DECODING EARLY IMMUNE EVENTS IN NON-HUMAN PRIMATES INFECTED WITH *MYCOBACTERIUM TUBERCULOSIS*

by

Anthony Michael Cadena

B.A. Human Biology, University of Virginia, 2010

Submitted to the Graduate Faculty of

the School of Medicine for the degree of

Doctor of Philosophy

University of Pittsburgh

2017

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE

This dissertation was presented

by

Anthony Michael Cadena

It was defended on

January 27, 2017

and approved by

Jennifer Bomberger, Ph.D., Assistant Professor of Microbiology and Molecular Genetics

Lawrence P. Kane, Ph.D., Associate Professor of Immunology

Elodie Ghedin, Ph.D., Professor of Biology, New York University

Karen A. Norris, Ph.D., Professor of Infectious Diseases, GRA Eminent Scholar in Translational Medicine and Immunology, College of Veterinary Medicine, University of Georgia

Dissertation Advisor: JoAnne L. Flynn, Ph.D., Professor of Microbiology and Molecular Genetics, Medicine, and Immunology

DECODING EARLY IMMUNE EVENTS IN NON-HUMAN PRIMATES INFECTED WITH *MYCOBACTERIUM TUBERCULOSIS*

Anthony Michael Cadena, Ph.D.

University of Pittsburgh, 2017

Tuberculosis (TB) continues to pose a significant health risk to morbidity and mortality worldwide. *Mycobacterium tuberculosis,* the causative agent of TB, is responsible for nearly 10 million new cases of active disease and 2 million deaths annually. While the majority of *M. tuberculosis* infected individuals are asymptomatic (termed latent TB) and contain the infection, a subset (~10%) of infected individuals either present initially with primary active disease or reactivate subsequently over the course of their lifetime. The precise immune mechanisms responsible for this observed spectrum remain unclear but recent evidence suggests that early events in *M. tuberculosis* infection influence host outcome. In this dissertation, we utilized established non-human primate (NHP) models of *M. tuberculosis* to examine the early immunologic, pathologic, and contextual responses following infection. The primary aim of this thesis was to develop a novel genomic barcoding approach to add to our *in vivo* toolbox permitting single-bacterial tracing to probe early events in a variety of infection contexts. Our work validated the use of these bacterial tags and provided a unique ability to quantitatively track individual founding bacilli and their descendants in infected macaques. We found that the majority of bacteria are able to establish infection (i.e. a primary granuloma) but only a subset of bacteria contributes to productive dissemination. In addition, our barcode strategy permitted reinfection studies in which primary and secondary infections are separately evaluated using library-specific identifiers. Our initial observations suggest that an ongoing primary infection

substantially limits secondary granuloma establishment and bacterial growth. By adapting our current NHP model of TB with new genomic barcoding tools, our work has provided insight into bacterial dissemination, reinfection, and host variability. Finally, our most recent studies are looking into the earliest context of the lung by probing the lung microbiome and its interaction with *M. tuberculosis*. Our latest observations suggest that the microbial lung landscape is highly variable across individuals, is distinct from the oral cavity, and undergoes significant alterations following infection. Overall, this body of work reiterates the importance of appreciating the influence that early infection and single lesion dynamics contributes to host outcome.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I started my doctoral adventure about five years ago and it has not disappointed. It has certainly been a challenging road but it also has been incredibly rewarding. I was very fortunate to have superb guidance and assistance along the way.

First and foremost, I would like to thank my mentor, Dr. JoAnne Flynn, for sharing her time, insight, and expertise. JoAnne has had an incredible impact on my career as a scientist and on my development as a responsible, critical, and thoughtful investigator. I have evolved considerably over the years and JoAnne has been instrumental in guiding me forward, helping me to network, and in establishing my publication record. It has truly been an honor and a privilege working in her laboratory and with such a wonderful scientist and individual.

Secondly, I would like to thank Dr. Sarah Fortune. She has been a superb collaborator and a great mentor as well. Sarah welcomed me into her Boston laboratory numerous times as we worked to barcode *M. tuberculosis*. She was also extremely helpful in helping me to finalize my postdoctoral endeavors and secure a position in Boston. I look forward to keeping in touch as I'll be just a few buildings away.

Lastly, I would like to thank my committee members, Karen, Elodie, Jen and Larry. They have been very helpful throughout my graduate experience. They have all provided valuable insight and critiques during my dissertation meetings and on my final thesis. I am very grateful for their input.

1.0 A BRIEF INTRODUCTION INTO *MYCOBACTERIUM TUBERCULOSIS,* TB BIOLOGY AND EARLY INFECTION

Sections of the chapter are adapted from the original publication:

Cadena AM, Flynn JL, & Fortune SM (2016) The Importance of First Impressions: Early Events in *Mycobacterium tuberculosis* Infection Influence Outcome. *mBio* 7(2):e00342-00316.

&

from the review in preparation:

Cadena AM, Fortune SM, & Flynn JL (2017) Heterogeinity in Tuberculosis: the Importance of Considering Local Conflicts in a Global Contex.

1.1 THE BUG: ITS PERSISTENCE, SUCCESS, AND HOST INTERACTION

1.1.1 *Mycobacterium tuberculosis,* an old threat remains

Tuberculosis remains a global health threat. In spite of an existing drug regimen and the *M. bovis* bacilli Calmette-Guerin (BCG) vaccine, there were 10.4 million incident cases of TB and 1.8 million deaths in 2015[1]. The greatest global burden for active TB lies in sub-Saharan Africa, India, and China where effective TB control is hampered by co-infection, poverty, inadequate disease surveillance, questionable BCG vaccine efficacy, and increasing drug resistance[1-4]. In

addition to these socio-economic barriers, our incomplete understanding of the human immune response further drives TB resurgence as a global infectious disease. The lack of true immune correlates of protection to *M. tuberculosis* infection and disease[5-7] coupled with the heterogeneity of the human immune response to TB[8,9], impinge on the development of novel vaccines and anti-mycobacterial drugs.

1.1.2 Host outcome in TB, beyond the two-state paradigm

Classically, *Mycobacterium tuberculosis* infection is thought to result in one of two states: latent infection or active disease. The former outcome is a clinically silent process defined by immunological containment and bacterial persistence and accounts for ninety percent of human infections with an estimated burden of approximately 2 billion individuals worldwide[10,11]. By contrast, active disease is classified as a failed host response manifesting primarily as pulmonary TB typified by chronic cough, fever, sustained weight loss, wasting, and hemoptysis[12]. In 2015, there were 10.4 million new cases of illness and 1.8 million deaths[1] due to TB; this active disease state is estimated to occur in approximately 5% of initial infections in the first 18 months with a remaining risk of 5% over the course of one's life[13]. More recently, the TB field has embraced a newer paradigm that recognizes a spectrum of clinical outcomes within these two states[7,14,15]. The biological and immunological underpinning for this variability is not well understood and remains an open question in human TB.

The large reservoir of asymptomatic latent infection in the human population represents a nuanced continuum of bacterial persistence and host containment ranging from cleared infection to low-grade tuberculosis[14,15] (**Figure 1**). This viewpoint extends the definition of subclinical, latent tuberculosis beyond a single status to better differentiate risk of reactivation, prioritize

preventive treatment, and emphasize host heterogeneity to *M. tuberculosis* infection. Individuals that have sterilized infection would be least likely to reactivate and presumably last to require intervention. By contrast, latently infected individuals who are harboring a low-grade, percolating infection would be at higher risk to reactivate and more likely to require treatment. In addition to better segregating individuals at higher risk of reactivation, this perspective provides greater appreciation for the dynamic nature of *M. tuberculosis* infection and the human immune response to TB, which, in turn, affords better explanation for those individuals who can control and clear infection as exhibited by persistent TST non-responders. These individuals, despite documented high-risk exposure to TB, are consistently TST negative[16-18] and suggest that intrinsic[19] and innate immune factors might be contributing to mycobacterial clearance in the absence of an adaptive immune response[20]. Lastly, this model reiterates the immunologic and pathologic variability underscoring all of these human outcomes, which is crucial to develop intervention strategies that recognize and target the unique, complex, and independent nature of the local host-pathogen interactions of this disease.

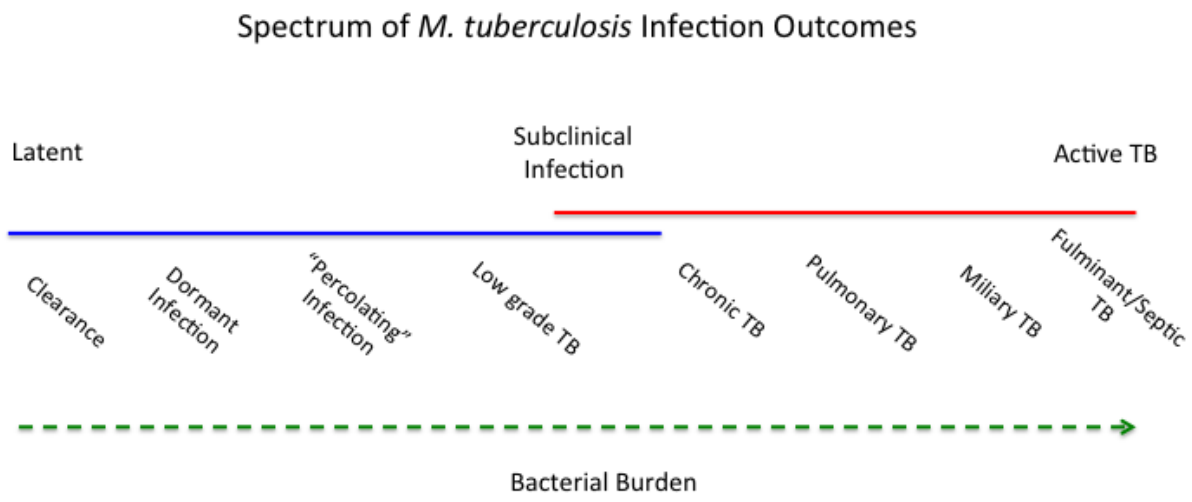


Figure 1. Tuberculosis presents along a spectrum of host outcomes.

Within the clinically-defined, binary classification of TB exists a more complex set of host outcomes to *M. tuberculosis* infection. This spectrum is driven by multiple factors but is thought to reflect differences in *M. tuberculosis* burden, a balance of pro (P)- and anti (A)-inflammatory cytokines, and pathology.

Evidence supporting this newer paradigm of human TB has been observed in human studies with 2-deoxy-2-[^{18}F]fluoro-D-glucose positron emission and computed tomography (FDG PET CT)[21]. The authors find that among 35 ART-naïve, HIV-1 positive LTBI adults, 10 patients with pulmonary irregularities indicative of subclinical disease had a significantly higher risk to present with active TB compared to the remaining 25 patients that had no subclinical pathology by FDG PET CT. These findings mirror data in the macaque model of tuberculosis in which FDG PET CT features including elevated lung inflammation and extrapulmonary sites of infection predicted reactivation risk in latently infected macaques prior to tumor necrosis factor (TNF) neutralization with 92% sensitivity and specificity[22]. Both of these studies validate the biological importance and subsequent clinical benefit in appreciating the variability of host outcomes in TB.

Clinical heterogeneity has also been observed in the context of recent studies examining blood transcriptional signatures related to host status. In 2010, Berry *et al* reported a 393-transcript signature distinct for active TB in both intermediate and high burden areas relative to latent TB subjects and healthy controls[23]. Interestingly, while the majority of latent individuals clustered independently, 10-25% of the LTBI subjects shared similar transcriptional profiles with active TB patients, and most likely highlighted individuals with subclinical, active disease. These shared transcriptional profiles observed in a subset of clinically-defined latent TB subjects reiterate the radiological findings above, and further emphasize the varied nature of this disease, which can even be distinguished in the periphery. More recently, additional gene expression analysis of pulmonary and extra-pulmonary TB patients found that transcriptional profiles are influenced by symptom status and site of disease[24]. Individuals with extra-pulmonary disease exhibited heterogeneous gene signatures in which the magnitude of response was related to the presence of symptoms and site; those individuals with the highest mean molecular distance to health had the highest percentage of patients within the group having one or more symptom of either fever, night sweats, chest pain, cough, or weight loss. These findings suggest that illness is likely linked to the site of infection, bacterial burden, and site host response, which in turn contribute to the resulting diversity of transcriptional profile and ultimately host status. In non-human primates, similar transcriptional studies have interrogated both lung-specific granuloma changes[25] as well as longitudinal changes in the blood[26]. In the first study, granulomas from rhesus macaques were found to undergo transcriptional reprogramming between early (4weeks) and late (13weeks) time points in infection[25]. In the second study, peripheral analyses in cynomolgus macaques revealed the greatest transcriptional change between 3 and 8 weeks post infection[26]. Both of these profiling studies reiterate the variable and dynamic states occurring

both locally and systemically throughout infection and reinforce the importance that early events play in shaping clinical outcome[27].

1.1.3 The granuloma, a dynamic structure of bacterial containment and persistence

The hallmark of human TB is the granuloma, an organized and localized aggregate of immune cells that consists of macrophages, lymphocytes, and other host immune cells and forms in response to persistent stimuli[28]. In TB, these structures arise at the sites of infection and represent an active interchange between the host and pathogen. The formation of a granuloma is crucial to adequately control and contain infection[29,30], but may also contribute to early *M. tuberculosis* proliferation and dissemination[31-33]. Human studies from over a half-century ago have revealed that in active disease and latent infection, granulomas exhibit morphological heterogeneity[7,34]. In addition to the classic caseous granuloma, granulomas can be non-necrotizing, neutrophil-rich, mineralized, or completely fibrotic[30]. In nearly all of these instances, the basic granuloma architecture exhibits the following layout: a central acellular necrotic core, termed caseum, surrounded by a diverse population of macrophages, which is itself circumscribed by a lymphocytic cuff of CD4 and $CD8^+$ T cells and B cells, and may have a peripheral fibrotic edge[29,32]. Granulomas primarily contain macrophages at various stages of activation and T and B cells, but can also contain neutrophils, dendritic cells (DCs), and fibroblasts[30,32] (**Figure 2**). The important lesson learned from appreciating the heterogeneity of human TB granulomas is that each separate lesion represents a localized microenvironment that can be independently influenced by the quality of the localized immune response, the pathogenicity, state, and prevalence of the bacteria, the extent of immunopathology, and the overall host disease status[35,36]. The interaction of all of these factors in the granuloma has yet to

be determined but does allow for speculation into how heterogeneous disease states can arise from varying lesion-specific dynamics, particularly in the early stages of infection.

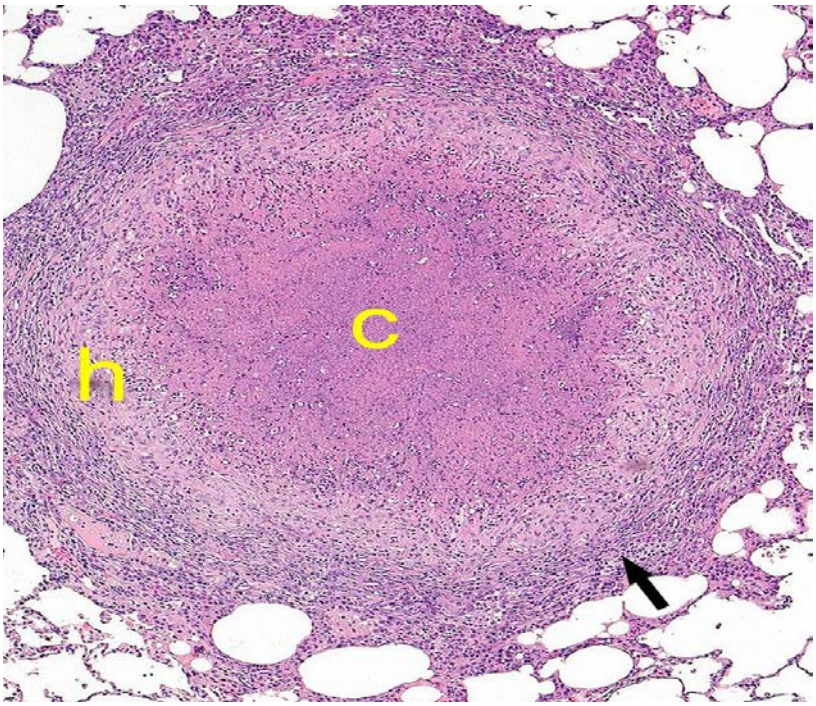


Figure 2. Archetypical caseous granuloma.

A classic caseous granuloma from an infected macaque. It is organized into an inner region (**c**) of necrotic caseum surrounded by a region (**h**) of histocytic macrophages that is encapsulated (granuloma periphery) by a cuff of T and B lymphocytes. This lesion has a peripheral region of fibrosis (black arrow). Image courtesy Edwin Klein, VDM.

1.2 FIRST IMPRESSIONS MATTER: EARLY EVENTS IN *MYCOBACTERIUM TUBERCULOIS* INFLUENCE OUTCOME

New vaccines against *M. tuberculosis* are essential for preventing infection, disease and transmission. However, the host immune responses induced by an effective vaccine remain

unclear[6]. Increasingly it has become clear that early events including early innate responses in infection are of major importance in the eventual outcome of the infection[37,38]. Studying such events in humans is challenging, as they are occurring within the lung and thoracic lymph nodes, and any clinical signs of early infection are relatively non-specific. Development of an effective vaccine requires a clear understanding of the successful (and detrimental) early host responses against *M. tuberculosis*, with the goal to improve upon natural immune responses and prevent infection or disease.

M. tuberculosis infection generally progresses quite slowly, with active TB being diagnosed most commonly within the first two years of infection. People often do not know when they are exposed or infected and therefore, studying early events of infection in humans can be challenging. Nonetheless, some studies support that the earliest events in *M. tuberculosis* infection are critically important in dictating clinical outcome. First and foremost, the inoculum dose in animal models from mice to non-human primates influences the severity of infection, so that more bacteria delivered to the lungs results in a worse outcome. In addition, data from nonhuman primate studies suggest that between 3-6 weeks post-infection, one can predict whether the animal will progress to active TB disease or remain latently infected 6 to 9 months later, based on the extent of early dissemination, serologic (ESR) and PET-CT features suggestive of more extensive inflammation[39]. Thus, clinical course is at least in part dictated by the early innate and adaptive immune responses. Indeed, efforts to understand the basis of protective immunity to *M. tuberculosis* infection are pushing us to look even earlier into the course of infection. Clinical studies suggest that there are individuals who are highly exposed to *M. tuberculosis* but remain persistently tuberculin skin test negative, and thus presumably uninfected[16,18]. These individuals raise the possibility that at still earlier points in the course of

infection, it is possible for the host to fully clear the bacteria and that this early bacteriocidal response could be harnessed through vaccination.

1.2.1 Clinical Data on Early Infection

There are major gaps in our understanding of the early events in *M. tuberculosis* infection in humans. These stem from the difficulty in identifying the onset of infection, for which there are no good diagnostic tools. The classic diagnostic test for *M. tuberculosis* infection is the tuberculin skin test (TST), which is a delayed type hypersensitivity response to a crude mixture of *M. tuberculosis* proteins and lipids known as purified protein derivative (PPD). A more recent test for infection is the Interferon-Gamma Release Assays (IGRA), in which blood is stimulated with two *M. tuberculosis*-specific antigens and then assayed for IFN-γ by ELISA or ELISPOT. Both the TST and IGRA measure the T cell response to mycobacterial antigens, and therefore are not positive until a measurable T cell response is induced. Induction of measurable T cell responses to *M. tuberculosis* infection can be quite slow in tuberculosis. It occurs in humans approximately 6 weeks post-infection[40]. In non-human primates and in guinea pigs, TST conversion occurs 4-8 weeks after exposure and infection[41,42]. Although TST is not usually performed in mice, an older study showed that PPD+ skin test responses were observed at 4-6 weeks post-infection[43]. However, *M. tuberculosis*-specific T cell responses in lymphoid tissues can be measured as early as 2 weeks post-infection[44]. Conversion of a negative to positive TST or IGRA result denotes recent exposure and infection. Frequent TST or IGRA testing is necessary in clinical settings to establish an approximate time of infection, but by the time a T cell response is measurable, the early events in infection have already occurred. The lack of simple diagnostic tests that can be used to identify infected individuals immediately after

infection makes it difficult to conduct rigorous studies on the course of early infection in humans.

Despite these limitations, landmark studies were conducted by Poulsen in the Faroe Islands in the pre-treatment era of the 1930s and 1940s that provided critical insights into initial infection in humans[40,45]. In these studies, a version of tuberculin skin testing was performed on nearly all of the 30,000 residents of the numerous villages in the Faroe Islands, and nearly all persons with active TB were known. Thus, the local epidemiology of TB in each village was well documented, and since most villages were small, tracking individuals and obtaining detailed histories was possible. New TST conversions were followed up closely, with particular attention paid to identifying the time of infection by determining the index case and the duration of exposure to that case by the newly infected individual, as well as clinical signs post-infection, including X-rays and fluoroscopy. The first detailed description of these studies is a fascinating series of "case reports" documenting the duration of and time since exposure to an index case, skin test conversion, subsequent clinical manifestations, and outcome[40]. In some cases, exposure to a person with active TB for less than 24 hours resulted in TST conversion and subsequent development of primary TB. TST conversion was generally evident by 6 weeks post-exposure. Interestingly, nearly all of the reported cases experienced a fever (termed "initial fever") around the time of skin test conversion, and follow up indicated hilar adenopathy, with a number of subsequent cases of active TB and deaths.

In a follow up study, Poulson rigorously analyzed early symptoms of infection and outcome[45]. A total of 232 subjects (children and adults) who had TST conversions within the 6 month interval of testing were followed for several years. An additional 285 subjects who were selected as presenting with an initial fever in conjunction with recent skin test conversion were

also followed. Of the 232 "unselected" subjects, 63% had an initial fever after infection, and the frequency was similar between adults and children. The fever was of variable intensity and duration, though this was not predictive of disease outcome. These findings suggest an initial inflammatory process in the majority of those infected with *M. tuberculosis*, regardless of infection trajectory (i.e. primary disease or containment). Poulson documented other signs of initial inflammation associated with infection including elevated erythrocyte sedimentation rate (ESR), a non-specific sign of inflammation and erythema nodosum, an inflammatory process often linked to mycobacterial infection. Elevated ESR was coincident with initial fever, and in most cases returned to normal within 2 months. Even with the X-ray and fluoroscopy technology available at the time, hilar adenopathy, presumably signifying thoracic lymph node enlargement, was identified in 55% of subjects unselected for fever in the first two months post-infection, with little increase in hilar lymph node involvement after this time. Pulmonary infiltrates were observed within the first year in 27% of converters, which agrees with a separate study in Norway[46]. These infiltrates were usually seen within 3 months of TST conversion and usually unilateral. Most of these infiltrates regressed over the next several months, and only 15% of those with infiltrates progressed to active TB. However, when separated between adults and children, the authors find that 2% of children with infiltrates developed active disease whereas 25% of adult converters with infiltrates progressed to TB suggesting that early pulmonary infiltrates in adult converters is linked with disease progression[45].

Importantly, the earliest stages of infection may also be associated with bacterial carriage in sputum. For example, a recent study employing active case finding to estimate the prevalence of TB in household contacts revealed a stunningly high rate of asymptomatic carriage of *M. tuberculosis* bacteria in the sputum in household contacts, presumably more recently infected by

the incident case[47]. Similarly, pediatric studies suggest that there can be an early period of bacterial "excretion" after infection, which subsequently can resolve and does not necessarily herald eventual tuberculosis disease[48]. This supports a model in which an early period of bacterial growth is relatively common, even where the infection will subsequently be controlled.

These studies provide evidence of an early evolution of the infection in the majority of those infected, including an inflammatory process, evidence of thoracic lymph node involvement and potentially also the presence of culturable bacilli in the airways or sputum. Thus, the initial events in humans are not usually "silent" and suggest that the host immune response to initial infection is relatively robust. As most infections do not progress to active TB, this immune response is often successful in restraining the infection, although it is apparently a matter of months before this containment is complete.

These findings from human studies are recapitulated by our studies in macaque models of tuberculosis. Cynomolgus macaques develop active TB or latent infection, defined clinically, following infection with <25 CFU *M. tuberculosis* strain Erdman[42,49]. Using PET/CT imaging with fluorodeoxyglucose (FDG) as a probe, immunological assays, and clinical assessments, we have demonstrated that all macaques infected with a low dose of M. *tuberculosis* have an evolution of infection within the lungs, with granulomas visible by 2-3 weeks by PET/CT imaging[39,50]. Thoracic lymph nodes are often enlarged or show FDG avidity within a few weeks of infection. Using culture of gastric aspirates and bronchoalveolar lavage (BAL) as surrogates for sputum cultures, 30% of infected monkeys were found to shed culturable *M. tuberculosis* within 2 months of infection, indicating bacilli in airways during acute infection[49]. This was loosely correlated with outcome, with 90% of macaques that progress to active TB showing a positive gastric aspirate or BAL in the first two months of infection, compared to 44% of those

monkeys that would present with latent infection (unpublished data). ESR is very low in uninfected monkeys (normal <2mm). Increased ESR within 60 days of infection was strongly correlated with eventual progression to active TB; ESR>15 in the first 2 months of infection correctly predicted outcome with 92% accuracy (unpublished data). Coleman, *et al.* showed that formation of new granulomas between 3-6 weeks post-infection, and increased PET avidity in those granulomas, was associated with eventual development of active TB[39]. Although radiographs are less useful in macaques than in humans and not as sensitive as PET CT, our early studies using x-rays suggested that early pulmonary infiltrates were observed in 60% of our macaques[42], most of which went on to develop active disease. The macaque data support the human data on early inflammatory responses following infection, but also suggest that the final outcome of primary infection is determined early by the ability of the host to control infection in the granulomas and prevent early dissemination.

A more in depth study of granulomas from macaques demonstrated that the bacterial burden in the initial granulomas at 4 weeks post-infection is relatively high (~5 x 10^4 CFU) with minimal bacterial killing occurring[36]. However, as the adaptive immune response is induced, bacterial killing in granulomas increases, dramatically reducing bacterial burden, and in some cases sterilizing the granuloma. Indeed, even in monkeys that progress to active TB, an average of 30% of the granulomas are sterile[36]. The specific factors that drive an individual granuloma towards a particular fate remain poorly understood, but are likely to reflect the contributions of both host and bacterial factors.

1.2.2 Early Granuloma Formation and Bacterial Dissemination

What is known about the course of granuloma formation in people? Our understanding of early granuloma formation is rooted in results from human autopsy series published by Georges Canetti in the 1955. Based on his histopathologic examinations of 1500 autopsies, Canetti proposed that all tuberculous lesions proceed initially through an exudative and then a caseous phase. He suggested that lesions then sit at a "crossroads"–some remaining solid and ultimately sclerosing and others softening, where softening of the caseum, not caseation per se, was the harbinger of active disease[51]. At the time that he wrote, Canetti did not understand the basis of softening although in about half the cases, it was associated with an influx of neutrophils. In a discussion that remains highly relevant today, he wrote[51]: "This poses the question of whether an unknown factor orients the lesion from the outset toward liquefaction; that is, determines its fate from the earliest phases when the lesion may not be distinguished from another whose evolution will be 'normal.' On the other hand, one also often observes softening of lesions of long-standing radiologic existence; this is evidence of the potential danger of old caseous lesions."

Researchers have sought to better define and identify determinants of granuloma fate in a variety of experimental models. Data from these models has supported many of Canetti's observations (**Figure 3**). Exploration of early granuloma organization and development in guinea pigs has shown that within 21 days there is evidence of caseation preceding initiation of T cell responses implicating both innate mechanisms and bacterial processes in early cellular death[52]. Similarly, in macaques, the earliest granulomas isolated from the lungs (at 3 and 4 weeks post-infection) are primarily caseous in nature[53], where some of these lesions are expected

to progress but most will not, and indeed, some will be fully sterilized after the onset of adaptive immunity.

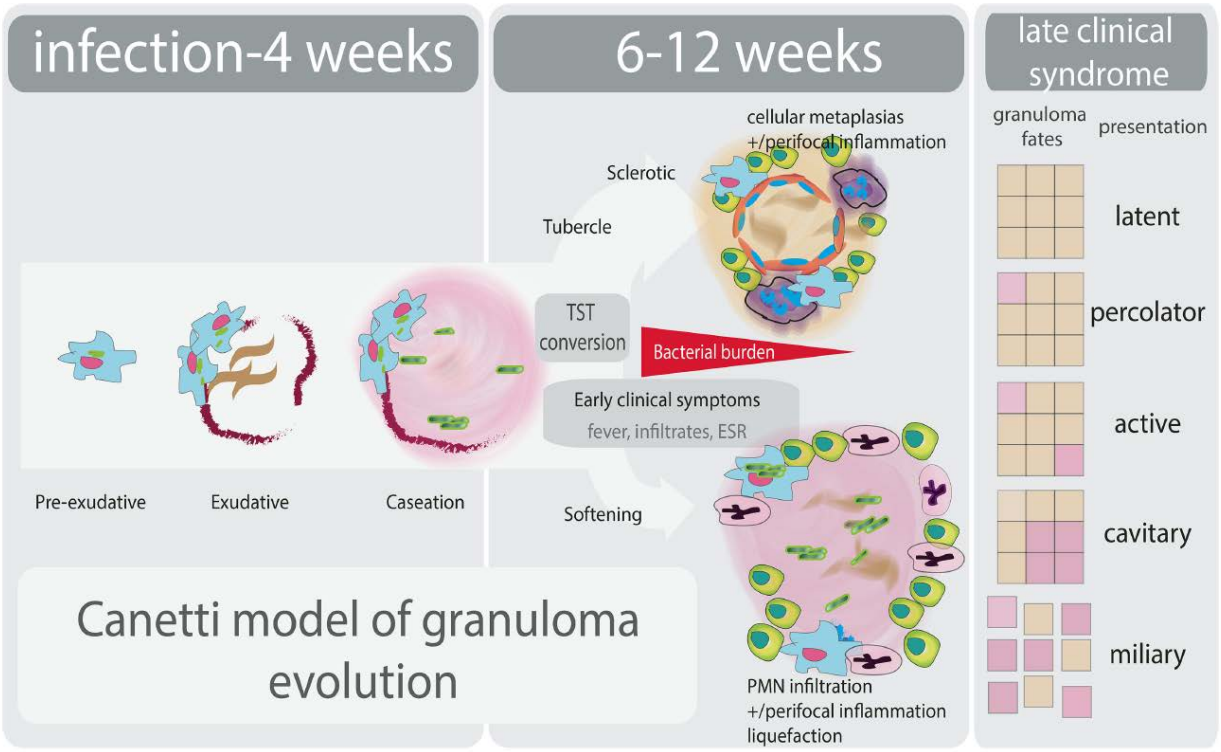


Figure 3. Canetti model of granuloma evolution.

Canetti granuloma formation and progression coupled with outcome of *M. tuberculosis* infection. Image used, with permission, Cadena *et al*[27]. Copyright (2016) American Association of Microbiology, USA.

The first steps in the development of the granuloma, and the path to early caseation, have been elucidated using *M. marinum* in zebrafish. These studies indicate that early macrophage death in the primary granuloma is linked to bacterial replication, dissemination, and secondary granuloma formation[31,32]. Specific modulation of macrophage apoptosis and necrosis by the ecosanoid pathway, PGE2 and LXA4, respectively, early in infection has been shown to influence infection outcome[54,55]. These studies extend the concept of the granuloma as a

protective host response to suggest that it also provides a niche for mycobacteria to establish robust infection.

The events that dictate the fate of the granuloma after early caseation are less well understood. *M. tuberculosis* expresses matrix metalloproteinases that can degrade collagen and facilitate tissue destruction[56-58]. The relative contribution of these bacterial virulence factors to the softening of the caseum that Canetti described, as opposed to caseation per se, remains unclear. It is possible that this reflects the effect of the enzymes responsible for the lipid mediators already implicated in TB disease, where alteration of activity is also expected to impact the linked lipid mediators of resolution of inflammation and neutrophil recruitment.

1.2.3 What are the Critical Cellular Events in Early Infection?

Our current understanding of the cellular events in early infection of *M. tuberculosis* draws on a number of *in vitro* model systems and *in vivo* small animal model studies. Inhalation of a small number of *M. tuberculosis* via droplet nuclei from a person with active tuberculosis results in deposition of the bacilli into the alveolar space and initial contact and phagocytosis by alveolar macrophages[59]. Following infection, an early influx of phagocytic cells including alveolar macrophages, neutrophils, and dendritic cells (DCs) arrive to this focus and begin to seed formation of a granuloma[7,60]. Studies using a *M. marinum* zebrafish model helped to characterize the earliest cells recruited to the site of infection and showed that macrophages, not neutrophils, were the dominant phagocytosing cells in the first 4 days post infection[61]. Murine studies with both *M. tuberculosis*[62] and BCG[63] suggest that there is early infection of alveolar macrophages, various populations of DCs, and neutrophils[64]. At 2 weeks post infection, the lungs of C57BL/6 mice infected with ~100 CFU of *M. tuberculosis* Erdman revealed that the

predominant populations in the lungs are Ly6G+ neutrophils and F4/80+ macrophages[64]. Further characterization of the kinetics of lung recruitment within the first 2 to 4 weeks of infection in mice infected with GFP labeled *M. tuberculosis* revealed a diverse and dynamic interplay between host cells and bacteria[62]. At 14 days post infection, myeloid DCs, alveolar macrophages, and neutrophils had the highest percentage of *M. tuberculosis* (GFP+) infected cells; by day 21, neutrophils and myeloid DCs had the highest percentages. Similar observations of the mediastinal LNs determined that myeloid DC migration was the primary infected source during the initial stages of infections[62]. An excellent overview of these early cellular interactions is reviewed and modeled by O'Garra *et al*[7] and summarized in **Figure 4**. These observations highlight the diverse range of early phagocytic responders to *M. tuberculosis* infection and implicate their initial influence on the progression of early disease and dissemination.

