

RESEARCH ARTICLE

# Systems genetics uncover new loci containing functional gene candidates in *Mycobacterium tuberculosis*-infected Diversity Outbred mice

Daniel M. Gatti<sup>1</sup>, Anna L. Tyler<sup>1</sup>, J Matthew Mahoney<sup>1</sup>, Gary A. Churchill<sup>1</sup>, Bulent Yener<sup>2</sup>, Deniz Koyuncu<sup>2</sup>, Metin N. Gurcan<sup>3</sup>, MK Khalid Niazi<sup>3</sup>, Thomas Tavorlara<sup>3</sup>, Adam Gower<sup>4</sup>, Denise Dayao<sup>5</sup>, Emily McGlone<sup>5</sup>, Melanie L. Ginese<sup>5</sup>, Aubrey Specht<sup>5</sup>, Anas Alsharaydeh<sup>6</sup>, Philippe A. Tessier<sup>7</sup>, Sherry L. Kurtz<sup>8</sup>, Karen L. Elkins<sup>8</sup>, Igor Kramnik<sup>9</sup>, Gillian Beamer<sup>6\*</sup>

**1** The Jackson Laboratory, Bar Harbor, Maine, United States of America, **2** Rensselaer Polytechnic Institute, Troy, New York, United States of America, **3** Wake Forest University School of Medicine, Winston Salem, North Carolina, United States of America, **4** Clinical and Translational Science Institute, Boston University, Boston, Massachusetts, United States of America, **5** Tufts University Cummings School of Veterinary Medicine, North Grafton, Massachusetts, United States of America, **6** Texas Biomedical Research Institute, San Antonio, Texas, United States of America, **7** Department of Microbiology and Immunology, Laval University School of Medicine, Quebec, Canada, **8** Center for Biologics Evaluation and Research, Food and Drug Administration, Silver Spring, Maryland, United States of America, **9** National Emerging Infectious Diseases Laboratories, Boston University, Boston, Massachusetts, United States of America

\* [GBeamer@txbiomed.org](mailto:GBeamer@txbiomed.org)



**OPEN ACCESS**

**Citation:** Gatti DM, Tyler AL, Mahoney JM, Churchill GA, Yener B, Koyuncu D, et al. (2024) Systems genetics uncover new loci containing functional gene candidates in *Mycobacterium tuberculosis*-infected Diversity Outbred mice. *PLoS Pathog* 20(6): e1011915. <https://doi.org/10.1371/journal.ppat.1011915>

**Editor:** Padmini Salgame, New Jersey Medical School, UNITED STATES

**Received:** December 22, 2023

**Accepted:** April 17, 2024

**Published:** June 11, 2024

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.ppat.1011915>

**Copyright:** © 2024 Gatti et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Data has been deposited in Gene Expression Omnibus (GEO), and assigned Series ID GSE179417.

## Abstract

*Mycobacterium tuberculosis* infects two billion people across the globe, and results in 8–9 million new tuberculosis (TB) cases and 1–1.5 million deaths each year. Most patients have no known genetic basis that predisposes them to disease. Here, we investigate the complex genetic basis of pulmonary TB by modelling human genetic diversity with the Diversity Outbred mouse population. When infected with *M. tuberculosis*, one-third develop early onset, rapidly progressive, necrotizing granulomas and succumb within 60 days. The remaining develop non-necrotizing granulomas and survive longer than 60 days. Genetic mapping using immune and inflammatory mediators; and clinical, microbiological, and granuloma correlates of disease identified five new loci on mouse chromosomes 1, 2, 4, 16; and three known loci on chromosomes 3 and 17. Further, multiple positively correlated traits shared loci on chromosomes 1, 16, and 17 and had similar patterns of allele effects, suggesting these loci contain critical genetic regulators of inflammatory responses to *M. tuberculosis*. To narrow the list of candidate genes, we used a machine learning strategy that integrated gene expression signatures from lungs of *M. tuberculosis*-infected Diversity Outbred mice with gene interaction networks to generate scores representing functional relationships. The scores were used to rank candidates for each mapped trait, resulting in 11 candidate genes: *Ncf2*, *Fam20b*, *S100a8*, *S100a9*, *Itgb5*, *Fstl1*, *Zbtb20*, *Ddr1*, *Ier3*, *Vegfa*, and *Zfp318*. Although all candidates have roles in infection, inflammation, cell migration, extracellular matrix remodeling, or intracellular signaling, and all contain single nucleotide polymorphisms (SNPs), SNPs in only four genes (*S100a8*, *Itgb5*, *Fstl1*, *Zfp318*) are predicted to have deleterious effects on protein functions. We performed methodological and candidate validations

**Funding:** • The work was supported by funding from the National Institutes of Health R21 AI115038 (GB); the National Institutes of Health R01 HL145411 (GB); the American Lung Association Biomedical Research Grant RG-349504 (GB); Tufts University's Cummings School of Veterinary Medicine Summer Research Program (AS), and the National Institutes of Health HL126066 (IK). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

to (i) assess biological relevance of predicted allele effects by showing that Diversity Outbred mice carrying PWK/PhJ alleles at the H-2 locus on chromosome 17 QTL have shorter survival; (ii) confirm accuracy of predicted allele effects by quantifying S100A8 protein in inbred founder strains; and (iii) infection of C57BL/6 mice deficient for the *S100a8* gene. Overall, this body of work demonstrates that systems genetics using Diversity Outbred mice can identify new (and known) QTLs and functionally relevant gene candidates that may be major regulators of complex host-pathogens interactions contributing to granuloma necrosis and acute inflammation in pulmonary TB.

### Author summary

We investigated the genetic basis of susceptibility to *Mycobacterium tuberculosis* using Diversity Outbred mice, a mouse population suited for studying complex genotype-phenotype relationships. We found five new genetic loci and three known loci. Three loci associated with multiple correlated disease traits likely contain genes that are major regulators of host inflammatory responses which favor *M. tuberculosis* growth. Further, these three loci contain four gene candidates with single nucleotide polymorphisms that are predicted to have deleterious effects upon protein functions.

## Introduction

The number of humans who develop active pulmonary tuberculosis (TB) is minor compared to those who eliminate or control *Mycobacterium tuberculosis* (5–10% vs 90–95%), yet morbidity and mortality from TB remain high. COVID-19 mortality temporarily surpassed global TB mortality, but TB has remained in the top two leading causes of death due to an infectious disease for decades, killing more people than HIV/AIDS and malaria. Pulmonary TB is the most common and most contagious form of TB, with mortality rates >40% if untreated and higher if caused by antibiotic resistant *M. tuberculosis* [1–8]. Human responses to *M. tuberculosis* range from fulminant pulmonary TB that develops within weeks to lifelong control or complete clearance of bacilli [9–11]. Further, a body of evidence shows an interesting paradox: Immune competence is necessary to restrict *M. tuberculosis* growth [12], but is not sufficient to prevent disease [13].

The variable responses to *M. tuberculosis* and lack of single gene defects in most patients indicate a complex genetic basis for pulmonary TB, and this has been investigated by linkage association mapping, genome wide association studies, and other methods, recently reviewed [14–17]. Reviews identify knowledge gaps attributable to using laboratory mouse strains that do not replicate key disease traits of human pulmonary TB, e.g., granuloma necrosis [18–23]. To address these gaps, we and others investigate genetic contribution, molecular pathogenesis, and vaccine efficacy in the Diversity Outbred mouse population, panels of Collaborative Cross inbred mice and their progeny, some of which do develop human-like pulmonary TB [24–29]. These mice are valuable resources to model complex genotype-phenotype associations; tools to dissect the genetic basis of disease; and a means to assess effects of candidate genetic polymorphisms *in vivo*.

The Diversity Outbred mouse population originated by breeding eight inbred founder strains together, resulting in a reproducible experimental population with balanced allele frequencies of one-eighth across the genome [30]. This is important for genetic mapping studies

because low allele frequencies in natural populations can diminish power and increase false positive associations [31]. Further, Diversity Outbred mice carry over 40 million variants [32], some of which alter regulatory elements, splice sites, and protein-coding sequences. This defined genetic architecture allows rigorous investigation of genotype-phenotype association in context of *M. tuberculosis* infection. Here, we selected phenotypes of pulmonary TB (i.e., quantified traits) for biological relevance, many of which are shared by humans, non-human primates, and Diversity Outbred mice [33]. Traits were selected to span clinical, granuloma, microbiological, and immune/inflammatory mediators that are known to positively correlate with disease, or are investigated as diagnostic biomarkers, or have been shown to activate or negatively regulate immune/inflammatory responses to *M. tuberculosis*. Briefly, the clinical trait of disease was weight loss. The granuloma trait was necrosis. The microbiological trait was *M. tuberculosis* lung burden. Chemokine correlates of disease were CXCL1, CXCL2, CXCL5 [24,34]. TH1 proinflammatory cytokines and acquired immune mediators were tumor necrosis factor (TNF) [35], interferon-gamma (IFN-gamma) [36], and interleukin (IL)-12 [37]. Biomarker candidates were S100A8, matrix metalloproteinase (MMP)8, and vascular endothelial growth factor (VEGF) [34] and the immune regulatory cytokine was IL-10 [38].

To find genetic loci associated with pulmonary TB traits, we used quantitative trait locus (QTL) mapping. Next we ranked candidate genes within the *Diversity Outbred tuberculosis susceptibility (Dots)* loci that were associated with correlated, colocalized traits by using a machine learning algorithm [39,40] to find genes functionally related to the mapped traits and the fit models scored each candidate [41]. All candidates contain a variety of SNPs as annotated in Mouse Variation Registry (MVAR). Seven of the eleven candidates contain missense SNPs in protein coding regions, and of those, the SNPs in four candidates (*S100a8*, *Itgb5*, *Fstl1*, and *Zfp318*) are predicted to have deleterious consequences on protein functions. Published *in vitro* and *in vivo* studies show roles for three candidates (*Itgb5*, *Fstl1*, *S100a8*) in bacterial lung infections including *M. tuberculosis* [34,42–47], although questions remain. To our knowledge, the other eight candidates have no published roles in *M. tuberculosis* infection but have been shown in other systems to contribute to cell stress responses, signaling pathways, adhesion and migration; extracellular matrix synthesis, tissue remodeling and angiogenesis; immune cell metabolism; macrophage inflammatory responses; and viral hepatitis [48–60]. Overall, ten candidate genes have roles in innate immune responses suggesting that genetically controlled responses of epithelial and endothelial cells, neutrophils, and monocytes, macrophages to *M. tuberculosis* bacilli are the primary drivers of susceptibility to *M. tuberculosis* and to disease progression in pulmonary TB. Only one candidate (*Zfp318*) has a direct role in acquired, antigen-specific immunity.

## Methods

### Ethics statement

Tufts University's Institutional Animal Care and Use Committee (IACUC) approved this work under protocols G2012-53; G2015-33; G2018-33; and G2020-121. Tufts University's Institutional Biosafety Committee (IBC) approved this work under registrations: GRIA04; GRIA10; GRIA17, and 2020-G61.

### Mice

Female Diversity Outbred mice from generations 15 16, 21, 22, 34, 35, 37 and 42 and the inbred founder strains: A/J, C57BL/6J, 129S1/SvImJ), NOD/LtJ, NZO/HILtJ, CAST/EiJ, PWK/PhJ, and WSB/EiJ mice were purchased from The Jackson Laboratory (Bar Harbor, ME) and group housed (n = 5–7 mice per cage) on Innovive (San Diego, CA) or Allentown Inc

(Allentown, NJ) ventilated, HEPA-filtered racks in the New England Regional Biosafety Laboratory (Tufts University, Cummings School of Veterinary Medicine, North Grafton, MA) or at The Ohio State University, Columbus, OH. The light cycle was 12 hours of light; 12 hours of dark. Two breeding pairs of female and male C57BL/6 inbred mice carrying null mutation for *S100a8* gene were a gift from Dr. Philippe Tessier, Department of Microbiology and Immunology, Faculty of Medicine, Université Laval. After quarantine, breeders were used to establish a colony of *S100a8* homozygous knock out (KO), heterozygous (HET) and wild-type (WT) C57BL/6 inbred mice, and genotypes confirmed (Transnetyx, Cordova, TN). Mice were housed in disposable sterile caging or re-usable autoclaved caging containing sterile corn-cob bedding, with sterile paper nestlets (Scotts Pharma Solutions, Marlborough, MA), and/or sterile enrichment paperboard or plastic “houses.” Cages were changed every other week or sooner if soiled. Mice were provided with sterile mouse chow (Envigo, Indianapolis, IA) and sterile, acidified water *ad libidum*.

### ***M. tuberculosis* Aerosol Infection**

Female Diversity Outbred mice and inbred founder strains were infected with aerosolized *M. tuberculosis* strain Erdman bacilli using a custom-built CH Technologies system [24,34,61] or a Glas-col (Terre Haute, ID) system [62,63] between eight and ten weeks of age. Male and female C57BL/6 *S100a8* KO, HET, and WT mice were infected between eight and sixteen weeks of age. For each aerosol infection, the retained lung dose was determined by euthanizing a cohort of four to twelve mice 24 hours after exposure, homogenizing the entire lungs in 5mL sterile phosphate buffered saline, and plating the entire homogenate onto OADC-supplemented 7H11 agar. After 3–4 weeks at 37°C, *M. tuberculosis* colony forming units were counted. Mice were infected with ~100 colony forming units in the first two experiments, and ~25 colony forming units in the subsequent eight experiments.

### **Quantification of TB-related Traits (Phenotyping)**

**Survival.** IACUC protocols disallowed natural death as an endpoint. Therefore, as a proxy of survival, we used the day of euthanasia due to any single criterion: Severe weakness/lethargy; respiratory distress; or body condition score < 2 [64]. We confirmed morbidity was due to pulmonary TB by finding: (i) Large nodular, or severe diffuse lung lesions; (ii) histopathology confirmation of severe granulomatous lung infiltrates; (iii) growth of viable *M. tuberculosis* colonies from lung tissue; and (iv) absence of other diseases. Twenty-one *M. tuberculosis* infected Diversity Outbred mice were excluded due to co-morbidity that developed during the in-life portions.

**Weight loss.** Mice were weighed 1 to 3 days prior to *M. tuberculosis* aerosol infection, at least once per week during infection, and immediately before euthanasia. For each mouse, weight loss was calculated as the percent loss from peak body weight.

**Lung granuloma necrosis.** Immediately after euthanasia, lung lobes were removed and inflated and fixed in 10% neutral buffered formalin (5–10 mL per lobe), processed, and embedded in paraffin, sectioned at 5µm, and stained with hematoxylin and eosin with or without carbol fuchsin for acid-fast bacilli at Tufts University, Cummings School of Veterinary Medicine, Core Histology Laboratory (North Grafton, MA). Hematoxylin and eosin-stained glass slides were magnified 400 times and digitally scanned by Aperio, LLC (Sausalito, CA) ScanScope scanners at 0.23 microns per pixel at The Ohio State University’s Comparative Pathology and Mouse Phenotyping Shared Resources Core resource (Columbus, OH) or by Aperio, LLC (Sausalito, CA) AT2 scanners at 0.23 microns per pixel at Vanderbilt University Medical Center’s Digital Histology Shared Resource (Nashville, TN). Lung granuloma necrosis was

quantified in one lung lobe per mouse by our previously validated, deep learning image analysis method [65] and reported here as a ratio of granuloma necrosis per lung tissue section area.

*M. tuberculosis* lung burden. Immediately after euthanasia, 2 or 3 lung lobes were removed from each mouse and homogenized in sterile phosphate buffered saline (1mL per lobe), serially diluted, plated onto OADC-supplemented 7H11 agar, incubated at 37°C for 3–4 weeks, after which colonies were counted, and *M. tuberculosis* lung burden in the lungs was calculated as described [66].

*Lung cytokines and chemokines.* Lung homogenates were stored at -80°C until the experiment ended. Lung homogenates were then thawed overnight at 4°C, serially diluted and tested for CXCL5, CXCL2, CXCL1, tumor necrosis factor (TNF), matrix metalloproteinase 8 (MMP8), S100A8, interferon-gamma (IFN- $\gamma$ ), interleukin (IL)-12p40, I-L12p70, IL-10, and vascular endothelial growth factor (VEGF) by sandwich ELISA using antibody pairs and standards from R&D Systems (Minneapolis, MN), Invitrogen (Carlsbad, CA), eBioscience (San Diego, CA), or BD Biosciences (San Jose, CA, USA), per kit instructions. Lung homogenate ELISA results from five of the experiments using Diversity Outbred mice have been published and analyzed for biomarkers previously [34].

## Phenotype correlation

We took the log of each phenotype after adding one (to ensure that zero was not converted to negative infinity) and regressed out the effect of the experimental batch. We then standardized the residuals and estimated the Pearson correlation between all pairs of phenotypes.

## Gene expression

Gene expression profiling was performed on a subset of 117 Diversity Outbred mice representing the spectrum of observed phenotypes. Detailed selection description, additional analyses, and released datasets are described in [34,67]. Briefly, one lung lobe was homogenized in TRIzol, stored at -80°C, and RNA was extracted using Pure Link mini-kits (Life Technologies, Carlsbad, CA). Boston University's Microarray and Sequencing Resource Core Facility (Boston, MA) confirmed quality and quantity were sufficient for microarray analyses. Mouse Gene 2.0 ST CEL files were normalized to produce gene-level expression values using the implementation of the Robust Multiarray Average (RMA) in the affy package (version 1.62.0) included in the Bioconductor software suite and an Entrez Gene-specific probeset mapping (24.0.0) from the Molecular and Behavioral Neuroscience Institute (Brainarray) at the University of Michigan. Array quality was assessed by computing Relative Log Expression (RLE) and Normalized Unscaled Standard Error (NUSE) using the affyPLM package (version 1.59.0). The CEL files were also normalized using Expression Console (build 1.4.1.46) and the default probesets defined by Affymetrix to assess array quality using an AUC metric computed from sets of negative and positive control probes; all samples used in this analysis had an AUC > 0.8. To remove microarray probes that intersected with Diversity Outbred SNPs, we intersected the Diversity Outbred founder strain SNPs [68] with the vendor-provided probes and removed probes containing SNPs. All microarray analyses were performed using the R environment for statistical computing (version 3.6.0). A related microarray dataset and other secondary analyses have been published elsewhere [34,67] and deposited in Gene Expression Omnibus (GEO), and assigned Series ID GSE179417.

## Genotyping

We collected tail tips from each Diversity Outbred mouse and sent them to Neogen (Lincoln, NE) for genomic DNA isolation and genotyping. Neogen genotyped the mice on the Illumina

GigaMUGA platform, which contains 143,259 markers [69]. Genotypes of *S100a8* KO, HET, and WT C57BL/6 inbred mice were confirmed by polymerase chain reaction (TransnetYX, Cordova, TN).

### Haplotype reconstruction and SNP imputation

We used 137,302 GigaMUGA marker positions located on the autosomes and chromosome X found at [https://github.com/kbroman/MUGAarrays/blob/main/UWisc/gm\\_uwisc\\_v1.csv](https://github.com/kbroman/MUGAarrays/blob/main/UWisc/gm_uwisc_v1.csv) and the R package *qtl2* to reconstruct the Diversity Outbred haplotypes using the founder and Diversity Outbred allele calls, and used the haplotype reconstructions to impute the founder SNPs onto the Diversity Outbred genomes [70].

