RESEARCH ARTICLE

A computational model tracks whole-lung *Mycobacterium tuberculosis* infection and predicts factors that inhibit dissemination

Timothy Wessler[1,2‡], Louis R. Joslyn[1,2‡], H. Jacob Borish[3], Hannah P. Gideon[3], JoAnne L. Flynn[3], Denise E. Kirschner[2]*, Jennifer J. Linderman[1]*

1 Department of Chemical Engineering, University of Michigan, Ann Arbor, Michigan, United States of America, **2** Department of Microbiology and Immunology, University of Michigan, Ann Arbor, Michigan, United States of America, **3** Department of Microbiology and Molecular Genetics, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, United States of America

‡ These authors share first authorship on this work.
* kirschne@umich.edu (DEK); linderma@umich.edu (JJL)

Abstract

Mycobacterium tuberculosis (Mtb), the causative infectious agent of tuberculosis (TB), kills more individuals per year than any other infectious agent. Granulomas, the hallmark of Mtb infection, are complex structures that form in lungs, composed of immune cells surrounding bacteria, infected cells, and a caseous necrotic core. While granulomas serve to physically contain and immunologically restrain bacteria growth, some granulomas are unable to control Mtb growth, leading to bacteria and infected cells leaving the granuloma and disseminating, either resulting in additional granuloma formation (local or non-local) or spread to airways or lymph nodes. Dissemination is associated with development of active TB. It is challenging to experimentally address specific mechanisms driving dissemination from TB lung granulomas. Herein, we develop a novel hybrid multi-scale computational model, *MultiGran*, that tracks Mtb infection within multiple granulomas in an entire lung. *MultiGran* follows cells, cytokines, and bacterial populations within each lung granuloma throughout the course of infection and is calibrated to multiple non-human primate (NHP) cellular, granuloma, and whole-lung datasets. We show that *MultiGran* can recapitulate patterns of *in vivo* local and non-local dissemination, predict likelihood of dissemination, and predict a crucial role for multifunctional CD8+ T cells and macrophage dynamics for preventing dissemination.

Author summary

Tuberculosis (TB) is caused by infection with *Mycobacterium tuberculosis* (Mtb) and kills 3 people per minute worldwide. Granulomas, spherical structures composed of immune cells surrounding bacteria, are the hallmark of Mtb infection and sometimes fail to contain the bacteria and disseminate, leading to further granuloma growth within the lung environment. To date, the mechanisms that determine granuloma dissemination events have not been characterized. We present a computational multi-scale model of granuloma

formation and dissemination within primate lungs. Our computational model is calibrated to multiple experimental datasets across the cellular, granuloma, and whole-lung scales of non-human primates. We match to both individual granuloma and granuloma-population datasets, predict likelihood of dissemination events, and predict a critical role for multifunctional CD8+ T cells and macrophage-bacteria interactions to prevent infection dissemination.

Introduction

Tuberculosis (TB) kills more individuals per year than any other infectious disease and treatment remains a global challenge [1]. Only a small fraction (5–10%) of those infected with *Mycobacterium tuberculosis* (Mtb) develop active symptomatic disease [2], while the remainder control but do not eliminate the infection, which is termed latent TB (LTBI). A hallmark of Mtb infection is the presence of lung granulomas (lesions), collections of immune cells that surround Mtb in an effort to contain and control an infection. Multiple granulomas can be present in humans and non-human primates (NHPs). In NHPs, each granuloma is initiated by a single bacillus [3]. Of key importance is that each granuloma within an individual has its own independent trajectory behavior. For example, the immune response in some granulomas eliminates all bacteria, resulting in sterilization. In other granulomas, immune cells only contain Mtb growth, resulting in stable granulomas that may persist for decades [4]. If Mtb growth is not contained, however, granulomas can grow and/or spread, allowing for dissemination of bacteria across the lungs leading to the formation of new granulomas, spread to the airways resulting in transmission of infection through aerosolized bacteria, and possibly death of the host if not treated. Understanding the collective behavior of granulomas within lungs leading to dissemination events is critical to the ultimate goal of controlling the global TB epidemic.

It is difficult to experimentally address specific mechanisms operating within lungs that drive different granuloma outcomes in primates, although it is known through interventional studies that certain factors, such as TNF, CD4+ T cells, and CD8+ T cells are important in controlling early and established Mtb infection [5–8]. As a complementary approach, mathematical modeling can generate hypotheses that can then be tested experimentally. Several mathematical and computational models for Mtb infection have been developed to explore the contributions of the innate and adaptive immune responses to granuloma formation and function [9–20]. These models are informed by studies in humans and in animal models of infection, especially NHPs, rabbits, pigs, and mice [21]. In particular, *GranSim*, our computational model that allows simulation of the formation and function of a single granuloma using a hybrid agent-based model framework, has offered strategies for drug treatment and vaccine development [12,14,22–24]. *GranSim*, which considers thousands of cells and bacteria as "agents" in the simulation and tracks diffusion of multiple immune mediators (e.g., cytokines), is computationally intensive, limiting our ability to simultaneously simulate multiple granulomas present in an entire lung during infection. In contrast, Prats et al. [18] utilized a bubble model to demonstrate the importance of local inflammation, dissemination, and coalescence of lesions as key factors leading to active TB, but did not specifically model events at the granuloma scale. However, following the formation of individual granulomas, the dissemination of those granulomas across the lungs over time, and, importantly, tracking events at the granuloma scale could provide an important window into infection dynamics and could lead to new insights for prevention or treatment.
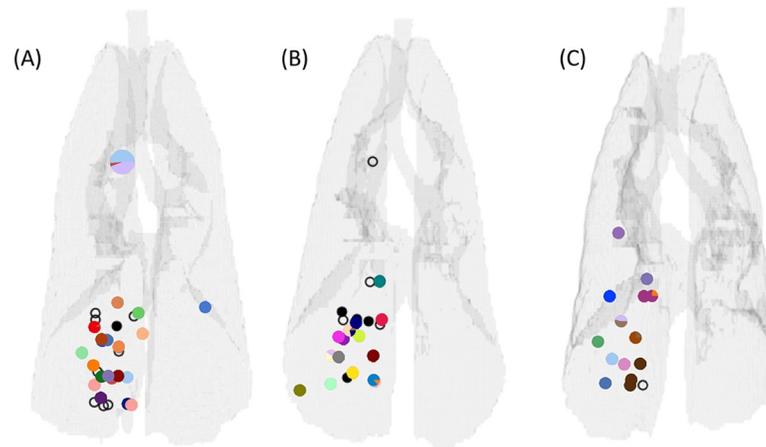
Fig 1. Three NHP lung maps illustrating the positioning of pulmonary granulomas and thoracic lymph nodes (data previously published in Martin et al. [25]). Gray outlines denote the extent of the lungs, bronchial tubes, and trachea. Small markers superimposed on the outlines represent the positions of pulmonary granulomas, while larger markers denote lymph nodes. Colors denote unique barcode tags. Some samples had more than one barcode tag present, and often these were doublet granulomas (i.e., two granulomas too close in proximity to distinguish at necropsy) and so are marked with a pie chart showing the relative abundance of each barcode tag. The black markers represent pulmonary granulomas for which no barcode tags were found. Filled black markers are granulomas which grew bacteria upon plating but barcodes could not be determined for technical reasons, while open markers are granulomas that did not grow bacteria upon plating (sterile).

In order to study the formation of new granulomas after initial establishment of infection, referred to as dissemination, the evolution of individual granulomas must be captured over time. Recently, research on Mtb-infected NHPs provided data on disseminating granulomas [25]. Of all animal models used to study Mtb infection, NHPs are most relevant to human TB disease because they recapitulate the full spectrum of clinical outcomes and pathologies seen in humans [26]. From PET CT imaging, the emergence of new granulomas was tracked and recorded. The authors genetically matched Mtb barcodes, assigned each inoculation Mtb a unique barcode ID, and associated each granuloma identified in the temporal PET CT images with the Mtb barcodes inside that granuloma (Fig 1). By identifying Mtb barcodes that were present in multiple granulomas, they were able to distinguish disseminating from non-disseminating granulomas. When identifying multiple bacterial barcodes within a single granuloma, it is surmised a merger of granulomas took place. While Martin et al. showed these distinctions, the mechanisms that lead to granuloma clustering or dissemination remain unanswered. We address these open questions using a hybrid computational-mathematical modeling framework.

Herein, we develop a novel multi-scale hybrid model, *MultiGran*, to track Mtb infection at the scale of the entire lung, including capturing multiple granulomas and their individual outcomes as well as the formation of new granulomas. *MultiGran* is an agent-based model that follows cells, cytokines, and bacterial populations across multiple lung granulomas throughout the course of infection. Each granuloma is now formulated as a single agent, and each agent contains within it a system of non-linear ordinary differential equations (ODEs) that capture individual granuloma dynamics. *MultiGran* follows the steps observed through the course of Mtb infection: (1) *initial granuloma establishment* with Mtb that have been virtually barcoded and placed within the lung environment, (2) *granuloma development* across time, (3) the possibility of *granuloma dissemination* with barcoded bacteria moving to a new location, and (4) *granuloma merging* by granulomas that have formed close together and whose individual

boundaries are indistinguishable, or those that grow in size and thus merge into a granuloma cluster (that may have multiple barcoded bacteria IDs). We use *MultiGran* to address three outstanding questions about dissemination: what mechanisms are consistent with granuloma dissemination and merging patterns seen *in vivo*? What is the likelihood of a granuloma to disseminate? Can we predict factors that lead to dissemination?

Methods

Ethics statement

All experimental manipulations, protocols, and care of the animals were approved by the University of Pittsburgh School of Medicine Institutional Animal Care and Use Committee (IACUC). The protocol assurance number for our IACUC is A3187-01. Our specific protocol approval numbers for this project are 1280653, 12126588, 11110045, 19024273, 15066174, 16017309 and 18124275. The IACUC adheres to national guidelines established in the Animal Welfare Act (7 U.S.C. Sections 2131–2159) and the Guide for the Care and Use of Laboratory Animals (8th Edition) as mandated by the U.S. Public Health Service Policy.

All macaques used in this study were housed at the University of Pittsburgh in rooms with autonomously controlled temperature, humidity, and lighting. Animals were singly housed in caging at least 2 square meters apart that allowed visual and tactile contact with neighboring conspecifics. The macaques were fed twice daily with biscuits formulated for NHPs, supplemented at least 4 days/week with large pieces of fresh fruits or vegetables. Animals had access to water *ad libitem*. Because our macaques were singly housed due to the infectious nature of these studies, an enhanced enrichment plan was designed and overseen by our nonhuman primate enrichment specialist. This plan has three components. First, species-specific behaviors are encouraged. All animals have access to toys and other manipulata, some of which will be filled with food treats (e.g., frozen fruit, peanut butter). These are rotated on a regular basis. Puzzle feeders foraging boards, and cardboard tubes containing small food items also are placed in the cage to stimulate foraging behaviors. Adjustable mirrors accessible to the animals stimulate interaction between animals. Second, routine interaction between humans and macaques are encouraged. These interactions occur daily and consist mainly of small food objects offered as enrichment and adhere to established safety protocols. Animal caretakers are encouraged to interact with the animals (by talking or with facial expressions) while performing tasks in the housing area. Routine procedures (e.g., feeding, cage cleaning) are done on a strict schedule to allow the animals to acclimate to a routine daily schedule. Third, all macaques are provided with a variety of visual and auditory stimulation. Housing areas contain either radios or TV/video equipment that play cartoons or other formats designed for children for at least 3 hours each day. The videos and radios are rotated between animal rooms so that the same enrichment is not played repetitively for the same group of animals.

All animals are checked at least twice daily to assess appetite, attitude, activity level, hydration status, etc. Following Mtb infection, the animals are monitored closely for evidence of disease (e.g., anorexia, weight loss, tachypnea, dyspnea, coughing). Physical exams, including weights, are performed on a regular basis. Animals are sedated prior to all veterinary procedures (e.g., blood draws) using ketamine or other approved drugs. Regular PET/CT imaging is conducted on most of our macaques following infection and has proved very useful for monitoring disease progression. Our veterinary technicians monitor animals especially closely for any signs of pain or distress. If any are noted, appropriate supportive care (e.g., dietary supplementation, rehydration) and clinical treatments (analgesics) are given. Any animal considered to have advanced disease or intractable pain or distress from any cause is sedated with ketamine and then humanely euthanized using sodium pentobarbital.