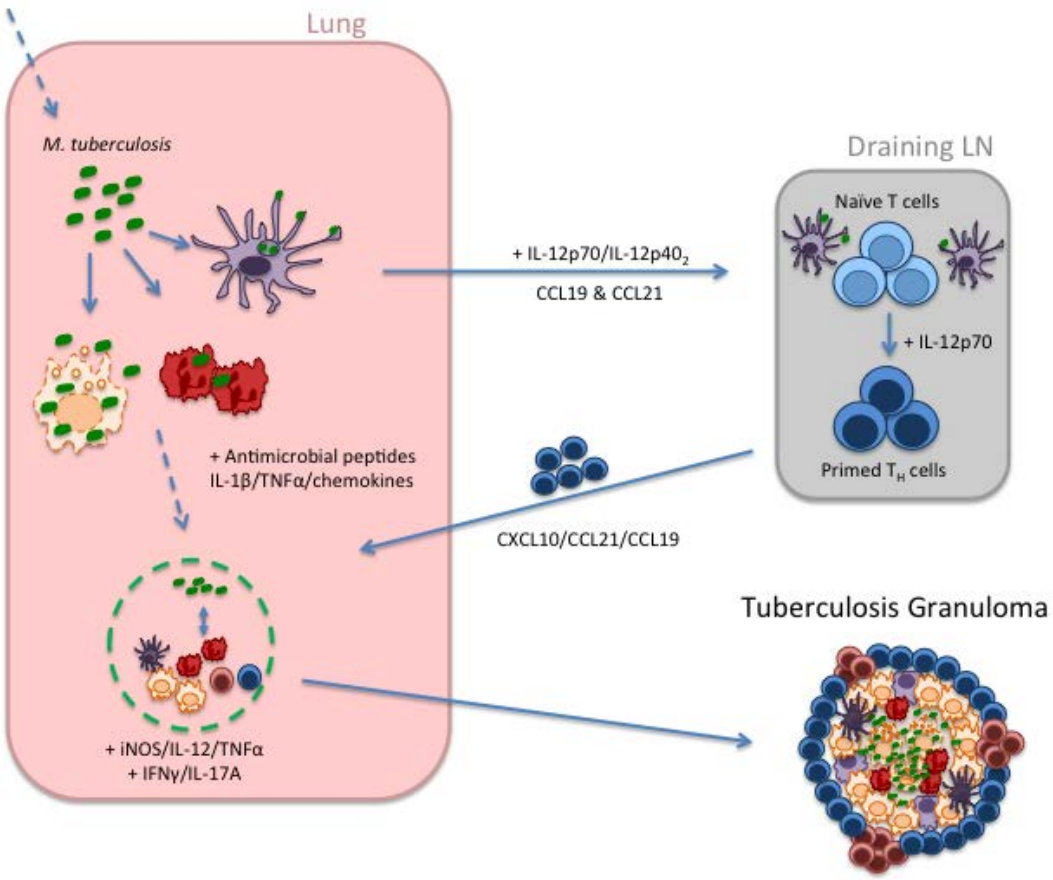


Figure 4. Early cellular response to *M. tuberculosis.*

An overview of the early cellular interactions between host and bacterium following *M. tuberculosis* infection that culminate in formation of the tuberculosis granuloma.

A recent review by Srivastava *et al* comprehensively assessed the diversity of these initial mononuclear cell subsets and detailed their contributions in tuberculosis[65]. The authors evaluated the relative differences of each subset's ability to both prime adaptive responses and control *M. tuberculosis* by reviewing each subset individually, drawing on *in vitro*, mouse, and human data. Ultimately, the authors posit that the inherent functional differences of each subset likely contribute to the overall outcome of infection following potential skewing of one subset over the others. Importantly, this hypothesis helps to bridge the observed diversity in cell types

with the resulting clinical variability observed in tuberculosis. Building on this hypothesis, we propose that the interplay of these diverse cell subsets and *M. tuberculosis* in the initial stages of infection contributes to granuloma fate and heterogeneity[36,66] which subsequently influences host outcome[39]. From this, we hypothesize that the local differences driven by early host-pathogen interactions manifest as a spectrum of granulomas with varying capacities for containment and bacterial killing, which ultimately leads to different clinical outcomes. As a point of speculation, it is possible that the activation status of the initial infected cell that seeds a granuloma could set the stage for inflammation, T cell recruitment, macrophage activation and eventual fate of that granuloma.

1.2.4 Neutrophil activity is a strong correlate of the outcome of human infection with M. tuberculosis

In recent years, it has become clear that neutrophil activity is a strong correlate of human TB disease state. Studies of gene expression in the blood of people with active and latent TB revealed a signature of neutrophil driven IFN-β production in those with active disease[23,67]. Furthermore, neutrophils were the primary cell infected with *M. tuberculosis* in samples of human BAL fluid, sputum, and pulmonary cavities from with active TB cases[68]. In addition to being a correlate of disease state, recent studies in small animal models suggest that neutrophil activity directly contributes to the progression to active disease. For example, Dorhoi et al. sought to understand the role of micro-RNA223, which had been identified as a correlate of *M. tuberculosis* infection state in a large human cohort[69]. Studying the function of this micro-RNA in mice, they found that it controlled neutrophil recruitment to the lung during infection by regulating the expression of key neutrophil chemoattractants including CXCL2, CCL3 as well as

IL-6. In spite of these observations, the role of neutrophils in early infection remains incompletely understood. In zebrafish, neutrophils contributed an early protective effect against *M. marinum* through NADPH-dependent oxidative killing[70]. In mice infected with *M. tuberculosis*, early depletion of neutrophils reduced migration of DCs to lymph nodes and further delayed priming of antigen-specific $CD4^+$ T cells[71]. In this same study, DCs that ingested infected neutrophils migrated better in an *in vitro* transwell chemotaxis system than those that simply ingested free *M. tuberculosis* bacilli. These findings suggest that neutrophil recruitment plays a causal role in disease progression with potentially both protective and destructive properties, which is likely dependent on their timing and magnitude of response.

Recent data suggest that the human gene expression signature of active tuberculosis, neutrophil derived IFN-β expression, reflects part of a broader signaling network that regulates the function of key immune players including macrophages, T cells and neutrophils[72-74]. The signaling network includes cytokines, most notably Type I IFNs, IL-10, IL-1 and IL-1 Receptor Antagonist (IL-1RA), as well as lipid derived small molecule regulators of inflammation such as the eicosanoid prostaglandin E2 (PGE2). Recent work by Mayer-Barber and colleagues suggests a cross-regulatory network in which Type I IFNs promote the production of IL-10 and IL-1RA, which in turn negatively regulate IL-1 expression and IL-1 dependent expression of PGE2[75]. In this model, PGE2 and IL-1 are protective and inhibit bacterial growth. Altering this signaling network, for example by supplementing with PGE2 or increasing PGE2 by inhibiting 5-lipoxygenase (5-LO) with the asthma drug, Zileuton, dramatically ameliorates infection outcome in mice. The mechanism(s) by which this signaling network controls the outcome of TB infection in infected animals or humans is unclear. These eicosanoids are important players in a cascade of lipid mediators that coordinate inflammation and the resolution of inflammation and

include the prostacyclins, leukotrienes, thromboxanes and resolvins[76]. Importantly these lipid mediators are both synthesized by ---and have substantial effects on-- other key immune cells in the granuloma, including neutrophils. The 5-LO product, LTB4, which has been implicated in zebrafish as an important driver of poor TB outcomes[77,78], is a powerful neutrophil chemoattractant that serves to amplify primary danger signals and coordinate neutrophil recruitment. Thus, it is likely that the profound effects of perturbing the Type I IFN-IL1-PGE2 signaling network in mice reflect not only the impact on macrophage fate but also the arrival and function of neutrophils at the site of infection.

1.2.5 T cells and beyond

For a granuloma to function, T cells are required. The interactions between *M. tuberculosis* and the cells of the innate immune system clearly have profound consequences on the subsequent adaptive response. The adaptive response is slow to emerge in *M. tuberculosis* infected hosts. Humans (and macaques) convert a tuberculin skin test, a measure of an adaptive T cell response, at approximately 6 weeks post-infection[40,42]. T cell responses in blood can be detected in macaques between 4-6 weeks[49], while a T cell response the lymphoid tissues of mice can be detected between 14-21 days post-infection[44,79]. This delay in T cell response has been attributed to several factors, including a delay in delivery of bacteria or antigens to the thoracic lymph node for T cell priming. Although some studies have suggested that the delay is in part due to the low numbers of bacilli delivered to the host[44], another study found that dose of infection did not appreciably influence the time to priming of an adaptive response[80]. There may be specific bacterial factors that inhibit delivery of *M. tuberculosis* to the lymph node. Evidence exists for both dendritic cells and CCR2+ macrophages as important players in transit of *M. tuberculosis* to

the lymph nodes for priming T cells[81,82]. Whether the bacteria that end up in the lymph nodes come from the airways during early infection, or the lung granulomas once they are established, is not yet clear. There is evidence from mice that dendritic cells can carry *M. tuberculosis* from airways to the thoracic lymph node[81]. It is less clear where the CCR2+ macrophages are encountering the bacilli for transport to lymph nodes, but it appears more likely that these are lung parenchymal bacteria. Nonetheless, most mouse studies have shown that *M. tuberculosis* bacilli must be in a lymph node to initiate priming of a T cell response[44,79], although mice devoid of lymph nodes and spleen were capable of priming T cell responses in the lungs[83].

Alteration of macrophage apoptosis, driven by the balance of LXA4 and PGE2 signals, alters $CD8^+$ T cell cross-priming by DCs[55]. These seminal findings provide insight into earlier observations of *M. tuberculosis*-specific impairment of antigen presentation and defects in Ag85-specific $CD4^+$ T cell expansion in spite of enhanced airway-LPS stimulated macrophage recruitment to the lung and increased migration of DCs to the draining LNs[44]. In addition to delaying potent T cell responses, recent work in mice has proposed that preliminary *M. tuberculosis* dissemination utilizes $CD11c^+$ DCs to seed new granuloma formation[84]. Investigating both intraperitoneal BCG and aerosol H37Rv infections in C57Bl/6 mice, Harding *et al* demonstrated that inflammatory DCs are a possible source of bacterial spread after acute infection as these cells are frequently arrested during their migration to the lymph nodes following interaction with antigen-specific T cells. These areas of infected DC-T cell capture generate new foci of inflammation that can either formulate new granulomas or extend preexisting structures depending on distance traveled. Collectively, these findings support the view that virulent *M. tuberculosis* actively subverts the early host immune response by modulating preliminary macrophage death to delay the onset of potent adaptive responses and

utilize trafficking DCs to further dissemination. These adaptations are likely crucial for *M. tuberculosis* to establish a foothold for infection given its slow growth.

In considering potential important drivers of granuloma resolution – or the softening of the caseum observed by Canetti-- it is also interesting to note that the T cell response directly regulates the innate inflammatory response. Nandi and Behar have shown in the mouse model that IFN-γ produced by CD4+ T cells inhibits neutrophil recruitment such that influx of neutrophils was reflective of a failed Th1 response[74]. These results were extended by Mishra and colleagues who showed that IFN-γ dependent nitric oxide production suppressed IL-1 production by inhibiting assembly of the NLRP3 inflammasome[85].

In addition to phagocytes, *M. tuberculosis* likely encounters other cell types, cytokines, and innate defense molecules in the airways during initial infection. Mucosal associated invariant T cells (MAITs) are CD3+CD8+ (or double negative) T cells that have T cell receptors encoded by the TRAV1/1 genes and are restricted not by the classical MHC molecules, but by a non-classical molecule MR-1 (reviewed in [86]). These cells are at higher frequency in blood and mucosal sites in humans compared to mice. MAITs emerge from the thymus with effector capabilities and thus can be considered to be early responders to bacterial, including *M. tuberculosis*, infections. MAITs respond to cells infected with bacterial pathogens without prior exposure to that pathogen, produce the cytokines IFN-γ and TNF, and are cytotoxic. Although the range of microbial ligands recognized by these "innate" T cells is not known, it was shown that MAITs recognize microbe derived riboflavin metabolites[87]. Recent studies suggest greater TCR diversity than originally appreciated, and thus MAITs are likely to recognize other microbial-derived ligands[88]. In MR-1 deficient murine models, MAIT cells were associated with early protection against bacterial pathogens, including mycobacteria[89]. Thus, these cells may act

as early sensors of *M. tuberculosis* infection in airways, and provide early cytokines to activate macrophages against this infection.

Other innate cells, including Natural Killer (NK) cells, may also play a role in early *M. tuberculosis* infection. NK cells are strong producers of IFN-γ and TNF and can also be cytolytic for *M. tuberculosis*-infected macrophages[38]. Mycolic acids are ligands for NK cells, and human studies have shown substantial variability in responses of NK cells to extracellular *M. tuberculosis*[90], suggesting that the capacity of NK cells to recognize and respond to *M. tuberculosis* could contribute to early innate resistance to infection.

In addition to cells, the airways also have molecules such as surfactants and hydrolases that have been proposed as potential modulators of *M. tuberculosis* infection. Human surfactant proteins A and D bind to *M. tuberculosis*[91]. Surfactant protein A up-regulates expression of the mannose receptor on human macrophages[92], an important receptor for binding to *M. tuberculosis*, and modulates the inflammatory response of macrophages[93]. Human surfactant protein D directly binds to *M. tuberculosis* reducing uptake by macrophages[94]. However, mice deficient in both surfactant proteins A and D were not impaired in control of low dose aerosol infection with *M. tuberculosis*[95]. Antimicrobial peptides, such as cathelicidin (LL-37), are also present in airways[96] and have been shown to increase the pro-inflammatory functions of macrophages and the killing of intracellular *M. tuberculosis*[97]. Intratracheal administration with synthetic peptides mimicking LL-37 in mice reduced *M. tuberculosis* bacterial burden[98]. There is substantial evidence that Vitamin D is important in resistance to tuberculosis[99-102], and this appears to be in part due to induction of LL-37[103,104].

Antibodies are an obvious acquired immune response that might modulate the course of infection in airways. As part of the acquired immune response, pathogen specific antibodies

cannot be predicted to prevent initial infection in previously unexposed hosts, though it is possible that they could serve this function in the case of repeat exposure or vaccination. More importantly, although *M. tuberculosis* infection induces strong antibody responses, there is only scant experimental evidence that antibodies can prevent the initial establishment of infection. Passive transfer of antibodies specific to some cell wall antigens has been reported to confer protection against disease in a mouse model but the effect was inconsistent. However, there are clear data that antibodies can change the interaction of the bacterium with macrophages in a variety of ways[105,106]; bacterial opsonization alters vesicular trafficking and macrophage signaling and interaction of antibodies with activatory or inhibitory Fc receptors on macrophages can modulate macrophage function[105]. Beyond their classical functions, antibodies have the capacity to mark the infected macrophage as aberrant and recruit the responses other innate immune cells—thus making them potential modulators of the local immune response. However, whether antibodies are present in the airways in sufficient quantities to modulate initial infection remains unclear, and a source of substantial investigation.

In summary, there are a variety of cells, cytokines, and molecules present in airways that can modulate the initial response of the host to *M. tuberculosis* infection. These factors may prevent infection completely, limit initial establishment of granulomas, modulate the local environment of newly emerging granulomas or increase the induction of T cell responses against *M. tuberculosis*. Changes in these factors could increase susceptibility to initial infection as well. Further studies will be necessary to more fully understand the relative contributions of these factors to modulation of initial infection.

1.2.6 Lessons from Clinical Isolates of *M. tuberculosis*

It is highly likely that bacterial factors drive differences in granuloma fate as well as host factors. Using barcoded bacteria to track to origins of the bacterial populations in individual lesions, we have shown that most pulmonary granuloma arise from one progenitor bacterium[36] and genetic polymorphisms arise and become fixed in the bacterial populations of isolated granuloma[107]. These data are consistent with historical data indicating that within a given individual, bacteria in one lesion can acquire drug resistance independently of the bacterial populations in other lesions[108]. These data reinforce the concept that granuloma evolve relatively independently within the same host.

Many bacterial virulence factors have been identified through forward and reverse genetic approaches in experimental systems. However, it is not clear whether any of these might be modulated to alter interactions with the host in a lesion specific fashion and thus contribute to the different lesional trajectories we and others have observed[36,39,109]. This question has not yet been addressed through lesion specific analyses, for example, transcriptional profiling of granuloma bacterial populations, which is experimentally challenging given the relatively small number of bacteria in many lesions.

It is likely that the different virulence manifestations of clinical strains will shed some light on the bacterial pathways that flexibly alter interactions with the host. Six distinct lineages have now been defined based on sequence differences[110]. There is mounting evidence that this genetic diversity generates clinically relevant phenotypic variation and impacts infection outcome. Strains have been shown to differ in terms of their mortality, pathologic manifestations and immune responses in mice and in human macrophages[111-114]. Despite the mounting evidence

that the genetic diversity of *M. tuberculosis* has clinical consequences, few concrete links between genotype and phenotype have been identified. The best studied association has been between the presence of phenolic glycolipid biosynthesis and the hypervirulent phenotype and immunosuppressive properties associated with Lineage 2 strains[115]. In *M. marinum* infection of zebrafish, phenolic glycolipid promotes the recruitment of permissive macrophages to the site of infection and is required to establish robust infection[116]. In modulating initial macrophage recruitment[114] as well as macrophage death pathways (discussed above), *M. tuberculosis* is likely influencing multiple early host interactions to affect inflammatory programs, granuloma fate, dissemination, and ultimately infection outcome.

It remains unclear whether bacterial expression of virulence lipids varies in a lesion dependent fashion. It is interesting to note, however, that production of another virulence lipid, PDIM, is one mechanism to resolve propionyl-CoA toxicity during growth on fatty acids[117,118]. PDIM biosynthesis is both required for bacterial survival in mice and macrophages[119,120] and PDIM has been proposed, based on work in *M. marinum*, to directly cloak TLR2 ligands[116]. Thus, it is possible that the regulatory effects of central carbon metabolism—where more or less PDIM may be produced depending on carbon source availability—provides an energy efficient mechanism to link host environment with bacterial virulence and reinforce the trajectory of any given lesion after it is established by very early host events.

Establishment and progression of *M. tuberculosis* remains somewhat of a mystery in humans. However, a deeper understanding of the early events in tuberculosis is essential to identifying new and effective strategies of preventing active TB. The best vaccine would prevent establishment of the infection, or at the very least, prevent early dissemination of individual granulomas. Understanding the early airway and lung responses to this infection is crucial, as

this is where control must occur. There are a variety of host cell types and molecules, as well as bacterial factors, which interact in early infection, as described here. Building on this knowledge will move the field of vaccines against tuberculosis forward. Without a clear understanding of the early processes that vaccines must prevent or limit, and the host responses that can be harnessed for protection, we cannot expect a vaccine to succeed against this complex and evolved pathogen.

1.3 MODELING HETEROGENEITY AND SINGLE LESION DYNAMICS IN THE MACAQUE MODEL OF *MYCOBACTERIUM TUBERCULOSIS*

Animal models of *M. tuberculosis* are critical to dissect features of infection pathogenesis, TB pathology and immunology, and bacterial virulence. Several animal models have been used over the past decades to study TB including zebrafish[121], mice[122], and nonhuman primates[123,124]. Each system possesses both benefits and limitations in modeling TB and while none perfectly capture human *M. tuberculosis* infection, they all contribute to our understanding of this disease. The model recognized to best recapitulate the multiple facets of variability in TB is the cynomolgus macaque. While ethical considerations, limited reagents, and cost may detract general use, the range of outcomes both clinically and pathologically is remarkably similar to that observed in humans[49,125]. There is a 50:50 ratio of latent infection and active disease in adult cynomolgus macaques infected with low dose (<25) virulent *M. tuberculosis* Erdman. Importantly, within these binary definitions, infected macaques recapitulate the entire spectrum of human disease both at the local, lung and overall host level including the ability to reactivate

latent infection following immune suppression with SIV[126], anti-TNF[127] or anti-CD4 treatment[128].

Given the recent seminal finding that the majority of *M. tuberculosis* lesions in NHPs arise from a single bacillus[36], genetically tagging individual bacilli provides an unparalleled ability to track the outcomes of discrete lesions within a single host. This dissertation utilizes this molecular barcoding tool to examine lesions in multiple dimensions by linking immunology, histology, and radiographic imaging to specific lesions in a novel manner. Collating these different facets of TB biology allows for the re-creation of infection on a lesion-by-lesion basis and will enhance understanding into the early events that occur post challenge by generating an overall picture of *M. tuberculosis* infection *in vivo*. Furthermore, barcoding individual *M. tuberculosis* permits the study of sequential infections and exposures in a single host. These are crucial aspects of *M. tuberculosis* infection that have yet to be thoroughly explored.

2.0 STATEMENT OF INTENT AND SPECIFIC AIMS

It is estimated that a third of the world's population is infected with *M. tuberculosis*, the acid-fast bacillus responsible for tuberculosis (TB). A tenth of this infected population will develop primary active disease with the remaining 90 percent containing the infection as latent TB. Of those individuals with latent TB, ~5-10% will reactivate to active TB disease over their lifetime. Epidemiologic studies of tuberculin skin test (TST) screening in highly endemic countries have further revealed persistent non-responders, despite high exposure to *M. tuberculosis*, hinting of a potentially TB resistant human reservoir. The complex heterogeneity that is responsible for this observed human variability to TB remains poorly understood. It is likely that the human TB spectrum stems from early immunologic and host-bacterium interactions that orchestrate both the susceptibility to infection, progression of disease, and risk of reactivation. In probing early events in *M. tuberculosis* infection, we hope to identify key phases of infection that control bacterial dissemination and correlate with protection, which can be subsequently exploited for novel vaccine development and drug design. In this dissertation, we examine how early events in *M. tuberculosis* infection relate to host outcome by addressing questions of bacterial dissemination and reinfection using individually barcoded bacteria in the non-human primate (NHP) model of TB. We are also evaluating the earliest microbial interactions in the lung by examining the influence of *M. tuberculosis* on the lung microbiome, particularly as a function of

pulmonary inflammation. Together, these studies will help elucidate the contribution of early events on host outcome.

Primary hypothesis: *We hypothesize that early interactions between* M. tuberculosis *and host immune components drive local lesion heterogeneity and that these initial events influence infection outcome.* To address this hypothesis, we propose the following Specific Aims:

2.1 SPECIFIC AIM 1: DETERMINE FEATURES CONTRIBUTING TO BACTERIAL DISSEMINATION IN *M. TUBERCULOSIS* INFECTION AND EXPLORE HOW EARLY INTERACTIONS BETWEEN HOST AND BACTERIA MODIFY BACTERIAL SPREAD

We will infect a cohort of cynomolgus macaques (*Macaca fascicularis)* with a library of individually barcoded *M. tuberculosis* Erdman (~25 CFU) to track bacterial fate post challenge. These identity tags enable us to correlate lesion development with *M. tuberculosis* dissemination, infection progression, disease presentation, and lesion heterogeneity. Coupling the roadmap of infection from tagged *M. tuberculosis* (identity) with serial 18[F]-fluorodeoxyglucose (FDG) PET/CT imaging (timing) and cumulative bacterial burden (granuloma fate), we aim to unravel events in early infection critical to productive dissemination.

Hypothesis: Early granulomas that are highly inflammatory (FDG avid) will be more likely to contribute to productive (CFU+) secondary lesions.

2.2 SPECIFIC AIM 2: DETERMINE HOW AN ONGOING *M. TUBERCULOSIS* INFECTION MODULATES SECONDARY INFECTION: PROTECTION OR EXACERBATION?

Surprisingly little is known about the protective or pathogenic potential of a prior, ongoing *M. tuberculosis* infection on a subsequent exposure or infection. Addressing this question is crucial to the development of effective vaccines. We will examine the impact of prior infection on reinfection in TB by infecting macaques with two different libraries of barcoded *M. tuberculosis.* Each library will have a unique library-specific tag to differentiate primary and secondary granulomas. In two successive, low-dose (>25 CFU) infections, we will investigate whether an ongoing infection assists host clearance by enhancing TB immunity or undermines host defense by exacerbating host immunopathology.

Hypothesis: Primary infection enhances protection to secondary infection by eliciting immune responses that limit early bacterial dissemination and augment bacterial killing in secondary granulomas.

2.3 SPECIFIC AIM 3: DETERMINE HOW *M. TUBERCULOSIS* INFLUENCES THE LUNG MICROBIOME IN THE MACAQUE MODEL OF TUBERCULOSIS

The specific interactions of the lung microbiota with *M. tuberculosis* early in infection are entirely unexplored. Studies on other infections that examine immunity and the microbiome suggest that there are important exchanges that occur between these entities that influence the immunological landscape by altering immune regulation and inflammation in both local and

systemic contexts[129]. We will assess if *M. tuberculosis* modifies the lung microbiome in 26 infected macaques by serially comparing infected and uninfected lobes over the course of infection looking for changes in abundance and diversity as a function of pulmonary inflammation. Since we have robust markers of disease progression using PET/CT as a measure of lung inflammation[39,109], we will determine whether there is a correlation between the duration of change in lung microbiome in response to *M. tuberculosis* infection and the outcome of infection, gaining power from the serial sampling of individual animals. We reasoned that the increased inflammation in animals that are progressing to active TB is likely to have a sustained interaction with the microbiome, while the animals that are controlling the infection early, and therefore have reduced inflammation, are less likely to have a sustained change in microbiome composition. Successful completion of this aim will demonstrate the relationship of the structure of microbial communities to *M. tuberculosis* infection and progression of infection, define changes in the structure of lung microbial communities over time, and determine the relationship of microbial communities to systemic inflammatory biomarkers. In examining some of the earliest relationships between *M. tuberculosis* and the lung microbe milieu, we aim to better understand how the initial lung context may influence host outcome and host variability in TB.

Hypothesis: We hypothesize that lung inflammation from M. tuberculosis *influences the variability and diversity of the lung microbiome in the infected lobe of macaques.*

3.0 DIGITALLY BARCODING *MYCOBACTERIUM TUBERCULOSIS* REVEALS *IN VIVO* INFECTION DYNAMICS IN THE MACAQUE MODEL OF TB

This chapter is adapted from the accepted manuscript:

Constance J. Martin*, **Anthony M. Cadena***, Vivian W. Leung, Philana Ling Lin, Pauline Maiello, Nathan Hicks, Michael R. Chase, JoAnne L. Flynn[#] & Sarah M. Fortune[#]. 2017. Digitally barcoding *Mycobacterium tuberculosis* reveals *in vivo* infection dynamics in the macaque model of tuberculosis. *mBio* 8:e00312-17. *Co-first authors, both authors contributed equally to this work; [#]Corresponding authors

3.1 IMPORTANCE

Classically, *M. tuberculosis* infection was thought to result in either latent infection or active disease. More recently, the field has recognized that there is a spectrum of *M. tuberculosis* infection clinical outcomes. Within a single host, this spectrum is recapitulated at the granuloma level, where there can simultaneously be lesional sterilization and poorly contained disease. To better understand the lesional biology of TB infection, we digitally barcoded *M. tuberculosis* to quantitatively track the fate of each infecting bacterium. By combining this technology with serial PET-CT imaging, we can dynamically track both bacterial populations and granuloma trajectories. We demonstrate that there is little constraint on the bacterial population at the time

of infection. However, the granuloma imposes a strong bottleneck on dissemination, and the subset of granulomas at risk of dissemination can be distinguished by physical features.

3.2 INTRODUCTION

Tuberculosis (TB) poses a threat to global health, responsible for more than 10 million new cases of active disease and nearly 2 million deaths in the last year[130]. However, these recognized cases of TB disease reflect only a small fraction of *M. tuberculosis* infections, most of which result in a spectrum of outcomes that are clinically silent and collectively referred to as latent TB (LTBI)[2,10,131]. These different courses of infection are characterized by a range of bacterial burdens and pathology, and correlate with reactivation risk[14,21,22,49]. Recently, the field has focused on early interactions between host and bacterium as potential drivers of the variable outcomes of *M. tuberculosis* infection[27]. However, new tools are needed to dissect the local biology of *M. tuberculosis* infection, especially in the non-human primate model (NHP), whose strengths are that it recapitulates the variable course of human infection and produces individual granulomas with pathology very similar to that found in humans[49,132,133].

Dissection of lesion course in macaques has been transformed by the use of positron emission tomography coregistered with computed tomography (PET/CT) and a ^{18}F-fluorodeoxyglucose radiotracer (^{18}[F]-FDG), and validated in humans[21,39,109]. Using this approach, we previously demonstrated that formation of disseminated lesions early after infection is associated with development of active disease whereas limited early dissemination is associated with maintenance of clinically latent infection, suggesting that early dissemination is critical in determining host outcome[39]. From these studies, it was unclear whether early

dissemination was the result of a global defect in immunity or due to loss of control by a single lesion. We have also shown that granulomas follow unique trajectories even within individual macaques as evidenced by differences in killing efficacy[132]. Thus, we hypothesized that individual granulomas might similarly vary in their risk of dissemination.

We have also shown that most lesions harbor the progeny of a single bacterial founder. However, this study was limited to very early time points (~4 weeks post-infection) and so did not assess the dynamics of dissemination[132]. Moreover, as the previous study only used a panel of 8 bacterial strains, it did not allow us to unambiguously resolve the subsequent fate of each bacterium following infection.

To address the lack of tools available to answer these critical biological questions, we developed a genome barcoding system, allowing us to track the fate of each infecting bacillus. The population dynamics of other pathogens have been tracked by assessing change in genetic composition of a population, leveraging either natural or artificially introduced genetic variation[134]. However, these infection models are typically characterized by a relatively large infecting inoculum and wide bottlenecks that make tracking of individual bacteria or viruses in a population infeasible, and instead investigators have tracked changes in the distribution of diversity in the population to estimate bottleneck size. By contrast, TB is a paucibacillary infection, where inocula of <20 organisms successfully establish infection, allowing us to track infection dynamics by directly following the fate of each bacterium that establishes infection. To do this, we developed a complex library of digitally barcoded *M. tuberculosis* such that each infecting organism carries a unique and quantifiable sequence identifier. In parallel, we developed a robust Python algorithm to reliably discriminate bacterial barcodes. By combining this barcoding system with serial ^{18}[F]-FDG PET/CT in a macaque model, we provide the first

quantitative map of within-host bacterial population dynamics focusing on *M. tuberculosis* infection and dissemination.

3.3 MATERIALS AND METHODS

3.3.1 Barcode generation

Primers (CM29/CM30 – see primer table, **Supplementary Table 6**) using a string of degenerate bases were used to amplify qTags[135] from parent vectors and were restriction enzyme-digested and cloned into pJeb402 with KpnI and XbaI. Each library (single qTag) was constructed in *E. coli* and plated at 4x coverage before plates were scraped, maxi-prepped (Qiagen) and transformed into *M. tuberculosis* Erdman or *M. smegmatis* mc[155] at 4x coverage. Libraries were then scraped, passed through a 5μM filter (Millex) and sonicated for single cell suspension and mixed at equal OD_{600} to generate infectious libraries containing a pool of three Library IDs.

3.3.2 Sequencing

Bead beaten and phenol:chloroform purified genomic DNA was diluted to 10ng/uL, and amplified using Phusion polymerase through two rounds of PCR of 12 cycles each. Primers are indicated in primer table. Fragments of the appropriate size were gel purified, quantified using Clonetech NGS kit and sequenced on an Illumina Miseq using V2 chemistry.

3.3.3 Macaque infections, PET/CT imaging, and tissue excision

Four adult cynomolgus macaques (*Macacca fasicularis*) were obtained from Valley Biosystems (Sacramento, California) and screened for *M. tuberculosis* and other comorbidities during a month-long quarantine. Each macaque had a baseline blood count and chemical profile and was housed according to the standards listed in the Animal Welfare Act and the Guide for the Care and Use of Laboratory Animals. All animals were infected with barcoded strain Erdman *M. tuberculosis* via bronchoscopic instillation as previously published[49,125] and received an inoculum of 11±5 CFU (determined by plating a sample of the inoculum and counting CFU after 3 weeks). All animals were followed with serial 2-deoxy-2-[^{18}F]-fluoro-D-glucose ([^{18}F]-FDG) PET/CT imaging as formerly described[39,109,132] to identify and track lesion formation and progression over time (**Table 1**). As before, lesions were individually characterized by their date of establishment (scan date), size (mm), and relative metabolic activity as a proxy for inflammation ([^{18}F]-FDG standard uptake normalized to muscle [SUVR]). Lesions $\geq$ 1mm can be discerned by our PET/CT equipment. To avoid barcode cross-contamination, individual granulomas were separately excised and processed using the animal's pre-necropsy scan as a guide for identity and location.

3.3.4 Isolation and preparation of bacteria from tissue samples

Following removal at necropsy, each lesion was homogenized and plated for bacterial burden on 7H11 agar supplemented with oleic albumin dextrose catalase (OADC). A small portion of homogenate was frozen for CEQ analysis (below). After a 3-week incubation, the plates were counted for CFU and colonies were pooled and scraped into 7H9 supplemented with OADC and

20% Tween 80. Each plate was scraped into a separate tube and frozen in 5mL of 7H9 at -80°C to await genomic DNA extraction and sequencing.

3.3.5 Extraction of *M. tuberculosis* genomes and chromosomal equivalent quantification (CEQ)

M. tuberculosis genomic extraction and CEQ quantification was performed as previously published[22,132]. Briefly, gDNA was extracted with phenol:chloroform:isoamyl alcohol (25:24:1, Invitrogen) with an intermediate bead beating step using 0.1mm zirconia-silica beads (BioSpec Products, Inc.). CEQ was assessed relative to a serially diluted standard curve of *M. tuberculosis* genomic DNA using quantitative real-time PCR preformed in triplicate on an iQ5 Multicolor Real-Time PCR Detection System (Bio-Rad Laboratories, Inc.) and a 384-well 7900HT Fast Real Time PCR System (Applied Biosystems). Quantification of CEQ used the previously published primer-probe mixture and TaqMan Universal Master Mix II (ThermoFisher Scientific). Acceptable real-time PCR efficiency for each run was kept between 90%-110%.

3.3.6 Bioinformatics analysis

Data were analyzed using a custom Python pipeline available on GitHub (https://github.com/sarahfortunelab/barcodetracking). The pipeline iterates through the raw FASTQ files generated from Illumina sequencing for each indexed sample. We first identify library ID, barcode and molecular counter features using constant “handle” sequences as search motifs. We then filter data to obtain high-quality reads that: 1-have all features present; and 2-have a maximum probability of base calling error of 0.001 (equivalent to a minimum base

quality score of least 30 or Phred 63). As an additional check, we grouped reads that passed quality control and counted molecular counter copies to check the extent of skewing. For each sample, we also tallied the number of unique molecular counters for each sequence combination and normalized counts to the total molecular counts in the sample. Using the aforementioned percentage counts as a metric for library-ID-barcode abundance, we computed the true-false-determining threshold in the pipeline using a recursive approach. In each round, we first computed and sorted the relative abundance in descending order. Second, we iterated through the barcodes that passed the previous threshold and, for each, eliminated less-abundant variants that differed by one position from the former. Third, we calculated the threshold dividing true and false barcodes using a modified concavity approach. Once the threshold reached a stable point, we applied the final threshold to return the set of true barcodes.

3.3.7 T cell flow cytometry and intracellular cytokine staining

At necropsy, granulomas were processed intro single cell suspensions with sterile PBS. Portions of these suspensions were used for T cell profiling and cytokine analysis. Tissue suspensions were stimulated with Mtb CFP10 and ESAT 6 (BEI Resources, Manassas, VA) and Brefeldin A (GolgiPlug, BD Biosciences) in RPMI (Lonza, Walkersville, MD) supplemented with 1% L-glutamine and 1% HEPES (Sigma, St. Louis, MO) for 3.5 hours. The final concentration of Mtb-specific peptides was 2.5μg/mL. Cells were first stained with LIVE/DEAD Fixable Blue Dead Cell Staining Kit (ThermoFisher Scientific) for viability and then stained with the T cell-specific marker, CD3, using the anti-human CD3 (clone SP34-2, BD Biosciences) antibody in a standard FACs buffer. Following surface staining described above, the cells were fixed and permeablized (BD, Cytofix/Cytoperm) and washed with BD Perm/Wash buffer (BD,

Biosciences). Incubations were preformed according to manufacturer recommendations. Cells were then stained with the following intracellular cytokine stains: anti-human IFN-γ (clone B27, BD Biosciences), anti-human TNFα (clone Mab11, eBiosciences), anti-human IL-2 (clone MQ1-17H12, BD Biosciences), IL-6 (clone MQ2-6A3, BD Biosciences), anti-human IL-10 (clone JES3-9D7, eBiosciences), and anti-human IL-17A (clone eBio64CAP17, eBiosciences). Flow cytometry was preformed on a LSR II (BD) and analyzed using FlowJo Software ver. 9.9.5 (Treestar Inc., Ashland, OR). Size (FSC) and granularity (SSC) were used to isolate the lymphocyte population at the cytometry. All cytokine data presented are gated on $CD3^+$ T cells.

