carbohydrate Metabolism	Metabolism	4.90E-19	1.60E-16
	Glyoxylate and dicarboxylate metabolism	Carbohydrate Metabolism	Metabolism	2.50E-18	8.20E-16
	Inositol phosphate metabolism	Carbohydrate Metabolism	Metabolism	5.30E-20	1.70E-17
	Pentose and glucuronate interconversions	Carbohydrate Metabolism	Metabolism	5.20E-18	1.70E-15
	Pentose phosphate pathway	Carbohydrate Metabolism	Metabolism	3.30E-18	1.10E-15

Enriched Group	Pathway	Sub Pathway	Super Pathway	p-value	Bonferroni p-value
	Propanoate metabolism	Carbohydrate Metabolism	Metabolism	1.80E-19	6.00E-17
	Pyruvate metabolism	Carbohydrate Metabolism	Metabolism	2.50E-21	8.40E-19
	Starch and sucrose metabolism	Carbohydrate Metabolism	Metabolism	1.10E-09	3.50E-07
	Carbon fixation in photosynthetic organisms	Energy Metabolism	Metabolism	5.60E-12	1.80E-09
	Carbon fixation pathways in prokaryotes	Energy Metabolism	Metabolism	1.00E-18	3.30E-16
	Methane metabolism	Energy Metabolism	Metabolism	2.50E-17	8.20E-15
	Nitrogen metabolism	Energy Metabolism	Metabolism	3.70E-19	1.20E-16
	Oxidative phosphorylation	Energy Metabolism	Metabolism	4.80E-19	1.60E-16
	Sulfur metabolism	Energy Metabolism	Metabolism	1.20E-18	4.00E-16
	Cytochrome P450	Enzyme Families	Metabolism	2.20E-10	7.10E-08
	Peptidases	Enzyme Families	Metabolism	3.30E-18	1.10E-15
	Protein kinases	Enzyme Families	Metabolism	2.40E-21	7.80E-19
	Glycosyltransferases	Glycan Biosynthesis and Metabolism	Metabolism	4.90E-18	1.60E-15
	Lipopolysaccharide biosynthesis	Glycan Biosynthesis and Metabolism	Metabolism	2.20E-19	7.10E-17
	Lipopolysaccharide biosynthesis proteins	Glycan Biosynthesis and Metabolism	Metabolism	8.40E-22	2.70E-19
	Peptidoglycan biosynthesis	Glycan Biosynthesis and Metabolism	Metabolism	5.80E-06	1.90E-03
	alpha-Linolenic acid metabolism	Lipid Metabolism	Metabolism	1.20E-17	4.10E-15
	Arachidonic acid metabolism	Lipid Metabolism	Metabolism	1.10E-20	3.70E-18
	Biosynthesis of unsaturated fatty acids	Lipid Metabolism	Metabolism	1.10E-18	3.70E-16
	Ether lipid metabolism	Lipid Metabolism	Metabolism	1.50E-14	5.00E-12
	Fatty acid biosynthesis	Lipid Metabolism	Metabolism	4.00E-21	1.30E-18
	Fatty acid metabolism	Lipid Metabolism	Metabolism	2.30E-18	7.40E-16

Enriched Group	Pathway	Sub Pathway	Super Pathway	p-value	Bonferroni p-value
	Glycerolipid metabolism	Lipid Metabolism	Metabolism	5.10E-22	1.70E-19
	Glycerophospholipid metabolism	Lipid Metabolism	Metabolism	2.10E-18	7.00E-16
	Linoleic acid metabolism	Lipid Metabolism	Metabolism	1.30E-18	4.30E-16
	Lipid biosynthesis proteins	Lipid Metabolism	Metabolism	5.90E-18	1.90E-15
	Primary bile acid biosynthesis	Lipid Metabolism	Metabolism	4.80E-17	1.60E-14
	Steroid biosynthesis	Lipid Metabolism	Metabolism	3.20E-15	1.10E-12
	Steroid hormone biosynthesis	Lipid Metabolism	Metabolism	3.80E-18	1.20E-15
	Synthesis and degradation of ketone bodies	Lipid Metabolism	Metabolism	6.60E-20	2.20E-17
	Biotin metabolism	Metabolism of Cofactors and Vitamins	Metabolism	3.70E-18	1.20E-15
	Folate biosynthesis	Metabolism of Cofactors and Vitamins	Metabolism	8.50E-18	2.80E-15
	Lipoic acid metabolism	Metabolism of Cofactors and Vitamins	Metabolism	2.60E-19	8.50E-17
	Nicotinate and nicotinamide metabolism	Metabolism of Cofactors and Vitamins	Metabolism	1.30E-19	4.40E-17
	One carbon pool by folate	Metabolism of Cofactors and Vitamins	Metabolism	8.80E-16	2.90E-13
	Pantothenate and CoA biosynthesis	Metabolism of Cofactors and Vitamins	Metabolism	1.00E-15	3.30E-13
	Porphyrin and chlorophyll metabolism	Metabolism of Cofactors and Vitamins	Metabolism	4.10E-19	1.30E-16
	Retinol metabolism	Metabolism of Cofactors and Vitamins	Metabolism	8.80E-18	2.90E-15
	Riboflavin metabolism	Metabolism of Cofactors and Vitamins	Metabolism	2.30E-21	7.60E-19
	Thiamine metabolism	Metabolism of Cofactors and Vitamins	Metabolism	1.40E-13	4.60E-11
	Ubiquinone and other terpenoid-quinone biosynthesis	Metabolism of Cofactors and Vitamins	Metabolism	3.10E-22	1.00E-19

Enriched Group	Pathway	Sub Pathway	Super Pathway	p-value	Bonferroni p-value
	Vitamin B6 metabolism	Metabolism of Cofactors and Vitamins	Metabolism	1.30E-17	4.40E-15
	beta-Alanine metabolism	Metabolism of Other Amino Acids	Metabolism	1.70E-18	5.70E-16
	Cyanoamino acid metabolism	Metabolism of Other Amino Acids	Metabolism	1.50E-18	4.80E-16
	D-Arginine and D-ornithine metabolism	Metabolism of Other Amino Acids	Metabolism	9.10E-07	3.00E-04
	D-Glutamine and D-glutamate metabolism	Metabolism of Other Amino Acids	Metabolism	5.40E-16	1.80E-13
	Glutathione metabolism	Metabolism of Other Amino Acids	Metabolism	3.00E-19	9.90E-17
	Phosphonate and phosphinate metabolism	Metabolism of Other Amino Acids	Metabolism	1.50E-22	4.90E-20
	Selenocompound metabolism	Metabolism of Other Amino Acids	Metabolism	2.00E-15	6.60E-13
	Taurine and hypotaurine metabolism	Metabolism of Other Amino Acids	Metabolism	5.70E-18	1.90E-15
	Biosynthesis of siderophore group nonribosomal peptides	Metabolism of Terpenoids and Polyketides	Metabolism	3.00E-19	9.80E-17
	Biosynthesis of type II polyketide products	Metabolism of Terpenoids and Polyketides	Metabolism	1.10E-16	3.50E-14
	Carotenoid biosynthesis	Metabolism of Terpenoids and Polyketides	Metabolism	1.40E-15	4.70E-13
	Geraniol degradation	Metabolism of Terpenoids and Polyketides	Metabolism	7.10E-18	2.30E-15
	Limonene and pinene degradation	Metabolism of Terpenoids and Polyketides	Metabolism	9.10E-19	3.00E-16
	Polyketide sugar unit biosynthesis	Metabolism of Terpenoids and Polyketides	Metabolism	6.20E-11	2.00E-08
	Prenyltransferases	Metabolism of Terpenoids and Polyketides	Metabolism	4.90E-10	1.60E-07
	Terpenoid backbone biosynthesis	Metabolism of Terpenoids and Polyketides	Metabolism	8.10E-14	2.60E-11

Enriched Group	Pathway	Sub Pathway	Super Pathway	p-value	Bonferroni p-value
	Tetracycline biosynthesis	Metabolism of Terpenoids and Polyketides	Metabolism	2.80E-24	9.10E-22
	Purine metabolism	Nucleotide Metabolism	Metabolism	9.50E-16	3.10E-13
	Aminobenzoate degradation	Xenobiotics Biodegradation and Metabolism	Metabolism	6.10E-19	2.00E-16
	Atrazine degradation	Xenobiotics Biodegradation and Metabolism	Metabolism	9.20E-19	3.00E-16
	Benzoate degradation	Xenobiotics Biodegradation and Metabolism	Metabolism	3.00E-19	9.70E-17
	Bisphenol degradation	Xenobiotics Biodegradation and Metabolism	Metabolism	4.30E-18	1.40E-15
	Caprolactam degradation	Xenobiotics Biodegradation and Metabolism	Metabolism	5.80E-19	1.90E-16
	Chloroalkane and chloroalkene degradation	Xenobiotics Biodegradation and Metabolism	Metabolism	7.10E-20	2.30E-17
	Chlorocyclohexane and chlorobenzene degradation	Xenobiotics Biodegradation and Metabolism	Metabolism	7.50E-18	2.50E-15
	Dioxin degradation	Xenobiotics Biodegradation and Metabolism	Metabolism	7.30E-18	2.40E-15
	Drug metabolism - cytochrome P450	Xenobiotics Biodegradation and Metabolism	Metabolism	1.10E-18	3.70E-16
	Fluorobenzoate degradation	Xenobiotics Biodegradation and Metabolism	Metabolism	2.10E-17	6.90E-15
	Metabolism of xenobiotics by cytochrome P450	Xenobiotics Biodegradation and Metabolism	Metabolism	7.80E-19	2.60E-16
	Naphthalene degradation	Xenobiotics Biodegradation and Metabolism	Metabolism	1.40E-19	4.50E-17
	Nitrotoluene degradation	Xenobiotics Biodegradation and Metabolism	Metabolism	6.30E-11	2.10E-08
	Polycyclic aromatic hydrocarbon degradation	Xenobiotics Biodegradation and Metabolism	Metabolism	1.10E-16	3.50E-14

Enriched Group	Pathway	Sub Pathway	Super Pathway	p-value	Bonferroni p-value
	Styrene degradation	Xenobiotics Biodegradation and Metabolism	Metabolism	6.70E-19	2.20E-16
	Toluene degradation	Xenobiotics Biodegradation and Metabolism	Metabolism	1.50E-17	4.90E-15
	Xylene degradation	Xenobiotics Biodegradation and Metabolism	Metabolism	2.00E-16	6.50E-14
	Cardiac muscle contraction	Circulatory System	Organismal Systems	3.90E-19	1.30E-16
	Mineral absorption	Digestive System	Organismal Systems	1.50E-12	4.90E-10
	Adipocytokine signaling pathway	Endocrine System	Organismal Systems	2.30E-16	7.40E-14
	GnRH signaling pathway	Endocrine System	Organismal Systems	3.80E-09	1.30E-06
	Insulin signaling pathway	Endocrine System	Organismal Systems	2.00E-14	6.50E-12
	Melanogenesis	Endocrine System	Organismal Systems	4.20E-08	1.40E-05
	PPAR signaling pathway	Endocrine System	Organismal Systems	1.10E-18	3.50E-16
	Progesterone-mediated oocyte maturation	Endocrine System	Organismal Systems	2.30E-15	7.70E-13
	Renin-angiotensin system	Endocrine System	Organismal Systems	1.60E-16	5.40E-14
	Circadian rhythm - plant	Environmental Adaptation	Organismal Systems	3.20E-15	1.10E-12
	Plant-pathogen interaction	Environmental Adaptation	Organismal Systems	1.30E-18	4.30E-16
	Proximal tubule bicarbonate reclamation	Excretory System	Organismal Systems	2.20E-17	7.30E-15
	Antigen processing and presentation	Immune System	Organismal Systems	2.30E-15	7.70E-13
	Fc gamma R-mediated phagocytosis	Immune System	Organismal Systems	3.80E-09	1.30E-06
	Glutamatergic synapse	Nervous System	Organismal Systems	2.30E-20	7.60E-18
	Cell division	Cellular Processes and Signaling	Unclassified	7.20E-20	2.40E-17
	Cell motility and secretion	Cellular Processes and Signaling	Unclassified	6.00E-19	2.00E-16

Enriched Group	Pathway	Sub Pathway	Super Pathway	p-value	Bonferroni p-value
	Inorganic ion transport and metabolism	Cellular Processes and Signaling	Unclassified	1.00E-18	3.30E-16
	Membrane and intracellular structural molecules	Cellular Processes and Signaling	Unclassified	4.40E-20	1.40E-17
	Other ion-coupled transporters	Cellular Processes and Signaling	Unclassified	5.10E-19	1.70E-16
	Other transporters	Cellular Processes and Signaling	Unclassified	1.20E-18	4.00E-16
	Pores ion channels	Cellular Processes and Signaling	Unclassified	2.80E-19	9.20E-17
	Signal transduction mechanisms	Cellular Processes and Signaling	Unclassified	9.20E-22	3.00E-19
	Protein folding, and associated processing	Genetic Information Processing	Unclassified	1.70E-18	5.50E-16
	Replication, recombination, and repair proteins	Genetic Information Processing	Unclassified	3.20E-17	1.00E-14
	Transcription related proteins	Genetic Information Processing	Unclassified	1.90E-21	6.10E-19
	Translation proteins	Genetic Information Processing	Unclassified	5.40E-17	1.80E-14
	Amino acid metabolism	Metabolism	Unclassified	2.00E-17	6.40E-15
	Biosynthesis and biodegradation of secondary metabolites	Metabolism	Unclassified	5.70E-19	1.90E-16
	Carbohydrate metabolism	Metabolism	Unclassified	5.10E-20	1.70E-17
	Energy metabolism	Metabolism	Unclassified	5.00E-19	1.70E-16
	Glycan biosynthesis and metabolism	Metabolism	Unclassified	2.70E-18	8.80E-16
	Lipid metabolism	Metabolism	Unclassified	9.70E-20	3.20E-17
	Metabolism of cofactors and vitamins	Metabolism	Unclassified	1.10E-18	3.60E-16
	Others	Metabolism	Unclassified	2.30E-20	7.50E-18
	Function unknown	Poorly Characterized	Unclassified	4.10E-19	1.30E-16
	General function prediction only	Poorly Characterized	Unclassified	2.30E-18	7.70E-16

Enriched Group	Pathway	Sub Pathway	Super Pathway	p-value	Bonferroni p-value
MCS2A	Phosphotransferase system (PTS)	Membrane Transport	Environmental Information Processing	4.40E-11	1.40E-08
	Ion channels	Signaling Molecules and Interaction	Environmental Information Processing	3.10E-14	1.00E-11
	RNA polymerase	Transcription	Genetic Information Processing	3.20E-08	1.00E-05
	RNA transport	Translation	Genetic Information Processing	1.50E-19	5.00E-17
	Bacterial invasion of epithelial cells	Infectious Diseases	Human Diseases	1.40E-05	4.40E-03
	Staphylococcus aureus infection	Infectious Diseases	Human Diseases	4.80E-18	1.60E-15
	Flavone and flavonol biosynthesis	Biosynthesis of Other Secondary Metabolites	Metabolism	4.10E-13	1.30E-10
	Galactose metabolism	Carbohydrate Metabolism	Metabolism	9.80E-10	3.20E-07
	D-Alanine metabolism	Metabolism of Other Amino Acids	Metabolism	5.60E-08	1.90E-05
	Ethylbenzene degradation	Xenobiotics Biodegradation and Metabolism	Metabolism	4.10E-17	1.30E-14
	Carbohydrate digestion and absorption	Digestive System	Organismal Systems	1.00E-11	3.40E-09
	RIG-I-like receptor signaling pathway	Immune System	Organismal Systems	1.10E-04	3.60E-02
	Sporulation	Cellular Processes and Signaling	Unclassified	2.80E-11	9.10E-09
	Nucleotide metabolism	Metabolism	Unclassified	1.50E-05	4.90E-03
MCS2B	Lysosome	Transport and Catabolism	Cellular Processes	1.00E-18	3.40E-16
	Amoebiasis	Infectious Diseases	Human Diseases	3.30E-11	1.10E-08
	Glycosaminoglycan degradation	Glycan Biosynthesis and Metabolism	Metabolism	6.20E-20	2.00E-17
	Glycosphingolipid biosynthesis - ganglio series	Glycan Biosynthesis and Metabolism	Metabolism	1.90E-20	6.30E-18

Enriched Group	Pathway	Sub Pathway	Super Pathway	p-value	Bonferroni p-value
	Glycosphingolipid biosynthesis - globo series	Glycan Biosynthesis and Metabolism	Metabolism	2.10E-20	6.80E-18
	N-Glycan biosynthesis	Glycan Biosynthesis and Metabolism	Metabolism	3.60E-18	1.20E-15
	Other glycan degradation	Glycan Biosynthesis and Metabolism	Metabolism	1.00E-16	3.40E-14
	Various types of N-glycan biosynthesis	Glycan Biosynthesis and Metabolism	Metabolism	2.50E-05	8.10E-03
	Sphingolipid metabolism	Lipid Metabolism	Metabolism	3.70E-17	1.20E-14
	Zeatin biosynthesis	Metabolism of Terpenoids and Polyketides	Metabolism	1.40E-18	4.50E-16
	Protein digestion and absorption	Digestive System	Organismal Systems	1.30E-20	4.40E-18
	NOD-like receptor signaling pathway	Immune System	Organismal Systems	2.20E-15	7.10E-13
	Restriction enzyme	Genetic Information Processing	Unclassified	4.80E-18	1.60E-15

Table S4. Metabolites enriched in each microbial state.