Experimental dataset

Experimental data specifically for this study were obtained from seven cynomolgus macaques (*Macaca fascicularis*), infected with low dose Mtb (Erdman strain, ~10 CFU by bronchoscopic instillation) as previously described [27–29]. Infection was confirmed by PET CT imaging. PET CT scans were performed monthly to quantify new granuloma formation or clustering, as well as disease progression. Necropsy was performed as previously described [28,29]. Briefly, an ^{18}F-FDG PET-CT scan was performed on every animal 1–3 days prior to necropsy to measure disease progression and identify individual granulomas and other pathologies as described [27–30]; this scan was used as a map for identifying individual lesions. At necropsy, each granuloma or other pathologies from lung and mediastinal lymph nodes were obtained for histological analysis, bacterial burden, and immunological studies, including flow cytometry, as previously described [27–30]. For bacterial burden, each granuloma homogenate was plated onto 7H11 medium, and the CFU were enumerated 21 days later to determine the number of bacilli in each granuloma [27,29].

To calibrate the individual granuloma computational model, we excised granulomas from macaques that were infected for 3 weeks (n = 2), 5 weeks (n = 2), 7 weeks (n = 2) and 9 weeks (n = 1). In addition, an animal without Mtb infection was also included in this study as a control. To obtain accurate cell number measurements, enzymatic digestion (Tumor dissociation kit, human; Miltenyi Biotec) was performed on excised granulomas using gentleMACS octo dissociator. The single cell suspension obtained by enzymatic digestion was processed for bacterial burden and cell numbers enumeration [27]. Single cell suspensions of individual granulomas were stained with cell surface antibodies to enumerate T cells (CD3) and macrophages (CD11b). The cells were further stained intracellularly with Calprotectin antibody to exclude CD11b+Calprotectin+ cells from macrophage population. Flow cytometry and data acquisition was performed using BD LSRII and analysis was performed using Flowjo Software v10 [27].

In addition, bacterial burden data of 623 granulomas from 38 NHP that were controls in other studies (previously published [20,27,31–33] and ongoing studies) at University of Pittsburgh (Flynn Lab) were included for evaluation. The timing of infection depended on the particular study (Table of CFU values and tables of cell counts located at http://malthus.micro.med.umich.edu/lab/movies/MultiGran/ Table: Granuloma_CFU_counts) and ranged from 4–17 weeks post Mtb infection.

Non-human primate lung lattice data

To create a virtual lung that replicates an NHP lung, we used a CT scan of an uninfected NHP to model the 3-dimensional lung space. Binary images mapping the cross section of the lungs were created for each CT slice by segmentation of CT image values below -320 Houndsfield units. The individual slices were then stacked into an array, and a polygon mesh outlining the lung volume was generated using the marching_cubes_classic function in the open source Python scikit-image package (v 0.14.1, [34]).

Identifying granuloma distributions in non-human primate lungs

To allow us to test whether the distribution of granulomas in our virtual lungs matched that observed in NHP lungs, we refer to the distribution of granulomas arising from barcoded bacteria derived from our previously published data in Martin et al. [25]. In that study, four cynomolgus macaques were infected with 11+/- 5 CFU barcoded Mtb Erdman. Barcoded libraries were generated where each bacterium has a different random 7-mer along with one of three 75-mer identifier tags inserted into the bacterial chromosome. This process created roughly

50,000 bacteria that are able to be uniquely identified by the random 7-mer tag with very small (< 2%) risk of duplication in an infection of <50 CFU (See Fig 1 in Martin et al. [25]). The animals were necropsied between 15 and 20 weeks post-infection. Animals were imaged at monthly intervals (or more frequently) to identify timing of granuloma establishment. Pulmonary granulomas were excised during necropsy, and their three-dimensional positions were recorded via matching to PET/CT imaging. Homogenates from excised pulmonary granulomas and infected thoracic lymph nodes were plated, scraped, and sequenced to identify the specific barcode(s) present in each granuloma. Matching the x, y, and z coordinates recorded for each granuloma with its determined barcode content led to a three-dimensional map of the locations of each barcode throughout the pulmonary space. Bacterial burden for each granuloma was determined by counting colonies on the plates.

Three of the four maps are shown in Fig 1 (the fourth was already presented in the original paper [25]). Lung outlines were calculated from terminal scans of each NHP by the process of creating a polygon mesh described above. Small markers represent pulmonary granulomas, while larger markers denote lymph nodes. Each color represents a unique barcode tag. Some samples had more than one barcode tag present, and often these were doublet granulomas (i.e., two granulomas too close in proximity to distinguish at necropsy) and so are marked with a pie chart showing the relative abundance of each barcode tag. The black markers represent pulmonary granulomas for which no barcode tags were found. Filled black markers are granulomas which grew bacteria upon plating but for which barcodes could not be determined, while open markers are granulomas that did not grow bacteria upon plating (sterile); in this study, only CFU+ granulomas were available for barcode determination.

Model overview

MultiGran is a novel multi-scale, hybrid agent-based model that describes the formation, function, and dissemination of lung granulomas containing Mtb (Fig 2). It uses sampling of non-homogeneous Poisson processes; rule-based agent placement; parameter randomization; solving systems of non-linear ODEs; and post-process agent groupings to perform *in silico* experiments that track the progress of infection in an individual host. Each granuloma (agent) is placed stochastically within the boundary of the lung environment based on a set of rules. Within each agent, a system of ODEs is linked internally and solved simultaneously to update concentrations of cells, cytokines, and bacterial burdens within each granuloma at every time step. Additionally, within every time step, each granuloma is given the opportunity to disseminate locally and non-locally. *Local dissemination* involves a new granuloma being initialized nearby, while *non-local dissemination* allows initialization anywhere within the lung environment. At the lung scale, the model tracks the development, location, and quantity of granulomas, and determines whether each granuloma is either alone or a member of a larger granuloma cluster. At the granuloma scale, dissemination-event decisions, rules for granuloma formation, and concentrations of all granuloma components are tracked and defined. As is occasionally done when a flexible agent size is needed [35], our agents (granulomas) are placed on a continuous grid. Agents are spherical with dynamically-changing sizes, and granuloma clustering depends on the geometry and position of each of the agents.

Each *in silico* experiment using *MultiGran* is designed to replicate an *in vivo* experiment. To replicate the studies by Martin et al. [25], our simulated NHP is infected with roughly 19 uniquely-identified (barcoded) Mtb that are randomly placed in a localized region of the lungs, similar to the typical inoculation process in the NHP experiments. Each Mtb is assumed to be immediately taken up by a resident lung macrophage, forming a single, unique new granuloma [25]. Each granuloma evolves independently. Whenever a granuloma is formed, it is
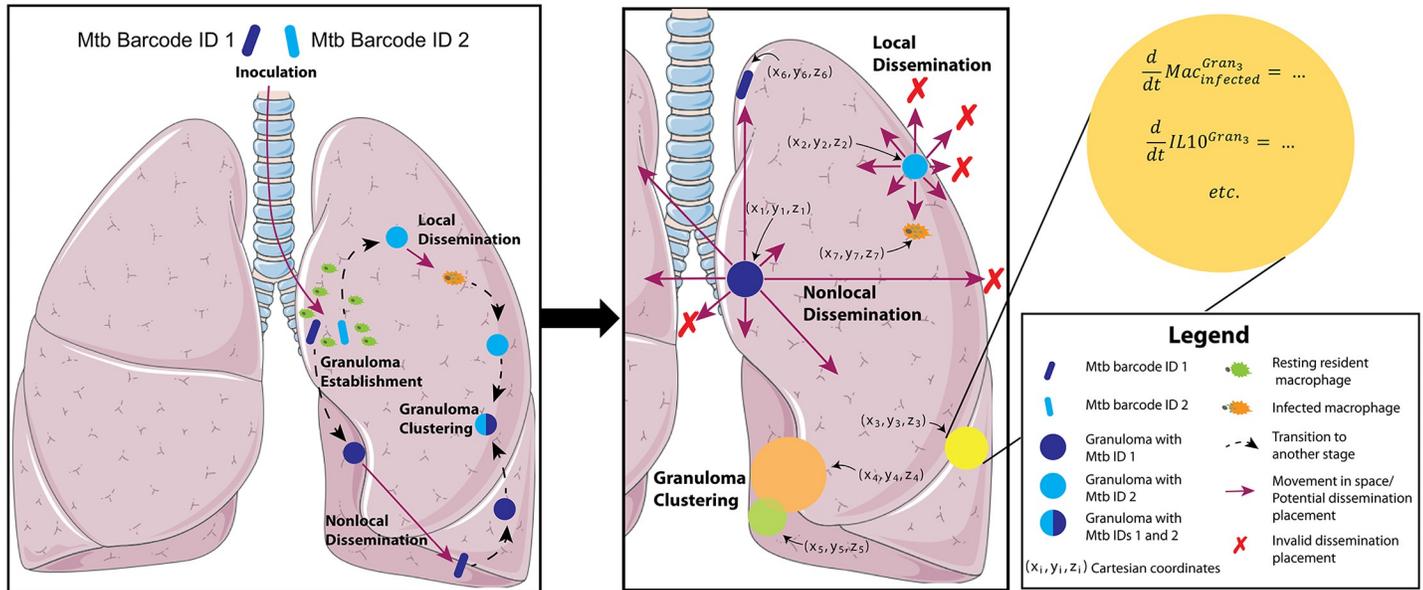
Fig 2. Process of Mtb infection and rules for granuloma dissemination and location within *MultiGran*. A non-human primate is inoculated with Mtb, here tracked using different "barcodes" or IDs. These Mtb are taken up by resident macrophages, initiating an innate immune response. This response includes the secretion of various cytokines and chemokines that help prime and/or recruit other immune cells to the site of the infection, resulting in the formation of lung granulomas. The dynamics within each granuloma are governed by systems of ordinary differential equations. Occasionally, as a granuloma develops, it may disseminate—either locally or non-locally. In local dissemination, an Mtb-infected macrophage moves to another nearby location within the same lung lobe. In non-local dissemination, a free extracellular Mtb reaches the airways or is carried to a draining lymph node and then deposited at a site not necessarily near the original location; i.e., in a different lung lobe. The (x,y,z) positions of disseminated granulomas depend on the method of dissemination (i.e., local or nonlocal) and the position of the parent granuloma. Granuloma clusters can form when granulomas develop near each other and may grow into each other, or when one granuloma forms immediately adjacent to the original granuloma via local dissemination [3]. Granuloma clusters may contain more than one Mtb ID.

initialized with parameter values that represent several characteristics that ultimately influence its future behavior, as well as the emergent outcomes of the system as a whole.

Simulation environment

Code is written in MATLAB, with Bash script for submission to run on computer clusters. ODEs are solved using MATLAB's ode15s with the NonNegative option for all terms, and we define the start and end time interval to be the size of the agent time step. To avoid complications with the random number generator seed being reset with the initialization of each MATLAB instance, the Bash script executes code that generates a randomized seed list for the simulation to use. The website http://malthus.micro.med.umich.edu/lab/movies/MultiGran/ has pseudocode and implementation descriptions, as well as simulation videos.

Granuloma establishment

A granuloma is initialized when Mtb is deposited into the lung environment. Based on our previous publications [3,25], we assume that each Mtb creates one granuloma [3,36]. The granulomas established during inoculation (S1 Fig–Case 1) are referred to as "founder" granulomas and are considered first-generation granulomas; all other granulomas that may emerge throughout the simulation originate from these founders.

Granulomas are agents, so at initialization we assign parameter values to each granuloma and its infecting Mtb, as well as counts and concentrations of all cell types and cytokines. Every granuloma is assigned unique identification markers. These include being given a unique individual granuloma ID *IndivGranID(i)*, which is assigned in chronological order of

initialization *i = 1,2,. . .N* (where *N* is the total number of granulomas), as well as the individual granuloma ID of its parent, so the lineage of each of the founder Mtb can be tracked throughout the course of infection. Each granuloma is also given a position on a continuous grid.

