

# **Cathepsin Z is a conserved susceptibility factor underlying tuberculosis severity**

**Short title:** Cathepsin Z and tuberculosis severity

## **Authors:**

Rachel K. Meade<sup>1,2,\*</sup>, Oyindamola O. Adefisayo<sup>1,\*</sup>, Marco T. P. Gontijo<sup>1,\*</sup>, Summer J. Harris<sup>1</sup>, Charlie J. Pyle<sup>1</sup>, Kaley M. Wilburn<sup>1</sup>, Alwyn M. V. Ecker<sup>1</sup>, Erika J. Hughes<sup>1,2</sup>, Paloma D. Garcia<sup>1</sup>, Joshua Ivie<sup>3</sup>, Michael L. McHenry<sup>4</sup>, Penelope H. Benchek<sup>4</sup>, Harriet Mayanja-Kizza<sup>5</sup>, Jadee L. Neff<sup>6</sup>, Dennis C. Ko<sup>1,2</sup>, Jason E. Stout<sup>7</sup>, Catherine M. Stein<sup>4,8</sup>, Thomas R. Hawn<sup>3</sup>, David M. Tobin<sup>1,2,9</sup>, Clare M. Smith<sup>1,2,#</sup>

## **Affiliations:**

<sup>1</sup> Department of Molecular Genetics and Microbiology, Duke University, Durham, NC, USA

<sup>2</sup> University Program in Genetics and Genomics, Duke University, Durham, NC, USA

<sup>3</sup> Department of Medicine, University of Washington, Seattle, WA, USA

<sup>4</sup> Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, OH, USA

<sup>5</sup> Uganda-CWRU Research Collaboration and Department of Medicine, School of Medicine, Makerere University, Kampala, Uganda

<sup>6</sup> Department of Pathology, Duke University, Durham, NC, USA

<sup>7</sup> Division of Infectious Disease and International Health, Department of Medicine, Duke University Medical Center, Durham, NC, USA

<sup>8</sup> Department of Medicine, Case Western Reserve University, Cleveland, OH, USA

<sup>9</sup> Department of Integrative Immunobiology, Duke University, Durham, NC, USA

\* Denotes co-first authors

# Correspondence: [clare.m.smith@duke.edu](mailto:clare.m.smith@duke.edu)

**Keywords:** Host-pathogen interactions, tuberculosis, mycobacteria, cathepsins, inflammation, disease, systems genetics, genomics, genetic diversity, natural variation, mouse models, quantitative trait locus, Collaborative Cross

**Abstract:** Tuberculosis (TB) outcomes vary widely, from asymptomatic infection to mortality, yet most animal models do not recapitulate human phenotypic and genotypic variation. The genetically diverse Collaborative Cross mouse panel models distinct facets of TB disease that occur in humans and allows identification of genomic loci underlying clinical outcomes. We previously mapped a TB susceptibility locus on mouse chromosome 2. Here, we identify cathepsin Z (*Ctsz*) as a lead candidate underlying this TB susceptibility and show that *Ctsz* ablation leads to increased bacterial burden, CXCL1 overproduction, and decreased survival in mice. *Ctsz* disturbance within murine macrophages enhances production of CXCL1, a known biomarker of TB severity. From a Ugandan household contact study, we identify significant associations between *CTSZ* variants and TB disease severity. Finally, we examine patient-derived TB granulomas and report *CTSZ* localization within granuloma-associated macrophages, placing human *CTSZ* at the host-pathogen interface. These findings implicate a conserved *CTSZ*-CXCL1 axis in humans and genetically diverse mice that mediates TB disease severity.

# Introduction

*Mycobacterium tuberculosis* (*Mtb*), the causative agent of tuberculosis (TB), is a prolific obligate pathogen that has threatened human health for millennia [1]. Through centuries of co-evolution, human hosts have developed a plethora of immunological mechanisms in response to *Mtb* infection [2]. Such host-bacterial interactions give rise to a spectrum of disease states, ranging from subclinical infection to fulminant disease [3]. The disease severity experienced by an individual is intricately connected to their genetic background. For example, monozygotic twins are at a demonstrably higher risk for TB concordance than dizygotic twins, highlighting shared genetic identity as a contributor to TB disease outcomes [4–6]. Human genome-wide association studies (GWAS) conducted in impacted geographic regions have also identified polymorphisms that modulate host TB immunity [7–13], indicating numerous immunological pathways involved in *Mtb* susceptibility.

One such gene is cathepsin Z (*CTSZ*), which has been associated with TB susceptibility in independent human studies conducted across Africa. *CTSZ* encodes a lysosomal cysteine protease with a known structure and several reported cellular functions [14–22]. The link between single nucleotide polymorphisms (SNPs) in *CTSZ* and human TB susceptibility was first established by sibling pair analysis in South African and Malawian populations and independent case-control studies in West Africa [23]. These findings were further validated in a South African case-control study [24] and in a Ugandan GWAS [25] and subsequent household contact study [26]. *CTSZ* is primarily expressed by monocytes and macrophages [27–30], and participates in central immune functions,

including dendritic cell maturation [31] and lymphocyte propagation and migration [32,33]. Although *in vitro* work has been undertaken to study the role of CTSZ in macrophage-driven protection against mycobacteria [34,35], CTSZ-linked TB susceptibility has not been explored *in vivo*. The functional role of CTSZ during *Mtb* infection remains unknown, despite growing genetic evidence of its association with TB disease outcomes.

Studying the mechanisms that underlie CTSZ-linked susceptibility in humans is complex [36]. Humans are genetically outbred, and the low- and middle-income countries that harbor 80% of the global TB burden face challenges in and outside of the healthcare sector that complicate diagnosis, research, and treatment [37]. The connection between TB severity and host background is not uniquely human. Classic studies measuring survival in inbred mice infected with *Mtb* illustrate the heritability of TB susceptibility [38,39], establishing mice as tractable models that demonstrate the vital impact of host genetic background on *Mtb* pathogenesis. However, because inbred mice are nearly genetically identical within strain [40], studies leveraging standard inbred strains omit the contributions of natural host genetic diversity to TB pathogenesis. Recombinant inbred panels like the biparental BXD [41–44], and octoparental Collaborative Cross (CC) [45–47] systematically model host genetic variation, allowing insight into a spectrum of immune profiles without compromising the reproducibility of inbred strains [48]. We previously reported *Mtb* infection screens of BXD [49] and CC [50] recombinant inbred strains, leveraging these diverse mammalian panels to expand the range of known TB disease complexes and host-pathogen interactions modeled by mice. Using a quantitative trait locus (QTL) mapping approach across a cohort of 52 CC genotypes, we identified a

QTL on chromosome 2 (174.29-178.25Mb) significantly associated with *Mtb* burden. Genetic inheritance from NOD/ShiLtJ (NOD), a CC panel founder, at the *Tuberculosis ImmunoPhenotype 5* (*Tip5*) QTL predicted elevated bacterial burden. CC strains that inherited the susceptible *Tip5* variant (*Tip5<sup>S</sup>*) from NOD succumbed to severe TB prior to the study endpoint. We therefore sought to determine whether genes found within the *Tip5* interval contribute to *Mtb* susceptibility in *Tip5<sup>S</sup>* CC strains.

