nature medicine

Sterilization of granulomas is common in active and latent tuberculosis despite within-host variability in bacterial killing

Philana Ling Lin[1,9], Christopher B Ford[2,3,9], M Teresa Coleman[4], Amy J Myers[5], Richa Gawande[2], Thomas Ioerger[6], James Sacchettini[7], Sarah M Fortune[2,3,8] & JoAnne L Flynn[5]

Over 30% of the world's population is infected with *Mycobacterium tuberculosis* (Mtb), yet only ~5–10% will develop clinical disease[1]. Despite considerable effort, researchers understand little about what distinguishes individuals whose infection progresses to active tuberculosis (TB) from those whose infection remains latent for decades. The variable course of disease is recapitulated in cynomolgus macaques infected with Mtb[2]. Active disease occurs in ~45% of infected macaques and is defined by clinical, microbiologic and immunologic signs, whereas the remaining infected animals are clinically asymptomatic[2,3]. Here, we use individually marked Mtb isolates and quantitative measures of culturable and cumulative bacterial burden to show that most lung lesions are probably founded by a single bacterium and reach similar maximum burdens. Despite this observation, the fate of individual lesions varies substantially within the same host. Notably, in active disease, the host sterilizes some lesions even while others progress. Our data suggest that lesional heterogeneity arises, in part, through differential killing of bacteria after the onset of adaptive immunity. Thus, individual lesions follow diverse and overlapping trajectories, suggesting that critical responses occur at a lesional level to ultimately determine the clinical outcome of infection. Defining the local factors that dictate outcome will be useful in developing effective interventions to prevent active TB.

A prevailing model in the TB field is one in which the immune response in individuals whose infection progresses to active disease is less effective than the response in individuals who develop a latent infection. This model suggests that most lesions will behave similarly within a given individual in response to the systemic immune response and that global differences in lesional dynamics between individuals ultimately result in different clinical outcomes. However, recent studies suggest that this view is overly simplistic[2,4–8]. Using [^{18}F]fluorodeoxyglucose ([^{18}F]FDG) positron emission tomography in conjunction with computed tomography (PET-CT) imaging, we find that the metabolic activities (measured as standard uptake volume ratio, SUVR) of individual lesions are independent and dynamic during infection, even within the same host (**Fig. 1a**). The SUVR of each lesion loosely correlates with the number of colony-forming units (CFUs) in that lesion (Spearman's ρ, 0.4431; $P < 0.0001$) (**Fig. 1b**), suggesting lesional variability in both host response and bacterial growth. We previously reported a wide spectrum of lesion types by histopathology within individual animals[2], which is similar to descriptions of lesions in TB patients in the preantibiotic era[9]. These findings together appear inconsistent with the idea of a host response that is globally restrictive or permissive[2,10].

We therefore sought to define the quantitative dynamics of lesion formation and progression. To understand the initial formation of multiple sites of infection, we used a sequenced panel of eight Mtb isolates in which each is distinguished by a single nucleotide polymorphism (SNP)[11] serving as a 'molecular barcode' for each isolate (**Supplementary Table 1**). We found no evidence that the individual SNPs altered bacterial fitness *in vitro* (**Supplementary Fig. 1a**). We infected two macaques with a small dose (~34 CFU) of this inoculum and followed individual lesions temporally by [^{18}F]FDG PET-CT. After 4 weeks, animals were necropsied, and we quantified the relative abundance of each barcode in individual granulomas by amplicon deep sequencing. The median CFU per lesion for these animals is similar to the median CFU per lesion at 4 weeks in animals infected with wild-type Mtb Erdman strain (**Supplementary Fig. 1b**).

Although all eight barcodes were represented in each animal, notably, 79% of individual granulomas contained only one barcode. There are two possible explanations for these data. Either most granulomas were founded by only a single bacterium (**Fig. 2a** and **Supplementary Table 2**), or in each granuloma a single isolate outcompeted the others.

[1]Department of Pediatrics, Children's Hospital of Pittsburgh, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA. [2]Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, Massachusetts, USA. [3]Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA. [4]Department of Radiology, Positron Emission Tomography Research Center, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA. [5]Department of Microbiology and Molecular Genetics, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA. [6]Department of Computer Science and Engineering, Texas A&M University, College Station, Texas, USA. [7]Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas, USA. [8]Ragon Institute of MGH, MIT, and Harvard, Boston, Massachusetts, USA. [9]These authors contributed equally to this work. Correspondence should be addressed to S.M.F. (sfortune@hsph.harvard.edu) or J.L.F. (joanne@pitt.edu).

Received 17 September; accepted 29 October; published online 15 December 2013; doi:10.1038/nm.3412

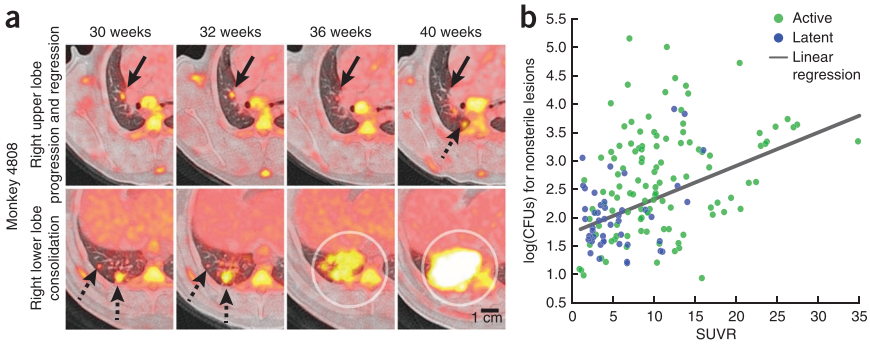

Figure 1 Serial imaging reveals the dynamic evolution of lesions in TB. (**a**) [^{18}F]FDG PET-CT images of progressing and regressing lesions shown from the same animal, which developed active disease following low-dose infection (at 30, 32, 36 and 40 weeks after infection). The images show a resolving lesion in the right upper lobe (top, solid black arrows) at the same time that new lesions appear (top, dashed black arrow). In the right lower lobe (bottom), lesions progress (dashed black arrows) and new lesions coalesce to form a consolidation (circles). (**b**) [^{18}F]FDG avidity (SUVR) is plotted against log(CFU) (Spearman's ρ, 0.4431). A linear regression model of SUVR versus CFU is shown (slope, 0.058 ± 0.0097; $P < 0.001$). Dots represent individual lesions ($n = 274$) from monkeys with active disease (green, $n = 15$) or latent infection (blue, $n = 10$).