Granuloma development

The development of each individual granuloma "agent" is captured by a set of ODEs with 16 equations for 16 state variables capturing bacterial, T cell, macrophage and cytokine dynamics (see S1 Equations for complete term-by-term description of the model). ODE model formulations build on our previous work [37–39] describing cells and levels of cytokines in a whole lung. The equations have been re-calibrated to NHP granuloma data (see section on Experimental dataset) to represent an individual granuloma (see section Model Calibration), and have been updated in several ways. First, we increased the role of IL-10, including it as a factor for downregulating macrophage activation and TNF-α production by activated macrophages, as well as allowing infected macrophages to produce IL-10, based on NHP data [40–42]. The other set of changes relates to intra- and extra-cellular Mtb to be consistent with recent findings on Mtb growth within macrophages [43–45]. Rather than releasing the entire carrying capacity of bacteria at the occurrence of each death of an infected macrophage, the amount of intracellular Mtb within an average infected macrophage is released (with the exception of a bursting infected macrophage, in which case the maximum amount of Mtb is released). Furthermore, only a fraction of intracellular Mtb released during the natural death of an infected macrophage survives to become an extracellular Mtb. The expression for intracellular Mtb replication was also changed along with the addition of an expression for the natural slow death of intracellular Mtb for model stability. We record granuloma sterilization when the count of Mtb drops below 0.5.

Granuloma dissemination

While the mechanisms behind dissemination are not yet well-understood [25], we have created rules such that the emergent outcomes are consistent with experiments (S1 Fig). We define a probability function for likelihood of a dissemination event, which we make dependent on the bacterial load (CFU) of the granuloma. We selected CFU because the data presented by Lin et al. [3] indicates that granuloma carrying capacity has a limit (approximately 10^5). Because NHP granulomas rarely exceed this limit [3,28], there is likely a link between granuloma CFU and dissemination. Because Mtb is by itself non-motile, we consider two routes of dissemination: 1) Mtb conveyance within an infected macrophage and 2) a single Mtb flowing through lung airways or deposited via a draining lymph node (LN). From these, we incorporated two types of dissemination events: local and non-local, the probabilities of each event being independent, and in the unlikely event that multiple dissemination events occur in the same time step, the order of events is randomized.

When a granuloma disseminates locally (S1 Fig–Case 3), an infected macrophage carrying intracellular Mtb is assumed to move from the parent granuloma position to a new position nearby. We assume the distance between the parent granuloma and a new position likely follows a normal distribution with respect to parent location and we calibrated the mean and variance of this location using the data presented in Martin et al. [25]. In Martin et al., the authors compute distances of each granuloma and granuloma clusters that they could identify via PET/CT, rather than every individual granuloma regardless of size and cluster affiliation. We also assume that a pre-determined quantity of T cells moves with an infected macrophage. After this dissemination event, the parent and daughter granulomas evolve independently from each other. When a granuloma disseminates non-locally (S1 Fig–Case 2), an extracellular

Mtb is simulated as if entering airways (or via a LN) and deposited with equal likelihood anywhere within the lungs, where it is immediately taken up by a macrophage. S1 Fig—Case 2 represents three realizations of trial coordinates wherein the trial coordinates represented by the red arrow do not satisfy our criteria, but the two black arrows would be acceptable placements for a bacterium in non-local dissemination.

We created two dissemination event probabilities describing local and non-local dissemination. In both, λ is the maximum probability of dissemination and is scaled by a Michaelis-Menten fraction, using a value of CFU at which the probability is half of the maximum value.

$$Prob_{Local}(t) = \lambda_{Local} \frac{CFU(t)}{CFU(t) + CFU_{half}^{Local}} \qquad \text{Eq1}$$

$$Prob_{Nonlocal}(t) = \lambda_{Nonlocal} \frac{CFU(t)}{CFU(t) + CFU_{half}^{Nonlocal}} \qquad \text{Eq2}$$

Granuloma merging

Experiments demonstrate that a subset of granulomas contain a more than one Mtb barcode [25]. Following inoculation or dissemination events, individual granulomas may merge, or are sufficiently close to each other, to form clusters. We identify granuloma clusters and their members when needed for plotting and computing statistics but allow them to evolve independently. Briefly, our algorithm evaluates all intersections of granulomas, and combines groups of granulomas that intersect in 3D space. These grouped granulomas are the granuloma clusters. A granuloma cluster may contain only descendants of a single founder Mtb ID, or may contain descendants of multiple founder Mtb IDs.

Model calibration

An effective strategy for calibration of a complex system such as our multiscale model is evaluating the model's ability (and model parameters) to reproduce multiple scenarios and datasets [46–48], reducing the likelihood of overfitting parameters to a single dataset. Herein, we have built our multi-scale model of granuloma formation and dissemination using *a priori* biological knowledge and utilized multiple datasets, across multiple scales from unique experiments, to calibrate and validate our model parameters. Table 1 displays the multiple distinct datasets that we used to calibrate and validate this multi-scale model. First, we identified the parameter space of the individual granuloma ODE model that best represents the individual granuloma datasets (CFU and cell counts). See Fig 3 for an example of our workflow. To determine an initial, wide range of parameter values to test, we examined experimental values from literature, previous ODE models of a single granuloma formation [37–39], and values from *GranSim* [14,16,19,22,37–40,49]. We then used a Latin Hypercube Sampling (LHS) algorithm [50] to sample this multi-dimensional parameter space 500 times. This initial wide range of simulations did not match the NHP data. We then narrowed the initial ranges and resampled the space in an iterative process until, out of the 500 simulations, ninety percent of the runs fell within the bounds of our experimental data on CFU, T cell counts, and macrophages within individual NHP granulomas. Fig 3 shows how we identify simulations that satisfy our criteria and isolate the parameters of those simulations to narrow parameter space. The parameter ranges for the calibrated model runs are in S1 Table, and represent the parameter space that we sampled in order to create the granulomas for our biorepository of 200 virtual monkey lungs and perform global sensitivity analysis techniques on.

Table 1. Distinct datasets used to calibrate, validate and inform prediction-making in *MultiGran*. We calibrated our single-granuloma model, and then calibrated parameters specific for dissemination at the whole-lung level, prior to validation and predictions of individual NHP outcomes.

Dataset	Use	Model (scale)
Experimental values from literature, parameter values from previous ODE models of a single granuloma formation [37–39], and values from *GranSim* [14,16,19,22,37–40,49].	Initial sampling of entire single-granuloma parameter space	Single-granuloma ODE model (granuloma-level)
CFU/Granuloma Dataset Includes 623 granulomas from 38 NHPs	Calibration	Single-granuloma ODE model (granuloma-level)
Individual Granuloma Cell Count Dataset Includes 26 granulomas from 7 NHPs including total T cell counts and macrophage counts	Calibration	Single-granuloma ODE model (granuloma-level)
Granuloma Location Data Mean and variance of granuloma dissemination presented in Martin et al. [25]	Initial sampling of entire dissemination parameter space	*MultiGran* model (lung-level)
Whole Lung Outcome Measures Four outcomes—(1) number of Mtb at time of inoculation (2) number of granulomas at necropsy (3) the percentage of Mtb barcodes found in multiple granulomas and (4) the percentage of granulomas containing multiple Mtb barcodes	Calibration of dissemination parameters	*MultiGran* model (lung-level)
Spread of Infection Outcome Measures Including 38 NHP and the location of 623 granulomas across both lungs	Model validation	*MultiGran* model (lung-level)
Individual NHP CFU/granuloma Individual CFU/granuloma for a single NHP	Model prediction	*MultiGran* predictions about dissemination (multi-scale)

Next, we identified the dissemination parameter space of *MultiGran* that matched the NHP whole lung outcome datasets (previously published [20,25,27,31–33] and ongoing studies). We again utilized LHS to sample this space and identify baseline parameter ranges that match the data (S3 Table). Fig 3 explains our workflow for calibrating the single-granuloma ODE model, which was similar to the approach for calibrating the lung-level dissemination equations. This calibrated parameter space was sampled to create a biorepository of *in silico* lungs to be used in Results. The entire repository of 200 virtual NHP CFU trends are displayed at http://malthus. micro.med.umich.edu/lab/movies/MultiGran/.

Sensitivity analysis

We then used Partial Rank Correlation Coefficient (PRCC), a global sensitivity analysis index [50], to identify significant correlations between single-granuloma ODE model parameter changes and variation in whole lung outputs. We excluded the dissemination parameters from our multi-scale PRCC analysis because they are phenomenological in nature and we are interested in identifying the mechanistic events that occur at the granuloma scale and lead to dissemination, a whole lung outcome.

Linking cellular scale and tissue time scales

We link the cell and cytokine scale events in the ODE model (single granuloma) with the tissue scale ABM (multiple granulomas) to form the multi-scale *MultiGran* model (Fig 2). Linking of timescales is important for proper model design [51]. We use an ABM time-step of 1 day. At each ABM time-step, dissemination events can occur. After each ABM time step, the system of ODEs is solved for each granuloma to update the states of all host cells, cytokines and Mtb populations over the next 24 hours. We run the ODEs using adaptive time steps for 1 agent iteration, for each granuloma, before proceeding to the next agent time step, as dissemination events at the agent time step depend on the dynamically-changing state of ODEs. Additionally,
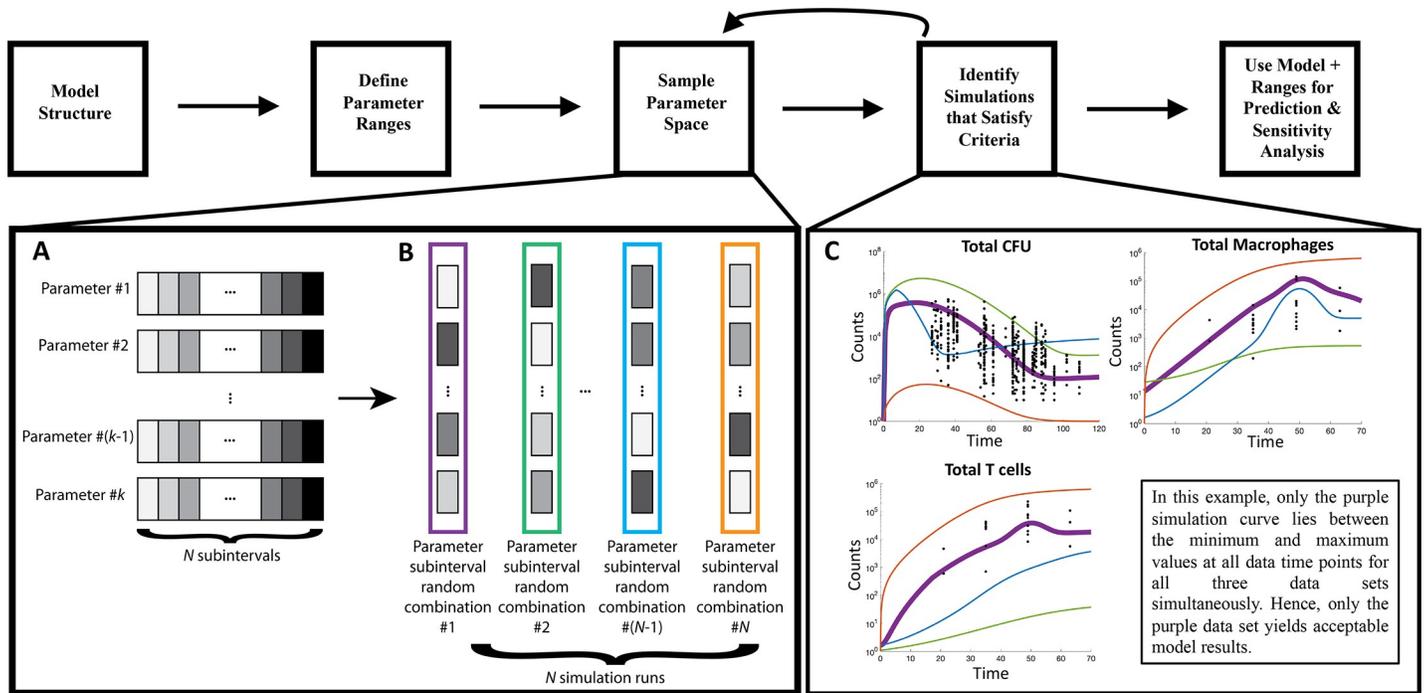
Fig 3. MultiGran calibration process. The model structure for the single granuloma ODE model and the broad parameter ranges (S1 Table) were based on previously published models and experimental data [14,16,19,22,37–40,49]. We sample from the initial parameter space using a Latin Hypercube Sampling (LHS) process. LHS divides each of the *k* varied parameters into *N* subintervals **(A)**, randomly selects a value from each subinterval, and randomly groups *k* values (one for each parameter) to create *N* parameter sets **(B)**. We next identify simulations that satisfy all our criteria **(C)**. Here, only the purple simulation curve (parameter set #1) has satisfied the criteria of passing within the minimum and maximum value at each timepoint for each of our three datasets simultaneously. Note that in a single LHS run, many simulations may satisfy all criteria, but for simplicity we have shown only a single one (purple curve). By inspecting values from parameter sets that pass all criteria, we are able to narrow each of the *k* broader parameter ranges. We iterate performing an LHS simulation followed by narrowing the *k* parameter ranges based on passing all of the criteria until 90% of the 500 simulations fall within the bounds of our experimental data on CFU, T cell counts, and macrophages within individual NHP granulomas.

https://doi.org/10.1371/journal.pcbi.1007280.g003

the ODE state variable concentrations can be affected by the occurrence of a dissemination event.