Here, we show that CC strains harboring the *Tip5<sup>S</sup>* locus produce lower levels of CTSZ protein while exhibiting higher bacterial burden than B6 mice following aerosol infection, validating *Tip5* as a susceptibility locus from the large-scale CC cohort screen. We report the first *in vivo* *Mtb* infections of mice lacking *Ctsz* (*Ctsz<sup>-/-</sup>*). We find that *Ctsz* ablation on a B6 background results in increased *Mtb* burden and an increased risk of mortality following infection. Moreover, *Ctsz<sup>-/-</sup>* mice overproduce CXCL1, a biomarker of active TB [51], at both acute and chronic timepoints. In *Ctsz<sup>-/-</sup>* bone-marrow derived macrophages (BMDMs), we find that CXCL1 is rapidly induced following mycobacterial infection. Leveraging published transcriptional data from genetically diverse mice, humans, macaques, and zebrafish, we find cathepsin Z expression is highest in macrophages following infection. We combine these findings with recent data from a Ugandan patient cohort, highlighting 5 variants in *CTSZ* as correlates of TB severity. Finally, we identify the presence of CTSZ in CD68<sup>+</sup> macrophages within patient-derived pulmonary granulomas, revealing that CTSZ is produced at the host-pathogen interface in human lungs. Collectively, this work establishes genetic variation in cathepsin Z as a determinant

of TB disease outcomes and places human CTSZ in a vital position within the pulmonary microenvironment to impact TB outcomes.

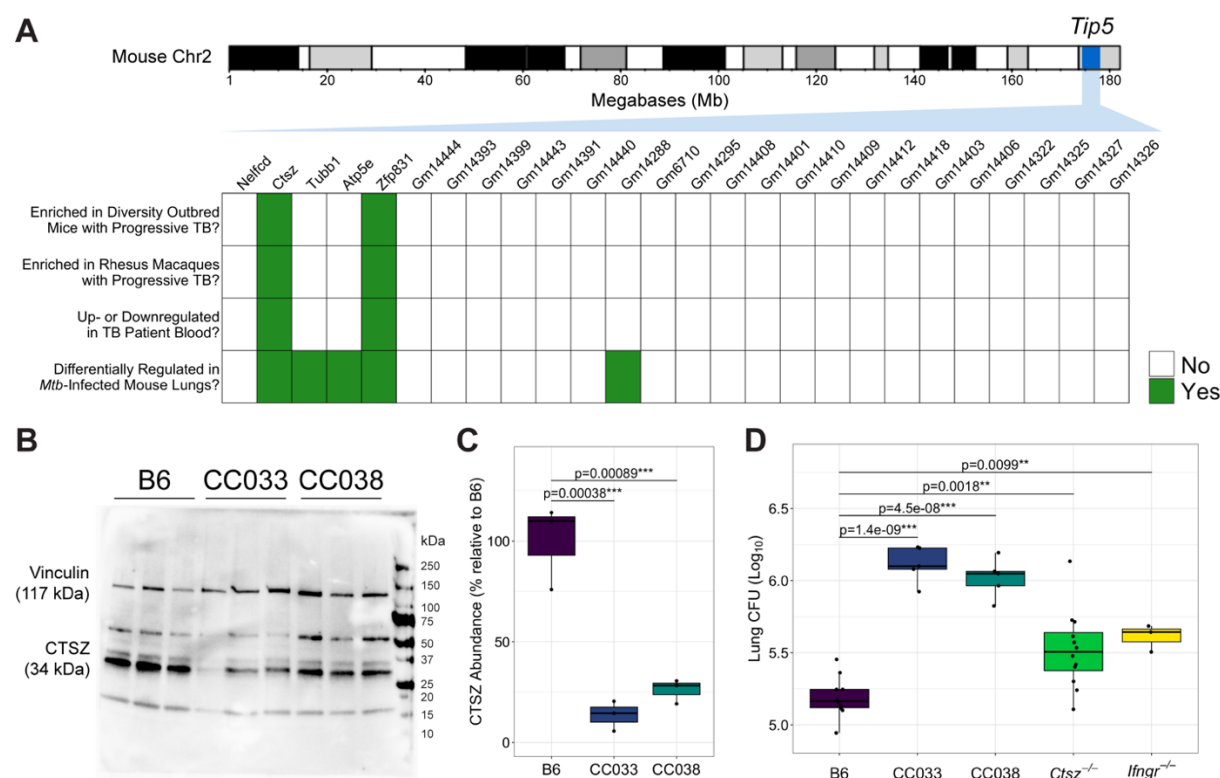
## Results

### *Comparative transcriptional analysis to prioritize candidate genes within the Tip5 locus*

We previously reported the *Tip5* QTL (Chr2, 174.29-178.25Mb) as a TB susceptibility locus across the genetically diverse CC panel [50]. To identify gene candidates within *Tip5*, we leveraged published transcriptomic data from *Mtb*-infected mammalian lungs [52,53] (**Figure 1A**). Within the *Tip5* interval, cathepsin Z (*Ctsz*; also, cathepsin X or P) and zinc finger protein 831 (*Zfp831*) were significantly induced in the lungs of genetically heterogeneous Diversity Outbred (DO) mice exhibiting progressive TB, characterized by elevated pulmonary *Mtb* burden and inflammation [52]. In rhesus macaques, animals with progressive TB disease produced significantly more *CTSZ* and *ZNF831* transcript in their lungs [52]. In the blood of patients with active TB, *CTSZ* transcription was significantly elevated while *ZNF831* transcription was significantly repressed [52,54]. In an additional lung transcriptomic study in inbred mice leveraging distinct *Mtb* strains and infectious doses, only 5 gene transcripts within *Tip5*, including *Ctsz*, were differentially regulated across all strains and doses [53]. Currently, there is no established association between human *ZNF831* SNPs and TB outcomes. Conversely, mutations in human *CTSZ* were previously associated with poorer TB outcomes [23,24,26]. From this analysis, *Ctsz* was identified as a lead candidate for further interrogation as a potential genetic cause of *Tip5*-linked TB susceptibility.

# *The susceptible NOD variant of Tip5 and ablation of Ctsz both impart TB susceptibility*

To evaluate *Ctsz* as a causal factor underlying *Tip5*-linked susceptibility, we measured CTSZ protein from the lungs of uninfected CC strains harboring the susceptible NOD *Tip5* variant (CC033, CC038). Compared to *Mtb*-resistant B6, the lungs of both CC033 and CC038 exhibited significantly lower baseline levels of CTSZ protein (**Figure 1B & C**). Collectively, these data suggest that the NOD *Tip5* haplotype contains a hypomorphic variant of *Ctsz*, resulting in reduced production of CTSZ protein in *Tip5<sup>S</sup>* CC strains.