If fitness differences were responsible for these observations, then the prevalence of more fit isolates should increase significantly within a host over their prevalence in the inoculum. We developed a stochastic mathematical simulation to predict the prevalence of each barcoded isolate *in vivo* on the basis of random sampling from the stock to generate the inoculum (**Supplementary Fig. 1c**,**d**). No isolates were significantly under-represented *in vivo*, but one barcode, I3, was significantly over-represented at 4 weeks after infection ($P = 0.023$), which is consistent with a 1% increase in fitness (**Supplementary Fig. 1d**,**e**). However, after removing all lesions containing I3 from our analysis (thus excluding 7 of 45 lesions), 88% of lesions still contained only one barcode. Therefore, our data suggest that most granulomas arise from an individual bacterium.

In most lesions where more than one barcode was identified, there was histologic evidence of 'multifocal lesions' (lesions composed of two or more granulomas). However, even among multifocal lung lesions, 64% contained a single barcode, suggesting that the majority arose via localized spreading (new granulomas forming from an existing granuloma) as opposed to coalescence between adjacent independent granulomas. In contrast to lung granulomas, thoracic lymph nodes often contained several different barcodes (**Fig. 2a**), possibly reflecting draining of infecting bacilli from lungs to lymph nodes. There was no significant relationship between bacterial burden and the number of founder bacteria in individual granulomas (**Fig. 2b**), such that initial bacterial load did not influence final bacterial burden. A range of bacterial burdens per lesion was seen in granulomas, multifocal lesions and lymph nodes; however, lesions arising at earlier time points had a significantly higher burden (approximately sixfold, $P = 0.0012$) than lesions arising 1–2 weeks later (**Fig. 2c**).

We next sought to assess the trajectory of the bacterial population in individual lesions from 4 to 11 weeks after infection, a period coinciding with the onset of the adaptive immune response. We infected seven macaques, tracked lesions using [^{18}F]FDG PET-CT and then measured CFU per lesion. At 4 weeks, the median bacterial burden in individual granulomas was ~1.8×10^4 CFU (**Fig. 3a**), but this was reduced significantly (~15-fold, $P = 0.0002$) by 11 weeks, suggesting significant killing after the onset of adaptive immunity. Although it is too early to distinguish between latent infection and active disease at 11 weeks, one animal had a significantly higher median bacterial burden than the other two, suggesting an inflection point in disease trajectory (i.e., progression to active disease or development of latent infection) (**Fig. 3b**, $P < 0.05$). However, there was also a notable range of bacterial burdens (1×10^2 to 1×10^6 CFU per lesion) within each macaque, suggesting that the extent of killing may vary dramatically between lesions.

To understand how these early phenotypes relate to disease outcome, we next assessed bacterial burden in lesions from animals with active disease or latent infection. Although animals with active TB harbor significantly more bacteria than animals with latent infection[2] ($P < 0.001$), at a lesional level, there is a broad and overlapping distribution of CFU per granuloma (**Fig. 3a** and **Supplementary Fig. 2a**,**b**). Notably, the median CFU of individual lesions from animals with active disease (4–12 months after infection) and animals with latent infection (>6 months after infection) is 10- to 100-fold lower than that of lesions in animals at 4 and 11 weeks after infection (**Fig. 3a**), and a substantial number of lesions in animals with active disease and animals with latent infection are sterile (no colonies were visible after 6 weeks on solid medium) (**Fig. 3c**). Although we cannot rule out the

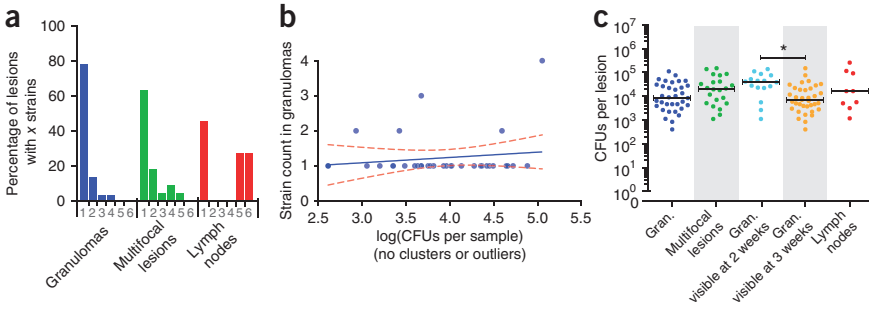

Figure 2 The majority of lesions in monkeys are initiated by a single bacterium. (**a**) The percentage of lesions containing the indicated number of barcodes *x* strains are shown, with percentage shown along the *y* axis. Lesions were classified as granulomas ($n = 57$), multifocal lesions ($n = 22$) or lymph nodes ($n = 13$). (**b**) For each granuloma, the number of strains (or barcodes) found in each lesion (*y* axis) is plotted against the log(CFU per sample) (*x* axis). A linear regression of total bacterial load in individual granulomas ($n = 37$) indicates no relationship between CFU and the number of barcodes recovered (Spearman's ρ, 0.24; slope, 0.15 ± 0.19; $P = 0.44$. (**c**) Bacterial load (CFU) per lesion in individual lesions, multifocal lesions and lymph nodes is shown. Granulomas were further divided by the timing of their formation based on [^{18}F]FDG PET-CT into lesions visible at 2 weeks ($n = 16$) or 3 weeks ($n = 36$). Bars represent medians. Gran., granulomas.