### Quantitative Trait Locus (QTL) mapping

We performed multiple rounds of genetic mapping to identify the least noisy dataset, ensure appropriate sample size for statistical power, and generate robust results. The final QTL mapping results include all available genotype and phenotype data from Diversity Outbred mice that survived *M. tuberculosis* infection for up to 250 days because this dataset produced the most statistically rigorous results, and minimized confounding effects of age-related comorbidities (e.g., lung tumors, lymphoma, uterine lesions). We used *qtl2* [70] to perform linkage mapping using the founder haplotypes and association mapping using the imputed SNPs. We calculated the kinship between mice using the leave-one-chromosome-out method, which excludes the current chromosome in kinship calculations [71]. We standardized each phenotype and mapped with the Diversity Outbred outbreeding generation as an additive covariate and used the linear mixed-effects model with one kinship matrix per chromosome. We estimated the genome-wide significance thresholds by permuting the samples 1,000 times and performed a genome scan with each permutation. We retained the maximum  $\log_{10}$  of the odds ratio (LOD) score from each permutation and estimated the genome-wide significance threshold of 7.6 from the 95<sup>th</sup> percentile of the empirical distribution of maximum LOD scores under permutation. We estimated the support interval around each peak using the 95% Bayesian Credible Interval.

For each peak with a LOD score above the genome-wide threshold  $> 7.6$ , we then searched for peaks associated with other traits that had LOD scores  $> 6$  and confidence intervals that overlapped [72]. Our rationale was that the probability that a peak is biologically relevant, given that another trait has a co-located peak, is greater than the probability that a peak is significant with no prior evidence.

### Candidate gene selection

Within each QTL interval, we imputed the founder SNPs onto the Diversity Outbred mouse genomes using *qtl2* and performed association mapping. We selected the SNPs that were within a 1 LOD drop of the peak SNP in the QTL interval and filtered them to retain ones with missense, splice, or stop codon effects as annotated by the Sanger Mouse Genome Project [68]. We considered the genes in which these polymorphisms occurred as candidate causal genes for the associated trait(s).

### Trait-related gene sets

Because causal variants within a QTL may exert their influence through mechanisms other than gene expression, identifying differentially expressed genes within the QTL may be insufficient for ranking causal genes. Here we took an alternative approach and ranked candidates in

each QTL based on their predicted association with the mapped traits. To do this, we trained a Support Vector Machine (SVM) classifier to classify trait-related genes, and then used the trained SVM to score each positional candidate gene as trait-related or not-trait-related. We defined the training set of trait-related genes for the SVM as those genes that were highly correlated to the measured trait using the gene expression data described above. We calculated the Pearson correlation between the abundance of each transcript, and each physiological trait using rank Z normalized gene expression and traits. For each trait, we defined the training set of trait-related genes as the 500 genes with the largest magnitude Pearson correlation to the trait. We have made these gene lists available as a set of zipped text files in [S1 File](#).

### Support Vector Machine classifier training

We trained SVMs to classify genes in each gene list as trait-related using features derived from the Functional Network of Tissues in Mouse [41]. The nodes in this network are genes, and the edges between them are weights between 0 and 1 that predict the likelihood that each pair of genes is annotated to the same Gene Ontology (GO) term or KEGG pathway [39]. Values closer to one indicate more certainty that the genes are more likely to be annotated to the same GO term or KEGG pathway and thus functionally related. The weights were derived using Bayesian integration of data sets from numerous sources of data, including gene expression, protein-protein interaction data, and phenotype annotations [41]. We used the top edges of the mouse lung network downloaded on March 31, 2021, from <http://fntm.princeton.edu>.

### Application of Support Vector Machine classifiers to identify genes functionally related to traits

We used SVMs to classify each positional candidate as trait-related or not-trait-related, as described previously [39,40]. Briefly, the expression-derived gene sets for each lung trait served as the *positive labeled set* of genes. We used the R package `e1071` [73] to train SVMs to distinguish this set of genes from a balanced set of genes drawn randomly from the remaining genes in the lung network. The randomly selected genes were the *unlabeled set*. We performed this training 100 times, each time with a new set of random unlabeled genes. The SVMs were trained to distinguish positive labeled genes from unlabeled genes using the connection weights to the positive labeled genes. It is expected that positively labeled genes have strong connections to each other because they are functionally related. It is further expected that randomly drawn genes will be unrelated to the trait and to the positive labeled set and will thus have lower connection weights to the positive labeled genes. The SVM learns to distinguish these two groups of genes, and the resulting model can be used to classify genes that have not been seen before based on their connection weights to the positively labeled genes. We initialized each run by tuning the SVM over a series of cost parameters, starting with the sequence  $10^{25}$  to  $10^2$  by factors of ten, and iteratively narrowing the range until we found a series of eight cost parameters that maximized accuracy. In running each SVM, we used a linear kernel and 10-fold cross-validation.

We calculated the area under the receiver operating characteristic curves (AUC) for each set of trait-related genes as follows. We defined labeled positives (LP) as positive labeled genes that were classified by the SVM as trait related. Unlabeled negatives (UN) were unlabeled genes that were classified by the SVM as not trait related. Unlabeled positives (UP) were unlabeled genes that were classified as trait-related, and labeled negatives (LN) were positive labeled genes that were classified as not trait-related. These terms are conceptually like true/false positive and true/false negative scores. However, because unlabeled genes may not be truly unrelated to the trait, we cannot call them true negatives. Instead, we call them *unlabeled*.

We generated ROC curves using the *Unlabeled Predicted Positive Rate* ( $UPPR = UP/(UP + UN)$ ), which is akin to the false positive rate, and the *labeled positive rate* ( $LPR = LP/(LP + UN)$ ), which is akin to the false negative rate, along a series of SVM scores from the minimum to the maximum. We then calculated the average AUC across all 100 SVMs.

### Positional candidate scoring

After training SVMs for each trait, we scored all positional candidate genes in each QTL, defined as the minimum to the maximum position across a set of overlapping QTLs. Each candidate gene received one score for each trait that mapped to that location. To compare scores across traits, we used the UPPR for each gene at its calculated SVM score. The UPPR varies between 0 and 1, allowing us to compare scores for candidate genes across models. To visually compare across models, we used the  $-\log_{10}(UPPR)$  such that genes with small UPPR (high-confidence) got large positive values. In contrast, the SVM scores cannot be used to compare across models because they are unbounded and vary from model to model. Within each pleiotropic QTL, each gene received a score from each trait that mapped to the QTL.

### Mouse genome build and database versions

We used mouse genome build GRCm38 and SNPs and Indels from the Sanger Mouse Genomes Project, version 7, which uses Ensembl version 97 gene models. We also used and cross-referenced candidates with the Mouse Phenome Database GenomeMUSter, the Mouse Genome Informatics databases [74], and Ensembl's Variant Effect Predictor tool.

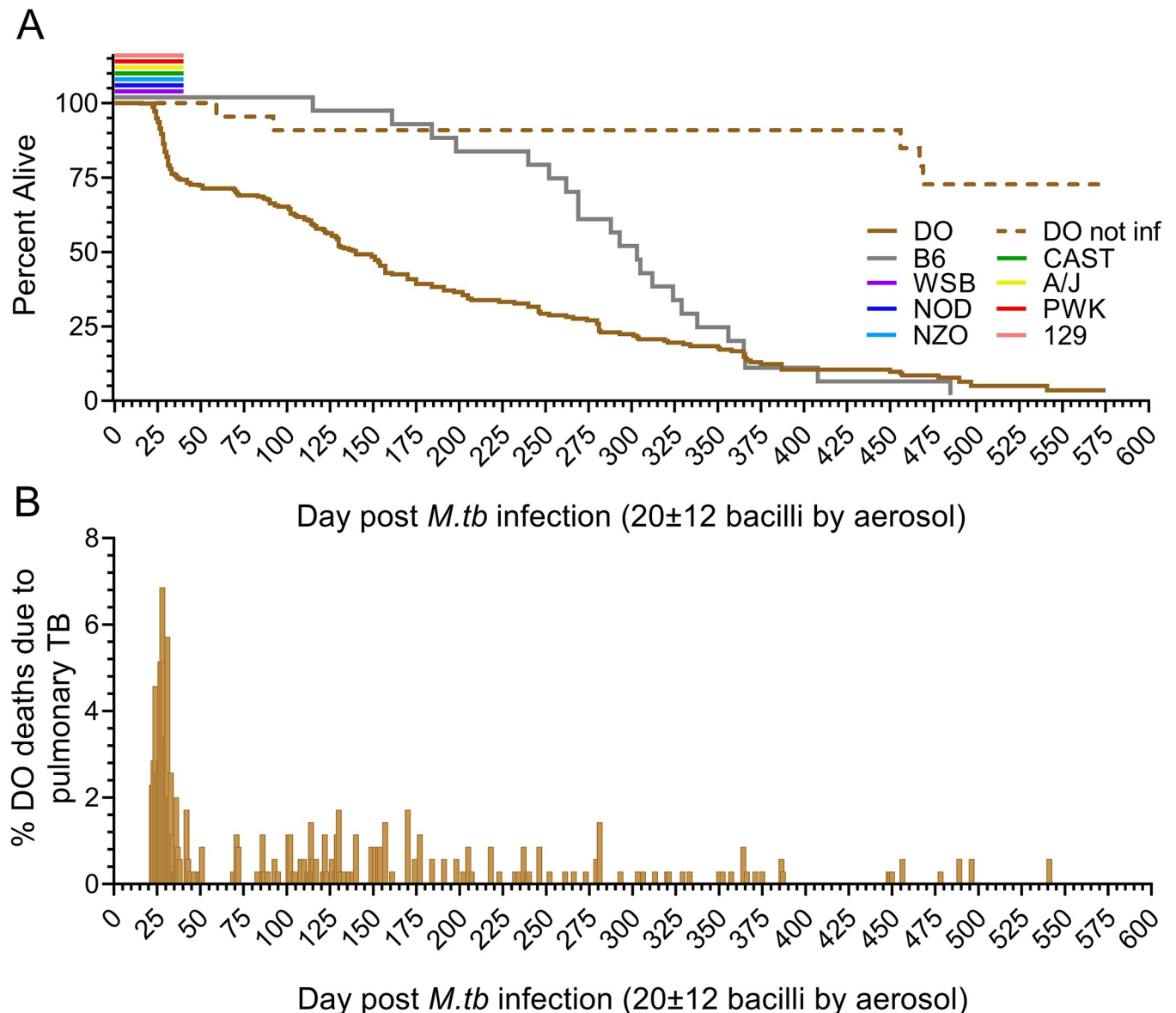
## Results

### Survival and body weight changes

We infected Diversity Outbred mice by aerosol with  $\sim 100$  *M. tuberculosis* colony forming units in the first two experiments ( $N = 167$ ), and  $\sim 25$  colony forming units in the subsequent eight experiments ( $N = 683$ ). Infection reduced survival of Diversity Outbred mice compared to identically housed, age-, gender-, and generation-matched uninfected Diversity Outbred mice and to identically housed age- and gender-matched infected C57BL/6J inbred mice (Fig 1A). Approximately one-third of infected Diversity Outbred mice succumbed prior to 60 days post infection (Fig 1A) reflecting early mortality between 20–56 days that peaked at 30 days (Fig 1B). This supersusceptible fraction of the Diversity Outbred population has been named Progressors [24,33,65,67]. Morbidity in all Progressors was due to pulmonary TB, confirmed by histology, recovery of viable *M. tuberculosis* bacilli from the lungs and absence of other diseases. After the first mortality wave subsided, cumulative survival declined slowly to nearly 600 days with no discernable mortality waves (Fig 1A and 1B). This resistant fraction of Diversity Outbred mice has been named Controllers [24,33,65,67]. The eight founder strains survived at least 40 days of *M. tuberculosis* infection, without early mortality (Fig 1A).

All mice were weighed prior to infection, during infection, and immediately before euthanasia. Non-infected Diversity Outbred mice gained weight until they developed other diseases or were euthanized at the experiment end (S1A Fig). Progressors gained weight for 2–3 weeks, and then lost weight rapidly (S1B Fig). Controllers and C57BL/6J inbred mice gained weight for long and variable durations through about 250 days of infection, and then most but not all slowly lost weight (S1C and S1D Fig). We questioned whether pre-infection body weight influenced differential susceptibility. Retrospective analysis identified no significant differences in pre-infection body weights of non-infected Diversity Outbred mice compared to Progressors; and a significant difference of on average 1.75 gm less in mean body weights of non-infected

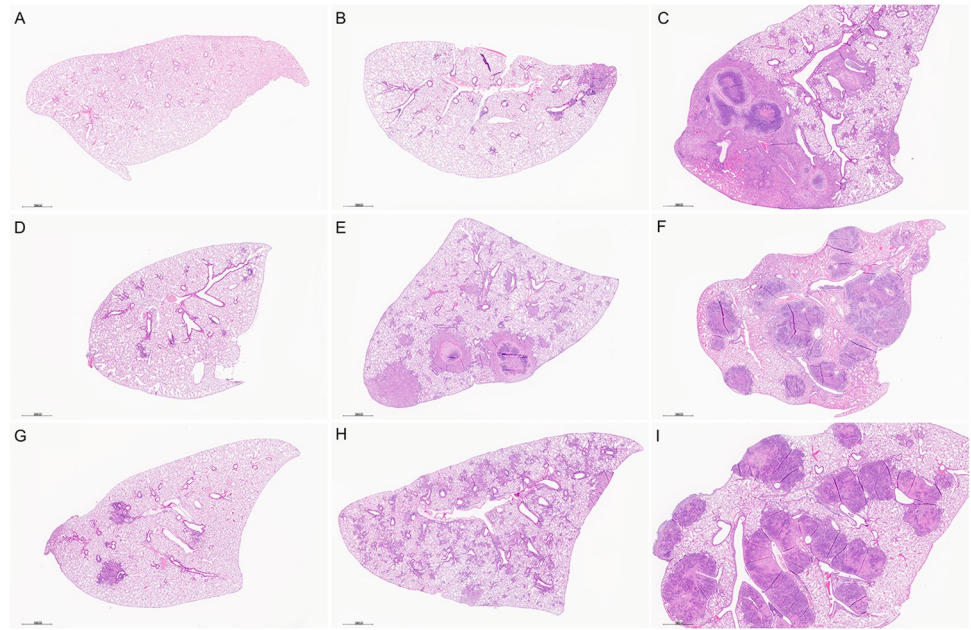




**Fig 1. Mouse survival following exposure to a low dose of aerosolized *M. tuberculosis* strain Erdman.** Diversity Outbred (DO) mice ( $n = 680$ , brown solid line), and the eight inbred founder strains ( $n = 15$  to  $78$ , colored lines) were infected with *M. tuberculosis* strain Erdman bacilli by aerosol. Panel A: Cumulative survival. Approximately 30% of the DO population succumbed to pulmonary TB by 60 days post infection and approximately 70% succumbed to between 60- and 600-days post infection. Of the eight inbred founder strains, survival studies were completed for the C57BL/6J inbred strain; the other seven inbred founder strains were euthanized 40 days after *M. tuberculosis* infection. No inbred founder or non-infected (NI) DO mice ( $n = 53$ , dashed line) showed mortality in that period. Panel B: Daily mortality of *M. tuberculosis*-infected DO mice, highlighting the early wave of mortality that peaked between 25- and 35-days post infection.

<https://doi.org/10.1371/journal.ppat.1011915.g001>

Diversity Outbred mice and Progressors compared to Controllers (S2A Fig). Whether this is spurious or biologically relevant (i.e., heavier pre-infection body weight is partially protective) remains to be determined. S2B, S2C, and S2D Fig show correlations between survival and eight clinical indicators. Seven indicators positively correlated with survival, including pre-infection body weight which was weakly positive. Only one indicator, the rate of body weight loss, negatively correlated with survival.



**Fig 2. Diversity Outbred (DO) mice develop a spectrum of histopathological lung lesions due to *M. tuberculosis* infection.** Lung lobes were formalin-fixed, paraffin-embedded, sectioned, and stained with hematoxylin & eosin. Panel A: Lung section from a non-infected DO mouse. Panels B through I: Lung sections from *M. tuberculosis*-infected DO mice euthanized eight weeks after infection show a spectrum of lesions from mild to severe (upper left to bottom right); focal lesions (e.g., Panel B) to diffuse infiltration (Panel H); and include necrotizing (Panels C, E, F, I) and non-necrotizing granulomas (Panels B, D, G, H). Low magnification (15X).

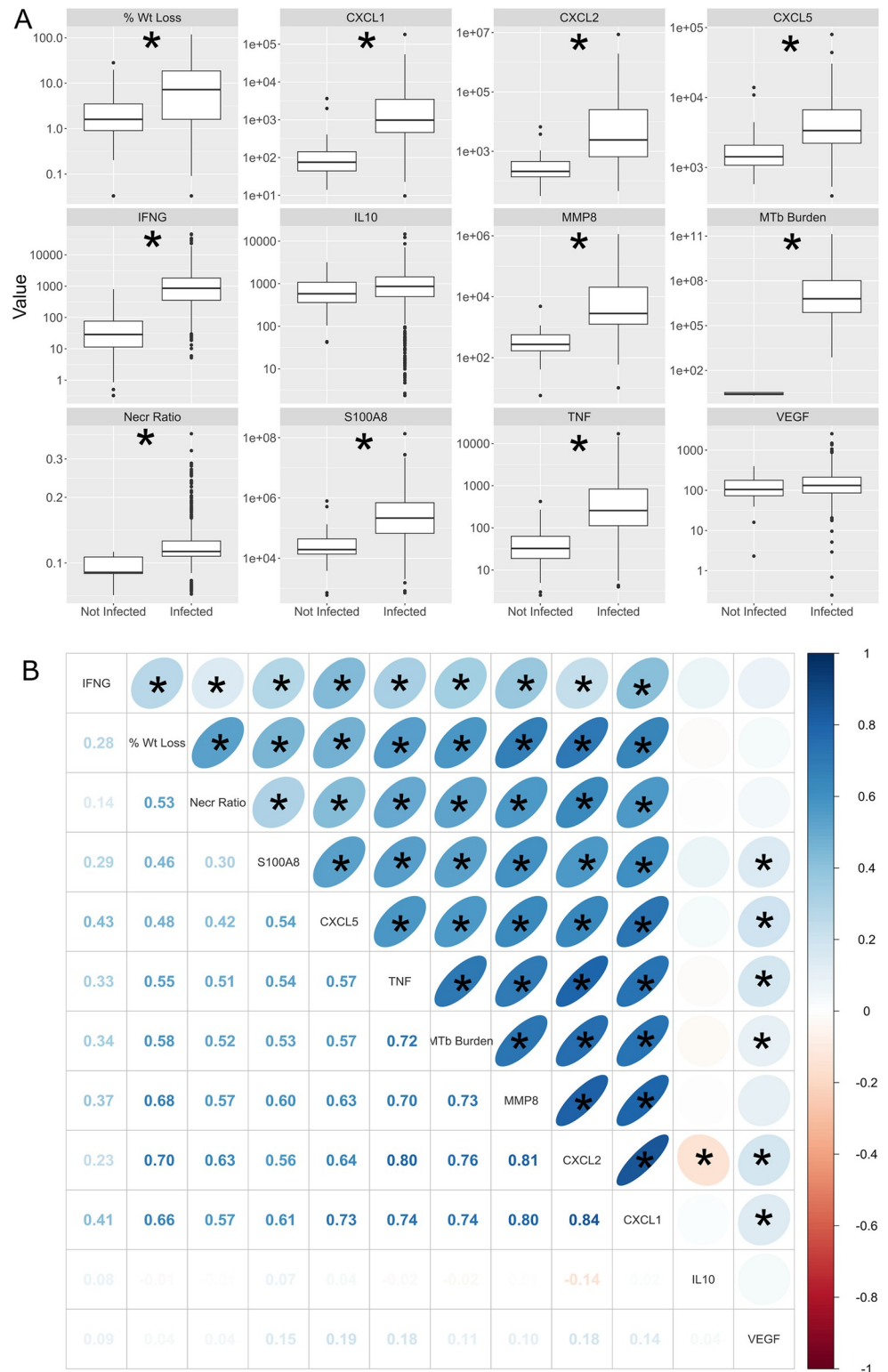
<https://doi.org/10.1371/journal.ppat.1011915.g002>

### Lung histology and automated image analysis of granuloma necrosis

By eight weeks of *M. tuberculosis* infection, Diversity Outbred mice showed a spectrum of lung lesions visible at low magnification (Fig 2) with variation in severity (minimal to marked); distribution of cellular infiltrates (focal, multifocal, and diffuse); and granuloma content (e.g., necrotizing, and non-necrotizing, shown in S3 Fig at higher magnification). Additional lesions included fibrin thrombosis with alveolar septal necrosis; cavities with peripheral fibrosis; foamy and multinucleated macrophages with cholesterol clefts; formation of secondary lymphoid follicles; alveolar septal fibrosis; and intra- and extracellular *M. tuberculosis* bacilli described elsewhere [19,24,28,67,75–77]. Since granuloma necrosis is a key feature of pulmonary TB in humans, we focused on this feature of disease and used our automated image analysis method [65] to quantify granuloma necrosis in lung tissue sections. The ratio of necrotic to non-necrotic lung tissue was then input as a trait for genetic mapping.