3.3.8 Statistical analysis

Statistical analysis was performed in Graphpad Prism, JMP and R, significance was found when $p < 0.05$.

3.4 RESULTS

3.4.1 Barcoded bacterial library generation

We engineered a library of digitally barcoded plasmids that we introduced into *M. smegmatis* and *M. tuberculosis* Erdman. The 'barcode' consists of a random 7-mer and adjacent 75-mer library identifier tag stably inserted into the bacterial chromosome[135] (**Figure 5A**). Using this approach, we were able to generate libraries of roughly 50,000 uniquely identifiable bacteria, ensuring a <2% chance that a barcode would be represented twice if 20 bacteria are randomly

selected. To quantitate barcodes via sequencing, the barcode is amplified from isolated genomic DNA in two rounds of nested PCR steps. In the first round, a pool of four primers with a degenerate variable length spacer anneals a random 9-mer 'molecular counter' that allows us to enumerate PCR templates rather than amplicons[136]. The degenerate variable length spacer, or "phasing region", introduces sequence variability necessary for Illumina based sequencing of relatively low complexity libraries [137]. The second round of PCR completes the addition of sequencing adapters and multiplexing indices. Using this sequencing approach, we confirmed the diversity of the *M.tuberculosis* Erdman library and found that most barcodes were present at similar abundances (**Figure 5B**). A few sequences appeared over-represented, however, these represented only 0.5% of the total library and were not identified in any NHP used in this study.

Concurrently, we developed a custom Python pipeline, BARTI, to identify barcodes from Illumina sequencing data in complex biological samples (i.e. granulomas), where both sequences and number of barcodes in a sample are unknown. The challenge in reading out randomized DNA barcodes is that errors inherent in sequencing create uncertainty in identifying unique barcode sequences. In previously published genomic barcoding approaches, this challenge was addressed by using only arrayed libraries of known barcodes, such that exact or near-exact sequence matching could be enforced[138]. However, this approach significantly constrains the number of unique individuals represented in a population and precludes unambiguously tracking individual bacteria. In sequencing approaches where unknown barcodes are used, error correction methods have been developed to discriminate false sequences from true[139,140] but the best approach to high precision barcode counting is unclear.

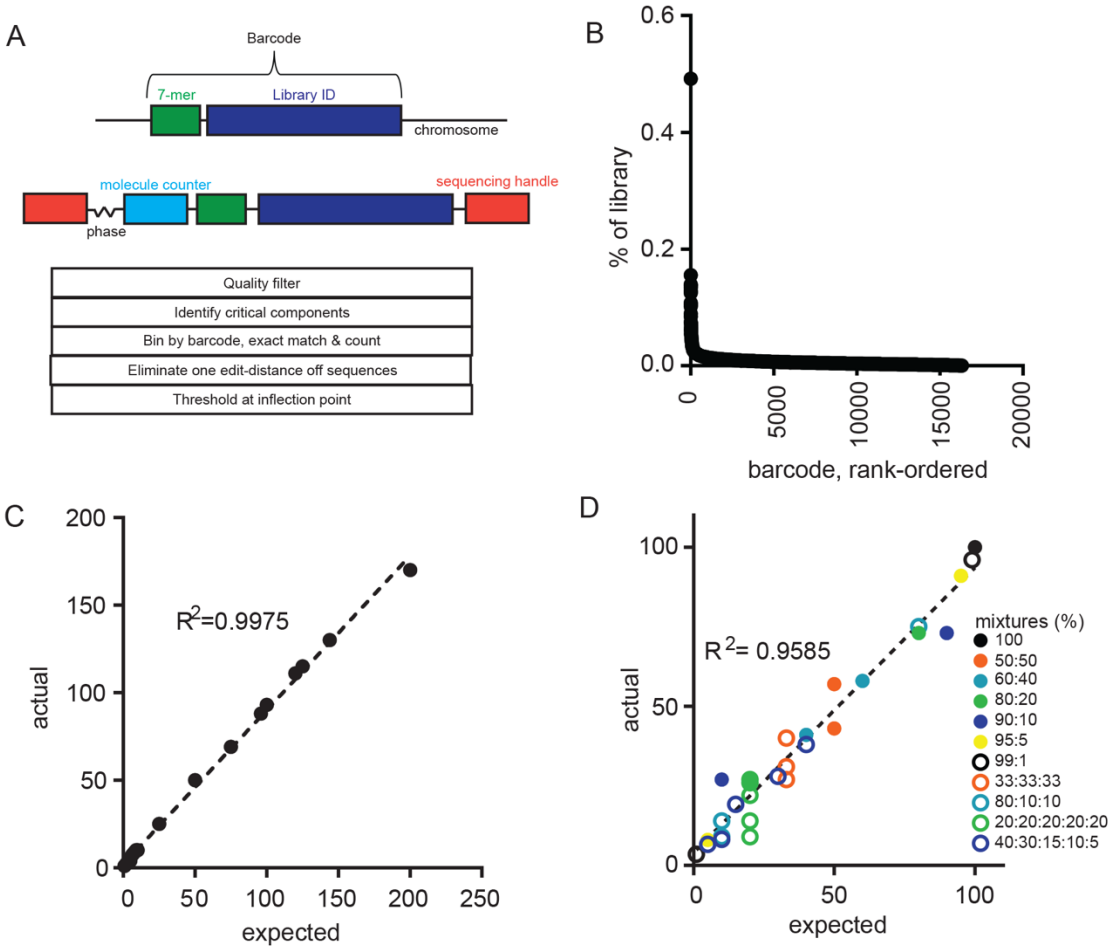


Figure 5. Barcode structure and pipeline validation.

A. Barcode structure in the *M. tuberculosis* genome (top), and after preparation for sequencing (middle). Schema of the BARTI pipeline for barcode identification, thresholding and counting. B. Abundance of individual barcodes in a single library. C. Number of barcodes counted using BARTI from known number of barcoded *Mycobacterium smegmatis* colonies at a read depth of >100,000. D. Ratios of barcodes counted using BARTI (actual) from known mixtures of barcode-containing plasmids mixed at indicated ratios (expected) and sequenced at a read depth of >100,000. Dashed line is linear regression, R2 is Spearman correlation.

To define the analytic path that provided the most accurate barcode quantitation, we arrayed single colonies of barcoded *M. smegmatis*, generated and sequenced pools of known numbers of colonies. We expected that the number of barcode sequences identified should be

equal to the number of colonies used in that sample's preparation. Without error correction, as with other iterations of genome barcoding technology[138,139], a much larger number of barcode

sequences were identifiable in the sequencing data than were present in the input library. Strict read-quality filtering eliminated many, but not all, suspected erroneous sequences. Distance errors, wherein one or more bases in a barcode sequence are erroneously called, are common to these types of sequencing data. Often algorithms are applied to correct distance errors by condensing sequences that differ by one base to the dominant sequence in order to maintain the greatest read depth possible[141,142]. However, we found that condensing barcodes that differed by a single base in some cases skewed the estimation of barcode number. Thus, we chose to discard sequences that differed by one base from a more abundant barcode. This processed pool of barcode sequences still contained two populations of barcodes, abundant and putatively "true" and rare and putatively "false" barcodes representing persistent erroneous sequencing. By rank ordering the barcode sequences by template counts and taking the second derivative of the slope between each barcode to the next lesser abundant barcode, we found that the maximally negative point along this line described an inflection point that accurately separated expected number of barcodes from sequencing artifact. Using this analytic approach, we accurately quantitated the number of known bacterial barcodes in the pooled mixtures of arrayed mycobacterial colonies across a range of sequencing read depths (**Figure 5C and Supplementary Figure 30**). The sequencing pipeline also accurately quantitated both the number and relative abundance of barcode sequences present at unequal ratios when the least abundant barcodes were $\geq$1% of the population (**Figure 5D**).

Table 1. Details of macaque infection and disease.

Animal ID	Strain	Recovered # Barcodes	4wk Gran Counts (PET/CT)	Time to Nx (wks)	Gross Pathology Score	Total CFU
17814	Barcoded Mtb Erdman	22	26	19	21	23,979
17914		21	27	19	24	24,489
18014		21	26	15	23	53,710
18114		16	12	16	14	14,060

3.4.2 Mapping *M. tuberculosis* infection dynamics in macaques.

To define bacterial population dynamics *in vivo*, we infected 4 macaques (**Table 1**) with a goal inoculum of ~20 CFU of barcoded Erdman, where the delivery inoculum as determined by plating was 11 ± 5 CFU, recognizing there is uncertainty in this number created by plating and counting very low numbers of a very dilute sample. We tracked temporal and spatial granuloma formation *in vivo* with serial ^{18}[F]-FDG PET/CT for 15-20 weeks. Lesions were identified, labeled, and characterized for size and PET avidity (SUV)[39]. Using the final pre-necropsy PET/CT as a lesion map, we excised each individual lesion, plated their homogenates, scraped and sequenced the resulting colonies to decode bacterial founder identity. By overlaying serial PET/CT scans with post-sequencing barcode identity, we were able to create a systematic history of productive—i.e. culture positive—lesion formation and dissemination in each animal (**Figure 6A**).

3.4.3 Concordance between estimated dose and number of bacteria that successfully establish infection.

In other pathogens including Salmonella and HIV, barcode tracking has demonstrated that only a tiny fraction of the delivered organisms successfully establish infection[143,144]. *M. tuberculosis* is thought to be a more efficient pathogen but the infective dose of *M. tuberculosis* has been difficult to define. We compared the number of bacteria in the inoculum as estimated by plating with the number of unique barcodes identified in that animal to estimate the fraction of the inoculum that successfully established infection. The number of bacteria in the inoculum closely matched the number of barcodes recovered from the animal. These numbers also paralleled the number of granulomas identified by PET-CT at 4 weeks post infection (**Table 1**). These data suggest that the early events of infection impose little to no bottleneck on the bacterial population and are consistent with the model that the infective dose of *M. tuberculosis* can be as low as one bacterium.

3.4.4 A subset of granulomas disseminate to form new lesions.

We then sought to define bacterial population dynamics through the first three months of infection, encompassing the window of dissemination that we previously identified between 3 and 8 weeks of infection. We have previously shown that at 3 weeks post infection, most granuloma harbor the progeny of a single founder[132]. Here, we find that that the majority of granulomas even at later time points contain the progeny of a single bacterial founder, as most lesions (85%) had only one barcode sequence (**Figure 6B**). Thus, we find little evidence of inter-lesional mixing of the bacterial population even after a period of dissemination. Thoracic

lymph nodes often contained more than one barcode (75%), as expected for sites draining the lung. However, only a minority of barcodes present in the lung were culturable from the lymph nodes, suggesting that not all granulomas productively disseminate to lymph nodes, at least at these later time points (**Figure 6C-F**).

Barcode tracking defines two populations of granulomas: lesions containing bacteria that are derived from the same founder, which we interpret as the product of productive dissemination, and lesions harboring unique bacterial barcodes, which were "contained"—i.e. did not appear to have disseminated to form new culture positive lesions. The pattern of dissemination varied across the four animals (**Figure 6C-F**). In 3 macaques the majority of lesions were contained. In one macaque, ID: 18014, ~80% of lesions were disseminated. This appeared to reflect failure of containment at multiple sites as opposed to widespread dissemination from a single initial lesion. On the whole, however, this analysis demonstrates that while nearly all bacteria in an infecting inoculum successfully establish a culture positive lesion, the majority do not contribute to further spread of infection. Across all 4 animals, only a small fraction (8.75%) of granulomas were able to seed multiple (≥3) new granulomas in lung and lymph nodes (**Figure 6C-F**).

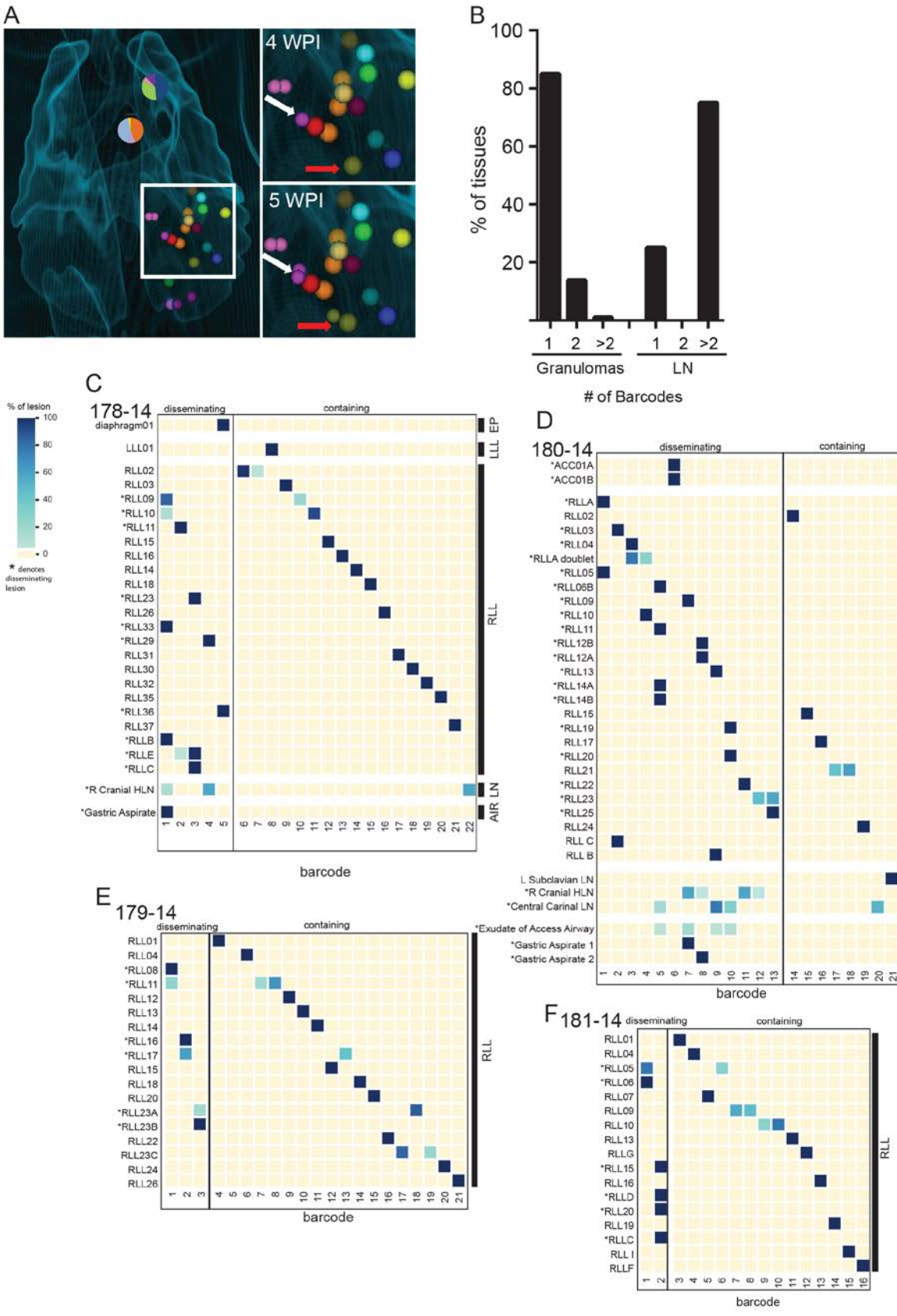

A
4 WPI
5 WPI
B
% of tissues
100
80
60
40
20
1 2 >2
1 2 >2
Granulomas
LN
# of Barcodes
C
% of lesion
100
80
60
40
20
0
* denotes disseminating lesion
178-14
disseminating
containing
diaphragm01
LLL01
RLL02
RLL03
*RLL09
*RLL10
*RLL11
RLL15
RLL16
RLL14
RLL18
*RLL23
RLL26
*RLL33
*RLL29
RLL31
RLL30
RLL32
RLL35
*RLL36
RLL37
*RLLB
*RLLE
*RLLC
*R Cranial HLN
*Gastric Aspirate
EP
LLL
RLL
LN
AIR
barcode
D
180-14
disseminating
containing
*ACC01A
*ACC01B
*RLLA
RLL02
*RLL03
*RLL04
*RLLA doublet
*RLL05
*RLL06B
*RLL09
*RLL10
*RLL11
*RLL12B
*RLL12A
*RLL13
*RLL14A
*RLL14B
RLL15
*RLL19
RLL17
*RLL20
RLL21
*RLL22
*RLL23
*RLL25
RLL24
RLL C
RLL B
L Subclavian LN
*R Cranial HLN
*Central Carinal LN
*Exudate of Access Airway
*Gastric Aspirate 1
*Gastric Aspirate 2
ACC
RLL
LN
AIR
barcode
E
179-14
disseminating
containing
RLL01
RLL04
*RLL08
*RLL11
RLL12
RLL13
RLL14
*RLL16
*RLL17
RLL15
RLL18
RLL20
*RLL23A
*RLL23B
RLL22
RLL23C
RLL24
RLL26
RLL
barcode
F
181-14
disseminating
containing
RLL01
RLL04
*RLL05
*RLL06
RLL07
RLL09
RLL10
RLL13
RLLG
*RLL15
RLL16
*RLLD
*RLL20
RLL19
*RLLC
RLL I
RLLF
RLL
barcode

Figure 6. Contained and disseminated lesions from 15-20 week infected macaques.

A. Thoracic CT scan from animal 180-14, with barcode composition in lymph nodes (pie charts) and granulomas (bubbles). Inset, 4 and 5 WPI (weeks post infection) with granulomas identifiable at those time points. White arrow marks dissemination of 'purple' barcode from founder lesion at 4 WPI to new lesion at 5 WPI. Red arrow marks dissemination of 'olive' barcode from founder lesion at 4 WPI to new lesion at 5 WPI. B. Quantification across all animals of number of unique barcode sequences found in granulomas and lymph nodes. C-F. Contained and disseminated barcode sequences in all recovered CFU+ lesions arrayed according to spatial distribution approximately from top to bottom of each lung lobe for the 4 macaques (macaque ID in top left of each graph). Blue shading reflects percentage of molecular counters out of total for a lesion for that barcode. Numbered lesions were identified pre-necropsy by scan, lettered lesions were found during necropsy and often do not have corresponding XYZ coordinates. Gastric aspirates in E were taken 13 days apart. Barcodes are arbitrarily numbered. RLL, right lower lobe; LLL, left lower lobe; ACC, accessory lung lobe; EP, extra-pulmonary; LN, lymph node; AIR, airway.

3.4.5 Dissemination occurs primarily through local spread.

We sought to better understand these patterns of dissemination by defining the spatial dynamics of bacterial spread. Using 3-dimensional lesional coordinates from pre-necropsy PET/CT scans, we interrogated the spatial relationships between disseminated lesions containing the same barcode compared to the population of contained lesions (**Supplemental Figure 31**). Lesions that share barcodes are closer together than the population of contained lesions (mean contained = 19.657mm, mean disseminated = 8.87mm) and are significantly different populations (D=0.615, p=2.82e-9) (**Figure 7A**). Thus, most dissemination appears to result from local rather than hematogenous or lymphatic spread.

Matching timing data from serial scans with barcode identity we also established the temporal dynamics of dissemination. Consistent with previous findings[39], there is substantial

dissemination within the first 6 weeks of infection. After this window of early dissemination, there is a period of apparent quiescence and then a second wave of granuloma spread (**Figure 7B**).

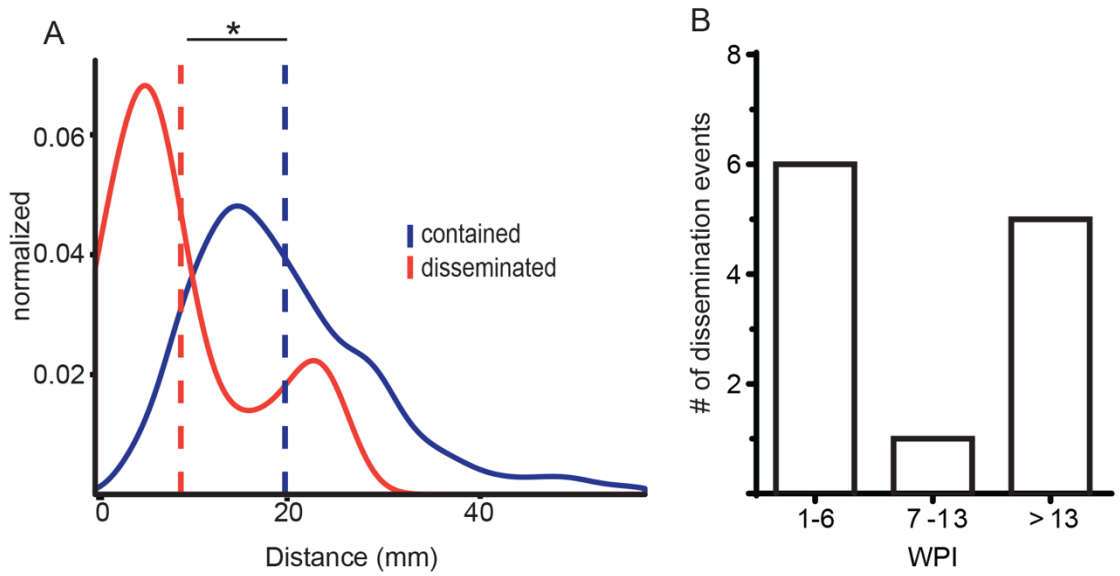


Figure 7. Spatial and temporal characteristics of dissemination.

A. Density histogram of the Euclidean distance of each disseminated lesion (red) to contained (blue) lesions with the same barcode sequence across all animals. Blue line is the Euclidean distance of all contained lesions to other lesions that do not contain the same barcode sequence. Data is only reflective of lesions for which XYZ coordinates are known. Dashed line is mean for each distribution, *; $p < 0.05$ by Welch's t-test. B. Number of lung disseminating events across all animals at the indicated time points (WPI; weeks post infection). Timing reflects definitive dissemination events matched with barcode and serial imaging.

3.4.6 Disseminated lesions are larger than contained lesions.

The variability in the patterns of dissemination was striking as even the animal allowing extensive dissemination was not identifiable by clinical course. These animals were also not distinguished by their total bacterial burden at necropsy (**Table 1**), which were similar to the

average total CFU as compared to our historical data (N=18 cynomolgus macaques[22,39] range = 0 - 222,864; mean = 34,338; median = 6613).

These data suggested that dissemination is driven by lesional as well as global features. We therefore sought to identify lesional characteristics that distinguish containing from disseminating lesions. We compared the populations of T cells between these sets of lesions, but found minor differences in the frequency of cytokines produced by T cells following ex vivo stimulation (**Supplemental Figure 32**). However, as dissemination and necropsy are separated by weeks-to-months, the quality of the immune response likely changed dramatically, and we cannot determine in this study T cell responses in granulomas at the time of dissemination.

PET/CT imaging allows us to interrogate lesional biology over the course of infection, not simply at the time of necropsy. To begin to identify lesional features that correlate with fate, we screened disseminated and contained granulomas for differences in bacterial burden (**Figure 8A**), cumulative bacterial load (CEQ) (**Figure 8B**), and bacterial killing (CFU/CEQ) (**Figure 8C**) at the time of necropsy. None of these parameters, determined by necessity only 15-20 weeks post-infection, predicted the risk of dissemination. Next, we evaluated ^{18}[F]-FDG PET/CT characteristics including FDG avidity (SUVR) (**Figure 8D**) and size (**Figure 8E)** for each granuloma over the course of infection. While SUV, a measure of metabolic uptake of ^{18}[F]-FDG and a proxy for inflammation, did not reveal any differences between the two lesion fates, granuloma size as measured by PET/CT early in infection (at 4-5wks) differentiated the two fates ($p=0.0154$); larger lesions early in infection were associated with a higher risk of dissemination and formation of new culture positive lesions.

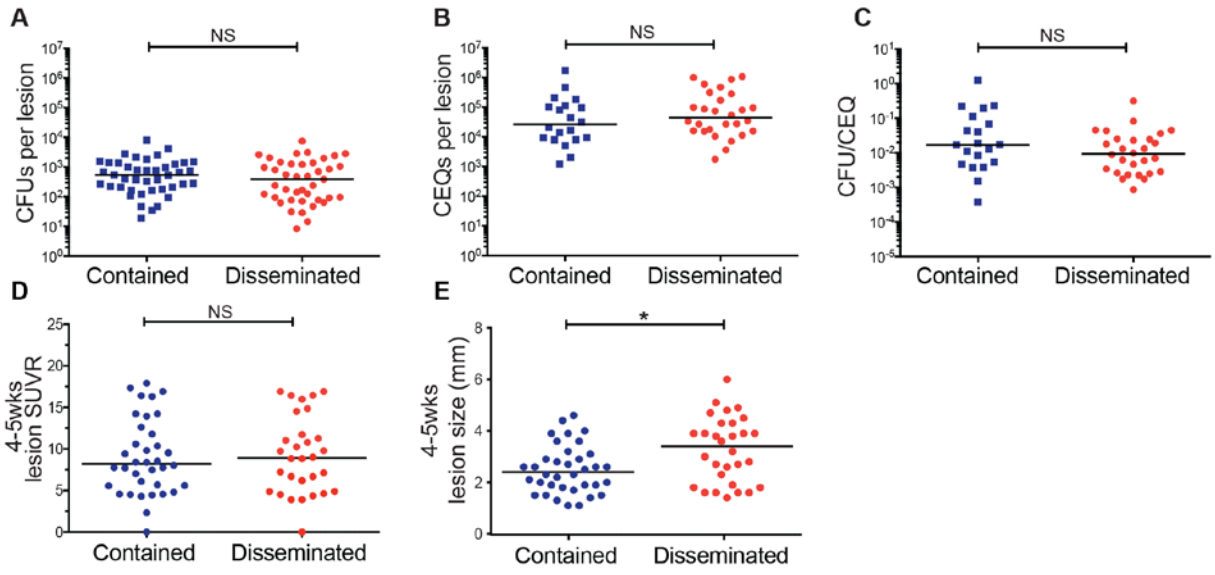


Figure 8. Granuloma size early differentiates dissemination.

A. Bacterial burden (CFU) from granulomas obtained at necropsy (15-20 WPI) of contained (n=41) and disseminated (n=42) lesions from four barcoded macaques. Data are reflective of lesions initiating dissemination as determined by temporal PET/CT analysis. In instances where the precise founder is not known, all involved lesions were classified as disseminated. B. Total bacterial load (live + dead; CEQ) of contained (n=20) and disseminating (n=28) lesions. C. Bacterial killing determined by CFU/CEQ ratio for contained (n=20) and disseminating lesions (n=28). D. Granuloma FDG avidity (SUVR) and E. size (mm) of contained (n=36) and disseminating (n=30) lesions at 4-5 weeks post infection as assessed by PET/CT (*p=0.0154). Panels A-E: each symbol is a granuloma. Statistics for A-E: Mann Whitney test.

3.5 PRELIMINARY DISSEMINATION IN EARLY INFECTION

To assess this commitment to fate in early infection, we infected two additional macaques with 14 CFU of barcoded Erdman and necropsied the animals between 4-5 weeks with PET/CT scans at 3 and 4 weeks. As we did in the late-term macaques, we isolated individual lesions at necropsy using pre-necropsy scans and sequenced barcode identity from plated CFU. Both

animals had fewer stable lesions than spreader lesions, with the majority of spread occurring between the lung and lymph nodes (**Figure 9**) rather than in formation of new lung granulomas. Differentiating disseminating and contained lesions in these early macaques not only confirmed our previous association between increased granuloma size and spread (p=0.0010) but also revealed a correlation between bacterial burden and spread (p=0.0130) (**Figure 10)**. There was no correlation with FDG avidity (SUVR) in these time points between the two types of lesions (**Figure 10**).

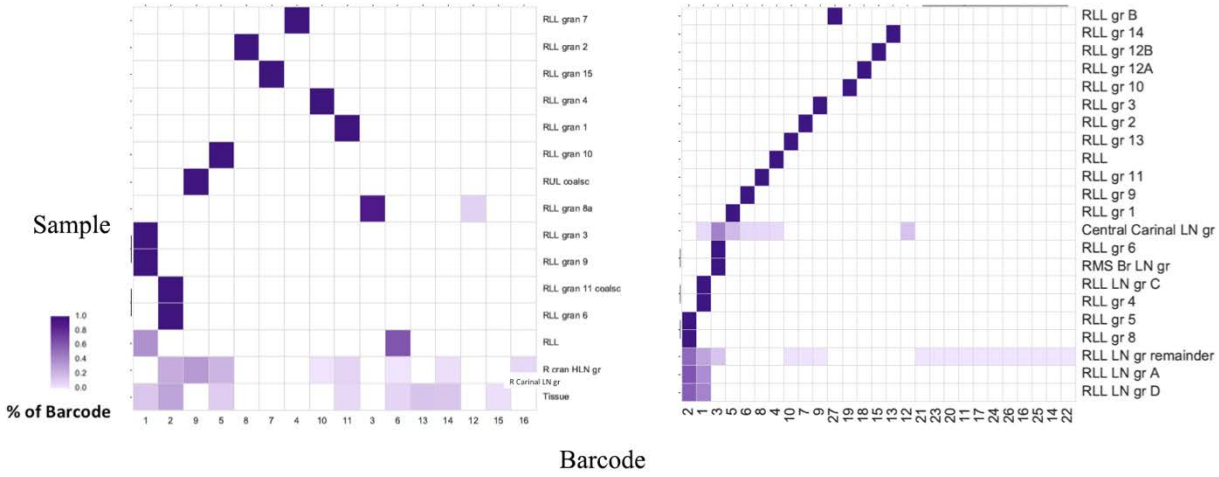


Figure 9. Early dissemination localizes to the lymph nodes.

Left panel: barcode map, macaque ID: 17513; right panel: barcode map, macaque ID 15614. Images courtesy Vivian Leung, HSPH.

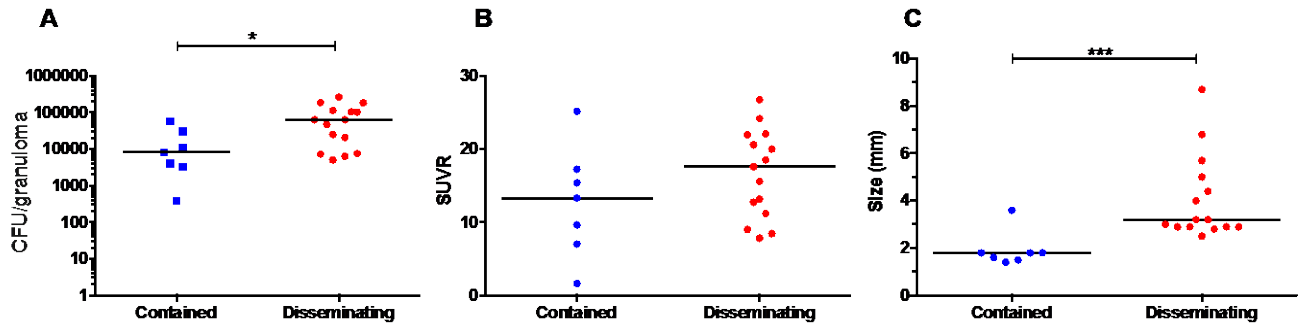


Figure 10. Granuloma size and bacterial burden differentiates dissemination in early infection.

A. Bacterial burden (CFU) from granulomas obtained at necropsy (4-5 WPI) of contained (n=7) and disseminating (n=15) lesions from four barcoded macaques (*p=0.0262). Data are reflective of lesions initiating dissemination as determined by temporal PET/CT analysis. In instances where the precise founder is not known, all involved lesions were classified as disseminated. B. Granuloma FDG avidity (SUVR) and C. size (mm) of contained (n=7) and disseminating (n=15) lesions at 4-5 weeks post infection as assessed by PET/CT (***p=0.0010). Panels A-C: each symbol is a granuloma. Statistics for A-C: Mann Whitney test.

3.6 DISCUSSION

To resolve the within-host bacterial population dynamics that contribute to the spectrum of *M. tuberculosis* infection, we generated a novel library of digitally barcoded *M. tuberculosis* and an associated Python pipeline for their identification and tracking in biological samples from macaques. This barcode library and accompanying pipeline overcome significant hurdles inherent in this type of analysis. By pairing information from serial PET/CT scanning with the sequence identities of bacteria from infected tissue, we have an unparalleled level of resolution of within-host infection dynamics at a lesional level. Genome tagging has been used successfully to ask many questions about the population behavior of bacterial and viral

infections, but to date all of these strategies involved tracking of pools of known barcodes via exact sequencing matching to distinguish barcodes. We have developed an approach to accurately identify and quantitate individuals pulled from a highly complex library where the sequence identifiers are not known *a priori*. These tools allow us to confidently track the descendants of individual bacteria and the population as a whole.

In this study, we sought to define constraints on the bacterial population from inoculation through the first five months post-infection. Importantly, we found no evidence of significant impediments to the establishment of *M. tuberculosis* infection, in that the number of individual bacteria identified by digital barcode as well as the number of initial granulomas observed by PET/CT is similar to the estimated infectious dose. While this study cannot assess all constraints on the bacterial population associated with transmission—especially those associated with aerosolization—it is consistent with the hypothesis that the natural infective dose of *M. tuberculosis* could be as low as a single organism[145,146]. These data highlight the limited capacity of the earliest innate responses to prevent establishment of infection.

Though the innate response does not appear to limit initial infection, it does appear to constrain early dissemination. Indeed, we found evidence of significant constraints on dissemination, indicating that there is a bottleneck at the level of the granuloma. In most animals, only a minority of granulomas seed a second, culture positive granuloma. Previous work in the zebrafish model found that in contrast, dissemination was the rule and not the exception[31].

Our data suggest additional facets to the lesional variability that we have previously documented in terms of bacteriocidal capacity[132]. We found that there is variability in the extent of dissemination that is an independent marker of disease course, distinct from total bacterial,

clinical, or radiographic measures of disease. We demonstrated that granuloma size assessed by ^{18}[F]-FDG PET/CT in the first few weeks post-infection correlates with risk of dissemination. Size is not a proxy for higher numbers of bacteria in a granuloma (data not shown), which might be a simple explanation for the relationship with dissemination risk. Instead, we speculate that size is an independent biologic feature and that physical expansion increases the bacterium's access to avenues of spread. Because most instances of productive dissemination that we identify are to sites less than 10mm away, it suggests that productive dissemination throughout the lung rarely occurs through the blood or lymph.