Enriched group	Metabolite	Super pathway	Sub pathway	p-value
MCS1	4-guanidinobutanoate	Amino Acid	Guanidino and Acetamido Metabolism	0.008
	cis-urocanate	Amino Acid	Histidine Metabolism	0.005
	trans-urocanate	Amino Acid	Histidine Metabolism	0.004
	xanthurenate	Amino Acid	Tryptophan Metabolism	0.048
	glucose	Carbohydrate	Glycolysis, Gluconeogenesis, and Pyruvate Metabolism	0.002
	1-methylnicotinamide	Cofactors and Vitamins	Nicotinate and Nicotinamide Metabolism	0.007
	biopterin	Cofactors and Vitamins	Tetrahydrobiopterin Metabolism	0.043
	15-HETE	Lipid	Eicosanoid	0.016
	leukotriene B4	Lipid	Eicosanoid	0.02
	eicosanodioate	Lipid	Fatty Acid, Dicarboxylate	0.048
	maleate (cis-Butenedioate)	Lipid	Fatty Acid, Dicarboxylate	0.044
	13-HODE + 9-HODE	Lipid	Fatty Acid, Monohydroxy	0.017
	scyllo-inositol	Lipid	Inositol Metabolism	0.038
	1-palmitoylglycerophosphate	Lipid	Lysolipid	0.043
	1-palmitoylglycerol (1-monopalmitin)	Lipid	Monoacylglycerol	0.038
	1-stearoylglycerol (1-monostearin)	Lipid	Monoacylglycerol	0.046
	2-palmitoylglycerol (2-monopalmitin)	Lipid	Monoacylglycerol	0.02
	glycerophosphoinositol	Lipid	Phospholipid Metabolism	0.016
	chenodeoxycholate	Lipid	Primary Bile Acid Metabolism	0.028
	glycochenodeoxycholate	Lipid	Primary Bile Acid Metabolism	0.022
	taurochenodeoxycholate	Lipid	Primary Bile Acid Metabolism	0.039
	glycocholenate sulfate	Lipid	Secondary Bile Acid Metabolism	0.017
	glycodeoxycholate	Lipid	Secondary Bile Acid Metabolism	0.041
	glycolithocholate sulfate	Lipid	Secondary Bile Acid Metabolism	0.021

Enriched group	Metabolite	Super pathway	Sub pathway	p-value
	glycoursodeoxycholate	Lipid	Secondary Bile Acid Metabolism	0.016
	taurocholenate sulfate	Lipid	Secondary Bile Acid Metabolism	0.05
	taurodeoxycholate	Lipid	Secondary Bile Acid Metabolism	0.013
	ursodeoxycholate	Lipid	Secondary Bile Acid Metabolism	0.03
	21-hydroxypregnenolone disulfate	Lipid	Steroid	0.012
	5alpha-pregnan-3(alpha or beta),20beta-diol disulfate	Lipid	Steroid	0.038
	pregnen-diol disulfate	Lipid	Steroid	0.04
	N6-methyladenosine	Nucleotide	Purine Metabolism, Adenine containing	0.031
	4-ureidobutyrate	Nucleotide	Pyrimidine Metabolism, Uracil containing	0.044
	glycylphenylalanine	Peptide	Dipeptide	0.03
	glycylvaline	Peptide	Dipeptide	0.039
	lysyltyrosine	Peptide	Dipeptide	0.006
	phenylalanylaspartate	Peptide	Dipeptide	0.005
	tyrosyllysine	Peptide	Dipeptide	0.002
	thymol sulfate	Xenobiotics	Food Component/Plant	0.017
	1-methylxanthine	Xenobiotics	Xanthine Metabolism	0.009
	7-methylxanthine	Xenobiotics	Xanthine Metabolism	0.037
	theobromine	Xenobiotics	Xanthine Metabolism	0.046
MCS2A	N-acetylphenylalanine	Amino Acid	Phenylalanine and Tyrosine Metabolism	0.022
	maltose	Carbohydrate	Glycogen Metabolism	0.035
	1-margaroylglycerophospho-ethanolamine	Lipid	Lysolipid	0.008

Enriched group	Metabolite	Super pathway	Sub pathway	p-value
	1-palmitoylglycerophospho-glycerol	Lipid	Lysolipid	0.024
	2-stearoylglycerophospho-ethanolamine	Lipid	Lysolipid	0.015
	5alpha-pregnan-3beta,20alpha-diol disulfate	Lipid	Steroid	0.023
	N4-acetylcytidine	Nucleotide	Pyrimidine Metabolism, Cytidine containing	0.028
MCS2B	3-methyl-2-oxobutyrate	Amino Acid	Leucine, Isoleucine and Valine Metabolism	0.016
	4-methyl-2-oxopentanoate	Amino Acid	Leucine, Isoleucine and Valine Metabolism	0.044
	methionine sulfone	Amino Acid	Methionine, Cysteine, SAM and Taurine Metabolism	0.043
	1-dihomo-linolenylglycerol (alpha, gamma)	Lipid	Monoacylglycerol	0.028
	1-myristoylglycerol (1-monomyristin)	Lipid	Monoacylglycerol	0.044
	inosine	Nucleotide	Purine Metabolism, (Hypo)Xanthine/Inosine containing	0.031
	phenylalanyltryptophan	Peptide	Dipeptide	0.047
MCS1 & MCS2A	phenylalanine	Amino Acid	Phenylalanine and Tyrosine Metabolism	0.024
	1-palmitoylglycerophospho-ethanolamine	Lipid	Lysolipid	0.048
MCS1 & MCS2B	glycolithocholate	Lipid	Secondary Bile Acid Metabolism	0.046
MCS2A & MCS2B	alpha-ketoglutarate	Energy	TCA Cycle	0.036

ACKNOWLEDGEMENTS

We acknowledge the patients and staff of Mulago Hospital for their contributions to the study. We thank the research team who enrolled the patients who participated in this study. We gratefully acknowledge Emily K. Cope and Kei E. Fujimura for assistance with sequencing analysis and figure design.

Chapter 3

Gut Microbiota is Related to Peripheral CD4 Counts, Lung Microbiota, and In Vitro Macrophage Dysfunction in HIV-Pneumonia Patients

Content in this chapter was modified from the following manuscript in review:

M.K. Shenoy, D.W. Fadrosh, D.L. Lin, W. Worodria, I. Ayakaka, P. Byanyima, S. Kaswabuli, M. McGing, M. Sommers, E. Chang, K. McCauley, L. Huang, and S.V. Lynch. Gut Microbiota is Related to Peripheral CD4 Counts, Lung Microbiota, and In Vitro Macrophage Dysfunction in HIV-Pneumonia Patients.

ABSTRACT

Bacterial pneumonia is a common and frequently fatal co-morbidity in HIV infected individuals in Africa. Prior investigations have described the lower airway microbiota in this population or the gut microbiota of HIV infected individuals; however, the relationship between the lower airway and gut microbiomes in HIV infected patients with bacterial pneumonia and their joint contribution to patient outcomes is unknown. We profiled paired BAL and stool bacterial and fungal communities from a large cohort (n=120) of Ugandan, HIV infected patients with acute pneumonia. We demonstrate that the lower airway microbiota of HIV-pneumonia patients stratify into distinct microbial community states that are significantly related to microbiological factors, but not to disease severity. In contrast, variation in gut microbiota significantly relates to patient mortality, disease severity as measured by CD4 count, and similarity between gut and lung microbial communities. We further establish that patients with low CD4 counts possess gut microbiota that are more similar to airway microbiota, are enriched for microbes shared with the airways, and elicit poorly activated, pro-inflammatory macrophages *in vitro*, when compared to gut microbiota from patients with high CD4 counts. These findings provide the first evidence that understanding and modulating the gut microbiota may be pivotal to improving outcomes in HIV infected patients with lower airway infections.

INTRODUCTION

Despite the efficacy of HAART in improving the health of HIV infected patients, pulmonary disease still presents a common co-morbidity within this population (*6, 7*). In HIV and TB co-endemic areas of Africa, the most common cause of inpatient hospital admission is pneumonia, resulting in permanent declines in pulmonary function and mortality rates between 5 and 30% despite antiretroviral and antibiotic treatment (*15-17*). Previous investigations of the HIV-associated pneumonia microbiome have focused on the airways, and have demonstrated that the lung microbiome of pneumonia patients differs based on HIV infection status and geography (*18, 19*). We recently identified three consistent, repeating patterns of microbial co-association within the lower airways of Ugandan HIV infected pneumonia patients, defined by the dominance of *Prevotellaceae*, *Streptococcaceae*, or *Pseudomonadaceae* and associated with local immune response, but only weakly related to patient outcomes (*20*). No study to date has demonstrated a consistent relationship between lung microbiota and circulating CD4 count, despite known shifts in lung microbiota of HIV infected patients with advanced disease (*11*).

In addition to changing the lower airway microbiome, HIV infection is known to shift gut microbiota, impair mucosal barrier function, and permit microbial translocation. A number of studies have established that HIV infection causes gut microbiota dysbiosis within United States populations, consistently characterized by increased *Prevotella* and *Proteobacteria* and decreased *Bacteroides* abundance (*14, 43, 44*). Vujkovic-Cvijin and colleagues showed that gut dysbiosis is only inconsistently and partially rescued with HAART treatment; while Dillon and colleagues demonstrated that gut dysbiosis is

associated with chronic gut mucosal inflammation. This microbial dysbiosis and associated inflammation causes increased barrier permeability and microbial translocation (*86, 87*). Overall, these studies have demonstrated that HIV infection results in consistent microbial dysbiosis and translocation, but investigations have yet to determine whether these large scale changes to the microbiome are related to HIV co-morbidities or disease severity. We hypothesized that the joint dysbiosis of gut and lung microbial communities and mucosal barrier breakdown would lead to increased microbial translocation between these two distal sites within HIV infected pneumonia patients. We further rationalized that this gut-lung microbial axis would influence immune response and patient outcomes. To investigate this hypothesis, we collected paired lower airway and stool samples from a large cohort of Ugandan, HIV infected patients with bacterial pneumonia, and asked whether gut and airway microbial community composition were related to one another, patient disease severity, and *in vitro* immune activation.

RESULTS

HIV infected pneumonia patients display bacterial anatomic site specificity

Paired BAL and stool samples were obtained from Ugandan HIV infected patients admitted to Mulago hospital for acute bacterial pneumonia (n=153). Study subjects possessed a median circulating CD4 count of 131 cells μl^{-1} and were administered antibiotics and antifungals (for full patient characteristics see Table S1). Bacterial profiles were generated using 16S V4 rRNA amplicon sequencing for lower airway and stool samples. Combined BAL and stool sample operational taxonomic unit (OTU) picking (representatively rarefied to 51,997 reads per sample) and PCoA (Bray Curtis dissimilarity) of 16S bacterial sequences revealed that BAL and stool samples from this severely immunocompromised population cluster based on sample type rather than patient sampled (Fig. S1A). OTU distribution modeling, as previously described (*88*), with validation by Random Forest was used to identify the most differential taxa between stool and BAL communities; taxa most enriched within the stool were traditional gut-associated microbes including *Bacteroides*, *Faecalibacterium*, and *Ruminococcaceae*, while BAL was enriched for airway-associated microbes including *Streptococcus, Veillonella,* and *Haemophilus* (Fig. S1B). These results demonstrate that despite severe illness, this HIV infected pneumonia study population retains overall microbial site specificity.

Lower airway microbiota stratify into distinct bacterial microbiota community states

To investigate patient lower airway microbial communities in greater depth, BAL bacterial 16S sequences were clustered into OTUs and representatively rarefied to 67,135 reads per sample for a final high-quality data set of n=117 samples. Lower airway bacterial microbiota displayed distinct compositional patterns [permutational multivariate analysis (PERMANOVA) with weighted UniFrac distance, R^2=0.658, p<0.001], with a gradient of *Prevotellaceae*- to *Streptococcaceae*-dominance over the majority of samples (n=95), or strong dominance by other airway pathogens consisting primarily of members of the *Gammaproteobacteria* (n=16; Fig. 1A). Delving into these patterns of microbiota demonstrated that *Prevotellaceae*- and *Streptococcaceae*-dominated communities were highly similar in composition; the two most abundant taxa in both groups were *Prevotellaceae* and *Streptococcaceae,* with consistent *Veillonellaceae* and *Paraprevotellaceae* co-colonization (Fig. 1B and 1C). *Gammaproteobacteria*-dominated samples were primarily characterized by *Pasteurellaceae*- (n=11) or *Pseudomonadaceae*- (n=3) dominance, appeared more compositionally variable by PCoA, and clustered distinctly from the majority of *Prevotellaceae* and *Streptococcaceae* dominated samples (Fig. 1B and 1C). Our data corroborate previous findings that HIV infected patients' lower airway microbiota stratify into specific clusters of microbial communities, and during HIV-associated bacterial pneumonia, lower airway communities are primarily dominated by a gradient of *Prevotellaceae* to *Streptococcaceae* or by members of the *Gammaproteobacteria* (*11, 20*).

To determine how our microbial community states compared to previous findings, we investigated their association with microbiological factors and disease severity. We

previously reported that *Prevotellaceae*- and *Streptococcaceae*-dominated lower airway samples were associated with higher bacterial diversity, administration of the first-line pneumonia antibiotic ceftriaxone, and the presence of culturable fungi. In our current study, *Prevotellaceae*- and *Streptococcaceae*-dominated samples possessed significantly higher Shannon's bacterial diversity than *Gammaproteobacteria*-dominated samples (Kruskal Wallis, $p<0.0001$; Fig. 1D). Additionally, there were far fewer non-*Prevotellaceae*- nor *Streptococcaceae*-dominated samples in this cohort compared to our previous study (19% versus 42% of cohort), which we attribute to the near universal administration of ceftriaxone (91%) in this cohort.

To address whether these bacterial community states differentially associated with fungi, we amplified and sequenced ITS2 rRNA from our BAL samples. BAL fungal community profiles (n=26) were representatively rarefied to 1044 reads per sample; the low number of samples with fungal amplification and rarefying depth demonstrate that this patient population is truly a bacterial, rather than fungal, pneumonia population. Fungal communities were primarily and highly dominated by *Candida* (n=19/26), and dominant fungal taxa was associated with variation in fungal composition (Bray Curtis PERMANOVA, $R^2=0.673$, $p<0.001$; Fig. S2A). CD4 count was also related to fungal composition (Bray Curtis PERMANOVA, $R^2=0.220$, $p<0.01$; Fig. S2B); this association was likely driven by the *Pneumocystis jiroveci*-dominated samples (n=3), all possessing low CD4 counts (cells/μl < 34). *Pneumocystis jiroveci* is a common cause of fungal pneumonia, so we assessed how each bacterial community state associated with fungal community profiles (Fig. S2C). While most of our BAL samples produced robust 16S amplification, *Pneumocystis*-dominated samples were among those that did not produce

a 16S amplicon, reaffirming that the patients with sequenced 16S bacterial amplicons were truly bacterial pneumonia patients. There was no association between lower airway fungal and bacterial composition, which may have been due to the high anti-fungal prophylaxis administration (83%); however, there was a pattern of increased fungal co-colonization with *Streptococcaceae*- and *Prevotellaceae*-dominated communities compared with only a single *Gammaproteobacteria*-dominated community with fungal co-colonization, though due to small fungal sample size this association was not significant (chi-square test, p=0.46; Fig. 1E). This pattern is corroborated by previous findings (*20*), but larger cohorts would be needed for validation.

Finally, we assessed whether lower airway bacterial community states were related to systemic patient health. CD4 count and gut microbial dysbiosis are known hallmarks of advanced HIV infection. Despite microbiological differences, circulating CD4 count (cells μl^{-1}) was not associated with microbiota community state (Kruskal Wallis, p=0.85; Fig. 1F). As a measure of how similar paired BAL and stool microbiota communities were within the same patient, we calculated paired distance using Bray Curtis dissimilarity, which weights higher abundance, rather than phylogenetically related (UniFrac) or lower abundance (Canberra) taxa. This distance allows assessment of the relative similarity of lower airway and gut microbial communities across the patient population. Using this metric, we determined that lower airway community states did not differ in their similarity to their paired stool samples (Kruskal Wallis, p=0.62; Fig. 1G), thus variation in pneumonia lower airway bacterial community is not related to systemic markers of patient disease severity nor to gut microbiota composition. Since lower airway bacterial community composition was not a good indicator of systemic

disease, we next asked whether gut microbiota would be a better indicator of disease severity and patient outcome.

Figure 1.

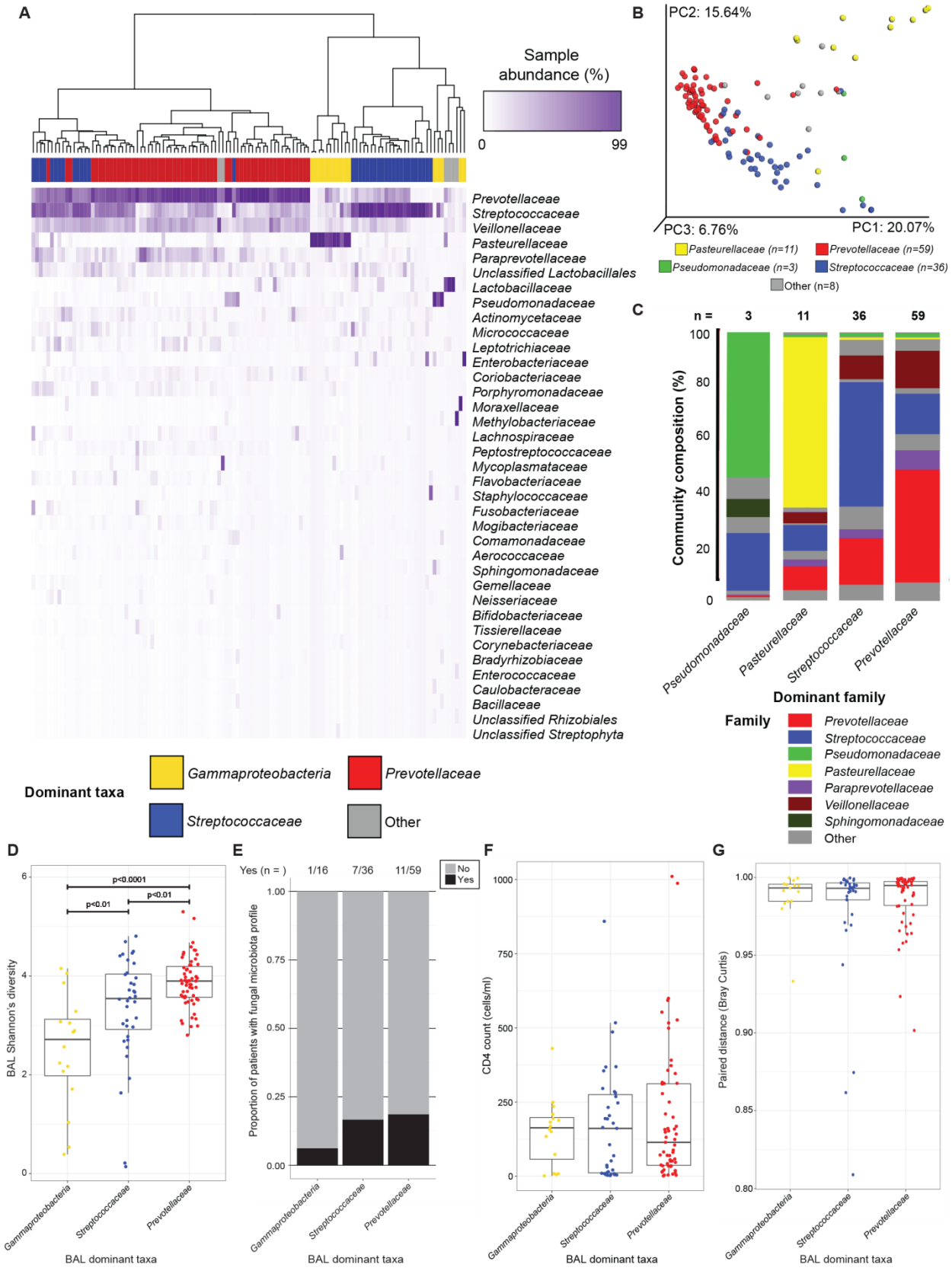

A
Sample abundance (%)
0
99
Prevotellaceae
Streptococcaceae
Veillonellaceae
Pasteurellaceae
Paraprevotellaceae
Unclassified Lactobacillales
Lactobacillaceae
Pseudomonadaceae
Actinomycetaceae
Micrococcaceae
Leptotrichiaceae
Enterobacteriaceae
Coriobacteriaceae
Porphyromonadaceae
Moraxellaceae
Methylobacteriaceae
Lachnospiraceae
Peptostreptococcaceae
Mycoplasmataceae
Flavobacteriaceae
Staphylococcaceae
Fusobacteriaceae
Mogibacteriaceae
Comamonadaceae
Aerococcaceae
Sphingomonadaceae
Gemellaceae
Neisseriaceae
Bifidobacteriaceae
Tissierellaceae
Corynebacteriaceae
Bradyrhizobiaceae
Enterococcaceae
Caulobacteraceae
Bacillaceae
Unclassified Rhizobiales
Unclassified Streptophyta
B
PC2: 15.64%
PC3: 6.76%
PC1: 20.07%
Pasteurellaceae (n=11)
Prevotellaceae (n=59)
Pseudomonadaceae (n=3)
Streptococcaceae (n=36)
Other (n=8)
C
n = 3 11 36 59
Community composition (%)
Pseudomonadaceae
Pasteurellaceae
Streptococcaceae
Prevotellaceae
Dominant family
Family
Prevotellaceae
Streptococcaceae
Pseudomonadaceae
Pasteurellaceae
Paraprevotellaceae
Veillonellaceae
Sphingomonadaceae
Other
Dominant taxa
Gammaproteobacteria
Prevotellaceae
Streptococcaceae
Other
D
p<0.0001
p<0.01
p<0.01
BAL Shannon's diversity
Gammaproteobacteria
Streptococcaceae
Prevotellaceae
BAL dominant taxa
E
Yes (n =) 1/16 7/36 11/59
No
Yes
Proportion of patients with fungal microbiota profile
BAL dominant taxa
F
CD4 count (cells/ml)
BAL dominant taxa
G
Paired distance (Bray Curtis)
BAL dominant taxa

Figure 1. Distinct bacterial community states are found in the lower airways of HIV infected pneumonia patients and relate to microbiological factors, but are not a good indicator of systemic immune response or gut microbiota. A. Unsupervised hierarchical clustering (using Bray Curtis dissimilarity (BC) and Ward 2 clustering) of BAL samples and abundance heat map of bacterial families present in at least 1 sample at ≥ 3% relative abundance (ordered from highest to lowest abundance) indicate patterns of microbial colonization in BAL. **B.** PCoA of weighted UniFrac distance for n = 117 BAL samples representatively rarefied to 67,135 reads/sample illustrates patterns of bacterial family dominance, which significantly explain variation in lower airway bacterial communities (PERMANOVA, $R^2 = 0.658$, $p < 0.001$). **C.** Mean community composition at the family level summarized by dominant family microbial community state. Bacterial family dominance is related to **D.** BAL Shannon's diversity (KW, $p<0.0001$; Mann-Whitney U Test p-values plotted) and potentially to **E.** fungal co-colonization (Chi-square, p=0.46), but not to **F.** CD4 count (cells/μl; KW p=0.85), or to **G.** similarity to paired stool sample (paired distance measured by BC; KW, p=0.62). PC = principal coordinate; PERMANOVA = permutational multivariate ANOVA; KW = Kruskal Wallis rank sum test.