Results

We present a whole lung model, *MultiGran*, that captures the behavior of Mtb infection leading to the development of multiple granulomas via initial infection and then dissemination of bacteria from existing granulomas. We calibrate and validate *MultiGran* with unique datasets derived from NHPs, the animal model that most closely mimics the features of human infection. We then use the model to identify dissemination rates and to predict mechanisms leading to dissemination within the calibrated parameter space.

Simulated individual granulomas recapitulate in vivo primate granuloma dynamics

We calibrated our single-granuloma model, comprised of a system of non-linear ODEs, to data derived from NHP studies. We compared bacterial load (CFU), T cell counts, and macrophage counts over time per granuloma. Our CFU dataset consists of 623 granulomas from 38 NHPs (previously published [20,27,31–33] and ongoing studies). T cell and macrophage counts, as well as additional CFU, were derived from a separate, new dataset of 26 granulomas from 7 Mtb-infected NHPs and baseline data from one uninfected macaque (see Methods).
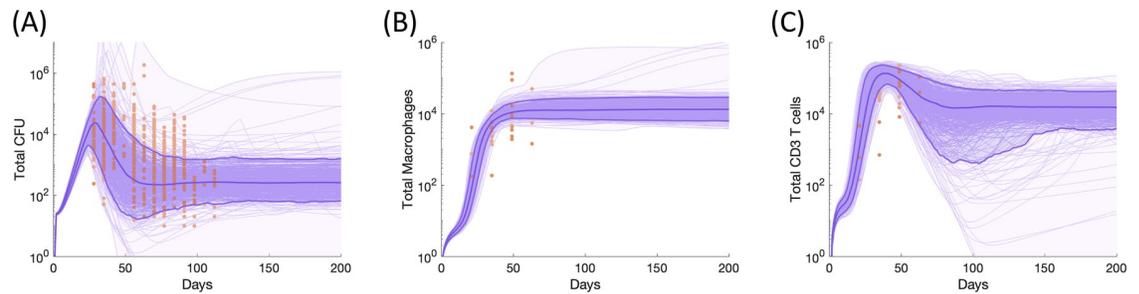
Fig 4. Bacteria, macrophage and T cell dynamics within an individual granuloma. Individual NHP granuloma bacteria **(A)**, macrophages **(B)**, and CD3+ T cells **(C)** shown as orange points across time. Each individual point represents data from a single NHP granuloma. Purple lines indicate simulation outputs from 500 simulations that match NHP data. Light purple shading shows the minimum and maximum of simulation runs, darker purple shading represents the 5th to 95th percentiles of the simulations, and dark purple lines represent the 5th, 50th, and 95th percentiles of simulations. Parameter ranges are listed in S1 Table.

https://doi.org/10.1371/journal.pcbi.1007280.g004

The data from these 7 NHPs capture the timing of the immune system during early events in infection (granulomas from all NHPs were collected between 3–9 weeks post infection) and were imperative for proper calibration of the model.

The single-granuloma model contains many parameters to represent various mechanisms that are part of innate and adaptive immune responses to Mtb infection, but we include few assumptions about the relative roles of each of these mechanisms; rather, we define equations and allow the relative influence of each mechanism emerge during calibration. We identify ranges of parameter values (S1 Table) that replicate CFU peaks at approximately 35 days and subsequent control of CFU after day 100 post-infection (Fig 4A), macrophage dynamics (Fig 4B), and T-cell dynamics (Fig 4C). These dynamics reflect the initial inability of the innate immune system to control Mtb replication, the eventual control provided by T cells that arrive from the lymph node around day 28, and the stabilization of Mtb counts around day 100. When isolating a suitable parameter range, we identified ranges that matched these overall trends and recapitulated the spread of granuloma outcomes outlined by the NHP datasets. Likely, our spread captures a fuller range of individual granuloma dynamics than a sample from a limited number of NHP can achieve.

MultiGran simulates the appearance of granulomas throughout the lung, as seen in vivo

By employing the calibrated single-granuloma model (Fig 4) within our *MultiGran* framework, we can now simulate the spread of infection within the lung. We inoculate with 16 to 21 individual bacteria, mimicking the protocol of Martin et al. [25], placing them within an inoculation region within one of the lower lung lobes, as is done in the NHP inoculations via bronchoscope (see Methods). Each initial granuloma in an NHP arises from a single bacterium in an inoculation event [25]. Therefore, we initially establish 16–21 granulomas. A sample simulation at the time-point of 250 days post-infection is shown in Fig 5. The blue lung mesh represents the dataset derived from NHPs for (x,y,z) coordinates of a lung. Placed on this mesh are simulation results–individual granulomas ("agents" in the model) and their location, size, and bacterial origin (barcode). Note that, as in the NHP images of Fig 1, infection is primarily within the inoculation region–but that 7 granulomas disseminated non-locally to the opposite lung. In this simulation, one granuloma cluster was found that contained more than one Mtb barcode, as is shown in the pie chart. Movies of disease progression using this 3D visualization are available on the website http://malthus.micro.med.umich.edu/lab/movies/MultiGran/.
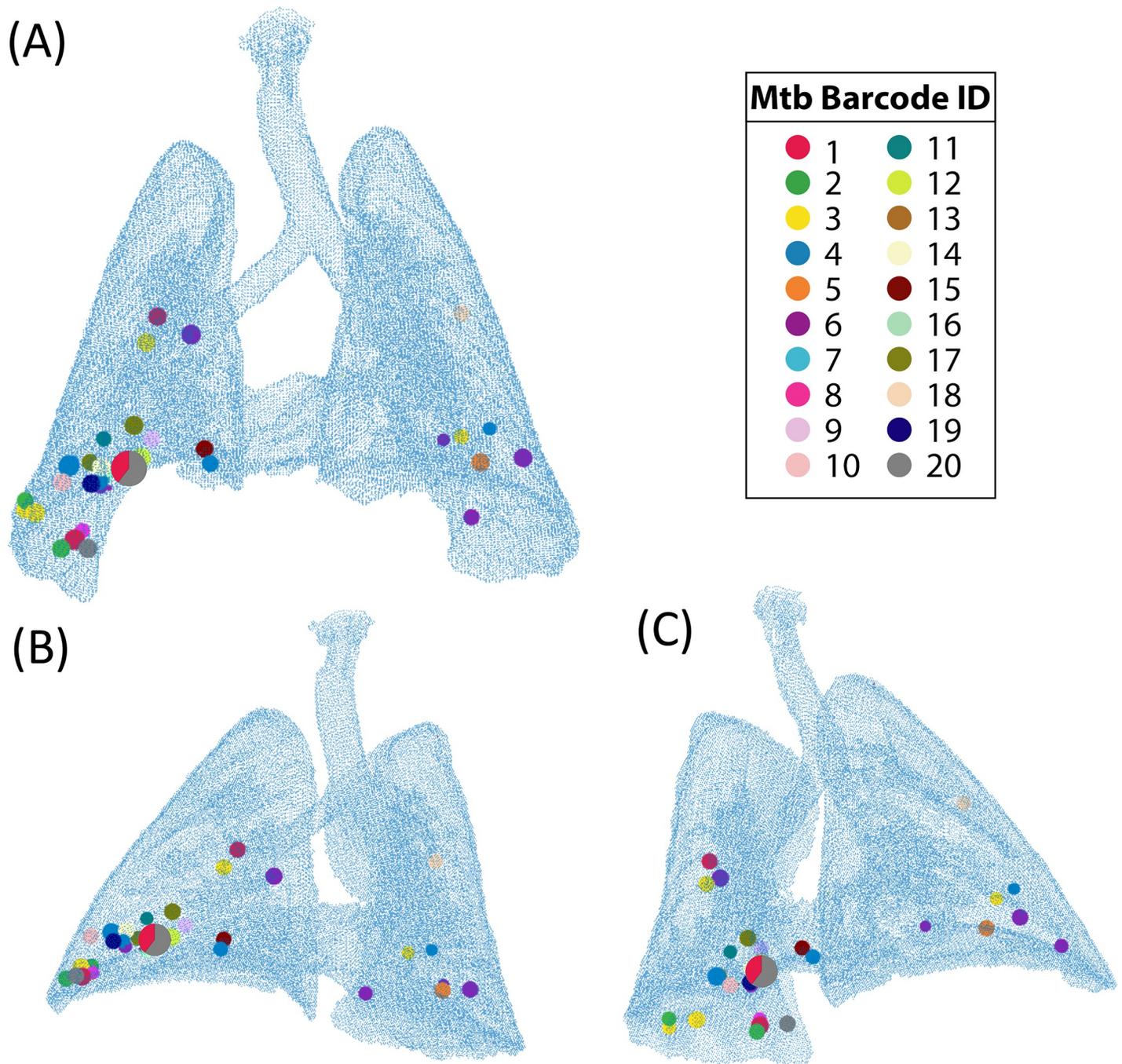
Fig 5. *MultiGran in silico* infection in a non-human primate lung. A single *in silico* simulation at 250 days post infection from three angles (**A**-anterior view, **B&C**-opposite posterior-lateral views), plotted over a data grid taken from PET/CT images of a single NHP. Granulomas are located within the lung in 3D space. Each circle of a single color represents a granuloma or granuloma cluster with a single Mtb barcode ID. The circle shown as a pie chart represents a granuloma cluster with two unique Mtb barcode IDs; each color represents the relative proportion of CFU of each ID compared to the total CFU of the granuloma cluster, while the overall size of the circle is proportional to the size of the cluster. Inoculation was in the lower right lung (bottom left in each image). Granulomas found in the upper right lung and the left lung result from non-local dissemination within the simulation.
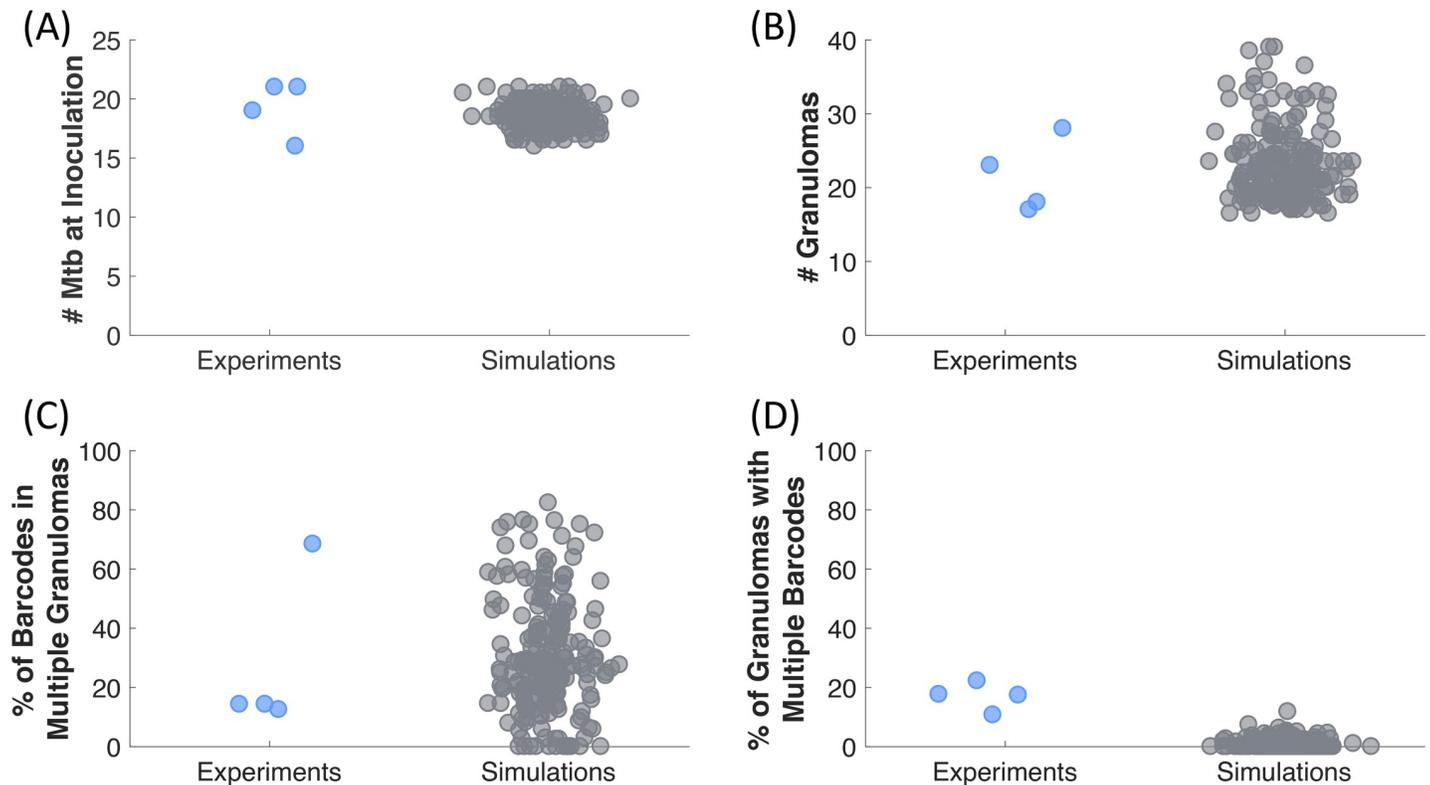
https://doi.org/10.1371/journal.pcbi.1007280.g005