**Figure 1: Identification and validation of *Ctsz* as the lead candidate gene underlying *Tip5*.** (A) Heatmap representation of the per-gene outcome of four distinct criteria for genes within the *Tip5* QTL (95% CI: 174.29-178.25Mb): i) whether the gene transcript is significantly enriched in the lungs of genetically heterogeneous Diversity Outbred (DO) mice experiencing elevated burden and inflammation after *Mtb* infection [52], ii) whether the gene transcript is significantly enriched in the lungs of rhesus macaques exhibiting clinical symptoms of severe TB disease [52], iii) whether the gene is significantly up- or downregulated in the blood of individuals with active TB [54], and iv) whether the gene is differentially expressed in inbred mouse lungs across variable host genotypes, *Mtb* strains, and infectious doses [53]. To be included in the heatmap, genes were required to encode proteins and to contain a known SNP



from the NOD inbred line [55]. Mouse chromosome 2 image generated in the R package karyoploteR. **(B)** CTSZ protein was measured from the lung homogenate of uninfected B6 and the *Tip5<sup>S</sup>* CC strains CC033 and CC038 (n=3 mice per genotype). Each lane is a separate biological replicate. Vinculin served as the loading control. **(C)** Relative abundance of the CTSZ protein between B6 and the *Tip5<sup>S</sup>* CC strains, quantified from **Figure 1B** by normalizing CTSZ levels for each biological replicate to its respective vinculin level. Values plotted as a percentage of the mean CTSZ to vinculin band intensity ratio relative to the average ratio for B6 mice. Hypothesis testing was performed by one-way ANOVA and Dunnett's *post hoc* test on individual ratios between CTSZ and vinculin band intensities by genotype. **(D)** Bacterial burden measured from lung homogenate 4 weeks after aerosol infection with *Mtb* H37Rv (n=3-12 mice per strain; all males except B6 and *Ctsz*<sup>-/-</sup> groups, which included both sexes in equal proportion). Hypothesis testing was performed by one-way ANOVA and Dunnett's *post hoc* test on log<sub>10</sub>-transformed values.

Considering the *Tip5* QTL was first identified in a large-scale *in vivo* screen, we next assessed whether *Tip5<sup>S</sup>* CC strains and *Ctsz* null mice (*Ctsz*<sup>-/-</sup>) (**Figure S1A-B**) are susceptible to aerosol infection, the natural route of *Mtb* infection. A cohort including B6, CC033, CC038, *Ctsz*<sup>-/-</sup>, and highly susceptible interferon gamma receptor null mice (*Ifngr*<sup>-/-</sup>) [56] was infected via aerosol route with *Mtb* H37Rv. The experiment terminated at 4 weeks post-infection, after the onset of adaptive immunity [57] and matching the initial CC screen endpoint [50]. Relative to B6, all infected strains exhibited significantly higher pulmonary *Mtb* burden (**Figure 1D**). The CC strains exhibited tenfold greater lung CFU than B6, surpassing the canonically susceptible *Ifngr*<sup>-/-</sup> mice. *Ctsz*<sup>-/-</sup> mice exhibited a twofold increase in lung burden relative to B6. No significant differences were identified in disseminated spleen burden at this timepoint (**Figure S1C**). We conclude that *Tip5<sup>S</sup>* CC strains and *Ctsz*<sup>-/-</sup> mice exhibit reduced pulmonary bacterial control at 4 weeks post-infection.

# *Ctsz* mediates lung CXCL1 levels early during *Mtb* infection

To characterize the impact of *Ctsz* on disease progression, we infected B6 and *Ctsz*<sup>-/-</sup> mice via aerosol, sacrificing cohorts of mice at 2, 3, 4, and 8 weeks post-infection to



capture innate and adaptive immune responses. *Ctsz*<sup>-/-</sup> mice exhibited higher lung burden at 2 weeks (4.09 log<sub>10</sub> CFU vs. 3.41 in B6; p<0.05) and 4 weeks (5.17 log<sub>10</sub> CFU vs. 4.09 in B6; p<0.05) post-infection (**Figure 2A**). Similarly, at 3 weeks post-infection, *Ctsz*<sup>-/-</sup> mice exhibited trends toward elevated spleen burden (2.68 log<sub>10</sub> CFU vs. 2.17 in B6; p=0.058), suggesting earlier dissemination and weaker bacterial containment in the lungs of *Ctsz*<sup>-/-</sup> mice (**Figure 2B**). However, by 4 weeks post-infection, spleen burden was indistinguishable between *Ctsz*<sup>-/-</sup> and B6.

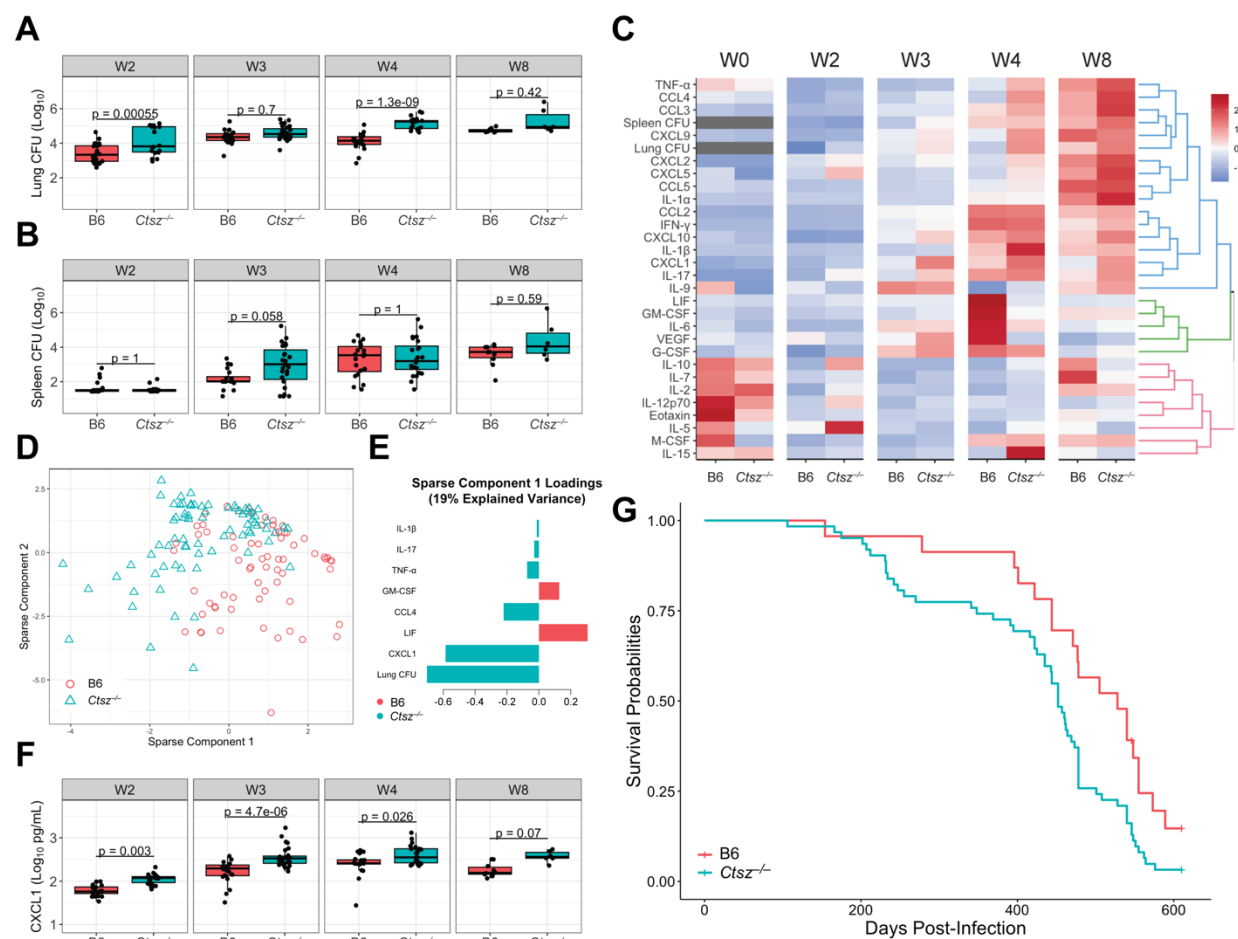
To profile the impact of *Ctsz* disturbance on the lung inflammatory response throughout the course of infection, we compared cytokine signatures of *Ctsz*<sup>-/-</sup> with B6 at assayed timepoints. At 4 weeks post-infection, *Ctsz*<sup>-/-</sup> mice exhibited higher concentrations of T<sub>H</sub>1-associated cytokines, like TNF-α (p=0.019) and IL-1β (p=0.016), and lower levels of GM-CSF (p=3.8e-06), IL-6 (p=5.9e-04), LIF (p=6.6e-07), and VEGF (p=6.6e-07) compared to B6 (**Figure 2C**).