Stool microbiota composition is related to clinical and immunological factors

To investigate patient gut microbial communities in greater depth, stool bacterial 16S sequences were clustered at 97% sequence identity into OTUs and rarefied to 120,665 reads per sample (n=106). Stool bacterial community composition was generally characterized by high relative abundance of *Ruminococcaceae*, *Bacteroidaceae*, *Enterococcaceae Enterobacteriaceae*, or *Prevotellaceae* with *Ruminococcaceae* (Fig. 2A). PERMANOVA analysis of stool samples revealed that variation in community composition was related to all clinical, demographic and microbiological factors, but not to how the samples were processed (Table 1). Of these factors, dominant family explained the most variation in community composition (Bray Curtis PERMANOVA, R^2=0.319, p<0.001; Fig. 2B) and was related to the percent dominance by the dominant family (Kruskal Wallis, p<0.001), suggesting that dominant bacterial families within gut microbiota are related to overall bacterial community composition.

Stool fungal microbiota composition was investigated by amplifying and sequencing ITS2 rRNA from stool. Fungal microbiota profiles (n=90) were representatively rarefied to 2,565 reads per sample. Stool fungal community composition was related to dominant family (Bray Curtis PERMANOVA, R^2=0.442, p<0.001; Fig. S3A), and, similar to BAL fungal communities, was frequently and highly dominated by *Candida* (72/90 samples with 94% mean dominance; Table S2). Stool fungal communities were not related to patient factors, bacterial microbiota, or paired airway fungal community composition (data not shown), and when stool and BAL were jointly clustered into OTUs, samples did not cluster based on anatomic site as bacterial

communities did (Fig. S3B). Since stool fungal communities were highly homogenous, we focused on understanding variation in the stool bacterial microbiota.

To better understand patient disease, variation in stool bacterial composition and microbiological characteristics were examined in relation to patient outcomes. While mortality was low within this cohort (6%), stool bacterial composition was related to mortality at time of hospital discharge (Bray Curtis PERMANOVA, R^2=0.018, p=0.017; Fig. 2C), and stool Shannon's diversity was significantly lower for in patients who went on to be deceased compared to survivors (Mann-Whitney, p<0.01; Fig. 2D). This supports the traditional premise that lower stool bacterial diversity is a sign of poor health. While stool microbiota was related to patient mortality, the ratio of deceased to non-deceased patients was too small to adequately interrogate this relationship. As such, we decided to examine circulating CD4 count as it is widely accepted as a surrogate for HIV-associated disease severity. This holds true within our HIV infected pneumonia population, with deceased patients having significantly lower CD4 counts (cells μl^{-1}) compared to survivors (Mann-Whitney, p<0.01; Fig. 2E). The cohort's mean CD4 count was 186, median was 131, and range was 1-1010 cells μl^{-1} (Fig. S4A). CD4 count differed significantly based on dominant family, with the lowest CD4 counts in *Enterobacteriaceae* dominated samples (Kruskal Wallis, p=0.02; Fig. S4B), and was significantly associated with variation in overall bacterial composition (Bray Curtis PERMANOVA, R^2=0.017, p=0.025; Fig. S4C). Since the CD4 distribution within the study population was severely skewed by high CD4 counts, the data were grouped by quartile into CD4 low (cells μl^{-1} < 34, quartile 1), CD4 intermediate (34 < cells μl^{-1} < 293, quartiles 2 and 3), and CD4 high (cells μl^{-1} > 293, quartile 4) patients to better

understand the extremes of patient disease within this population and detect consistent patterns in gut microbiota based on CD4 count. As with continuous CD4 count, CD4 groups were significantly associated with stool bacterial composition (Bray Curtis PERMANOVA, R^2=0.052, p=0.002; Fig. S4D) and the increase in R^2 (3.5%) gained by grouping CD4 count illustrates that consistent compositional patterns are associated with low versus high CD4 counts. Additionally, patients with high CD4 counts possessed significantly higher Faith's phylogenetic diversity than patients with low CD4 counts (Mann-Whitney, p<0.005; Fig. 2F), confirming that stool bacterial microbiota are related to CD4 count and disease severity.

Since stool bacterial microbiota was related to patient outcomes and disease severity, we next asked whether it was also related to airway composition. To investigate this, the similarity between paired BAL and stool bacterial microbiota was examined using Bray Curtis paired distance. The distribution of paired distances was skewed towards lower distances, with mean distance 0.981, median 0.994, and range 0.810-0.999 (Fig. S5A). In contrast to airway microbiota, variation in stool composition was significantly related to paired distance (Bray Curtis PERMANOVA, R^2=0.026, p=0.007; Fig. S5B), indicating that increased similarity between patient stool and airway microbiota is determined by variation in stool, rather than airway, bacterial composition. Not only was paired distance related to stool composition, it was also related to mortality at hospital discharge, with deceased patients possessing more similar airway and gut microbiota (Mann-Whitney, p=0.01; Fig. 2G). To better understand the relationship between paired distance and stool composition, patients were stratified based on paired distance quartiles into low [paired distance (PD) < 0.9816, quartile 1], intermediate

(0.9815 < PD < 0.9972, quartiles 2 and 3), and high (PD > 0.9971, quartile 4) PD. PD groups significantly explained variation in stool bacterial composition (PERMANVOA with Bray Curtis dissimilarity, R^2=0.054, p=0.002; Fig. S5C), and patients with low PD possessed significantly lower Faith's phylogenetic diversity than those with high PD (Mann-Whitney, p=0.03, Fig. 2H). Decreased gut diversity has previously been associated with gastrointestinal disease and dysbiosis, meaning that within this cohort, patients with more severe gut dysbiosis possess bacterial communities that more closely resemble the airways. Finally, administration of the antibiotic ceftriaxone was related to stool composition (Bray Curtis PERMANOVA, R^2=0.019, p=0.01; Fig. S6A) and associated with more dominated stool bacterial microbiota (Mann-Whitney, p=0.08; Fig. S6B). Ceftriaxone has previously been shown to relate to airway microbiota composition in the HIV infected pneumonia population, suggesting that similarity between lower airway and gut microbiota may in part be due to antibiotic administration in addition to mucosal barrier breakdown.

Table 1. Clinical and microbiological features are significantly associated with variation in stool microbial composition in HIV infected patients with bacterial pneumonia.

Category	Factor	PERMANOVA[a]	
Clinical		R^2	p-value
	CD4 count (grouped quartile 1 versus 4)	0.052	0.002
	Ceftriaxone at time of sampling	0.019	0.014
	Mortality at hospital discharge	0.018	0.017
	CD4 count (cells/μl)	0.017	0.025
	Pulmonary tuberculosis	0.009	0.499
	Ceftriaxone within 3 months prior to hospitalization	0.008	0.634
Microbiological			
	Dominant family	0.319	0.001
	Shannon diversity	0.145	0.001
	Paired BAL dominant family	0.122	0.033
	Total observed species[b]	0.111	0.001
	Faith's phylogenetic diversity	0.104	0.001
	Percent dominance	0.072	0.001
	Bray Curtis distance to paired BAL sample (grouped quartile 1 versus 4)[c]	0.054	0.002
	Bray Curtis distance to paired BAL sample (continuous)	0.026	0.007
	Faith's phylogenetic diversity of paired BAL sample[d]	0.021	0.028
	Fungal percent dominance	0.022	0.1
	Fungal dominant genus	0.095	0.284
Processing			
	Sequenced on NextSeq run 1	0.012	0.132
	Sequenced on NextSeq run 2	0.012	0.132
	ITS2 fungal amplicon sequenced	0.013	0.162
	16S amplification plate	0.022	0.221
	16S primer plate	0.022	0.221
	Extraction plate	0.007	0.824

[a]Permutational multivariate ANOVA results for Bray Curtis distance shown; representative of weighted and unweighted UniFrac and Canberra distances.
[b]Similar results for chao1 richness estimate
[c]Similar results for weighted and unweighted UniFrac and Canberra distances.
[d]Stool composition is not significantly related to BAL richness or Shannon diversity.

Figure 2.

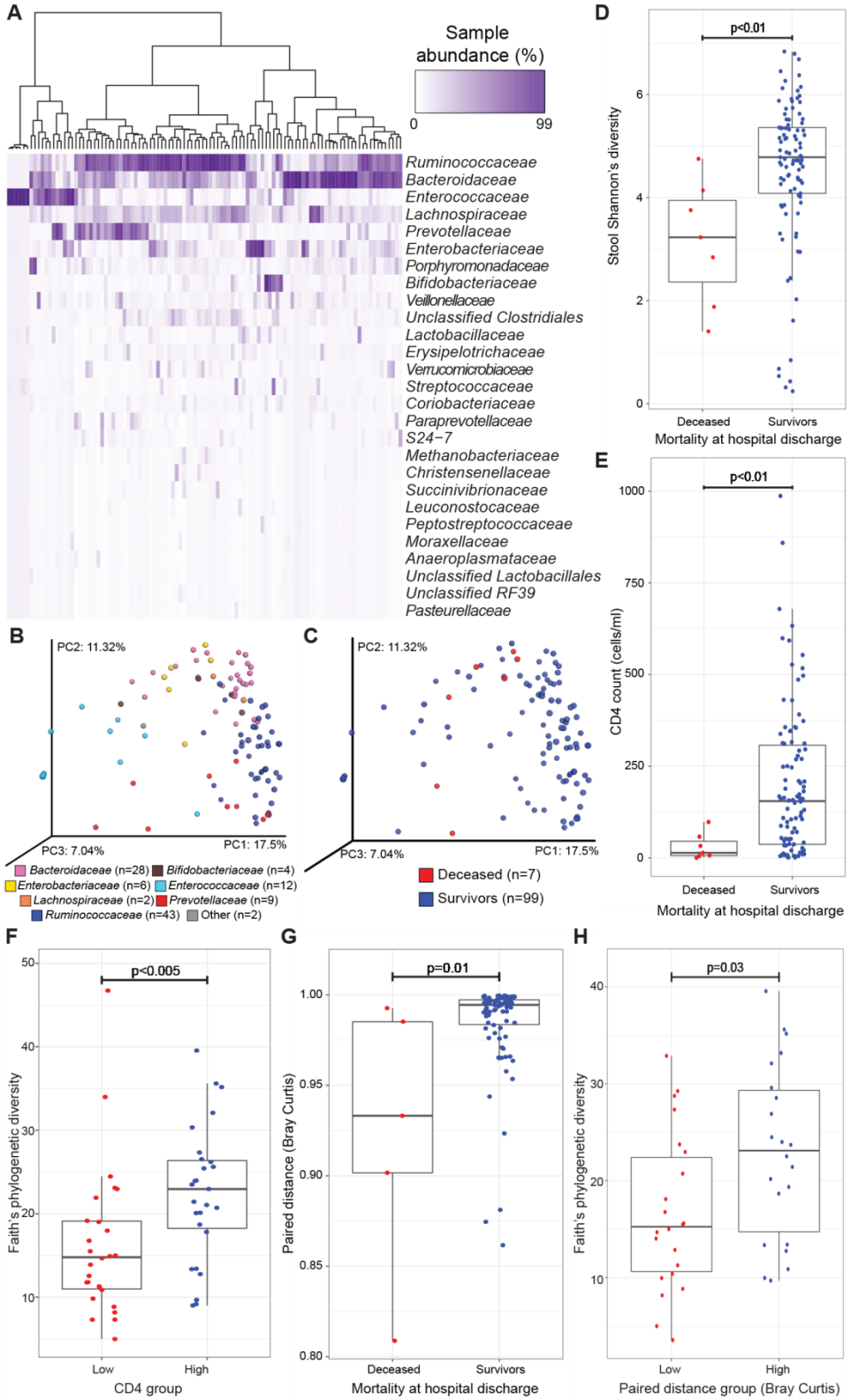

A
Sample abundance (%)
0
99
Ruminococcaceae
Bacteroidaceae
Enterococcaceae
Lachnospiraceae
Prevotellaceae
Enterobacteriaceae
Porphyromonadaceae
Bifidobacteriaceae
Veillonellaceae
Unclassified Clostridiales
Lactobacillaceae
Erysipelotrichaceae
Verrucomicrobiaceae
Streptococcaceae
Coriobacteriaceae
Paraprevotellaceae
S24−7
Methanobacteriaceae
Christensenellaceae
Succinivibrionaceae
Leuconostocaceae
Peptostreptococcaceae
Moraxellaceae
Anaeroplasmataceae
Unclassified Lactobacillales
Unclassified RF39
Pasteurellaceae
B
PC2: 11.32%
PC3: 7.04%
PC1: 17.5%
Bacteroidaceae (n=28)
Bifidobacteriaceae (n=4)
Enterobacteriaceae (n=6)
Enterococcaceae (n=12)
Lachnospiraceae (n=2)
Prevotellaceae (n=9)
Ruminococcaceae (n=43)
Other (n=2)
C
PC2: 11.32%
PC3: 7.04%
PC1: 17.5%
Deceased (n=7)
Survivors (n=99)
D
p<0.01
Stool Shannon's diversity
Deceased
Survivors
Mortality at hospital discharge
E
p<0.01
CD4 count (cells/ml)
Deceased
Survivors
Mortality at hospital discharge
F
p<0.005
Faith's phylogenetic diversity
Low
High
CD4 group
G
p=0.01
Paired distance (Bray Curtis)
Deceased
Survivors
Mortality at hospital discharge
H
p=0.03
Faith's phylogenetic diversity
Low
High
Paired distance group (Bray Curtis)

Figure 2. Stool microbiota composition is related to clinical and immunological factors and is a better indicator of HIV infected pneumonia patient health than lower airway microbiota. A. Unsupervised hierarchical clustering (using Bray Curtis dissimilarity (BC) and Ward 2 clustering) and abundance heat map of bacterial families present in at least 1 sample at ≥ 3% relative abundance (ordered from highest to lowest abundance) indicate patterns of microbial colonization in stool. PCoA of BC dissimilarity for n = 106 stool samples representatively rarefied to 120,665 reads/sample illustrates that **B.** dominant bacterial family [PERMANOVA, $R^2 = 0.319$, $p < 0.001$; Other = *Streptococcaceae* (covered by *Enterococcaceae* on left) and *Porphyromonadaceae*] and **C.** mortality outcome at hospital discharge (PERMANOVA, $R^2 = 0.018$, $p = 0.017$) significantly explain variation in stool bacterial microbiota. Mortality at hospital discharge is significantly associated with **D.** stool Shannon's diversity (Mann-Whitney U Test, $p<0.01$) and **E.** CD4 count (cells/μl; MW, $p<0.01$). **F.** CD4 count grouped (by quartile 1 versus 4) is significantly related to stool Faith's phylogenetic diversity. Similarity between lower airway and stool samples (Bray Curtis Paired Distance) is significantly related to **G.** mortality at hospital discharge (MW, $p=0.01$), and **H.** when grouped (by quartile 1 versus 4), to stool Faith's phylogenetic diversity (MW, $p<0.03$). PC = principal coordinate; PERMANOVA = permutational multivariate ANOVA; MW = Mann-Whitney U Test.

Gut microbiota of HIV-pneumonia patients with low CD4 counts are enriched for shared airway microbes and depleted of healthy gut microbes

To validate whether CD4 count is directly related to similarity between the airway and gut microbiota, the proportion of patients with low and high paired distances was compared between patients with low and high CD4 counts. Patients with low CD4 counts were significantly more likely to have low PD, while patients with high CD4 counts were more likely to have high PD (Fisher's exact test, $p<0.001$; Fig. 3A). OTU distribution modeling was used to understand how CD4 low and CD4 high patient stool microbiota differ and why this might be related to paired distance (Fig. 3B), and was validated with Random Forest (Fig. S7A). Traditional airway-associated microbes including *Streptococcus, Veillonella, unclassified Lactobacillales,* and *Megasphaera* were the top enriched taxa in the stool of patients with low CD4 counts, while *Prevotella copri*, *Faecalibacterium prausnitzii*, *unclassified Ruminococcaceae, Bacteroides,* and *Ruminococcus* were the top enriched taxa in patients with high CD4 counts. *Prevotella* copri, *Faecalibacterium prausnitzii*, *Ruminococcaceae,* and *Bacteroides* are traditional, healthy constituents of the gut microbiota, which would explain why they are enriched in the stool of healthier, CD4 high patients and depleted as CD4 count decreases. Interestingly, identical *Streptococcus* and *Megasphaera* OTU sequences to those enriched in stool of CD4 low patients were also found in patient BAL using independent OTU clustering, suggesting that these microbes are translocating between the airways and the gut.

Supporting the idea of microbial translocation between the airways and gut, patients with low paired distances were highly enriched for *Streptococcus*,

Lactobacillus, and *unclassified Enterobacteriaceae* within the gut that were shared with airways, while high PD was primarily distinguished by enrichment of *Prevotella copri* within the gut (Fig. 3C and Fig. S7B). As a final validation of microbial translocation, the most highly enriched OTUs were compared between PD and CD4 groups (Fig. 3D). A number of CD4 and PD low stool enriched taxa overlapped, almost all of which were *Streptococcus* or *Veillonella* OTUs. Two *Prevotella copri* OTUs overlapped between CD4 and PD high stool samples, while no enriched OTUs overlapped between CD4 high and PD low stool, nor between CD4 low and PD high stool, demonstrating that low CD4 count in HIV-pneumonia patients is strongly associated with enrichment of microbes within the gut that are shared with and potentially translocated from the airways. These same analyses within the lower airways yielded far fewer taxa, though *Streptococcus* was enriched within CD4 low airways (Fig. S8A), *Lactobacillus* was enriched within PD low airways (Fig. S8B), and *unclassified Enterobacteriaceae* was mutually enriched within CD4 and PD low airways (Fig. S8C). While gut microbiota varied greatly based on patient disease severity, variation in airway microbiota did not strongly relate to CD4 count or gut microbiota similarity, suggesting that the gut microbiota may be a better indicator of patient outcomes in this population. Overall, these data demonstrate that HIV infected pneumonia patients with severe disease experience gut bacterial dysbiosis resulting in depletion of healthy gut microbes and enrichment of microbes shared with pneumonia patient airways.