To assess dissemination early, we infected two monkeys with barcoded *M. tuberculosis* and necropsied them at 4-5 weeks post infection. Our preliminary findings suggested that dissemination is linked to granuloma size, as we saw in our longer-term monkeys, and, more interestingly, bacterial burden. We hypothesize that prior to the onset of a robust adaptive response at around 4-6 weeks[27,36], a proportion of granulomas support robust bacterial growth, enabling early spread, perhaps following inadequate innate containment and bacterial evasion. Recalling that the quantity of these early spreading events influences host outcome[39], we propose that a critical axis exists between a lesion's bacterial burden, physical capacity, and granuloma fate that relies on early bacterial control to minimize spread. A more in-depth analysis of early and late dissemination is planned to better understand the basis for dissemination and how it may relate to early infection.

Interestingly, spread to lymph nodes is also highly variable across infected macaques[9,132,133]. Only 15% of barcodes were represented in the lymph nodes that were culture positive (2 macaques). One current limitation of this work is that we are only able to confidently assess and sequence barcodes from live bacteria present at necropsy, i.e. those that grow on

plates, and not directly from tissue homogenate. This is in part due to the paucity of bacterial genomic material in these samples, which makes the distinction between true and erroneous barcodes in the amplified library challenging. Clearly, productive infection of the lymph node by every bacterial population in the animal is not required for the lymph node to sample and present the antigens from these populations. However, these data do raise questions about the extent of compartmentalization during *M. tuberculosis* infection, and how that might influence generation of adaptive immune responses.

Questions like those of antigenic compartmentalization are only important if there are important functional differences between the bacteria in different sites within the host. Our work to date has demonstrated that lesional bacterial populations are isolated from the earliest points in infection. Our previous studies of bacterial mutability[35] have demonstrated that the rates of genetic mutation are too low to generate sufficient variation to account for observed differences in lesional course at these early time points during infection – though they don't preclude genetic differentiation over longer time periods. Furthermore, we cannot rule out the possibility of distinct patterns of transcriptional adaptation or epigenetic inheritance in these isolated bacterial populations that will be the subject of future studies. One important lesson from these studies is the fact that the barcodes recovered from gastric and bronchoalveolar lavages, which are taken as proxies for the sputum in NHP, represent only a fraction of the bacterial barcodes (3.75%) present in the animal (**Figure 6C & 6D**). These data are consistent with clinical studies demonstrating that organisms with different drug resistance patterns can be isolated from the sputum of individual patients and highlight the need to be cautious in interpreting the sputum bacteria as representative of the entire bacterial population within an individual[147,148].

Previously we reported heterogeneity of granuloma features and bacterial control within a single infected animal and here we reveal another layer of heterogeneity and complexity between different granulomas. Our findings reveal novel avenues of research to probe both host and bacterial factors that influence such disparate granuloma fates. These new tools will allow further characterization into the dynamic local interactions of bacteria and host that ultimately govern both lesion fate and patient outcome. Furthermore, the use of genome barcoding and tagging technologies has broad application, from development to cancer, and the new tools presented here will provide solutions to digitally tracking cells in a variety of fields.

3.7 ACKNOWLEDGEMENTS

We thank the members, laboratory technicians and veterinary staff of the Flynn and Fortune laboratories for their technical expertise and assistance with animal care, sample processing, and study design. This work was made possible by support from Aeras (JLF, SMF) and the National Institutes of Health, R01 AI114674 (JFL, SMF) and T32 AI089443 (AMC). Support was also provided by the Burroughs Wellcome Foundation (SMF), and the Harvard Center for AIDS Research P30 A1060354 (CJM, SMF).

4.0 PRIMARY INFECTION WITH *MYCOBACTERIUM TUBERCULOSIS* CONFERS PROTECTION TO SECONDARY INFECTION IN THE MACAQUE MODEL OF TUBERCULOSIS

This chapter is adapted from the manuscript in preparation:

Anthony M. Cadena, Constance J. Martin, Vivian W. Leung, Philana Ling Lin, Pauline Maiello, Nathan Hicks, Michael R. Chase, Sarah M. Fortune[#] & JoAnne L. Flynn[#], Primary infection with *Mycobacterium tuberculosis* confers protection to secondary infection in the macaque model of tuberculosis, manuscript in preparation, 2017. [#]Corresponding authors.

4.1 INTRODUCTION

It is increasingly recognized that individuals are repeatedly infected with *M. tuberculosis*. Estimates place reinfection TB as high as >50% in individuals that have successfully completed treatment for an initial episode of TB[149-152] emphasizing its prevalence and highlighting the discordance between a previous infection, clearance, and renewed susceptibility[152,153]. In addition, recent studies have indicated that cases of multiple isolate infection (mixed infection) can occur in around 2-18% of active TB patients[151,152] and that such cases can result in disparate drug sensitivities complicating disease management[150,151,154]. Further confounding these epidemiological estimates is the unknown burden of reinfection among individuals with latent

TB infection (LTBI)[155]. How all these previous infections with *M. tuberculosis* influence subsequent infection is poorly understood and whether such instances confer protection or exacerbate pathology remains a key question in TB.

Early observations among entering nursing students in the 1920s suggested that those students that were initially PPD negative were at higher risk to develop active TB relative to those that entered with PPD positivity, and thus implicate a level of protection in LTBI. Subsequent studies have attempted to characterize the level of protection but have had variable results[13,156,157]. A more recent comprehensive survey of the epidemiology literature by Andrews *et al* suggests that latent TB infection in humans may reduce the risk of progression to active TB after re-exposure by as much as 79% compared to uninfected individuals[13], however these epidemiological studies are limited in that the variables are notoriously difficult to control. Appropriately evaluating the effect of LTBI in this context is crucial as drug treatment within this group, which occurs in some regions of the world[10,158], may alter or subvert this apparent protection. A recent LTBI study in South African gold miners by Churchyard *et al* demonstrated diminished protection following termination of preventive therapy implicating a potential requirement for an ongoing infection for protection[159]. In an experimental setting with drug treatment, a mouse model of reinfection using the virulent W-Beijing *M. tuberculosis* strain HN878 found that mice were transiently resistant to reinfection in this animal model[160]. Additional study is needed to adequately determine whether previous exposure or infection with *M. tuberculosis* provides protection against establishment of a second infection, which is very difficult to assess in humans unless there is progression to disease, as well as against disease.

4.2 METHODS AND MATERIALS

4.2.1 Macaque infections, PET/CT imaging, and tissue excision

Fourteen adult cynomolgus macaques (*Macacca fasicularis*) were obtained from Valley Biosystems (Sacramento, California) and screened for *M. tuberculosis* and other comorbidities during a month-long quarantine. Each macaque had a baseline blood count and chemical profile and was housed according to the standards listed in the Animal Welfare Act and the Guide for the Care and Use of Laboratory Animals. The animals were separated into two cohorts: 8 macaques were assigned to reinfection and 6 were assigned to 4-week only controls. All animals were infected with barcoded strain Erdman *M. tuberculosis* via bronchoscopic instillation as previously published[49,125]. The infection schema is provided below in **Figure 11** with the 4-week only animals receiving only barcoded library B in a series of matched infections with their reinfection counterparts. All animals received an inoculum of >15 CFU (determined by plating a sample of the inoculum and counting CFU after 3 weeks) with the precise details listed in **Table 2**. The animals were further subdivided such that 4 reinfection animals were directly paired with 2, 4-week only control animals; 2 additional 4-week only animals were infected separately to supplement early infection data (**Table 2**). Each macaque was followed with serial 2-deoxy-2-[^{18}F]-fluoro-D-glucose ([^{18}F]-FDG) PET/CT imaging as previously described[39,109,132] to identify and track lesion formation and progression over time. As before, lesions were individually characterized by their date of establishment (scan date), size (mm), and relative metabolic activity as a proxy for inflammation ([^{18}F]-FDG standard uptake normalized to muscle [SUVR]). Lesions $\geq$ 1mm can be discerned by our PET/CT imaging analysis. To avoid barcode cross-contamination, individual granulomas were separately excised and processed using the animal's

prenecropsy scan as a guide for identity and location. All lesions observed post second infection in the reinfection cohort were initially labeled as secondary lesions pending library qTag verification.

4.2.2 Isolation and preparation of bacteria from tissue samples

Following removal at necropsy, each lesion was homogenized and plated for bacterial burden on 7H11 agar supplemented with oleic albumin dextrose catalase (OADC). A small portion of homogenate was frozen for qTag sequencing as well as chromosomal equivalent (CEQ) analysis. Genomic DNA was extracted using

4.2.3 T cell flow cytometry profiling and intracellular cytokine staining

At necropsy, granulomas, lung tissues, and thoracic lymph nodes were processed into single cell suspensions with sterile PBS. A portion of these suspensions were used for T cell profiling and characterization. T cells were initially stained for viability using the LIVE/DEAD Fixable Blue Dead Cell Stain Kit (ThermoFisher Scientific) and then stained for T cell-specific markers using the anti-human CD3 (clone SP34-2, BD Pharmigen), anti-human CD4 (clone L200, BD Biosciences), and anti-human CD8 (clone RPA-T8, BD Biosciences) antibodies in a standard FACs buffer.

For intracellular cytokine staining, tissue suspensions were stimulated with *M. tuberculosis* CFP10 and ESAT6 (BEI Resources, Manassas, VA) and Brefeldin A (GolgiPlug, BD Biosciences) in RPMI (Lonza, Walkersville, MD) supplemented with 1% L-glutamine and 1% HEPES (Sigma, St Louis, MO) for 3.5 hours. The final concentration of the *M. tuberculosis*

specific peptides was 2.5μg/mL. Following surface staining as detailed above, the cells were fixed and permeablized (BD Cytofix/Cytoperm) and washed with BD Perm/Wash buffer (BD Biosciences). Incubations were all performed according to manufacturer recommendations. Cells were then stained with the following intracellular cytokines: anti-human IFNγ (clone B27, BD Biosciences), anti-human TNFα (clone Mab11, eBiosciences), anti-human IL-2 (clone MQ1-17H12, BD Biosciences), IL-6 (clone MQ2-6A3, BD Biosciences), anti-human IL-10 (clone JES3-9D7, eBiosciences), and anti-human IL-17A (clone eBio64CAP17, eBiosciences).

Flow cytometry was preformed on a LSR II (BD) and analyzed using FlowJo Software v.9.8 (Treestar Inc., Ashland, OR). Size (FSC) and granularity (SSC) were used to isolate the lymphocyte population at the cytometer. All cytokine data presented are gated on $CD3^+$ T cells.

4.2.4 Statistical analysis

Statistical analysis was performed in Graphpad Prism, JMP and R, significance was defined as $p < 0.05$.

4.3 APPROACH

To evaluate the contribution of an ongoing infection on subsequent infection in TB, 14 macaques were separated into two sets, a reinfection set and a matched 4 week only set (**Table 2**). The 8 animals in the reinfection set had a primary infection for 16 weeks followed by a secondary infection for 4 weeks; each infection was carried out with a separate barcoded library. The 6 animals in the four week-only controls animals received only the second barcoded library

(**Figure 11**). The infections of the second group of monkeys were performed in parallel with the second infection of the reinfection cohort, except for 2, 4 week only animals that formed a third infection group used to supplement our early infection data (**Table 2**). This study design was critical to compare reinfection in a setting controlling for dose and inoculum variation such that 4 week secondary granulomas could be directly evaluated against 4 week primary granulomas, a critical time point where granuloma bacterial burden is highest[36], and, thus, a potential phenotype most evident. As of the date of this writing, primary and secondary granulomas were distinguished using [^{18}F]-FDG PET/CT imaging[39,109,132], pending library qTag validation through sequencing. Consequently, any new lesions that were observed after the second infection were termed "secondary lesions" resulting in an unbiased but perhaps overestimated contribution of granulomas arising from the second infection in some animals. We expect that sequence analysis will eventually show that some lesions were actually due to spread from the primary infection in some animals.

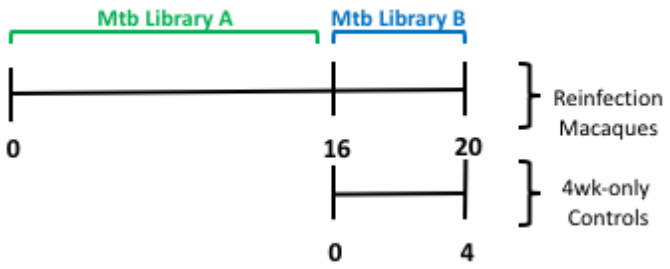


Figure 11. Schematic of Reinfection Study Design.

14 animals were separated into reinfection and 4wk only controls with 8 and 6 macaques, respectively. Animals in the first group were infected with two successive libraries of barcoded *M. tuberculosis* (A and B) with a 16 week interval between them. Animals in the second, 4 week only control group were infected with only library B in a series of matched infections with the reinfected animals. (All numbers above in weeks).

Table 2. Parameters of macaque infections, imaging, and disease.

Animal ID	Strain	Infection Dose (CFU)	Time to Nx (wks)	Gross Pathology Score	Total CFU	PET/CT Scans (wks)
18915	Library B 2.0 Barcoded Mtb Erdman	6±1	4	7	318,039	2, 3
19015		6±1	4	17	1,738,747	2, 4
19915		5±1	4	19	564,310	2, 4
20015		5±1	5	11	104,809	3, 5
5616		10±1	5	32	370,409	4, 5
5716		10±1	4	20	461,740	3, 4
19115	1°: Library A 2.0 Barcoded Mtb Erdman 2°: Library B 2.0 Barcoded Mtb Erdman	1° = 5±1 2° = 5±1	20	23	37,756	4, 6, 8, 12, 15, 18, 20
19215		1° = 5±1 2° = 5±1	20	22	6719	4, 6, 8, 12, 15, 18, 20
19315		1° = 5±1 2° = 5±1	21	12	12,196	4, 6, 8, 12, 15, 18, 21
19415		1° = 5±1 2° = 5±1	21	22	4,334	4, 7, 8, 13, 15, 18, 21
19515		1° = 8±1 2° = 10±1	20	38	361,420	4, 8, 12, 16, 20
19615		1° = 8±1 2° = 10±1	20	57	122,872	4, 8, 12, 16, 20
19715		1° = 8±1 2° = 10±1	21	34	10,600	4, 6, 8, 12, 16, 19, 21
19815		1° = 8±1 2° = 10±1	21	28	3,331	5, 6, 8, 13, 16, 19, 21

4.4 RESULTS

4.4.1 Primary infection protects against secondary infection by limiting granuloma formation and reducing granuloma bacterial burden

Using [^{18}F]-FDG PET/CT to count the numbers of new, secondary lesions (i.e. those that only become apparent subsequent to infection with Library B), we observed a trend towards an overall reduction in the numbers of new granulomas formed from the second *M. tuberculosis* library (+p=0.2135) (**Figure 12A**). When examined by monkey and separated by infection

group (**Figure 12B**), the granuloma reduction becomes more apparent with the most surprising finding that four macaques in the reinfection set exhibit nearly complete to complete protection (ID: 19315, 19415, 19715 & 19815) against establishment of new granulomas. In the first infection group, the estimated CFU for the second library was 5±1, which is close to the threshold of infection despite a conservative, calculated error rate. A more robust reduction was appreciated in the second infection group where the estimated CFU for the second library was 10±1. Here, in spite of an infection dose of 10, we have two macaques that have two and zero new lesions, respectively (**Figure 12B**). The other two monkeys in that group have higher secondary granuloma counts than their 4 week only counterparts, which is likely due, in part, to an actively disseminating primary infection resulting in overestimation by [^{18}F]-FDG PET/CT and await qTag verification for definitive library placement. To our knowledge, this is the first reported instance of complete protection against the establishment of infection observed in a macaque model of tuberculosis, including in vaccine studies.

In addition to assessing the numbers of granulomas formed following a second infection in the context of reinfection, we interrogated the overall bacterial burden of the secondary granulomas that were able to successfully establish at 4 weeks post second infection. We found that bacterial load in these secondary granulomas were several orders of magnitude lower than age-matched primary granulomas, with a significant number having sterilized altogether (53% vs. 15%) (**Figure 13A &13B**). Collectively, our two observations that primary infection limits both the establishment and bacterial growth of secondary granulomas implicates the presence of a specific and capable immune response elicited by an ongoing *M. tuberculosis* infection.

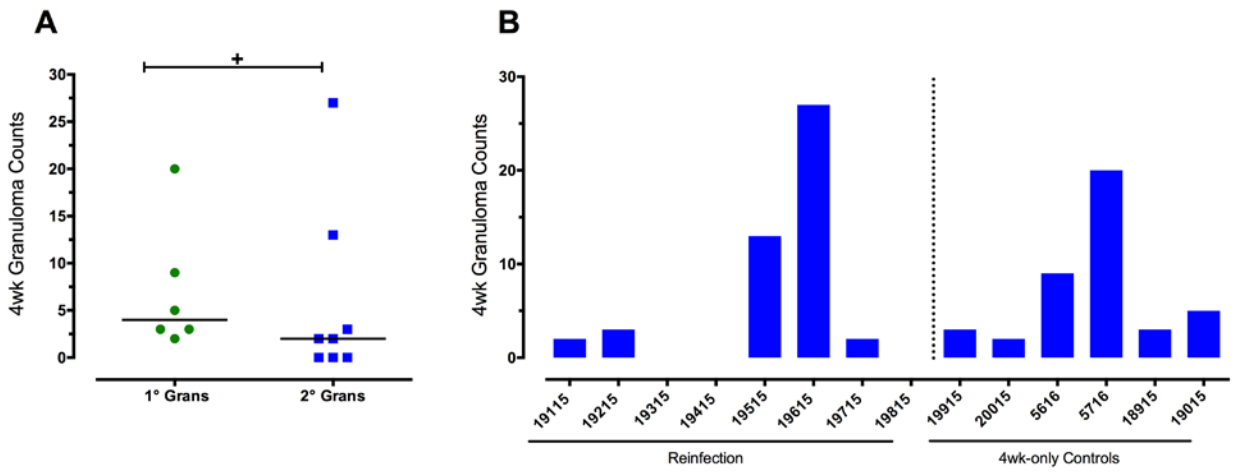


Figure 12. Primary infection limits secondary granuloma development.

A. Granuloma counts of age-matched primary (n=6 animals) and secondary (n=8) lesions (4-5weeks post infection) (p=0.2153); each symbol is a monkey. B. 4 week granuloma counts separated by monkey and infection group. In each group, the animals received the same dose of the second library (*M. tuberculosis* Erdman Library B); group 1 = 5±1 CFU, group 2 = 10±1 CFU, and group 3 = 6±1 CFU. The last group only had 2, 4 week only control animals.

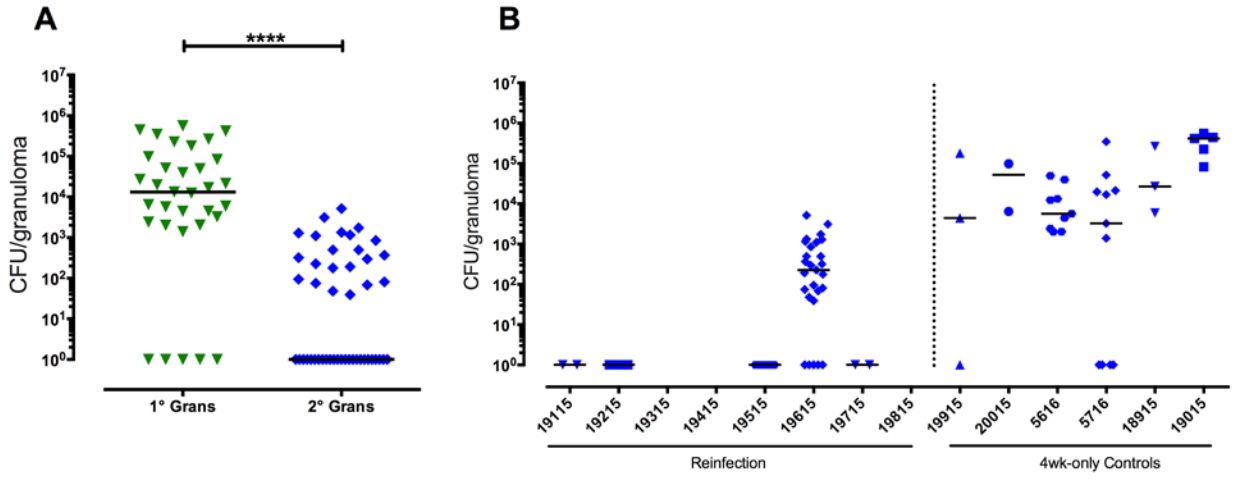


Figure 13. Secondary granulomas have significantly fewer bacteria than primary granulomas.

A. CFU of 4 week, age-matched primary (n=33) and secondary (n=47) granulomas (****p<0.0001), each symbol is a granuloma. B. 4 week granuloma CFU separated by monkey and infection group, as above in **Figure 12**, each symbol is a granuloma and all are from barcoded *M. tuberculosis* library B.

4.4.2 Secondary lesions have more robust cytokine expression than age-matched primary lesions

To begin to dissect a possible mechanism driving the protection observed in secondary infection, we examined age-matched primary and secondary lesions using *ex vivo* stimulated, intracellular cytokine staining. We found that a higher frequency of $CD3^+$ T cells in the 4 week secondary granulomas produced IL-2, IL-6, and IL-10, and a trend towards higher production for IL-17A and TNFα (**Figure 14**). Only IFN-γ did not appear to be upregulated relative to 4 week primary granulomas.

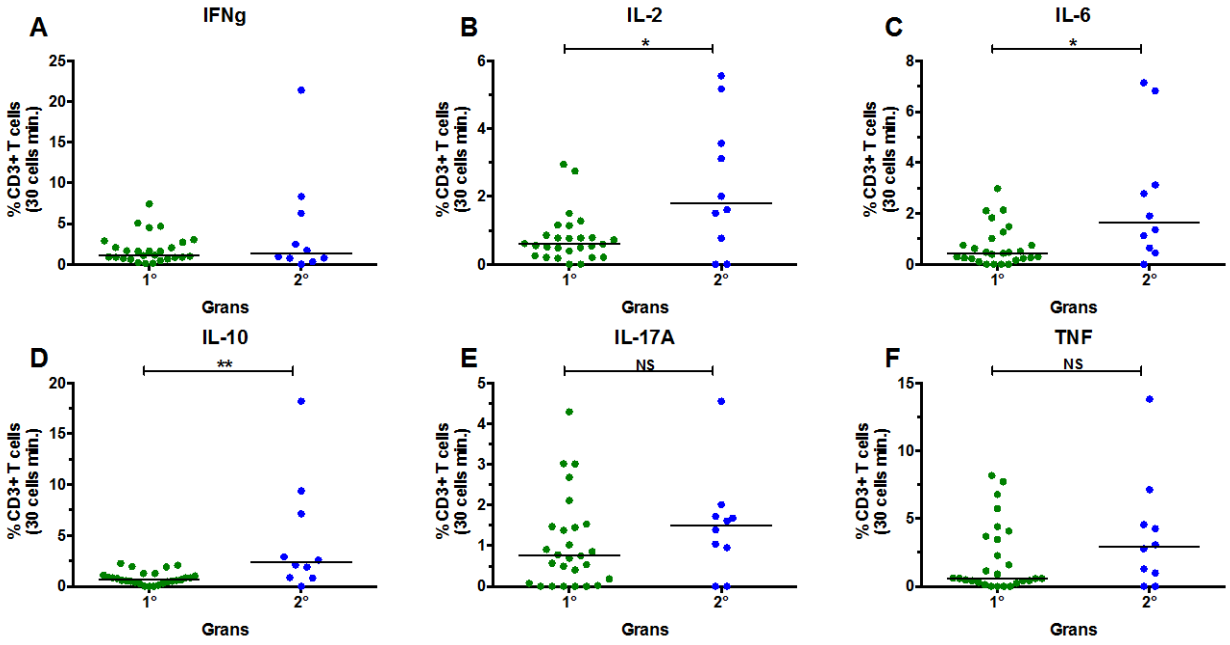


Figure 14. T cells in secondary granulomas produce greater amounts of pro- and anti-inflammatory cytokines.

A-F. $CD3^+$ T cell cytokine expression for IFN-γ, IL-2, IL-6, IL-10, IL-17A, and TNFα, respectively, between 4 week primary (n=27) and 4 week secondary (n=10) granulomas (*p>0.05 and **p>0.01). All plots use a 30 T cell minimum. Each symbol is a granuloma.

4.5 DISCUSSION

In this study, we addressed reinfection in the macaque model of TB by directly comparing infection establishment, bacterial burden, and granuloma immune profiles in a paired set of macaque cohorts. Our data support enhanced protection against early (4 week) secondary infection following a primary infection. Importantly, this is a critical time point in infection as clearance of *M. tuberculosis* is least effective, bacterial dissemination is prominent, and bacterial load is highest[7,27,36]. Most surprisingly, we report the novel finding of complete protection against the establishment of secondary infection in several macaques (**Figure 12**) suggestive of highly effective mycobacterial immunity in the lung and airway. Moreover, in the secondary granulomas that were able to form, we report >1000 fold bacterial killing further indicative of a potent, antibacterial state in these granulomas, with the median bacterial burden at 0 (**Figure 13**). The finding that a large portion of 4 week granulomas was sterile is surprising as our previous data has shown that granuloma bacterial load in primary infection is greatest in the first several weeks of infection[36], most likely due to the delay in adaptive immunity[27]. We hypothesized that the enhanced killing in the reinfected animals was a result of the local lung immunity generated by a primary infection and our initial T cell cytokine profiling of the 4 week, secondary granulomas revealed significant upregulation of both pro- and anti-inflammatory cytokines (**Figure 14**). We speculate that the subsequent, pan-hyperresponsive cytokine state promoted a balanced immune microenvironment in which both antibacterial responses and immunopathology were maintained appropriately[30]; two immune components that have previously been shown to best correlate with granuloma sterility[133]. Subsequent studies in reinfection will carefully probe the immune profiles and functionality of both innate and adaptive factors in the lung and airway to better understand the immune signatures providing this

protection. Further study is required to determine the durability of protection against *M. tuberculosis* seen here as well as the apparent necessity for live *M. tuberculosis* to elicit the appropriate immunity, as the current BCG vaccine does not engender such protection[2,3]. Ultimately, our work in reinfection provides exciting clues for novel immunological biomarkers of protection that will inform more effective vaccine development[6,161].

4.6 ACKNOWLEDGEMENTS

We thank the members, laboratory technicians and veterinary staff of the Flynn and Fortune laboratories for their technical expertise and assistance with animal care, sample processing, and study design. This work was made possible by support from Aeras (JLF, SMF) and the National Institutes of Health, R01 AI114674 (JFL, SMF) and T32 AI089443 (AMC). Support was also provided by the Burroughs Wellcome Foundation (SMF), and the Harvard Center for AIDS Research P30 A1060354 (CJM, SMF).

5.0 THE INFLUENCE OF *M. TUBERCULOSIS* ON THE LUNG MICROBIOME IN THE MACAQUE MODEL OF TUBERCULOSIS

This chapter is partially adapted from the accepted NIH grant RO3 AI122067 (Flynn JL, Lin PL, and Ghedin E, 2015) and represents an ongoing study probing the relationship of the lung microbiome and *M. tuberculosis* infection in the macaque model of tuberculosis that I initiated.

5.1 INTRODUCTION AND INNOVATION

Microbial dysbiosis is increasingly being recognized as a significant factor influencing human disease in a variety of contexts spanning allergy, autoimmunity, cancer, and infectious disease[162]. The gut microbiome, in particular, has had the greatest amount of literature linking the immune system and an organ's microbiome[163] but it is increasingly becoming clear that other human immune surfaces and interfaces including the lung are shaped by host-microbe interactions[164,165]. Specific human studies linking lung microbiome changes and bacterial infections are limited, but earlier work in *Pseudomonas aeruginosa*[166] has identified an association between airway microbial loss and pathogen colonization suggesting that microbial diversity can influence pathogen selection and disease at this site.

Here we explored the effects of *M. tuberculosis* infection on the lung microbiota over the first several months of infection in a macaque model of TB, using serial bronchoalveolar lavage (BAL) samples obtained from animals funded on other studies. This initial exploration of the dynamic changes in microbiota in response to *M. tuberculosis* infection, and preliminary correlations with outcome measures of infection, will provide novel data and set the stage for more focused future investigations of the influence of lung microbiota on *M. tuberculosis* infection.

In this study, we coupled our cynomolgus macaque model of TB[125] with the unique ability to serially sample the airways of animals throughout the course of infection. We exploited the technologies that we have developed with respect to PET/CT imaging, immunologic analyses, clinical assessments, prediction of outcome, and genomic sequencing analyses[39,49,53,167-170] to determine the duration of effect of *M. tuberculosis* infection on the lung microbiome. This is the first targeted exploration of the lung microbiome in a model of TB that replicates human infection outcomes and pathology.

5.1.1 Tuberculosis and the microbiome

Almost nothing is known about the interaction of the lung microbiome and *M. tuberculosis*. *M. tuberculosis* causes a chronic lung infection as primary disease can take up to 2 years to be clinically diagnosed and humans with latent TB can remain infected for their lifetime. Reactivation of clinically latent infection can occur years to decades after initial infection. Cross-sectional human studies examining the interaction of *M. tuberculosis* infection and lung microbiota (measured by sputum) have been limited but suggest that TB patients have altered diversity in their sputum microbiota compared to uninfected humans[171,172]. The distribution of

some genera (*Stenotrophomonas* and *Phenylobacterium)* was unique to TB patients[172]. Certain patterns of microbiota (e.g., presence of *Pseudomonas spp.*) were observed in patients with recurrent TB and treatment failure suggesting that the lung microbiota interacts with both host and pathogen[173]. Specifically, alterations in *Treponema* and *Atopobium* were associated with recurrent TB suggesting that an altered "normal" microbiome is associated with ongoing susceptibility to TB[173]. Studies examining the microbiota of the oropharynx between health controls and TB patients showed differences in diversity and abundance of particular organisms although no differences in major phyla were observed[174]. Distinct changes in the gut microbiome were observed after experimental *M. tuberculosis* infection in a mouse model[175], although the effects of these changes on course of infection are not clear. Perry *et al* showed that seroprevalence of *H. pylori* in humans was associated with a decreased risk of progression to TB disease and this pattern was recapitulated in our cynomologus macaque model of *M. tuberculosis* infection[176].

5.2 METHODS AND MATERIALS

5.2.1 Microbiome sample collection and *M. tuberculosis* infections

To sample the oral flora, we first obtained a 5mL rinse of the cheek pouches by administering 5ml of PBS by syringe and then recovering as much liquid as we could. We followed the oral cavity washing with a 5mL saline wash of the sterilized bronchoscope to detect any residual DNA within the scope. To minimize subsequent lower airway broncheoalveolar lavage (BAL) sampling, we swabbed each animal's mouth with an antiseptic agent (chlorohexane). A 7mL

saline lavage of the uninfected lower lung (often left lung) was performed, followed by sterilization of the scope with antiseptic solution (Cidex, Civco Medical Solutions) before sampling the opposite infected lung. From these lavages, we collected approximately 3-5mL from each lobe and immediately aliquoted into 2-4.5mL cryovials and froze the vials at -80°C. Thus, from each monkey at each time point, we had 4 samples (mouth, scope, right lower lobe, left lower lobe) (**Figure 15**). Our serial sampling was done at pre-infection (0), 1, 4 and 5-months post-infection (several macaques had sampling only out to 4 months due to the onset of drug treatment). Monkeys were infected with <25 CFU *M. tuberculosis* strain Erdman via bronchoscopic instillation into the right lower lung, as we have previously described[42,49]. The animals for this project are from separately funded studies of Dr. Flynn and Dr. Lin, and are un-manipulated except for *M. tuberculosis* infection. Samples are analyzed using target gene sequencing of the V4 hypervariable region of 16S rRNA for bacteria[169] (see below). To date, we have obtained serial samples from 26 *M. tuberculosis* infected cynomolgus macaques and a further 10 will be sampled in the following year.

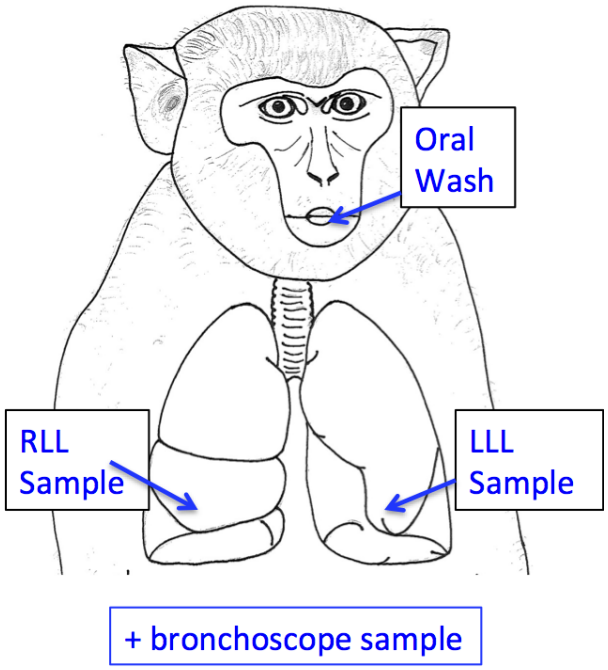


Figure 15. Schematic of macaque airway sampling.

Ongoing clinical assessments (e.g., cough, appetite or weight loss), erythrocyte sedimentation rate (ESR, marker of systemic inflammation) and gastric aspirate and BAL for *M. tuberculosis* growth were performed to obtain clinical outcome measures as described[42,49]. More importantly, PET/CT imaging of lung disease in all of these animals will be obtained over the course of infection, including at the 4 and 5-month time points to assess disease progression. As noted above, we have identified several parameters of PET/CT change that are correlated with extent of disease.

5.2.2 DNA extraction from BAL samples

To avoid potential contamination, all DNA extractions are performed in a dedicated biosafety level 3 cabinet that is first exposed to a UV light source for 15 minutes. Samples were prepared in batches by monkey, where each sample is thawed and centrifuged, and the PowerSoil® DNA

extraction kit (MO BIO Laboratories, Inc.), which includes a bead-beating step, is used to extract the DNA.

5.2.3 Analysis of microbial community structures

Dr. Elodie Ghedin, an expert in genomic characteristics of the microbiome as they relate to lung infection, and her laboratory preformed the taxonomic analysis of samples. The 16S amplicon sequence data from the samples were processed using the following pipeline: first, we used the Mothur pipeline[177,178] modified for the MiSeq platform, which includes removal of low quality reads, chimeras (UCHIME), non-mate-paired reads and merging of paired-end reads[179]. Taxonomic assignments of the sequences were made using the RDP (Ribosomal Database Project) classifier[180], which classifies sequences to the genus level. Second, we utilized taxonomy-independent analyses through the generation of Operational Taxonomic Units (OTUs), which provide classifications at approximately the species level by clustering sequences based on nucleotide identities of ≥97% using the UCLUST program [181] Assignments of taxonomy were further refined through phylogenetic methods including using PyNAST [182] to generate an alignment for each cluster from which a phylogenetic tree was built using the FastTree approximately-maximal-likelihood method[183]. Third, we used the taxonomic and OTU-based profiles in a series of ordination, clustering, and community structure (measuring richness and evenness) analyses designed to compare and identify significant shifts in 16S rRNA profiles between samples and for further statistical analyses (see Statistical Analyses). Finally, we determined the stability in the structure of the microbiome over time using the ordination, clustering and diversity analyses and comparing data across time points.