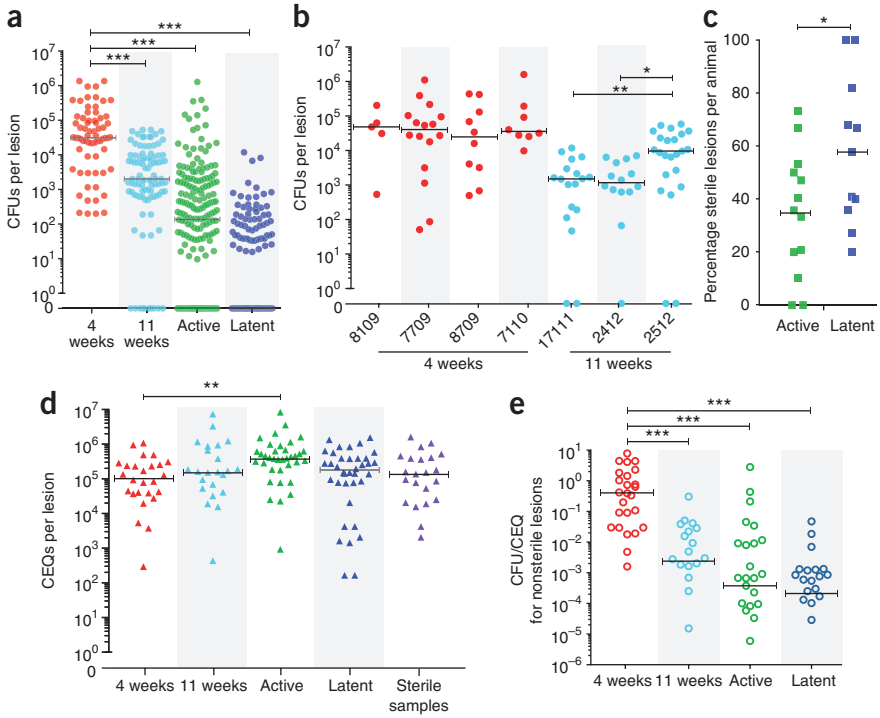

Figure 3 CFU and CEQs reflect viable and total bacterial burden in individual lesions. (**a**) CFUs per lesion from monkeys at 4 weeks (4 animals, 68 lesions), 11 weeks (3 animals, 98 lesions), with active disease (13 animals, 222 lesions) and with clinically latent infection (11 animals, 145 lesions) are shown. All groups are significantly different by pairwise comparison. Circles represent individual lesions. The median for animals with latent infection is 0. (**b**) The CFUs per lesion for individual monkeys necropsied at 4 and 11 weeks are shown; circles represent individual lesions. By 11 weeks, there were significant differences between animals. (**c**) The percentage of sterile lesions in individual monkeys is calculated for animals with latent infection ($n = 11$) or active disease ($n = 13$). Squares represent individual animals. (**d**) CEQs per lesion are shown for animals following 4 weeks of infection ($n = 26$) and 11 weeks of infection ($n = 23$) and for animals with active disease ($n = 36$) or latent infection ($n = 37$). CEQs from individual sterile lesions from all disease states ($n = 22$) are also shown. Triangles represent individual lesions. (**e**) The ratio of CFU to CEQs in individual nonsterile lesions was determined for each lesion in animals necropsied after 4 weeks ($n = 26$) and 11 weeks of infection ($n = 22$), as well as in animals with active disease ($n = 29$) or latent infection ($n = 28$). Open circles represent individual lesions. For all panels in all figures, $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$; statistical tests as indicated in Online Methods. Bars represent medians.

possibility of viable but nonculturable bacteria in these sterile lesions, we have previously shown that lesional CFU do not increase after extended culture (~1 year)[2]. Moreover, although immunosuppression should allow reactivation of quiescent Mtb, animals with latent infection treated with tumor necrosis factor–specific antibody had a proportion of sterile lesions (54%) similar to that of animals with latent infection that were not treated (58%) (refs. 10,12). These data indicate that sterile lesions develop in animals with both active disease and latent infection but do not occur until after 4 weeks of infection.

The comparison of lesional CFU suggests that the efficacy of bacterial killing dictates lesional bacterial burden and that there is marked lesion-to-lesion variability in bacterial killing. However, sterile lesions