To identify the inflammatory signature of *Ctsz*<sup>-/-</sup> mice in an unsupervised manner, we performed sparse partial least squares discriminant analysis (sPLS-DA) across measured phenotypes (**Figure 2D**). Higher lung burden and CXCL1 levels in *Ctsz*<sup>-/-</sup> mice were the strongest features underlying sparse component 1 (**Figure 2E**). Although component 1 explains 19% of variance in the data compared to 23% variance explained by component 2 (**Figure S2A**), component 1 better captures the variance attributable to genotype. CXCL1 has previously been identified as a biomarker of active TB disease in genetically diverse mice [51] and in humans [58]. From 2 to 4 weeks post-infection, *Ctsz*<sup>-/-</sup> mice

exhibited significantly higher lung CXCL1 levels (**Figure 2F**), suggesting that *Ctsz* ablation increases disease severity. However, by 8 weeks post-infection, although mean CXCL1 levels in *Ctsz*<sup>-/-</sup> lungs were elevated, the difference was no longer significant. Enhanced production of CXCL1 was consistent throughout infection, suggesting that this effect may occur independent of differences in *Mtb* burden.

To explore the possibility that elevated CXCL1 levels may occur independent of infection in *Ctsz*<sup>-/-</sup> mice, we sacrificed uninfected mice of both sexes. From lung homogenate, we found elevated levels of CXCL1 in *Ctsz*<sup>-/-</sup> compared to B6 (**Figure 2C**; p=0.007), suggesting that the connection between *Ctsz* and CXCL1 extends beyond the context of infection. Notably, the total CXCL1 levels in uninfected mice were comparable to levels measured at 2 weeks post-infection.

To determine whether *Ctsz* ablation alone is sufficient to confer susceptibility to aerosolized *Mtb* H37Rv, we conducted two longitudinal challenges of B6 and *Ctsz*<sup>-/-</sup> mice in which mice were sacrificed when IACUC-approved humane endpoints were reached. *Ctsz* ablation was associated with a significant reduction in survival time (p=0.008; **Figure 2G**), which was driven by male mice (**Figure S2B & C**). Thus, disease progression in a host lacking *Ctsz* is characterized by increased lung *Mtb* burden, lung CXCL1 levels, and overall mortality risk.



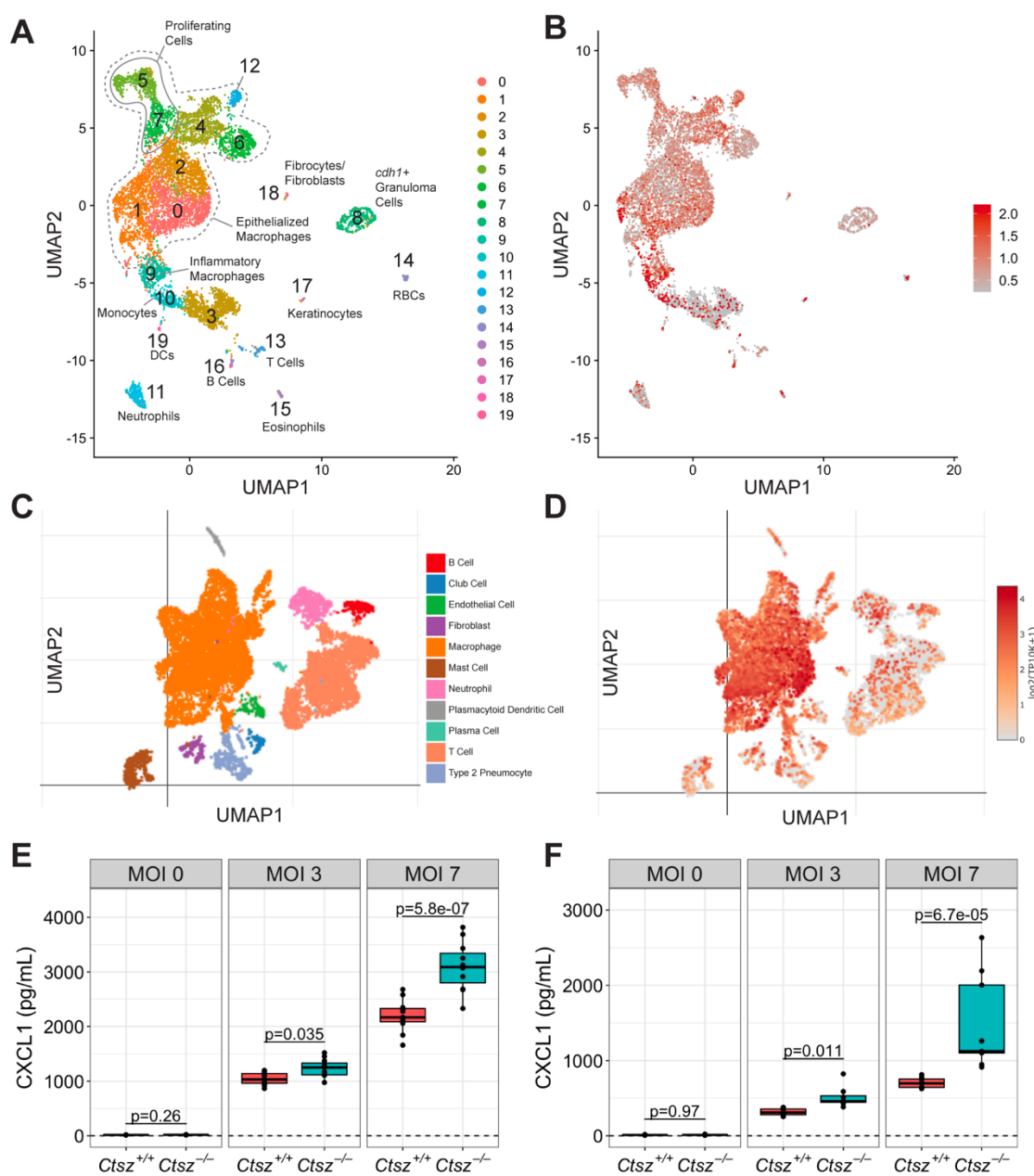
**Figure 2: *Ctsz*<sup>-/-</sup> mice have higher lung burden and earlier spleen dissemination during acute infection followed by greater chronic inflammation and mortality risk.** *Mtb* burden measured from (A) lung and (B) spleen homogenate by dilution plating. (C) Heatmap depicting scaled and centered phenotypes, hierarchically clustered and separated into 3 k clusters. (D) Individual mice plotted against the first two sPLS-DA components, which explained the greatest variance in the data after optimization. (E) Phenotype loadings contributing to component 1. Component 2 loadings shown in **Figure S2A**. (F) CXCL1 levels measured from lung homogenate by multiplex ELISA. For panels A, B, and F, hypothesis testing was performed by two-way ANOVA and Tukey's *post hoc* test on log<sub>10</sub>-transformed values. For panels A-F, mice were sacrificed at 2, 3, 4, and 8 weeks following aerosolized *Mtb* H37Rv infection. Data are from two independent experiments with n=6-14 mice per genotype, representative of both sexes, at each timepoint. In panel C, age-matched, uninfected mice (n=3-4 per genotype and sex) were assayed for comparison. (G) Kaplan-Meier survival estimates of aerosol-infected B6 (n=23) and *Ctsz*<sup>-/-</sup> mice (n=62) across two independent experiments. Equal proportions of both sexes were included. Mice that were not moribund at time of sacrifice were censored for analysis.