Figure 3.

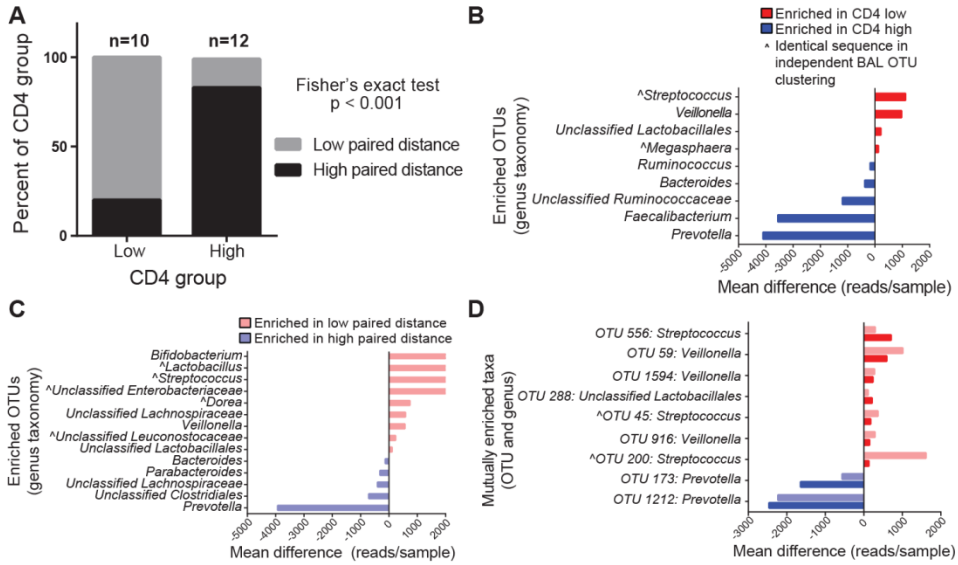


Figure 3. Patients with low CD4 counts possess stool microbiota that are more similar to their paired airway samples, are enriched for microbes shared with their lower airways, and are depleted for members of the healthy gut microbiota. A. Patients with low CD4 counts are significantly more likely to have low paired distance (quartile 1 versus 4 for each factor, Fisher's Exact Test, $p<0.001$). Mean difference in reads/sample are plotted for the top taxa (OTUs summarized to genus taxonomy) differentiating **B.** CD4 low and high, as well as **C.** paired distance low and high, stool bacterial communities, demonstrating enrichment of shared airway taxa and depletion of healthy-associated gut microbes in CD4 and paired distance low patients. **D.** Mean difference in reads/sample are plotted for top differential taxa shared between CD4 and paired distance enrichment comparisons. All data plotted were filtered based on $q < 0.05$, read difference ≥ 100, and presence in $\geq 50\%$ of the enriched group; no OTUs with the same genus level taxonomy assignment were significantly enriched in two different directions for any comparison. OTU = operational taxonomic unit; ^ = identical 16S sequence shared with independent BAL sample OTU picking.

Gut microbial products differentially modulate healthy macrophage phenotype based on patient CD4 count

Macrophage dysfunction is a hallmark of advanced HIV infection and characterized by decreased circulating monocytes and increased pro-inflammatory macrophages within the gut, but the cause of this aberrant behavior is unknown (*89*). Peripheral monocyte-derived macrophages are necessary for controlling lower airway infections, and resolving inflammation in pneumonia, yet in HIV infected individuals, pneumonia is not only more common, but it is also frequently fatal due to rampant, systemic inflammation. We hypothesized that the gut microbiota, which is a known modulator of systemic metabolism and airway immunity (*33, 90, 91*), contributes to macrophage dysfunction within HIV infected individuals and relates to patient disease severity. To investigate this, the THP-1 macrophage cell line was differentiated for two days with phorbol 12-myristate 13-acetate (PMA) prior to treatment with patient sterile fecal water (SFW) for 24 hours; SFW is made by homogenizing stool in buffer and removing all live cells, in order to treat macrophages with sterile fecal microbial metabolites and ligands. Following SFW treatment, THP-1 cells were assessed via flow cytometry for differentiation (CD14+CD68+), activation (CD80+CD86+), pro-inflammatory (IL-1β+), and tissue repair (CD206+IL-10+) markers. Treatment with SFW from CD4 or PD low patients resulted in significantly fewer activated macrophages compared to CD4 or PD high SFW (Mann-Whitney; CD4 $p=0.047$, Fig. 4A; PD $p<0.01$, Fig. 4B), though total macrophages did not differ between treatments (data not shown). CD4 or PD high SFW induced significantly fewer activated (CD4 $p<0.0001$, Fig. 4C; PD $p<0.0001$, Fig. 4D) and total (CD4 $p<0.001$, Fig. S9A; PD $p<0.001$, Fig. S9B) IL-1β+

pro-inflammatory macrophages compared to CD4 and PD low samples, and conversely, significantly more activated (CD4 $p<0.01$, Fig. 4E; PD $p<0.001$, Fig. 4F) and total (CD4 $p<0.03$, Fig. S9C; PD $p<0.0001$, Fig. S9D) CD206+IL-10+ tissue repair macrophages. Metabolites and ligands within stool, the majority of which are microbially derived, differentially modified activation and differentiation of a healthy macrophage cell line based on patient disease severity and degree of microbial translocation, indicating that changes to the gut microbiota associated with severe disease in HIV infected pneumonia patients could have meaningful consequences for macrophage dysfunction and pneumonia patient outcomes.

Figure 4.

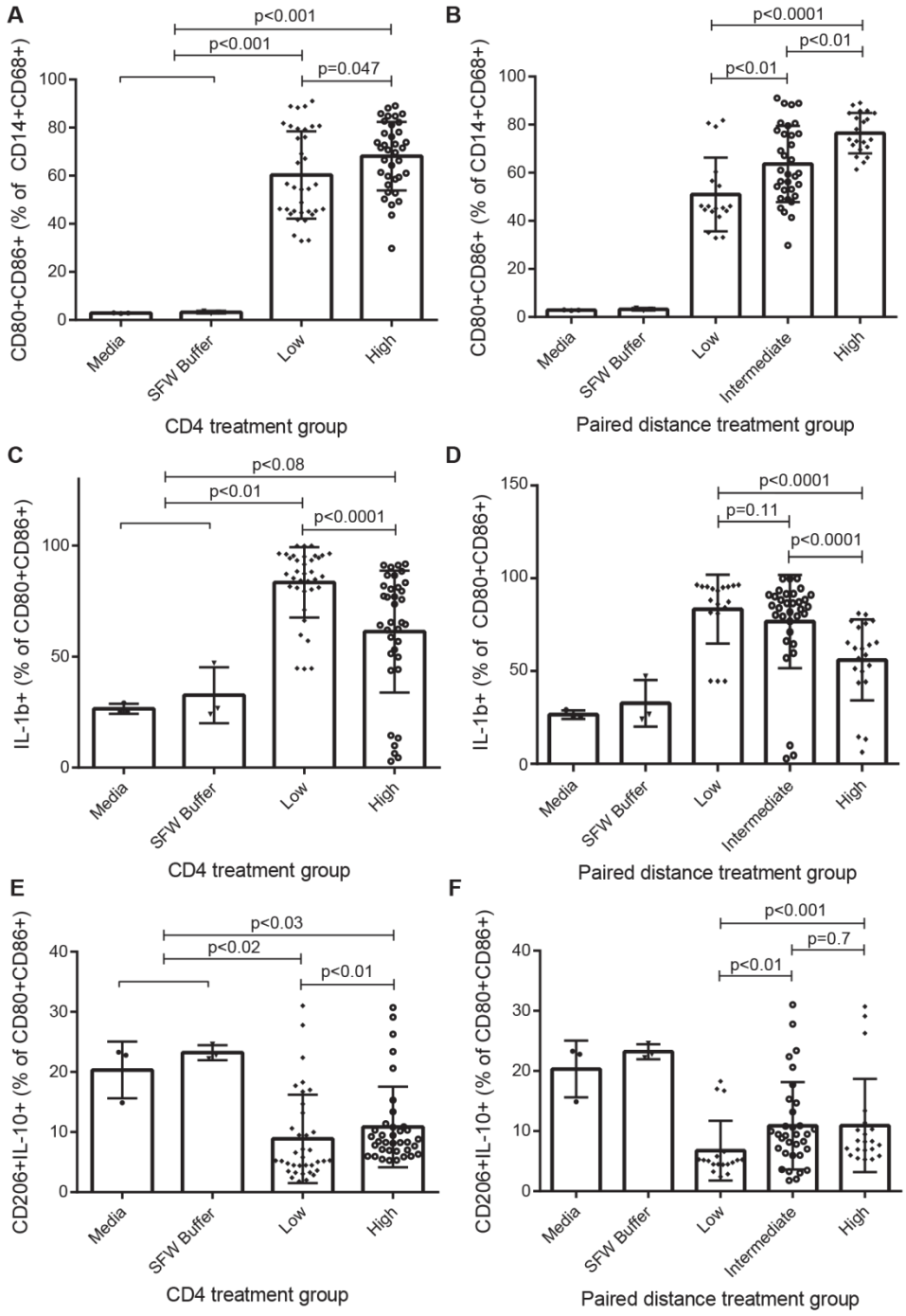


Figure 4. Sterile microbial stool metabolites and ligands from patients with low CD4 counts or paired distance induce increased pro-inflammatory and decreased tissue repair macrophage differentiation. Fewer activated CD80+ CD86+ THP-1-

derived macrophages are induced more frequently by **A.** CD4 high compared to CD4 low (KW, $p<0.001$), or **B.** paired distance high compared to paired distance low or medium (KW, $p<0.0001$; all treatments versus media or buffer $p<0.001$), patient SFW. Increased pro-inflammatory IL-1β+, activated THP-1-derived macrophages are induced by **C.** CD4 low compared to CD4 high (KW, $p<0.0001$), or **D.** paired distance low or medium compared to paired distance high (KW, $p<0.0001$; low or intermediate versus media or buffer $p<0.01$, high versus media or buffer $p<0.08$), patient SFW. Decreased tissue repair associated CD206+ IL-10+, activated THP-1-derived macrophages are induced by **E.** CD4 low compared to CD4 high (KW, $p<0.001$), or **F.** paired distance low compared to paired distance high or medium (KW, $p<0.001$; low versus media or buffer $p<0.05$, intermediate or high versus media or buffer $p<0.06$), patient SFW. Low and high groups are quartiles 1 and 4 respectively, with paired distance quartile 2 and 3 combined into an intermediate group. Data representative of 2 independent experiments; 24 biological replicates with 3 technical replicates each are plotted. Plotted p-values from Mann-Whitney test. KW = Kruskal Wallis; SFW = sterile fecal water.

DISCUSSION

The factors that determine disease severity and outcome in HIV infected patients with bacterial pneumonia are poorly understood. Previous work has focused on the role of the lower airway microbiome, and this is the first study to date to investigate the airway-gut microbial axis in this population. We hypothesized that the activities of the gut microbiota and their relationship with the airway microbiota are key to understanding pneumonia patient outcomes in the HIV infected population. This study demonstrates that HIV infected pneumonia patients with low CD4 counts have both increased lower airway to gut microbial translocation and community similarity, and the collective products of these gut microbial communities induce *in vitro* macrophage dysfunction that has previously been observed within this population. Previous work has demonstrated that lower airway microbial communities are related to local but not systemic immune response in HIV infected patients with and without pneumonia (*20, 40*). Here we elucidate a relationship between gut microbiota composition and systemic immune response and describe a potential mode by which gut microbiota may modulate distal pneumonia inflammation and patient outcomes.

Previous investigations of the HIV-pneumonia lower airway microbiome have described distinct microbiota community states dominated by *Prevotellaceae, Streptococcaceae*, or *Pseudomonadaceae*, a member of the *Gammaproteobacteria (20)*. Here we confirm the finding of distinct microbiota states characterized by dominance of *Prevotellaceae*, *Streptococcaceae*, or *Gammaproteobacteria*. *Prevotellaceae* and *Streptococcaceae* dominated communities possessed similar

microbes at varying abundances, and were significantly more diverse and had more fungal co-colonization compared to *Gammaproteobacteria* dominated communities.

This study built on prior lower airway microbiota framework, to investigate the airway-gut microbial axis in relation to disease severity. This study is the first to interrogate paired lung and gut microbial communities in the HIV infected pneumonia population and shows a direct relationship between overall stool microbial composition and circulating CD4 count. We attribute the ability to detect this overall shift in composition and patterns of consistently enriched or depleted taxa to the large cohort and rigorous quality filtering employed in this study. Here we report that *Prevotella copri* is enriched in patients with higher CD4 counts and less severe gut dysbiosis, while previous studies have reported an increased ratio of *Prevotella copri* to *Bacteroides* in HIV infected patients (*14, 43, 44*). It is important to note that these studies were conducted solely within the United States, a population that consumes a high protein, high fat diet which is associated with a low *Prevotella* to *Bacteroides* ratio in healthy individuals, whereas Asian and African populations with diets rich in starch or vegetables have a higher ratio of *Prevotella* to *Bacteroides* in healthy individuals (*45, 92, 93*). Our study demonstrated increased *Prevotella copri* in healthier Ugandan, HIV infected patients, revealing that geography may also play a role in understanding HIV-associated gut dysbiosis.

Monocyte-derived macrophages are necessary for resolving lower airway infection and inflammation (*94, 95*), and progressive HIV infection results in depletion of circulating monocytes and build-up of poorly activated, pro-inflammatory macrophages within the gut (*89, 96*). The gut microbiome is known to modulate local and systemic

immunity (*33, 90, 91, 97*), and microbiome dysbiosis is a hallmark of HIV infection (*14, 43, 46*), so we investigated whether sterile, fecal microbial products from HIV infected pneumonia patients differentially modulated macrophage function based on patient disease severity. Using the THP-1 cell line, we demonstrated that sterile fecal water (SFW) from patients with low CD4 counts is poorer at activating (CD80+CD86+) macrophages and differentiating tissue repair (CD206+IL-10+) macrophages compared to CD4 high patient SFW, but better at inducing pro-inflammatory (IL-1β+) macrophages. These results indicate that the gut microbiota from HIV infected pneumonia patients differentially modulates *in vitro* macrophage activation and differentiation based on HIV disease severity, and that this may contribute to *in vivo* macrophage dysfunction and pneumonia patient outcomes; however, follow-up animal models are necessary to prove whether manipulation of the gut microbiota independent of the airways can modulate pneumonia inflammation.

Microbial translocation is an established phenomenon within severely ill HIV infected individuals with mucosal barrier breakdown and CD4+ T cell depletion; in this study we demonstrate that the most severely ill patients with the lowest CD4 counts have gut microbiota that is enriched for microbes shared with the airways. While this study is cross-sectional and the direction of microbial translocation was not definitively determined, the significant relationship between paired distance and variation in gut, but not airway, microbiota suggests that the gut microbiota is gaining airway microbes, rather than vice versa. Additionally, the probable increase in lung mucosal barrier breakdown from pneumonia may explain how more microbes would travel distally from the lungs than from the gut. Nonetheless, longitudinal sampling of the lower airways and

gut are necessary to determine whether specific microbes are initially colonizing the respiratory or gastrointestinal tract. Further follow-up is also required to know whether these results can be generalized to the broader HIV infected population with and without pneumonia since this study focused solely on Ugandan HIV infected patients with bacterial pneumonia.

Despite these limitations, this study is the first to reveal a potential role for the gut microbiome in understanding HIV-associated pneumonia and a relationship between the gut-lung microbial axis and patient outcomes. Bacterial pneumonia in HIV infected individuals is frequently fatal, yet the determinants of patient outcomes are poorly understood. We demonstrate that gut microbiota composition is significantly related to pneumonia patient outcome and that the products of the gut microbiome may modulate macrophage phenotype. Overall, these investigations suggest that in the future, profiling and modulating the gut microbiome may provide a better avenue for developing patient interventions than the airway microbiome in HIV infected pneumonia patients.

MATERIALS AND METHODS

Study Design

We enrolled subjects infected with HIV admitted to Mulago Hospital in Kampala, Uganda for acute pneumonia from June 2012 to November 2015 as part of the Lung MicroCHIP (Lung Microbiome in Cohorts of HIV Infected Persons) study. Patients without a prior diagnosis of TB underwent bronchoscopy with BAL for clinical diagnosis, as previously described (*20*), with 10 ml set aside for microbiome analysis. Bronchoscopy was performed a median of 1 day after hospital admission (interquartile range, 1–3 d) with concurrent collection of stool. Patients underwent two sputum acid-fast bacilli smear examinations to diagnose pulmonary TB. Clinical data were collected and diagnoses were assigned as previously described (*19*). Study endpoint was hospital discharge.

Ethics Statement

The Makerere University School of Medicine Research Ethics Committee, the Mulago Hospital Research and Ethics Committee, the Uganda National Council for Science and Technology, and the University of California San Francisco Committee on Human Research approved the protocol. Subjects provided written, informed consent.

Bacterial and fungal community profiling

DNA extraction

Individual BAL and stool samples were placed into lysing matrix E (INTEGRA Biosciences) tubes pre-aliquoted with 500 µl of hexadecyltrimethylammonium bromide

(CTAB; Sigma Aldrich) DNA extraction buffer and incubated at 65º C for 15 minutes. An equal volume of phenol:chloroform:isoamyl alcohol (25:24:1; Sigma Aldrich) was added to each tube and samples were homogenized in a Fast Prep-24 homogenizer at 5.5 m/s for 30 seconds. Tubes were centrifuged for 5 minutes at 16,000 x g and the aqueous phase was transferred to individual wells of a deep-well 96-well plate. An additional 500 µl of CTAB buffer was added to the LME tubes, the previous steps were repeated, and the aqueous phases were combined. An equal volume of chloroform was added to each sample and mixed followed by centrifugation at 3000 x g for 10 minutes to remove excess phenol. The aqueous phase (600 µl) was transferred to a deep-well 96-well plate, combined with 2 volume-equivalents of polyethylene glycol (PEG) and stored overnight at 4º C to precipitate DNA. Plates were centrifuged for 60 min at 3000 x g to pellet DNA and the PEG solution was removed. DNA pellets were washed twice with 300 µl of 70% ethanol, air-dried for 30 minutes and suspended in 100 µl of sterile water. DNA samples were quantitated using the Qubit dsDNA HS Assay Kit and diluted to 10 or 2 ng/µl for stool or BAL, respectively. No template controls (NTCs, one per extraction plate) were processed similar to samples.