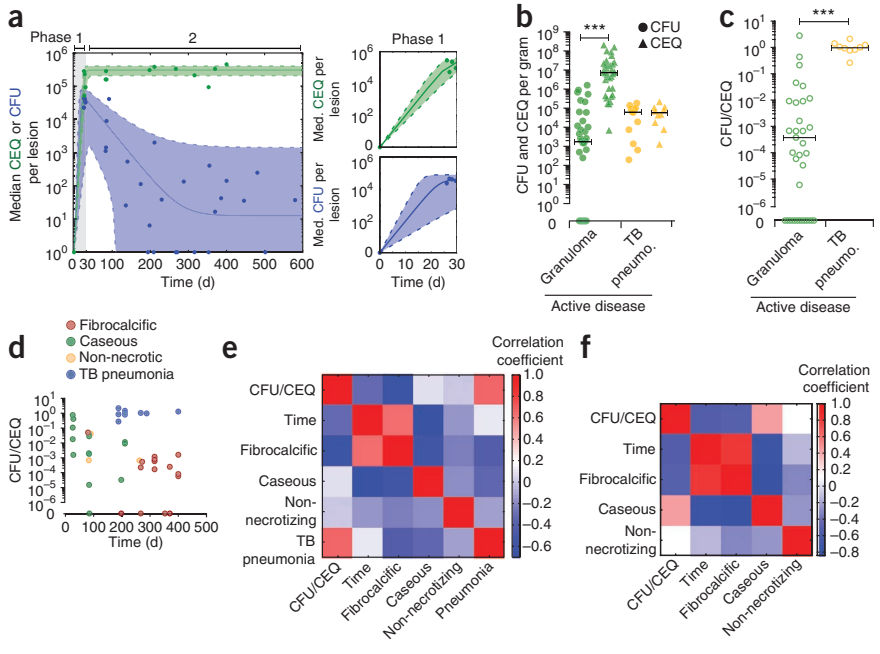

Figure 4 Models and correlates of replication and killing. (**a**) Median (Med.) CEQs per lesion for each animal (green, $n = 16$) and CFU per lesion for each animal (blue, $n = 29$) were fit with ordinary differential equations describing growth and death in individual lesions (solid line, with 95% confidence intervals shown as dashed lines). t_0 data were inferred from our barcode data, which suggests that most granulomas start with a single bacterium. CFU data were modeled as two phases: phase 1 represents initial growth ($t = 0$–30 d; phase 2 represents a subsequent killing ($t = 30$–600 d). Doubling times in individual lesions were inferred from CEQ and CFU fitted curves (**Supplementary Table 4**). Phase 1 is shown in detail in panels on the right and validated against 3-week CFU samples in **Supplementary Figure 3**. (**b**) CFU (circles) and CEQs (triangles) per gram of tissue are shown for granulomas from animals with active disease or sites with TB pneumonia (pneumo.). Bars represent medians. (**c**) The ratio of CFU per gram to CEQs per gram (open circles) calculated for TB pneumonia samples to estimate killing at these sites. Bars represent medians. (**d**) Plotting time, histopathology and CFU/CEQ reveals associations between these variables. (**e**) A correlation matrix reveals relationships between variables. Color represents the value of the correlation coefficient (**f**). We removed TB pneumonia samples from the analysis to focus on granulomatous sites. Color represents the value of the correlation coefficient.

do not necessarily reflect true bacterial killing activity. To assess the degree of bacterial killing in individual lesions, we took advantage of the finding that genomes (chromosomal equivalents, CEQs) from nonviable Mtb are not degraded but instead persist in tissues in the mouse model of TB[13]. The number of genomes reflects the cumulative bacterial burden, and the ratio of CFU to CEQs reflects the degree of killing. To determine whether genomes of killed Mtb are similarly stable in macaque lesions, we compared CEQs from animals treated with isoniazid for 2 months to CEQs from untreated animals over a similar time frame. Median CEQs in lesions from treated animals were orders of magnitude higher than CFU, though only ~20-fold lower than median CEQs in lesions from untreated animals (**Supplementary Fig. 2c**). We conclude that although genomes may degrade slowly over time in macaque lesions, the rate of decay is slow enough that CEQs can be used to estimate cumulative bacterial burdens per lesion across animals.

We assessed CEQs in lesions from animals at 4 and 11 weeks after infection and from those with active disease and animals with latent infection (information on infection times shown in **Supplementary Table 3**). Notably, CEQs were stable across disease states, and CEQs from sterile lesions were indistinguishable from CEQs from nonsterile lesions (**Fig. 3d**). The only significant difference in CEQs was between lesions in animals at 4 weeks after infection and those in animals with active disease (~3.5-fold greater in animals with active disease, $P < 0.01$), although the number of viable bacteria in lesions (CFU) from every group of animals was significantly different (**Fig. 3a**, $P < 0.001$). The CFU/CEQ ratio reflects the proportion of the bacterial population that is still viable, that is, the population that has not been killed. There was minimal killing at 4 weeks, but by 11 weeks the CFU/CEQ ratio was significantly reduced compared to that at 4 weeks and even lower in animals with active disease or latent infection (**Fig. 3e**, $P < 0.001$). Thus, as the adaptive immune response is induced[14], killing of bacteria in individual lesions occurs, even to the point of lesional sterilization. Furthermore, both animals with latent infection and animals with active disease have the capacity to sterilize lesions, which demonstrates that the capacity to sterilize a lesion is not a hallmark of latent infection alone.

To estimate the rate of Mtb growth and killing in lesions, we developed and fit an ordinary differential equation–based logistic model of bacterial dynamics in the host (**Fig. 4a**, **Supplementary Table 4** and **Supplementary Fig. 3**). Our data are best fit by a model in which there is an initial rapid growth phase, after which bacterial growth is drastically slowed and bacterial death leads to a reduction in CFU. Our CEQ data suggest that lesions reach a maximum bacterial burden or 'carrying capacity' coincident with the onset of adaptive immunity. Carrying capacity is probably driven by multiple factors, including the immune response and the maximum physical capacity of a granuloma; lesions that progress beyond this capacity may spread locally or lead to TB pneumonia[2]. Indeed, after measuring CFU and CEQs per gram in regions of TB pneumonia from animals with active disease, we found no evidence of bacterial killing (median CFU/CEQ = 1.05), suggesting that these sites are permissive environments for Mtb survival (**Fig. 4b**,**c**).

These data suggest that the histopathology of a lesion may predict local killing capacity. We sought to define the associations between histopathology, time after infection and CFU/CEQ ratio (**Fig. 4d**). A correlation matrix reveals that TB pneumonia is positively correlated with CFU/CEQ ratio ($r = 0.66$, $P = 9.24 \times 10^{-7}$) and poorly correlated with time after infection, suggesting that at any time during infection, there is little killing in areas of TB pneumonia (**Fig. 4e**). By excluding TB pneumonia samples, we found that fibrocalcific lesions negatively correlate with CFU/CEQ ratio ($r = -0.56$, $P = 0.002$) and represent sites of bacterial control, whereas caseous sites are associated with poor control ($r = 0.44$, $P = 0.0086$) (**Fig. 4f**). However, both fibrocalcific and caseous lesions significantly correlate with time after infection ($r = 0.80$, $P = 7.19 \times 10^{-9}$ and $r = -0.74$, $P = 4.96 \times 10^{-7}$, respectively), which suggests that, over time, killing increases and the dominant lesion type switches from caseous to fibrocalcific. These data are consistent with our previous findings in which animals with active TB were treated with short-course anti-TB treatment and residual bacterial growth after treatment occurred primarily in caseous granulomas and TB pneumonia[10].