5.2.4 Neutral Model Analysis

The data from the oral wash and the bronchoscope control were used to determine which microbe (identified in the taxonomic analysis described above) was a true resident of the lung rather than a microbe from the mouth. To do this, we applied the neutral model of community ecology [169,184]. This model examines whether distribution of organisms in the lung results from dispersal from the mouth or from active environmental selection in the lungs. This analysis allowed us to determine true residents of the lungs and how their representation varies over time (see Statistical Analyses).

5.2.5 Statistical Analyses

Whole community structure changes were analyzed using Adonis, a multivariate ANOVA based on dissimilarities, available in QIIME[185]. Individual OTUs that differ between infected and uninfected lobes were identified with repeated measures ANOVA. To examine the relationship of microbial communities to both systemic marker of inflammation (ESR) and lung TB-specific inflammation by PET/CT, we will use a random intercept model, or linear mixed model, with autoregressive correlation structure; a LASSO penalty will be added to this regression to account for the large number of potential fixed effects (OTUs). On categorical variables, the linear mixed model will be replaced with a logistic mixed model with the same autoregressive correlation structure and penalty parameters. To test the shifts in abundance for specific taxa due to *M. tuberculosis* infection, we also employed smoothing spline ANOVA (SSANOVA) methods (http://cbcb.umd.edu/software/metagenomeSeq). We have previously used these methods in a

separate study on the progression of differential abundance of specific oral microbial taxa during the course of a SHIV infection[186].

5.3 PRELIMINARY RESULTS

Lung microbiome analyses have been performed on the first cohort of 10 macaques with a total of 162 samples (**Table 3**). The second set of 16 macaques is in progress. Sequencing and analysis was performed for all oral wash and BAL samples using the bronchoscope and reagent controls for appropriate thresholding. The data below is representative of the first set of 10 animals. Infection status of each lung lobe is determined by PET/CT at each sampling time point[36,39,109] allowing for direct comparisons of infected and uninfected lobes over time.

Table 3. Lung Microbiome Samples

Cohort	Macaques	Lung BAL	Oral Wash	Bronchoscope Controls	Reagent Controls	Total Samples
1	10	70	36	36	20	162
2	16	114	58	58	32	262

5.3.1 The oral microbiome is distinct from the lung microbiome

Using principle component analysis (PCA), comparison of OTUs between the oral wash and BAL reveals that the oral cavity clusters separately from the lung lobes, irrespective of time point post infection (**Figure 16**). While the infected and uninfected lobes (red and blue, respectively) are spread along the first component (axis 1), the oral samples are tightly clustered with very little spread indicating that the microbial composition of this compartment is distinct from the lower airway.

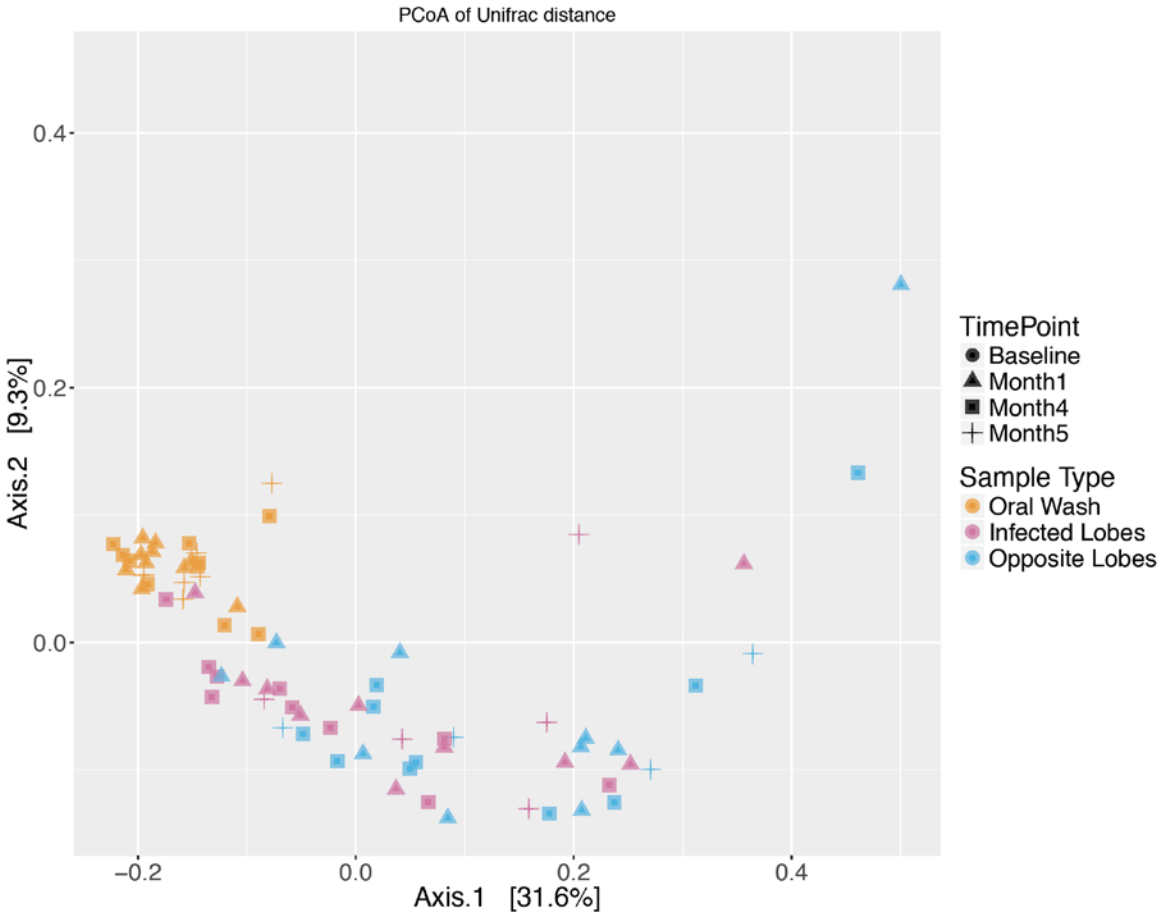


Figure 16. Principle component analysis reveals discrete microbial landscapes between the oral and lung environment.

Figure courtesy Elodie Ghedin, NYU.

5.3.2 The lung microbial community changes over time following infection

To look for differences over time following *M tuberculosis* infection, we compared the alpha-diversity of the OTUs from infected and uninfected lobes (determined by PET/CT) from all macaque at each time point. We find that at 4 months post infection, there is a trend towards greater microbial diversity in the infected lobe (p=0.08) (**Figure 17**). This shift appears to be transient as the diversity then begins to decrease in the following month.

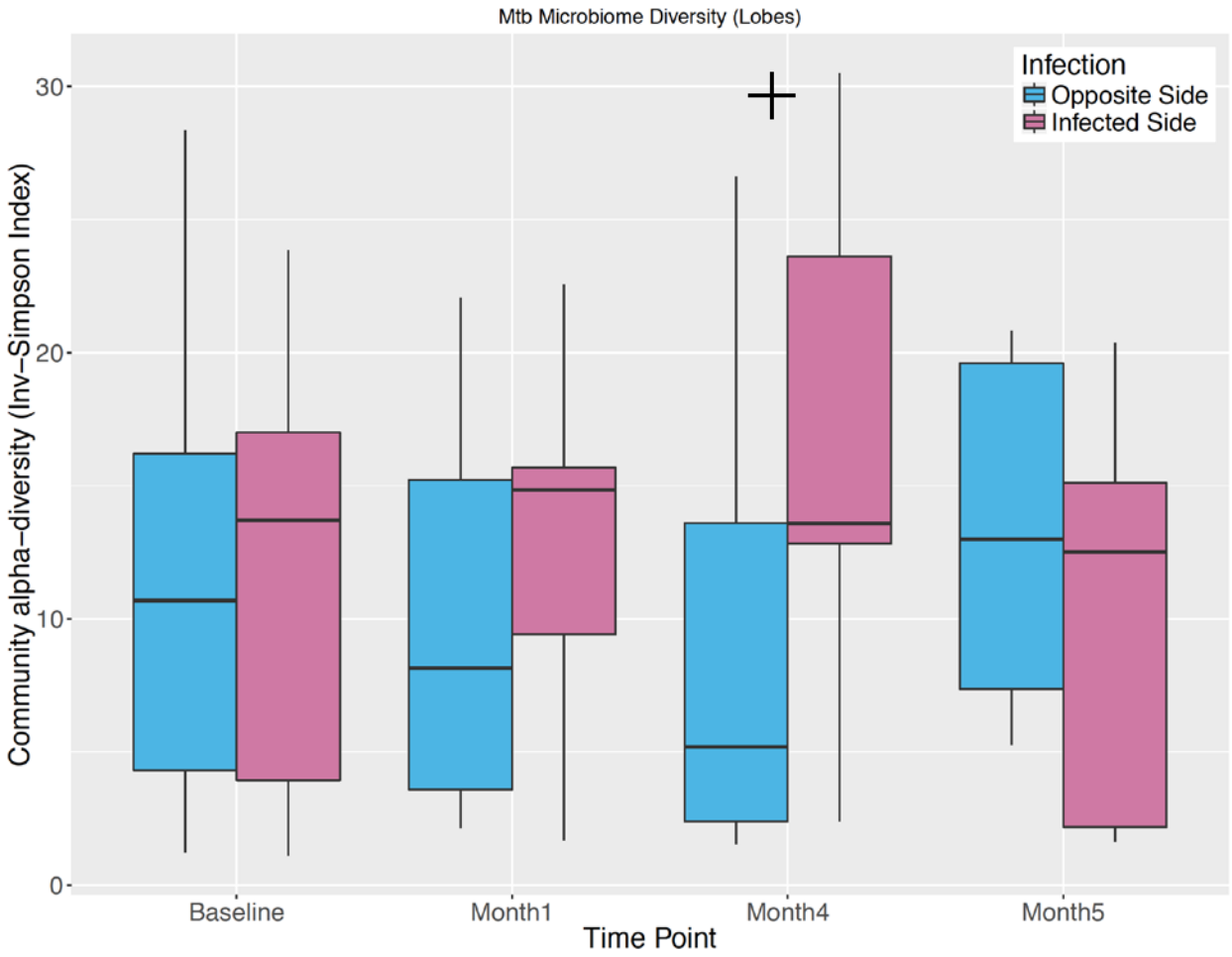


Figure 17. Lung microbial diversity shifts over time with *M. tuberculosis* infection.

Alpha-level diversity of BAL samples throughout infection between uninfected (blue) and infected (red) lobes ($^{+}$p=0.08). Figure courtesy Elodie Ghedin, NYU.

To better understand the specific microbial communities contributing to the change at 4 months, we analyzed the data with the linear discriminate analysis (LDA) effect size (LEfSe) method[187]. This approach helps to determine features responsible for the differences found between class comparisons on a metagenomic scale. Using this method, we determined that in this first macaque set, the bacterial genera *Fusobacterium, Neisseria,* and *Aggregatibacter* were associated with *M. tuberculosis* in the infected lobe and *Burkholderia* in the uninfected lobe (**Table 4**). Subsequent time-dependent Smoothing Spline ANOVA (SSANOVA) analyses validated all of the associations except for *Neisseria* (**Table 4**).

Table 4. Discriminating features between infected and uninfected lobes in *M. tuberculosis* infection

Genus	Lobe	LDA Score	p-value (LEfSe)	p-value (SSANOVA)
Fusobacterium	Infected	4.363	0.003	0.010
Neisseria	Infected	4.221	0.015	-
Aggregatibacter	Infected	4.113	0.004	0.019
Burkholderia	Uninfected	4.336	0.0003	0.010

5.3.3 The oral and lung microbiome are highly variable across macaques

By separating out the ten bacterial genera with the highest relative abundance of OTUs in the infected lobe for each macaque during *M. tuberculosis* infection, we see that there is great diversity across animals (**Figure 18**). Several macaques display little change in the lung

microbiome following infection, whereas others display a high degree of variability, both with respect to magnitude and duration.

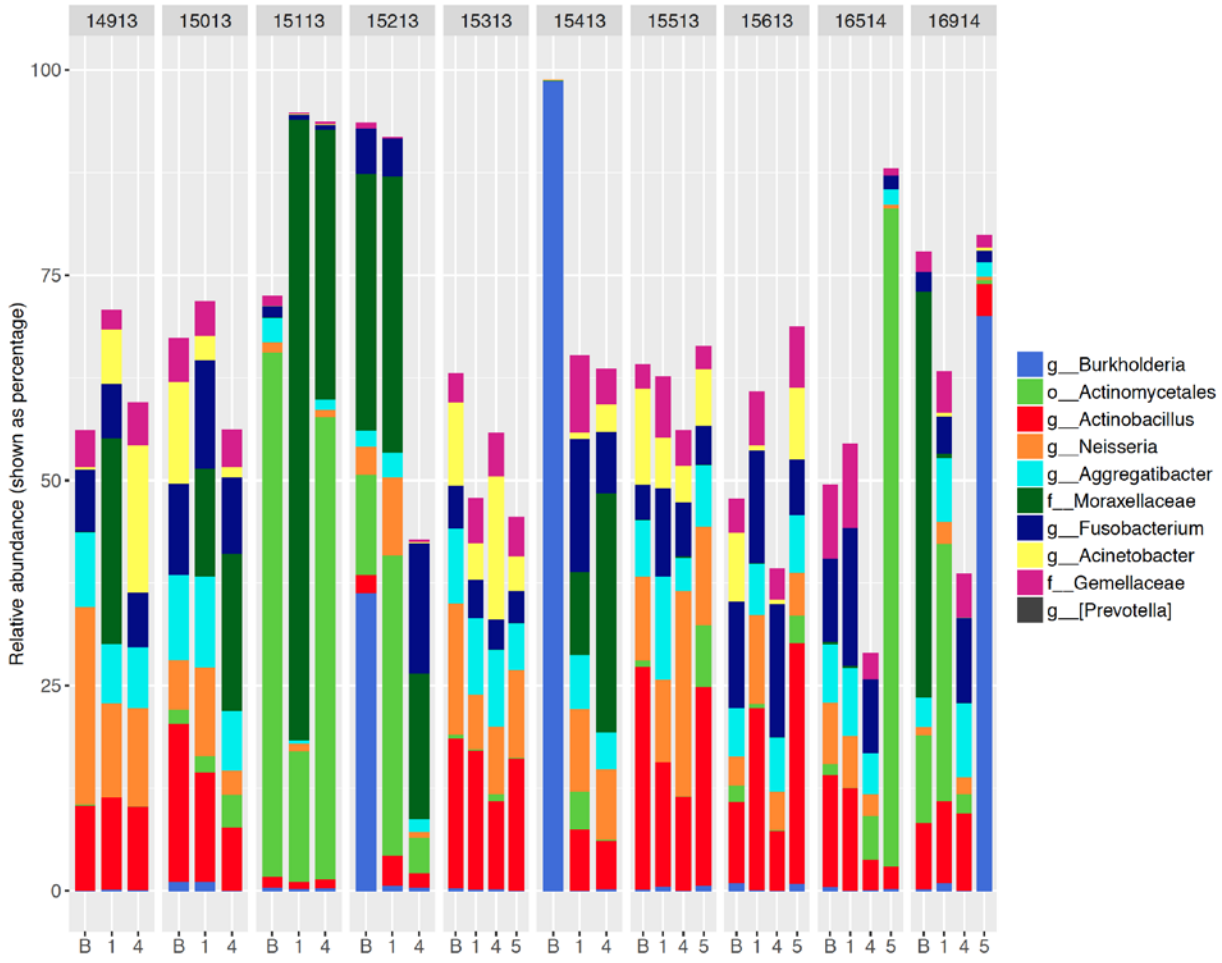


Figure 18. There is a high degree of variability across macaques during infection.

The relative abundance of taxa of the infected lobe for each monkey (ID across top) over the course of *M. tuberculosis* infection. (B=pre-infection, 1=1 month post infection, 4=4 months post infection, and 5=5 months post infection.) Figure courtesy Elodie Ghedin, NYU.

5.3.4 Pulmonary inflammation is highly variable across macaques

To begin to relate pulmonary inflammation with the changes in the lung microbiota, we quantified lung inflammation, as measured by [^{18}F]-FDG avidity (Total PET hot), in each macaque at 1, 4, and 5 months post *M. tuberculosis* infection (**Figure 19**). Each time point coincides with the BAL lung microbiome sampling that was done above. There is a wide range of lung inflammation observed in these animals. Several macaques exhibited large changes in total PET hot at 4 months whereas others animals have relatively little pulmonary inflammation throughout infection. A more precise analysis will be performed on each macaque in which each lobe will be evaluated for lung inflammation, infection status, and microbial diversity.

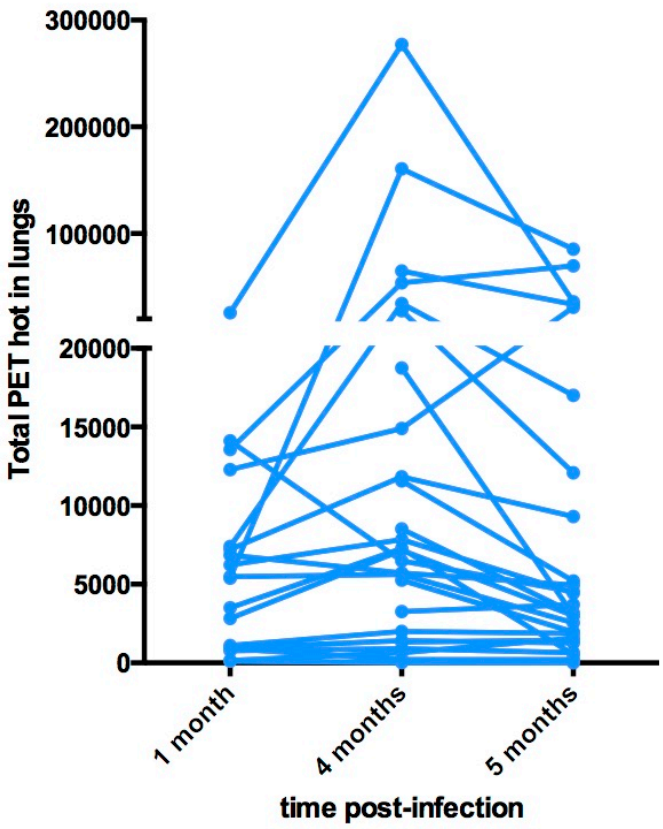


Figure 19. Total pulmonary lung inflammation in infected macaques.

Using [^{18}F]-FDG avidity (total PET hot), we assessed lung inflammation in all macaques (n=26) over the course of *M. tuberculosis* infection.

5.4 DISCUSSION/FUTURE DIRECTIONS

This primary aim of this study is to determine whether *M. tuberculosis* influences the lung microbiome following infection. Analysis of the first cohort of ten macaques reveals that *M. tuberculosis* can induce changes in the lung microbiome of infected macaques, but that the alteration is variable across macaques. Importantly, our study strategy is designed to compare

infected lobes from uninfected lobes within the same macaque across infection to minimize the masking effect of animal-to-animal variability. We find that the changes are specific to the infected lobe, relative to the opposite, uninfected lobe, and greatest in the 4th month post infection (**Figure 17**). We subsequently preformed metagenomic LDA effect size analyses (LEfSe) to better discern the specific microbial genera most likely driving the difference in lobe status and discovered that Fusibacterium, *Neisseria,* and *Aggregatibacter* are associated with *M. tuberculosis* (**Table 4**). Interestingly, these microbes are normally associated with the oropharynx[188] and may suggest opportunistic, migratory colonization following *M. tuberculosis* infection that stems from inflammatory dysbiosis in the lung[129,188]. Dickson *et al* present a model of airway dysbiosis and inflammation in which an insult (infection, allergy, or environmental agent) triggers an inflammatory state that perpetuates a cycle of microbial alteration and respiratory inflammation driven by changes in growth conditions and distorted pulmonary architecture[129] (**Figure 20**). While the functional significance of the increase in several oropharynx species is not known, it implies that there may be less separation between the upper and lower respiratory tract during infection, and, when coupled with lung inflammation, may drive differential host responses in infection. Going forward, we will finalize our full data set of 26 animals and rerun all analyses to validate and finalize our observations and conclusions. Ultimately, we aim to relate the microbiome changes to pulmonary inflammation using our serial FDG PET/CT imaging[39,109]. We have quantified the total lung inflammation in the 26 animals (**Figure 19**) and observed a wide range of inflammation with several macaques exhibiting an increase in inflammation at 4 months, others had no changes in inflammation at all; this heterogeneity matches the range of infection outcomes that we see in this animal model[49,125] and in humans. Interestingly, the total PET hot increases at 4 months coincide with the increased

microbial diversity observed above (**Figure 17**). This relationship will be further examined on a lobe-by-lobe basis for each macaque. We will also look for evidence suggesting that the degree of inflammation influences the duration and magnitude of lung microbiome alteration. Finally, we will generate a microbial interaction network to visualize how the microbes influence each other over the course of *M. tuberculosis* infection. A better appreciation of the interactome between host immunity, lung microbiota, and pathogen is important as it presents a fresh domain to study that likely contributes to host heterogeneity and a new potential for therapy. It is intriguing to speculate that certain microbial species could be utilized in tandem with chemotherapy or vaccination to minimize inflammation or enhance/polarize specific immune cells for better bacterial killing.

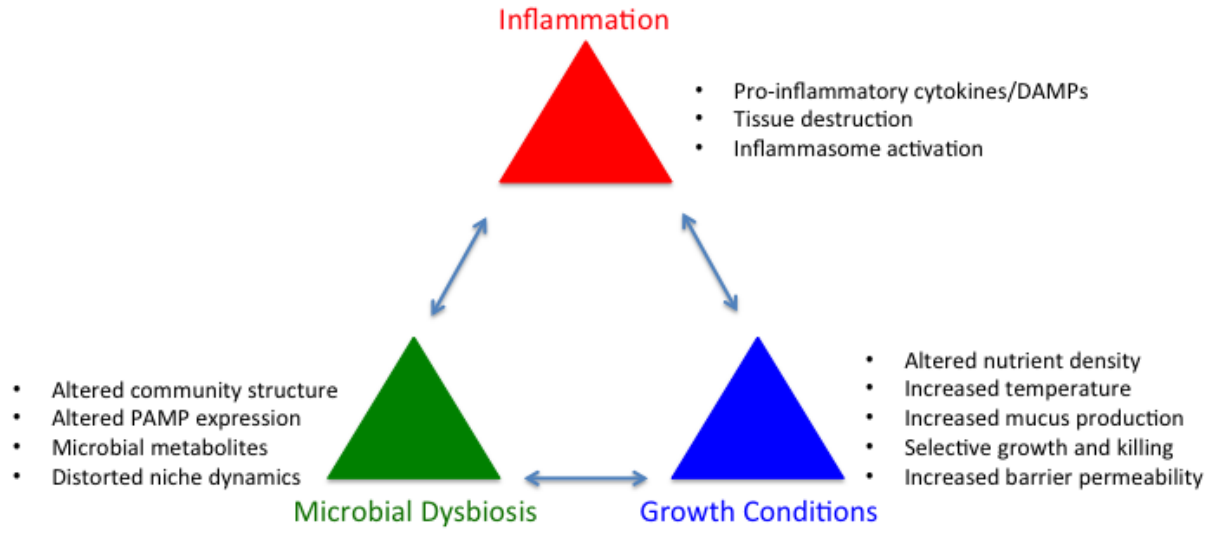


Figure 20. Cycle of lung microbiome dysbiosis and airway inflammation.

A pulmonary insult initiates localized inflammation that modifies airway function and architecture leading to altered growth conditions that in turn promotes microbe dysbiosis through various selective pressures.

5.5 ACKNOWLEDGEMENTS

We thank the members, laboratory technicians and veterinary staff of the Flynn and Ghedin laboratories for their technical expertise and assistance with animal care, sample processing, bioninformatics analysis, and study design. This work was made possible by support from the National Institutes of Health, R03 AI122067 (JLF, PLL & EG) and T32 AI089443 (AMC).

6.0 IMPLICATIONS OF DISSERTATION: SIGNIFICANCE, CONCLUSIONS AND FUTURE STUDIES

In this dissertation, we have coupled our well characterized macaque model of tuberculosis, which appropriately captures and recapitulates the full spectrum of human infection and disease[42,49,109,126,189], with novel molecular, imaging, and bioinformatics tools to address several important questions in early *M. tuberculosis* infection. Early events in this bacterial infection are critical junctures during which key components of the immune system are poised to either interact favorably to promote containment or interact poorly and allow for early dissemination. These initial interactions influence host outcome[27,39] and likely contribute to the overall heterogeneity observed in human infection. In this body of work, we have specifically tackled questions involving dissemination, reinfection, and the lung microbiome. Each of these three domains are uniquely tied to early infection and their exploration has provided new insight into how to exploit the earliest exchange between bacterium and host in favor of man.

6.1 SIGNIFICANCE

In 2015, TB overtook AIDS as the leading infectious disease killer worldwide with an estimated 1.4 million deaths, and a further 0.4 million deaths among human immunodeficiency virus-positive (HIV+) individuals[1]. Overall incidence for the same year was an estimated 10.4 million

cases, of which there was approximately 580,000 cases of drug-resistant tuberculosis[1]. These numbers persist in spite of widespread infant vaccination with *M. bovis* Bacille-Calmette et Guerin (BCG) in endemic areas and inexpensive and effective antibiotic treatment[2,190,191]. This is, in part, due to poor patient compliance and poor access to health care, which limits the efficacy of chemotherapy despite an estimated 90-95% cure rate in TB treatment control programs in drug-sensitive cases[2,158]. Moreover, the BCG vaccine is reputed to protect against the worst forms of disseminated TB in children but has a highly variable rate of protection in adults likely stemming from multiple factors including differences in geographic location, route, substrain, and the local, lung environment[191]; newer, more reliable and effective vaccines are needed.

6.2 CHAPTER CONCLUSIONS AND FUTURE STUDIES

6.2.1 Barcoded *M. tuberculosis* reveals a bottleneck of secondary dissemination that is predicted by early granuloma size.

To probe dissemination and track early bacterial dynamics, we generated a novel *M. tuberculosis* barcoding system with an associated Python pipeline that has given us unparalleled ability to track the fate of individual bacilli in infected macaques. Moreover, when paired with our serial [^{18}F]-FDG PET/CT imaging, we were able to reconstruct infection maps in which we could segregate disseminated or contained lesions and begin to examine individual lesions for features relating to bacterial spread. Interestingly, we found that only a subset of granulomas disseminated to form new, productive (CFU^{+}) granulomas, and that granuloma size early

(between 4-5 weeks post infection) correlated with dissemination. When we looked at two additional early infection monkeys, we confirmed the early feature of size and found that those larger lesions had higher CFU relative to their contained counterparts. We speculate that there is a critical crossroads for granuloma fate in early infection in which early cellular responses, bacterial growth, and granuloma size interact to dictate further dissemination or containment (**Figure 21**). Deciphering the key players that direct either endpoint will be an interesting question for future study, particularly as these decisions likely rely on components of the innate response given its timing. Notably, the two early macaques had significant dissemination to the thoracic lymph nodes and only minimal early lung spread. Whether this is indicative of a more general phenotype of early spread or an isolated instance found in these two animals is also unknown and warrants further study. Ultimately, this new barcoding scheme has not only provided novel depth into disease progression and individual granuloma fate mapping, it has also reaffirmed an unappreciated role for early granuloma dynamics in shaping host outcome.

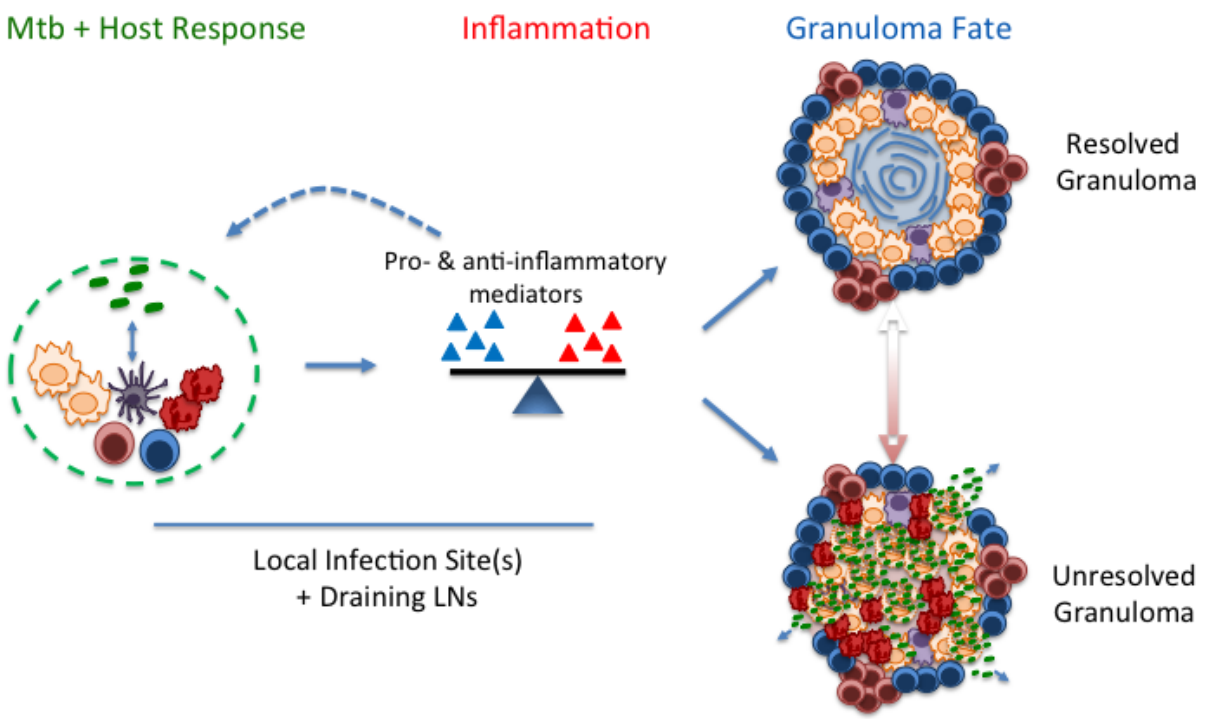


Figure 21. Granuloma fate is influenced by a complex and dynamic exchange of host and bacterial features.

6.2.2 Reinfection in TB: primary infection protects against early secondary infection.

How an ongoing *M. tuberculosis* infection influences a subsequent, secondary infection is entirely unexplored. Here we leveraged the macaque model of TB to experimentally assess the effect of reinfection on the course of early secondary infection. Utilizing ^{18}[F]-FDG PET/CT imaging and separate library qTags, we differentiated primary and secondary lesions in a matched set of macaque experiments where 4 week primary granulomas were directly compared against 4 week secondary granulomas. Excitingly, our data suggested that previous infection is protective against reinfection in that it completely protected several animals from developing any new secondary granulomas and, of the granulomas that did form, a large majority (>50%) were sterile (**Figure 12 & Figure 13**). Those with detectable CFU were all significantly diminished

versus primary granulomas of the same age (~4 weeks post infection) demonstrating a uniformly strong anti-tuberculous response. Importantly, the level of protection exhibited here is greater than that of BCG vaccination[191] and any vaccine candidate tested in our macaques to date. To provide better perspective into the bacterial killing seen in reinfection, we compared granuloma CFU data from early (5-6 weeks post infection) vaccinated animals that received BCG, an H56 boost, and a CAF01 adjuvant prior to *M. tuberculosis* infection (**Figure 22**) [data courtesy Robert DiFazio, unpublished]. The bacterial burden in the secondary granulomas in reinfection was considerably lower than the primary granulomas from the vaccinated animals further highlighting the magnitude of protection observed in our reinfected animals. (Infectious dose was similar for both studies.) As a result, the immunologic features engendering the protection observed in our reinfection studies offer unique potential to identify novel immunological biomarkers for better vaccine development. Our initial investigations into some of the protective immune responses suggested that the $CD3^+$ T cells from secondary granulomas at 4 weeks post infection were poised to produce greater amounts of both pro- and anti-inflammatory cytokines than the primary granulomas (**Figure 14**). We believe this rapid, balanced cytokine milieu contributed to the substantial killing by appropriately pairing macrophage activation and bactericidal responses with immune regulatory programs and tissue preserving responses[7,30,133] (**modeled in Figure 23**). T_H1 responses are likely necessary for appropriate activation of macrophages to kill internalized *M. tuberculosis*[32], and most vaccines entering clinical trials have relied on induction of IFN-γ^+ T cell subsets as a measure of vaccine success[6]. However, our reinfection data above suggested that multiple cytokines are necessary to achieve protection, at least in the context of an early secondary infection. Intriguingly, IFN-γ was the only cytokine not found to be upregulated when we compared primary and secondary granulomas. It should be

noted that this data does not discount the importance of IFN-γ but rather suggests that a broader landscape of cytokines is needed to initiate a protective response in TB. Future studies in reinfection are needed to fully understand the immunological basis providing this protection, particularly in light of a renewed focus on resident memory T cells[192]. In addition, the durability of this observed protection as well as the apparent needed for viable *M. tuberculosis* will have to be thoroughly tested.

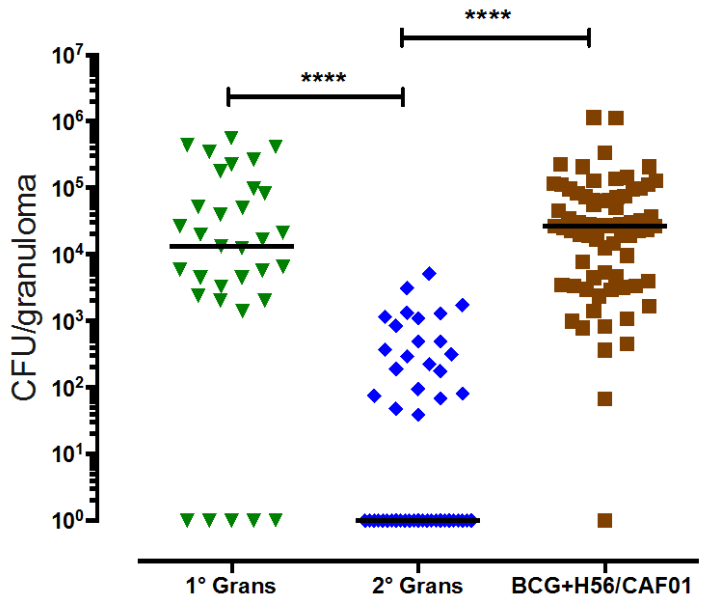


Figure 22. Putting in perspective: bacterial killing in reinfection vs. vaccination.