16S DNA amplification and sequencing

The V4 region of the 16S rRNA gene was amplified in triplicate, as previously described (*91*). Triplicate reactions each possessed one no-template control to assess background contamination. Triplicate reactions were combined and purified using the SequalPrep Normalization Plate Kit (Invitrogen) according to manufacturer's specifications. Purified amplicons were quantitated using the Qubit dsDNA HS Assay Kit

and pooled at equimolar concentrations. For NTCs, total volume was pooled because burden was too low for equimolar concentrations. The amplicon library was concentrated using the Agencourt AMPure XP system (Beckman-Coulter) quantitated using the KAPA Library Quantification Kit (KAPA Biosystems) and diluted to 2nM. Equimolar PhiX was added at 40% final volume to the amplicon library and sequenced on the Illumina NextSeq 500 Platform on a 153bp x 153bp sequencing run.

ITS2 DNA Amplification and Sequencing

The internal transcribed spacer region 2 (ITS2) of the rRNA gene was amplified in triplicate, as previously described (*91*). Amplification and sequencing were performed with the same protocols as above with the following modifications: PhiX (5 pM final) to the amplicon library and sequenced on the Illumina MiSeq Platform on a 290bp x 290bp sequencing run.

16S OTU Table Generation

Raw sequence data were converted from bcl to fastq format using bcl2fastq v2.16.0.10. Paired sequencing reads with a minimum overlap of 25 bp were merged using FLASH v1.2.11. Successfully merged reads were identified, had index sequences extracted and were demultiplexed in the absence of quality filtering using QIIME (Quantitative Insights Into Microbial Ecology, v1.9.1). Reads were then quality filtered using USEARCH's fastq filter (v7.0.1001) to remove reads having >2 expected errors. Following this step, BAL and stool samples were both processed individually, as well as part of a combined dataset. Quality filtered reads were dereplicated at 100%

identity, clustered at 97% sequence identity into operational taxonomic units (OTUs) if they had ≥2 reads and had chimeras removed, and mapped back to the resulting OTUs using USEARCH v8.0.1623. The Greengenes database (May 2013) was used to assign taxonomy to OTUs and QIIME was used to make the phylogenetic tree. OTUs were filtered by 1) removing any known common contaminant OTU present in more than half of the Negative Extraction Controls for this study, and 2) removing any remaining OTU that had a total read count across all samples less than 1/5000th of a percent of the total read counts across all samples for total OTU picking, or removing any OTU with less than 70 reads for BAL samples and less than 250 reads for stool samples. Finally, sequencing reads were representatively normalized by rarefying 100 times and using sample centroids as described previously (*91*). Total sample OTU picking was rarefied to 51,997, BAL alone to 67,135, and stool alone to 120,665 reads per sample. Total reads from each NTC were below the level of rarefying, and NTCs for BAL samples possessed a maximum of 2 reads/NTC at dereplication.

ITS2 OTU Table Generation

ITS2 OTU tables were generated using the same protocol as above with the following modifications and notes: 1) chimeras were removed and taxonomy was assigned using the UNITE database [January 2016, (*98*)]; 2) no phylogenetic tree was generated; 3) no NTC successfully sequenced; 4) any OTU was removed if it possessed a total read count less than 1/5000th of a percent of the total read counts across all samples for total, BAL, or stool OTU picking; and 5) samples were

representatively rarefied to 1000, 1044, or 2565 reads/sample for total, BAL, or stool OTU picking, respectively.

THP-1 sterile fecal water (SFW) assay and flow cytometry.

Stool samples from 12 CD4 low patients and 12 CD4 high patients were used as biological replicates, and also stratified based on Bray Curtis paired distance into 6 low, 11 intermediate, and 7 high samples. Samples were chosen based on CD4 low or high status and sufficient material. Stool samples were homogenized 1 g ml^{-1} (w:v) in pre-warmed phosphate-buffered saline (PBS) containing 20% heat-inactivated fetal bovine serum (FBS). Samples were vigorously vortexed, incubated (37° C water bath, 10 min), and centrifuged (14,000 x g, 10 min). Supernatant was removed to a new tube and centrifuged again (16,000 x g, 1 hour). Supernatant was filter-sterilized through a 0.4-µm filter, followed by a 0.2-µm filter, before sterile fecal water (SFW) was used in the THP-1 assay described below. SFW buffer (PBS with 20% FBS) and R10 media [Roswell Park Memorial Institute media 1640 with 10% heat-inactivated FBS (antigen activator) and 2 mM L-glutamine and 100 U ml^{-1} penicillin–streptomycin added; Life Technologies] were used as the negative controls.

The THP-1 monocyte cell line (passages 2-3) was cultured for 48 hours with 10ng/ml Phorbol 12-Myristate 13-Acetate (PMA, Thermo Fisher Scientific) added to R10 media to induce macrophage differentiation. Half of the culture media was changed to new R10 with SFW (final SFW concentration 8%) for 24 hours. Controls included R10 and SFW buffer. To assess cytokine production, the cultures were mixed with GolgiPlug (Gplug; BD Biosciences) for 12 h before staining for flow cytometry. Experiment was

replicated twice due to fecal material scarcity, with three technical replicates per biological replicate.

For flow cytometry, single-cell suspensions were stained using a panel of antibodies, including anti-CD14 (63D3, 1:100; BV711, BioLegend), anti-CD68 (Y1/82A, 1:40; PE-Cy7, BioLegend), anti-CD80 (2D10, 1:50; FITC, BioLegend), anti-CD86 (IT2.2, 1:50; BV650, BioLegend); anti-CD206 (15-2, 1:50; APC-Cy7, BioLegend); anti-IL-1β (H1b-98, 1:20; Pacific Blue, BioLegend); and anti-IL-10 (JES3-9D7, 1:20; PE, Miltenyi Biotec). Validation for each primary antibody is provided on the manufacturers' websites. Dead cells were stained positive with LIVE–DEAD Aqua Dead Cell Stain (Life Technologies). Permeabilization buffer (BD Biosciences) was used to permeabilize cells before staining for the intracellular markers CD68, IL-1β, and IL-10. For flow analysis, live THP-1 macrophages were gated as $CD14^{+}CD68^{+}$ cells. Among macrophage subpopulations, activated cells were $CD80^{+}CD86^{+}$, pro-inflammatory cells were $IL\text{-}\beta^{+}$, and tissue repair cells were both $CD206^{+}$ and $IL\text{-}10^{+}$. Stained cells were assayed via a flow cytometer on a BD LSR II (BD Biosciences).

Statistical analysis.

Diversity and richness indices were calculated using QIIME. Distance matrices [weighted UniFrac (*99*) for BAL comparison with previous lung studies; Bray–Curtis for stool and shared taxa] were calculated in QIIME (*71*) to assess compositional dissimilarity between samples, and visualized using PCoA plots constructed in Emperor (*83*). Bray Curtis paired distance was determined using whole dataset OTU picking. Permutational multivariate analysis of variance (PERMANOVA) was performed

using *Adonis* in the R environment (*vegan* package) to determine factors that explained variation in microbiota composition, and was used to confirm that BAL and stool composition were not related to processing factors ($p \geq 0.1$). Unsupervised hierarchical clustering was performed using the *hclust* function in the R environment (*stats, heatmap3* packages) with Bray Curtis distance (*vegan* package) and Ward 2 cluster method on bacterial families present in ≥1 sample at ≥3% abundance.

To determine which OTUs differed in relative abundance between patient CD4 or paired distance groups, the distribution of each OTU was modeled using Poisson, negative binomial, or zero-inflated negative binomial distributions and the best fit distribution was assigned based on Akaike information criterion (AIC), as previously described (*88*). This modeling approach is appropriate for sequence-count data, and was corrected for multiple testing using false-discovery rate [$q < 0.05$, (*100*)]. Results were presented if ≥50% of the enriched group possessed the OTU and mean OTU difference between patient groups was ≥100 reads/sample.

When examining the association between clinical or microbiological factors and patient groups, *P* values were calculated on the basis of covariate distribution by Kruskal–Wallis (numerical, skewed distribution, >2 groups), Mann-Whitney (numerical, skewed, 2 groups), chi-square (categorical), or Fisher's exact (sparse categorical) tests. To test for THP-1 differences based on patient group, Kruskal Wallis was performed in GraphPad Prism 6. Except where indicated, all analyses were conducted in the R statistical programming environment.

SUPPLEMENTAL FIGURES AND TABLES

Figure S1.

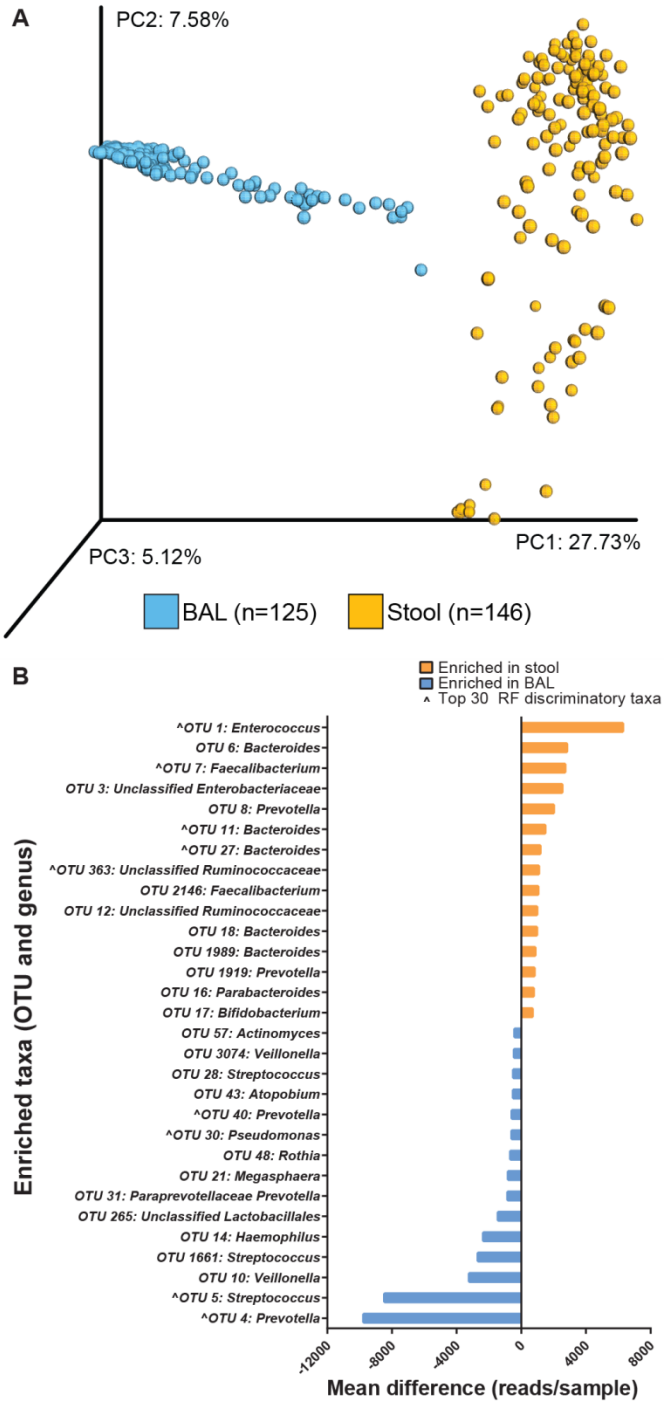


Figure S1. HIV infected pneumonia patient stool and BAL exhibit site specificity.

A. PCoA using Bray Curtis dissimilarity of n = 271 BAL (n = 125) and stool (n = 146)

bacterial community profiles (including 120 BAL-stool patient paired samples) representatively rarefied to 51,997 reads/sample. **B.** Mean difference in reads/sample are plotted for the top 30 taxa differentiating BAL and stool bacterial communities, using linear mixed effects and filtering based on $q < 0.05$, read difference ≥ 100, and presence in ≥ 50% of the enriched group. PC = principal coordinate; ^ = one of top 30 predictive taxa for random forest prediction of BAL and stool microbiota.

Figure S2.

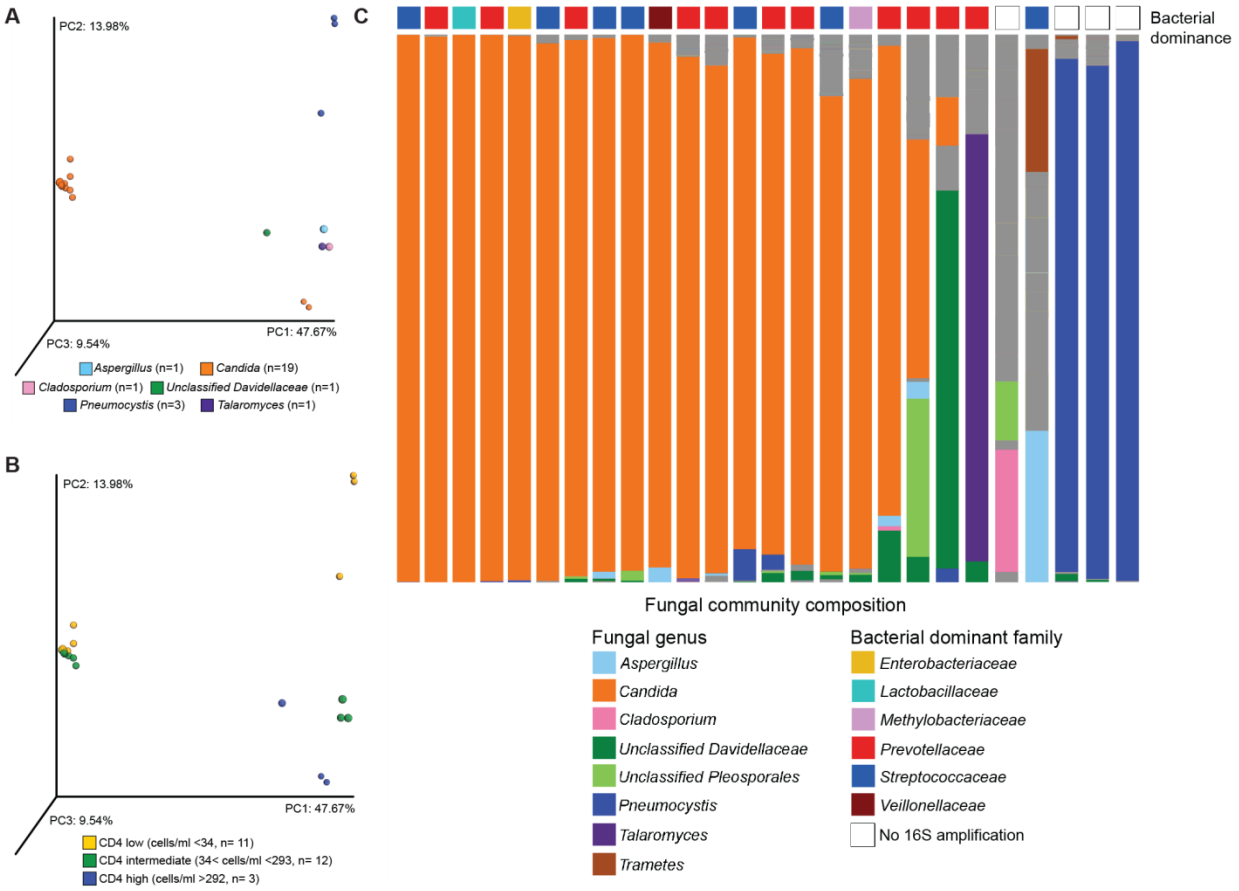


Figure S2. Lower airway fungal communities are primarily dominated by *Candida*. PCoA of Bray Curtis dissimilarity for n = 26 BAL fungal communities representatively rarefied to 1044 reads/sample demonstrate that **A.** the majority of samples with fungal profiles are dominated by *Candida* (PERMANOVA for dominant genus, $R^2 = 0.673$, $p < 0.001$), and **B.** CD4 count (grouped by quartile with quartiles 2 and 3 combined) is significantly related to fungal composition with *Pneumocystis* dominated patients all possessing low CD4 counts (PERMANOVA, $R^2 = 0.220$, $p < 0.01$). **C.** Taxonomic profiles of fungal microbiota with bacterial dominant family indicated above. *Pneumocystis* dominated samples consistently lacked lower airway 16S bacterial amplification (in white), indicating that patients examined for bacterial pneumonia do not

have confounding *Pneumocysis* pneumonia. PC = principal coordinate; PERMANOVA = permutational multivariate ANOVA.

Figure S3.

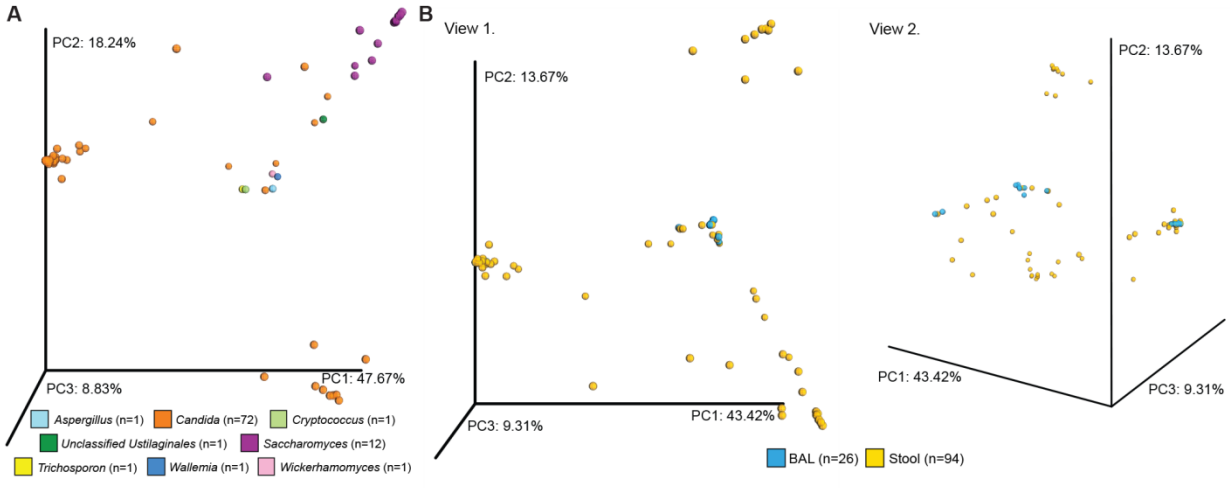


Figure S3. Similar to lower airways, stool of HIV infected pneumonia patients is consistently dominated by *Candida* and does not show anatomic site distinction from the lower airways. A. PCoA of Bray Curtis dissimilarity (BC) for n = 90 stool fungal communities representatively rarefied to 2565 reads/sample demonstrate that the majority of samples with fungal profiles are dominated by *Candida* (PERMANOVA for dominant genus, $R^2 = 0.442$, $p < 0.001$). *Candida* cluster on the left is dominated by OTU 1, cluster on the bottom right by OTU2, and samples in the middle by OTU4. **B.** PCoA of BC for n=26 BAL and n = 94 stool fungal communities representatively rarefied to 1000 reads/sample demonstrate samples are not clustered by anatomic site. Two views are provided to fully display BAL microbiota distribution. PC = principal coordinate; PERMANOVA = permutational multivariate ANOVA; BAL = bronchoalveolar lavage.

Figure S4.