Ultimately, these data imply that the difference in bacterial numbers is driven largely by killing efficacy, which varies considerably among lesions even within a single animal. Animals with active disease and those with latent infection have similar capacities for bacterial killing in individual lesions (**Fig. 3e**). However, animals with active disease can simultaneously have sites of extensive tissue pathology (i.e., TB pneumonia) where there is little bacterial killing. Given this heterogeneity, it is not surprising that relevant predictors of disease outcome have been hard to identify from peripheral measures of immune response[15,16]. As virtually all macaques appear capable of eliminating infection at the lesional level, defining the mechanism of lesion sterilization may provide a path to a TB vaccine capable of generating sterilizing immunity.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

These studies were funded by the Bill & Melinda Gates Foundation (grants to J.L.F. and P.L.L.), the Otis Childs Trust of the Children's Hospital of Pittsburgh Foundation (P.L.L.), US National Institutes of Health, National Institute of Allergy and Infectious Diseases DAIT BAA-05-10 (J.L.F.), US National Institutes of Health grants HL106804 (J.L.F.), AI094745 (J.L.F.), HL110811 (J.L.F.), DP2 0D001378 (S.M.F.) and AI076217 (S.M.F.), the Howard Hughes Medical Institute, the Physician Scientist Early Career Award (S.M.F.), the Harvard Merit Fellowship (C.B.F.), the Burroughs Wellcome Foundation Investigator in the Pathogenesis of Infectious Diseases Fellowship (S.M.F.), the Robert A. Welch Foundation (J.S.) and the Melvin J. and Geraldine L. Glimcher Associate Professorship (S.M.F.). We are grateful to E. Klein and C. Janssen for conducting necropsies, C. Scanga for coordination of studies, M. Rodgers, C. Cochran and C. Bigbee for excellent technical assistance, M. O'Malley, P. Johnston, J. Tomko, D. Fillmore and J. Frye for outstanding veterinary technical assistance, the members of the Flynn and Fortune labs for helpful discussions and B. Bloom, D. Young and E. Rubin for helpful comments on the manuscript.

AUTHOR CONTRIBUTIONS

P.L.L., C.B.F., S.M.F. and J.L.F. conducted experiments, analyzed results and drafted the manuscript. R.G., T.I. and J.S. performed Illumina sequencing and edited the manuscript. M.T.C. performed PET-CT analysis and assisted in necropsies. A.J.M. conducted experiments.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Zumla, A., Raviglione, M., Hafner, R. & von Reyn, C.F. Tuberculosis. *N. Engl. J. Med.* **368**, 745–755 (2013).
2. Lin, P.L. *et al.* Quantitative comparison of active and latent tuberculosis in the cynomolgus macaque model. *Infect. Immun.* **77**, 4631–4642 (2009).

3. Capuano, S.V. III *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).
4. Barry, C.E. III *et al.* The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nat. Rev. Microbiol.* **7**, 845–855 (2009).
5. Lin, P.L. & Flynn, J.L. Understanding latent tuberculosis: a moving target. *J. Immunol.* **185**, 15–22 (2010).
6. Kaplan, G. *et al. Mycobacterium tuberculosis* growth at the cavity surface: a microenvironment with failed immunity. *Infect. Immun.* **71**, 7099–7108 (2003).
7. Via, L.E. *et al.* Infection dynamics and response to chemotherapy in a rabbit model of tuberculosis using [^{18}F]2-fluoro-deoxy-d-glucose positron emission tomography and computed tomography. *Antimicrob. Agents Chemother.* **56**, 4391–4402 (2012).
8. Davis, J.M. & Ramakrishnan, L. The role of the granuloma in expansion and dissemination of early tuberculous infection. *Cell* **136**, 37–49 (2009).
9. Canetti, G. *The Tubercle Bacillus* (Springer Publishing, New York, 1955).
10. Lin, P.L. *et al.* Metronidazole prevents reactivation of latent *Mycobacterium tuberculosis* infection in macaques. *Proc. Natl. Acad. Sci. USA* **109**, 14188–14193 (2012).
11. Ford, C.B. *et al.* Use of whole genome sequencing to estimate the mutation rate of *Mycobacterium tuberculosis* during latent infection. *Nat. Genet.* **43**, 482–486 (2011).
12. 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 Rheum.* **62**, 340–350 (2010).
13. Munoz-Elías, E.J. *et al.* Replication dynamics of *Mycobacterium tuberculosis* in chronically infected mice. *Infect. Immun.* **73**, 546–551 (2005).
14. Wallis, R.S. *et al.* Tuberculosis biomarkers discovery: developments, needs, and challenges. *Lancet Infect. Dis.* **13**, 362–372 (2013).
15. Lin, P.L. *et al.* Early events in *Mycobacterium tuberculosis* infection in cynomolgus macaques. *Infect. Immun.* **74**, 3790–3803 (2006).
16. Walzl, G., Ronacher, K., Hanekom, W., Scriba, T.J. & Zumla, A. Immunological biomarkers of tuberculosis. *Nat. Rev. Immunol.* **11**, 343–354 (2011).

ONLINE METHODS

Macaque infection and clinical classification. Healthy adult (>4 years, males) cynomolgus macaques (*Macacca fasicularis*) (Valley Biosystems, Sacramento, California) were infected via bronchoscopic instillation of low-dose (~25 CFU) *M. tuberculosis* (Erdman strain) as previously described[2,3] (**Supplementary Table 3**). All animals received a similar low dose for infection without specific randomization or blinding. For the SNP strain infections, the inoculum was 34 CFU per animal. After infection, microbiologic (Mtb growth by gastric aspirate (GA) or bronchoalveolar lavage (BAL)), clinical (for example, weight loss, cough, anorexia and activity) and immunologic (erythrocyte sedimentation rate (ESR)) assessments were conducted every 2–4 weeks. Latent infection was defined at 6 months after infection in animals without clinical signs of disease, normal ESR and no evidence of Mtb growth by serial GA or BAL. Active disease was established based on evidence of clinical deterioration, persistent growth of Mtb by GA or BAL and increased ESR at 3 months after infection as previously defined[2]. In some cases, data were obtained from previously published studies[10,17]. Samples were chosen for each experiment to ensure both a broad sampling of disease states and an in-depth analysis of individual animals. For short-term experiments, macaques were infected for 4 weeks ($n = 4$) or 11 weeks ($n = 3$). Macaques in the barcoded Erdman strain[11] experiments were infected for 4 weeks ($n = 2$) (**Supplementary Table 3**). Animal studies were approved by the University of Pittsburgh's Institutional Animal Care and Use Committee, which followed the Public Health Service Policy on Humane Care and Use of Laboratory Animals and is approved by the Association for Assessment and Accreditation of Laboratory Animal Care International.