207

208 *Disturbance of Ctsz enhances CXCL1 induction in macrophages*

To explore the expression of cathepsin Z across species and mycobacterial infection models, we analyzed two previously published single-cell RNA sequencing (scRNA-Seq) datasets. In zebrafish infected with *Mycobacterium marinum* (*Mm*), *ctsz* was most highly expressed in inflammatory macrophages (cluster 9) after 14 days of infection (**Figure 3A & B**) [59]. *CTS*Z in cynomolgus macaques was most highly expressed in macrophages 4 weeks after *Mtb* infection compared to other assayed cell types (**Figure 3C & D**) [60]. These results agree with literature establishing the presence of CTSZ in monocytes and macrophages [27–30] and further highlight that cathepsin Z expression in these cell types following mycobacterial infection is conserved across diverse host species.

As cathepsin Z is consistently expressed in macrophages across several species following mycobacterial infection, we sought to characterize the impact of *Ctsz* ablation on the initial macrophage response to mycobacterial exposure. To test if macrophages contribute to the increased production of CXCL1 during infection in *Ctsz*<sup>−/−</sup> mice, we generated BMDMs from *Ctsz*<sup>+/+</sup> and *Ctsz*<sup>−/−</sup> sibling pairs. When infected with either non-pathogenic *Mycobacterium bovis* (Bacillus Calmette-Guérin; BCG) (**Figure 3E**) or *Mtb* (**Figure 3F**), *Ctsz*<sup>−/−</sup> macrophages produced greater amounts of CXCL1 than *Ctsz*<sup>+/+</sup> by 24 hours post-infection. In both infection models, this effect scaled with increasing multiplicity of infection (MOI). Thus, the elevated CXCL1 we observed in *Ctsz*<sup>−/−</sup> lungs may be driven by macrophages, especially during the early stages of infection, and appears to be independent of mycobacterial pathogenicity. These results from *Mtb*-infected *Ctsz*<sup>−/−</sup> mice and BMDMs suggest an interaction between CTSZ and CXCL1 following bacterial exposure.



**Figure 3: Cathepsin Z is highly expressed in macrophages across species following mycobacterial infection and mediates levels of CXCL1 in murine macrophages.** (A) UMAP showing scRNA-Seq results from zebrafish granulomas infected with *M. marinum* (*Mm*), first published by Cronan, Hughes et al., 2021 [59]. Cells are colored by cluster and assigned in an unsupervised approach from transcriptional signatures. Clusters were annotated by the authors. (B) ScRNA-Seq data colored by relative expression of zebrafish *ctsz* in *Mm*-infected granulomas, with highest expression levels observed in cluster 9. (C) UMAP showing scRNA-Seq of granulomas extracted from *Mtb*-infected cynomolgus macaques at 4 weeks post-infection, published by Gideon et al., 2022 [60]. Data accessed at the Broad Institute Single Cell online repository on October 3, 2023 ([https://singlecell.broadinstitute.org/single\\_cell/study/SCP1749/](https://singlecell.broadinstitute.org/single_cell/study/SCP1749/)). (D) ScRNA-Seq data colored by relative expression of *CTSZ* in *Mtb*-infected macaque granulomas. CXCL1 levels measured

in triplicate at 24 hours post-infection from (E) BCG-infected and (F) *Mtb*-infected BMDMs by ELISA. For panels E and F, BMDMs were differentiated from independent pairs of *Ctsz*<sup>+/+</sup> and *Ctsz*<sup>-/-</sup> sibling males for each infection (N=3 infections per pathogen). Dashed threshold denotes the limit of detection for the ELISA. Statistical significance was determined by two-way ANOVA and Tukey's *post hoc* test on batch-corrected, log<sub>10</sub>-transformed values.

### 233 Variants in human CTSZ are associated with TB severity

234 To investigate the impact of natural *CTSZ* variation on human TB disease outcomes, we  
 235 examined whether human *CTSZ* variants are associated with TB disease severity in a  
 236 household contact study in Uganda (n=328 across two independent cohorts) [61]. Of 81  
 237 observed *CTSZ* SNPs, 20 SNPs were associated with differences in Bandim TBScore, a  
 238 TB severity index (**Table S1**; unadjusted p<0.05, linear model with sex, HIV status, and  
 239 genotypic principal components 1 and 2 as covariates) [62]. After performing a Bonferroni  
 240 adjustment for multiple comparisons, 4 SNPs and 1 INDEL maintained associations with  
 241 TB disease severity (**Table 1**). These variants are in strong linkage disequilibrium (LD)  
 242 with one another ( $R^2 > 0.8$ ), suggesting that they represent a single haplotype block  
 243 (**Figure 4A**). For the most significant SNP (rs113592645), the minor T allele is associated  
 244 with decreased TB disease severity (**Figure 4B**, results for other haplotype SNPs  
 245 included in **Figure S3A-C**). To investigate the potential impact of the TB severity SNPs  
 246 on *CTSZ* expression, we used published RNA-Seq data [63] to compare *CTSZ* transcript  
 247 levels across *Mtb*- and mock-infected monocytes between genotypes at each *CTSZ* SNP.  
 248 In the cohort of human-derived monocytes, *CTSZ* was highly expressed at baseline and  
 249 was downregulated following *Mtb* infection (**Figure 4C**). Conversely, the rs113592645  
 250 minor T allele was associated with increased *CTSZ* expression following *Mtb* infection  
 251 (p=0.0395; **Figure 4D**; other haplotype SNP results in **Figure S3D-F**). This effect was not  
 252 observed following mock infection conditions (p=0.108; **Figure 4D**). Together, these data

suggest that *CTSZ* variants are associated with both TB disease severity and divergent transcription of *CTSZ* following *Mtb* infection.

**Table 1: *CTSZ* SNPs significantly associated with TB severity in Ugandan household contact study cohorts, sorted by ascending p-value.** Included SNPs were significantly associated with Bandim TBscore after Bonferroni correction for multiple comparisons ( $p < 0.05$ ). Complete collection of 81 SNPs can be found in **Table S1**. SNPs are annotated as described in McHenry et al., 2023 [61]. Allele effects were assessed using a linear mixed effect model in the R package *lme4* to account for sex, HIV status, and genotypic principal component 1 and 2. Cohort 1 and 2 are independent cohorts of culture-confirmed adult TB cases. Abbreviations: SNP, single nucleotide polymorphism; CHR, chromosome; BP, base pair from GRCh38 build; Adj., adjusted; MAF, minor allele frequency.