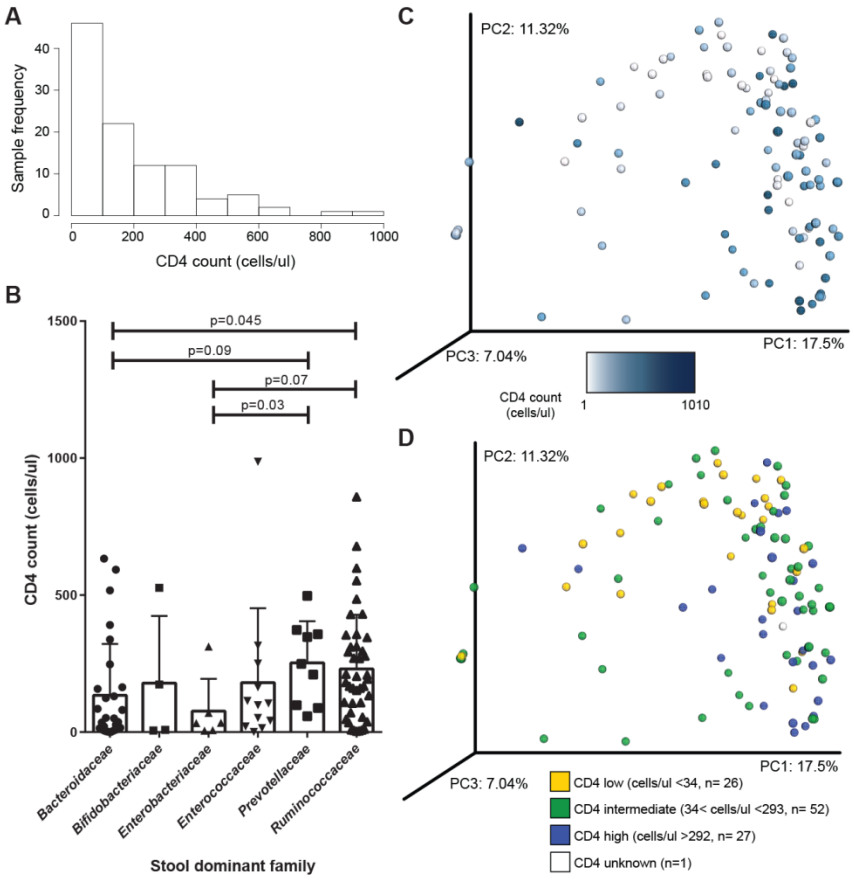


Figure S4. Stool bacterial microbiota is related to circulating CD4 count. A. Histogram of CD4 count (cells/μl) distribution across patients with stool bacterial profiles. **B.** CD4 count (cells/μl) is significantly related to stool dominant bacterial family (plotted for families with > 2 samples; KW, p=0.02; plotted p-values are with Mann Whitney U test). PCoA of Bray Curtis dissimilarity for n = 106 stool samples representatively rarefied to 120,665 reads/sample demonstrates that CD4 count significantly explains variation in stool composition, both as **C.** a continuous variable as measured from patients (cells/μl; PERMANOVA, R^2 = 0.017, p = 0.025) and as **D.** grouped by quartile (second and third combined into intermediate group; PERMANOVA,

R^2 = 0.052, p = 0.002). PC = principal coordinate; PERMANOVA = permutational multivariate ANOVA.

Figure S5.

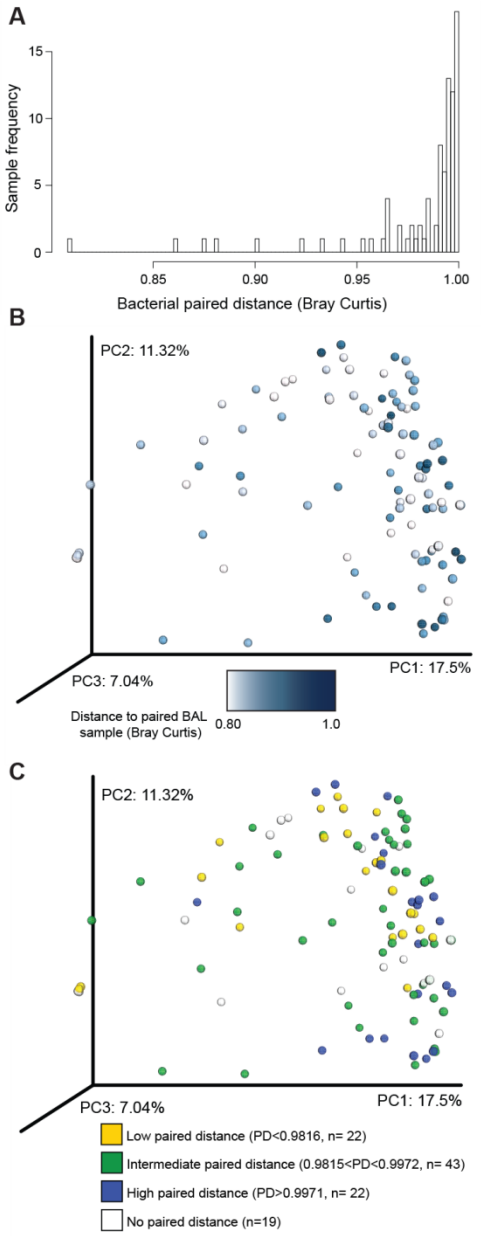


Figure S5. Stool bacterial microbiota composition differs based on similarity to paired BAL sample. A. Histogram of paired BAL-stool microbiota distance [Bray Curtis (BC)] distribution (outlier distance = 0.8). PCoA of BC for n = 106 stool samples demonstrates that distance to paired BAL sample (BC) significantly explains variation in stool composition, both as **B.** a continuous variable (PERMANOVA, R^2 = 0.026, p =

0.007) and **C.** grouped by quartile (second and third combined into intermediate group; PERMANOVA, R^2 = 0.054, p = 0.002). PC = principal coordinate; PERMANOVA = permutational multivariate ANOVA; BAL = bronchoalveolar lavage.

Figure S6.

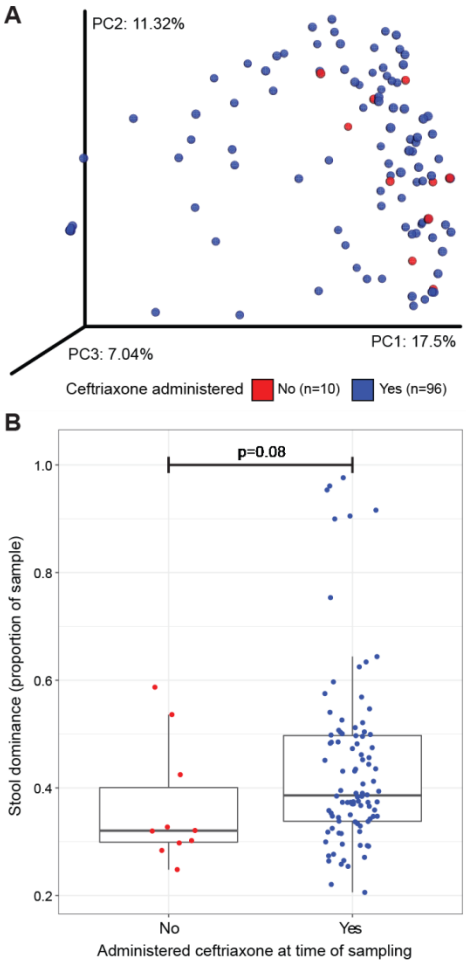


Figure S6. The antibiotic ceftriaxone is significantly related to stool bacterial microbiota. A. PCoA of Bray Curtis dissimilarity for n = 106 stool samples demonstrates that ceftriaxone administration at time of sample collection significantly explains variation in stool composition (PERMANOVA, R^2 = 0.019, p = 0.01). **B.** Ceftriaxone administration trends towards an association with lower stool bacterial diversity (Mann Whitney U Test, p = 0.08). PC = principal coordinate; PERMANOVA = permutational multivariate ANOVA**.**

Figure S7.

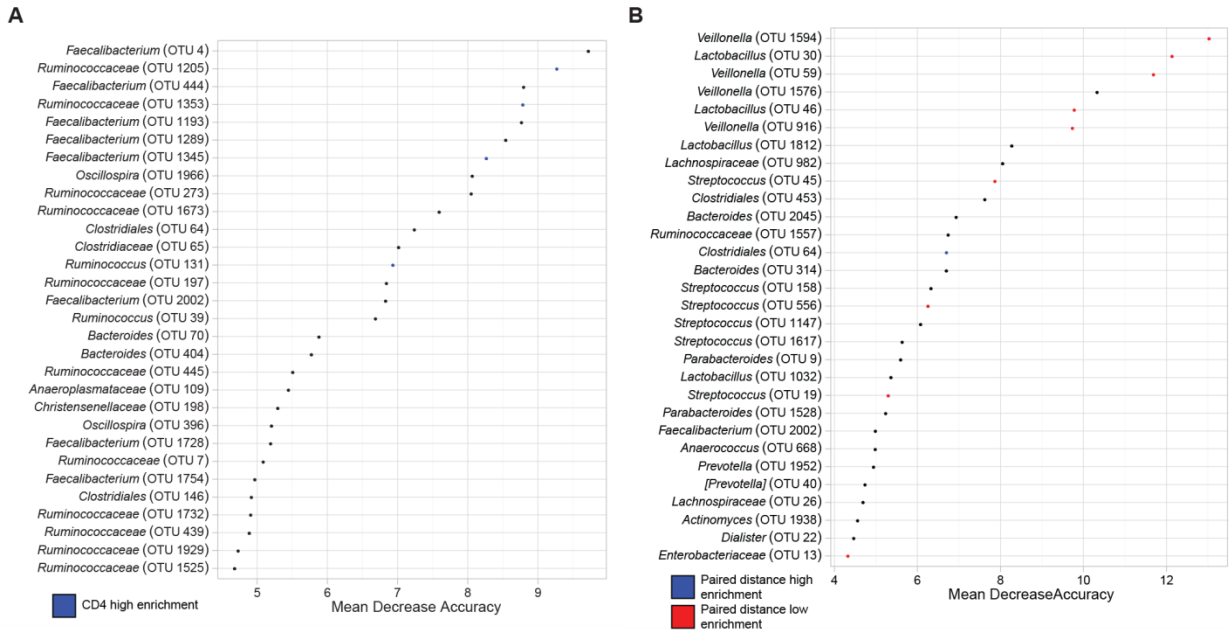


Figure S7. Random Forest confirms CD4 and paired distance associated enrichments. Random Forest mean decrease accuracy plotted for the top 30 most predictive taxa of **A.** CD4 group (quartile 1 versus 4), confirms enrichment of *Faecalibacterium*, *Bacteroides*, and *Ruminococcaceae* in stool bacterial communities of patients with high CD4 counts, and of **B.** paired distance group (Bray Curtis, quartile 1 versus 4), confirms enrichment of *Veillonella* and *Streptococcus* in stool bacterial communities of patients with more similar stool and lower airway samples. *[Prevotella]* = *Prevotella* within the family *Paraprevotellaceae*.

Figure S8.

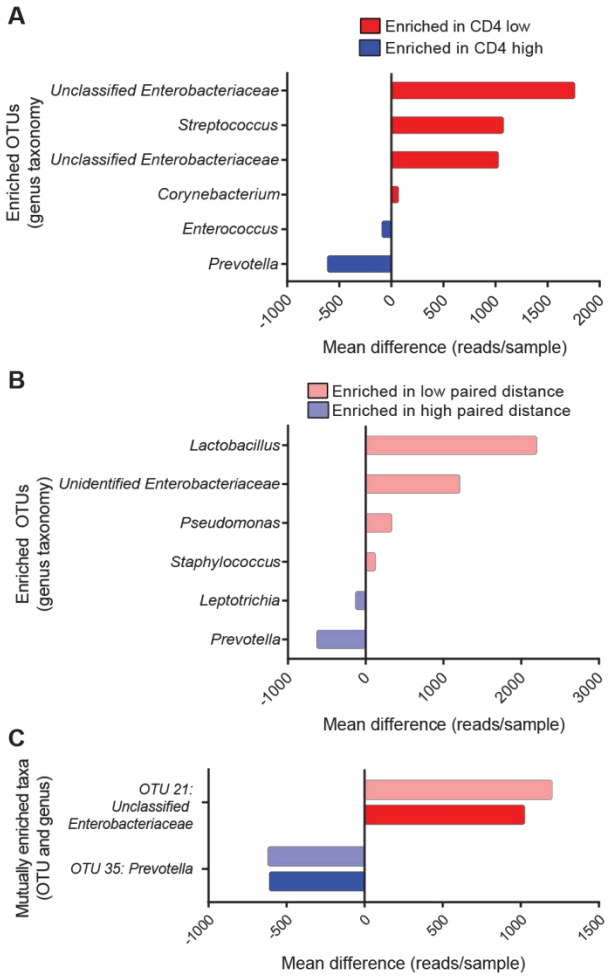


Figure S8. *Enterobacteriaceae* is consistently enriched in the lower airways of patients with low CD4 counts and paired distance. Mean difference in reads/sample are plotted for the top taxa (OTUs summarized to genus taxonomy) differentiating **A.** CD4 low and high, as well as **B.** paired distance low and high, BAL bacterial communities. **C.** Mean difference in reads/sample are plotted for top differential taxa shared between CD4 and paired distance enrichment comparisons. All data plotted were filtered based on $q < 0.05$, read difference ≥ 100, and presence in ≥ 50% of the enriched group; no OTUs with the same genus level taxonomy assignment were

significantly enriched in two different directions for any comparison. BAL = bronchoalveolar lavage; OTU = operational taxonomic unit.

Figure S9.

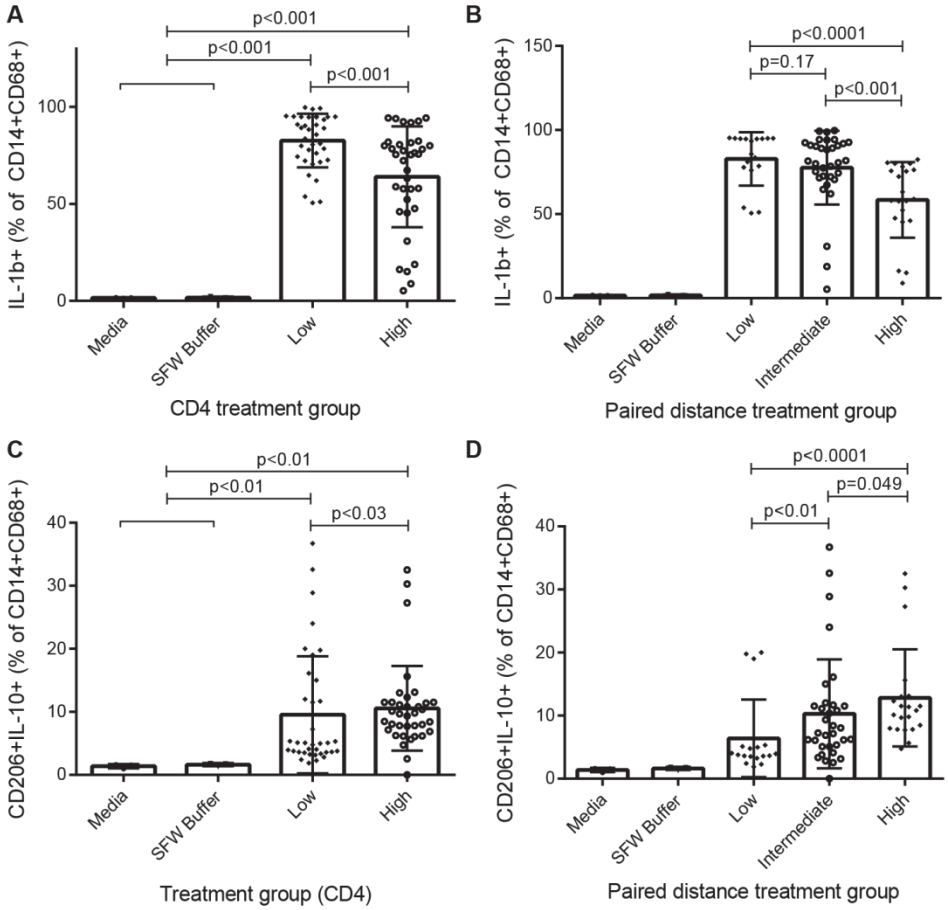


Figure S9. Sterile microbial stool metabolites and ligands from patients with low CD4 counts or paired distance induce increased pro-inflammatory and decreased tissue repair macrophage differentiation. Increased pro-inflammatory IL-1β+, total THP-1-derived macrophages are induced by **A.** CD4 low compared to CD4 high (KW, p<0.0001), or **B.** paired distance low or medium compared to paired distance high (KW, p<0.0001; all treatments versus buffer or media p<0.0001), patient SFW. Decreased tissue repair associated CD206+ IL-10+, total THP-1-derived macrophages are induced by **C.** CD4 low compared to CD4 high (KW, p<0.001), or **D.** paired distance low compared to paired distance high or medium (KW, p<0.001; all treatments versus buffer or media p<0.01), patient SFW. Low and high groups are quartiles 1 and 4 respectively,

with paired distance quartile 2 and 3 combined into an intermediate group. Plotted p-values from Mann Whitney test. KW = Kruskal Wallis; SFW = sterile fecal water.

Table S1. Clinical and demographic features of the HIV infected pneumonia patient cohort.

Variable	Sample (n)	Yes/No[a]	Min-Max (Median)
Gender	153	77|76 (F|M)	
Smoker	153	35|118	
Age	153		19-69 (34)
Temperature (ºC)	153		34.7-41 (36.9)
Chest pain	153	97|56	
Cough	153	153|0	
Dyspnea	153	83|79	
Wheeze	153	34|119	
Previous TB diagnosis	153	19|134	
Pulmonary Kaposi's Sarcoma	146	6|140	
Pneumocystis prophylaxis	123	101|22	
Antiretrovirals at admission	123	64|59	
Antibiotics at sample collection	146	146|0	
Ceftriaxone at sample collection	146	136|10	

[a] Unless otherwise noted

Table S2. HIV infected pneumonia patient stool fungal composition is consistently dominated by *Candida* or *Saccharomyces*.

Dominant Genus	Sample number (n =)	Mean dominance (%)
No fungal profile[a]	57	*NA*
Candida	72	94
Saccharomyces	12	82
Trichosporon	1	66
Aspergillus	1	87
Cryptococcus	1	85
Unclassified Ustilaginales	1	50
Wallemia	1	58
Wickerhamomyces	1	24

[a]Either less amplification than no template controls or did not reach multiply rarefying depth of 2565 reads/sample

ACKNOWLEDGEGMENTS

The authors thank the patients and staff of Mulago Hospital for their contributions to the study. They also thank the research team who enrolled the patients who participated in this study.

This work was supported by National Institutes of Health/NHLBI awards U01 HL098964, K24 HL087713, and R01 HL090335.

Author Contributions: L.H. and S.V.L. designed the study. W.W., I.A., P.B., S.K., E.C., M.M., M.S., and L.H. trained personnel in Uganda, enrolled patients, collected and shipped BAL and stool samples, and obtained demographic and clinical data. M.K.S. and D.F. performed microbiota processing. D.L.L. and M.K.S. designed immune assays. M.K.S. performed immune assays, microbiota profiling, and analyses. D.F. provided microbiota profiling guidance, and K.M. provided statistical guidance. M.K.S. and S.V.L. wrote the manuscript. All authors reviewed and approved the final manuscript.