Fig 6. *MultiGran* **recapitulates non-human primate dissemination outcomes.** Martin et al. [25] infected 4 NHP with 16–21 different Mtb barcodes (**A**), and after 120 days the NHP immune system formed 16–28 non-sterilized granuloma clusters (**B**). We replicated these experiments by simulating 200 NHP, which started with 16–21 different Mtb. Of the 16–21 Mtb in NHP, 10%-70% were found in multiple granuloma clusters, meaning at least 10%-70% of Mtb were disseminating. Similar to the NHP data, our simulations have 0%-90% of Mtb barcodes disseminated to multiple granuloma clusters (**C**). Within the NHP experiments, of the 16–28 non-sterilized granuloma clusters, 10%-25% had multiple Mtb IDs within them, meaning at least 10%-25% of observed granulomas are clusters involving multiple sources of Mtb infection. Our 200 *MultiGran* simulations demonstrate a similar range of granuloma clusters with multiple Mtb barcodes (**D**). Simulations are shown in gray whereas NHP experiment outcomes are shown in blue. Each point represents a single NHP or *in silico* simulated granuloma.

https://doi.org/10.1371/journal.pcbi.1007280.g006

Simulations are consistent with in vivo infection and predict dissemination likelihood rates

MultiGran allows both local and non-local dissemination of bacteria to initiate new granulomas, tracks the origin (Mtb ID) of each granuloma, and allows for merging of nearby granulomas to form a cluster. Each granuloma has a unique parameter set chosen from the ranges in S1 Table according to an LHS design. To determine what leads to different dissemination patterns *in vivo*, we use our dataset consisting of four NHPs in that were inoculated with uniquely identifiable Mtb (Fig 1; Martin et al. [25]). Outcome measures from these experiments include: (1) the number of Mtb at time of inoculation (16–21 Mtb), (2) the number of granuloma (or granuloma clusters) at necropsy (17–28 granulomas), (3) the percentage of Mtb barcodes found in multiple granulomas (12.5–68.4%), and (4) the percentage of granulomas containing multiple Mtb barcodes (~10–20%). We calibrated *MultiGran* dissemination dynamics to this dataset by varying the seven dissemination parameters (S3 Table). Our whole lung simulations and the NHP dataset are shown in Fig 6. Notice that the simulations capture the full heterogeneity of the *in vivo* results across each NHP. Additionally, the experimental data are from only four NHPs, while our simulations represent a larger, more diverse set of possible outcomes.

To more directly test for non-local dissemination events, we validate our simulations against a second dataset of 38 NHPs (Fig 7). Within this NHP dataset, we identified the lung
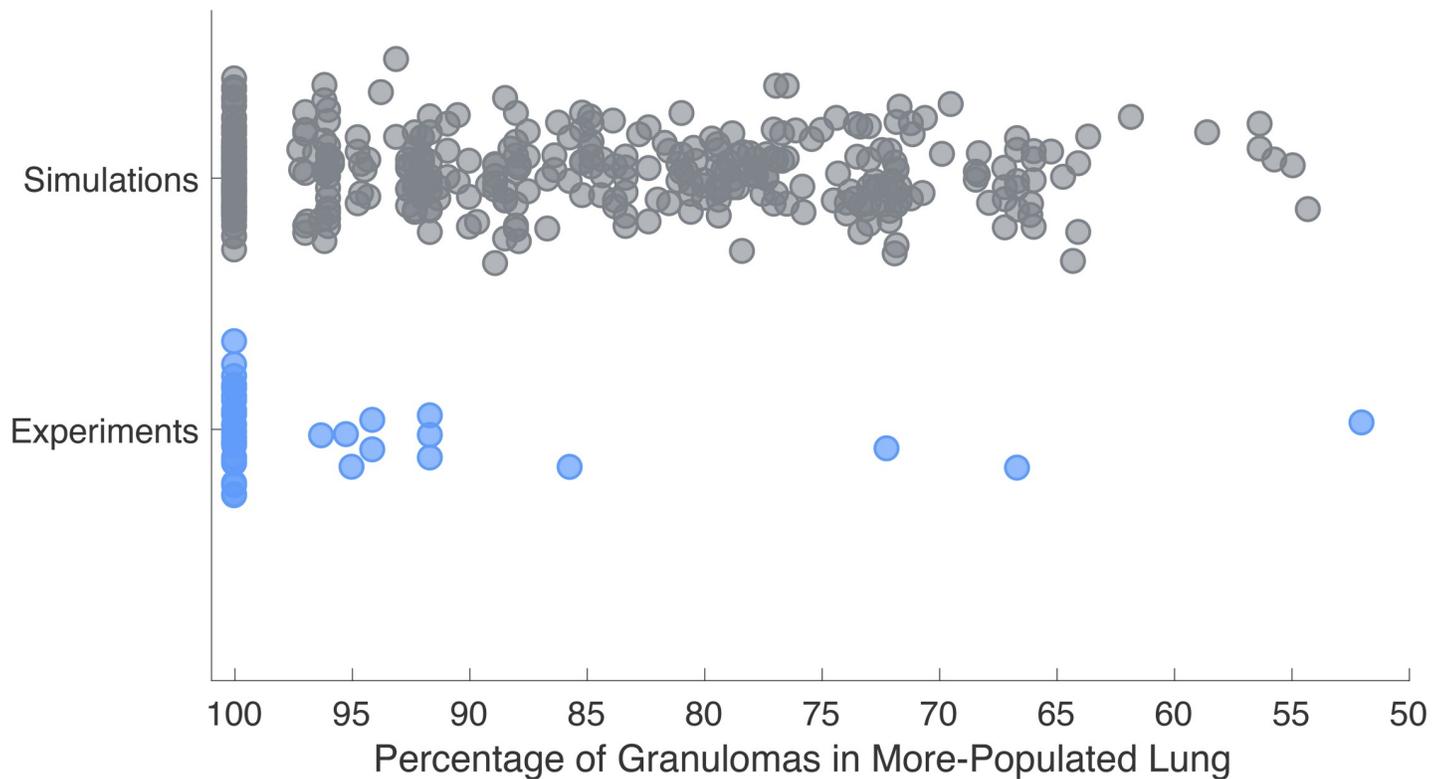
Fig 7. *MultiGran* recapitulates spread of infection data. At necropsy of 38 NHP experiments, we identified the lung that contained the most granulomas for each NHP. Next, we calculated the percentage of granulomas that resided in the more-populated lung out of the total number of granulomas. We found that 52–100% of granulomas formed resided within the more-populated lung. Blue dots represent each NHP experiment. We ran 200 *in silico* simulations that capture a similar range to the NHP spread of infection from lung to lung, ranging from 54.3% to 100%. Gray dots represent each simulated lung.

that contained the most granulomas for each NHP, and termed this lung the more-populated lung. Next, we calculated the percentage of granulomas that resided in the more-populated lung out of the total number of granulomas across both lungs. We found that the 38 NHPs exhibited a range of 52%-100% of granulomas in the lung that was more-populated. Results from the same simulations used to create Fig 6 give a range ~54%-100%, providing additional support for the model in its ability to capture the range of data offered by NHP experiments. As this result is a validation step of our model, if *MultiGran* could not exhibit a similar range of granulomas across both lungs, the employed approach would have invalidated the model.

When examining *in vivo* data, the total number of dissemination events may be under-counted due to sterilization and granuloma clustering. In contrast, our model is able to count every dissemination event, and thereby provides a predicted frequency of local and non-local dissemination. We found that, on average, the rate of dissemination is about 1/24 dissemination events per granuloma per month for simulations run out to 250 days. Most dissemination occurs earlier in the infection, as noted in Martin, et al. [25]. Further, *MultiGran* predicts that local dissemination events occur about twice as frequently as non-local dissemination events.