### Quantification of lung traits: *M. tuberculosis* burden and immune responses

We quantified *M. tuberculosis* lung burden by counting colonies from lung tissue homogenate and used the remaining lung homogenate to quantify cytokines and chemokines by ELISA [24,34]. Most lung traits were significantly greater in mice infected with *M. tuberculosis* compared to non-infected mice (Fig 3A), including neutrophil and monocyte/macrophage chemokines (CXCL1, CXCL2, CXCL5); mediators of innate immunity (S100A8, Tumor Necrosis Factor (TNF), interleukin (IL)-10, matrix metalloproteinase-8 (MMP8)); mediators of acquired immunity (interferon-gamma (IFN- $\gamma$ )) and *M. tuberculosis* burden. Pairwise Pearson correlation of all traits in infected mice showed that all the lung traits except IL10 and VEGF



**Fig 3. *M. tuberculosis* infection of Diversity Outbred (DO) mice induces clinical, microbiological, and immune/inflammatory phenotypes that are positively and significantly correlated.** (A) Boxplots of phenotypes in non-infected (n = 77) and infected (n = 930) mice at euthanasia, which ranged from 5 to 553 days post-infection. Each box shows the median (central line) and the inter-quartile range in the box. The whiskers show 90% of the data and the dots beyond the whiskers show data points outside of the 90th percentile of the distribution. Significant differences at a

Bonferroni corrected p-value of 0.004 are shown with an asterisk. (B) Pearson correlation of selected phenotypes with  $n = 930$  infected mice. Each cell in the upper triangle contains an ellipse in which the color and shape show the strength of the correlation. Positive correlations are shown in blue and negative correlations are shown in red. Strong correlations are shown by narrow ellipses, and weak correlations are shown by round ellipses. The more circular the ellipse, the closer the correlation is to zero. Numerical values of the Pearson correlation are shown in the lower triangle and range from -0.14 to 0.84. Phenotypes are ordered by hierarchical clustering of the similarity in correlations. Absolute Pearson correlations greater than 0.105 are significant at a Bonferroni-corrected p-value of 0.0007 and are shown with an asterisk.

<https://doi.org/10.1371/journal.ppat.1011915.g003>

positively correlated with each other (Fig 3B). Like previous findings in a small study of Diversity Outbred mice [24], correlations were strongest between *M. tuberculosis* lung burden and mediators of acute neutrophilic inflammation, innate immunity, and extracellular matrix degradation: CXCL1, CXCL2, TNF, and MMP8 with a mean correlation of 0.75.

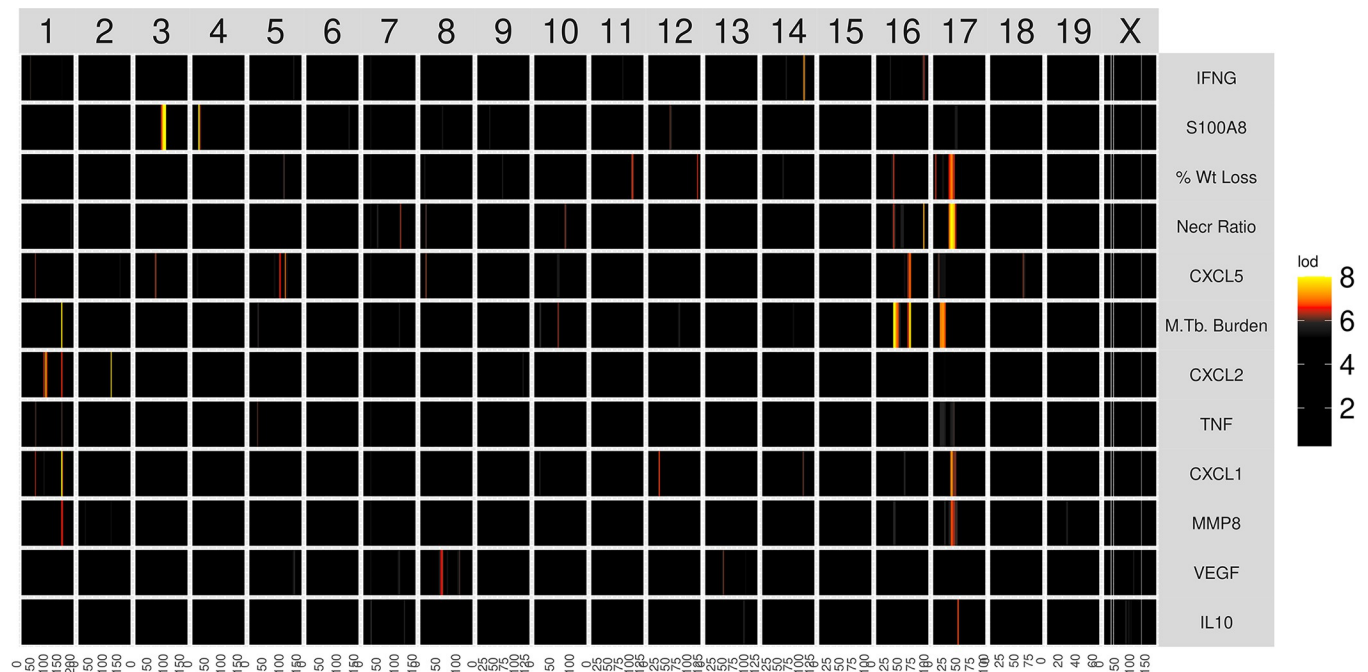
### Overview of genetic mapping and gene prioritization within QTLs

We performed linkage mapping on each trait by regressing each on the additive founder allele dosage at each locus using the R package qtl2 [70]. We selected peaks with a permutation-derived significance threshold of 7.62 ( $p_{GW} \leq 0.05$ ) and found eight peaks associated with multiple traits on chromosomes 1, 2, 3, 4, 16, and 17 (Table 1 and Fig 4). We observed that correlated traits colocalized to shared QTLs, and had similar patterns of allele effects, so we used a two-step procedure in which we recorded the confidence interval for peaks with a LOD  $\geq 7.62$  and then looked for peaks of colocalized traits with a LOD  $\geq 6.0$  ( $p_{GW} \leq 0.6$ ). We reasoned that once we had found the first significant peak for one trait, the threshold for colocalized peaks with the same pattern of founder allele effects should be less, allowing refinement of the loci. Fig 5 shows a flow diagram of the types of input data for genetic mapping to identify QTLs, and the subsequent methods of gene prioritization.

**Table 1. Quantitative Trait Locus (QTL) mapping results of lung traits in *M. tuberculosis*-infected DO mice.** Six QTLs with LOD scores  $> 7.62$  ( $p_{GW} < 0.05$ ) and two with lower thresholds were identified by linkage mapping. Correlated traits with lower LOD thresholds ( $\geq 6$ ) mapped to the same positions. The columns from left to right show: Lung trait(s), chromosome, LOD score of the peak, the position of the peak in megabases (Mb), lower and upper confidence intervals (C.I.) in Mb, and the width of the confidence interval in Mb. Five QTLs are new, and two QTLs are not new. *Dots3* overlaps with the QTL *tbs1*, previously identified by crossing A/Sn and I/St inbred mouse strains. *Dots7* and *Dots8* overlap with *sst5* and *sst6*, previously identified by crossing C3HeB/Fej and C57BL/6J inbred mouse strains.

QTL	Trait(s)	Chr	LOD	Peak	C.I. Lo	C.I. Hi	C.I. Width	New
<i>Dots1</i>	<i>M.tb</i> burden CXCL1 CXCL2 MMP8	1	7.68 7.87 $\geq 6$ $\geq 6$	155.36	154.25	156.71	2.45	Yes
<i>Dots2</i>	CXCL2	2	7.69	118.65	118.05	118.86	0.82	Yes
<i>Dots3</i>	S100A8	3	16.57	90.69	90.52	92.02	1.50	No
<i>Dots4</i>	S100A8	4	7.64	22.43	22.18	23.79	1.61	Yes
<i>Dots5</i>	<i>M.tb</i> burden Wt loss Granuloma necrosis	16	8.45 $\geq 6$ $\geq 6$	38.28	33.28	43.28	10.00	Yes
<i>Dots6</i>	<i>M.tb</i> burden CXCL5	16	6.68	52.23	37.97	57.67	19.70	Yes
<i>Dots7</i>	<i>M.tb</i> burden	17	$\sim 6.5$	$\sim 20$	$\sim 12.5$	$\sim 25$	$\sim 12.5$	No
<i>Dots8</i>	Granuloma necrosis Wt loss CXCL1 MMP8 IL-10	17	8.12 $\geq 6$ $\geq 6$ $\geq 6$ $\geq 6$	35.02	33.94	41.06	7.12	No

<https://doi.org/10.1371/journal.ppat.1011915.t001>

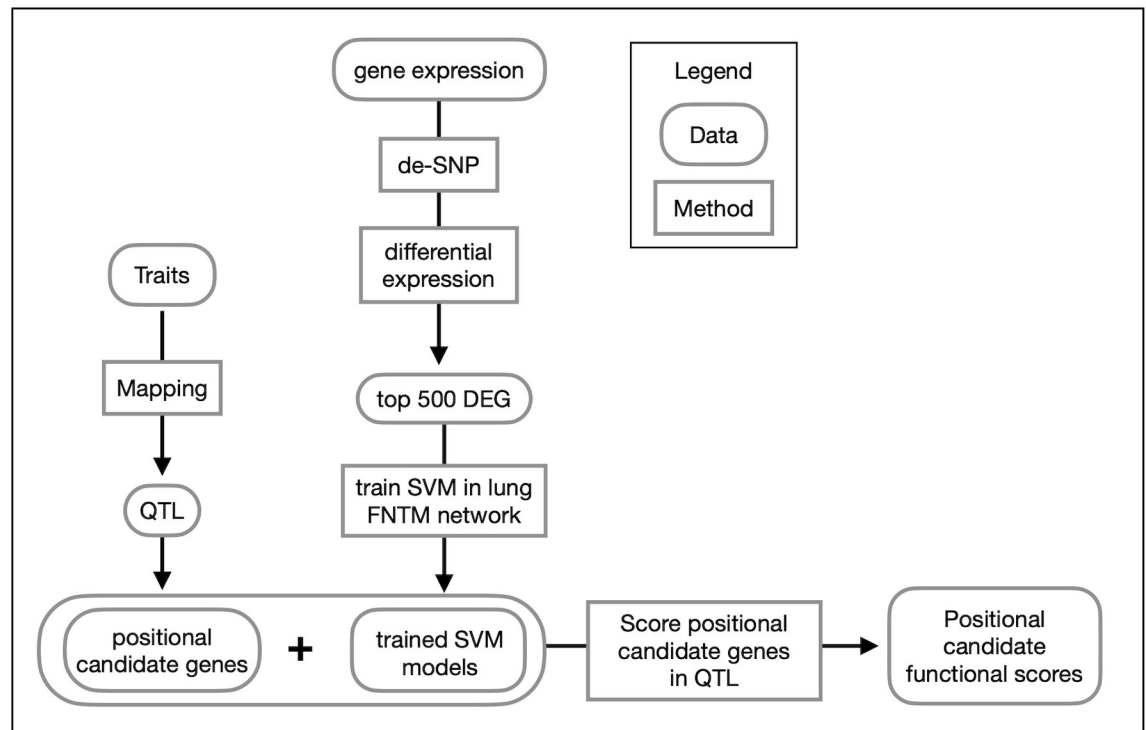


**Fig 4. The heatmap of linkage mapping peaks shows patterns of common genetic regulation.** Quantitative trait locus (QTL) mapping results from *M. tuberculosis* infected Diversity Outbred mice that were euthanized on or prior to 250 days post infection (if morbidity developed), and complete genotype and phenotypic trait data which also passed QA/QC ( $n = 853$ ). The mouse genome, from chromosome 1 through X, is shown on the horizontal axis. Phenotypes are shown on the vertical axis. Each cell shows the logarithmic of the odds (LOD) score on one chromosome for the phenotype listed on the left, colored by the color scale. Higher LOD scores (oranges and yellows) meet or exceed statistical thresholds and show the significant genotype-phenotype associations. The phenotypes are hierarchically clustered based on the correlation between LOD curves, i.e., phenotypes with similar LOD curves are clustered next to each other.

<https://doi.org/10.1371/journal.ppat.1011915.g004>

### Chromosome 1: Diversity Outbred Tuberculosis Susceptibility locus 1 (Dots1)

*Dots1* is a new QTL on chromosome 1 with a peak LOD  $>7.6$  ( $p_{GW} < 0.05$ ) at 155.36 M and interval 154.25–156.71 Mb shared by two correlated traits, *M. tuberculosis* burden and CXCL1 (Table 1). Two correlated traits (CXCL2 and MMP8) that mapped to the same position had LOD thresholds  $\geq 6.0$  ( $p_{GW} \leq 0.6$ ) (Fig 4). Notably, these four correlated traits shared patterns of founder allele effects (S5 Fig), suggesting this QTL contains an important mechanism of genetic regulation for neutrophil-mediated activities, extracellular matrix remodeling, and *M. tuberculosis* growth. To refine the locus, we calculated the first principal component of those four traits and plotted the LOD curve, which also peaked between 154–156 Mb (Fig 6A) and plotted the founder allele effects. A/J, C57BL/6J and WSB/Eij alleles contributed to greater values of principal component 1 and CAST/Eij alleles contributed to lesser values (Fig 6B). We next imputed the founder SNPs onto the Diversity Outbred genomes and performed association mapping in a 10 Mb region around the peak (Fig 6C). Interestingly, the SNPs with highest LOD scores were outside of the peak, and none of the SNPs with the highest LOD scores were missense, stop, or splice site SNPs. This suggested the SNPs in the confidence interval could regulate expression of nearby genes, including some of the 47 protein-coding genes in the interval (Fig 6D and S1 File). To find and prioritize trait-related gene candidates within *Dots1*, we used the trained SVM model to rank gene candidates based on the strength of their functional relationship in gene expression network modules (S2 File). *Fam20b* and *Ncf2* ranked highest by functional scoring (Fig 6E). Table 2 shows known protein functions, MVAR genome annotations, allele effects, founder alleles containing SNPs, and predicted effects of missense SNPs in *Fam20b* and *Ncf2* genes on protein functions.



**Fig 5. Overview of gene prioritization methods.** Traits were mapped to identify positional candidate genes in QTLs. Gene expression data were analyzed for differential gene expression. The top 500 differentially expressed genes (DEG) were used to train SVMs to distinguish these trait-related genes from other genes in the genome using the FNTM mouse lung network. The fitted models were used to score positional candidates in each trait QTL. Positional candidates were then ranked as trait-related based on their functional scores.

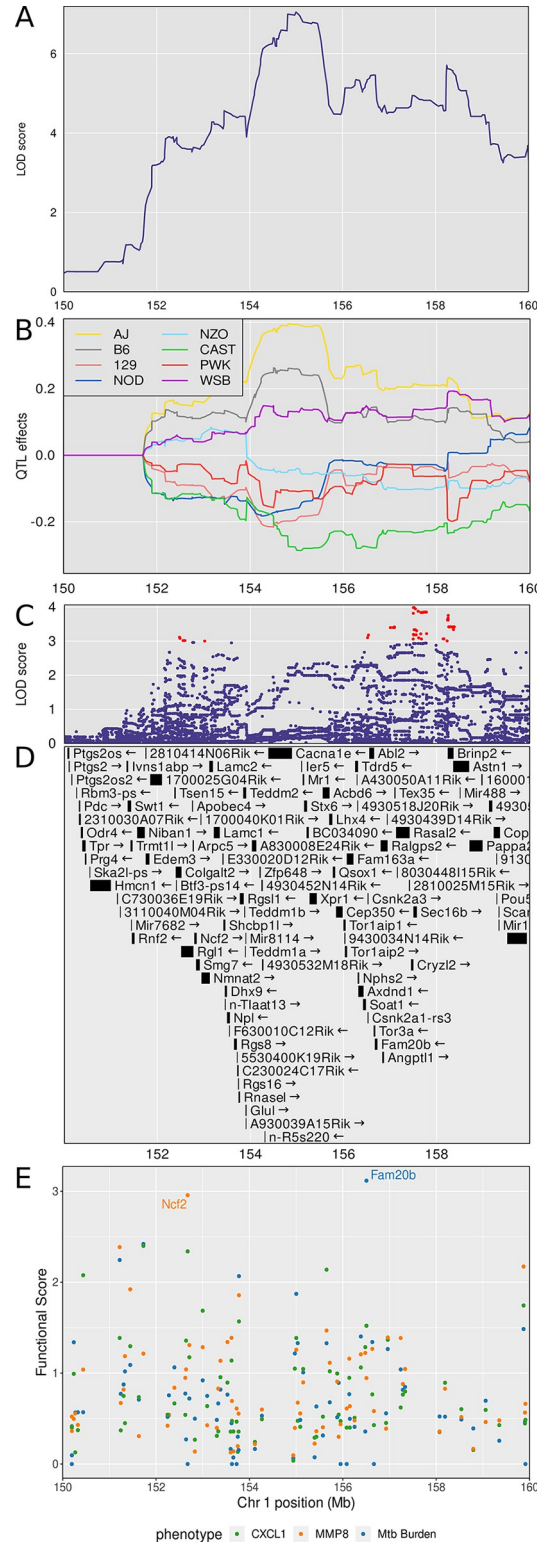
<https://doi.org/10.1371/journal.ppat.1011915.g005>

### Chromosome 2: Diversity Outbred Tuberculosis Susceptibility locus 2 (*Dots2*)

*Dots2* is a new QTL, not shared by correlated traits (Table 1), and has a peak LOD of 7.69 ( $p_{GW} < 0.05$ ) at 22.43 Mb that was associated with lung CXCL2 protein levels (Fig 4). *Dots2* contains 19 protein coding genes (S1 File). Because this QTL was associated with only one trait, gene prioritization by functional scoring was not pursued.