Isoniazid treatment. To assess chromosomal equivalents after isoniazid treatment among lesions for which bacteria could not be cultured, samples from previously published studies were obtained[17]. These animals had active disease and initiated treatment with isoniazid ~13 weeks after infection.

[^{18}F]fluorodeoxyglucose positron emission tomography and computed tomography imaging and analysis. Positron emission tomography using [^{18}F]fluorodeoxyglucose co-registered with computed tomography (PET-CT) was performed using a hybrid microPET Focus 220 preclinical PET scanner (Siemens Molecular Solutions, Knoxville, Tennessee) and an eight-slice helical CT scanner (Neurologica Corp, Danvers, Massachusetts)[17]. For each specific scan, macaques were sedated, intubated and placed on a ventilator during each [^{18}F]FDG PET-CT. For 4-week experiments, animals were scanned at baseline 2, 3 and 4 weeks after infection. [^{18}F]FDG PET-CT was performed only once immediately before necropsy in animals infected for 11 weeks and a select group of animals with active disease and latent infection. Each lung granuloma was characterized by size (in mm) and metabolic activity (as measured by standard uptake volume, SUV) on each scan, and lesions were tracked over time by serial scans. An SUV ratio (SUVR) was developed to normalize the degree of SUV variability between scans[17]. To determine the relationship between SUVR and CFU, we used the linear regression function of Prism (GraphPad, La Jolla, California) to determine the slope for the line of best fit including confidence intervals and goodness of fit (r^2).

Bacterial burden estimates at necropsy. [^{18}F]FDG PET-CT was performed 1–2 d before necropsy. At necropsy, scan-specific lesions were prioritized for harvest, and the remaining granulomas in the lung and mediastinal lymph nodes were harvested. Size of each granuloma was measured at necropsy and by prenecropsy scan. Granulomas were dissected, and one-half of each was submitted for bacterial burden and the other half for histopathology as previously described[2,10]. Samples were plated on 7H10 agar supplemented with oleic albumin dextrose catalase (OADC) and incubated for 3 weeks before enumeration of colonies. The number of bacilli per individual granuloma was determined using the CFU per gram of granuloma tissue multiplied by the weight of the granuloma, based on a standard curve of weight for granulomas of various sizes. To determine the accuracy of our standard curve, we measured CFU per granuloma using both weight and extrapolation from size for 64 granulomas from 5 randomly chosen animals. The results were highly correlative (Spearman's $\rho = 0.99$, $P < 0.0001$), which supported the use of size as a proxy for weight. Areas of TB pneumonia were identified grossly, harvested and weighed, and a predetermined portion was taken for bacterial burden estimates (measured as CFU per gram) as previously described[2]. Granuloma histopathological characteristics were assessed by a certified primate pathologist. Specifically, the granuloma type (for example, caseous or non-necrotizing) and pattern of granulomas (focal, coalescing or disseminated) were used for analysis in these studies to determine whether lesions had evidence of spreading[2].

Determination of sterility. Sterile lesions were defined as lesions with no growth after 6 weeks, with a limit of detection of 10, as previously published[2]. To confirm that plates with no growth do not actually represent plates with slowly growing bacteria, we use an extended incubation time of 6 weeks. Any plate that failed to produce colonies after 6 weeks was recorded as sterile for the purposes of these analyses. We previously incubated plates for a period of 1 year, checking each plate at 1-month intervals[2]. Only a small fraction (~1%) had new growth after 6 weeks, consistent with previous reports from resected lesions in humans[18,19]. Additionally, as a more stringent test of the potential for 'dormant' or nonculturable but still potentially infectious bacilli in lesions, we analyzed data from a separate set of studies performed in our lab where latently infected macaques were administered antibody to TNF (adalimumab, Abbott Labs, 4 mg per kg body weight per dose, weekly, for 6–8 weeks) to induce reactivation[10,12]. Antibody to TNF is known to be a trigger for reactivation in humans[20]. We previously published that TNF-specific antibody treatment for 6–8 weeks resulted in reactivation of ~70% of macaques with latent infection[10,12]. We reasoned that this treatment would result in regrowth of bacilli in lesions, and therefore, if lesions were truly sterile, they would not grow bacilli after treatment with TNF-specific antibody. We found that 54% of the individual granulomas identified at necropsy did not grow Mtb (i.e., were sterile) among nine macaques treated with anti-TNF antibody, similar to the results reported in **Figure 3c** for animals with latent infection (58%). This stringent test provides strong evidence that these lesions do not contain viable bacilli, even under conditions of severe immunosuppression.

Creation and validation of pooled barcoded library for infection. Strains (described in **Supplementary Table 1**) were previously sequenced[11] and found to differ from one another by only a single nucleotide polymorphism. Individual cultures of each strain were expanded in broth culture of 7H9 supplemented with 10% Middlebrook OADC, 0.0005% Tween 80 and 0.005% glycerol, with shaking at 37 °C to an optical density of 1.0 at a wavelength of 600 nm. The inoculum for infection was prepared by mixing each strain in equal parts. To determine the potential fitness effects of each single nucleotide polymorphism, the library was expanded *in vitro* over the course of 9 d, with sampling every 3 d.