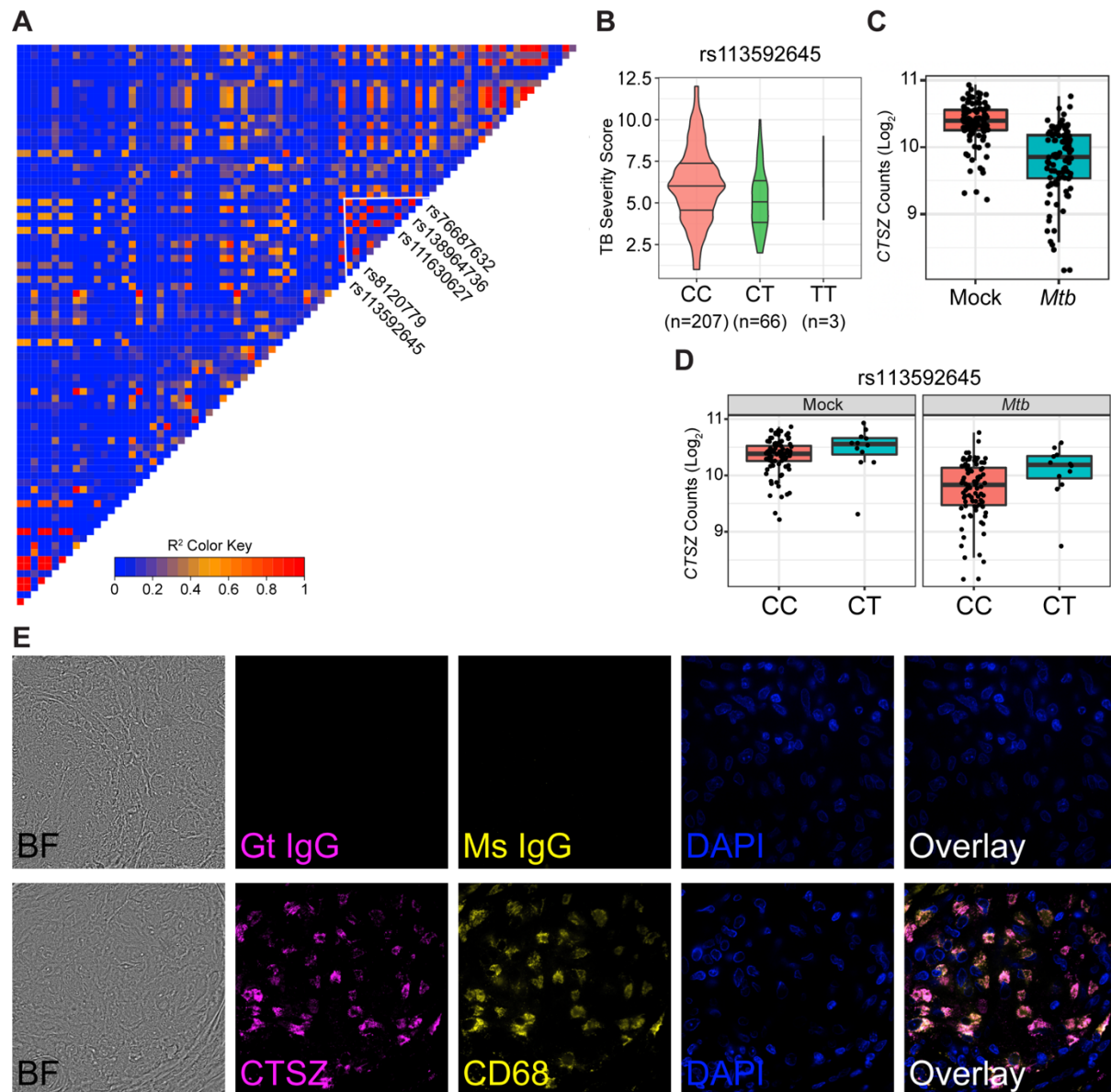
| SNP         | CHR:BP      | Effect Allele | Adj. p-value | $\beta$ | MAF in Cohort 1 (n=149) | MAF in Cohort 2 (n=179) |
|-------------|-------------|---------------|--------------|---------|-------------------------|-------------------------|
| rs113592645 | 20:59001340 | T             | 0.0001814    | -1.0036 | 0.18                    | 0.061                   |
| rs111630627 | 20:59002589 | G             | 0.0003077    | -0.9268 | 0.18                    | 0.075                   |
| rs138964736 | 20:59002671 | ACTTTG        | 0.0003077    | -0.9268 | 0.18                    | 0.075                   |
| rs76687632  | 20:59002905 | G             | 0.0003077    | -0.9268 | 0.18                    | 0.075                   |
| rs8120779   | 20:59001977 | G             | 0.0003942    | -0.8671 | 0.18                    | 0.095                   |

*CTSZ* is produced in macrophages associated with human pulmonary granulomas

The human granuloma is an organized structure that can develop within the host to contain and restrict *Mtb* infection and is composed of heterogeneous immune cell populations, predominantly macrophages [64]. To investigate whether *CTSZ* production in macrophages is conserved between mice and humans, we performed immunostaining



on granulomas biopsied from the lungs of patients with culture-confirmed TB. We positively identified CTSZ within TB granuloma-associated CD68<sup>+</sup> macrophages from this lung tissue (**Figure 4E**). Thus, CTSZ is produced at the site of host-pathogen interaction in humans, suggesting that native functions at this interface could be interrupted should CTSZ production or localization be impeded. Combined with the results from *Ctsz* null mice, these data suggest that balancing cathepsin Z levels is required to regulate lung inflammation and reduce risk of mortality. Collectively, these data establish an association between human CTSZ variants and TB disease severity and reveal CTSZ as a granuloma-associated protein in human lungs for the first time.



**Figure 4: Human *CTSZ* variants are associated with TB severity, and *CTSZ* protein is produced within the human pulmonary granuloma. (A)** LD plot of human *CTSZ*, highlighting a haplotype block of 4 identified SNPs and 4 SNPs and 1 INDEL associated with TB disease severity. **(B)** Comparison of TB severity, measured using Bandim TBScore, by genotype for the lead TB severity SNP, rs113592645. TB severity score by genotype for remaining SNPs can be found in **Figure S3A-C**. For panels **C** and **D**, *CTSZ* expression was profiled by RNA-Seq in monocytes from 100 Ugandan individuals. Human-derived monocytes were subjected to 6-hour *Mtb* and mock infection conditions. **(C)** Counts of *CTSZ* transcript (log<sub>2</sub> counts per million) collected following mock and *Mtb* infection. **(D)** Counts of *CTSZ* transcript (log<sub>2</sub> counts per million) according to genotype for the lead TB severity SNP, rs113592645, following mock and *Mtb* infection. Measurements for homozygous minor allele (TT) were excluded due to low sample size. Counts of *CTSZ* transcript by genotype of remaining SNPs can be found in **Figure S3D-F**. **(E)** Brightfield (BF) images and immunofluorescent staining of *CTSZ* and CD68 within a granuloma biopsy from an individual with pulmonary TB. Goat (Gt) and mouse (Ms) IgG isotype control staining is depicted in the top row. DAPI staining indicates cell nuclei. Images were captured at 100X magnification.

## Discussion

Over fifteen years have passed since the initial discovery that human *CTSZ* is linked with TB disease susceptibility in West and South Africa. However, the relationship between *Mtb* susceptibility and *CTSZ* had yet to be experimentally determined. We show that genetic interruption of *Ctsz* in mice causes a failure of bacterial restriction and overproduction of CXCL1 during early *Mtb* infection, precipitating an increased risk of mortality. We further show a cell-autonomous effect of CXCL1 overproduction in macrophages during *Mtb* infection. We report 4 *CTSZ* SNPs and 1 INDEL significantly associated with TB severity in Ugandan individuals and show higher *CTSZ* expression in infected monocytes from this cohort. Finally, we find that *CTSZ* protein is present within the CD68<sup>+</sup> macrophages in human granulomas, the pulmonary structure that contains and restricts *Mtb* growth.