CHAPTER 4

Implications and Future Directions

The research described above highlights the importance of the lung and gut microbiomes in HIV infected patients with bacterial pneumonia, demonstrating that types of microbial communities and their functions can be used to better understand patient outcomes. Despite high morbidity and mortality within the HIV infected pneumonia population, little is known about factors that influence heterogeneity in patient outcomes, and, specifically, whether variation in airway or gut microbiota composition and function are related to patient immune response, disease severity, and survival. The goal of this thesis was to determine whether specific microbial communities and their functions are clinically meaningful to understanding patient disease and outcomes within HIV infected patients with bacterial pneumonia. While previous studies have emphasized that HIV infection results in changes to the gut or lung microbiota (*8, 11, 43*), the studies herein describe the relationship between microbial community composition and function with HIV-associated pneumonia morbidity and mortality and outline a potential mechanism by which HIV-induced changes to the microbiota may lead to pneumonia outcomes.

This thesis builds on previous investigations of the lung microbiome that associated HIV infection and co-morbidities including pneumonia, pulmonary tuberculosis, or COPD with increased abundances of *Prevotella, Streptococcus,* and *Veillonella* within the lower airways (*10, 11, 18, 40*). The studies described in Chapter 2 demonstrate that Ugandan HIV-pneumonia patients colonized by *Prevotellaceae*-dominated communities are enriched for products of branched chain amino acid (BCAA) metabolism within their circulation, and express a number of T helper cell 2 (Th2) and Th17 cytokines within their lower airways, including interleukin 17 (IL-17A). BCAA

pathway products can be bacterial derived (*35*) and positively regulate mechanistic target of rapamycin (mTOR) signaling (*36*), a key component of Th17 differentiation (*37*), suggesting that *Prevotellaceae*-dominated communities may induce Th17 inflammation through BCAA metabolism, though the direct production of BCAA by *Prevotellaceae* in these patients still needs to be demonstrated. Nonetheless, association between lung Th17 inflammation and *Prevotella* and its co-colonizer, *Veillonella,* is corroborated by multiple studies: the presence and abundance of these oral commensals within the lungs is associated with increased Th17 inflammation in healthy individuals (*38*) and can induce Th17 inflammation *in vitro* and in mouse models (*39*). The studies within this thesis illustrate a shift in the airway microbiome field towards considering airway microbes within the context of host-wide metabolism and immune response, rather than as isolated microbial infections. Little is known about the causes of patient survival and recovery within this population, so stratifying patients into clinically meaningful groups based on lower airway microbial communities provides a novel basis for understanding and treating this patient population.

While airway infections such as pneumonia are traditionally studied in the context of microbial colonization of the airways, in Chapter 3 a novel relationship is demonstrated between gut microbial composition and HIV-pneumonia patient disease severity. Recently, the gut microbiome has been shown to influence and modulate local and systemic immune responses through metabolic products such as short chain fatty acids or linoleic acid metabolites, even in the context of airway disease such as asthma (*33, 91, 97*). The studies of the gut-lung axis described above demonstrate a novel role for the gut microbiome with airway disease in HIV infected patients; increased gut

microbial dysbiosis and airway microbial translocation to the gut in HIV-pneumonia patients was associated with increased disease severity as measured by lower CD4 counts, more airway-like gut communities, and *in vitro* immune response. These studies emphasize the systemic impact of HIV infection and HIV-associated co-morbidities and suggest that mucosal barrier breakdown in pneumonia patients may lead to airway pathogen colonization within the gut. The products of these altered gut communities induced dysfunctional, pro-inflammatory macrophages *in vitro*, suggesting that the gut microbiome may participate in innate immune priming in this patient population and dysbiotic microbiota may help drive systemic inflammation and severe outcomes within this patient population. These studies illustrate how the gut microbiome may afford a second avenue for understanding this patient population and tailoring anti- and pro-biotic therapies to treat the systemic microbial and immune dysfunction within HIV infected patients with pneumonia.

Despite the utility of HAART in treating HIV infection and improving patient health, pulmonary disease such as pneumonia still presents a challenging co-morbidity with this population (*6, 17*). These studies generated the first evidence that types of lung and gut microbial communities are tied to patient outcomes within the HIV-associated bacterial pneumonia population and have the ability to drive innate immune dysfunction *in vitro* that mirrors patient immune responses. Future studies can build on these early indications of a systemic microbial-immune interaction by investigating whole microbial meta-genomes and –transcriptomes in large cohorts of patients, as well as incorporating *in vitro* and animal model investigations to dissect the influence of specific microbial-microbial interactions on host immune response. The gut and lung

microbiomes present a novel area for understanding and treating disease in HIV infected patients with bacterial pneumonia. This thesis identified distinct community states within the lower airways and airway microbial translocation to the gut which are associated with patient outcomes, demonstrating that patient stratification based on gut and lung microbial communities may lead to more targeted and effective disease treatments for HIV infected pneumonia patients in the future.

References

1. S. V. Lynch, O. Pedersen, The Human Intestinal Microbiome in Health and Disease. *New England Journal of Medicine* **375**, 2369-2379 (2016).
2. L. Hall-Stoodley, J. W. Costerton, P. Stoodley, Bacterial Biofilms: from the Natural environment to Infectious Diseases. *Nat Rev Micro* **2**, 95-108 (2004).
3. M. Burmølle, D. Ren, T. Bjarnsholt, S. J. Sørensen, Interactions in Multispecies Biofilms: do they actually matter? *Trends in Microbiology* **22**, 84-91 (2014).
4. C. M. Bassis, J. R. Erb-Downward, R. P. Dickson, C. M. Freeman, T. M. Schmidt, V. B. Young, J. M. Beck, J. L. Curtis, G. B. Huffnagle, Analysis of the Upper Respiratory Tract Microbiotas as the Source of the Lung and Gastric Microbiotas in Healthy Individuals. *mBio* **6**, (2015).
5. R. P. Dickson, J. R. Erb-Downward, C. M. Freeman, N. Walker, B. S. Scales, J. M. Beck, F. J. Martinez, J. L. Curtis, V. N. Lama, G. B. Huffnagle, Changes in the Lung Microbiome following Lung Transplantation Include the Emergence of Two Distinct Pseudomonas Species with Distinct Clinical Associations. *PLoS One* **9**, (2014).
6. A. Morris, K. Crothers, J. M. Beck, L. Huang, An Official ATS Workshop Report: Emerging Issues and Current Controversies in HIV-associated Pulmonary Diseases. *Proc Am Thorac Soc* **8**, 17-26 (2011).
7. K. Crothers, B. W. Thompson, K. Burkhardt, A. Morris, S. C. Flores, P. T. Diaz, R. E. Chaisson, G. D. Kirk, W. N. Rom, L. Huang, H. I. V. S. for the Lung, HIV-Associated Lung Infections and Complications in the Era of Combination Antiretroviral Therapy. *Proceedings of the American Thoracic Society* **8**, 275-281 (2011).
8. C. Lozupone, A. Cota-Gomez, B. E. Palmer, D. J. Linderman, E. S. Charlson, E. Sodergren, M. Mitreva, S. Abubucker, J. Martin, G. Yao, T. B. Campbell, S. C. Flores, G. Ackerman, J. Stombaugh, L. Ursell, J. M. Beck, J. L. Curtis, V. B. Young, S. V. Lynch, L. Huang, G. M. Weinstock, K. S. Knox, H. Twigg, A. Morris, E. Ghedin, F. D. Bushman, R. G. Collman, R. Knight, A. P. Fontenot, Widespread Colonization of the Lung by Tropheryma whipplei in HIV Infection. *American Journal of Respiratory and Critical Care Medicine* **187**, 1110-1117 (2013).
9. J. M. Beck, P. D. Schloss, A. Venkataraman, H. Twigg, K. A. Jablonski, F. D. Bushman, T. B. Campbell, E. S. Charlson, R. G. Collman, K. Crothers, Multicenter Comparison of Lung and Oral microbiomes of HIV Infected and HIV Uninfected Individuals. *Am J Respir Crit Care Med* **192**, (2015).
10. A. Morris, J. N. Paulson, H. Talukder, L. Tipton, H. Kling, L. Cui, A. Fitch, M. Pop, K. A. Norris, E. Ghedin, Longitudinal Analysis of the Lung Microbiota of Cynomolgous Macaques during Long-term SHIV Infection. *Microbiome* **4**, 38 (2016).
11. H. L. Twigg III, K. S. Knox, J. Zhou, K. A. Crothers, D. E. Nelson, E. Toh, R. B. Day, H. Lin, X. Gao, Q. Dong, D. Mi, B. P. Katz, E. Sodergren, G. M. Weinstock, Effect of Advanced HIV Infection on the Respiratory Microbiome. *American Journal of Respiratory and Critical Care Medicine*, (2016).

12. E. Zaura, B. J. Keijser, S. M. Huse, W. Crielaard, Defining the Healthy "Core Microbiome" of Oral Microbial Communities. *BMC Microbiol* **9**, 259 (2009).
13. T. Takeshita, Y. Nakano, T. Kumagai, M. Yasui, N. Kamio, Y. Shibata, S. Shiota, Y. Yamashita, The Ecological Proportion of Indigenous Bacterial Populations in Saliva is Correlated with Oral Health Status. *Isme j* **3**, 65-78 (2009).
14. I. Vujkovic-Cvijin, R. M. Dunham, S. Iwai, M. C. Maher, R. G. Albright, M. J. Broadhurst, R. D. Hernandez, M. M. Lederman, Y. Huang, M. Somsouk, S. G. Deeks, P. W. Hunt, S. V. Lynch, J. M. McCune, Dysbiosis of the Gut Microbiota is Associated with HIV Disease Progression and Tryptophan Catabolism. *Sci Transl Med* **5**, 193ra191 (2013).
15. O. S. Sogaard, N. Lohse, J. Gerstoft, G. Kronborg, L. Ostergaard, C. Pedersen, G. Pedersen, H. T. Sørensen, N. Obel, Hospitalization for Pneumonia among Individuals With and Without HIV Infection, 1995–2007: A Danish Population-Based, Nationwide Cohort Study. *Clinical Infectious Diseases* **47**, 1345-1353 (2008).
16. A. M. Morris, L. Huang, P. Bacchetti, J. Turner, P. C. Hopewell, J. M. Wallace, P. A. Kvale, M. J. Rosen, J. Glassroth, L. B. Reichman, J. D. Stansell, Permanent Declines in Pulmonary Function following Pneumonia in Human Immunodeficiency Virus-Infected Persons. The Pulmonary Complications of HIV Infection Study Group. *Am J Respir Crit Care Med* **162**, 612-616 (2000).
17. C. S. Marshall, A. J. Curtis, T. Spelman, D. P. O'Brien, J. Greig, L. Shanks, P. du Cros, E. C. Casas, M. S. da Fonseca, E. Athan, J. H. Elliott, Impact of HIV-associated Conditions on Mortality in People Commencing Anti-retroviral Therapy in Resource Limited Settings. *PLoS One* **8**, e68445 (2013).
18. S. Iwai, M. Fei, D. Huang, S. Fong, A. Subramanian, K. Grieco, S. V. Lynch, L. Huang, Oral and Airway Microbiota in HIV Infected Pneumonia Patients. *J Clin Microbiol* **50**, 2995-3002 (2012).
19. S. Iwai, D. Huang, S. Fong, L. G. Jarlsberg, W. Worodria, S. Yoo, A. Cattamanchi, J. L. Davis, S. Kaswabuli, M. Segal, L. Huang, S. V. Lynch, The Lung Microbiome of Ugandan HIV Infected Pneumonia Patients is Compositionally and Functionally Distinct from that of San Franciscan Patients. *PLoS One* **9**, e95726 (2014).
20. M. K. Shenoy, S. Iwai, D. L. Lin, W. Worodria, I. Ayakaka, P. Byanyima, S. Kaswabuli, S. Fong, S. Stone, E. Chang, J. L. Davis, A. A. Faruqi, M. R. Segal, L. Huang, S. V. Lynch, Immune Response and Mortality Risk Relate to Distinct Lung Microbiomes in Patients with HIV and Pneumonia. *Am J Respir Crit Care Med* **195**, 104-114 (2017).
21. K. Crothers, A. A. Butt, C. L. Gibert, M. C. Rodriguez-Barradas, S. Crystal, A. C. Justice, Increased COPD among HIV-positive Compared to HIV-negative Veterans. *Chest J* **130**, (2006).
22. R. P. Dickson, F. J. Martinez, G. B. Huffnagle, The Role of the Microbiome in Exacerbations of Chronic Lung Diseases. *Lancet* **384**, 691-702 (2014).
23. R. K. Albert , J. Connett , W. C. Bailey , R. Casaburi , J. A. D. J. Cooper , G. J. Criner , J. L. Curtis , M. T. Dransfield , M. K. Han , S. C. Lazarus , B. Make , N. Marchetti , F. J. Martinez , N. E. Madinger , C. McEvoy , D. E. Niewoehner , J. Porsasz , C. S. Price , J. Reilly , P. D. Scanlon , F. C. Sciurba , S. M. Scharf , G.

R. Washko , P. G. Woodruff , N. R. Anthonisen Azithromycin for Prevention of Exacerbations of COPD. *New England Journal of Medicine* **365**, 689-698 (2011).
24. H. M. Kling, T. W. Shipley, S. P. Patil, J. Kristoff, M. Bryan, R. C. Montelaro, A. Morris, K. A. Norris, Relationship of Pneumocystis jiroveci Humoral Immunity to Prevention of Colonization and Chronic Obstructive Pulmonary Disease in a Primate Model of HIV Infection. *Infection and Immunity* **78**, 4320-4330 (2010).
25. T. W. Shipley, H. M. Kling, A. Morris, S. Patil, J. Kristoff, S. E. Guyach, J. E. Murphy, X. Shao, F. C. Sciurba, R. M. Rogers, T. Richards, P. Thompson, R. C. Montelaro, H. O. Coxson, J. C. Hogg, K. A. Norris, Persistent Pneumocystis Colonization Leads to the Development of Chronic Obstructive Pulmonary Disease in a Nonhuman Primate Model of AIDS. *The Journal of Infectious Diseases* **202**, 302-312 (2010).
26. A. H. Limper, A. Adenis, T. Le, T. S. Harrison, Fungal Infections in HIV/AIDS. *The Lancet Infectious Diseases*, (2017).
27. J. M. Wallace, J. Hannah, Cytomegalovirus Pneumonitis in Patients with AIDS. *Chest* **92**, 198-203 (1987).
28. L. Cui, L. Lucht, L. Tipton, M. B. Rogers, A. Fitch, C. Kessinger, D. Camp, L. Kingsley, N. Leo, R. M. Greenblatt, S. Fong, S. Stone, J. C. Dermand, E. C. Kleerup, L. Huang, A. Morris, E. Ghedin, Topographic Diversity of the Respiratory Tract Mycobiome and Alteration in HIV and Lung Disease. *Am J Respir Crit Care Med* **191**, 932-942 (2015).
29. A. Morris, F. C. Sciurba, I. P. Lebedeva, A. Githaiga, W. M. Elliott, J. C. Hogg, L. Huang, K. A. Norris, Association of Chronic Obstructive Pulmonary Disease Severity and Pneumocystis Colonization. *American Journal of Respiratory and Critical Care Medicine* **170**, 408-413 (2004).
30. A. Chowdhary, K. Agarwal, S. Kathuria, P. K. Singh, P. Roy, S. N. Gaur, G. S. de Hoog, J. F. Meis, Clinical Significance of Filamentous Basidiomycetes Illustrated by Isolates of the Novel Opportunist Ceriporia lacerata from the Human Respiratory Tract. *Journal of Clinical Microbiology* **51**, 585-590 (2013).
31. J. C. Young, C. Chehoud, K. Bittinger, A. Bailey, J. M. Diamond, E. Cantu, A. R. Haas, A. Abbas, L. Frye, J. D. Christie, F. D. Bushman, R. G. Collman, Viral Metagenomics Reveal Blooms of Anelloviruses in the Respiratory Tract of Lung Transplant Recipients. *American journal of transplantation : official journal of the American Society of Transplantation and the American Society of Transplant Surgeons* **15**, 200-209 (2015).
32. H. L. Twigg, G. M. Weinstock, K. S. Knox, Lung Microbiome in Human Immunodeficiency Virus Infection. *Translational Research* **179**, 97-107 (2017).
33. A. Trompette, E. S. Gollwitzer, K. Yadava, A. K. Sichelstiel, N. Sprenger, C. Ngom-Bru, C. Blanchard, T. Junt, L. P. Nicod, N. L. Harris, B. J. Marsland, Gut Microbiota Metabolism of Dietary Fiber Influences Allergic Airway Disease and Hematopoiesis. *Nat Med* **20**, 159-166 (2014).
34. S. K. Cribbs, K. Uppal, S. Li, D. P. Jones, L. Huang, L. Tipton, A. Fitch, R. M. Greenblatt, L. Kingsley, D. M. Guidot, E. Ghedin, A. Morris, Correlation of the Lung Microbiota with Metabolic Profiles in Bronchoalveolar Lavage Fluid in HIV Infection. *Microbiome* **4**, 1-11 (2016).

35. L. K. Massey, J. R. Sokatch, R. S. Conrad, Branched-chain Amino Acid Catabolism in Bacteria. *Bacteriological Reviews* **40**, 42-54 (1976).
36. O. Zhenyukh, E. Civantos, M. Ruiz-Ortega, M. S. Sánchez, C. Vázquez, C. Peiró, J. Egido, S. Mas, High Concentration of Branched-Chain Amino Acids Promotes Oxidative Stress, Inflammation and Migration of Human Peripheral Blood Mononuclear Cells via mTORC1 Activation. *Free Radical Biology and Medicine* **104**, 165-177 (2017).
37. S. Nagai, Y. Kurebayashi, S. Koyasu, Role of PI3K/Akt and mTOR Complexes in Th17 cell Differentiation. *Annals of the New York Academy of Sciences* **1280**, 30-34 (2013).
38. L. N. Segal, J. C. Clemente, J.-C. J. Tsay, S. B. Koralov, B. C. Keller, B. G. Wu, Y. Li, N. Shen, E. Ghedin, A. Morris, P. Diaz, L. Huang, W. R. Wikoff, C. Ubeda, A. Artacho, W. N. Rom, D. H. Sterman, R. G. Collman, M. J. Blaser, M. D. Weiden, Enrichment of the Lung Microbiome with Oral Taxa is Associated with Lung Inflammation of a Th17 Phenotype. *Nature Microbiology* **1**, 16031 (2016).
39. J. M. Larsen, The Immune Response to Prevotella Bacteria in Chronic Inflammatory Disease. *Immunology* **151**, 363-374 (2017).
40. L. N. Segal, J. C. Clemente, Y. Li, C. Ruan, J. Cao, M. Danckers, A. Morris, S. Tapyrik, B. G. Wu, P. Diaz, G. Calligaro, R. Dawson, R. N. van Zyl-Smit, K. Dheda, W. N. Rom, M. D. Weiden, Anaerobic Bacterial Fermentation Products Increase Tuberculosis Risk in Antiretroviral-Drug-Treated HIV Patients. *Cell Host & Microbe* **21**, 530-537.e534 (2017).
41. P. Nowak, M. Troseid, E. Avershina, B. Barqasho, U. Neogi, K. Holm, J. R. Hov, K. Noyan, J. Vesterbacka, J. Svard, K. Rudi, A. Sonnerborg, Gut Microbiota Diversity Predicts Immune Status in HIV-1 Infection. *Aids* **29**, 2409-2418 (2015).
42. E. A. Mutlu, A. Keshavarzian, J. Losurdo, G. Swanson, B. Siewe, C. Forsyth, A. French, P. DeMarais, Y. Sun, L. Koenig, S. Cox, P. Engen, P. Chakradeo, R. Abbasi, A. Gorenz, C. Burns, A. Landay, A Compositional Look at the Human Gastrointestinal Microbiome and Immune Activation Parameters in HIV Infected Subjects. *PLOS Pathogens* **10**, e1003829 (2014).
43. Catherine A. Lozupone, M. Li, Thomas B. Campbell, Sonia C. Flores, D. Linderman, Matthew J. Gebert, R. Knight, Andrew P. Fontenot, Brent E. Palmer, Alterations in the Gut Microbiota Associated with HIV-1 Infection. *Cell Host & Microbe* **14**, 329-339 (2013).
44. S. M. Dillon, E. J. Lee, C. V. Kotter, G. L. Austin, Z. Dong, D. K. Hecht, S. Gianella, B. Siewe, D. M. Smith, A. L. Landay, C. E. Robertson, D. N. Frank, C. C. Wilson, An Altered Intestinal Mucosal Microbiome in HIV-1 Infection is Associated with Mucosal and Systemic Immune Activation and Endotoxemia. *Mucosal Immunol* **7**, 983-994 (2014).
45. T. Yatsunenko, F. E. Rey, M. J. Manary, I. Trehan, M. G. Dominguez-Bello, M. Contreras, M. Magris, G. Hidalgo, R. N. Baldassano, A. P. Anokhin, A. C. Heath, B. Warner, J. Reeder, J. Kuczynski, J. G. Caporaso, C. A. Lozupone, C. Lauber, J. C. Clemente, D. Knights, R. Knight, J. I. Gordon, Human Gut Microbiome Viewed across Age and Geography. *Nature* **486**, 222-227 (2012).
46. Cynthia L. Monaco, David B. Gootenberg, G. Zhao, Scott A. Handley, Musie S. Ghebremichael, Efrem S. Lim, A. Lankowski, Megan T. Baldridge, Craig B.