MultiGran simulations match individual NHP infections

From our repository of 200 *MultiGran* simulated lungs, we isolated the five simulations that yielded the closest match to the median values of Mtb inoculation (20), the median number of granulomas at necropsy (20.5), the median percentage of Mtb barcodes that were found in
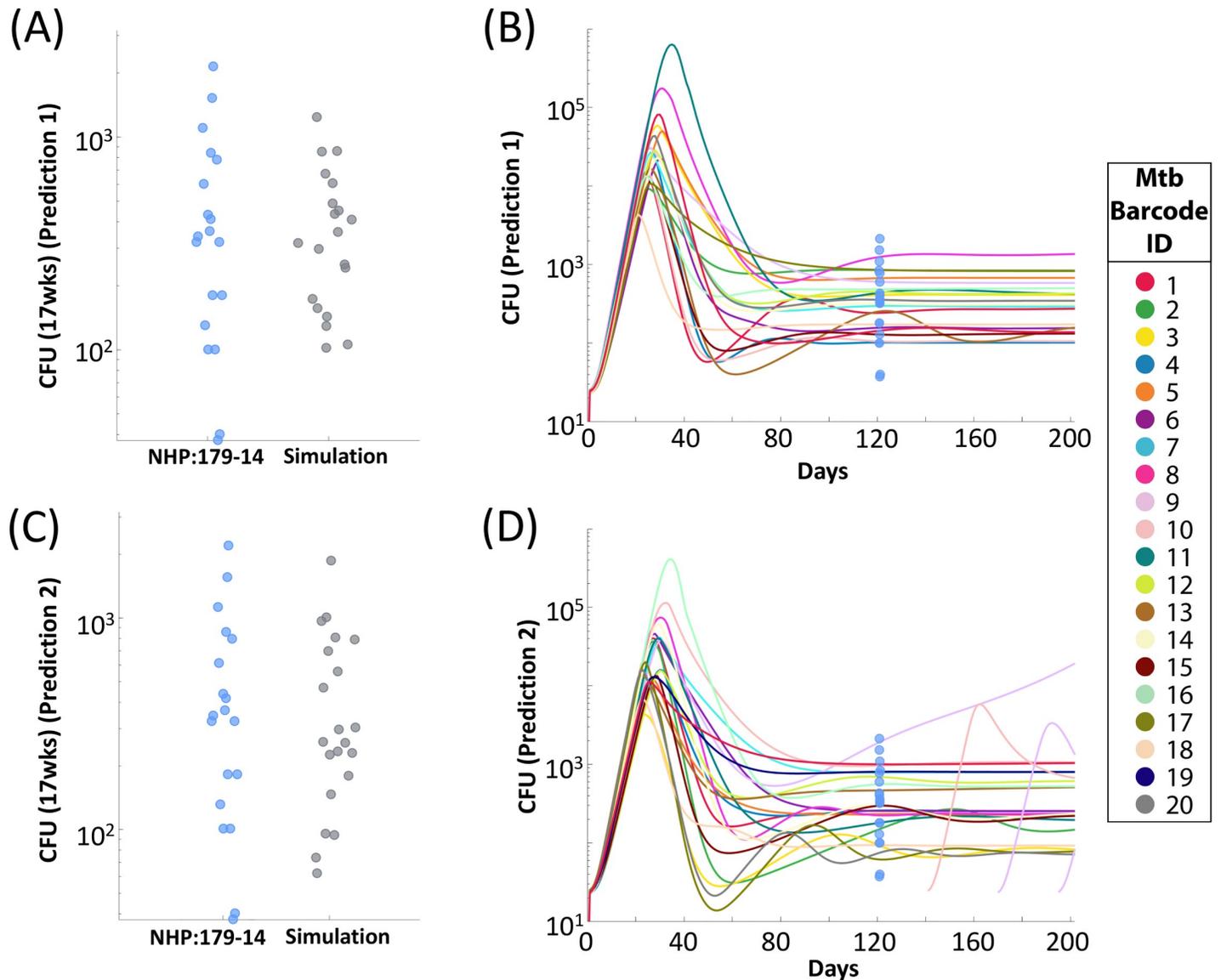
Fig 8. *MultiGran* **matches individual NHP granuloma dynamics and predicts CFU burden across time.** We compared the CFU/granuloma at necropsy for NHP:179–14 (**A&C**) to two separate simulations that matched these outcomes. Blue dots represent single-granuloma values taken from NHP:179–14; gray dots represent simulation values at comparable timepoints. Simulation predictions diverged after 17 weeks. One simulation predicted stability–i.e., granuloma containment of bacteria (**B**). The other simulation (**D**) predicted uncontrolled growth of bacteria within one granuloma, leading to dissemination and the formation of other granulomas across time. Each line in (**B&D**) represents one granuloma realization within *MultiGran* across time. Blue dots represent NHP:179 granuloma CFU values. Simulation behavior to the right of the blue dots should be considered a prediction.

multiple granulomas (14.3%), and the median percentage of granulomas that contained multiple Mtb barcodes (17.5%) across the four NHP from Martin et al. [25].

These five simulations represent the best matches to the NHP used in Martin et al. [25]. We compare two of these simulations to the CFU/granuloma at necropsy from NHP:179–14 (Fig 8A & 8C). Both lung simulations display satisfactory matches to the NHP CFU data; both simulations cover the spread of the experimental data while lying within the bounds of the dataset. However, while both simulations match the CFU data at 17 weeks, we are able to predict what could have happened beyond the necropsy date by running the simulation for a longer time

period. Shown are two distinct possible outcomes with the same parameter set: note they diverge when predicting later dissemination events. Fig 8B shows one simulation predicts bacterial control across all the granulomas within that simulation. Fig 8D shows another outcome. Here, a single granuloma within the lung exhibits uncontrolled bacterial growth leading to dissemination and there is also formation of new granulomas via both local and non-local dissemination (at days 145, 166, and 193). These simulations suggest that NHP:179–14 was either containing the bacteria (i.e., LTBI) (our prediction in Fig 8B) or could have had a subclinical infection that was on the edge of leading to multiple dissemination events (our prediction in Fig 8D). Simulations that match the other three NHP are shown on our website, and reveal similar trends and predictions. Additionally, the entire repository of 200 virtual NHP CFU trends are displayed at: http://malthus.micro.med.umich.edu/lab/movies/MultiGran/.

Sensitivity analysis reveals important mechanisms responsible for dissemination

To predict the mechanisms that lead to dissemination events within lungs, we perform global sensitivity analysis using our repository of 200 *MultiGran* simulated lungs on four whole-lung outcomes of interest: the number of dissemination events, the total number of granuloma clusters at the end of the simulation, the percentage of granuloma clusters that contain multiple barcodes, and the percentage of granulomas that occupy the initially-inoculated lung at the end of the simulation. We quantify the contributions of each model parameters to the outcomes of interest by calculating partial rank correlation coefficients (PRCC) at the end of the simulation (250 days). This analysis provides information about the contribution of each model parameter within our parameter space that was calibrated to multiple NHP datasets across cellular, granuloma, and whole-lung datasets. Our analysis reveals one parameter as the main driver of these four whole lung outcomes (Table 2). Parameter *CD8MultiFunc* describes the multi-functional nature of CD8+T cells, i.e., the amount of overlap of cytotoxic function and cytokine expression in CD8+ T cells. While the single-granuloma model was designed to be agnostic prior to calibration about the primary drivers behind bacterial killing and Mtb containment, we found that *CD8MultiFunc* is significantly correlated with each of the four outcomes. If *CD8MultiFunc* is increased so that a greater proportion of CD8+ T cells exhibits multi-functionality, then a larger percentage of granulomas will reside within a single lung (less non-local dissemination) and there will be fewer dissemination events and fewer granulomas overall. CD8+ T cells are a key host cell in a functional immune response to Mtb infection, and if the subpopulation that can perform multiple roles within the complex microenvironment of a granuloma increased, it would certainly benefit the host.

If we exclude parameter *CD8MultiFun* from the analysis, we reveal secondary contributions of other parameters to the whole lung outcomes (Table 2). These 12 parameters represent both adaptive and innate immune dynamics: we found both the adaptive and innate immune system were associated with non-local and local dissemination events. Notably, the role of macrophage-bacteria interactions is found to be important. *k18* represents the base rate of killing of extracellular bacteria by macrophages. If this rate is high, there are fewer dissemination events and fewer granulomas across the simulation. Additionally, *k17* represents the maximum bursting rate of infected macrophages. This parameter is positively correlated with the number of dissemination events and the number of granulomas across a simulation. If bursting occurs at a high rate within a granuloma, our model predicts that a granuloma is more likely to disseminate both locally and non-locally. Taken together, these two parameters identify an important role for macrophage dynamics within the granuloma: if macrophages cannot adequately respond to Mtb, the likelihood of dissemination increases. Altogether, the results of this

Table 2. Sensitivity Analysis reveals global drivers of dissemination outcomes. PRCCs are shown for each parameter, showing a significant impact on each of the 4 *MultiGran* whole lung simulation outcomes at the end of the simulation. **Excluding parameter *CD8MultiFunc*, the overlap of cytotoxic function and cytokine expression in CD8+ T cells, 12 other parameters were also identified as having a significant impact on each of 4 *MultiGran* whole lung simulation outcomes at the end of the simulation. All PRCCs shown are significant to $p < .05$.

Parameter Name	Parameter Description	Number of Dissemination Events	Number of Granulomas and Clusters	Percentage of Granulomas with Multiple Barcodes	Percentage of Granulomas in More-Populated Lung
CD8MultiFunc **	overlap of cytotoxic function and cytokine expression in CD8+ T cells	-0.39	-0.38	-0.14	0.32
k18	Extracellular bacteria killed by macrophages	-0.11	-0.11	-0.041	0.11
nuI10	decay rate of IL-10 cytokine	-0.088	-0.087	-0.068	0.089
Sr1b	TNF based recruitment of primed CD4+ T cells	-0.075	-0.074	-0.044	0.06
k6	rate of differentiation from primed to Th1 CD4+ T cells	-0.084	-0.073	-0.047	0.071
s12	cell production of IL-12	-0.058	-0.056	-0.025	0.056
w	contribution of intracellular bacteria to resting macrophage activation	-0.037	-0.04	-0.021	0.04
s2	half-saturation of IL-4	-0.024	-0.021	-0.025	0.02
Sr3b	TNF based recruitment of Th2 CD4 + T cells	-0.036	-0.033	-0.021	0.025
alpha30	TNF production by infected macrophages	0.032	0.028	0.022	-0.037
nuTg	IFNg induced apoptosis of Th1 CD4 + T cells	0.057	0.055	0.037	-0.04
s4b	half-saturation of TNF on local resting macrophage recruitment	0.042	0.043	0.04	-0.043
k17	max rate of infected macrophages bursting	0.14	0.14	0.076	-0.12

analysis represent a multi-scale impact: events governing cell function at the cellular scale impact local and non-local dissemination outcomes across the lungs and predict the difference between dissemination and control across the lung environment.

Discussion

Tuberculosis is a complex and heterogeneous disease with a spectrum of outcomes, and the myriad of mechanisms that influence outcomes of initial infection are poorly defined. Our data in NHP models, and bolstered by data in humans, support the notion that each individual granuloma in a host is independent and dynamic, in terms of immunologic composition and function, ability to kill or restrain Mtb bacilli, and risk for dissemination or reactivation [52,53]. However, it can be challenging in NHP models to determine the full range of host mechanisms that play a role in initial containment and prevention of dissemination, both of which are essential to limiting development of active TB. In the pursuit of a better understanding of the collective behavior of lung granulomas in individuals infected with Mtb, we performed a systems biology approach pairing NHP experiments and computational/ mathematical modeling. Specifically, we explored events that lead to dissemination and new granuloma formation, and several studies have recently explored this biological phenomenon [25,36,54,55]. In particular, the barcoding technique introduced by Martin et al. showed that dissemination varies widely among macaques despite initial infection conditions being similar, and that in individual macaques, each granuloma had a different dissemination risk, from no

dissemination by most granulomas, even though these granulomas were CFU+, to multiple dissemination events from a single granuloma. The barcoding analysis provided critical new information about bacterial spread within the lung. However, identifying mechanisms that leading to granuloma dissemination, which is linked to development of active TB [36], is important in designing more effective vaccines and therapeutics against TB. Systems biology approaches can address these mechanisms and more generally contribute to our still limited understanding of Mtb infection dynamics.

In this work, we combine experimental data from NHPs with a novel multi-scale, hybrid agent-based model of granuloma formation, function and dissemination within the lung, called *MultiGran*. We calibrate and validate *MultiGran* against multiple distinct NHP datasets that span cellular, bacterial, granuloma, and whole-lung scales. This calibration and validation allowed us to make predictions about dissemination within Mtb infected lungs. We report that the likelihood of local dissemination is approximately two times greater than non-local dissemination, which supports the in vivo data reported in Martin, et al. [25], and we used sensitivity analysis techniques to identify that dissemination is intertwined with the role of CD8+ T cells in granulomas. Specifically, we predict that the functionality of CD8+ T cells is critically important: if a greater percentage of CD8+ T cells can perform dual functions of cytokine expression (IFNγ, TNF, and IL-10) and cytotoxicity, then the likelihood of dissemination significantly decreases.

The role of CD8+ T cell multi-functionality within the granuloma is controversial (for reviews of CD8+ T cells in TB, see [38,56]. While the majority of T cells within a granuloma are single cytokine producers [27], multifunctional CD8+ T cells have been demonstrated in the blood of Mtb-infected humans and the proliferation and response rate of these cells differed between active and latent infection [57,58]. Together, these studies and our current work suggest a need for increased focus on this specific cell type to evaluate the potential that CD8 + multifunctional T cells may offer.