Determination of barcoded strain identity. Two cynomolgus macaques were infected as described above with the above-described barcoded library (**Supplementary Table 1**). At 4 weeks, animals were killed, and individual lesions were plated for CFU. Plates were scraped, colonies were pooled and genomic DNA was extracted as follows. Cells were harvested from plates into PBS and centrifuged at 4,000 r.p.m. for 10 min at 4 °C to separate the supernatant. The cell pellet was resuspended in Tris-EDTA buffer (TE, 0.1 M Tris and 1 mM EDTA, pH 9.0). An equal volume of chloroform/methanol (2:1) was added to the cell suspension and thoroughly mixed for 5 min. The suspension was centrifuged again for 10 min at 4,000 r.p.m. at 4 °C, and both the organic and aqueous phase were discarded. After allowing the pellet to air dry for 1 h, it was resuspended in TE buffer and supplemented with lysozyme to final concentration of 100 μg mL^{-1}. The suspension was incubated overnight at 37 °C, after which the suspension was supplemented with one-tenth volume of 10% SDS and proteinase K to a final concentration of 100 μg mL^{-1}. After mixing by vortex, this suspension was incubated at 50 °C for 2 h. An equal volume of phenol/chloroform (1:1) was added and mixed well for 30 min before separation by centrifugation at 10,000 r.p.m. for 15 min at 4 °C. The aqueous phase was removed after centrifugation and supplemented with one-half volume of chloroform and again centrifuged at 14,000 r.p.m. for 15 min at 4 °C. The aqueous phase was again removed and then supplemented with RNase A (Life Technologies) to a final concentration of 25 μg mL^{-1} and incubated at 37 °C for 1 h. The samples were then stored at 4 °C for 10 min before being

doi:10.1038/nm.3412

supplemented with an equal volume of phenol/chloroform and again centrifuged at 14,000 r.p.m. for 15 min at 4 °C. The aqueous layer was again removed and supplemented with one-half volume of chloroform, after which samples were centrifuged at 14,000 r.p.m. for 15 min at 4 °C, and the aqueous layer was removed. DNA was precipitated from the aqueous layer by adding one volume of isopropanol and incubating at −20 °C for 12 h. Mtb genomic DNA was washed using 70% ethanol and dissolved in sterile distilled deionized water.

Each sample was used as template for eight PCR reactions (**Supplementary Table 5**) to amplify the eight loci listed in **Supplementary Table 1.** Each primer pair contained a sequence tag from which a second PCR reaction was initiated using an additional set of primers to attach the sequence content necessary for Illumina sequencing and barcoding (**Supplementary Table 5**). Pooled libraries were subsequently sequenced using an Illumina GenomeAnalyzer IIx, with a read length of 54 bp. A mean of 840,000 reads was collected per sample. The reads were aligned to the eight targeted loci, yielding approximately 100,000 reads covering each polymorphic site. The frequency of each allele was calculated based on the proportion of nucleotides observed. To distinguish between background signal and noise resulting from both sequencing and experimental variation, we required a percentage greater than 2.5. Strain counts were determined for each sample, and the relationship to $\log_{10}$(CFU) per granuloma was determined by linear regression, Spearman correlation and comparison between groups (where groups were samples with x strains) by Kruskal-Wallis one-way analysis of variance. All tests were performed using Prism (GraphPad, La Jolla, California).

Determining barcode prevalence and fitness *in vivo*. To determine the probability of observing any given barcode at the end of infection, we generated a stochastic simulation of sampling from the inoculum stock. To simulate the sample, we ran 1 million random simulations i = 1:1,000,000, where the number of each barcode was determined by a random number (X_i), where X_i is a binomial random variable: X_i ~B(34, S_i). B represents the binomial distribution described by the parenthetical parameters. S_i, the probability of success of each trial, was estimated on the basis of the prevalence of each barcode in the inoculum stock, and the number of trials (34) was estimated by a plating of the inoculum. However, the sum of X_i varies stochastically with each sampling. To determine whether the prevalence after infection differed significantly from the prevalence before infection, the number of runs in which the simulated sampling exceeded the prevalence after infection was divided by the total number of runs (1 million). Significant deviations from the inoculum were defined as those occurring in less than 5% of runs, suggesting that this observation was unlikely to occur by chance. Notably, we did not correct for repeated testing of each barcode (n = 8), leading us to conservatively 'over call' strains that significantly deviated from the inoculum. Barcode I3 is the only isolate whose prevalence was significantly different from the expected value (P = 0.023). We reasoned that fitness changes could explain this result. To better understand the fitness change necessary to skew the observed distribution of barcodes after 28 d of growth, we expanded our simulation. The prevalence of each barcode after 28 d of infection was determined by

$$pr\left(i_{t(28)}\right) = \frac{N_{i0} \times e^{fr}}{N_{\text{total}}} \tag{1}$$

where $pr(i_{t(28)})$ is the prevalence of each barcode (i) after 28 d of infection based on the initial population (N_{i0}, simulated previously by random sampling) times the exponential growth equation divided by the total population (N_{total}). The rate parameter (r) of the exponential growth equation was parameterized on the basis of estimates determined by CEQs (detailed below), and time (t) was determined on the basis of length of infection (28 d). Fitness (f) was iterated from no change in fitness (f = 1.00) to 15% (f = 1.15) by an interval of 1% (0.01) for barcode I3 only. For all other barcodes, f = 1.00. To determine the fitness level at which the probability of the observed prevalence of strain I3 was greater than 5%, the number of runs in which the simulated growth exceeded the prevalence after infection was divided by the total number of runs (one million) for each fitness iteration.