*CTSZ* participates in several known immunological pathways [29,32,33,65]. Here, we show that lung CXCL1 levels are consistently elevated in *Ctsz*<sup>-/-</sup> mice prior to and throughout infection. Moreover, compared to wildtype siblings, *Ctsz*<sup>-/-</sup> macrophages produce more CXCL1 in response to pathogenic and non-pathogenic mycobacterial infection, suggesting a broad immunological response to bacterial exposure. CXCL1, a cytokine associated with severe TB disease in mice [51] and in humans [58], is primarily known as a neutrophil chemoattractant. Both *Ctsz* and *Cxcl1* are induced in *Mtb*-infected mice [66], but, to the best of our knowledge, this is the first study to directly link *CTSZ* and CXCL1 during TB pathogenesis. Future studies are needed to delineate the implications

of this CTSZ-CXCL1 axis with other known roles of CTSZ, including cellular adhesion, migration, and antigen presentation [29]. For example, CTSZ is known to interact with cell surface integrins that mediate immune cell activity, including lymphocyte function-associated antigen-1 (LFA-1) [32,33] and macrophage-1 (Mac-1) antigen [65], which regulates *Mtb* phagocytosis [67] and phagocyte migration.

Deeper understanding of how the functions of CTSZ impact disease severity could prove vital to developing therapeutic strategies for both endogenous and infectious diseases beyond TB. For example, CTSZ has been implicated as a mediator of host response during *Helicobacter pylori* infections of patient-derived monocytes and *Salmonella* Typhimurium infections of murine BMDMs [68,69]. Further, mouse and human studies have investigated CTSZ for roles in aging [70,71] and in a number of endogenous conditions, including multiple sclerosis [72], primary biliary cholangitis [73,74], osteoporosis [75], and Alzheimer's [76]. CTSZ has also been explored for prognostic value and roles in tumor progression across many cancers [77], including breast [78], colorectal [79], gastric [68], and prostate cancers [80], and hepatocellular carcinoma [81]. Increased CTSZ expression was associated with poor patient prognoses in some studies [79,81], with one study proposing CTSZ as a putative oncogene [81]. Given the importance of CTSZ in a spectrum of human disease categories, continued study of CTSZ may yield insights on the human response to departures from immune homeostasis.

While much remains unknown about the molecular roles of CTSZ during *Mtb* infection, this study is the first to identify cathepsin Z as a molecular correlate of TB severity in mice

and humans. This study is also the first to report CTSZ localization within granuloma-associated CD68<sup>+</sup> macrophages in *Mtb*-infected human lungs. Host genetic diversity is a central predictor of TB severity, and consideration of genetic diversity is essential to combat human pathogens as enduring and prolific as *Mtb*.

### **Data Availability Statement**

Relevant data is included in the manuscript and supplemental files. Additional data supporting this manuscript will be made available upon request.

### **Acknowledgements**

The authors acknowledge Martin Ferris, Rachel Lynch, and Ginger Shaw for coordination of experimental cohorts of CC mice through the UNC Systems Genetics Core Facility (SGCF); Christopher Sassetti, Douglas Marchuk, and Craig Lowe for thoughtful critiques and suggestions; and the staff of the Regional Biosafety Laboratory at Duke University for ongoing support of personnel safety in BSL-3 biocontainment. We acknowledge the staff of the Duke University BioRepository & Precision Pathology Center (BRPC) for identifying the human clinical cases and collecting and preparing the paraffin tissue sections. The authors also acknowledge the vital contribution of the patients whose samples provided data for this paper and the contributions made by senior physicians, medical officers, health visitors, laboratory personnel, and data personnel working with the Uganda-CWRU Research Collaboration. This study would not be possible without the generous participation of Ugandan patients and families.

## Funder Information

This work was funded by an NIH Director's New Innovator Award AI183152 (C.M.Sm.), a Pew Scholars award (C.M.Sm.), and the following NIH grants: AI166304 (D.M.T.), AI127715 (D.M.T. and C.M.Sm.), AI181898 (T.R.H. and C.M.Sm.), AI162583 (T.R.H., C.M.St., P.H.B., H.M.-K.), N01-AI95383 (C.M.St.), and T32HL007567 (M.L.M.). M.T.P.G. was supported by a grant (88881.625374/2021-01) from the Fulbright Association and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Research reported in this publication was supported in part by the Duke University Center for AIDS Research (CFAR), an NIH funded program (5P30 AI064518). The Duke University BRPC is supported in part by the NIH (P30CA014236). Biocontainment work performed in the Duke Regional Biocontainment Laboratory received partial support for construction and renovation from NIAID (UC6-AI058607 and G20-AI167200) and facility support from the NIH (UC7-AI180254). The sponsors or funders did not play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Conflict of Interest

The authors declare no conflicts of interest.

## Materials & Methods

### *Ethics Statement*

All animal studies were conducted in accordance with the guidelines issued in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health and the Office of Laboratory Animal Welfare. Mouse studies were conducted at Duke University

using protocols approved by the Duke Institutional Animal Care and Use Committee (IACUC) (Animal Welfare Assurance #A221-20-11 and #A204-23-10) in a manner designed to minimize pain and suffering in *Mtb*-infected animals. Any animal exhibiting signs of severe disease was immediately euthanized in accordance with IACUC-approved endpoints. Use of patient samples was approved by the Duke University Medical Center Institutional Review Board (IRB) under Protocol #00107795 and the University of Washington IRB under Protocol STUDY00001537. Patient sample processing at Duke University was carried out by Drs. Jadee Neff, Charlie Pyle, and Jason Stout. The human genetic data were obtained from the Kawempe Community Health Study in Uganda, which was approved by the National HIV/AIDS Research Committee of Makerere University (Protocol #014) and the University Hospitals Cleveland Medical Center IRB (Protocol #10-01-25). Final clearance was given by the Uganda National Council for Science and Technology (Ref #658).

### *Mice*

Male and female C57BL/6J (#000664) and male B6.129S7-*Ifngr*<sup>1<sup>tm1Agt</sup></sup>/J (*Ifngr*<sup>-/-</sup>; #003288) mice were purchased from The Jackson Laboratory. Male CC033/GeniUncJ (CC033) and CC038/GeniUnc (CC038) mice were purchased from the University of North Carolina (UNC) Systems Genetics Core Facility (SGCF). *Ctsz*<sup>-/-</sup> mice were generously provided by Robin Yates (University of Calgary, Calgary, AB, Canada) [82]. *Ctsz*<sup>+/+</sup> and *Ctsz*<sup>-/-</sup> mice were subsequently bred at Duke. All mice were housed in a specific pathogen-free facility within standardized living conditions (12-hour light/dark, food and water *ad libitum*). Aerosol-infected mice were matched at 8-12 weeks of age at the time



of *Mtb* infection. Mice were individually identified for weighing and wellness assessment throughout infection using Bio Medic Data Systems implantable electronic ID transponders (TP-1000) implanted subcutaneously at the back of the neck prior to infection.