The NHP datasets generated within this study are unique and critical to the predictions of *MultiGran*. In addition, these data also present new insights into early events occurring during Mtb infection. In particular, the ability to capture data on Mtb infection during early time points for CFU, T cell counts, and macrophage numbers is instrumental in elaborating timing of early immune response events. These early events in primates have been understudied, and knowledge of the role that timing plays in granuloma establishment, formation, and development is critical to early intervention strategies.

Using *MultiGran*, we were able to match to granuloma population data coming from multiple monkeys (Figs 6 & 7) and granulomas (Fig 4). we were also able to match experimental data from a single NHP (Fig 8). In the era of precision medicine [59], the ability of *MultiGran* to match to individual data could help predict, in real time, whether the granulomas within that individual are likely to disseminate. This could happen when paired with PET/CT images of individually lung granulomas. However, more realistically, this provides an impetus for identifying biomarkers that are associated with granulomas at risk of dissemination, which could be more widely used to identify persons at risk of developing active TB following infection.

There are a few limitations of our study and model. First, the driving dissemination probability rules are somewhat phenomenological. Our goal in this first study was to rely on as few assumptions as possible; the only granuloma characteristic that is explicitly used in the dissemination rules is the total bacterial burden. As a consequence, the model allows for even a stable, mature granuloma to disseminate (with small probability). We addressed this by allowing T cells to leave the parent granuloma to travel to a daughter granuloma in a local dissemination event, expecting this to sterilize new granulomas. Surprisingly, this was largely ineffective.

Instead, it is more likely that the lung parenchyma in infected individuals has increased numbers of Mtb-specific T cells and possibly activated macrophages, so that new granulomas form in a completely different immune environment, compared to the initial granulomas that form in an immunologically naïve environment. This notion is supported by our data in NHP models demonstrating that primary ongoing infection protects against reinfection [32]. *MultiGran* could be refined to test this in future iterations. Second, we restrict dissemination to be within the boundary of the lungs, but the actual environment within the lungs is very complicated and also could include airways and blood. Third, while we acknowledge thoracic lymph nodes as a source of non-local dissemination, and include adaptive immune cell recruitment in our ODE model, we currently do not explicitly model lymph node compartments. In future work, we plan to address the role of lymph nodes in Mtb infection and dissemination. Finally, while *MultiGran* was developed based on extensive NHP and human data, it does not contain all the various cell types and mechanisms in the complex environment of the granuloma, primarily because the functions and importance of certain cell types and factors remain obscure. As data become available, *MultiGran* can evolve to include additional factors for mechanistic test. Potentially, other modeling techniques might be better suited to answer specific questions surrounding granuloma dissemination. For example, if one was interested in exactly the path an infected macrophage might take when escaping the granuloma and initiating local dissemination, a spatial model at the site of infection, such as *GranSim* [60], could be more appropriate. However, such a modeling approach necessitates a loss of total lung-level dynamics that were critical to address our questions about dissemination events (both local and non-local) across NHP lungs.

In summary, we utilized a systems biology approach that combined computational modeling and NHP datasets to better understand mechanisms of granuloma dissemination. We present *MultiGran*, the first multi-scale model of granuloma dissemination and formation, that was calibrated and validated to NHP data and we make predictions about the rate of dissemination and the role of specific immune cells in granuloma dissemination. In particular, we discovered roles for multifunctional CD8+ T cells and macrophage dynamics in preventing local and non-local dissemination within the lungs. Altogether, we argue that *MultiGran*, together with NHP experimental approaches, offers great potential to understand and predict dissemination events within Mtb infected lungs.

Supporting information

S1 Equations. The ordinary differential equations that describe the cellular, cytokine, and bacterial dynamics within a developing granuloma.
(PDF)

S1 Table. ODE model parameters that govern individual granuloma formation and growth across time.
(DOCX)

S2 Table. Other parameters that represent the size of granulomas and are necessary for runtime execution.
(DOCX)

S3 Table. Parameters that govern dissemination in *Multigran*.
(DOCX)

S1 Fig. The rules of granuloma establishment and dissemination within *MultiGran*. Case 1 –inoculation. Inoculation deposits bacteria in a specific lung region at position *(xTrial, yTrail,*

zTrial). The black box designates inoculation region (row 1), wherein the specific within-lung region destined for inoculation is highlighted in green (row 2). The third row demonstrates successful inoculation of a single bacterium–the black box was sampled randomly until the sampled coordinates lie within the green region. Cases 2 and 3 define granuloma placement following dissemination. **Case2 –non-local dissemination**. When non-local dissemination occurs, a bacterium escapes a single granuloma (row 1) and can be placed in any region (shown in black in row 2) that encompasses the entire lung. The green highlighted region is the area in which the bacterial placement will be accepted. Row 3 shows three trial placements: two realizations of accepted bacterial placement (black arrows) and one unaccepted placement (red arrow) at *(xTrial*, *yTrial*, *zTrial)*. **Case 3– local dissemination**. Local dissemination is the only form of granuloma placement which does not utilize random placement within a region of lung space. Rather, an infected macrophage from the parent granuloma is placed in a random direction away from the parent granuloma. Row 2 shows several options for granuloma infected macrophage placement. Note that the arrows are of different length to represent our assumption that local dissemination likely follows a normal distribution with respect to parent granuloma location. Here, the green and black arrows show valid directions for the new placement for the infected macrophage, while red arrows show invalid directions. A new granuloma will begin to develop in the chosen (green) valid location (Row 3). Note that in both **(A)** and **(B)** bacteria, granulomas, and infected macrophages are not to scale. Lung image from Servier Medical Art.
(TIF)

Acknowledgments

Author Contributions

Conceptualization: Timothy Wessler, Louis R. Joslyn, JoAnne L. Flynn, Denise E. Kirschner, Jennifer J. Linderman.

Data curation: H. Jacob Borish, Hannah P. Gideon.

Formal analysis: Timothy Wessler, Louis R. Joslyn, Hannah P. Gideon.

Funding acquisition: JoAnne L. Flynn, Denise E. Kirschner, Jennifer J. Linderman.

Investigation: Timothy Wessler, Louis R. Joslyn, Denise E. Kirschner, Jennifer J. Linderman.

Methodology: Timothy Wessler, Louis R. Joslyn, H. Jacob Borish, Hannah P. Gideon.

Project administration: JoAnne L. Flynn, Denise E. Kirschner, Jennifer J. Linderman.

Resources: JoAnne L. Flynn, Denise E. Kirschner, Jennifer J. Linderman.

Software: Timothy Wessler, Louis R. Joslyn.

Supervision: JoAnne L. Flynn, Denise E. Kirschner, Jennifer J. Linderman.

Validation: Timothy Wessler, Louis R. Joslyn, Hannah P. Gideon.

Visualization: Timothy Wessler, Louis R. Joslyn, H. Jacob Borish.

Writing – original draft: Timothy Wessler, Louis R. Joslyn, H. Jacob Borish, Hannah P. Gideon.

Writing – review & editing: Timothy Wessler, Louis R. Joslyn, H. Jacob Borish, Hannah P. Gideon, JoAnne L. Flynn, Denise E. Kirschner, Jennifer J. Linderman.

References

1. WHO. WHO Global tuberculosis report 2016. World Heal Organ Press. 2016. ISBN 978 92 4 156539 4

2. Houben RMGJ, Esmail H, Emery JC, Joslyn LR, McQuaid CF, Menzies NA, et al. Spotting the old foe—revisiting the case definition for TB. Lancet Respir Med. 2019; 7: 199–201. https://doi.org/10.1016/S2213-2600(19)30038-4 PMID: 30823971

3. Lin PL, Ford CB, Coleman MT, Myers AJ, Gawande R, Ioerger T, et al. Sterilization of granulomas is common in active and latent tuberculosis despite within-host variability in bacterial killing. Nat Med. 2014; 20: 75–79. https://doi.org/10.1038/nm.3412 PMID: 24336248

4. Gideon HP, Flynn JL. Latent tuberculosis: What the host "sees"? Immunol Res. 2011; 50: 202–212. https://doi.org/10.1007/s12026-011-8229-7 PMID: 21717066

5. Lin PL, Myers A, Smith L, Bigbee C, Bigbee M, Fuhrman C, 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 Rheum. 2010; 62: 340–350. https://doi.org/10.1002/art.27271 PMID: 20112395

6. Diedrich CR, Mattila JT, Klein E, Janssen C, Phuah J, Sturgeon TJ, 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. 2010; 5. https://doi.org/10.1371/journal.pone.0009611 PMID: 20224771

7. Mattila JT, Diedrich CR, Lin PL, Phuah J, Flynn JL. Simian Immunodeficiency Virus-Induced Changes in T Cell Cytokine Responses in Cynomolgus Macaques with Latent Mycobacterium tuberculosis Infection Are Associated with Timing of Reactivation. J Immunol. 2011; 186: 3527–3537. https://doi.org/10.4049/jimmunol.1003773 PMID: 21317393

8. Chen CY, Huang D, Wang RC, Shen L, Zeng G, Yao S, et al. A critical role for CD8 T cells in a nonhuman primate model of tuberculosis. PLoS Pathog. 2009; 5. https://doi.org/10.1371/journal.ppat.1000392 PMID: 19381260

9. Sershen CL, Plimpton SJ, May EE. Oxygen Modulates the Effectiveness of Granuloma Mediated Host Response to Mycobacterium tuberculosis: A Multiscale Computational Biology Approach. Front Cell Infect Microbiol. 2016. https://doi.org/10.3389/fcimb.2016.00006 PMID: 26913242

10. Gough M, May E. An in silico model of the effects of Vitamin D3 on mycobacterium infected macrophage. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS. 2016. https://doi.org/10.1109/EMBC.2016.7590980

11. Linderman JJ, Kirschner DE. In silico models of M. Tuberculosis infection provide a route to new therapies. Drug Discov Today Dis Model. 2015; 15: 37–41. https://doi.org/10.1016/j.ddmod.2014.02.006 PMID: 26904139

12. Kirschner D, Pienaar E, Marino S, Linderman JJ. A review of computational and mathematical modeling contributions to our understanding of Mycobacterium tuberculosis within-host infection and treatment. Curr Opin Syst Biol. 2017; 3: 170–185. https://doi.org/10.1016/j.coisb.2017.05.014 PMID: 30714019

13. Hao W, Schlesinger LS, Friedman A. Modeling granulomas in response to infection in the lung. PLoS One. 2016; 11. https://doi.org/10.1371/journal.pone.0148738 PMID: 26986986

14. Warsinske HC, Pienaar E, Linderman JJ, Mattila JT, Kirschner DE. Deletion of TGF-β1 increases bacterial clearance by cytotoxic t cells in a tuberculosis granuloma model. Front Immunol. 2017; 8. https://doi.org/10.3389/fimmu.2017.01843 PMID: 29326718

15. Linderman JJ, Cilfone NA, Pienaar E, Gong C, Kirschner DE. A multi-scale approach to designing therapeutics for tuberculosis. Integr Biol (United Kingdom). 2015; 7: 591–609. https://doi.org/10.1039/c4ib00295d PMID: 25924949

16. Cilfone NA, Perry CR, Kirschner DE, Linderman JJ. Multi-Scale Modeling Predicts a Balance of Tumor Necrosis Factor-α and Interleukin-10 Controls the Granuloma Environment during Mycobacterium tuberculosis Infection. PLoS One. 2013; 8. https://doi.org/10.1371/journal.pone.0068680 PMID: 23869227

17. Fallahi-Sichani M, Kirschner DE, Linderman JJ. NF-??B signaling dynamics play a key role in infection control in tuberculosis. Front Physiol. 2012; 3 JUN. https://doi.org/10.3389/fphys.2012.00170 PMID: 22685435

18. Prats C, Vilaplana C, Valls J, Marzo E, Cardona PJ, López D. Local inflammation, dissemination and coalescence of lesions are key for the progression toward active tuberculosis: The bubble model. Front Microbiol. 2016; 7. https://doi.org/10.3389/fmicb.2016.00033 PMID: 26870005

19. Marino S, Linderman JJ, Kirschner DE. A multifaceted approach to modeling the immune response in tuberculosis. Wiley Interdiscip Rev Syst Biol Med. 2011; 3: 479–489. https://doi.org/10.1002/wsbm.131 PMID: 21197656