### Chromosome 3: Diversity Outbred Tuberculosis Susceptibility locus 3 (*Dots3*)

*Dots3* is not a new QTL (Table 1) and overlaps with *tbs1*, a QTL previously identified by crossing A/Sn and I/St inbred mouse strains [78]. These strains are not founder strains of the Diversity Outbred population. *Dots3* was identified by a single peak with a high LOD of 16.57 at 90.69 Mb and interval 90.52–92.02 ( $p_{GW} < 10^{-8}$ ) associated with lung S100A8 (calgranulin A) protein levels (Figs 4 and 7A). CAST/EiJ alleles effects were high and PWK/PhJ allele effects were low (Fig 7B). SNPs with the highest LOD scores within the peak are shown (Fig 7C). The interval contains 12 protein coding genes (S1 File) including the *S100a8* gene (Fig 7D), suggesting that genetic variants which affect *S100a8* transcription regulate S100A8 (calgranulin A) protein levels in *M. tuberculosis* infection. Further, based on the strength of the functional relationship in gene expression network modules (S2 File) the trained SVMs identified *S100a8* as the gene with the highest functional score in *Dots3* (Fig 7E). *Dots3* also contains the



**Fig 6. Quantitative Trait Locus (QTL) mapping results of first principal component (PC1) of CXCL1, CXCL2, *M. tuberculosis* burden, and MMP8 identifies *Dots1* on chromosome 1, containing the gene candidates *Fam20b* and *Ncf2*.** Panel A: LOD curve for PC1 between 150 and 160 Mb on chromosome 1 with peak near 155.36 Mb. Panel B: Founder allele effects for PC1 in the same genomic interval. Each colored line is the best linear unbiased predictor for one of the founder alleles. Founder colors are shown in the upper left. Panel C: LOD score of the imputed SNPs in the

same genomic interval. Each point represents the LOD score of one imputed SNP. Panel D: Genes in the confidence interval. Panel E: Functional scores for genes in the chromosome 1 QTL. Each dot represents a single gene. Its position on the x axis is its position within the QTL. Its position on the y axis is the functional  $-\log_{10}(\text{UPPR})$  derived from the SVM. Points are colored based on its correspondence with the trait—green with CXCL1, orange with MMP8, and blue with *M. tuberculosis* burden. *Fam20b* and *Ncf2* genes had the highest functional scores. QTL mapping results from  $n = 853$  *M. tuberculosis* infected Diversity Outbred mice up to 250 days post infection.

<https://doi.org/10.1371/journal.ppat.1011915.g006>

gene *S100a9*, which encodes S100A9 (calgranulin B), a protein binding partner of S100A8 (calgranulin A) required to form the heterodimer, calprotectin. Table 2 shows known protein functions, MVAR genome annotations, allele effects, founder alleles containing SNPs, and predicted effects of missense SNPs in *S100a8* and *S100a9* genes on protein functions.

### Chromosome 4: Diversity Outbred Tuberculosis Susceptibility locus (*Dots4*)

*Dots4* is a new QTL, not shared by correlated traits (Table 1), and has a peak LOD of 7.64 ( $p_{\text{GW}} < 0.05$ ) at 22.43 Mb and interval 22.18–23.79 Mb associated with lung S100A8 (calgranulin A) protein levels (Fig 4). *Dots4* contains two protein coding genes (S1 File). Because this QTL was associated with only one trait and contained few protein coding genes, prioritization by functional scoring was not pursued.

### Chromosome 16: Diversity Outbred Tuberculosis Susceptibility locus (*Dots5*)

*Dots5* is a new QTL on chromosome 16, shared by three correlated traits: lung *M. tuberculosis* burden (LOD = 8.45,  $p_{\text{GW}} \leq 0.01$ ), weight loss, and granuloma necrosis with a peak at 38.3 Mb

**Table 2. Loci, candidate genes, protein functions, genome annotations, founder alleles with single nucleotide polymorphisms (SNPs), and predicted effects of SNPs on protein functions.** Predicted effects of missense SNPs on proteins were identified using Ensembl's Variant Effect Predictor and Sorting Intolerant from Tolerant (SIFT) databases. SIFT produces scores based on sequence homology and physicochemical similarity between amino acids encoded by SNPs. Normalized scores are based on the probability that the amino acid change is tolerated, with scores  $< 0.05$  are predicted to be deleterious and all other scores are predicted 'tolerated'.

QTL	Candidate	Known protein functions	MVAR genome annotations*	Alleles with SNPs**	Predicted effects of missense SNPs
<i>Dots1</i>	<i>Ncf2</i>	superoxide production	Cn, Cs, I, U3, U5	129, <b>cast</b> , <i>nod</i> , <i>nzo</i> , <i>pwk</i> , <i>wsb</i>	2 of 2 tolerated
<i>Dots1</i>	<i>Fam20b</i>	glycosaminoglycan synthesis	Cn, Cs, I, U3, U5	129, <b>aj</b> , <b>cast</b> , <i>nod</i> , <i>nzo</i> , <i>pwk</i> , <i>wsb</i>	No missense SNPs
<i>Dots3</i>	<i>S100a8</i>	calgranulin A neutrophil responses	Cn, Cs, I, U3, U5	129, <b>cast</b> , <i>nod</i> , <i>nzo</i> , <b>pwk</b> , <i>wsb</i>	4 of 15 deleterious
<i>Dots3</i>	<i>S100a9</i>	calgranulin B neutrophil responses	Cs, I	<b>cast</b> , <i>nod</i> , <i>nzo</i> , <b>pwk</b> , <i>wsb</i>	1 of 1 tolerated
<i>Dots5</i>	<i>Itgb5</i>	integrin subunit binds fibronectin	Cn, Cs, I, U3, U5	129, <i>cast</i> , <b>pwk</b> , <i>wsb</i>	2 of 6 deleterious
<i>Dots5</i>	<i>Fstl1</i>	angiogenesis, wound-healing, fibrosis	Cn, Cs, I, U3, U5	<i>aj</i> , <i>cast</i> , <b>pwk</b> , <i>wsb</i>	3 of 4 deleterious
<i>Dots5</i>	<i>Zbtb20</i>	transcriptional repressor	Not annotated	<i>aj</i> , <i>cast</i> , <i>nod</i> , <b>pwk</b> , <b>wsb</b>	Not available
<i>Dots8</i>	<i>Ddr1</i>	interacts with collagen, fibrosis	Cn, Cs, I, U3, U5	129, <i>aj</i> , <i>cast</i> , <i>nod</i> , <b>nzo</b> , <b>pwk</b> , <i>wsb</i>	No results retrieved
<i>Dots8</i>	<i>Ier3</i>	signaling pathways, apoptosis	Cn, Cs, I, U3	<i>aj</i> , <i>nod</i> , <b>pwk</b> , <i>wsb</i>	No results retrieved
<i>Dots8</i>	<i>Vegfa</i>	angiogenesis, endothelial activation	Cs, I, U3, U5	129, <i>cast</i> , <i>nod</i> , <b>nzo</b> , <b>pwk</b> , <i>wsb</i>	No missense SNPs
<i>Dots8</i>	<i>Zfp318</i>	transcriptional repressor in B cells	Cn, Cs, I, U3, U5	129, <i>cast</i> , <i>nod</i> , <b>pwk</b>	4 of 16 deleterious

\* Mouse Variation Registry (MVAR) annotations

Cn Non-synonymous variant in codon, including missense.

Cs Synonymous variant in codon

I Variant in an intron

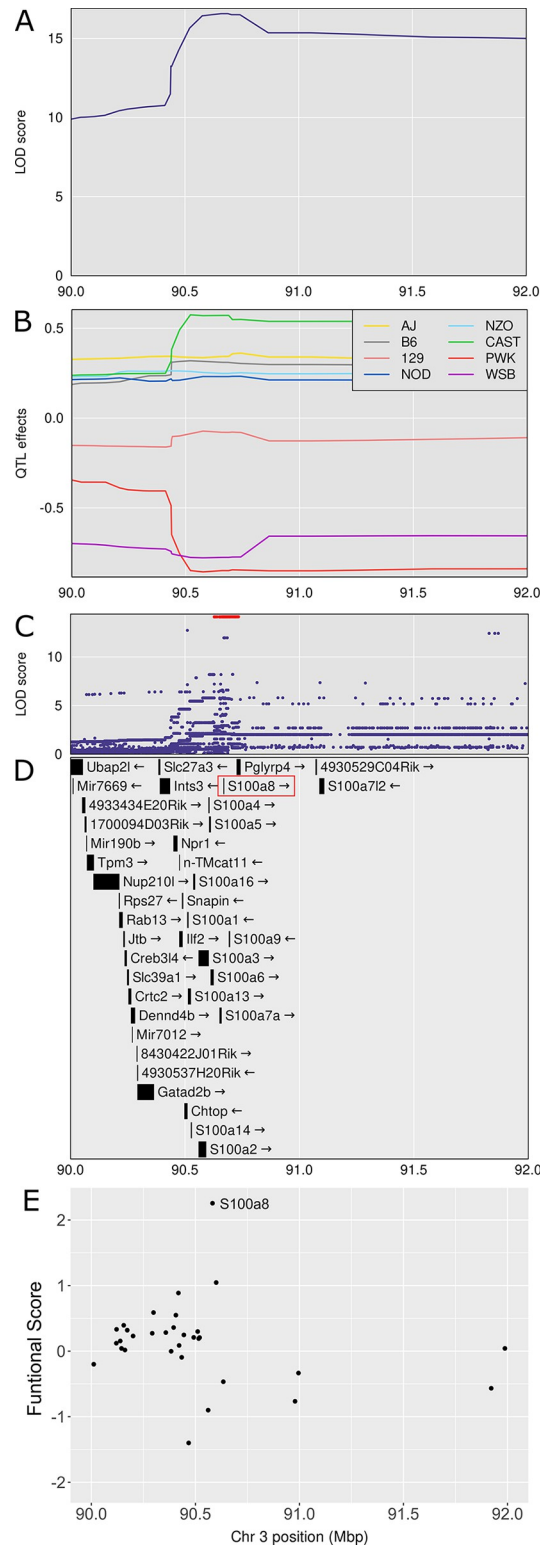
U3 UTR variant of the 3' UTR

U5 UTR variant of the 5' UTR

\*\* Bold alleles have high or low effects

<https://doi.org/10.1371/journal.ppat.1011915.t002>





**Fig 7. Quantitative Trait Locus (QTL) mapping of lung S100A8 identifies *Dots3* on chromosome 3, containing the gene candidates *S100a8* and *S100a9*.** Panel A: LOD score in the confidence interval from 90 to 92 Mb on chromosome 3. Panel B: Founder allele effects within the confidence interval. Panel C: LOD score of the imputed SNPs with the highest SNPs colored in red. Panel D: Genes in the same interval. The gene *S100a8* is directly under the SNPs with the highest LOD scores. Panel E: Functional scores for genes in the chromosome 3 QTL. Each dot represents a

single gene. Its position on the x axis is its position within the QTL. Its position on the y axis is the functional  $-\log_{10}$  (UPPR) derived from the SVM. The gene *S100a8* had the highest functional score. QTL mapping results from  $n = 853$  *M. tuberculosis* infected Diversity Outbred mice up to 250 days post infection.

<https://doi.org/10.1371/journal.ppat.1011915.g007>

and interval 33.28–43.28 Mb (Table 1 and Fig 4). We calculated the first principal component of these traits and plotted the LOD curve showing its peak (Fig 8A). The founder allele effects indicate that C57BL/6J alleles contribute to greater values of principal component 1 and that PWK/PhJ and NZO/HILtJ alleles contribute to lesser (Fig 8B). We imputed the founder SNPs onto the Diversity Outbred genomes and performed association mapping around the peak, showing the SNPs with the highest LOD scores (Fig 8C). The SNPs with high LOD scores (Fig 8D) were not missense, stop, or splice site SNPs in the 75 protein coding genes within the interval (S1 File). By prioritizing genes based on functional relationships in network modules, we identified *Fstl1* and *Itgb5* as functional candidates associated with weight loss, and *Zbtb20* as a functional candidate associated with *M. tuberculosis* burden (Fig 8E and S2 File). Table 2 shows known protein functions, MVAR genome annotations, allele effects, founder alleles containing SNPs, and predicted effects of missense SNPs in *Fstl1*, *Itgb5*, and *Zbtb20* genes on protein functions.

### Chromosome 16: Diversity Outbred Tuberculosis Susceptibility locus (Dots6)

*Dots6* is a new QTL on chromosome 16 and shared by two correlated traits: *M. tuberculosis* burden and CXCL5 with a peak at 52.23 Mb and interval 37.97–57.67 Mb (Table 1 and Fig 4). The interval contains 101 protein coding genes (S1 File). Because the LOD score was less than LOD threshold 7.64 for significance ( $p_{GW} < 0.05$ ), prioritization by functional scoring was not pursued.

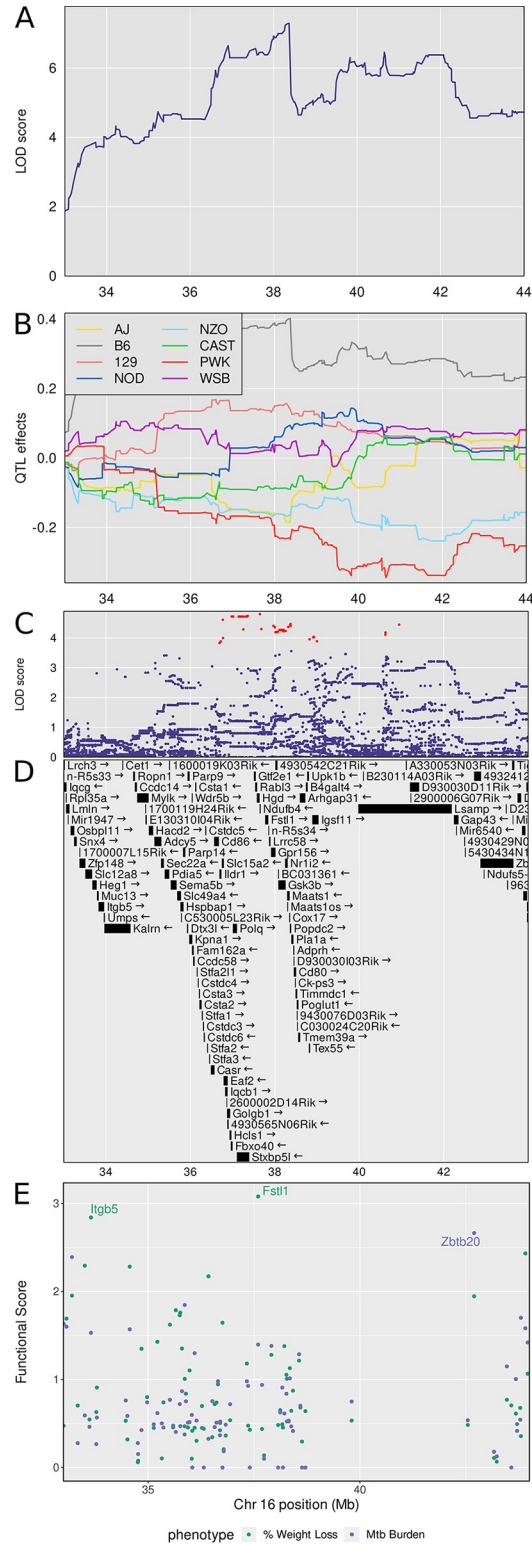
### Chromosome 17: Diversity Outbred Tuberculosis Susceptibility locus (Dots7)

*Dots7* on chromosome 17 is not new, and overlaps with *sst5* and *sst6*, QTLs that were previously identified by crossing C3HeB/FeJ and C57BL/6J inbred mouse strains [79]. *Dots7* is associated with *M. tuberculosis* burden and the LOD peaks ~20 Mb (Table 1 and Fig 4). The interval contains 198 protein coding genes (S1 File). Because the LOD score was less than the LOD threshold 7.64 for significance ( $p_{GW} < 0.05$ ), prioritization by functional scoring was not pursued.

### Chromosome 17: Diversity Outbred Tuberculosis Susceptibility locus (Dots8)

*Dots8* is not a new QTL and also overlaps with *sst5* and *sst6*, two QTLs that were previously identified by crossing C3HeB/FeJ and C57BL/6J inbred mouse strains [79]. Five traits: lung granuloma necrosis (“Necr Ratio”), weight loss, MMP8, CXCL1, and IL-10 mapped to *Dots8*. Lung granuloma necrosis had the highest LOD score (LOD = 8.12,  $p_{GW} \leq 0.02$ ) at 35.02 Mb and interval 33.94–41.06 Mb. Of those five traits, four positively correlated with each other (Fig 3) and had similar patterns of allele effects, and one, IL-10 had weak correlation and different pattern of founder allele effects.

We calculated the first principal component of the correlated traits and again performed QTL mapping. Principal component 1 mapped to a wide interval 30–50 Mb with a peak near 38 Mb (Fig 9A). The founder allele effects indicate PWK/PhJ alleles contribute to greater trait



**Fig 8. Quantitative Trait Locus (QTL) mapping results of first principal component (PC1) of *M. tuberculosis* burden, weight loss, and granuloma necrosis identifies *Dots5* on chromosome 16, containing gene candidates *Fstl1*, *Itgb5*, and *Zbtb20*.** *M. tuberculosis* lung burden, weight loss, and granuloma necrosis map to a region on chromosome 16 near 38 Mb. Panel A: LOD profile for *M. tuberculosis* in the confidence interval. The genomic position on chromosome 16 is on the horizontal axis and the LOD score is on the vertical axis. Panel B: Founder allele effect in

the confidence interval. The vertical axis shows the estimates effect of gaining one founder allele. Panel C shows the SNP LOD score for association mapping using imputed SNPs. Panel D: Genes in the confidence interval. Panel E: Functional scores for genes in the chromosome 16 QTL. Each dot represents a single gene. Its position on the x axis is its position within the QTL. Its position on the y axis is the functional  $-\log_{10}(\text{UPPR})$  derived from the SVM. Points are colored based on the trait—green with weight loss and blue with *M. tuberculosis* burden. *Fstl1* was the top ranked gene overall followed by *Igfb5* and *Zbtb20*. QTL mapping results from  $n = 853$  *M. tuberculosis* infected Diversity Outbred mice up to 250 days post infection.