(Adapted from **Figure 13**). CFU of 4 week primary (n=33), 4 week secondary(n=47), and vaccinated, 5-6 week primary (n=71) granulomas (****p<0.0001), each symbol is a granuloma. (Early BCG+H56/CAF01 granuloma data curtesy Robert DiFazio, unpublished).

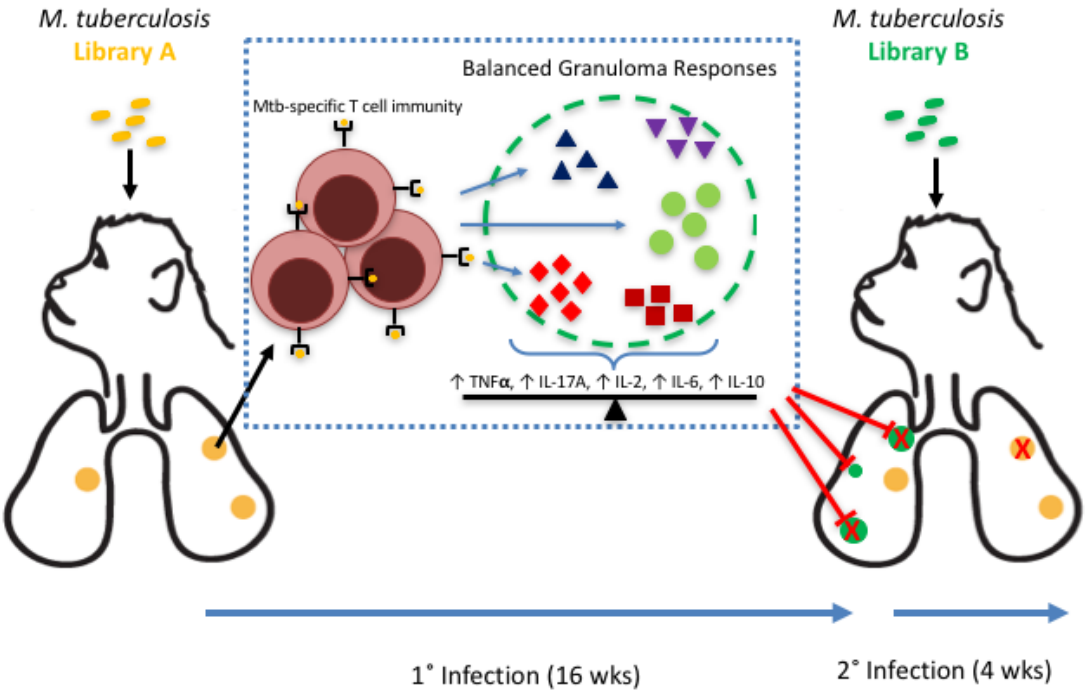


Figure 23. A balanced cytokine milieu elicits anti-tuberculous immunity and immune regulation.

Model of immunity in reinfection. A primary infection with low dose *M. tuberculosis* Erdman Library A initiates and maintains adaptive immunity including antigen-specific T cells that migrate to granulomas and help promote bacterial control and clearance. While several primary granulomas are sterilized and resolved over 16 weeks, this local T cell immunity is rapidly poised to respond to a secondary infection (Library B). The early production of pro- and anti-inflammatory cytokines helps prevent establishment of secondary granulomas and greatly enhances bacterial killing whilst diminishing excessive lung inflammation and immunopathology.

6.2.3 *M. tuberculosis* alters the lung microbiome.

The final arm of this dissertation was an initial survey into microbial environment of the lung and its relationship with *M. tuberculosis* infection. The lung microbiome is arguably among the

earliest interactions between host and pathogen, and the role that the microbiota plays in influencing host immunity and host response at mucosal sites is becoming increasingly apparent[129,162,193,194]. By serially sampling the BAL of both infected and uninfected lobes in macaques over the course of *M. tuberculosis* infection, we have provided the first evidence of a dynamic interplay between bacterium and host microbiota in the lung. Specifically, we see an increase in microbial diversity in the 4th month post infection in the infected lobe relative to the opposite, uninfected lobe that is most likely driven by species of the oropharynx[188], as well as a concomitant increase in total lung inflammation at the same time point in several macaques (**Figure 17, Figure 19** & **Table 4**). While our data analysis is still preliminary, this shift is intriguing as it suggests that the lower airway may undergo increased cross-contamination with the upper respiratory tract following infection and inflammation. The significance of this shift is still unclear and requires further study. In addition to finishing analysis on the remaining 16 macaques, we are working to associate this lung microbiome alteration with pulmonary inflammation in a more detailed, lobe-specific basis. Such analysis would provide a functional relationship to the observed interactions by tying inflammation driven by infection with the observed dysbiosis in the lung flora (**modeled in Figure 24**). Understanding this cycle of inflammation and microbial dysbiosis[129] in TB is very exciting as it posits an additional component that likely influences the variable responses to host outcome. Moreover, if precise microbe-host-pathogen relationships are identified, it presents an entirely new avenue for therapeutic intervention in which particular lung microbes could be used to engage specific immune programs and minimize overall inflammation in the hopes of steering more favorable outcomes to *M. tuberculosis*.

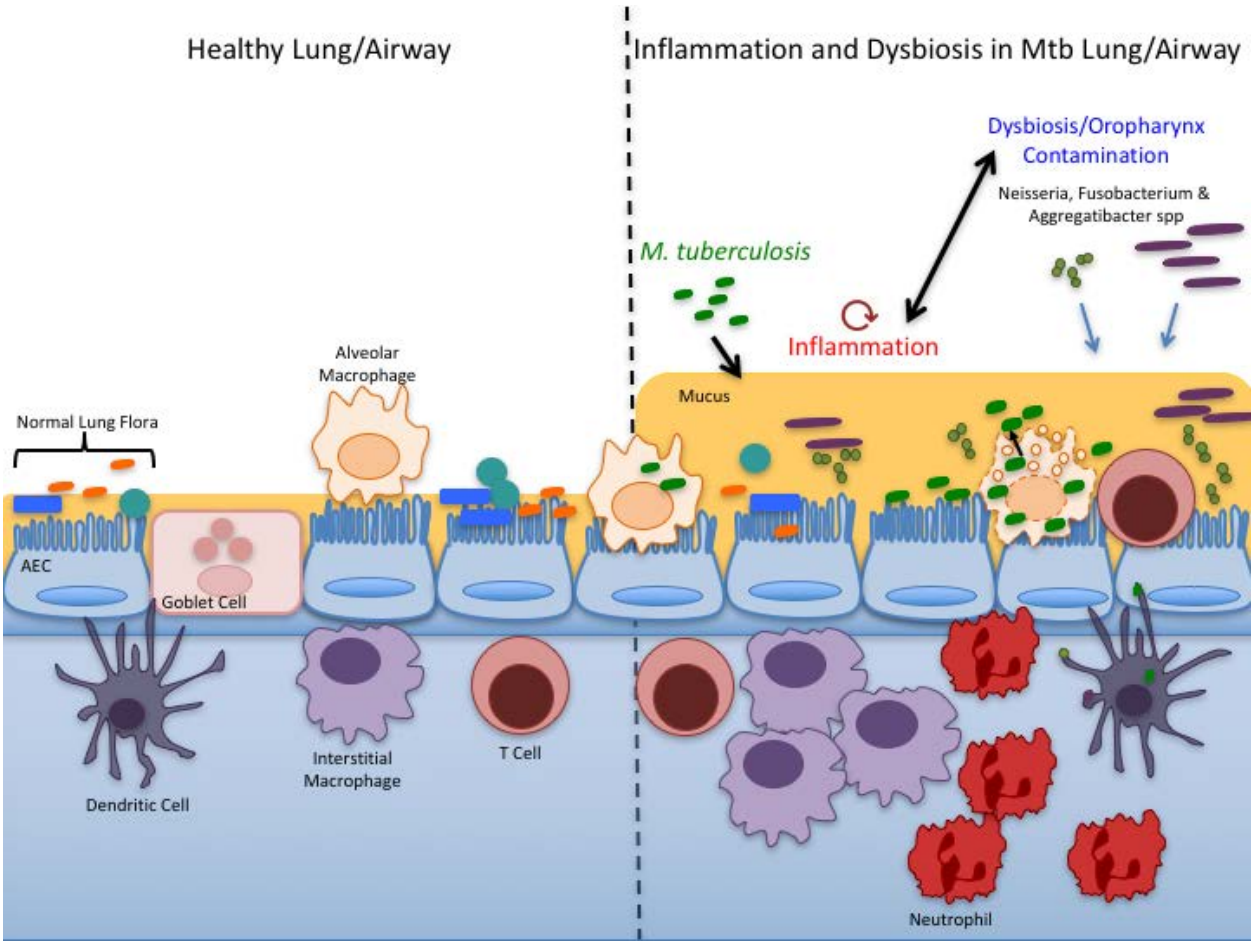


Figure 24. Model of *M. tuberculosis* alteration of the lung microbiome.

Left panel: normal lung at steady state. There is a dynamic interplay between the airway and local microbiota that maintains an immunological equilibrium. Right panel: inflammation and dysbiosis in *M. tuberculosis* infected lung. Following successful infection, there is perturbation of the airway concomitant with pulmonary inflammation and immune cell recruitment. These changes alter the local environment allowing for increased colonization with *Neisseria, Fusobacterium,* & *Aggregatibacter* spp. from the upper respiratory tract and further dysbiosis and inflammation.

6.3 FINAL THOUGHTS

This dissertation probed early events in three separate and unique contexts, early dissemination and disease progression, reinfection, and the lung microbiome. To address the first two aims, we developed and validated a novel tool for use in our macaque model of TB that permitted tracking of individual bacilli to determine bacterial fate and granuloma dynamics. We coupled this information with real-time serial [^{18}F]-FGD PET/CT imaging and recreated maps of infection that gave us new insight into events that contribute to granuloma outcome. We observed that early features, including granuloma size and bacterial burden, were associated with bacterial dissemination. These two features in tandem with early immune responses are likely key components of the initial interactions between host and pathogen that influence host outcome. Our new barcoded *M. tuberculosis* also allowed us to examine how primary infection impacts a secondary infection by utilizing library specific qTags that differentiate parental strain. Although we are awaiting final confirmation of library identity from our sequencing collaborators, our imaging data suggested that we revealed an exciting phenotype of protection in which we observed complete protection against new granulomas in several macaques and considerable anti-mycobacterial killing in the first weeks in all of the remaining animals. Importantly, this protection occurred across the spectrum of host outcomes in primary infection in our macaques implicating that significant disease burden in the first infection is not required. This is critical as the majority of humans (~90%) dwell within the latent end of the TB spectrum. Confirming the apparent requirement for viable *M. tuberculosis* is equally critical as this would be among the first design steps in generating a vaccine (a macaque reinfection study with sterilizing drug treatment has just started). While on the surface the administration of live *M. tuberculosis* might be unpalatable as a vaccine strategy, there are ideas of kill-switch and other

severely attenuated strains in the pipeline that offer encouragement for such an approach. Finally, and perhaps most readily, this model of macaque reinfection can be used as a robust platform to fully dissect the precise immune mechanisms eliciting protection, which can be directly translated for use and testing in human TB vaccine settings. The final segment of this thesis set the stage for an interesting set of new studies that will interrogate the interface of the local, lung microbiota and *M. tuberculosis*. Our first stab has yielded though-provoking suggestions on how a trifecta of microbe-host-pathogen interactions may influence infection outcome, particularly as a result of inflammation and dysbiosis. The influence of the microbiome on infectious disease and immunity is a very active area of research and by applying this fresh perspective to our model of macaque TB, we may uncover new understanding of host variability in TB and potentially open new opportunities for early treatment and intervention.

7.0 PUBLICATION RECORD

Martin, C.J.*, **A.M. Cadena***, V.W. Leung, P.L. Lin, P. Maiello, N. Hicks, M.R. Chase, J.L. Flynn, and S.M. Fortune. 2017. Digitally Barcoding *Mycobacterium tuberculosis* Reveals *in vivo* Infection Dynamics in the Macaque Model of Tuberculosis. *mBio* 8:e00312-17. *Co-first authors.

Cadena, A.M., J.L. Flynn, and S.M. Fortune. 2016. The Importance of First Impressions: Early Events in *Mycobacterium tuberculosis* Infection Influence Outcome. *MBio* 7:e00342-00316.

Cadena, A.M., E.C. Klein, A.G. White, J.A. Tomko, C.L. Chedrick, D.S. Reed, L.E. Via, P.L. Lin, and J.L. Flynn. 2016. Very Low Doses of *Mycobacterium tuberculosis* Yield Diverse Host Outcomes in Common Marmosets (*Callithrix jacchus*). *Comp Med* 66:412-419.

Lin, P.L., P. Maiello, H.P. Gideon, M.T. Coleman, **A.M. Cadena**, M.A. Rodgers, R. Gregg, M. O'Malley, J. Tomko, D. Fillmore, L.J. Frye, T. Rutledge, R.M. DiFazio, C. Janssen, E. Klein, P.L. Anderson, S.M. Fortune, and J.L. Flynn. 2016. PET CT Identifies Reactivation Risk in Cynomolgus Macaques with Latent *M. tuberculosis*. *PLoS Pathogens* 12:e1005739.

Cadena, A.M., C.J. Martin, V.W. Leung, P.L Lin, P. Maiello, N. Hicks, M.R. Chase, S.M. Fortune & J.L. Flynn. Primary infection with *Mycobacterium tuberculosis* Confers Protection to Secondary Infection in the Macaque Model of Tuberculosis. Manuscript in preparation.

Cadena, **A.M.**, S.M. Fortune, and J.L. Flynn. 2017. Heterogeneity in Tuberculosis: the Importance of Considering Local Conflicts in a Global Context. Review in submission to *Nature Reviews Immunology*.

8.0 APPENDIX A: VERY LOW DOSES OF *MYCOBACTERIUM TUBERCULOSIS* YIELD DIVERSE HOST OUTCOMES IN COMMON MARMOSETS (*CALLITHRIX JACCHUS*)

This appendix chapter is adapted from the original publication and represents a side project ancillary to my primary thesis in early events:

Cadena AM, *et al.* (2016) Very Low Doses of Mycobacterium tuberculosis Yield Diverse Host Outcomes in Common Marmosets (Callithrix jacchus). *Comp Med* 66(5):412-419.

8.1 INTRODUCTION

Tuberculosis remains a risk to global public health, with 9.6 million new cases of active tuberculosis and 1.5 million deaths in 2014[195]. Although most humans contain Mycobacterium tuberculosis in a clinically asymptomatic infection termed 'latent tuberculosis,' a smaller subset (approximately 10%) of patients present initially with primary active disease or subsequent disease reactivation over the course of their lifetime[7]. The biologic basis for the disease spectrum of *M. tuberculosis* in humans is unknown and continues to be an active area of research.

Several animal models have been adapted for experimental M. tuberculosis infection, greatly contributing to our understanding of tuberculosis biology[196]. NHP models of tuberculosis, particularly cynomolgus macaques (*Macaca fascicularis*)[49,125] and rhesus macaques (*Macaca*

mulatta)[197,198] are recognized as the most faithful animal models in replicating the human spectrum of disease in terms of both pathology and infection outcome[15,30,125,199]. In addition, the recently developed common marmoset model of tuberculosis (*Callithrix jacchus*)[200] replicates crucial facets of human tuberculosis, including cavitary disease. This model demonstrated divergent rates of disease progression after challenge with 3 strains—a Beijing isolate, the less virulent strain CDC1551, and *M. africanum*—and 2 different inocula (250 and 25 CFU). Notably, each infection resulted in rapid pulmonary disease presentation and weight loss, and all animals died by 75 days after challenge, thus reflecting the inherent susceptibility of common marmosets to this infection.

The influence of *M. tuberculosis* dose and strain on host outcome has been examined in several animal models. For example, BALB/c mice demonstrated a broad range of virulence, bacterial load, and pathology after infection with 19 different *M. tuberculosis* complex strains of 11 major genotype families[113]. These 19 strains were segregated according to their virulence, bacterial burden, pathology, and delayed-type hypersensitivity responses as high, intermediate, and low responders. Strains that were classified as high responders induced the most pathology, greatest bacterial burden and highest mortality and included the isolates Beijing 2 and 3, Africa 2, and Somalia 2. At the other end of the spectrum, those strains that elicited the least severe pathologic scores caused no or minimal mortality after 112 days, yielded the lowest bacterial load, and included the isolates H37Rv, Canetti, and Beijing 1[113]. A recent review examined the virulence and immunogenicity of several genotypic lineages of *M. tuberculosis* in mice and related these findings to observed human epidemiologic data[201]. Overall, virulence and immune response vary extensively within each genotype, with many genotypes having both high- and low-virulence variants. Furthermore, rabbits (like mice) demonstrate an influence of dose, strain,

and growth phase on tubercle formation and virulence, and *M. tuberculosis* strain Erdman had the greatest virulence in rabbits, requiring the fewest inhaled bacilli to generate visible tubercles at 5 weeks[202].

Whether common marmosets have any capacity to control *M. tuberculosis* infection or are universally susceptible to developing active tuberculosis, even with low-dose inocula, has not been determined to date. In this study, we aimed to better characterize the disease phenotypes and host outcomes of common marmosets after *M. tuberculosis* infection by challenging animals with very low doses (less than 15 CFU) of the Erdman and CDC1551 strains.

8.2 MATERIALS AND METHODS

8.2.1 Animals.

Common marmosets (*Callithrix jacchus*) were obtained from the Wisconsin National Primate Research Center (Madison, WI). Prior to shipment, the animals were tested for *Campylobacter*, *Shigella*, *Salmonella*, *Giardia*, *Cryptosporidium*, *Entamoeba*, and *Balantidium*. On arrival at the University of Pittsburgh, the marmosets were screened for *M. tuberculosis* infection and other comorbidities during a month-long quarantine. More specifically, each macaque underwent complete baseline blood and biochemical analyses, 2 tape tests to screen for pinworms, and *Giardia* ELISA with 2 pooled fecal sets. All marmosets were housed and maintained according to the practices and standards detailed in the Animal Welfare Act and the *Guide for the Care and Use of Laboratory Animals* within a biosafety level 3 facility. These NHP were housed in a 4 × 4 macaque cage system, with the internal dividers removed to allow for open access and

opportunities for climbing to the top and bottom of the cage. A nest box was included in the upper part of the cage, and the animals were fed a specialized marmoset diet (as advised by Dr Saverio Capuano, Wisconsin Primate Center). In some cases, infected animals were cohoused with naïve cagemates for a separate transmission study. The IACUC at the University of Pittsburgh approved all protocols and experiments.

8.2.2 Infection and necropsy.

Common marmosets were infected with either the virulent Erdman strain of *M. tuberculosis* or CDC1551 at doses of 1 to 12 CFU, as determined by plating the inoculum. One animal in the Erdman cohort was infected by aerosol, and the remaining marmosets were infected through bronchoscopic instillation, as previously described[125]. In brief, a disinfected bronchoscope was placed into the desired bronchus and 0.2 mL of sterile saline containing the appropriate infection inoculum was instilled. For the marmoset infected by aerosol, an infection inoculum was administered through a 10-min exposure to aerosolized bacilli created by a 3-jet Collison nebulizer (BGI, Waltham, MA) controlled by the AeroMP bioaerosol exposure system (Biaera Technologies, Hagerstown, MD) in a head-only exposure chamber (CH Technologies, Westwood, NJ) within a class III biologic safety cabinet (Baker, Sanford, ME). After infection, all marmosets were clinically monitored for signs of infection according to previously published methods[49,125], and ^{18}F-fluorodeoxyglucose positron-emission tomography–CT (PET–CT) scans were performed monthly, as previously described[39,109,200]. Marmosets were euthanized when they lost more than 20% of their preinfection body weight or when they exhibited substantial clinical signs of disease (for example, lethargy and anorexia). At necropsy, a gross pathology score for each marmoset was determined according to our published scoring system[49], which evaluates

tuberculosis-specific disease in all lung lobes, thoracic lymph nodes, and extrapulmonary sites. Overall bacterial burden scores and CFU counts for individual granulomas, regions of complex pathology (for example, tuberculosis pneumonia), and thoracic lymph nodes were quantified and calculated as previously described[36,49]. Briefly, multiple individual tissue samples in each marmoset were identified by using PET–CT imaging as a guide, excised, and plated at necropsy, and bacterial colonies were counted 21 d after necropsy. The CFU score for each animal was determined by summation of each tissue's log-transformation of bacterial burden, thus providing a relative comparator for all marmosets. In comparison, the total CFU count was the sum of the actual number of bacilli in all samples taken from the animal, and thus provided a more accurate representation of the bacterial burden.

8.3 RESULTS

8.3.1 Host outcome after *M. tuberculosis* infection in marmosets

To determine the clinical progression and disease pathology of very low-dose infection, 9 marmosets were challenged with 2 different strains of *M. tuberculosis* at doses of 1 to 12 CFU (**Table 5**). Inocula were determined by plating at the time of challenge. All but one of the marmosets in this study were infected by direct instillation of bacteria into the airways by using a bronchoscope. The marmoset that received 12 CFU of the Erdman strain (animal no. 3012) was infected through aerosol dispersion. All marmosets infected with the Erdman strain (dose range, 1 to 12 CFU, with most animals infected with 2 CFU) presented with rapidly progressing disease regardless of dose. Clinically, all of these marmosets exhibited physical signs of illness,

including inactivity, reclusiveness, and diminished appetite. The animals exhibited precipitous weight loss after infection and arrived at the weight-loss threshold of 20% by 45 days after infection (median, 39 days), requiring euthanasia (**Figure 25**). Coughing was observed in a subset of animals (marmosets 13210, 13310, 13410, 13610, 13710, and 3412). At necropsy, all animals infected with strain Erdman had gross pathology scores that exceeded 30 (median score, 41; **Table 5**).

Table 5. Survival, disease presentation, and bacterial burden and dissemination after infection of common marmosets with very low doses of *M. tuberculosis* infection.

Animal ID	Strain	Infection Dose (CFU[a])	Survival (days)	Gross Pathology Score	CFU Score	% Tissue Positive
13210	Mtb Erdman	2	39	49	142	100
13310		2	35	56	164	100
13410		2	37	41	158	100
13610		2	38	31	123	100
13710		2	40	33	137	95.2
3012		12	32	53	144	100
3412		1	43	33	107	100
3713	CDC1551	1	307	25	26	32
3913		7	89	66	53	100

[a]CFU = Colony forming units

By contrast, the 2 marmosets challenged with strain CDC1551 (dose, 1 or 7 CFU) had divergent outcomes associated with dose but demonstrated reduced overall disease compared with that of marmosets infected with strain Erdman. The marmoset that received 7 CFU of the

CDC1551 strain (animal no. 3913) had mild to moderate disease progression and steady weight loss and survived to 89 days (**Table 5** and **Figure 26**). Surprisingly, the marmoset that received 1 cfu of CDC1551 (animal no. 3713) had mild disease progression, with a sustained period of weight loss in this animal beginning on day 76 that was not accompanied by any other

clinical signs of disease. By day 174, this marmoset was regaining weight and returned to its initial weight by day 282 (**Figure 25**). This marmoset was euthanized after the conclusion of the study on day 307; euthanasia was not due to weight loss or clinical morbidity.

The marmoset inoculated with 7 CFU had a gross pathology score of 66, whereas the marmoset inoculated with 1 CFU had a score of 25 (**Table 5**).

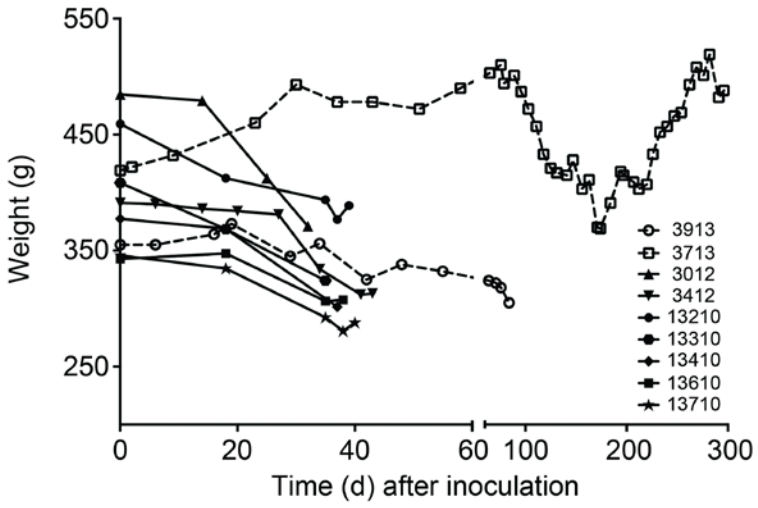


Figure 25. Weight loss after infection with very low doses of *M. tuberculosis* varied according to strain and dose.

All animals except one had sustained weight loss that prompted euthanasia by 90 days. The marmoset infected with 1 CFU of CDC1551 (animal no. 3713) had a period of weight loss by recovered its preinfection body weight by day 282. This animal was euthanized at 307 days after infection due to the conclusion of the study.

8.3.2 In vivo PET–CT imaging.

To compare in vivo disease progression between strains and doses, all marmosets were followed by using serial PET–CT imaging[39,109]; we here present 2 representative animals from each cohort

for comparison. For strain Erdman, 2 marmosets (nos. 3012 and 3412) were imaged and compared with 2 animals infected with CDC1551 (nos. 3913 and 3713). Both sets of animals represented the highest and lowest dose administered for each strain (**Table 5**). In agreement with clinical and weight observations, these marmosets had differing rates of disease progression, as demonstrated by ^{18}F-fluorodeoxyglucose uptake. Relative to their baseline scans, both Erdman-infected marmosets had activity indicative of pulmonary and lymphatic disease on their 4 week postinfection PET–CT scans (**Figure 26**, yellow arrows), whereas the 2 animals challenged with CDC1551 had normal scans without disease, thus reflecting the slower progression associated with this strain. The prenecropsy scan for each animal was obtained just a few days prior to euthanasia. Given the rapid loss of body weight that necessitated early euthanasia in the 2 Erdman-infected marmosets, the 4 week scans served as their prenecropsy scans. By contrast, the disease in the CDC1551-infected marmosets progressed more slowly (prenecropsy scans at weeks 12 and 44). In particular, marmoset 3913 had significant disease in the right lower lung lobe and thoracic lymph nodes (**Figure 26**, yellow arrows), and marmoset 3713 (dose, 1 CFU of CDC1551) had very little ^{18}F-fluorodeoxyglucose uptake during his prenecropsy scan at 300 days after infection.

8.3.3 Overall bacterial burden and extent of dissemination.

Comparisons of bacterial burden (CFU scores)[49] and the number of bacteria per tissue sample (granuloma, lymph node, tuberculosis pneumonia)[36] revealed strain- and dose-dependent differences dependent (Table 1 and Figure 3).

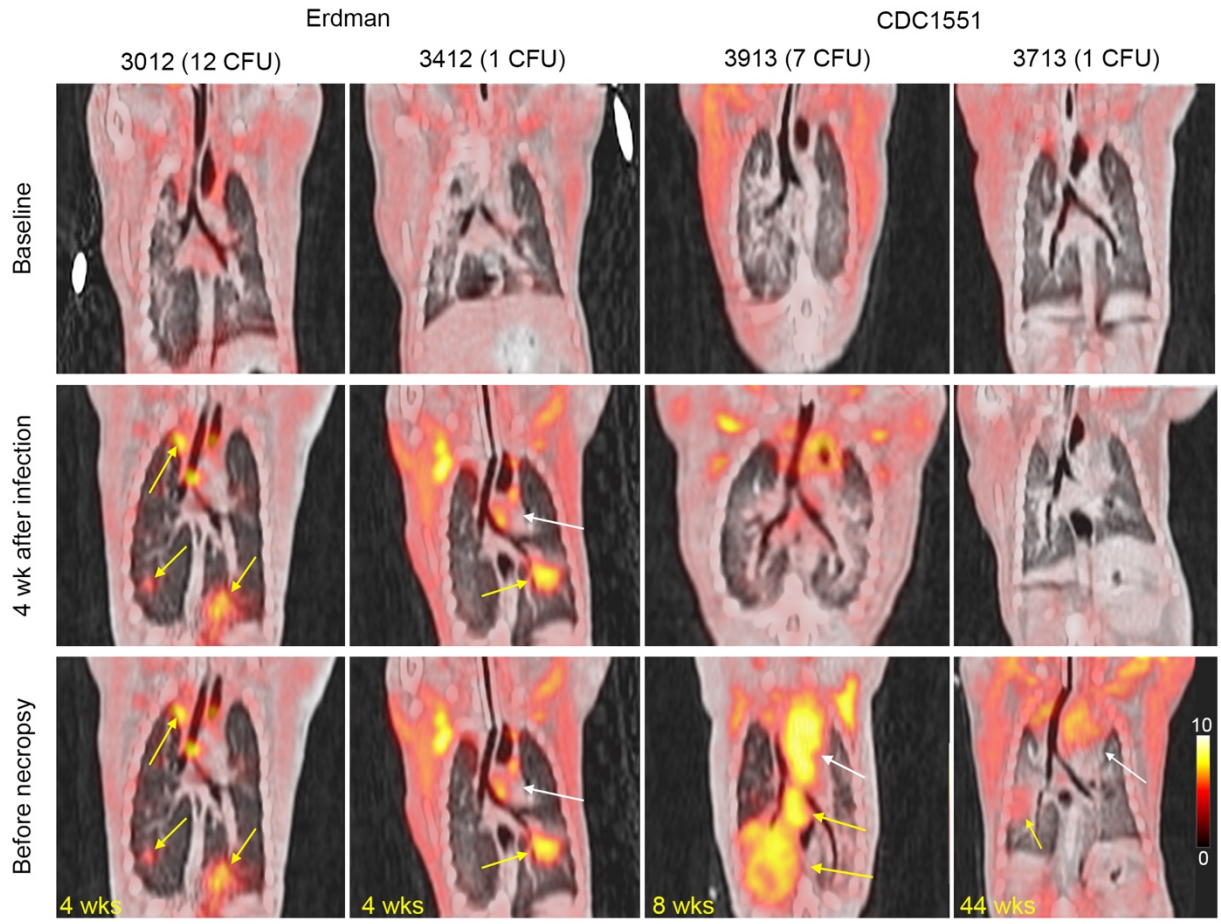


Figure 26. Serial PET/CT imaging of infected marmosets revealed variable disease progression that was dependent on dose.

These images are representative PET/CT coronal cross-sections from marmosets infected with very low doses of *M. tuberculosis;* 18F-flurodeoxyglucose was used as a probe for inflammation. Left panel: 2 marmosets infected with the Erdman strain (doses, 12 and 1 CFU) were followed over the course of infection. For both animals, their 4 week scans served as their prenecropsy scans, in light of their rapid disease progression and clinical decline. Right panel: 2 marmosets infected with CDC1551 (doses, 7 and 1 CFU) were followed throughout infection. For all animals, the date of each prenecropsy scan (in weeks) is labeled at the lower left. Yellow arrows denote diseased areas of lung and thoracic lymph nodes, whereas the white arrow highlights noninflammatory uptake in the heart.

All of the Erdman-infected marmosets (including the one that received a dose of 1 CFU) had CFU scores above 100 (median, 142) and more than 95% of the tissues cultured were positive for *M. tuberculosis* (**Table 5**), indicative of prominent and widespread disease. By contrast, the 2 marmosets challenged with 7 and 1 CFU of CDC1551 had CFU scores of 53 and 26, respectively, supporting the finding that bacterial burden is related to dose for this strain (**Figure 27A**). The inoculation dose of 7 CFU of CDC1551 resulted in 100% of tissues that were positive for tubercular bacilli, whereas the infection due to the 1-cfu dose was contained to a much greater extent, with only 32% of tissues positive for bacteria (**Table 5**). Two marmosets from the Erdman group (no. 3012, inoculated with 12 CFU, and no. 3412, which received 1 CFU) were selected for more precise comparison of tissue burden with the 2 animals infected with CDC1551 (marmosets 3913 and 3713). Comparisons of each animal's cumulative bacterial burden (total CFU counts among all lung and lymph node samples) revealed that 3 of the marmosets had cumulative burdens that exceeded 5×10^6 CFU, with little difference between the lung and lymph-node compartments (**Figure 27B**). In contrast, the total bacterial load in the marmoset infected with 1 CFU of CDC1551 (no. 3713) was nearly 100-fold lower than that in the other 3 (**Figure 27B**), with the vast majority of the burden contributed by infected thoracic lymph nodes (**Figure 27C**). Assessing each tissue separately revealed that the median lung or lymph node bacterial burden was higher in the 2 animals that received the higher dose within each strain (**Figure 27C**). Lymph nodes had the greatest variability across the different tissues, and the highest burdens occurred in regions of lung tuberculosis pneumonia. Most notably, the tissue bacterial burden in the marmoset that received 1 CFU of CDC1551 ranged from primarily sterile sites to the moderately infected superior pretracheal lymph node, which had 6.4×10^4 CFU.

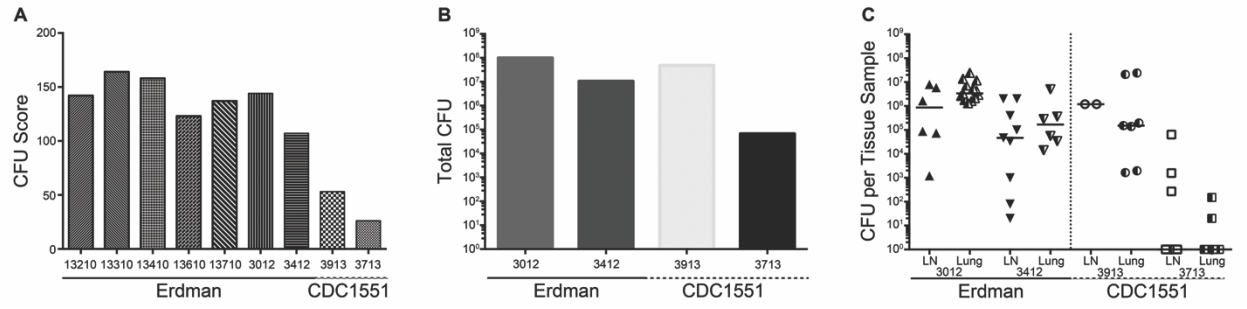


Figure 27. Very low dose infection with CDC1551 results in diminished bacterial burden.

(A) Comparisons of CFU scores for all 7 marmosets reveal the decreased scores for the 2 animals that received CDC1551. (B) Total CFU counts (the sum of all bacilli from every plated sample) and (C) the number of cfu per tissue sample from a subset of marmosets reveal the reduced bacterial loads in the 2 marmosets infected with CDC1551. In the marmoset infected with 1 CFU of CDC1551 (animal no. 3713), the total bacterial burden was approximately 2 logs lower than that in any other animal, and there were fewer bacilli in individual tissues, with a majority of samples being sterile.