Isolation and quantification of *M. tuberculosis* genomes from lesions. Following necropsy, samples were stored in PBS at −80 °C until processing. Investigators were blinded as to the clinical status of each sample. Volumes were tracked throughout genomic extraction to correct for volume loss. Samples were resuspended in 1 mL Tris-EDTA buffer, pH 8.0, and supplemented with 300 µL of 70 °C phenol/chloroform/isoamyl alcohol (25:24:1, Sigma-Aldrich). Samples were mixed by inversion and incubated at room temperature for 10 min before being transferred to tubes containing 250 µL of 0.1-mm zirconia-silica beads (MP Biomedicals) and pulsed twice for 30 s each using the FastPrep 24 system (MP Biomedicals), with a 5-min break between pulses. The supernatant was then separated by centrifugation at 14,000 r.p.m. and supplemented with 50 µL of 5 M sodium chloride. The phenol/chloroform/isoamyl alcohol (25:24:1, Sigma-Aldrich) extraction was repeated, and samples were again centrifuged at 14,000 r.p.m. to separate the aqueous phase. Genomic DNA was precipitated using one volume of isopropanol and one-tenth volume of 3 M sodium acetate (pH 5.2) and washed with one volume of 70% ethanol. A subset of samples was further purified by dialysis. Mtb genomic DNA was then quantified using the previously described primer-probe combination[13] (**Supplementary Table 5**) (Integrated DNA Technologies). Chromosomal equivalents (CEQs) were quantified in relation to a standard curve derived from serial dilution of Mtb genomic prepared from liquid culture. Quantitative RT-PCR was performed in triplicate on the Viia 7 (Life Technologies) using TaqMan Universal Master Mix II (Life Technologies). Researchers conducting sample processing and quantification of CEQs were blinded to the live bacterial counts until the time of analysis.

Mathematical model of population dynamics and estimates of replication rate. To determine the rate of growth and death of the Mtb population in individual lesions, we developed a set of ordinary differential logistic equations describing population dynamics (equations (1) and (2)) in each lesion. To accurately describe both the expansion of the viable bacterial population (CFU) and the decline of the viable bacterial population following the onset of adaptive immunity after 4 weeks, these data were divided into two separate phases (I and II) and analyzed by the following equations ((2) and (3), respectively). The total bacterial population (represented by chromosomal equivalents, CEQs) was modeled by equation (2) only

$$\frac{dN_1}{dt} = (r_1 \times N_1) \times \frac{(K_1 - N_1)}{K_1} \tag{2}$$

$$\frac{dN_2}{dt} = (r_2 \times N_2) \times \frac{(N_2 - K_2)}{K_2} \tag{3}$$

Here, r represents the growth rate of the population in replications per day, N_1 and N_2 represent the population size and K represents carrying capacity. The parameters for equations (2) and (3) were optimized to best fit the medians for CEQs and CFU at each time point where data were available using Matlab (Natick, Massachusetts). Specifically, scripts were created to fit the equations above using the *nlinfit* function with maximum iterations set to 10,000 and the weight function for robust fitting set to 'fair'. The initial values were chosen on the basis of estimates derived from the data and adjusted to provide the fit and final parameter estimates that produced the smallest residuals. To determine the replication rate (in days per duplication), we began with the discrete form of the logistic equation

$$x_t = x_0\left(1 + r \times \frac{K_{\text{CEQ}} - \text{fit}_{\text{CEQ}}(t)}{K_{\text{CEQ}}}\right)^t \tag{4}$$

By setting x_t equal to $2 \times x_0$ and solving for t, we derived the time required for the population to double, t_d.

$$t_d = \log(2)/\log(1 + r_{\text{CEQ}} \times \frac{K_{\text{CEQ}} - \text{fit}_{\text{CEQ}}(t)}{K_{\text{CEQ}}}) \tag{5}$$

where t_d represents the time for the population to double, r_{CEQ} is the estimate of r_1 derived by fitting equation (2) to the median CEQs for each time point,

K_{CEQ} is the estimate of K_1 derived by fitting equation (2) to the median CEQs for each time point and $fit_{CEQ}(t)$ (with fit as the fitted variable) is the modeled population size of CEQs derived from equation (2). Equation (5) was used to determine both average and instantaneous doubling times, where instantaneous doubling times were estimated for each time point from $t(1)$ to $t(600)$, and average doubling time for a given time point was estimated by summing across all previous instantaneous doubling times and dividing by the total time elapsed. For estimates of doubling after fit_{CEQ} exceeds K_{CEQ}, K_{CEQ} was set as the maximum value for fit_{CEQ}; the modeled value of fit_{CEQ} actually fluctuates slightly around K_{CEQ}, consistent with logistic models of population growth. As instantaneous estimates of slope fail to capture the stable nature of population size, K_{CEQ} was substituted for these time points. Notably, the estimates of population growth derived from equation (2) parameterized by CEQ data were used to derive doubling times, which are applicable to the viable population (CFU).

Statistical analyses. Pairwise comparison was performed by Student's *t*-test for normally distributed data, and nonparametric equivalent analyses were used for non-normally distributed data. Normal distribution was confirmed by D'Agostino-Pearson omnibus test. To determine statistically significant differences in CFU, groups (or animals) were compared using the Kruskal-Wallis one-way analysis of variance function in Prism with *post hoc* pair-wise Mann-Whitney comparison of each possible pair of groups and Dunn's correction for determining significance. Percentage sterilization was compared between animals with active disease and those with latent infection, with significance determined by Student's *t*-test. Statistical computations were performed with Prism software (GraphPad, La Jolla, California).

To determine statistically significant differences in CEQs, groups were compared using the Kruskal-Wallis one-way analysis of variance function in Prism (GraphPad, La Jolla, California) with *post hoc* pair-wise Mann-Whitney comparison of each possible pair of groups and Dunn's correction for determining significance. To determine correlates between time, CFU/CEQ and histopathological characteristic of each lesion, these data were correlated using the 'corrcoef' function in Matlab (Natick, Massachusetts) and plotted according to strength of correlation. CFU/CEQ were log normalized before analysis.

17. Lin, P.L. *et al.* Radiologic responses in cynomolgus macaques for assessing tuberculosis chemotherapy regimens. *Antimicrob. Agents Chemother.* **57**, 4237–4244 (2013).
18. Kennedy, H.E., Vandiviere, H.M. & Willis, H.S. The effects of extended incubation on propagability of tubercle bacilli. *Am. Rev. Tuberc.* **77**, 802–814 (1958).
19. 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 (1956).
20. Keane, J. *et al.* Tuberculosis associated with infliximab, a tumor necrosis factor α-neutralizing agent. *N. Engl. J. Med.* **345**, 1098–1104 (2001).

doi:10.1038/nm.3412