<https://doi.org/10.1371/journal.ppat.1011915.g008>

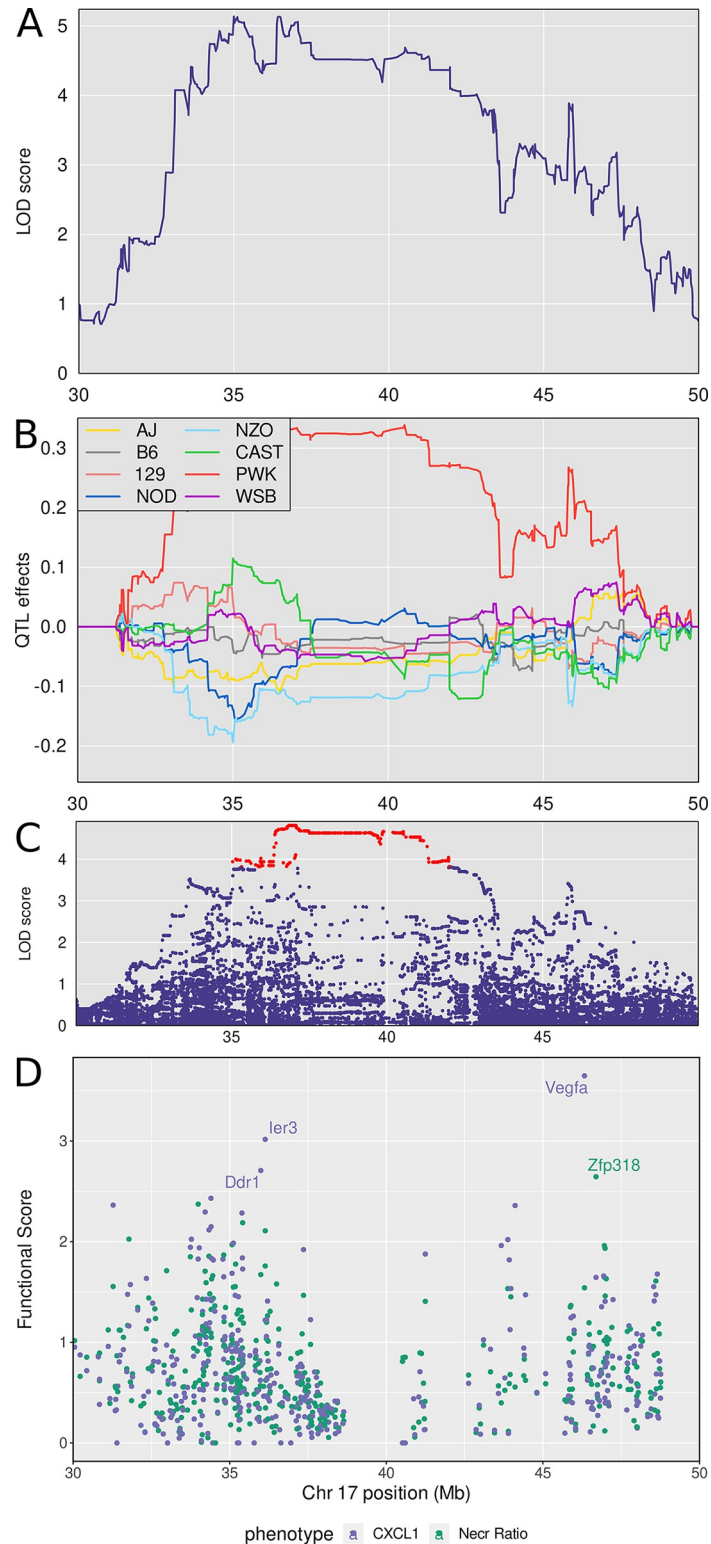
values, and NZO/HILtJ and NOD/ShiLtJ alleles contribute to lesser values (Fig 9B). We expected to find polymorphisms in the proximal peak of *Dots8* at 34–38 Mb because it contains the mouse histocompatibility-2 (H-2; or Major Histocompatibility Complex-II MHCII). This locus contains genes known to regulate innate and adaptive immunity and is known to be highly polymorphic. Indeed, the highest SNP association mapping LOD scores were over the mouse H-2 locus, located 36–38 Mb (Fig 9C), and there were 27 SNPs with protein-coding or splice site variation which occurred in 15 genes (S3 File). Among these were several histocompatibility genes (*H2-M1*, *H2-M5*, *H2-M9*, *H2-M11*) and several tripartite motif (TRIM) family genes (*Trim10*, *Trim26*, *Trim31*, *Trim40*).

The genes within the broad interval of *Dots8* are difficult to show in the figure and are provided in S1 File. There were 361 protein-coding genes within the 30–50 Mb locus. We prioritized positional candidate genes based on their functional relationships in network modules to the correlated traits (S2 File). This identified candidates *Ddr1*, *Ier3*, and *Vegfa* associated with CXCL1; and *Zfp318* associated with granuloma necrosis (Fig 9D). Table 2 shows known protein functions, MVAR genome annotations, allele effects, founder alleles containing SNPs, and predicted effects of missense SNPs in *Ddr1*, *Ier3*, *Vegfa*, and *Zfp318* genes on protein functions.

## Methodological and gene candidate validation

We performed three validation types shown in S6 Fig. This included (i) survival analysis of Diversity Outbred mice carrying PWK/PhJ alleles at the H-2 locus in *Dots8* on chromosome 17; (ii) quantification of S100A8 protein levels in lungs of *M. tuberculosis* infected PWK/PhJ and CAST/EiJ inbred founder strains; and (iii) infection of *S100a8* C57BL/6 gene deficient mice. Notably, infected Diversity Outbred mice carrying at least one copy of the PWK/PhJ allele at the mouse H-2 locus had shorter survival than mice carrying other alleles at the H-2 locus (S6A Fig).

The *Dots3* QTL on chromosome 3, associated with S100A8, had the highest LOD score, narrowest interval, distinct founder allele effects, and the gene candidate was the *S100a8* gene itself. For these reasons, the locus was selected for additional study. To help validate the method and ensure the founder allele effects generated by QTL mapping were biologically relevant, S100A8 was measured in lungs of the two founder inbred strains with the highest and lowest allele effects at the locus (Fig 7). As predicted by the allele effects, lungs from *M. tuberculosis* infected CAST/EiJ inbred mice contained significantly more S100A8 protein than lungs from *M. tuberculosis* infected PWK/PhJ inbred mice (S6B Fig). To evaluate *in vivo* effects of *S100a8* genetic deficiency, we obtained C57BL/6 breeding pairs to generate knockout, heterozygous, and wild-type mice. Littermates with null mutation (knockout), heterozygous, and wild-type C57BL/6 *S100a8* alleles were infected with *M. tuberculosis* for up to 50 days. The absence of one or both copies of the C57BL/6 *S100a8* allele had minimal impact on *M. tuberculosis* lung burden (S6C Fig). The lung granulomas of knockout, heterozygous, and wild type C57BL/6 mice were similar and non-necrotizing, composed of small dense aggregates of lymphocytes and macrophages, without neutrophils and lacking interalveolar pyknotic debris



**Fig 9. Quantitative Trait Locus (QTL) mapping results of first principal component (PC1) of granuloma necrosis, *M. tuberculosis* burden, weight loss, CXCL1 and MMP8 identifies *Dots7* on chromosome 17, which contains gene candidates *Vegfa*, *Ier3*, *Ddr1*, and *Zfp318*.** Panel A: LOD score for PC1 of lung granuloma necrosis ratio, *M. tuberculosis* burden, MMP8, CXCL1, and weight loss in the interval where the phenotypes map. Panel B: Founder allele effects for the two peaks. Panel C: LOD score of the imputed SNPs in the interval, with the highest scoring SNPs

colored in red. Panel D: Functional scores for genes in the chromosome 17 QTL. Each dot represents a single gene. Its position on the x axis is its position within the QTL. Its position on the y axis is the functional  $-\log_{10}(\text{UPPR})$  derived from the SVM. Points are colored based on correspondence to the traits—blue with CXCL1 and green with granuloma necrosis. *Vegfa* was the top ranked gene overall followed by *Ier3*, *Ddr1*, and *Zpf318*. QTL mapping results from  $n = 853$  *M. tuberculosis* infected Diversity Outbred mice up to 250 days post infection.

<https://doi.org/10.1371/journal.ppat.1011915.g009>

(S6D, S6E and S6F Fig). This indicates that a lack or reduction of S100A8 has no major impact on early *M. tuberculosis* growth or granuloma formation in C57BL/6 inbred mice. These findings are unlikely to generalize to all genetic backgrounds, however, especially those genetic backgrounds predisposed to early accumulation of neutrophils and necrotizing granulomas with relatively high S100A8 levels, like the Diversity Outbred mice that are highly susceptible to *M. tuberculosis* [24,34].

## Discussion

TB remains a major public health concern in the United States and across the globe, with an estimated 2 billion people infected with *M. tuberculosis*; 8–9 million patients diagnosed each year, and 1–1.5 million deaths annually [2]. Fortunately, most humans (~90%) are highly resistant to *M. tuberculosis* and clear or control infection [11,80]. In susceptible adults, active pulmonary TB develops a few years following exposure and tends to occur in young to middle-aged adults in the prime years of their lives [81]. The disease is usually restricted to the lungs and is characterized by granuloma necrosis and cavitation, neutrophilic infiltration, and cachexia [82,83]. The genetic basis of pulmonary TB is complex and not attributable to single-gene defects that cause severe immune deficiency (i.e., Mendelian susceptibility to mycobacterial disease does not explain pulmonary TB) [15,84–89]. Although genome-wide association studies have identified loci, gene candidates, and SNPs associated with increased or decreased odds ratios for pulmonary TB, only a few (e.g., *Ipr1*/SP110b and HLA variants/I-A Major Histocompatibility genes) have been validated [16,90–96]. This has led investigators to seek alternative experimental mouse models such as Diversity Outbred mice and Collaborative Cross recombinant inbred strains to examine effects of genetics on host responses to *M. tuberculosis* [24–26,47].

An advantage of the Diversity Outbred mouse population is that infection with *M. tuberculosis* induces phenotypes that are rare in common laboratory inbred strains of mice [15–17]. Further, a growing body of evidence shows similarities in *M. tuberculosis*-infected Diversity Outbred mice and humans in biomarkers, gene expression signatures, and BCG vaccination [28,33,34,47,97]. The phenotypic similarities suggest that humans and Diversity Outbred mice may share underlying genetic pathways of immunity and disease. And, because SNP variants in the Diversity Outbred mouse genomes are dense, with balanced allele frequencies, any gene that plays a role in disease is theoretically detectable [98]. This eliminates a problem common to human genetic studies where under-represented alleles cannot be confidently associated with disease phenotypes because they are low-frequency genetic events.

We performed genetic mapping in *M. tuberculosis*-infected Diversity Outbred mice and used orthogonal methods to rationally select candidate genes. We first used DOQTL mapping, which relies entirely on phenotypic and genetic variation, to find eight QTLs on six different chromosomes named *Dots1* through *Dots8*. To refine loci, we then subjected the QTLs on chromosomes 1, 16, and 17 (*Dots1*, *Dots5*, and *Dots8*) by mapping the first principal component of the correlated traits with similar patterns of allele effects that colocalized to the same interval. Finally, we applied a gene-based machine-learning SVM to identify and rank gene candidates based on functional scores. The sequential methods narrowed the candidate gene

list to eleven polymorphic, protein coding genes. Finally, the SNPs were critically examined using publicly available databases to find four candidates (*S100a8*, *Itgb5*, *Fstl1*, *Zfp318*) where missense SNPs are predicted to have deleterious effects on protein function.

All eleven candidates have roles in infection, inflammation, cell migration, extracellular matrix remodeling, or intracellular signaling. Of those, only one (mouse *Ncf2* in *Dots1* on chromosome 1) has a human homologue where a single SNP (G nucleotide in human NCF2 rs10911362) may provide a protective effect because this SNP was associated with a lower odds ratio for pulmonary TB in humans [99]. Absence of *Ncf2* in C57BL/6 inbred mice temporarily impairs resistance to *M. tuberculosis* infection by abrogating superoxide production, but the defect does not affect overall survival due to compensation by T cell mediated immunity [100]. Experimental validation of mouse *Ncf2* and human NCF2 polymorphisms remains to be confirmed.

We identified *Fam20b* in *Dots1* QTL on chromosome 1. The gene encodes a xylosylkinase that functions in glycosaminoglycan synthesis to produce extracellular matrix components in tissues. Deficiencies are embryologically lethal or cause cranioskeletal malformations [52,53,101,102]. A functionally homologous enzyme phosphorylates cadherins [103] which regulate immune cell migration [104] by interacting with extracellular matrix components, and since cell migration is required to form mycobacterial granulomas, *Fam20b* polymorphisms may alter host susceptibility to *M. tuberculosis* by changing extracellular matrix.

We identified *S100a8* and *S100a9* genes in *Dots3* QTL on chromosome 3, which encode S100A8 (calgranulin A) and S100A9 (calgranulin B). The proteins form monomers, homodimers, heterodimers, and multimers in inflammation, host defense, and nociception [105–108]. Some forms activate Toll-Like receptor 4; some activate the receptor for advanced glycation end-products (RAGE); and some sequester calcium, zinc, and manganese metal ions [105,109,110]. In pulmonary TB, S100A9 contributes to neutrophil localization to granulomas, and both S100A8 and S100A9 are biomarkers of TB-related lung damage [34,47,97,110–112]. Our results showing *S100a8* deficiency does not impact on *M. tuberculosis* lung burden, or granulomas, during early infection of C57BL/6 inbred mouse strain aligns with findings by Scott *et al.* [97]. More specifically, the authors showed that S100A9 was elevated during chronic *M. tuberculosis* infection of C57BL/6 mice where it promotes neutrophil recruitment by CD11b, and C57BL/6 mice lacking S100A9 had lower *M. tuberculosis* CFU by 100 days post infection [97]. Whether these findings and mechanisms generalize to all genetic backgrounds and to all forms of S100A8 and S100A9 remain open questions, as S100A8 has anti-inflammatory functions [113] that appear independent from S100A9.

We identified *Itgb5* in *Dots5* QTL on chromosome 16 as a gene candidate. *Itgb5* encodes the beta 5 ( $\beta 5$ ) integrin subunit which dimerizes with the alpha v subunit to mediate cell adhesion and signaling by binding to fibronectin and vitronectin [49]. Notably, the  $\beta 5$  subunit is on the surface of cancer cells, and normal epithelial cells and activated endothelial cells but not on lymphoid or myeloid cells [49,114–120]. To our knowledge, neither the mouse nor human gene, nor subunit  $\beta 5$ , nor the  $\alpha v \beta 5$  integrin heterodimer have been investigated in pulmonary TB.

We identified *Fstl1* in *Dots5* QTL on chromosome 16. The primary transcript encodes microRNA (miR)-198. The product is a secreted glycoprotein, FSTL1 with activities in angiogenesis, cell proliferation, differentiation, embryogenesis, metastasis, and wound healing; specifically reducing inflammation and fibrosis in cardiovascular disease [51,121–124]. Notably, *Fstl1* affects survival of *M. tuberculosis*-infected macrophages [43,44,125]. Given the central roles of macrophages, inflammatory mediators, and fibrosis in *M. tuberculosis* infection, understanding how *Fstl1* polymorphisms and FSTL1 function *in vivo* may inform TB pathogenesis, and host-directed therapy.

We identified *Zbtb20* in *Dots5* QTL on chromosome 16. The gene encodes a transcriptional repressor involved in glucose homeostasis; growth; hematopoiesis; innate immunity; neurogenesis; and B cell development and long-term survival of plasma cells [54–57,126–130]. Natural mutations occur in humans with Primrose Syndrome, although immune deficiencies are not reported [131]. To our knowledge, there are no studies on *Zbtb20* and *M. tuberculosis* infection or pulmonary TB. However, in *Listeria monocytogenes* infection, *Zbtb20*-deficiency improved CD8 T cell memory functions due to efficient use of diverse fuel sources [57]. Whether the same is true in pulmonary TB is unknown.

We identified *Ddr1* in *Dots8* QTL on chromosome 17. *Ddr1* encodes for the discoidin domain receptor 1 (DDR1), which interacts with collagen [50]. Initial studies suggested DDR1 function was restricted to epithelial cells; however, recent work shows expression on solid tumors, metastatic cells, and mouse histiocytic cancer cell lines, J774 and Raw264.7 [132–138]. DDR1 has additional roles in demyelination, fibrosis, vitiligo, and wound healing, and it is also a promising target for anti-fibrotic therapy [139–143]. Whether *Ddr1* (mouse) or DDR1 (human) gene polymorphisms contribute to pulmonary TB, or whether it could be a target for anti-fibrotic therapy in TB are areas open for investigation.

We identified the immediate early response gene, *Ier3* in *Dots8* QTL on chromosome 17. *Ier3* transcription is triggered by cytokines, hormones, DNA damage, and infections. The protein, IER3, regulates apoptosis, DNA repair, differentiation, and proliferation by interfering with NF- $\kappa$ B, MAPK/ERK and PI3K/Akt signaling pathways [58–60,144–146]. Mice lacking *Ier3* are more susceptible to *Leishmania* [147], an intracellular pathogen that shares some similar immune responses profiles with those induced by *M. tuberculosis* but we did not find studies showing that mutated *Ier3* also increases susceptibility to *M. tuberculosis*. One *in vitro* study of human macrophages, however, had high levels of IER3 mRNA following infection with a hypervirulent strain of *M. tuberculosis* [148] indicating the transcriptional pathway is triggered.

We identified *Vegfa* in *Dots8* QTL on chromosome 17. Mouse *Vegfa* and human VEGFA, encode for a heparin-binding protein and essential growth factor that induces proliferation, migration, and permeability changes in vascular endothelial cells by binding VEGFR1 and VEGFR2 [149–152]. Roles for VEGF in pathogenesis and diagnostics for extrapulmonary TB, cavitary TB, and active TB have been published [153–155]. Myeloid-specific gene deletion of *Vegfa* extended survival of C57BL/6J inbred mice infected with *M. tuberculosis* [156], highly noteworthy because very few gene deletions improve survival. Whether *Vegfa* or VEGFA polymorphisms have the same effect is unknown.

Lastly, we identified *Zfp318* in *Dots8* QTL on chromosome 17. The gene encodes the transcription factor, zinc finger protein 318 and it is expressed in testes, hematopoietic, and lymph nodes [157]. In B cells, the protein represses transcription required for class switching, helping to maintain B cell anergy and prevent autoimmunity [158–160]. Database and literature searches identified no publications on mouse *Zfp318* or human ZFP318 in infectious diseases.

When we compared genetic mapping results from Diversity Outbred mice to results from colleagues using Collaborative Cross inbred strains [26], QTLs and gene candidates did not overlap although we measured a few of the same traits by standard laboratory methods (e.g., body weight changes, lung *M. tuberculosis* burden, and lung CXCL1 by immunoassays). This suggests that phenotype-genotype relationships in the Collaborative Cross strains may be fundamentally different than in the Diversity Outbred mouse population, possibly due to heterozygosity in the Diversity Outbred population which is absent from Collaborative Cross inbred strains. Other reasons could be differences in routes of infection that change the host cell types first encountering *M. tuberculosis* bacilli which alters antigen presentation, T-cell, and B-cell priming. Here, we modeled natural aerosol exposure by delivering a low dose of *M.*



*tuberculosis* bacilli to the lungs of Diversity Outbred mice in nebulizer-delivered aerosol mist, and then focused on quantification of lung disease. In contrast, Smith *et al* [26] took a different approach by using intravenous infection with  $1 \times 10^6$  bacilli to take advantage of their rich library of transposon mutants, allowing detailed assessment of pathogen-associated QTLs. As the intravenous route of infection favors rapid induction of acquired immunity by delivering bacilli directly to lymphoid organs (i.e., spleen, thoracic, and abdominal lymph nodes by portal and systemic circulation), this approach maximized identification of unique Host-Interacting-with Pathogen QTLs and resulted in a prioritized list of candidate genes involved in immunity.

Overall, by using a systems genetics approach focused on the lungs, we identified multiple new and existing QTLs, and 11 candidate genes. Of those, gene products for five (*Ncf2*, *Fstl1*, *Zbtb20*, *Vegfa*, *Zfp318*) have known roles in recruitment, activation, or regulation of effector functions of immune cells (e.g., neutrophils, monocytes, macrophages and CD8 T cells). The gene products for three candidates (*Fam20b*, *Itgb5*, *Ddr1*) have known roles in epithelial cell, endothelial cell, and (possibly) macrophage adhesion to extracellular matrix glycoproteins or participate in remodeling of extracellular matrix. The gene products for two candidates (*S100a8* and *S100a9*) have complex and context-dependent roles in innate immune response signaling and in host defenses. Finally, the gene product of one candidate (*Ier3*) controls early stress responses of cells, including cell survival and death pathways. Ten of the eleven candidates have annotated polymorphisms; six have missense SNPs in protein coding regions; and the SNPs in four candidates (*S100a8*, *Itgb5*, *Fstl1*, and *Zfp318*) are predicted to have deleterious consequences on protein functions. Together, these results yield a concise list of candidates that may be major regulators of host necrotizing and inflammatory responses during *M. tuberculosis* infection and pulmonary TB disease progression. Future studies will focus on testing effects of these gene candidates and polymorphisms *in vivo* and identification of pathogenic molecular and cellular mechanisms.