8.3.4 Gross and histologic disease pathology after challenge with very low doses of *M. tuberculosis*.

Comparison of gross pathology at necropsy (**Figure 28**) revealed marked differences in the magnitude and extent of disease progression. For example, marmoset 3412, which was infected with 1 CFU of the Erdman strain, demonstrated pulmonary disease that was grossly limited to the left lower lung (**Figure 28A**), with concomitant enlargement and effacement of the left thoracic lymph nodes (**Figure 28B**). The left lower lobe showed extensive tuberculous pneumonia, with a central focal area of severe necrotizing consolidation (2 cm); the lobe completely lacked any normal, aerated lung parenchyma. Gross examination of the spleen and

liver revealed enlargement of both organs, with numerous (20 or more) granulomas (diameter, pinpoint to 1 mm or greater) throughout (**Figure 28C**).

By contrast, gross examination of tissues from the marmoset infected with 1 CFU of CDC1551 (no. 3713) revealed 2 relatively well-circumscribed lesions in the left (lesion diameter, 1.5 mm) and right (10 mm) lower lobes (**Figure 28D**); there were 4 additional discrete lesions (diameter, 2 mm or less) in the right lower lobe. The inferior and superior pretracheal lymph nodes were both enlarged, with significant effacement in the superior node (**Figure 28E**). The liver (**Figure 28F**) and spleen had no readily identifiable gross lesions. The liver was markedly enlarged, extending several centimeters below the costal region, and had an accentuated lobular appearance, with several large, irregular, superficial regions of parenchymal pallor. This discoloration was likely due to differential blood settling, euthanasia-solution–induced hepatocellular artifact, or glycogenation; there was no gross indication of tuberculosis throughout this organ.

Histopathology of lung, thoracic lymph nodes, spleen, and liver largely confirmed the findings at necropsy. For the marmoset challenged with 1 CFU of the Erdman strain (animal 3412), the region of tuberculosis pneumonia in the left lower lobe consisted of extremely neutrophil-rich inflammatory infiltrate, extending from alveolus to alveolus. Centrally, there was a large area of coagulative necrosis with near total effacement of all recognizable tissue architecture. There was minimal evidence of the formation of architecturally organized granuloma structures per se within the entirety of the large necrotizing region (**Figure 29A**). The left cranial hilar lymph node displayed extensive nodal effacement, with regions of necrotizing and nonnecrotizing inflammation with poor structural organization (**Figure 29B**). Examination

of the spleen (**Figure 29C**) and liver (**Figure 29D**) revealed multiple regions of caseous and nonnecrotizing granulomas, indicative of widespread disseminated disease.

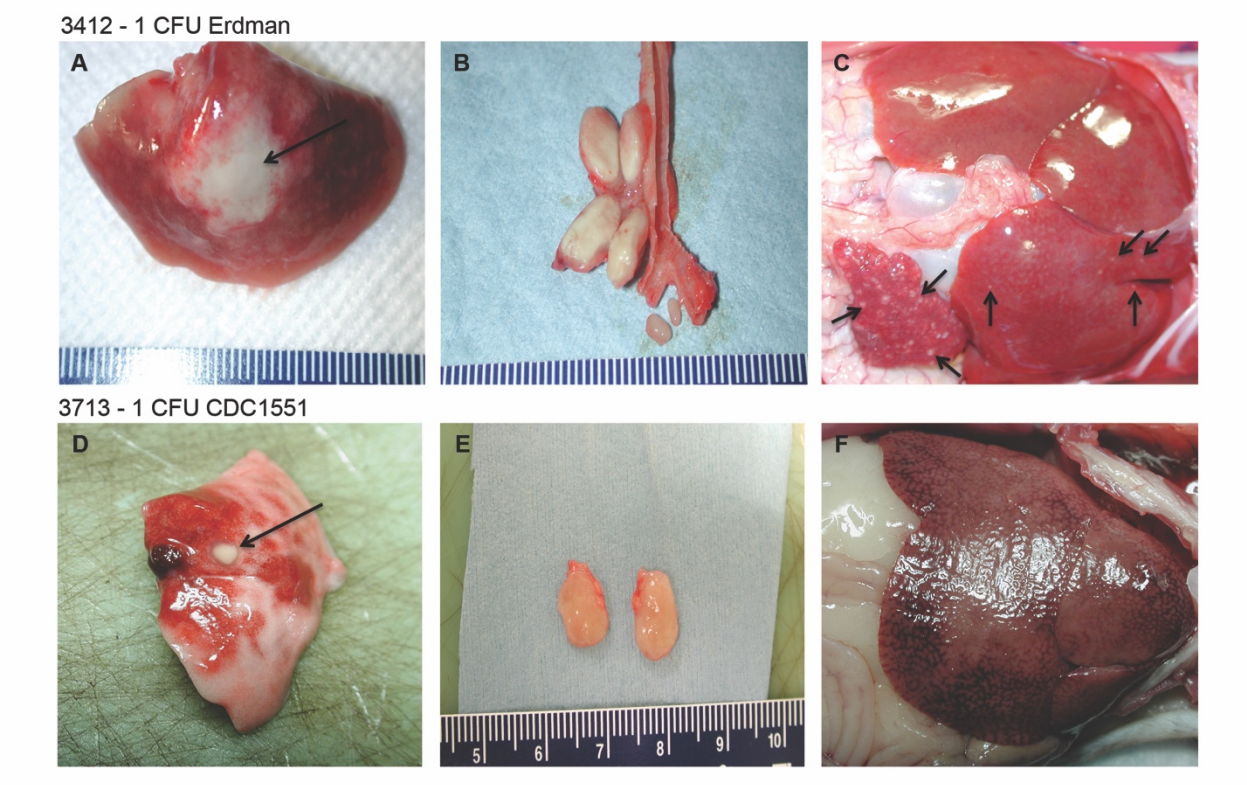


Figure 28. Gross pathology after infection with very low doses of CDC1551 was reduced relative to that after Erdman.

Representative gross pathology from 2 marmosets infected with 1 CFU of the Erdman strain (top row) and 1 CFU of CDC1551 (bottom row). These images show the major disease-associated differences between the 2 animals. (A) Left lower lobe of lung with a central consolidation. (B) Enlarged and effaced left cranial hilar and left mainstem bronchial nodes, with mildly swollen central carinal lymph nodes. (C) Enlarged spleen and liver, with many disseminated pinpoint granulomas. (D) Granuloma in right lower lobe of lung. (E) Enlarged superior pretrachael lymph nodes. (F) Liver without gross evidence of tuberculous disease. Black arrows highlight areas of tuberculous disease.

In the marmoset challenged with 1 CFU of CDC1551 (animal 3713), the left lower lung lobe contained a single, caseous granuloma that was well circumscribed and had a prominent lymphocytic cuff (**Figure 29D**). In contrast to lung, the superior pretracheal lymph node had prominent coalescing and effacing caseous granulomas, with evidence of early collagen fibril formation within the necrotic matrix (**Figure 29E**). Despite the lack of gross disease at necropsy, this marmoset had microscopic evidence of splenic tuberculosis, with multiple, small minimally necrotizing granulomas (**Figure 29F**). In addition, sections of liver exhibited several focal areas containing caseous granulomas and small, nonnecrotizing sinusoidal granulomas (**Figure 29F**). Overall, the gross and microscopy pathology mirrored the clinical differences in these 2 marmosets and reiterate the distinction between the 2 strains, even at very low doses. Similar extensive gross and microscopic disease was present in all marmosets infected with the Erdman strain as well as in the animal infected with 7 CFU of strain CDC1551. Importantly, despite the marked difference in the general histologic appearance between the 2 CDC1551-infected animals and the more immunologically contained response in the marmoset challenged with 1 CFU, this marmoset had numerous microscopic findings that suggested that the disease was entering a more widely disseminating stage. The presence of numerous small nonnecrotizing granulomas and epithelioid cell aggregates within thoracic lymph nodes as well as the microscopic granulomas in the hepatic and splenic parenchyma indicated the onset of more fulminant and progressive disease prior to necropsy.

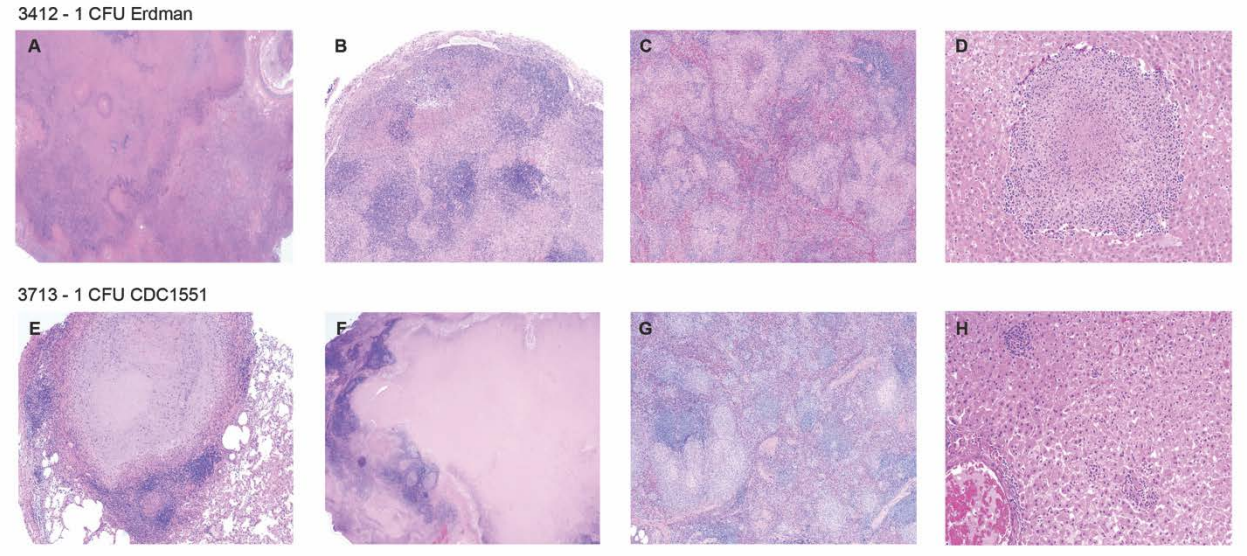


Figure 29. Histopathology after infection with very low doses of the Erdman and CDC1551 strains.

Representative histopathology from the 2 marmosets shown in Figure 4, which were infected with 1 CFU of the Erdman strain (top row) and 1 CFU of CDC1551 (bottom row). These images illustrate the differences between the 2 animals, particularly in the organization and structure of the lung granulomas. (A) Tuberculosis pneumonia with a large area of necrosis in the left lower lung lobe; magnification, 1.25×. (B) Left cranial hilar lymph node with areas of necrotizing and nonnecrotizing granulomatous inflammation; magnification, 4×. (C) Spleen with disseminating necrotizing and nonnecrotizing foci; magnification, 4×. (D) Liver with caseous granuloma; magnification, 10×. (E) Circumscribed caseous granuloma with a well-defined lymphocytic cuff in the left lower lung lobe; magnification, 4×. (F) Superior pretracheal lymph node with coalescing and effacing caseous granulomas; magnification, 4×. (G) Spleen with multiple areas of minimally necrotizing granulomas dispersed throughout the parenchyma; magnification, 4×. (H) Liver with several, small sinusoidal nonnecrotizing granulomas; magnification, 10×. Hematoxylin and eosin stain.

8.4 DISCUSSION

Here we report the diverse disease progression and host outcome patterns that occurred after the infection of common marmosets with very low doses (1 to 12 CFU) of 2 strains of *M. tuberculosis*. The susceptibility of marmosets to the Erdman strain was unrelated to dose or route. However, the inoculation dose may be an important determinant of outcome after challenge with the less virulent CDC1551 strain. All of the marmosets challenged with low doses of strain Erdman presented with fulminant, disseminating active tuberculosis, had a median survival time of 39 days, and had evidence of invasive necrotizing alveolitis and tuberculosis pneumonia, with little evidence of well-circumscribed granulomas. In contrast, infection with CDC1551 promoted an more slowly progressing infection overall; the marmoset that received 7 CFU survived to 89 days, and the animal given 1 CFU survived more than 300 days. To our knowledge, this is the first reported instance of apparent recovery, stabilization, and survival beyond 300 days of a marmoset infected with a strain of *M. tuberculosis*.

Our initial findings recapitulated those reported by our colleagues, who also described differential disease progression between several strains within the *M. tuberculosis* complex[200]. This previous study compared CDC1551, *M. africanum* N0091, and a Beijing K04 isolate at doses of 25 and 250 cfu; all 3 strains at both doses produced fulminant disease and prompted euthanasia by 80 d after infection due to weight loss[200]. The rapid disease progression, high bacterial burden, and invasive necrotizing pathology associated with the Beijing K04 isolate is very similar to the disease evolution, bacterial load, and disease presentation in our marmosets that were infected with the virulent Erdman strain. The median survival time in the previous study was 37 days[200], similar to the 39 days for Erdman-infected marmosets in the current study.

Moreover, the higher dose of the K04 isolate was the only case for which the rate of weight loss was significantly increased. Likewise, among the Erdman-infected marmosets in our study, the animal that received 12 cfu had the most rapid weight decline and was euthanized 32 days after infection. In contrast, infection with CDC1551 resulted in the slowest disease progression among the 3 strains evaluated previously and yielded a median survival time of 59 days[200]. Notably, half of the marmosets infected with the CDC1551 strain previously developed cavitary lesions, a particular manifestation of human tuberculosis that is associated with erosion of lung parenchyma into the airway, facilitating bacterial transmission[203]. In agreement with these findings, our marmoset that received 7 CFU of CDC1551 (no. 3913) demonstrated several sites of cavitation. A separate study performed by these collaborators[200,204] followed marmosets infected with various low aerosol doses of CDC1551 for weight loss in infected animals prior to drug treatment at 7 weeks. Little to no weight loss was seen in the animals exposed to 1 to 2 CFU, whereas infection with higher doses significantly increased weight loss percentage (data not shown) suggesting that the route of infection was not directly responsible for the difference observed in the marmoset infected with 1 CFU of CDC1551 presented here.

Extending the published findings, we present data supportive of long-term control of infection in a single marmoset. Infection with approximately 1 CFU of CDC 1551 presented as 2 primary lesions, one in each of the lower lung lobes, as determined by PET–CT imaging[39,109] (Figure 2). Except for a period of transient weight loss beginning at day 76 after infection, this marmoset showed no other sign of clinical morbidity. Close observation by PET–CT imaging revealed that the lesion in the right lower lobe gradually increased in tracer uptake, whereas the signal in the lesion in the opposite lobe gradually decreased, suggesting at least partial immunologic control and disease resolution. This apparent discrepancy between the 2 individual

lesions is characteristic of the independent nature of granulomas even within the same NHP host[36,39]. These results were confirmed at necropsy, in that the lesion in the left lower lobe was sterile (that is, yielded no organisms by plating), whereas the lesion in the right lower lobe had approximately 2.6×10^3 bacteria. However, the microscopic pathology in the nodal lymph nodes and the presence of microscopic granulomas in the liver and spleen suggest that the infection may have been entering a more invasive stage, albeit at a slower rate than reported for all other marmosets to date. The histologic characteristics of the 2 lung lesions from this animal were distinct from the granulomas excised from the marmoset infected with 7 CFU of the same strain and from any of marmosets challenged with the Erdman strain. The left lower lobe lesion was an archetypical caseous granuloma with a well-circumscribed lymphocytic cuff, whereas the right lower lobe lesion had areas of tuberculosis pneumonia and nonnecrotizing granulomas. All of the Erdman-infected animals had a predominance of invasive, neutrophil-rich granulomatous alveolitis in their pulmonary sites that exhibited occasional organization and circumcision but did not form traditional defined granulomatous structures. Overall, the recovery of body weight and the lack of clinical manifestations coupled with the postmortem observations of decreased total cfu counts, diminished bacterial dissemination, and less extensive gross pathology suggest that the marmoset's greater capacity to limit disease progression was likely due to the reduced virulence of CDC1551 and the very low infection dose.

The primary limitation of our study is that the stable but chronic infection phenotype has only been observed in the single marmoset described. However, we found a strong relationship between the infectious dose of CDC1551 and weight loss in marmosets. Further infections with 1 to 3 CFU of CDC1551 are warranted in this small animal model to better define the mechanisms of control of low-dose infection. If this NHP species does exhibit relative resistance to *M.*

tuberculosis in a persistent manner, it would broaden their applicability as a small animal model of tuberculosis to include chronic infection studies yet retain their inherent benefits as a NHP model that offers decreased cost, small size[205], dizygotic twinning[206], pliability for drug studies[204], and the ability to replicate aspects of human disease[200,205]. In addition, marmosets might be developed as a model for vaccine-induced protection against tuberculosis.

In conclusion, optimizing the common marmoset (*C. jacchus*) model of *M. tuberculosis* infection requires identification of the range of host outcomes for this NHP species. In the current study, we noted variable disease progression in marmosets challenged with very low doses (1 to 12 CFU) of 2 strains of *M. tuberculosis*. Both the dose and strain of *M. tuberculosis* influenced the outcomes after challenge. Very low-dose challenge with the virulent Erdman strain did not ameliorate the rate of disease progression, because all of these marmosets (even those infected with 1 to 2 CFU) presented with rapidly disseminating active disease, resulting in clinical decline that prompted euthanasia by 43 days after inoculation. Infection with the less virulent CDC1551 strain resulted in delayed disease progression that was somewhat dependent on the inoculation dose. One of the most striking findings was that challenge with approximately 1 CFU of CDC1551 produced one case in which a marmoset effectively controlled the *M. tuberculosis* infection in a subclinical state for more than 300 days. At necropsy, this marmoset had reduced bacterial burden in its involved lymph nodes and lungs, reduced tissue dissemination, less overt gross pathology, and mildly progressing histology. These findings are in stark contrast to all previously reported *M. tuberculosis* infections of common marmosets, thus potentially extending this small-animal model beyond studies of acute infection while maintaining the clinical spectrum observed in the human disease.

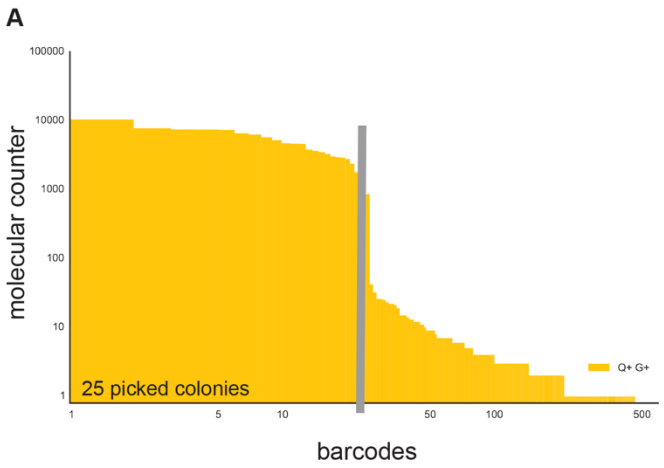
8.5 ACKNOWLEDGMENTS

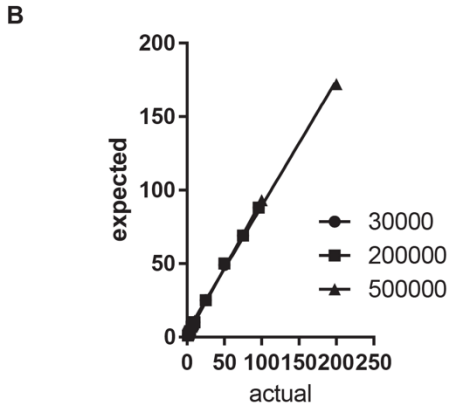
We thank Mark Rodgers, Carolyn Bigbee, Catherine Cochran, and Charles Scanga for technical assistance. We also thank Daniel Filmore for veterinary and animal assistance. We gratefully acknowledge Pauline Maiello and Teresa Coleman for imaging support and Danielle Weiner, Daniel Schimel, and Manny Dayao for veterinary assistance at NIAID. These studies were funded in part by a NIH training grant (T32 AI089443 [to AMC]), the Bill and Melinda Gates Foundation (grant no. OPP1034408 [to JLF]), and the intramural research program of NIAID, NIH (to LEV).

9.0 APPENDIX B: SUPPLEMENTARY TABLES & FIGURES

Table 6. Sequencing Primer Table.

Name	*Sequence*
Barcode Generation Primers	
CM29	GTACGAGGTACCCGANNNCNNNNaattcgatggcctagctgg
CM30	GGCCTATCTAGAgaccacaacggtttccatatg
Illumina Library Primers	
step 1	
Ftotal1	CCCTACACGACGCTCTTCCGATCTNCNNNCNNNCNNN
Ftotal2	CCCTACACGACGCTCTTCCGATCTNNCNNNCNNNCNNN
Ftotal3	CCCTACACGACGCTCTTCCGATCTNNNCNNNCNNNCNNN
Ftotal4	CCCTACACGACGCTCTTCCGATCTCNNNCNNNCNNN
Rtotal1	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCNGagaccacaacggtttccatatg
Rtotal2	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCNNNGAgaccacaacggtttccatatg
Rtotal3	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCNNNNGAgaccacaacggtttccatatg
Rtotal4	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCGAgaccacaacggtttccatatg
step 2	
F501	AATGATACGGCGACCACCGAGATCTACACTAT AGCCTACACTCTTTCCCTACACGACGCTCTTCC
F502	AATGATACGGCGACCACCGAGATCTACACATA GAGGCACACTCTTTCCCTACACGACGCTCTTCC
F503	AATGATACGGCGACCACCGAGATCTACACCCT ATCCTACACTCTTTCCCTACACGACGCTCTTCC
F504	AATGATACGGCGACCACCGAGATCTACACGGC TCTGAACACTCTTTCCCTACACGACGCTCTTCC
F505	AATGATACGGCGACCACCGAGATCTACACAGG CGAAGACACTCTTTCCCTACACGACGCTCTTCC
F506	AATGATACGGCGACCACCGAGATCTACACTAATCTTAACACTCTTTCCCTACACGACGCTCTTCC
F507	AATGATACGGCGACCACCGAGATCTACACCAGGACGTACACTCTTTCCCTACACGACGCTCTTCC
F508	AATGATACGGCGACCACCGAGATCTACACGTACTGACACACTCTTTCCCTACACGACGCTCTTCC
R701	CAAGCAGAAGACGGCATACGAGATCGAGTAATGTGACTGGAGTTCAGACGTGTGC
R702	CAAGCAGAAGACGGCATACGAGATTCTCCGGAGTGACTGGAGTTCAGACGTGTGC
R703	CAAGCAGAAGACGGCATACGAGATAATGAGCGGTGACTGGAGTTCAGACGTGTGC
R704	CAAGCAGAAGACGGCATACGAGATGGAATCTCGTGACTGGAGTTCAGACGTGTGC
R705	CAAGCAGAAGACGGCATACGAGATTTCTGAATGTGACTGGAGTTCAGACGTGTGC
R706	CAAGCAGAAGACGGCATACGAGATACGAATTCGTGACTGGAGTTCAGACGTGTGC
R707	CAAGCAGAAGACGGCATACGAGATAGCTTCAGGTGACTGGAGTTCAGACGTGTGC
R708	CAAGCAGAAGACGGCATACGAGATGCGCATTAGTGACTGGAGTTCAGACGTGTGC
R709	CAAGCAGAAGACGGCATACGAGATCATAGCCGGTGACTGGAGTTCAGACGTGTGC
R710	CAAGCAGAAGACGGCATACGAGATTTCGCGGAGTGACTGGAGTTCAGACGTGTGC
R711	CAAGCAGAAGACGGCATACGAGATGCGCGAGAGTGACTGGAGTTCAGACGTGTGC
R712	CAAGCAGAAGACGGCATACGAGATCTATCGCTGTGACTGGAGTTCAGACGTGTGC





read depth	30000	200000	500000
R square	0.9986	0.9981	0.9975

Figure 30. Barcode counting.

A. Pictorial example of BARTI thresholding 'real' barcodes from 'noise' arising from sequencing artifact. Unique barcodes from 25 picked *M. smegmatis* colonies are rank-ordered by number of molecular counters. The inflection point can be visually described as the largest drop-off between two barcode sequences and is mathematically described and found in the methods section (**Section 3.4.6**). B. BARTI has good accord between expected and actual found sequences across a range of read depths.

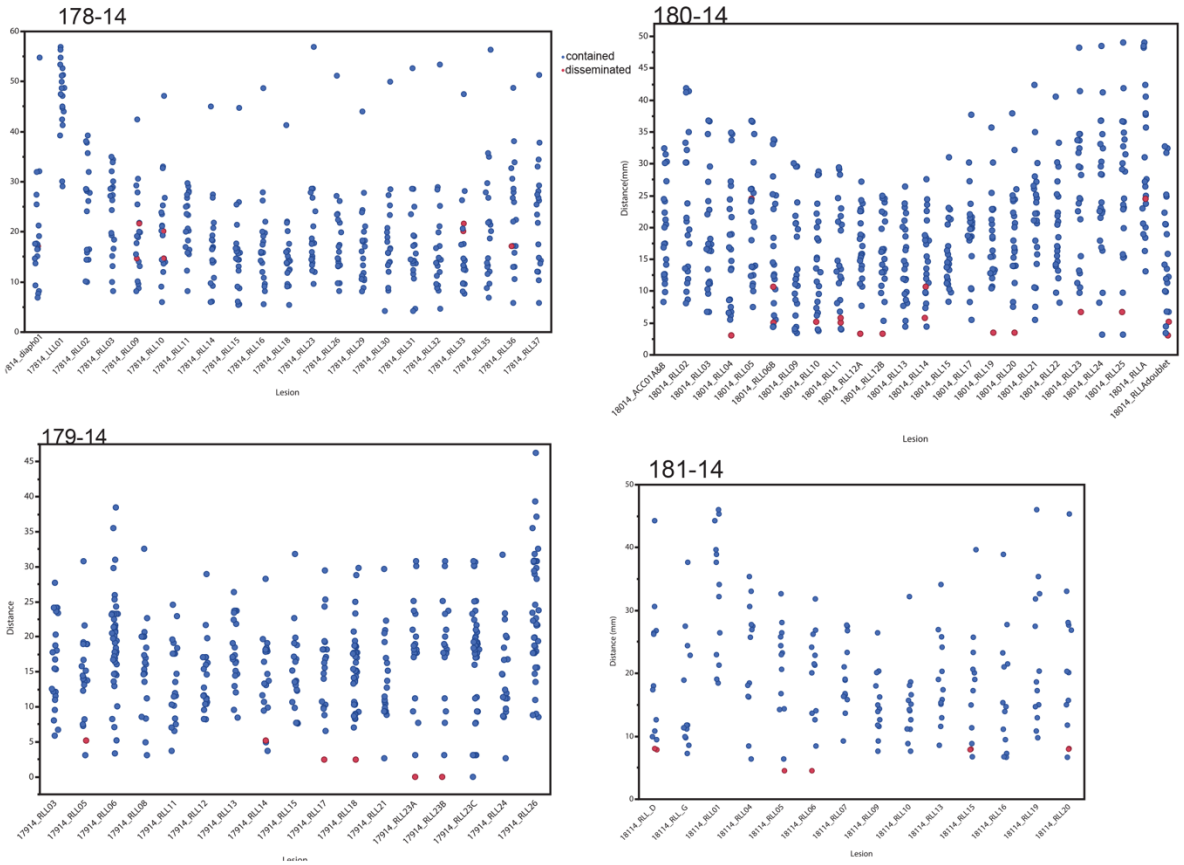


Figure 31. Distance between lesions.

Euclidean distance between each lung granuloma to every other lesion in that animal. Contained lesions in blue, disseminated lesions in red.

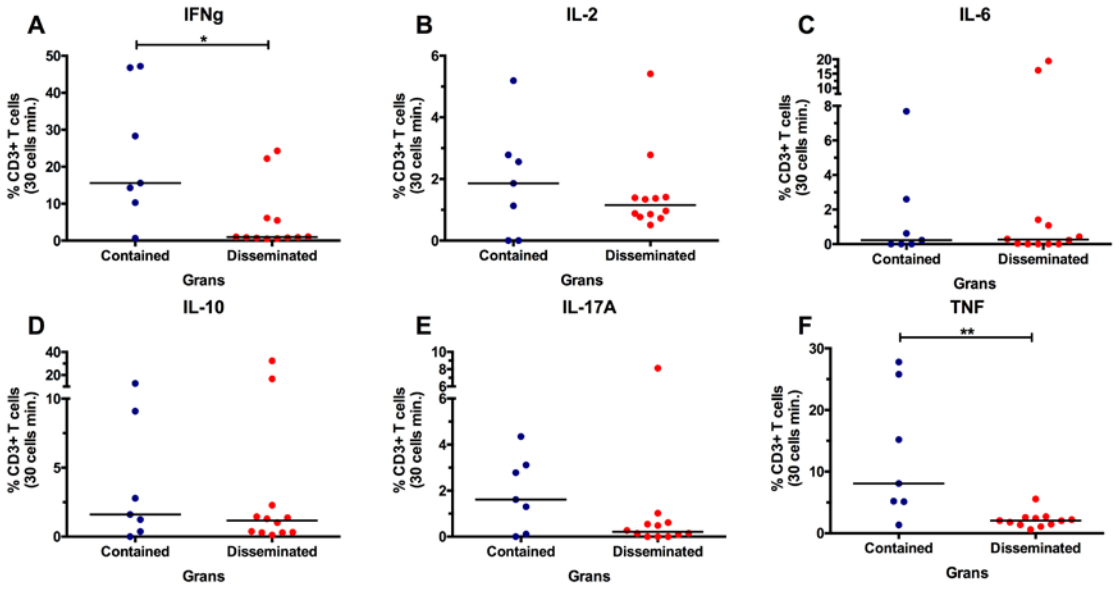


Figure 32. Frequency of cytokine producing T cells for IFN- (A), IL-2 (B), IL-6 (C), IL-10 (D), IL-17A (F), and TNF show minor differences between contained and disseminated lesions.

Each symbol is a granuloma. Analyzed contained (n=7) and disseminated (n=12) lesions had a minimum of 30 CD3+ T cells after processing and gating in FlowJo (ver. 9.9.5) for analysis (*p=0.0.283, **p=0.0098). Statistics for A-F: Mann Whitney.

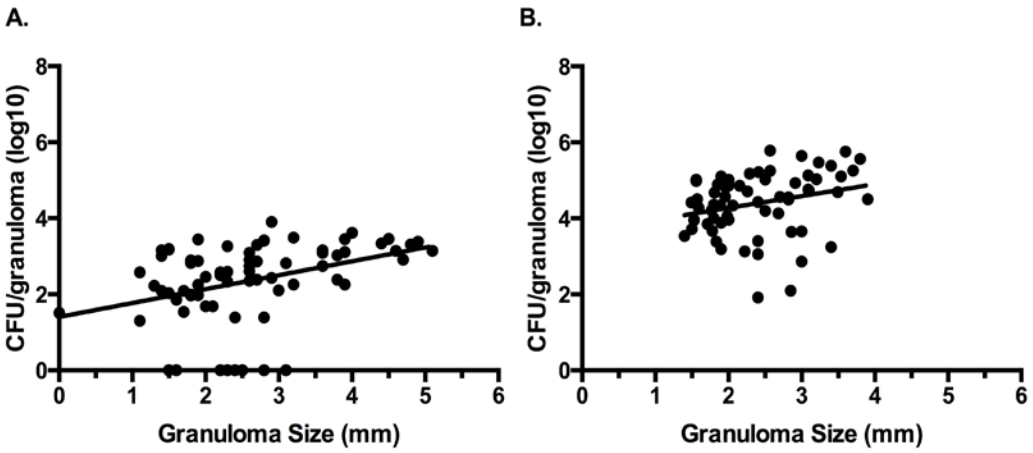


Figure 33. Size is not associated with granuloma bacterial burden.

There is a very weak correlation between granuloma size, as measured by PET/CT at 4-5 weeks post infection and CFU in granulomas necropsied at both 15-19 weeks post infection (A) (r^2=0.140, p=0.001, n=71) and 5-6 weeks (B) (r^2=0.066, p=0.0391) post infection. Linear regressions performed in Prism 6.

BIBLIOGRAPHY

1. World Health Organization. *Global tuberculosis report 2015*, (World Health Organization, Geneva, 2015).

2. Zumla, A., Raviglione, M., Hafner, R. & von Reyn, C.F. Tuberculosis. *The New England journal of medicine* **368**, 745-755 (2013).

3. Ottenhoff, T.H. & Kaufmann, S.H. Vaccines against tuberculosis: where are we and where do we need to go? *PLoS pathogens* **8**, e1002607 (2012).

4. Koul, A., Arnoult, E., Lounis, N., Guillemont, J. & Andries, K. The challenge of new drug discovery for tuberculosis. *Nature* **469**, 483-490 (2011).

5. Kagina, B.M., *et al.* Specific T cell frequency and cytokine expression profile do not correlate with protection against tuberculosis after bacillus Calmette-Guerin vaccination of newborns. *American journal of respiratory and critical care medicine* **182**, 1073-1079 (2010).

6. Walzl, G., Ronacher, K., Hanekom, W., Scriba, T.J. & Zumla, A. Immunological biomarkers of tuberculosis. *Nature reviews. Immunology* **11**, 343-354 (2011).

7. O'Garra, A., *et al.* The immune response in tuberculosis. *Annual review of immunology* **31**, 475-527 (2013).

8. Lenzini, L., Rottoli, P. & Rottoli, L. The spectrum of human tuberculosis. *Clin Exp Immunol* **27**, 230-237 (1977).

9. Lenaerts, A., Barry, C.E., 3rd & Dartois, V. Heterogeneity in tuberculosis pathology, microenvironments and therapeutic responses. *Immunological reviews* **264**, 288-307 (2015).

10. Getahun, H., Chaisson, R.E. & Raviglione, M. Latent Mycobacterium tuberculosis Infection. *The New England journal of medicine* **373**, 1179-1180 (2015).

11. Corbett, E.L., *et al.* The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. *Arch Intern Med* **163**, 1009-1021 (2003).

12. Lawn, S.D. & Zumla, A.I. Tuberculosis. *Lancet* **378**, 57-72 (2011).

13. Andrews, J.R., *et al.* Risk of progression to active tuberculosis following reinfection with Mycobacterium tuberculosis. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **54**, 784-791 (2012).

14. Barry, C.E., 3rd, *et al.* The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nature reviews. Microbiology* **7**, 845-855 (2009).

15. Lin, P.L. & Flynn, J.L. Understanding latent tuberculosis: a moving target. *J Immunol* **185**, 15-22 (2010).

16. Kassim, S., *et al.* Tuberculin skin testing to assess the occupational risk of Mycobacterium tuberculosis infection among health care workers in Abidjan, Cote d'Ivoire. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* **4**, 321-326 (2000).

17. Alcais, A., Fieschi, C., Abel, L. & Casanova, J.L. Tuberculosis in children and adults: two distinct genetic diseases. *The Journal of experimental medicine* **202**, 1617-1621 (2005).

18. Stead, W.W. Management of health care workers after inadvertent exposure to tuberculosis: a guide for the use of preventive therapy. *Ann Intern Med* **122**, 906-912 (1995).

19. Fliegauf, M., Sonnen, A.F., Kremer, B. & Henneke, P. Mucociliary clearance defects in a murine in vitro model of pneumococcal airway infection. *PloS one* **8**, e59925 (2013).