Wilen, M. Flagg, Jason M. Norman, Brian C. Keller, Jesús M. Luévano, D. Wang, Y. Boum, Jeffrey N. Martin, Peter W. Hunt, David R. Bangsberg, Mark J. Siedner, Douglas S. Kwon, Herbert W. Virgin, Altered Virome and Bacterial Microbiome in Human Immunodeficiency Virus-Associated Acquired Immunodeficiency Syndrome. *Cell Host & Microbe* **19**, 311-322.
47. S. J. Salter, M. J. Cox, E. M. Turek, S. T. Calus, W. O. Cookson, M. F. Moffatt, P. Turner, J. Parkhill, N. J. Loman, A. W. Walker, Reagent and Laboratory Contamination can Critically Impact Sequence-based Microbiome Analyses. *BMC Biology* **12**, 87 (2014).
48. M. Laurence, C. Hatzis, D. E. Brash, Common Contaminants in Next-Generation Sequencing That Hinder Discovery of Low-Abundance Microbes. *PLOS ONE* **9**, e97876 (2014).
49. WHO, "The Gap Report," (Joint United Nations Programme on HIV/AIDS (UNAIDS), 2014).
50. WHO, "UNAIDS Report on the Global AIDS Epidemic 2013," *Global Report* (Joint United Nations Programme on HIV/AIDS (UNAIDS), 2013).
51. M. Badri, R. Ehrlich, R. Wood, T. Pulerwitz, G. Maartens, Association between Tuberculosis and HIV Disease Progression in a High Tuberculosis Prevalence Area. *Int J Tuberc Lung Dis* **5**, 225-232 (2001).
52. R. E. Chaisson, N. A. Martinson, Tuberculosis in Africa—Combating an HIV-driven Crisis. *N Engl J Med* **358**, 1089-1092 (2008).
53. N. French, S. B. Gordon, T. Mwalukomo, S. A. White, G. Mwafulirwa, H. Longwe, M. Mwaiponya, E. E. Zijlstra, M. E. Molyneux, C. F. Gilks, A Trial of a 7-valent Pneumococcal Conjugate Vaccine in HIV Infected Adults. *N Engl J Med* **362**, 812-822 (2010).
54. J. L. Flanagan, E. L. Brodie, L. Weng, S. V. Lynch, O. Garcia, R. Brown, P. Hugenholtz, T. Z. DeSantis, G. L. Andersen, J. P. Wiener-Kronish, J. Bristow, Loss of Bacterial Diversity during Antibiotic Treatment of Intubated Patients Colonized with Pseudomonas aeruginosa. *J Clin Microbiol* **45**, 1954-1962 (2007).
55. M. J. Anderson, A New Method for Non-parametric Multivariate Analysis of Variance. *Austral Ecology* **26**, 32-46 (2001).
56. C. Lozupone, R. Knight, UniFrac: A New Phylogenetic Method for Comparing Microbial Communities. *Appl Environ Microbiol* **71**, 8228-8235 (2005).
57. S. T. Rutherford, B. L. Bassler, Bacterial Quorum Sensing: Its Role in Virulence and Possibilities for its Control. *Cold Spring Harb Perspect Med* **2**, (2012).
58. V. Antonic, A. Stojadinovic, B. Zhang, M. J. Izadjoo, M. Alavi, Pseudomonas aeruginosa Induces Pigment Production and Enhances Virulence in a White Phenotypic Variant of Staphylococcus aureus. *Infect Drug Resist* **6**, 175-186 (2013).
59. I. Holmes, K. Harris, C. Quince, Dirichlet Multinomial Mixtures: Generative Models for Microbial Metagenomics. *PLoS One* **7**, e30126 (2012).
60. M. G. I. Langille, J. Zaneveld, J. G. Caporaso, D. McDonald, D. Knights, J. A. Reyes, J. C. Clemente, D. E. Burkepile, R. L. Vega Thurber, R. Knight, R. G. Beiko, C. Huttenhower, Predictive Functional Profiling of Microbial Communities using 16S rRNA Marker Gene Sequences. *Nat Biotechnol* **31**, 814-821 (2013).

61. M. Kanehisa, S. Goto, KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* **28**, 27-30 (2000).
62. M. Kanehisa, S. Goto, Y. Sato, M. Kawashima, M. Furumichi, M. Tanabe, Data, Information, Knowledge and Principle: Back to Metabolism in KEGG. *Nucleic Acids Research* **42**, D199-D205 (2014).
63. C. K. Rahman AN, Mujib S, Fong IW, Ostrowski MA, TIM-3 and Its Immunoregulatory Role in HIV Infection. *Journal of Clinical & Cellular Immunology* **7**, 007 (2012).
64. R. B. Jones, L. C. Ndhlovu, J. D. Barbour, P. M. Sheth, A. R. Jha, B. R. Long, J. C. Wong, M. Satkunarajah, M. Schweneker, J. M. Chapman, G. Gyenes, B. Vali, M. D. Hyrcza, F. Y. Yue, C. Kovacs, A. Sassi, M. Loutfy, R. Halpenny, D. Persad, G. Spotts, F. M. Hecht, T.-W. Chun, J. M. McCune, R. Kaul, J. M. Rini, D. F. Nixon, M. A. Ostrowski, Tim-3 Expression Defines a Novel Population of Dysfunctional T cells with Highly Elevated Frequencies in Progressive HIV-1 Infection. *J Exp Med* **205**, 2763-2779 (2008).
65. L. Golden-Mason, B. E. Palmer, N. Kassam, L. Townshend-Bulson, S. Livingston, B. J. McMahon, N. Castelblanco, V. Kuchroo, D. R. Gretch, H. R. Rosen, Negative Immune Regulator Tim-3 is Overexpressed on T cells in Hepatitis C Virus Infection and its Blockade Rescues Dysfunctional CD4+ and CD8+ T cells. *J Virol* **83**, 9122-9130 (2009).
66. S. Muller, M. Radic, Citrullinated Autoantigens: From Diagnostic Markers to Pathogenetic Mechanisms. *Clin Rev Allergy Immunol*, (2014).
67. L. Cha, C. M. Berry, D. Nolan, A. Castley, S. Fernandez, M. A. French, Interferon-alpha, Immune Activation and Immune Dysfunction in Treated HIV Infection. *Clin Trans Immunol* **3**, e10 (2014).
68. K. E. Burhop, W. M. Selig, A. B. Malik, Monohydroxyeicosatetraenoic Acids (5-HETE and 15-HETE) Induce Pulmonary Vasoconstriction and Edema. *Circ Res* **62**, 687-698 (1988).
69. K. Mobraten, T. Haugbro, E. Karlstrom, C. R. Kleiveland, T. Lea, Activation of the Bile Acid Receptor TGR5 Enhances LPS-induced Inflammatory Responses in a Human Monocytic Cell Line. *J Recept Signal Transduct Res*, 1-8 (2014).
70. J. C. Gower, Generalized procrustes analysis. *Psychometrika* **40**, 33-51 (1975).
71. J. G. Caporaso, J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Peña, J. K. Goodrich, J. I. Gordon, G. A. Huttley, S. T. Kelley, D. Knights, J. E. Koenig, R. E. Ley, C. A. Lozupone, D. McDonald, B. D. Muegge, M. Pirrung, J. Reeder, J. R. Sevinsky, P. J. Turnbaugh, W. A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld, R. Knight, QIIME Allows Analysis of High-throughput Community Sequencing Data. *Nat Methods* **7**, 335-336 (2010).
72. N. Mantel, The Detection of Disease Clustering and a Generalized Regression Approach. *Cancer Res* **27**, 209-220 (1967).
73. E. Kolwijck, F. L. van de Veerdonk, The Potential Impact of the Pulmonary Microbiome on Immunopathogenesis of Aspergillus-related Lung Disease. *Eur J Immunol* **44**, 3156-3165 (2014).
74. T. Homma, A. Kato, B. Bhushan, J. E. Norton, L. A. Suh, R. G. Carter, D. S. Gupta, R. P. Schleimer, Aspergillus fumigatus Activates PAR-2 and Skews Toward a Th2 Bias in Airway Epithelial Cells. *Am J Respir Cell Mol Biol*, (2015).

75. M. Urb, B. D. Snarr, G. Wojewodka, M. Lehoux, M. J. Lee, B. Ralph, M. Divangahi, I. L. King, T. K. McGovern, J. G. Martin, R. Fraser, D. Radzioch, D. C. Sheppard, Evolution of the Immune Response to Chronic Airway Colonization with Aspergillus fumigatus Hyphae. *Infect Immun* **83**, 3590-3600 (2015).
76. I. Sada-Ovalle, L. Chávez-Galán, L. Torre-Bouscoulet, L. Nava-Gamiño, L. Barrera, P. Jayaraman, M. Torres-Rojas, M. A. Salazar-Lezama, S. M. Behar, The Tim3-Galectin 9 Pathway Induces Antibacterial Activity in Human Macrophages Infected with Mycobacterium tuberculosis. *J Immunol* **189**, 5896-5902 (2012).
77. O. Kilsgård, P. Andersson, M. Malmsten, S. L. Nordin, H. M. Linge, M. Eliasson, E. Sörenson, J. S. Erjefält, J. Bylund, A. I. Olin, O. E. Sørensen, A. Egesten, Peptidylarginine Deiminases Present in the Airways during Tobacco Smoking and Inflammation can Citrullinate the Host Defense Peptide LL-37, Resulting in Altered Activities. *Am J Respir Cell Mol Biol* **46**, 240-248 (2012).
78. T. Magoč, S. L. Salzberg, FLASH: Fast Length Adjustment of Short Reads to Improve Genome Assemblies. *Bioinformatics* **27**, 2957-2963 (2011).
79. B. J. Haas, D. Gevers, A. M. Earl, M. Feldgarden, D. V. Ward, G. Giannoukos, D. Ciulla, D. Tabbaa, S. K. Highlander, E. Sodergren, B. Methé, T. Z. DeSantis, H. M. Consortium, J. F. Petrosino, R. Knight, B. W. Birren, Chimeric 16S rRNA Sequence Formation and Detection in Sanger and 454-pyrosequenced PCR Amplicons. *Genome Res* **21**, 494-504 (2011).
80. R_Development_Core_Team, R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, (2008).
81. T. Z. DeSantis, P. Hugenholtz, N. Larsen, M. Rojas, E. L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu, G. L. Andersen, Greengenes, a Chimera-checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Appl Environ Microbiol* **72**, 5069-5072 (2006).
82. T. D. Schmittgen, K. J. Livak, Analyzing Real-time PCR Data by the Comparative CT Method. *Nat. Protocols* **3**, 1101-1108 (2008).
83. Y. Vázquez-Baeza, M. Pirrung, A. Gonzalez, R. Knight, EMPeror: A Tool for Visualizing High-throughput Microbial Community Data. *Gigascience* **2**, 16 (2013).
84. J. R. Bray, J. T. Curtis, An Ordination of Upland Forest Communities of Southern Wisconsin. *Ecological Monographs* **27**, 325-349 (1957).
85. C. Lozupone, M. Hamady, R. Knight, UniFrac—An Online Tool for Comparing Microbial Community Diversity in a Phylogenetic Context. *BMC Bioinformatics* **7**, 371 (2006).
86. J. D. Estes, L. D. Harris, N. R. Klatt, B. Tabb, S. Pittaluga, M. Paiardini, G. R. Barclay, J. Smedley, R. Pung, K. M. Oliveira, V. M. Hirsch, G. Silvestri, D. C. Douek, C. J. Miller, A. T. Haase, J. Lifson, J. M. Brenchley, Damaged Intestinal Epithelial Integrity Linked to Microbial Translocation in Pathogenic Simian Immunodeficiency Virus Infections. *PLoS Pathog* **6**, e1001052 (2010).
87. A. Nazli, O. Chan, W. N. Dobson-Belaire, M. Ouellet, M. J. Tremblay, S. D. Gray-Owen, A. L. Arsenault, C. Kaushic, Exposure to HIV-1 Directly Impairs Mucosal Epithelial Barrier Integrity Allowing Microbial Translocation. *PLoS Pathog* **6**, e1000852 (2010).

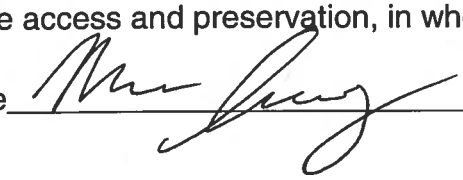
88. R. Romero, S. S. Hassan, P. Gajer, A. L. Tarca, D. W. Fadrosh, L. Nikita, M. Galuppi, R. F. Lamont, P. Chaemsaithong, J. Miranda, T. Chaiworapongsa, J. Ravel, The Composition and Stability of the Vaginal Microbiota of Normal Pregnant Women is Different from that of Non-pregnant Women. *Microbiome* **2**, 4 (2014).
89. K. Allers, M. Fehr, K. Conrad, H. J. Epple, D. Schurmann, A. Geelhaar-Karsch, K. Schinnerling, V. Moos, T. Schneider, Macrophages Accumulate in the Gut Mucosa of Untreated HIV Infected Patients. *J Infect Dis* **209**, 739-748 (2014).
90. K. E. Fujimura, T. Demoor, M. Rauch, A. A. Faruqi, S. Jang, C. C. Johnson, H. A. Boushey, E. Zoratti, D. Ownby, N. W. Lukacs, S. V. Lynch, House Dust Exposure Mediates Gut Microbiome Lactobacillus Enrichment and Airway Immune Defense Against Allergens and Virus Infection. *Proc Natl Acad Sci U S A* **111**, 805-810 (2014).
91. K. E. Fujimura, A. R. Sitarik, S. Havstad, D. L. Lin, S. Levan, D. Fadrosh, A. R. Panzer, B. LaMere, E. Rackaityte, N. W. Lukacs, G. Wegienka, H. A. Boushey, D. R. Ownby, E. M. Zoratti, A. M. Levin, C. C. Johnson, S. V. Lynch, Neonatal Gut Microbiota Associates with Childhood Multisensitized Atopy and T cell Differentiation. *Nat Med* **22**, 1187-1191 (2016).
92. J. Nakayama, K. Watanabe, J. Jiang, K. Matsuda, S.-H. Chao, P. Haryono, O. La-Ongkham, M.-A. Sarwoko, I. N. Sujaya, L. Zhao, K.-T. Chen, Y.-P. Chen, H.-H. Chiu, T. Hidaka, N.-X. Huang, C. Kiyohara, T. Kurakawa, N. Sakamoto, K. Sonomoto, K. Tashiro, H. Tsuji, M.-J. Chen, V. Leelavatcharamas, C.-C. Liao, S. Nitisinprasert, E. S. Rahayu, F.-Z. Ren, Y.-C. Tsai, Y.-K. Lee, Diversity in Gut Bacterial Community of School-Age Children in Asia. *Sci Rep* **5**, 8397 (2015).
93. S. Ruengsomwong, O. La-Ongkham, J. Jiang, B. Wannissorn, J. Nakayama, S. Nitisinprasert, Microbial Community of Healthy Thai Vegetarians and Non-Vegetarians, Their Core Gut Microbiota, and Pathogen Risk. *Journal of microbiology and biotechnology* **26**, 1723-1735 (2016).
94. L. Chen, Z. Zhang, K. E. Barletta, M. D. Burdick, B. Mehrad, Heterogeneity of Lung Mononuclear Phagocytes during Pneumonia: Contribution of Chemokine Receptors. *American Journal of Physiology - Lung Cellular and Molecular Physiology* **305**, L702-L711 (2013).
95. L. K. Johnston, C. R. Rims, S. E. Gill, J. K. McGuire, A. M. Manicone, Pulmonary Macrophage Subpopulations in the Induction and Resolution of Acute Lung Injury. *American Journal of Respiratory Cell and Molecular Biology* **47**, 417-426 (2012).
96. Z. D. Swan, A. L. Bouwer, E. R. Wonderlich, S. M. Barratt-Boyes, Persistent Accumulation of Gut Macrophages with Impaired Phagocytic Function Correlates with SIV Disease Progression in Macaques. *Eur J Immunol*, (2017).
97. M. Levy, C. A. Thaiss, D. Zeevi, L. Dohnalová, G. Zilberman-Schapira, J. A. Mahdi, E. David, A. Savidor, T. Korem, Y. Herzig, M. Pevsner-Fischer, H. Shapiro, A. Christ, A. Harmelin, Z. Halpern, E. Latz, R. A. Flavell, I. Amit, E. Segal, E. Elinav, Microbiota-Modulated Metabolites Shape the Intestinal Microenvironment by Regulating NLRP6 Inflammasome Signaling. *Cell* **163**, 1428-1443 (2015).

98. U. Kõljalg, R. H. Nilsson, K. Abarenkov, L. Tedersoo, A. F. S. Taylor, M. Bahram, S. T. Bates, T. D. Bruns, J. Bengtsson-Palme, T. M. Callaghan, B. Douglas, T. Drenkhan, U. Eberhardt, M. Dueñas, T. Grebenc, G. W. Griffith, M. Hartmann, P. M. Kirk, P. Kohout, E. Larsson, B. D. Lindahl, R. Lücking, M. P. Martín, P. B. Matheny, N. H. Nguyen, T. Niskanen, J. Oja, K. G. Peay, U. Peintner, M. Peterson, K. Põldmaa, L. Saag, I. Saar, A. Schüßler, J. A. Scott, C. Senés, M. E. Smith, A. Suija, D. L. Taylor, M. T. Telleria, M. Weiss, K.-H. Larsson, Towards a Unified Paradigm for Sequence-based Identification of Fungi. *Molecular Ecology* **22**, 5271-5277 (2013).
99. C. Lozupone, R. Knight, UniFrac: A New Phylogenetic Method for Comparing Microbial Communities. *Appl Environ Microbiol* **71**, (2005).
100. Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B Methodol* **57**, (1995).

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Author Signature  Date 10/20/2017