## Supporting information

**S1 Fig. Mouse body weight following a low dose of aerosolized *M. tuberculosis* strain Erdman.** Mice were infected with a low dose of *M. tuberculosis* strain Erdman by aerosol and infection progressed unmanipulated until mice were euthanized due to IACUC-approved morbidity criteria. Panel A: Body weight changes of identically housed, age-, gender-, and generation-matched non-infected Diversity Outbred (DO) mice (n = 49) compared to baseline. Panels B, C, and D: Body weight changes of Progressor DO mice (n = 195); Controller DO mice (n = 145); and C57BL/6J inbred founder strain mice that succumbed to pulmonary TB (n = 39), are shown over time compared to pre-infection baseline. All mice were weighed 1 to 3 days prior to *M. tuberculosis* infection, at least twice per week during infection, and immediately before euthanasia. Each line is the body weight expressed as a percent change compared to initial pre-infection body weight. (TIF)

**S2 Fig. Clinical correlates of survival due to pulmonary TB in Diversity Outbred (DO) mice following exposure to a low dose of aerosolized *M. tuberculosis* strain Erdman.** Age-, gender-, and generation-matched DO mice were assigned to cages at random, and infected (or not infected) with a low dose of *M. tuberculosis* strain Erdman by aerosol exposure. All mice were initially weighed 1–3 days prior to infection, at least twice per week during infection, and immediately before euthanasia. Panel A: shows retrospective analysis of pre-infection body weights of Non-infected mice (n = 76) compared to pre-infection body weights of Progressors (n = 298) and pre-infection body weights of Controllers (n = 195), shown as box-and-whisker plots with the line at the mean for each group, and whiskers at the minimum and maximum.

Data were analyzed by 1-way ANOVA with Tukey's multiple comparisons test \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ . Panels B, C, D: Infection progressed unmanipulated until mice were euthanized due to IACUC-approved morbidity criteria. Panel B: The rate of weight loss (gm/day) and duration of body weight (BW) loss in days negatively correlate with each other. Panel C: Duration of BW loss was strongly, positively, and linearly correlated with survival by Spearman correlation analysis ( $r = 0.848$  with dashed lines indicating the 95% confidence interval, 0.8204 to 0.8717,  $p < 0.0001$ ). Panel D: Correlation matrix to show how survival and eight clinical indicators correlate with each other. Only correlations with  $p$ -values  $< 0.00001$  are shown on the matrix. Cells marked by an "X" were not significantly correlated.

(TIF)

**S3 Fig. Examples of necrotizing and non-necrotizing lesions in lungs of *M. tuberculosis* infected Diversity Outbred (DO) mice by light microscopy at high magnification.** Lung lobes were formalin-fixed, paraffin-embedded, sectioned, and stained with carbol fuchsin and counterstained with hematoxylin & eosin. Panels A and B: High magnification images of necrotizing lung lesions. One example contains abundant pyknotic nuclear debris (A) and one example contains abundant fibrin, eosinophilic cellular debris, and less nuclear debris (B). Panels C and D: High magnification images of non-necrotizing lung lesions. Both examples contain mostly viable cells, including macrophages, foamy macrophages, and foci of lymphocytes (400X).

(TIF)

**S4 Fig. Similar patterns of founder allele effects for four traits that map to distal chromosome 1 QTL.** The pattern of founder allele effects of four correlated traits all with  $\text{LOD} > 6.0$  on chromosome 1 at 155.36 Mb is similar. Panel A: Founder allele effects for CXCL1. Panel B: Founder allele effects for CXCL2. Panel C: Founder allele effects for MMP8. Panel D: Founder allele effects for *M. tuberculosis* burden. Founder strain names are on the horizontal axis and the standardized allele effect are on the vertical axis. Allele effects from QTL mapping results from  $n = 853$  *M. tuberculosis* infected Diversity Outbred mice up to 250 days post infection.

(TIF)

**S5 Fig. Receiver operator characteristic (ROC) curves for SVM training on traits used in gene prioritization.** Each panel shows the true positive rate of the trained SVM as a function of the false positive rate for each trait. The area under the curve (AUC) is noted for each panel.

(TIF)

**S6 Fig. Validation of QTL mapping results.** Panel A: *M. tuberculosis* infected Diversity Outbred (DO) mice with one or more PWK/PhJ alleles (red) at the H-2 locus on chromosome 17 in *Dots8* (near 36 Mb) have significantly reduced survival compared to DO mice carrying other alleles (grey) ( $p = 0.00075$ , Cox-PH test). Days of survival are shown on the horizontal axis and the proportion of mice surviving is shown on the vertical axis. Panel B: Lungs from *M. tuberculosis* infected CAST/EiJ inbred mice contain significantly more S100A8 protein (calgranulin A) than lungs of PWK/PhJ inbred mice at the time points indicated on the horizontal axis. PWK/PhJ (red) and CAST/EiJ (green) inbred founder strains with 4–6 mice per strain per time point, analyzed by Mann-Whitney  $t$ -tests within each time point, \* $p < 0.05$ . Panel C: *M. tuberculosis* infected *S100a8* knockout (KO), heterozygotes (HET), wild-type (WT) C57BL/6 inbred mice were euthanized at the time points indicated, and *M. tuberculosis* lung burden assessed by CFUs with total combined 15–22 mice per genotype per time point from 2 independent experiments, shown as average and standard error of the mean. No significant (ns) differences were identified within each time point by mixed effects ANOVA with Tukey's post-test ( $p < 0.05$ ). Panel D, E, F: Representative examples of lymphohistiocytic non-

necrotizing granulomas day 50 post infection from *S100a8* knockout (D), heterozygous (E), and wild type (F) C57BL/6 inbred mice magnified 400x normal.  
(TIF)

**S1 File. This file is an Excel workbook containing worksheets that list all protein-coding genes in each QTL with the functional candidates highlighted.**

(XLSX)

**S2 File. This is an Excel file that lists the top ten functional candidates for each trait in each QTL.**

(XLSX)

**S3 File. This is an Excel file that lists genes within chromosome 17 QTL.**

(XLSX)

## Acknowledgments

We thank Ms. Julie Tzipori, Mr. Curtis Rich, Mr. Donald Girouard, and Dr. Sam Telford III for study support at the New England Regional Biosafety Laboratory at Tufts University Cummings School of Veterinary Medicine, North Grafton, MA. Ms. Frances Brown, Ms. Linda Wrijil, Ms. Sarah Ducat, Ms. Gina Scarglia, Ms. Dian Taylor provided histology services at Tufts University's Cummings School of Veterinary Medicine Comparative Pathology and Genomics Shared Resource, directed by Dr. Amanda Martinot and Dr. Heather Gardner. The Digital Histology Shared Resource at Vanderbilt University Medical Center performed whole slide imaging. The following reagents were obtained through BEI Resources: ESAT-6, Recombinant Protein Reference Standard, NR-49424; CFP-10, Recombinant Protein Reference Standard, NR49425; Plasmid pMRLB.7 Containing Gene Rv3875 (Protein ESAT-6) from *Mycobacterium tuberculosis*, NR-50170; Plasmid pMRLB.46 Containing Gene Rv3874 (Protein Cfp10) from *Mycobacterium tuberculosis*, NR-13297; *Mycobacterium tuberculosis*, Strain H37Rv, Culture Filtrate Proteins, NR-14825; and *Mycobacterium tuberculosis*, Strain H37Rv, Cell Wall Fraction, NR-14828. All microarray protocols were conducted by the Boston University Microarray and Sequencing Resource (BUMSR), and we thank Mr. Eduard Drizik of the BUMSR for initial microarray analyses. The following individuals are thanked for their technical expertise: Ms. Victoria Mello at Tufts Cummings School of Veterinary Medicine, North Grafton, MA; Mr. Austin Hossfeld at The Ohio State University, Columbus, OH; and Dr. Joanne Turner at Texas Biomedical Research Institute, San Antonio, TX.

Boston University Microarray and Sequencing Resource received support from the National Institutes of Health UL1TR001430. The New England Regional Biosafety Laboratory at Tufts University's Cummings School of Veterinary Medicine received support from the National Institutes of Health UC6A1066843. Tufts University's Cummings School of Veterinary Medicine Comparative Pathology and Genomics Shared Resource received support from the Massachusetts Life Sciences Center.

## Author Contributions

**Conceptualization:** Daniel M. Gatti, Sherry L. Kurtz, Karen L. Elkins, Igor Kramnik, Gillian Beamer.

**Data curation:** Daniel M. Gatti, Deniz Koyuncu, Thomas Tavolara, Adam Gower, Gillian Beamer.

**Formal analysis:** Daniel M. Gatti, Anna L. Tyler, J Matthew Mahoney, Adam Gower, Gillian Beamer.

**Funding acquisition:** Daniel M. Gatti, Bulent Yener, Metin N. Gurcan, Gillian Beamer.

**Investigation:** Daniel M. Gatti, Anna L. Tyler, Melanie L. Ginese, Anas Alsharaydeh, Gillian Beamer.

**Methodology:** Daniel M. Gatti, Anna L. Tyler, J Matthew Mahoney, Gary A. Churchill, Denise Dayao, Emily McGlone, Melanie L. Ginese, Aubrey Specht, Philippe A. Tessier, Gillian Beamer.

**Project administration:** Gillian Beamer.

**Resources:** Gary A. Churchill, Philippe A. Tessier, Gillian Beamer.

**Supervision:** MK Khalid Niazi, Gillian Beamer.

**Visualization:** Daniel M. Gatti, Gillian Beamer.

**Writing – original draft:** Daniel M. Gatti, Anna L. Tyler, J Matthew Mahoney, Deniz Koyuncu, Adam Gower, Anas Alsharaydeh, Gillian Beamer.

**Writing – review & editing:** Daniel M. Gatti, Anna L. Tyler, J Matthew Mahoney, Bulent Yener, Deniz Koyuncu, Metin N. Gurcan, MK Khalid Niazi, Thomas Tavolara, Adam Gower, Sherry L. Kurtz, Gillian Beamer.

## References

1. WHO. 14.9 million excess deaths associated with the COVID-19 pandemic in 2020 and 2021 2022 [updated 5/5/2022. Available from: <https://www.who.int/news/item/05-05-2022-14.9-million-excess-deaths-were-associated-with-the-covid-19-pandemic-in-2020-and-2021>.
2. Organization WH. Global tuberculosis report 2021. Geneva, Switzerland: World Health Organization; 2021.
3. WHO. Global Tuberculosis Report: Multidrug-/ rifampicin-resistant TB (MDR/RR-TB):Update 2017. 2017.
4. WHO. Global Tuberculosis Report. 2013.
5. Soleimanpour S, Yaghoubi A, Sadat Seddighinia F, Rezaee SAR. A century of attempts to develop an effective tuberculosis vaccine: Why they failed? International immunopharmacology. 2022; 109:108791. <https://doi.org/10.1016/j.intimp.2022.108791> PMID: 35487086
6. Heidary M, Shirani M, Moradi M, Goudarzi M, Pouriran R, Rezaeian T, et al. Tuberculosis challenges: Resistance, co-infection, diagnosis, and treatment. Eur J Microbiol Immunol (Bp). 2022. <https://doi.org/10.1556/1886.2021.00021> PMID: 35420996
7. Kassa GM, Tadesse A, Gelaw YA, Alemayehu TT, Tsegaye AT, Tamirat KS, et al. Predictors of mortality among multidrug-resistant tuberculosis patients in central Ethiopia: a retrospective follow-up study. Epidemiol Infect. 2020; 148:e258. <https://doi.org/10.1017/S0950268820002514> PMID: 33054897
8. Tiemersma EW, van der Werf MJ, Borgdorff MW, Williams BG, Nagelkerke NJ. Natural history of tuberculosis: duration and fatality of untreated pulmonary tuberculosis in HIV negative patients: a systematic review. PLoS One. 2011; 6(4):e17601. <https://doi.org/10.1371/journal.pone.0017601> PMID: 21483732
9. Vilaplana C, Cardona PJ. The lack of a big picture in tuberculosis: the clinical point of view, the problems of experimental modeling and immunomodulation. The factors we should consider when designing novel treatment strategies. Frontiers in microbiology. 2014; 5:55. <https://doi.org/10.3389/fmicb.2014.00055> PMID: 24592258
10. Achkar JM, Jenny-Avital ER. Incipient and subclinical tuberculosis: defining early disease states in the context of host immune response. J Infect Dis. 2011;204 Suppl 4:S1179–86. <https://doi.org/10.1093/infdis/jir451> PMID: 21996700

11. Verrall AJ, Netea MG, Alisjahbana B, Hill PC, van Crevel R. Early clearance of *Mycobacterium tuberculosis*: a new frontier in prevention. *Immunology*. 2014; 141(4):506–13. <https://doi.org/10.1111/imm.12223> PMID: 24754048
12. Flynn JL, Chan J. Immunology of tuberculosis. *Annu Rev Immunol*. 2001; 19:93–129. <https://doi.org/10.1146/annurev.immunol.19.1.93> PMID: 11244032
13. Hunter RL, Jagannath C, Actor JK. Pathology of postprimary tuberculosis in humans and mice: contradiction of long-held beliefs. *Tuberculosis (Edinb)*. 2007; 87(4):267–78. <https://doi.org/10.1016/j.tube.2006.11.003> PMID: 17369095
14. Azad AK, Sadee W, Schlesinger LS. Innate immune gene polymorphisms in tuberculosis. *Infect Immun*. 2012; 80(10):3343–59. <https://doi.org/10.1128/IAI.00443-12> PMID: 22825450
15. Gibbs KD, Schott BH, Ko DC. The Awesome Power of Human Genetics of Infectious Disease. *Annu Rev Genet*. 2022. <https://doi.org/10.1146/annurev-genet-080320-010449> PMID: 35697043
16. Medhasi S, Chantratita N. Human Leukocyte Antigen (HLA) System: Genetics and Association with Bacterial and Viral Infections. *J Immunol Res*. 2022; 2022:9710376. <https://doi.org/10.1155/2022/9710376> PMID: 35664353
17. Varshney D, Singh S, Sinha E, Mohanty KK, Kumar S, Kumar Barik S, et al. Systematic review and meta-analysis of human Toll-like receptors genetic polymorphisms for susceptibility to tuberculosis infection. *Cytokine*. 2022; 152:155791. <https://doi.org/10.1016/j.cyto.2021.155791> PMID: 35158257
18. Williams A, Orme IM. Animal Models of Tuberculosis: An Overview. *Microbiol Spectr*. 2016; 4(4). <https://doi.org/10.1128/microbiolspec.TBTB2-0004-2015> PMID: 27726810
19. Kramnik I, Beamer G. Mouse models of human TB pathology: roles in the analysis of necrosis and the development of host-directed therapies. *Semin Immunopathol*. 2016; 38(2):221–37. <https://doi.org/10.1007/s00281-015-0538-9> PMID: 26542392
20. Clark S, Hall Y, Williams A. Animal models of tuberculosis: Guinea pigs. *Cold Spring Harb Perspect Med*. 2015; 5(5):a018572.
21. Scanga CA, Flynn JL. Modeling tuberculosis in nonhuman primates. *Cold Spring Harb Perspect Med*. 2014; 4(12):a018564. <https://doi.org/10.1101/cshperspect.a018564> PMID: 25213189
22. Dharmadhikari AS, Nardell EA. What animal models teach humans about tuberculosis. *Am J Respir Cell Mol Biol*. 2008; 39(5):503–8. <https://doi.org/10.1165/rcmb.2008-0154TR> PMID: 18556589
23. Helke KL, Mankowski JL, Manabe YC. Animal models of cavitation in pulmonary tuberculosis. *Tuberculosis (Edinb)*. 2006; 86(5):337–48. <https://doi.org/10.1016/j.tube.2005.09.001> PMID: 16359922
24. Niazi MK, Dhulekar N, Schmidt D, Major S, Cooper R, Abeijon C, et al. Lung necrosis and neutrophils reflect common pathways of susceptibility to *Mycobacterium tuberculosis* in genetically diverse, immune-competent mice. *Dis Model Mech*. 2015; 8(9):1141–53. <https://doi.org/10.1242/dmm.020867> PMID: 26204894
25. Smith C, Proulx M, Olive A, Laddy D, Mishra B, Moss C, et al. Tuberculosis Susceptibility and Vaccine Protection Are Independently Controlled by Host Genotype. *mBio*. 2016; 7(5):e01516. <https://doi.org/10.1128/mBio.01516-16> PMID: 27651361
26. Smith CM, Baker RE, Proulx MK, Mishra BB, Long JE, Park SW, et al. Host-pathogen genetic interactions underlie tuberculosis susceptibility in genetically diverse mice. *Elife*. 2022; 11. <https://doi.org/10.7554/eLife.74419> PMID: 35112666
27. Kurtz SL, Rydén P, Elkins KL. Transcriptional signatures measured in whole blood correlate with protection against tuberculosis in inbred and outbred mice. *PLoS One*. 2023; 18(8):e0289358. <https://doi.org/10.1371/journal.pone.0289358> PMID: 37535648
28. Kurtz SL, Rossi AP, Beamer GL, Gatti DM, Kramnik I, Elkins KL. The Diversity Outbred Mouse Population Is an Improved Animal Model of Vaccination against Tuberculosis That Reflects Heterogeneity of Protection. *mSphere*. 2020 ;5(2).
29. Lai R, Gong DN, Williams T, Ogunsola AF, Cavallo K, Lindestam Arlehamn CS, et al. Host genetic background is a barrier to broadly effective vaccine-mediated protection against tuberculosis. *J Clin Invest*. 2023; 133(13). <https://doi.org/10.1172/JCI1167762> PMID: 37200108
30. Churchill GA, Gatti DM, Munger SC, Svenson KL. The Diversity Outbred mouse population. *Mammalian genome: official journal of the International Mammalian Genome Society*. 2012; 23(9–10):713–8. <https://doi.org/10.1007/s00335-012-9414-2> PMID: 22892839
31. Tabangin ME, Woo JG, Martin LJ. The effect of minor allele frequency on the likelihood of obtaining false positives. *BMC Proc*. 2009; 3 Suppl 7(Suppl 7):S41. <https://doi.org/10.1186/1753-6561-3-S7-S41> PMID: 20018033
32. Gatti DM, Svenson KL, Shabalin A, Wu LY, Valdar W, Simecek P, et al. Quantitative trait locus mapping methods for diversity outbred mice. *G3*. 2014; 4(9):1623–33. <https://doi.org/10.1534/g3.114.013748> PMID: 25237114