20. Marino S, Gideon HP, Gong C, Mankad S, McCrone JT, Lin PL, et al. Computational and Empirical Studies Predict Mycobacterium tuberculosis-Specific T Cells as a Biomarker for Infection Outcome. PLoS Comput Biol. 2016; 12. https://doi.org/10.1371/journal.pcbi.1004804 PMID: 27065304

21. Flynn JL, Tsenova L, Izzo A, Kaplan G. Experimental Animal Models of Tuberculosis. Handbook of Tuberculosis. 2017. pp. 389–426. https://doi.org/10.1002/9783527611614.ch32

22. Wong EA, Joslyn L, Grant NL, Klein E, Lin PL, Kirschner DE, et al. Low Levels of T Cell Exhaustion in Tuberculous Lung Granulomas. Infect Immun. 2018;86. https://doi.org/10.1128/iai.00426-18 PMID: 29891540

23. Pienaar E, Linderman JJ, Kirschner DE. Emergence and selection of isoniazid and rifampin resistance in tuberculosis granulomas. PLoS One. 2018; 13. https://doi.org/10.1371/journal.pone.0196322 PMID: 29746491

24. Sarathy J, Blanc L, Alvarez-Cabrera N, O'Brien P, Dias-Freedman I, Mina M, et al. Fluoroquinolone Efficacy against Tuberculosis Is Driven by Penetration into Lesions and Activity against Resident Bacterial Populations. Antimicrob Agents Chemother. 2019; 63. https://doi.org/10.1128/aac.02516-18 PMID: 30803965

25. Martin CJ, Cadena AM, Leung VW, Lin PL, Maiello P, Hicks N, et al. Digitally Barcoding Mycobacterium tuberculosis Reveals In Vivo Infection Dynamics in the Macaque Model of Tuberculosis. MBio. 2017; 8. https://doi.org/10.1128/mbio.00312-17 PMID: 28487426

26. Flynn JL, Gideon HP, Mattila JT, Lin P ling. Immunology studies in non-human primate models of tuberculosis. Immunol Rev. 2015; 264: 60–73. https://doi.org/10.1111/imr.12258 PMID: 25703552

27. Gideon HP, Phuah JY, Myers AJ, Bryson BD, Rodgers MA, Coleman MT, et al. Variability in Tuberculosis Granuloma T Cell Responses Exists, but a Balance of Pro- and Anti-inflammatory Cytokines Is Associated with Sterilization. PLoS Pathog. 2015; 11: 1–28. https://doi.org/10.1371/journal.ppat.1004603 PMID: 25611466

28. Lin PL, Rodgers M, Smith L, Bigbee M, Myers A, Bigbee C, et al. Quantitative comparison of active and latent tuberculosis in the cynomolgus macaque model. Infect Immun. 2009; 77: 4631–4642. https://doi.org/10.1128/IAI.00592-09 PMID: 19620341

29. Lin PL, Pawar S, Myers A, Pegu A, Fuhrman C, Reinhart TA, et al. Early events in Mycobacterium tuberculosis infection in cynomolgus macaques. Infect Immun. 2006. https://doi.org/10.1128/IAI.00064-06 PMID: 16790751

30. Lin PL, Coleman T, Carney JPJ, Lopresti BJ, Tomko J, Fillmore D, et al. Radiologic Responses in Cynomolgus Macaques for Assessing Tuberculosis Chemotherapy Regimens. Antimicrob Agents Chemother. 2013; 57: 4237–4244. https://doi.org/10.1128/AAC.00277-13 PMID: 23796926

31. Phuah J, Lin PL, Flynn JL, Chan J, Hendricks MR, Maiello P, et al. Effects of B Cell Depletion on Early Mycobacterium tuberculosis Infection in Cynomolgus Macaques. Infect Immun. 2016; 84: 1301–1311. https://doi.org/10.1128/IAI.00083-16 PMID: 26883591

32. Cadena AM, Hopkins FF, Maiello P, Carey AF, Wong EA, Martin CJ, et al. Concurrent infection with Mycobacterium tuberculosis confers robust protection against secondary infection in macaques. PLoS Pathog. 2018; 14. https://doi.org/10.1371/journal.ppat.1007305 PMID: 30312351

33. Darrah PA, DiFazio RM, Maiello P, Gideon HP, Myers AJ, Rodgers MA, et al. Boosting BCG with proteins or rAd5 does not enhance protection against tuberculosis in rhesus macaques. npj Vaccines. 2019; 4. https://doi.org/10.1038/s41541-019-0113-9 PMID: 31149352

34. van der Walt S, Schönberger JL, Nunez-Iglesias J, Boulogne F, Warner JD, Yager N, et al. scikit-image: image processing in Python. PeerJ. 2014; 2: e453. https://doi.org/10.7717/peerj.453 PMID: 25024921

35. Wiktorski T. Spark. Adv Inf Knowl Process. 2019; 85–97. https://doi.org/10.1007/978-3-030-04603-3_9

36. Coleman MT, Maiello P, Tomko J, Frye LJ, Fillmore D, Janssen C, et al. Early changes by 18Fluoro-deoxyglucose positron emission tomography coregistered with computed tomography predict outcome after Mycobacterium tuberculosis infection in cynomolgus macaques. Infect Immun. 2014; 82: 2400–2404. https://doi.org/10.1128/IAI.01599-13 PMID: 24664509

37. Wigginton JE, Kirschner D. A Model to Predict Cell-Mediated Immune Regulatory Mechanisms During Human Infection with Mycobacterium tuberculosis. J Immunol. 2001; 166: 1951–1967. https://doi.org/10.4049/jimmunol.166.3.1951 PMID: 11160244

38. Sud D, Bigbee C, Flynn JL, Kirschner DE. Contribution of CD8+ T Cells to Control of Mycobacterium tuberculosis Infection. J Immunol. 2014; 176: 4296–4314. https://doi.org/10.4049/jimmunol.176.7.4296 PMID: 16547267

39. Guzzetta G, Kirschner D. The Roles of Immune Memory and Aging in Protective Immunity and Endogenous Reactivation of Tuberculosis. PLoS One. 2013; 8. https://doi.org/10.1371/journal.pone.0060425 PMID: 23580062

40. Cilfone NA, Ford CB, Marino S, Mattila JT, Gideon HP, Flynn JL, et al. Computational Modeling Predicts IL-10 Control of Lesion Sterilization by Balancing Early Host Immunity–Mediated Antimicrobial Responses with Caseation during Mycobacterium tuberculosis Infection. J Immunol. 2015; 194: 664–677. https://doi.org/10.4049/jimmunol.1400734 PMID: 25512604

41. Marino S, Myers A, Flynn JL, Kirschner DE. TNF and IL-10 are major factors in modulation of the phagocytic cell environment in lung and lymph node in tuberculosis: A next-generation two-compartmental model. J Theor Biol. 2010; 265: 586–598. https://doi.org/10.1016/j.jtbi.2010.05.012 PMID: 20510249

42. Redford PS, Murray PJ, O'Garra A. The role of IL-10 in immune regulation during M. tuberculosis infection. Mucosal Immunol. 2011; 4: 261–270. https://doi.org/10.1038/mi.2011.7 PMID: 21451501

43. Mahamed D, Boulle M, Ganga Y, Mc Arthur C, Skroch S, Oom L, et al. Intracellular growth of Mycobacterium tuberculosis after macrophage cell death leads to serial killing of host cells. Elife. 2017; 6. https://doi.org/10.7554/eLife.22028 PMID: 28130921

44. Huang L, Nazarova E V., Tan S, Liu Y, Russell DG. Growth of Mycobacterium tuberculosis in vivo segregates with host macrophage metabolism and ontogeny. J Exp Med. 2018; 215: 1135–1152. https://doi.org/10.1084/jem.20172020 PMID: 29500179

45. Hoal-Van Helden EG, Hon D, Lewis LA, Beyers N, Van Helden PD. Mycobacterial growth in human macrophages: Variation according to donor, inoculum and bacterial strain. Cell Biol Int. 2001; 25: 71–81. https://doi.org/10.1006/cbir.2000.0679 PMID: 11237410

46. Read MN, Alden K, Rose LM, Timmis J. Automated multi-objective calibration of biological agent-based simulations. J R Soc Interface. 2016; 13. https://doi.org/10.1098/rsif.2016.0543 PMID: 27628175

47. Read MN, Alden K, Timmis J, Andrews PS. Strategies for calibrating models of biology. Brief Bioinform. 2018. https://doi.org/10.1093/bib/bby092 PMID: 30239570

48. Menzies NA, Soeteman DI, Pandya A, Kim JJ. Bayesian Methods for Calibrating Health Policy Models: A Tutorial. Pharmacoeconomics. 2017; 35: 613–624. https://doi.org/10.1007/s40273-017-0494-4 PMID: 28247184

49. Fallahi-Sichani M, El-Kebir M, Marino S, Kirschner DE, Linderman JJ. Multiscale Computational Modeling Reveals a Critical Role for TNF- Receptor 1 Dynamics in Tuberculosis Granuloma Formation. J Immunol. 2011; 186: 3472–3483. https://doi.org/10.4049/jimmunol.1003299 PMID: 21321109

50. Marino S, Hogue IB, Ray CJ, Kirschner DE. A methodology for performing global uncertainty and sensitivity analysis in systems biology. Journal of Theoretical Biology. 2008. pp. 178–196. https://doi.org/10.1016/j.jtbi.2008.04.011 PMID: 18572196

51. Cilfone NA, Kirschner DE, Linderman JJ. Strategies for Efficient Numerical Implementation of Hybrid Multi-scale Agent-Based Models to Describe Biological Systems. Cellular and Molecular Bioengineering. 2015. pp. 119–136. https://doi.org/10.1007/s12195-014-0363-6 PMID: 26366228

52. Cadena AM, Fortune SM, Flynn JL. Heterogeneity in tuberculosis. Nat Rev Immunol. 2017; 17: 691–702. https://doi.org/10.1038/nri.2017.69 PMID: 28736436

53. Cadena AM, Flynn JL, Fortune SM. The importance of first impressions: Early events in mycobacterium tuberculosis infection influence outcome. mBio. 2016. https://doi.org/10.1128/mBio.00342-16 PMID: 27048801

54. Lin PL, Maiello P, Gideon HP, Coleman MT, Cadena AM, Rodgers MA, et al. PET CT Identifies Reactivation Risk in Cynomolgus Macaques with Latent M. tuberculosis. PLoS Pathog. 2016; 12: e1005739. https://doi.org/10.1371/journal.ppat.1005739 PMID: 27379816

55. Maiello P, DiFazio RM, Cadena AM, Rodgers MA, Lin PL, Scanga CA, et al. Rhesus macaques are more susceptible to progressive tuberculosis than cynomolgus macaques: A quantitative comparison. Infect Immun. 2018; 86. https://doi.org/10.1128/IAI.00505-17 PMID: 28947646

56. Lin PL, Flynn JL. CD8 T cells and Mycobacterium tuberculosis infection. Semin Immunopathol. 2015; 37: 239–249. https://doi.org/10.1007/s00281-015-0490-8 PMID: 25917388

57. Rozot V, Vigano S, Mazza-Stalder J, Idrizi E, Day CL, Perreau M, et al. Mycobacterium tuberculosis-specific CD8+ T cells are functionally and phenotypically different between latent infection and active disease. Eur J Immunol. 2013; 43: 1568–1577. https://doi.org/10.1002/eji.201243262 PMID: 23456989

58. Commandeur S, Lin MY, van Meijgaarden KE, Friggen AH, Franken KLMC, Drijfhout JW, et al. Double- and monofunctional CD4 + and CD8 + T-cell responses to Mycobacterium tuberculosis DosR antigens

and peptides in long-term latently infected individuals. Eur J Immunol. 2011; 41: 2925–2936. https://doi.org/10.1002/eji.201141602 PMID: 21728172

59. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med. 2015; 372: 793–5. https://doi.org/10.1056/NEJMp1500523 PMID: 25635347

60. Segovia-Juarez JL, Ganguli S, Kirschner D. Identifying control mechanisms of granuloma formation during M. tuberculosis infection using an agent-based model. J Theor Biol. 2004; 231: 357–376. https://doi.org/10.1016/j.jtbi.2004.06.031 PMID: 15501468