20. Young, D.B., Gideon, H.P. & Wilkinson, R.J. Eliminating latent tuberculosis. *Trends in microbiology* **17**, 183-188 (2009).

21. Esmail, H., *et al.* Characterization of progressive HIV-associated tuberculosis using 2-deoxy-2-[18F]fluoro-D-glucose positron emission and computed tomography. *Nature medicine* **22**, 1090-1093 (2016).

22. Lin, P.L., *et al.* PET CT Identifies Reactivation Risk in Cynomolgus Macaques with Latent M. tuberculosis. *PLoS pathogens* **12**, e1005739 (2016).

23. Berry, M.P., *et al.* An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature* **466**, 973-977 (2010).

24. Blankley, S., *et al.* The Transcriptional Signature of Active Tuberculosis Reflects Symptom Status in Extra-Pulmonary and Pulmonary Tuberculosis. *PloS one* **11**, e0162220 (2016).

25. Mehra, S., *et al.* Transcriptional reprogramming in nonhuman primate (rhesus macaque) tuberculosis granulomas. *PloS one* **5**, e12266 (2010).

26. Gideon, H.P., Skinner, J.A., Baldwin, N., Flynn, J.L. & Lin, P.L. Early Whole Blood Transcriptional Signatures Are Associated with Severity of Lung Inflammation in Cynomolgus Macaques with Mycobacterium tuberculosis Infection. *J Immunol* **197**, 4817-4828 (2016).

27. Cadena, A.M., Flynn, J.L. & Fortune, S.M. The Importance of First Impressions: Early Events in Mycobacterium tuberculosis Infection Influence Outcome. *mBio* **7**, e00342-00316 (2016).

28. Williams, G.T. & Williams, W.J. Granulomatous inflammation--a review. *J Clin Pathol* **36**, 723-733 (1983).

29. Russell, D.G., Cardona, P.J., Kim, M.J., Allain, S. & Altare, F. Foamy macrophages and the progression of the human tuberculosis granuloma. *Nature immunology* **10**, 943-948 (2009).

30. Flynn, J.L., Chan, J. & Lin, P.L. Macrophages and control of granulomatous inflammation in tuberculosis. *Mucosal immunology* **4**, 271-278 (2011).

31. Davis, J.M. & Ramakrishnan, L. The role of the granuloma in expansion and dissemination of early tuberculous infection. *Cell* **136**, 37-49 (2009).

32. Ramakrishnan, L. Revisiting the role of the granuloma in tuberculosis. *Nature reviews. Immunology* **12**, 352-366 (2012).

33. Flynn, J.L. Mutual attraction: does it benefit the host or the bug? *Nature immunology* **5**, 778-779 (2004).

34. Canetti, G. *The tubercle bacillus in the pulmonary lesion of man; histobacteriology and its bearing on the therapy of pulmonary tuberculosis*, (Springer Pub. Co., New York,, 1955).

35. Ford, C.B., *et al.* Use of whole genome sequencing to estimate the mutation rate of Mycobacterium tuberculosis during latent infection. *Nature genetics* **43**, 482-486 (2011).

36. Lin, P.L., *et al.* Sterilization of granulomas is common in active and latent tuberculosis despite within-host variability in bacterial killing. *Nature medicine* **20**, 75-79 (2014).

37. Ernst, J.D. The immunological life cycle of tuberculosis. *Nature reviews. Immunology* (2012).

38. Sia, J.K., Georgieva, M. & Rengarajan, J. Innate Immune Defenses in Human Tuberculosis: An Overview of the Interactions between Mycobacterium tuberculosis and Innate Immune Cells. *J Immunol Res* **2015**, 747543 (2015).

39. Coleman, M.T., *et al.* Early Changes by (18)Fluorodeoxyglucose positron emission tomography coregistered with computed tomography predict outcome after Mycobacterium tuberculosis infection in cynomolgus macaques. *Infection and immunity* **82**, 2400-2404 (2014).

40. Poulsen, A. Some clinical features of tuberculosis. 1. Incubation period. *Acta Tuberc Scand* **24**, 311-346 (1950).

41. Weldingh, K. & Andersen, P. ESAT-6/CFP10 skin test predicts disease in M. tuberculosis-infected guinea pigs. *PloS one* **3**, e1978 (2008).

42. Capuano, S.V., 3rd, *et al.* Experimental Mycobacterium tuberculosis infection of cynomolgus macaques closely resembles the various manifestations of human M. tuberculosis infection. *Infect Immun* **71**, 5831-5844 (2003).

43. Pelletier, M., Forget, A., Bourassa, D. & Skamene, E. Histological and immunopathological studies of delayed hypersensitivity reaction to tuberculin in mice. *Infect Immun* **46**, 873-875 (1984).

44. Wolf, A.J., *et al.* Initiation of the adaptive immune response to Mycobacterium tuberculosis depends on antigen production in the local lymph node, not the lungs. *The Journal of experimental medicine* **205**, 105-115 (2008).

45. Poulsen, A. Some clinical features of tuberculosis. *Acta Tuberc Scand* **33**, 37-92; concl (1957).

46. Gedde-Dahl, T. Tuberculous infection in the light of tuberculin matriculation. *Am J Hyg* **56**, 139-214 (1952).

47. Shapiro, A.E., *et al.* Community-based targeted case finding for tuberculosis and HIV in household contacts of patients with tuberculosis in South Africa. *Am J Respir Crit Care Med* **185**, 1110-1116 (2012).

48. Marais, B.J., *et al.* The prevalence of symptoms associated with pulmonary tuberculosis in randomly selected children from a high burden community. *Arch Dis Child* **90**, 1166-1170 (2005).

49. Lin, P.L., *et al.* Quantitative comparison of active and latent tuberculosis in the cynomolgus macaque model. *Infection and immunity* **77**, 4631-4642 (2009).

50. Lin, P.L., *et al.* Radiologic responses in cynomolgous macaques for assessing tuberculosis chemotherapy regimens. *Antimicrobial agents and chemotherapy* **57**, 4237–4244 (2013).

51. Canetti, G. *The Tubercle Bacillus*, (Springer Publishing Co, Inc., New York, NY, 1955).

52. Turner, O.C., Basaraba, R.J. & Orme, I.M. Immunopathogenesis of pulmonary granulomas in the guinea pig after infection with Mycobacterium tuberculosis. *Infect Immun* **71**, 864-871 (2003).

53. Lin, P.L., *et al.* Early events in Mycobacterium tuberculosis infection in cynomolgus macaques. *Infection and immunity* **74**, 3790-3803 (2006).

54. Chen, M., *et al.* Lipid mediators in innate immunity against tuberculosis: opposing roles of PGE2 and LXA4 in the induction of macrophage death. *The Journal of experimental medicine* **205**, 2791-2801 (2008).

55. Divangahi, M., Desjardins, D., Nunes-Alves, C., Remold, H.G. & Behar, S.M. Eicosanoid pathways regulate adaptive immunity to Mycobacterium tuberculosis. *Nature immunology* **11**, 751-758 (2010).

56. Elkington, P., *et al.* MMP-1 drives immunopathology in human tuberculosis and transgenic mice. *J Clin Invest* **121**, 1827-1833 (2011).

57. Al Shammari, B., *et al.* The Extracellular Matrix Regulates Granuloma Necrosis in Tuberculosis. *J Infect Dis* **212**, 463-473 (2015).

58. Volkman, H.E., *et al.* Tuberculous granuloma induction via interaction of a bacterial secreted protein with host epithelium. *Science* **327**, 466-469 (2010).

59. Dannenberg, A.M., Jr. Immune mechanisms in the pathogenesis of pulmonary tuberculosis. *Rev Infect Dis* **11 Suppl 2**, S369-378 (1989).

60. Flynn, J.L., J. Chan. Immunology of tuberculosis. *Ann Rev Immunol* **19**, 93-129 (2001).

61. Clay, H., *et al.* Dichotomous role of the macrophage in early Mycobacterium marinum infection of the zebrafish. *Cell Host Microbe* **2**, 29-39 (2007).

62. Wolf, A.J., *et al.* Mycobacterium tuberculosis infects dendritic cells with high frequency and impairs their function in vivo. *J Immunol* **179**, 2509-2519 (2007).

63. Humphreys, I.R., *et al.* A role for dendritic cells in the dissemination of mycobacterial infection. *Microbes and infection / Institut Pasteur* **8**, 1339-1346 (2006).

64. Tsai, M.C., *et al.* Characterization of the tuberculous granuloma in murine and human lungs: cellular composition and relative tissue oxygen tension. *Cellular microbiology* **8**, 218-232 (2006).

65. Srivastava, S., Ernst, J.D. & Desvignes, L. Beyond macrophages: the diversity of mononuclear cells in tuberculosis. *Immunological reviews* **262**, 179-192 (2014).

66. Matty, M.A., Roca, F.J., Cronan, M.R. & Tobin, D.M. Adventures within the speckled band: heterogeneity, angiogenesis, and balanced inflammation in the tuberculous granuloma. *Immunol Rev* **264**, 276-287 (2015).

67. Bloom, C.I., *et al.* Transcriptional blood signatures distinguish pulmonary tuberculosis, pulmonary sarcoidosis, pneumonias and lung cancers. *PloS one* **8**, e70630 (2013).

68. Eum, S.Y., *et al.* Neutrophils are the predominant infected phagocytic cells in the airways of patients with active pulmonary TB. *Chest* **137**, 122-128 (2010).

69. Dorhoi, A., *et al.* MicroRNA-223 controls susceptibility to tuberculosis by regulating lung neutrophil recruitment. *J Clin Invest* **123**, 4836-4848 (2013).

70. Yang, C.T., *et al.* Neutrophils exert protection in the early tuberculous granuloma by oxidative killing of mycobacteria phagocytosed from infected macrophages. *Cell host & microbe* **12**, 301-312 (2012).

71. Blomgran, R. & Ernst, J.D. Lung neutrophils facilitate activation of naive antigen-specific CD4+ T cells during Mycobacterium tuberculosis infection. *J Immunol* **186**, 7110-7119 (2011).

72. Lowe, D.M., Redford, P.S., Wilkinson, R.J., O'Garra, A. & Martineau, A.R. Neutrophils in tuberculosis: friend or foe? *Trends Immunol* **33**, 14-25 (2012).

73. Martineau, A.R., *et al.* Neutrophil-mediated innate immune resistance to mycobacteria. *J Clin Invest* **117**, 1988-1994 (2007).

74. Nandi, B. & Behar, S.M. Regulation of neutrophils by interferon-gamma limits lung inflammation during tuberculosis infection. *The Journal of experimental medicine* **208**, 2251-2262 (2011).

75. Mayer-Barber, K.D., *et al.* Host-directed therapy of tuberculosis based on interleukin-1 and type I interferon crosstalk. *Nature* **511**, 99-103 (2014).

76. Bannenberg, G. & Serhan, C.N. Specialized pro-resolving lipid mediators in the inflammatory response: An update. *Biochim Biophys Acta* **1801**, 1260-1273 (2010).

77. Tobin, D.M., *et al.* Host genotype-specific therapies can optimize the inflammatory response to mycobacterial infections. *Cell* **148**, 434-446 (2012).

78. Tobin, D.M., *et al.* The lta4h locus modulates susceptibility to mycobacterial infection in zebrafish and humans. *Cell* **140**, 717-730 (2010).

79. Chackerian, A.A., Alt, J.M., Perera, T.V., Dascher, C.C. & Behar, S.M. Dissemination of Mycobacterium tuberculosis is influenced by host factors and precedes the initiation of T-cell immunity. *Infect Immun* **70**, 4501-4509 (2002).

80. Myers, A.J., Marino, S., Kirschner, D.E. & Flynn, J.L. Inoculation dose of Mycobacterium tuberculosis does not influence priming of T cell responses in lymph nodes. *Journal of immunology* **190**, 4707-4716 (2013).

81. Bhatt, K., Hickman, S.P. & Salgame, P. Cutting edge: a new approach to modeling early lung immunity in murine tuberculosis. *J Immunol* **172**, 2748-2751 (2004).

82. Samstein, M., *et al.* Essential yet limited role for CCR2(+) inflammatory monocytes during Mycobacterium tuberculosis-specific T cell priming. *Elife* **2**, e01086 (2013).

83. Day, T.A., *et al.* Secondary lymphoid organs are dispensable for the development of T-cell-mediated immunity during tuberculosis. *Eur J Immunol* **40**, 1663-1673 (2010).

84. Harding, J.S., Rayasam, A., Schreiber, H.A., Fabry, Z. & Sandor, M. Mycobacterium-Infected Dendritic Cells Disseminate Granulomatous Inflammation. *Sci Rep* **5**, 15248 (2015).

85. Mishra, B.B., *et al.* Nitric oxide controls the immunopathology of tuberculosis by inhibiting NLRP3 inflammasome-dependent processing of IL-1beta. *Nature immunology* **14**, 52-60 (2013).

86. Napier, R.J., Adams, E.J., Gold, M.C. & Lewinsohn, D.M. The Role of Mucosal Associated Invariant T Cells in Antimicrobial Immunity. *Front Immunol* **6**, 344 (2015).

87. Kjer-Nielsen, L., *et al.* MR1 presents microbial vitamin B metabolites to MAIT cells. *Nature* **491**, 717-723 (2012).

88. Gold, M.C., *et al.* MR1-restricted MAIT cells display ligand discrimination and pathogen selectivity through distinct T cell receptor usage. *J Exp Med* **211**, 1601-1610 (2014).

89. Gold, M.C., Napier, R.J. & Lewinsohn, D.M. MR1-restricted mucosal associated invariant T (MAIT) cells in the immune response to Mycobacterium tuberculosis. *Immunol Rev* **264**, 154-166 (2015).

90. Portevin, D., Via, L.E., Eum, S. & Young, D. Natural killer cells are recruited during pulmonary tuberculosis and their ex vivo responses to mycobacteria vary between healthy human donors in association with KIR haplotype. *Cell Microbiol* **14**, 1734-1744 (2012).

91. Hall-Stoodley, L., *et al.* Mycobacterium tuberculosis binding to human surfactant proteins A and D, fibronectin, and small airway epithelial cells under shear conditions. *Infect Immun* **74**, 3587-3596 (2006).

92. Beharka, A.A., *et al.* Pulmonary surfactant protein A up-regulates activity of the mannose receptor, a pattern recognition receptor expressed on human macrophages. *J Immunol* **169**, 3565-3573 (2002).

93. Gold, J.A., *et al.* Surfactant protein A modulates the inflammatory response in macrophages during tuberculosis. *Infect Immun* **72**, 645-650 (2004).

94. Ferguson, J.S., Voelker, D.R., McCormack, F.X. & Schlesinger, L.S. Surfactant protein D binds to Mycobacterium tuberculosis bacilli and lipoarabinomannan via carbohydrate-lectin interactions resulting in reduced phagocytosis of the bacteria by macrophages. *J Immunol* **163**, 312-321 (1999).

95. Lemos, M.P., McKinney, J. & Rhee, K.Y. Dispensability of surfactant proteins A and D in immune control of Mycobacterium tuberculosis infection following aerosol challenge of mice. *Infect Immun* **79**, 1077-1085 (2011).

96. Cakir, E., *et al.* Cathelicidin and human beta-defensin 2 in bronchoalveolar lavage fluid of children with pulmonary tuberculosis. *Int J Tuberc Lung Dis* **18**, 671-675 (2014).

97. Sonawane, A., *et al.* Cathelicidin is involved in the intracellular killing of mycobacteria in macrophages. *Cell Microbiol* **13**, 1601-1617 (2011).

98. Rivas-Santiago, B., *et al.* Activity of LL-37, CRAMP and antimicrobial peptide-derived compounds E2, E6 and CP26 against Mycobacterium tuberculosis. *Int J Antimicrob Agents* **41**, 143-148 (2013).

99. Martineau, A.R., *et al.* Reciprocal seasonal variation in vitamin D status and tuberculosis notifications in Cape Town, South Africa. *Proc Natl Acad Sci U S A* **108**, 19013-19017 (2011).

100. Martineau, A.R., *et al.* A single dose of vitamin D enhances immunity to mycobacteria. *Am J Respir Crit Care Med* **176**, 208-213 (2007).

101. Fabri, M., *et al.* Vitamin D is required for IFN-gamma-mediated antimicrobial activity of human macrophages. *Science translational medicine* **3**, 104ra102 (2011).

102. Coussens, A.K., *et al.* Vitamin D accelerates resolution of inflammatory responses during tuberculosis treatment. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 15449-15454 (2012).

103. Martineau, A.R., *et al.* IFN-gamma- and TNF-independent vitamin D-inducible human suppression of mycobacteria: the role of cathelicidin LL-37. *J Immunol* **178**, 7190-7198 (2007).

104. Liu, P.T., Stenger, S., Tang, D.H. & Modlin, R.L. Cutting edge: vitamin D-mediated human antimicrobial activity against Mycobacterium tuberculosis is dependent on the induction of cathelicidin. *J Immunol* **179**, 2060-2063 (2007).

105. Chan, J., *et al.* The role of B cells and humoral immunity in Mycobacterium tuberculosis infection. *Semin Immunol* **26**, 588-600 (2014).

106. Achkar, J.M., Chan, J. & Casadevall, A. B cells and antibodies in the defense against Mycobacterium tuberculosis infection. *Immunol Rev* **264**, 167-181 (2015).

107. Ford, C.B., *et al.* Use of whole genome sequencing to estimate the mutation rate of Mycobacterium tuberculosis during latent infection. *Nature genetics* **43**, 482-486 (2011).

108. Vandiviere, H.M., Loring, W.E., Melvin, I. & Willis, S. The treated pulmonary lesion and its tubercle bacillus. II. The death and resurrection. *Am J Med Sci* **232**, 30-37; passim (1956).

109. Coleman, M.T., *et al.* PET/CT imaging reveals a therapeutic response to oxazolidinones in macaques and humans with tuberculosis. *Sci Transl Med* **6**, 265ra167 (2014).

110. Gagneux, S., *et al.* The competitive cost of antibiotic resistance in Mycobacterium tuberculosis. *Science* **312**, 1944-1946 (2006).

111. Manca, C., *et al.* Differential monocyte activation underlies strain-specific Mycobacterium tuberculosis pathogenesis. *Infect Immun* **72**, 5511-5514 (2004).

112. Lopez, B., *et al.* A marked difference in pathogenesis and immune response induced by different Mycobacterium tuberculosis genotypes. *Clin Exp Immunol* **133**, 30-37 (2003).

113. Dormans, J., *et al.* Correlation of virulence, lung pathology, bacterial load and delayed type hypersensitivity responses after infection with different Mycobacterium tuberculosis genotypes in a BALB/c mouse model. *Clin Exp Immunol* **137**, 460-468 (2004).

114. Portevin, D., Gagneux, S., Comas, I. & Young, D. Human macrophage responses to clinical isolates from the Mycobacterium tuberculosis complex discriminate between ancient and modern lineages. *PLoS Pathog* **7**, e1001307 (2011).

115. Reed, M.B., *et al.* A glycolipid of hypervirulent tuberculosis strains that inhibits the innate immune response. *Nature* **431**, 84-87 (2004).

116. Cambier, C.J., *et al.* Mycobacteria manipulate macrophage recruitment through coordinated use of membrane lipids. *Nature* **505**, 218-222 (2014).

117. Lee, W., VanderVen, B.C., Fahey, R.J. & Russell, D.G. Intracellular Mycobacterium tuberculosis exploits host-derived fatty acids to limit metabolic stress. *J Biol Chem* **288**, 6788-6800 (2013).

118. Griffin, J.E., *et al.* Cholesterol catabolism by Mycobacterium tuberculosis requires transcriptional and metabolic adaptations. *Chem Biol* **19**, 218-227 (2012).

119. Cox, J.S., Chen, B., McNeil, M. & Jacobs, W.R., Jr. Complex lipid determines tissue-specific replication of Mycobacterium tuberculosis in mice. *Nature* **402**, 79-83 (1999).

120. Sassetti, C.M. & Rubin, E.J. Genetic requirements for mycobacterial survival during infection. *Proc Natl Acad Sci U S A* **100**, 12989-12994 (2003).

121. Myllymaki, H., Bauerlein, C.A. & Ramet, M. The Zebrafish Breathes New Life into the Study of Tuberculosis. *Frontiers in immunology* **7**, 196 (2016).

122. Kramnik, I. & Beamer, G. Mouse models of human TB pathology: roles in the analysis of necrosis and the development of host-directed therapies. *Semin Immunopathol* **38**, 221-237 (2016).

123. Scanga, C.A. & Flynn, J.L. Modeling tuberculosis in nonhuman primates. *Cold Spring Harb Perspect Med* **4**, a018564 (2014).

124. Flynn, J.L., Gideon, H.P., Mattila, J.T. & Lin, P.L. Immunology studies in non-human primate models of tuberculosis. *Immunological reviews* **264**, 60-73 (2015).

125. Capuano, S.V., *et al.* Experimental Mycobacterium tuberculosis Infection of Cynomolgus Macaques Closely Resembles the Various Manifestations of Human M. tuberculosis Infection. *Infection and immunity* **71**, 5831-5844 (2003).

126. Diedrich, C.R., *et al.* Reactivation of latent tuberculosis in cynomolgus macaques infected with SIV is associated with early peripheral T cell depletion and not virus load. *PloS one* **5**, e9611 (2010).

127. Lin, P.L., *et al.* Tumor necrosis factor neutralization results in disseminated disease in acute and latent Mycobacterium tuberculosis infection with normal granuloma structure in a cynomolgus macaque model. *Arthritis and rheumatism* **62**, 340-350 (2010).

128. Lin, P.L., *et al.* CD4 T cell depletion exacerbates acute Mycobacterium tuberculosis while reactivation of latent infection is dependent on severity of tissue depletion in cynomolgus macaques. *AIDS Res Hum Retroviruses* **28**, 1693-1702 (2012).

129. Dickson, R.P., Erb-Downward, J.R., Martinez, F.J. & Huffnagle, G.B. The Microbiome and the Respiratory Tract. *Annu Rev Physiol* **78**, 481-504 (2016).

130. Global tuberculosis report 2015. (ed. Global, W.) (World Health Organization, 2015).

131. Modlin, R.L. & Bloom, B.R. TB or not TB: that is no longer the question. *Sci Transl Med* **5**, 213sr216 (2013).

132. Lin, P.L., *et al.* Sterilization of granulomas is common in active and latent tuberculosis despite within-host variability in bacterial killing. *Nature medicine* **20**, 75-79 (2014).

133. Gideon, H.P., *et al.* Variability in tuberculosis granuloma T cell responses exists, but a balance of pro- and anti-inflammatory cytokines is associated with sterilization. *PLoS pathogens* **11**, e1004603 (2015).

134. Abel, S., Abel zur Wiesch, P., Davis, B.M. & Waldor, M.K. Analysis of Bottlenecks in Experimental Models of Infection. *PLoS Pathog* **11**, e1004823 (2015).

135. Blumenthal, A., Trujillo, C., Ehrt, S. & Schnappinger, D. Simultaneous analysis of multiple Mycobacterium tuberculosis knockdown mutants in vitro and in vivo. *PloS one* **5**, e15667 (2010).

136. Casbon, J.A., Osborne, R.J., Brenner, S. & Lichtenstein, C.P. A method for counting PCR template molecules with application to next-generation sequencing. *Nucleic Acids Res* **39**, e81 (2011).

137. Wu, L., *et al.* Phasing amplicon sequencing on Illumina Miseq for robust environmental microbial community analysis. *BMC Microbiol* **15**, 125 (2015).

138. Abel, S., *et al.* Sequence tag-based analysis of microbial population dynamics. *Nat Methods* **12**, 223-226, 223 p following 226 (2015).

139. Lu, R., Neff, N.F., Quake, S.R. & Weissman, I.L. Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding. *Nat Biotechnol* **29**, 928-933 (2011).

140. Blundell, J.R. & Levy, S.F. Beyond genome sequencing: lineage tracking with barcodes to study the dynamics of evolution, infection, and cancer. *Genomics* **104**, 417-430 (2014).

141. Buschmann, T. & Bystrykh, L.V. Levenshtein error-correcting barcodes for multiplexed DNA sequencing. *BMC Bioinformatics* **14**, 272 (2013).

142. Bystrykh, L.V. Generalized DNA barcode design based on Hamming codes. *PLoS One* **7**, e36852 (2012).

143. Grant, A.J., *et al.* Modelling within-host spatiotemporal dynamics of invasive bacterial disease. *PLoS Biol* **6**, e74 (2008).

144. Joseph, S.B., Swanstrom, R., Kashuba, A.D. & Cohen, M.S. Bottlenecks in HIV-1 transmission: insights from the study of founder viruses. *Nature reviews. Microbiology* **13**, 414-425 (2015).

145. Dean, G.S., *et al.* Minimum infective dose of Mycobacterium bovis in cattle. *Infect Immun* **73**, 6467-6471 (2005).

146. Turner, R.D. & Bothamley, G.H. Cough and the transmission of tuberculosis. *J Infect Dis* **211**, 1367-1372 (2015).

147. Post, F.A., *et al.* Genetic polymorphism in Mycobacterium tuberculosis isolates from patients with chronic multidrug-resistant tuberculosis. *J Infect Dis* **190**, 99-106 (2004).

148. Lieberman, T.D., *et al.* Genomic diversity in autopsy samples reveals within-host dissemination of HIV-associated Mycobacterium tuberculosis. *Nature medicine* (2016).

149. Wang, J.Y., *et al.* Prediction of the tuberculosis reinfection proportion from the local incidence. *J Infect Dis* **196**, 281-288 (2007).

150. Cohen, T. & Murray, M. Incident tuberculosis among recent US immigrants and exogenous reinfection. *Emerg Infect Dis* **11**, 725-728 (2005).

151. Cohen, T., *et al.* Mixed-strain mycobacterium tuberculosis infections and the implications for tuberculosis treatment and control. *Clin Microbiol Rev* **25**, 708-719 (2012).

152. Charalambous, S., *et al.* Contribution of reinfection to recurrent tuberculosis in South African gold miners. *Int J Tuberc Lung Dis* **12**, 942-948 (2008).

153. Verver, S., *et al.* Rate of reinfection tuberculosis after successful treatment is higher than rate of new tuberculosis. *Am J Respir Crit Care Med* **171**, 1430-1435 (2005).

154. Theisen, A., *et al.* Mixed-strain infection with a drug-sensitive and multidrug-resistant strain of Mycobacterium tuberculosis. *Lancet* **345**, 1512 (1995).

155. Chiang, C.Y. & Riley, L.W. Exogenous reinfection in tuberculosis. *Lancet Infect Dis* **5**, 629-636 (2005).

156. Sutherland, I., Svandova, E. & Radhakrishna, S. The development of clinical tuberculosis following infection with tubercle bacilli. 1. A theoretical model for the development of clinical tuberculosis following infection, linking from data on the risk of tuberculous infection and the incidence of clinical tuberculosis in the Netherlands. *Tubercle* **63**, 255-268 (1982).

157. Vynnycky, E. & Fine, P.E. The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection. *Epidemiology and infection* **119**, 183-201 (1997).

158. Zumla, A., Nahid, P. & Cole, S.T. Advances in the development of new tuberculosis drugs and treatment regimens. *Nature reviews. Drug discovery* **12**, 388-404 (2013).

159. Churchyard, G.J., *et al.* A trial of mass isoniazid preventive therapy for tuberculosis control. *The New England journal of medicine* **370**, 301-310 (2014).

160. Henao-Tamayo, M., *et al.* A mouse model of tuberculosis reinfection. *Tuberculosis (Edinb)* **92**, 211-217 (2012).

161. Wallis, R.S., *et al.* Tuberculosis biomarkers discovery: developments, needs, and challenges. *The Lancet Infectious Diseases* **13**, 362-372 (2013).

162. Round, J.L. & Mazmanian, S.K. The gut microbiota shapes intestinal immune responses during health and disease. *Nature reviews. Immunology* **9**, 313-323 (2009).

163. Rooks, M.G. & Garrett, W.S. Gut microbiota, metabolites and host immunity. *Nature reviews. Immunology* **16**, 341-352 (2016).

164. Marsland, B.J. & Gollwitzer, E.S. Host-microorganism interactions in lung diseases. *Nature reviews. Immunology* **14**, 827-835 (2014).

165. Dickson, R.P., Martinez, F.J. & Huffnagle, G.B. The role of the microbiome in exacerbations of chronic lung diseases. *Lancet* **384**, 691-702 (2014).

166. Flanagan, J.L., *et al.* Loss of bacterial diversity during antibiotic treatment of intubated patients colonized with Pseudomonas aeruginosa. *Journal of clinical microbiology* **45**, 1954-1962 (2007).

167. Cui, L., *et al.* Topographical Diversity of the Respiratory Tract Mycobiome and Alteration in HIV and Lung Disease. *American journal of respiratory and critical care medicine* (2015).

168. Cui, L., *et al.* The microbiome and the lung. *Annals of the American Thoracic Society* **11 Suppl 4**, S227-232 (2014).

169. Morris, A., *et al.* Comparison of the respiratory microbiome in healthy nonsmokers and smokers. *American journal of respiratory and critical care medicine* **187**, 1067-1075 (2013).

170. Twigg, H.L., 3rd, *et al.* Use of bronchoalveolar lavage to assess the respiratory microbiome: signal in the noise. *The Lancet. Respiratory medicine* **1**, 354-356 (2013).

171. Cheung, M.K., *et al.* Sputum microbiota in tuberculosis as revealed by 16S rRNA pyrosequencing. *PloS one* **8**, e54574 (2013).

172. Cui, Z., *et al.* Complex sputum microbial composition in patients with pulmonary tuberculosis. *BMC microbiology* **12**, 276 (2012).

173. Wu, J., *et al.* Sputum microbiota associated with new, recurrent and treatment failure tuberculosis. *PloS one* **8**, e83445 (2013).

174. Botero, L.E., *et al.* Respiratory tract clinical sample selection for microbiota analysis in patients with pulmonary tuberculosis. *Microbiome* **2**, 29 (2014).

175. Winglee, K., *et al.* Aerosol Mycobacterium tuberculosis infection causes rapid loss of diversity in gut microbiota. *PloS one* **9**, e97048 (2014).

176. Perry, S., *et al.* Infection with Helicobacter pylori is associated with protection against tuberculosis. *PloS one* **5**, e8804 (2010).

177. Schloss, P.D., Gevers, D. & Westcott, S.L. Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PloS one* **6**, e27310 (2011).

178. Schloss, P.D., *et al.* Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and environmental microbiology* **75**, 7537-7541 (2009).

179. Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K. & Schloss, P.D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and environmental microbiology* **79**, 5112-5120 (2013).

180. Cole, J.R., *et al.* The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic acids research* **37**, D141-145 (2009).

181. Edgar, R.C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460-2461 (2010).

182. Caporaso, J.G., *et al.* PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* **26**, 266-267 (2010).

183. Price, M.N., Dehal, P.S. & Arkin, A.P. FastTree 2--approximately maximum-likelihood trees for large alignments. *PloS one* **5**, e9490 (2010).

184. Hubbell, S.P. Neutral theory in community ecology and the hypothesis of functional equivalence. *Functional Ecology* **19**, 166-172 (2005).

185. Caporaso, J.G., *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nature methods* **7**, 335-336 (2010).

186. Morris, A., *et al.* Longitudinal analysis of the lung microbiota of cynomolgous macaques during long-term SHIV infection. *Microbiome* **4**, 38 (2016).

187. Segata, N., *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol* **12**, R60 (2011).

188. Huffnagle, G.B., Dickson, R.P. & Lukacs, N.W. The respiratory tract microbiome and lung inflammation: a two-way street. *Mucosal immunology* (2016).

189. Lin, P.L., *et al.* Tumor necrosis factor neutralization results in disseminated disease in acute and latent Mycobacterium tuberculosis infection with normal granuloma structure in a cynomolgus macaque model. *Arthritis and Rheumatism* **62**, 340-350 (2010).

190. Egelund, E.F., Alsultan, A. & Peloquin, C.A. Optimizing the clinical pharmacology of tuberculosis medications. *Clin Pharmacol Ther* **98**, 387-393 (2015).

191. Moliva, J.I., Turner, J. & Torrelles, J.B. Prospects in Mycobacterium bovis Bacille Calmette et Guerin (BCG) vaccine diversity and delivery: why does BCG fail to protect against tuberculosis? *Vaccine* **33**, 5035-5041 (2015).

192. Clark, R.A. Resident memory T cells in human health and disease. *Science translational medicine* **7**, 269rv261 (2015).

193. Honda, K. & Littman, D.R. The microbiome in infectious disease and inflammation. *Annual review of immunology* **30**, 759-795 (2012).

194. Thaiss, C.A., Zmora, N., Levy, M. & Elinav, E. The microbiome and innate immunity. *Nature* **535**, 65-74 (2016).

195. Zumla, A., *et al.* The WHO 2014 global tuberculosis report--further to go. *Lancet Glob Health* **3**, e10-12 (2015).

196. Flynn, J.L. Lessons from experimental Mycobacterium tuberculosis infections. *Microbes and infection / Institut Pasteur* **8**, 1179-1188 (2006).

197. Chen, C.Y., *et al.* A critical role for CD8 T cells in a nonhuman primate model of tuberculosis. *PLoS pathogens* **5**, e1000392 (2009).

198. Verreck, F.A., *et al.* MVA.85A boosting of BCG and an attenuated, phoP deficient M. tuberculosis vaccine both show protective efficacy against tuberculosis in rhesus macaques. *PloS one* **4**, e5264 (2009).

199. Kaushal, D., Mehra, S., Didier, P.J. & Lackner, A.A. The non-human primate model of tuberculosis. *Journal of medical primatology* (2012).

200. Via, L.E., *et al.* Differential Virulence and Disease Progression following Mycobacterium tuberculosis Complex Infection of the Common Marmoset (Callithrix jacchus). *Infection and immunity* **81**, 2909-2919 (2013).

201. Hernandez-Pando, R., Marquina-Castillo, B., Barrios-Payan, J. & Mata-Espinosa, D. Use of mouse models to study the variability in virulence associated with specific genotypic lineages of Mycobacterium tuberculosis. *Infect Genet Evol* **12**, 725-731 (2012).

202. Manabe, Y.C., *et al.* Different strains of Mycobacterium tuberculosis cause various spectrums of disease in the rabbit model of tuberculosis. *Infection and immunity* **71**, 6004-6011 (2003).

203. Ong, C.W., Elkington, P.T. & Friedland, J.S. Tuberculosis, pulmonary cavitation, and matrix metalloproteinases. *American journal of respiratory and critical care medicine* **190**, 9-18 (2014).

204. Via, L.E., *et al.* A sterilizing tuberculosis treatment regimen is associated with faster clearance of bacteria in cavitary lesions in marmosets. *Antimicrobial agents and chemotherapy* **59**, 4181-4189 (2015).

205. Mansfield, K. Marmoset models commonly used in biomedical research. *Comp Med* **53**, 383-392 (2003).

206. Abbott, D.H., Barnett, D.K., Colman, R.J., Yamamoto, M.E. & Schultz-Darken, N.J. Aspects of common marmoset basic biology and life history important for biomedical research. *Comp Med* **53**, 339-350 (2003).