33. Ahmed M, Thirunavukkarasu S, Rosa BA, Thomas KA, Das S, Rangel-Moreno J, et al. Immune correlates of tuberculosis disease and risk translate across species. *Science translational medicine*. 2020; 12(528). <https://doi.org/10.1126/scitranslmed.aay0233> PMID: 31996462
34. Koyuncu D, Niazi MKK, Tavolara T, Abeijon C, Ginese ML, Liao Y, et al. CXCL1: A new diagnostic biomarker for human tuberculosis discovered using Diversity Outbred mice. *PLoS Pathog*. 2021; 17(8): e1009773. <https://doi.org/10.1371/journal.ppat.1009773> PMID: 34403447
35. Flynn JL, Goldstein MM, Chan J, Triebold KJ, Pfeffer K, Lowenstein CJ, et al. Tumor necrosis factor-alpha is required in the protective immune response against *Mycobacterium tuberculosis* in mice. *Immunity*. 1995; 2(6):561–72. [https://doi.org/10.1016/1074-7613\(95\)90001-2](https://doi.org/10.1016/1074-7613(95)90001-2) PMID: 7540941
36. Flynn JL, Chan J, Triebold KJ, Dalton DK, Stewart TA, Bloom BR. An essential role for interferon gamma in resistance to *Mycobacterium tuberculosis* infection. *J Exp Med*. 1993; 178(6):2249–54. <https://doi.org/10.1084/jem.178.6.2249> PMID: 7504064
37. Cooper AM, Magram J, Ferrante J, Orme IM. Interleukin 12 (IL-12) is crucial to the development of protective immunity in mice intravenously infected with *mycobacterium tuberculosis*. *J Exp Med*. 1997; 186(1):39–45. <https://doi.org/10.1084/jem.186.1.39> PMID: 9206995
38. Wong EA, Evans S, Kraus CR, Engelman KD, Maiello P, Flores WJ, et al. IL-10 Impairs Local Immune Response in Lung Granulomas and Lymph Nodes during Early *Mycobacterium tuberculosis* Infection. *J Immunol*. 2020; 204(3):644–59. <https://doi.org/10.4049/jimmunol.1901211> PMID: 31862711
39. Guan Y, Ackert-Bicknell CL, Kell B, Troyanskaya OG, Hibbs MA. Functional genomics complements quantitative genetics in identifying disease-gene associations. *PLoS Comput Biol*. 2010; 6(11): e1000991. <https://doi.org/10.1371/journal.pcbi.1000991> PMID: 21085640
40. Tyler AL, Raza A, Kremontsov DN, Case LK, Huang R, Ma RZ, et al. Network-Based Functional Prediction Augments Genetic Association To Predict Candidate Genes for Histamine Hypersensitivity in Mice. *G3*. 2019; 9(12):4223–33. <https://doi.org/10.1534/g3.119.400740> PMID: 31645420
41. Goya J, Wong AK, Yao V, Krishnan A, Homilius M, Troyanskaya OG. FNTM: a server for predicting functional networks of tissues in mouse. *Nucleic acids research*. 2015; 43(W1):W182–7. <https://doi.org/10.1093/nar/gkv443> PMID: 25940632
42. Crouser ED, White P, Caceres EG, Julian MW, Papp AC, Locke LW, et al. A Novel In Vitro Human Granuloma Model of Sarcoidosis and Latent Tuberculosis Infection. *Am J Respir Cell Mol Biol*. 2017; 57(4):487–98. <https://doi.org/10.1165/rcmb.2016-0321OC> PMID: 28598206
43. Zhang ZM, Zhang AR, Xu M, Lou J, Qiu WQ. TLR-4/miRNA-32-5p/FSTL1 signaling regulates mycobacterial survival and inflammatory responses in *Mycobacterium tuberculosis*-infected macrophages. *Experimental cell research*. 2017; 352(2):313–21. <https://doi.org/10.1016/j.yexcr.2017.02.025> PMID: 28215633
44. Deng Q, Huang J, Yan J, Mao E, Chen H, Wang C. Circ\_0001490/miR-579-3p/FSTL1 axis modulates the survival of mycobacteria and the viability, apoptosis and inflammatory response in *Mycobacterium tuberculosis*-infected macrophages. *Tuberculosis (Edinb)*. 2021; 131:102123. <https://doi.org/10.1016/j.tube.2021.102123> PMID: 34555658
45. Duan XY, Sun Y, Zhao ZF, Shi YQ, Ma XY, Tao L, et al. Baicalin attenuates LPS-induced alveolar type II epithelial cell A549 injury by attenuation of the FSTL1 signaling pathway via increasing miR-200b-3p expression. *Innate Immun*. 2021; 27(4):294–312. <https://doi.org/10.1177/17534259211013887> PMID: 34000873
46. Henkel M, Dutta JA, Partyka J, Eddens T, Hirsch R, Kolls JK, et al. Regulation of Pulmonary Bacterial Immunity by Follistatin-Like Protein 1. *Infect Immun*. 2020; 89(1). <https://doi.org/10.1128/IAI.00298-20> PMID: 33077624
47. Gopal R, Monin L, Torres D, Slight S, Mehra S, McKenna KC, et al. S100A8/A9 proteins mediate neutrophilic inflammation and lung pathology during tuberculosis. *Am J Respir Crit Care Med*. 2013; 188(9):1137–46. <https://doi.org/10.1164/rccm.201304-0803OC> PMID: 24047412
48. Qiu J, Peng P, Xin M, Wen Z, Chen Z, Lin S, et al. ZBTB20-mediated titanium particle-induced peri-implant osteolysis by promoting macrophage inflammatory responses. *Biomater Sci*. 2020; 8(11):3147–63. <https://doi.org/10.1039/d0bm00147c> PMID: 32363359
49. Pasqualini R, Bodorova J, Ye S, Hemler ME. A study of the structure, function and distribution of beta 5 integrins using novel anti-beta 5 monoclonal antibodies. *Journal of cell science*. 1993; 105 (Pt 1):101–11. <https://doi.org/10.1242/jcs.105.1.101> PMID: 7689573
50. Vogel W. Discoidin domain receptors: structural relations and functional implications. *The FASEB journal*. 1999; 13(9001):S77–S82. <https://doi.org/10.1096/fasebj.13.9001.s77> PMID: 10352148
51. Mattiotti A, Prakash S, Barnett P, van den Hoff MJB. Follistatin-like 1 in development and human diseases. *Cell Mol Life Sci*. 2018; 75(13):2339–54. <https://doi.org/10.1007/s00018-018-2805-0> PMID: 29594389

52. Kuroda Y, Murakami H, Enomoto Y, Tsurusaki Y, Takahashi K, Mitsuzuka K, et al. A novel gene (FAM20B encoding glycosaminoglycan xylosylkinase) for neonatal short limb dysplasia resembling Desbuquois dysplasia. *Clinical genetics*. 2019; 95(6):713–7. <https://doi.org/10.1111/cge.13530> PMID: 30847897
53. Koike T, Izumikawa T, Tamura J, Kitagawa H. FAM20B is a kinase that phosphorylates xylose in the glycosaminoglycan-protein linkage region. *The Biochemical journal*. 2009; 421(2):157–62. <https://doi.org/10.1042/BJ20090474> PMID: 19473117
54. Zhu C, Chen G, Zhao Y, Gao XM, Wang J. Regulation of the Development and Function of B Cells by ZBTB Transcription Factors. *Frontiers in immunology*. 2018; 9:580.
55. Sutherland AP, Zhang H, Zhang Y, Michaud M, Xie Z, Patti M-E, et al. Zinc finger protein Zbtb20 is essential for postnatal survival and glucose homeostasis. *Molecular and cellular biology*. 2009; 29(10):2804–15. <https://doi.org/10.1128/MCB.01667-08> PMID: 19273596
56. Lightman SM, Utley A, Lee KP. Survival of Long-Lived Plasma Cells (LLPC): Piecing Together the Puzzle. *Frontiers in immunology*. 2019; 10:965. <https://doi.org/10.3389/fimmu.2019.00965> PMID: 31130955
57. Sun Y, Preiss NK, Valenteros KB, Kamal Y, Usherwood YK, Frost HR, et al. Zbtb20 Restrains CD8 T Cell Immunometabolism and Restricts Memory Differentiation and Antitumor Immunity. *J Immunol*. 2020; 205(10):2649–66. <https://doi.org/10.4049/jimmunol.2000459> PMID: 32998985
58. Schilling D, Pittelkow MR, Kumar R. IEX-1, an immediate early gene, increases the rate of apoptosis in keratinocytes. *Oncogene*. 2001; 20(55):7992–7. <https://doi.org/10.1038/sj.onc.1204965> PMID: 11753682
59. Im H-J, Pittelkow MR, Kumar R. Divergent Regulation of the Growth-promoting Gene IEX-1 by the p53 Tumor Suppressor and Sp1. *Journal of Biological Chemistry*. 2002; 277(17):14612–21. <https://doi.org/10.1074/jbc.M109414200> PMID: 11844788
60. Arlt A, Schäfer H. Role of the immediate early response 3 (IER3) gene in cellular stress response, inflammation and tumorigenesis. *Eur J Cell Biol*. 2011; 90(6–7):545–52. <https://doi.org/10.1016/j.ejcb.2010.10.002> PMID: 21112119
61. Harrison DE, Astle CM, Niazi MK, Major S, Beamer GL. Genetically diverse mice are novel and valuable models of age-associated susceptibility to *Mycobacterium tuberculosis*. *Immun Ageing*. 2014; 11(1):24. <https://doi.org/10.1186/s12979-014-0024-6> PMID: 25606048
62. Cyktor JC, Carruthers B, Kominsky RA, Beamer GL, Stromberg P, Turner J. IL-10 inhibits mature fibrotic granuloma formation during *Mycobacterium tuberculosis* infection. *J Immunol*. 2013; 190(6):2778–90.
63. Vesosky B, Rottinghaus EK, Stromberg P, Turner J, Beamer G. CCL5 participates in early protection against *Mycobacterium tuberculosis*. *J Leukoc Biol*. 2010; 87(6):1153–65.
64. Ullman-Cullere MH, Foltz CJ. Body condition scoring: a rapid and accurate method for assessing health status in mice. *Lab Anim Sci*. 1999; 49(3):319–23. PMID: 10403450
65. Tavolara TE, Niazi MKK, Ginese M, Piedra-Mora C, Gatti DM, Beamer G, et al. Automatic discovery of clinically interpretable imaging biomarkers for *Mycobacterium tuberculosis* supersusceptibility using deep learning. *EBioMedicine*. 2020; 62:103094. <https://doi.org/10.1016/j.ebiom.2020.103094> PMID: 33166789
66. Beamer GL, Flaherty DK, Vesosky B, Turner J. Peripheral blood gamma interferon release assays predict lung responses and *Mycobacterium tuberculosis* disease outcome in mice. *Clin Vaccine Immunol*. 2008; 15(3):474–83.
67. Tavolara TE, Niazi MKK, Gower AC, Ginese M, Beamer G, Gurcan MN. Deep learning predicts gene expression as an intermediate data modality to identify susceptibility patterns in *Mycobacterium tuberculosis* infected Diversity Outbred mice. *EBioMedicine*. 2021; 67:103388. <https://doi.org/10.1016/j.ebiom.2021.103388> PMID: 34000621
68. Keane TM, Goodstadt L, Danecek P, White MA, Wong K, Yalcin B, et al. Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature*. 2011; 477(7364):289–94. <https://doi.org/10.1038/nature10413> PMID: 21921910
69. Morgan AP, Fu CP, Kao CY, Welsh CE, Didion JP, Yadgary L, et al. The Mouse Universal Genotyping Array: From Substrains to Subspecies. *G3*. 2015; 6(2):263–79. <https://doi.org/10.1534/g3.115.022087> PMID: 26684931
70. Broman KW, Gatti DM, Simecek P, Furlotte NA, Prins P, Sen S, et al. R/qt2: Software for Mapping Quantitative Trait Loci with High-Dimensional Data and Multiparent Populations. *Genetics*. 2019; 211(2):495–502. <https://doi.org/10.1534/genetics.118.301595> PMID: 30591514

71. Cheng R, Abney M, Palmer AA, Skol AD. QTLRel: an R package for genome-wide association studies in which relatedness is a concern. *BMC Genet.* 2011; 12:66. <https://doi.org/10.1186/1471-2156-12-66> PMID: 21794153
72. Van Ooijen JW. LOD significance thresholds for QTL analysis in experimental populations of diploid species. *Heredity (Edinb).* 1999; 83 (Pt 5):613–24.
73. David Meyer ED, Kurt Hornik, Andreas Weingessel and Friedrich Leisch. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. 1.7–7 ed2021. p. R package.
74. Blake JA, Baldarelli R, Kadin JA, Richardson JE, Smith CL, Bult CJ, et al. Mouse Genome Database (MGD): Knowledgebase for mouse-human comparative biology. *Nucleic acids research.* 2021; 49 (D1):D981–D7. <https://doi.org/10.1093/nar/gkaa1083> PMID: 33231642
75. Kus P, Gurcan MN, Beamer G. Automatic Detection of Granuloma Necrosis in Pulmonary Tuberculosis Using a Two-Phase Algorithm: 2D-TB. *Microorganisms.* 2019; 7(12). <https://doi.org/10.3390/microorganisms7120661> PMID: 31817882
76. Niazi MKK, Beamer G, Gurcan MN. An application of transfer learning to neutrophil cluster detection for tuberculosis: efficient implementation with nonmetric multidimensional scaling and sampling: SPIE; 2018.
77. Niazi MKK, Beamer G, Gurcan MN. A computational framework to detect normal and tuberculosis infected lung from H and E-stained whole slide images: SPIE; 2017.
78. Sanchez F, Radaeva TV, Nikonenko BV, Persson AS, Sengul S, Schalling M, et al. Multigenic control of disease severity after virulent *Mycobacterium tuberculosis* infection in mice. *Infect Immun.* 2003; 71 (1):126–31. <https://doi.org/10.1128/IAI.71.1.126-131.2003> PMID: 12496157
79. Yan BS, Kirby A, Shebzukhov YV, Daly MJ, Kramnik I. Genetic architecture of tuberculosis resistance in a mouse model of infection. *Genes Immun.* 2006; 7(3):201–10. <https://doi.org/10.1038/sj.gene.6364288> PMID: 16452998
80. Dodd CE, Schlesinger LS. New concepts in understanding latent tuberculosis. *Curr Opin Infect Dis.* 2017; 30(3):316–21. <https://doi.org/10.1097/QCO.0000000000000367> PMID: 28177961
81. Behr MA, Edelstein PH, Ramakrishnan L. Revisiting the timetable of tuberculosis. *Bmj.* 2018; 362: k2738. <https://doi.org/10.1136/bmj.k2738> PMID: 30139910
82. Loddenkemper R, Lipman M, Zumla A. Clinical Aspects of Adult Tuberculosis. *Cold Spring Harb Perspect Med.* 2015; 6(1):a017848.
83. Hunter RL, Actor JK, Hwang SA, Karev V, Jagannath C. Pathogenesis of post primary tuberculosis: immunity and hypersensitivity in the development of cavities. *Annals of clinical and laboratory science.* 2014; 44(4):365–87. PMID: 25361920
84. Naranbhai V. The Role of Host Genetics (and Genomics) in Tuberculosis. *Microbiol Spectr.* 2016; 4 (5). <https://doi.org/10.1128/microbiolspec.TB2-0011-2016> PMID: 27787193
85. Abel L, El-Baghdadi J, Bousfiha AA, Casanova JL, Schurr E. Human genetics of tuberculosis: a long and winding road. *Philos Trans R Soc Lond B Biol Sci.* 2014; 369(1645):20130428. <https://doi.org/10.1098/rstb.2013.0428> PMID: 24821915
86. Kondratieva E, Logunova N, Majorov K, Averbakh M, Apt A. Host genetics in granuloma formation: human-like lung pathology in mice with reciprocal genetic susceptibility to *M. tuberculosis* and *M. avium*. *PLoS One.* 2010; 5(5):e10515. <https://doi.org/10.1371/journal.pone.0010515> PMID: 20463893
87. Bustamante J, Boisson-Dupuis S, Abel L, Casanova JL. Mendelian susceptibility to mycobacterial disease: genetic, immunological, and clinical features of inborn errors of IFN-gamma immunity. *Semin Immunol.* 2014; 26(6):454–70.
88. Qu HQ, Fisher-Hoch SP, McCormick JB. Molecular immunity to mycobacteria: knowledge from the mutation and phenotype spectrum analysis of Mendelian susceptibility to mycobacterial diseases. *International journal of infectious diseases: IJID: official publication of the International Society for Infectious Diseases.* 2011; 15(5):e305–13. <https://doi.org/10.1016/j.ijid.2011.01.004> PMID: 21330176
89. Cottle LE. Mendelian susceptibility to mycobacterial disease. *Clinical genetics.* 2011; 79(1):17–22. <https://doi.org/10.1111/j.1399-0004.2010.01510.x> PMID: 20718793
90. Kramnik I. Genetic dissection of host resistance to *Mycobacterium tuberculosis*: the *sst1* locus and the *lpr1* gene. *Current topics in Microbiology and Immunology.* 2008; 321:123–48. [https://doi.org/10.1007/978-3-540-75203-5\\_6](https://doi.org/10.1007/978-3-540-75203-5_6) PMID: 18727490
91. Tosh K, Campbell SJ, Fielding K, Sillah J, Bah B, Gustafson P, et al. Variants in the SP110 gene are associated with genetic susceptibility to tuberculosis in West Africa. *Proc Natl Acad Sci U S A.* 2006; 103(27):10364–8. <https://doi.org/10.1073/pnas.0603340103> PMID: 16803959



92. Pan H, Yan BS, Rojas M, Shebzukhov YV, Zhou H, Kobzik L, et al. Ipr1 gene mediates innate immunity to tuberculosis. *Nature*. 2005; 434(7034):767–72. <https://doi.org/10.1038/nature03419> PMID: 15815631
93. Kramnik I, Demant P, Bloom BB. Susceptibility to tuberculosis as a complex genetic trait: analysis using recombinant congenic strains of mice. *Novartis Found Symp*. 1998; 217:120–31; discussion 32–7. <https://doi.org/10.1002/0470846526.ch9> PMID: 9949805
94. Apt AS, Avdienko VG, Nikonenko BV, Kramnik IB, Moroz AM, Skamene E. Distinct H-2 complex control of mortality, and immune responses to tuberculosis infection in virgin and BCG-vaccinated mice. *Clin Exp Immunol*. 1993; 94(2):322–9. <https://doi.org/10.1111/j.1365-2249.1993.tb03451.x> PMID: 8222323
95. Barrera LF, Kramnik I, Skamene E, Radzioch D. I-A beta gene expression regulation in macrophages derived from mice susceptible or resistant to infection with *M. bovis* BCG. *Mol Immunol*. 1997; 34(4):343–55. [https://doi.org/10.1016/s0161-5890\(97\)00018-7](https://doi.org/10.1016/s0161-5890(97)00018-7) PMID: 9244347
96. Logunova N, Korotetskaya M, Polshakov V, Apt A. The QTL within the H2 Complex Involved in the Control of Tuberculosis Infection in Mice Is the Classical Class II H2-Ab1 Gene. *PLoS genetics*. 2015; 11(11):e1005672. <https://doi.org/10.1371/journal.pgen.1005672> PMID: 26618355
97. Scott NR, Swanson RV, Al-Hammadi N, Domingo-Gonzalez R, Rangel-Moreno J, Kriel BA, et al. S100A8/A9 regulates CD11b expression and neutrophil recruitment during chronic tuberculosis. *J Clin Invest*. 2020; 130(6):3098–112. <https://doi.org/10.1172/JCI130546> PMID: 32134742
98. Yang H, Bell TA, Churchill GA, Pardo-Manuel de Villena F. On the subspecific origin of the laboratory mouse. *Nat Genet*. 2007; 39(9):1100–7. <https://doi.org/10.1038/ng2087> PMID: 17660819
99. Jiao L, Song J, Ding L, Liu T, Wu T, Zhang J, et al. A Novel Genetic Variation in NCF2, the Core Component of NADPH Oxidase, Contributes to the Susceptibility of Tuberculosis in Western Chinese Han Population. *DNA and cell biology*. 2020; 39(1):57–62. <https://doi.org/10.1089/dna.2019.5082> PMID: 31794672
100. Cooper AM, Segal BH, Frank AA, Holland SM, Orme IM. Transient loss of resistance to pulmonary tuberculosis in p47(phox<sup>-/-</sup>) mice. *Infect Immun*. 2000; 68(3):1231–4. <https://doi.org/10.1128/IAI.68.3.1231-1234.2000> PMID: 10678931
101. Wu J, Tian Y, Han L, Liu C, Sun T, Li L, et al. FAM20B-catalyzed glycosaminoglycans control murine tooth number by restricting FGFR2b signaling. *BMC Biol*. 2020; 18(1):87. <https://doi.org/10.1186/s12915-020-00813-4> PMID: 32664967
102. Saiyin W, Li L, Zhang H, Lu Y, Qin C. Inactivation of FAM20B causes cell fate changes in annulus fibrosus of mouse intervertebral disc and disc defects via the alterations of TGF-beta and MAPK signaling pathways. *Biochim Biophys Acta Mol Basis Dis*. 2019; 1865(12):165555.
103. Ishikawa HO, Takeuchi H, Haltiwanger RS, Irvine KD. Four-jointed is a Golgi kinase that phosphorylates a subset of cadherin domains. *Science*. 2008; 321(5887):401–4. <https://doi.org/10.1126/science.1158159> PMID: 18635802
104. Perez TD, Nelson WJ. Cadherin adhesion: mechanisms and molecular interactions. *Handb Exp Pharmacol*. 2004(165):3–21. [https://doi.org/10.1007/978-3-540-68170-0\\_1](https://doi.org/10.1007/978-3-540-68170-0_1) PMID: 20455088
105. Zygiel EM, Nolan EM. Transition Metal Sequestration by the Host-Defense Protein Calprotectin. *Annual review of biochemistry*. 2018; 87:621–43. <https://doi.org/10.1146/annurev-biochem-062917-012312> PMID: 29925260
106. Wang S, Song R, Wang Z, Jing Z, Wang S, Ma J. S100A8/A9 in Inflammation. *Frontiers in immunology*. 2018; 9:1298. <https://doi.org/10.3389/fimmu.2018.01298> PMID: 29942307
107. Dale CS, Altier C, Cenac N, Giorgi R, Juliano MA, Juliano L, et al. Analgesic properties of S100A9 C-terminal domain: a mechanism dependent on calcium channel inhibition. *Fundam Clin Pharmacol*. 2009; 23(4):427–38. <https://doi.org/10.1111/j.1472-8206.2009.00686.x> PMID: 19515068
108. Dale CS, Pagano Rde L, Paccola CC, Pinotti-Guirao T, Juliano MA, Juliano L, et al. Effect of the C-terminus of murine S100A9 protein on experimental nociception. *Peptides*. 2006; 27(11):2794–802. <https://doi.org/10.1016/j.peptides.2006.07.002> PMID: 16920226
109. Hudson BI, Lippman ME. Targeting RAGE Signaling in Inflammatory Disease. *Annu Rev Med*. 2018; 69:349–64. <https://doi.org/10.1146/annurev-med-041316-085215> PMID: 29106804
110. Garcia V, Perera YR, Chazin WJ. A Structural Perspective on Calprotectin as a Ligand of Receptors Mediating Inflammation and Potential Drug Target. *Biomolecules*. 2022; 12(4).
111. Ryckman C, Vandal K, Rouleau P, Talbot M, Tessier PA. Proinflammatory activities of S100: proteins S100A8, S100A9, and S100A8/A9 induce neutrophil chemotaxis and adhesion. *The Journal of Immunology*. 2003; 170(6):3233–42. <https://doi.org/10.4049/jimmunol.170.6.3233> PMID: 12626582

112. Liu Q, Li R, Li Q, Luo B, Lin J, Lyu L. High levels of plasma S100A9 at admission indicate an increased risk of death in severe tuberculosis patients. *Journal of Clinical Tuberculosis and Other Mycobacterial Diseases*. 2021; 25:100270. <https://doi.org/10.1016/j.jctube.2021.100270> PMID: 34849408
113. Cesaro A, Defrene J, Lachhab A, Page N, Tardif MR, Al-Shami A, et al. Enhanced myelopoiesis and aggravated arthritis in S100a8-deficient mice. *PLoS One*. 2019; 14(8):e0221528. <https://doi.org/10.1371/journal.pone.0221528> PMID: 31437241
114. Li HM, Huang Q, Tang F, Zhang TP. Altered NCF2, NOX2 mRNA Expression Levels in Peripheral Blood Mononuclear Cells of Pulmonary Tuberculosis Patients. *Int J Gen Med*. 2021; 14:9203–9.
115. Leung K. (99m)Tc-Diamine dioxime-Lys-Cys-Arg-Gly-Asp-Cyc-Phe-Cys-polyethylene glycol. *Molecular Imaging and Contrast Agent Database (MICAD)*. Bethesda (MD): National Center for Biotechnology Information (US); 2004.
116. Lin Z, He R, Luo H, Lu C, Ning Z, Wu Y, et al. Integrin- $\beta$ 5, a miR-185-targeted gene, promotes hepatocellular carcinoma tumorigenesis by regulating  $\beta$ -catenin stability. *Journal of Experimental & Clinical Cancer Research*. 2018; 37:1–13.
117. Shi W, He J, Huang Y, Zeng Z, Feng Z, Xu H, et al. Integrin beta5 enhances the malignancy of human colorectal cancer by increasing the TGF-beta signaling. *Anticancer Drugs*. 2021; 32(7):717–26.
118. Zhuang H, Zhou Z, Ma Z, Li Z, Liu C, Huang S, et al. Characterization of the prognostic and oncologic values of ITGB superfamily members in pancreatic cancer. *J Cell Mol Med*. 2020; 24(22):13481–93. <https://doi.org/10.1111/jcmm.15990> PMID: 33073486
119. Nurzat Y, Su W, Min P, Li K, Xu H, Zhang Y. Identification of Therapeutic Targets and Prognostic Biomarkers Among Integrin Subunits in the Skin Cutaneous Melanoma Microenvironment. *Front Oncol*. 2021; 11:751875. <https://doi.org/10.3389/fonc.2021.751875> PMID: 34660316
120. Zhang LY, Guo Q, Guan GF, Cheng W, Cheng P, Wu AH. Integrin Beta 5 Is a Prognostic Biomarker and Potential Therapeutic Target in Glioblastoma. *Front Oncol*. 2019; 9:904. <https://doi.org/10.3389/fonc.2019.00904> PMID: 31616629
121. Ouchi N, Asaumi Y, Ohashi K, Higuchi A, Sono-Romanelli S, Oshima Y, et al. DIP2A functions as a FSTL1 receptor. *J Biol Chem*. 2010; 285(10):7127–34. <https://doi.org/10.1074/jbc.M109.069468> PMID: 20054002
122. Jin X, Nie E, Zhou X, Zeng A, Yu T, Zhi T, et al. Fstl1 promotes glioma growth through the BMP4/Smad1/5/8 signaling pathway. *Cellular Physiology and Biochemistry*. 2017; 44(4):1616–28. <https://doi.org/10.1159/000485759> PMID: 29212066
123. Murakami K, Tanaka M, Usui T, Kawabata D, Shiomi A, Iguchi-Hashimoto M, et al. Follistatin-related protein/follistatin-like 1 evokes an innate immune response via CD14 and toll-like receptor 4. *FEBS letters*. 2012; 586(4):319–24. <https://doi.org/10.1016/j.febslet.2012.01.010> PMID: 22265692
124. Horak M, Fairweather D, Kokkonen P, Bednar D, Bienertova-Vasku J. Follistatin-like 1 and its paralogs in heart development and cardiovascular disease. *Heart Fail Rev*. 2022; 27(6):2251–65. <https://doi.org/10.1007/s10741-022-10262-6> PMID: 35867287
125. Zhang Y, Luo D, Tang MI, Jiang DI, Yi H. Circ-WDR27 regulates mycobacterial vitality and secretion of inflammatory cytokines in *Mycobacterium tuberculosis*-infected macrophages via the miR-370-3p/FSTL1 signal network. *J Biosci*. 2022;47.
126. Zhang W, Mi J, Li N, Sui L, Wan T, Zhang J, et al. Identification and characterization of DPZF, a novel human BTB/POZ zinc finger protein sharing homology to BCL-6. *Biochem Biophys Res Commun*. 2001; 282(4):1067–73. <https://doi.org/10.1006/bbrc.2001.4689> PMID: 11352661
127. Xie Z, Zhang H, Tsai W, Zhang Y, Du Y, Zhong J, et al. Zinc finger protein ZBTB20 is a key repressor of alpha-fetoprotein gene transcription in liver. *Proceedings of the National Academy of Sciences*. 2008; 105(31):10859–64. <https://doi.org/10.1073/pnas.0800647105> PMID: 18669658
128. Xie Z, Ma X, Ji W, Zhou G, Lu Y, Xiang Z, et al. Zbtb20 is essential for the specification of CA1 field identity in the developing hippocampus. *Proceedings of the National Academy of Sciences*. 2010; 107(14):6510–5. <https://doi.org/10.1073/pnas.0912315107> PMID: 20308569
129. Zhang Y, Xie Z, Zhou L, Li L, Zhang H, Zhou G, et al. The zinc finger protein ZBTB20 regulates transcription of fructose-1, 6-bisphosphatase 1 and  $\beta$  cell function in mice. *Gastroenterology*. 2012; 142(7):1571–80. e6.
130. Liu X, Zhang P, Bao Y, Han Y, Wang Y, Zhang Q, et al. Zinc finger protein ZBTB20 promotes Toll-like receptor-triggered innate immune responses by repressing I $\kappa$ B $\alpha$  gene transcription. *Proc Natl Acad Sci U S A*. 2013; 110(27):11097–102.
131. Arora V, Ferreira CR, Dua Puri R, Verma IC. Primrose Syndrome. In: Adam MP, Mirzaa GM, Pagon RA, Wallace SE, Bean LJH, Gripp KW, et al., editors. *GeneReviews*. Seattle (WA): University of Washington, Seattle; 1993.

132. Kim SH, Lee S, Suk K, Bark H, Jun CD, Kim DK, et al. Discoidin domain receptor 1 mediates collagen-induced nitric oxide production in J774A.1 murine macrophages. *Free radical biology & medicine*. 2007; 42(3):343–52. <https://doi.org/10.1016/j.freeradbiomed.2006.10.052> PMID: 17210447
133. Wang H, Wen Y, Wang L, Wang J, Chen H, Chen J, et al. DDR1 activation in macrophage promotes IPF by regulating NLRP3 inflammasome and macrophage reaction. *International immunopharmacology*. 2022; 113(Pt A):109294. <https://doi.org/10.1016/j.intimp.2022.109294> PMID: 36257259
134. Yeh YC, Lin HH, Tang MJ. Dichotomy of the function of DDR1 in cells and disease progression. *Biochim Biophys Acta Mol Cell Res*. 2019; 1866(11):118473. <https://doi.org/10.1016/j.bbamcr.2019.04.003> PMID: 30954568
135. Taciak B, Białasek M, Braniewska A, Sas Z, Sawicka P, Kiraga Ł, et al. Evaluation of phenotypic and functional stability of RAW 264.7 cell line through serial passages. *PLoS One*. 2018; 13(6):e0198943. <https://doi.org/10.1371/journal.pone.0198943> PMID: 29889899
136. Ralph P, Prichard J, Cohn M. Reticulum cell sarcoma: an effector cell in antibody-dependent cell-mediated immunity. *J Immunol*. 1975; 114(2 pt 2):898–905. PMID: 1089721
137. Zhang X, Hu Y, Pan Y, Xiong Y, Zhang Y, Han M, et al. DDR1 promotes hepatocellular carcinoma metastasis through recruiting PSD4 to ARF6. *Oncogene*. 2022; 41(12):1821–34. <https://doi.org/10.1038/s41388-022-02212-1> PMID: 35140331
138. Su H, Yang F, Fu R, Trinh B, Sun N, Liu J, et al. Collagenolysis-dependent DDR1 signalling dictates pancreatic cancer outcome. *Nature*. 2022; 610(7931):366–72. <https://doi.org/10.1038/s41586-022-05169-z> PMID: 36198801
139. Ma R, Xie X, Zhao L, Wu Y, Wang J. Discoidin domain receptors (DDR1): Potential implications in periodontitis. *J Cell Physiol*. 2022; 237(1):189–98. <https://doi.org/10.1002/jcp.30560> PMID: 34431091
140. Moll S, Desmoulière A, Moeller MJ, Pache JC, Badi L, Arcadu F, et al. DDR1 role in fibrosis and its pharmacological targeting. *Biochim Biophys Acta Mol Cell Res*. 2019; 1866(11):118474. <https://doi.org/10.1016/j.bbamcr.2019.04.004> PMID: 30954571
141. Sirvent A, Espie K, Papadopoulou E, Naim D, Roche S. New functions of DDR1 collagen receptor in tumor dormancy, immune exclusion and therapeutic resistance. *Front Oncol*. 2022; 12:956926.
142. Marchioro HZ, Silva de Castro CC, Fava VM, Sakiyama PH, Dellatorre G, Miot HA. Update on the pathogenesis of vitiligo. *An Bras Dermatol*. 2022; 97(4):478–90. <https://doi.org/10.1016/j.abd.2021.09.008> PMID: 35643735
143. Vilella E, Gas C, Garcia-Ruiz B, Rivera FJ. Expression of DDR1 in the CNS and in myelinating oligodendrocytes. *Biochim Biophys Acta Mol Cell Res*. 2019; 1866(11):118483. <https://doi.org/10.1016/j.bbamcr.2019.04.010> PMID: 31108116
144. Wu M. Roles of the stress-induced gene IEX-1 in regulation of cell death and oncogenesis. *Apoptosis*. 2003; 8:11–8. <https://doi.org/10.1023/a:1021688600370> PMID: 12510147
145. Arlt A, Grobe O, Sieke A, Kruse M-L, Foëlsch UR, Schmidt WE, et al. Expression of the NF-κB target gene IEX-1 (p22/PRG1) does not prevent cell death but instead triggers apoptosis in HeLa cells. *Oncogene*. 2001; 20(1):69–76.
146. Sebens Mürköster S, Rausch A, Isberner A, Minkenber J, Blaszczyk E, Witt M, et al. The apoptosis-inducing effect of gastrin on colorectal cancer cells relates to an increased IEX-1 expression mediating NF-κB inhibition. *Oncogene*. 2008; 27(8):1122–34.
147. Akilov OE, Ustyugova IV, Zhi L, Hasan T, Wu MX. Enhanced susceptibility to Leishmania infection in resistant mice in the absence of immediate early response gene X-1. *J Immunol*. 2009; 183(12):7994–8003. <https://doi.org/10.4049/jimmunol.0900866> PMID: 19923449
148. Leisching G, Pietersen RD, van Heerden C, van Helden P, Wiid I, Baker B. RNAseq reveals hypervirulence-specific host responses to *M. tuberculosis* infection. *Virulence*. 2017; 8(6):848–58. <https://doi.org/10.1080/21505594.2016.1250994> PMID: 27763806
149. Coultas L, Chawengsaksophak K, Rossant J. Endothelial cells and VEGF in vascular development. *Nature*. 2005; 438(7070):937–45. <https://doi.org/10.1038/nature04479> PMID: 16355211
150. Ferrara N, Gerber HP. The role of vascular endothelial growth factor in angiogenesis. *Acta Haematol*. 2001; 106(4):148–56. <https://doi.org/10.1159/000046610> PMID: 11815711
151. Ferrara N, Alitalo K. Clinical applications of angiogenic growth factors and their inhibitors. *Nature medicine*. 1999; 5(12):1359–64. <https://doi.org/10.1038/70928> PMID: 10581076
152. Shibuya M. Vascular Endothelial Growth Factor (VEGF) and Its Receptor (VEGFR) Signaling in Angiogenesis: A Crucial Target for Anti- and Pro-Angiogenic Therapies. *Genes Cancer*. 2011; 2(12):1097–105. <https://doi.org/10.1177/1947601911423031> PMID: 22866201
153. Davuluri KS, Singh AK, Kumar V, Singh SV, Singh AV, Kumar S, et al. Stimulated expression of ELR+ chemokines, VEGFA and TNF-AIP3 promote mycobacterial dissemination in extrapulmonary

- tuberculosis patients and *Cavia porcellus* model of tuberculosis. *Tuberculosis (Edinb)*. 2022; 135:102224. <https://doi.org/10.1016/j.tube.2022.102224> PMID: 35763913
154. Delemarre EM, van Hoorn L, Bossink AWJ, Drylewicz J, Joosten SA, Ottenhoff THM, et al. Serum Biomarker Profile Including CCL1, CXCL10, VEGF, and Adenosine Deaminase Activity Distinguishes Active From Remotely Acquired Latent Tuberculosis. *Frontiers in Immunology*. 2021; 12:725447. <https://doi.org/10.3389/fimmu.2021.725447> PMID: 34691031
  155. Golubinskaya EP, Filonenko TG, Kramar TV, Yermola YA, Kubyskin AV, Gerashenko AV, et al. Dysregulation of VEGF-dependent angiogenesis in cavernous lung tuberculosis. *Pathophysiology*. 2019; 26(3–4):381–7. <https://doi.org/10.1016/j.pathophys.2019.11.004> PMID: 31791834
  156. Harding JS, Herbath M, Chen Y, Rayasam A, Ritter A, Csoka B, et al. VEGF-A from Granuloma Macrophages Regulates Granulomatous Inflammation by a Non-angiogenic Pathway during Mycobacterial Infection. *Cell Rep*. 2019; 27(7):2119–31.e6. <https://doi.org/10.1016/j.celrep.2019.04.072> PMID: 31091450
  157. Tao R-H, Kawate H, Ohnaka K, Ishizuka M, Hagiwara H, Takayanagi R. Opposite effects of alternative TZF spliced variants on androgen receptor. *Biochemical and biophysical research communications*. 2006; 341(2):515–21. <https://doi.org/10.1016/j.bbrc.2005.12.213> PMID: 16446156
  158. Pioli PD, Debnath I, Weis JJ, Weis JH. Zfp318 regulates IgD expression by abrogating transcription termination within the Ighm/Ighd locus. *J Immunol*. 2014; 193(5):2546–53. <https://doi.org/10.4049/jimmunol.1401275> PMID: 25057009
  159. Enders A, Short A, Miosge LA, Bergmann H, Sontani Y, Bertram EM, et al. Zinc-finger protein ZFP318 is essential for expression of IgD, the alternatively spliced Igh product made by mature B lymphocytes. *Proceedings of the National Academy of Sciences*. 2014; 111(12):4513–8. <https://doi.org/10.1073/pnas.1402739111> PMID: 24616512
  160. Schwickert TA, Tagoh H, Schindler K, Fischer M, Jaritz M, Busslinger M. Ikaros prevents autoimmunity by controlling anergy and Toll-like receptor signaling in B cells. *Nat Immunol*. 2019; 20(11):1517–29. <https://doi.org/10.1038/s41590-019-0490-2> PMID: 31591571