### *Genotyping*

In-house confirmation of *Ctsz*<sup>-/-</sup> genotype was performed using forward primer 5'-TTG CTG TTG GCG AGT GCG-3' and reverse primer 5'-CTT GTC ACC AGA TTC CAG C-3' to detect wildtype *Ctsz* and forward primer 5'-GCT ACC TGC CCA TTC GAC-3' and reverse primer 5'-ACA GTA GGA CTG GCC AGC-3' to detect knockout product. Primer sequences were generously provided by Robin Yates (University of Calgary, Calgary, AB, Canada). DNA was extracted from tissue samples using the DNEasy Blood & Tissue Kit (Qiagen). DNA products were prepared for PCR using Q5 High-Fidelity Master Mix (New England BioLabs) and amplified. Protocol included initial 98°C (30s), then 34 cycles of denaturation (98°C, 10s), annealing (68°C, 30s), and extension (72°C, 90s), and a final 72°C (180s), resting at 10°C ∞ until stopped. Amplified products were separated on a 1% agarose-TAE gel using SYBR Safe stain (Thermo Fisher Scientific) and 1kb Plus DNA Ladder (New England BioLabs). *Ctsz*<sup>+/+</sup> and *Ctsz*<sup>-/-</sup> mice were genotyped at the time of weaning from ear and tail tissue biopsies by TransnetYX (Cordova, TN, USA) using proprietary RT-PCR primers designed to detect both *lacZ*, present in the IRES vector disturbing the second exon of *Ctsz* [82], and wildtype *Ctsz*.

### *Bacterial Strains & Culture*

All infections were performed with either *Mtb* H37Rv genotype or *M. bovis* BCG (Bacillus Calmette-Guérin) Danish (gift from Sunhee Lee, University of Texas Medical Branch, Galveston, TX, USA), which was transformed with pTEC-15 wasabi fluor and possesses a hygromycin resistance marker for selection [83]. Aerosol infections were performed using an *Mtb* H37Rv strain confirmed to be positive for the cell wall lipid and virulence factor phthiocerol dimycocerosate (PDIM; gift from Kyu Y. Rhee, Weill Cornell Medicine, New York, NY, USA). Bacteria were cultured in Middlebrook 7H9 medium supplemented with oleic acid-albumin-dextrose catalase (OADC), 0.2% glycerol, and 0.05% Tween 80 (or 0.005% tyloxapol for macrophage infections) to log-phase with shaking (200rpm) at 37°C. Hygromycin (50 µg/mL) was added when necessary. Prior to all *in vivo* infections, cultures were washed and resuspended in phosphate-buffered saline (PBS) containing 0.05% Tween 80 (hereafter PBS-T). Bacterial aggregates were then broken into single cells using a blunt needle before diluting to desired concentration for infection.

### *Mouse Infections*

Mice were infected with ~150-350 *Mtb* CFU via aerosol inhalation (Glas-Col). On the day following each infection, one cage was sacrificed to enumerate lung CFU as an approximation of infectious dose. For all infections, mice were euthanized in accordance with approved IACUC protocols, and lung and spleen were harvested into PBS-T and processed in a FastPrep-24 Homogenizer (MP Biomedicals, 4.0m/s, 45s, 2-3x). *Mtb* burden was quantified by dilution plating onto Middlebrook 7H10 agar supplemented with OADC, 0.2% glycerol, 50 mg/mL Carbenicillin, 10 mg/mL Amphotericin B, 25 mg/mL Polymyxin B, and 20 mg/mL Trimethoprim. Lung homogenate was centrifuged through a

0.2µm filter to collect decontaminated filtrate, and cytokines and chemokines were assayed using the pro-inflammatory focused 32-plex assay (Eve Technologies, Calgary, AB, Canada).

#### *Human Tissue Immunofluorescent Staining*

Patient tissue samples containing *Mtb* granulomas were identified at the Duke University School of Medicine. Clinical tissue specimens were obtained from the Duke Pathology Department, and 5µm paraffin sections for antibody staining were cut by the Research Histology Laboratory within the BioRepository & Precision Pathology Center (BRPC). Paraffin was dissolved using two xylene washes followed by washes with ethanol of increasing dilution (100% twice, 95% twice, 70% once, 50% once), three washes with deionized water, and a final wash in PBS. Sample was placed in antigen retrieval buffer (10 mM Tris/1 mM EDTA, pH 9.0) and processed in a pressure cooker for 10min. Following a cooling step, samples were blocked for an hour in 2.5% normal donkey serum. Samples were incubated overnight at 4°C with Goat anti-Human/Mouse/Rat Cathepsin X/Z/P Polyclonal Antibody (R&D Systems, AF934, 0.185 mg/mL) and Mouse anti-Human CD68 Monoclonal Antibody (Agilent Dako, M081401-2, 0.185 mg/mL) in 2.5% serum in a humidified chamber. Immunoglobulin G (IgG) isotype controls for background staining (Goat: Biotechne, AB-108-C, 1 mg/mL stock; Mouse: GenScript, A01007, 1 mg/mL stock; Rabbit: Invitrogen, 10500C, 3 mg/mL provided) were also used. Primary antibody was removed with three washes of PBS and two of deionized water. Samples were incubated in Alexa Fluor (AF) conjugated secondary antibody (Thermo Fisher Scientific, 1:500; Donkey anti-Goat IgG AF Plus 647: A32849; Donkey anti-Mouse

IgG AF 488: A-21202; Donkey anti-Rabbit IgG AF 555: A-31572) in 2.5% serum for 1-3 hours. Following three PBS washes, the samples were mounted for imaging in DAPI Fluoromount-G (Southern Biotech, 0100-20) on glass slides (Fisher Scientific, 22-035813). All antibodies used for staining were centrifuged at 10000 RCF (4°C) for 10min to remove antibody precipitate prior to use.

### *Microscopy Analysis*

Human samples were imaged at 100X on a Zeiss Axio Observer Z1 inverted microscope with an X-Light V2 spinning disk confocal imaging system (Biovision). Images were processed identically within Fiji software (v2.14.0/1.54f) for image clarity.

### *Bone Marrow-Derived Macrophage Infections*

*Ctsz*<sup>+/+</sup> and *Ctsz*<sup>-/-</sup> sibling pairs were sacrificed in accordance with approved IACUC protocols between 10-12 weeks of age. For BCG infections, bone marrow was flushed from hip and leg bones with DMEM (Corning) and cultured for a week at 37°C in a sterile solution of DMEM with 10% heat-inactivated fetal bovine serum (Corning), 18% 3T3-derived M-CSF, 1X Pen/Strep (Corning), and 25mM HEPES (gibco). For *Mtb* infections, bone marrow was flushed from hip and leg bones with sterile DMEM (Corning) and frozen in 10% DMSO in heat-inactivated fetal bovine serum (Corning). Aliquots were later thawed and cultured for a week at 37°C in a sterile solution of DMEM with 10% heat-inactivated fetal bovine serum (Corning), 30 µg/mL recombinant M-CSF (PeproTech), 1X Pen/Strep (Corning), and 25mM HEPES (gibco). Differentiated macrophages were then plated at a concentration of 4x10<sup>5</sup> cells/well in a 24-well plate and cultured at 37°C

overnight in a DMEM solution as above but without Pen/Strep. BMDMs were infected with BCG or transported to BSL-3 biocontainment for infection with *Mtb* at MOI 3 or 7 or left uninfected. Wells were tested for even infection by CFU plating. At 24 hours post-infection, supernatants were collected and filtered using a 0.2µm filter to remove bacteria. Cytokines and chemokines were assayed from using the high-sensitivity 18-plex discovery assay (Eve Technologies, Calgary, AB, Canada).

### *Western Blotting*

To compare the presence of mouse CTSZ in *Ctsz*<sup>+/+</sup> and *Ctsz*<sup>-/-</sup> mice, protein was extracted from lung and spleen homogenate (12-14 weeks of age) using RIPA buffer (Sigma-Aldrich) and 1X Protease Inhibitor Cocktail (Sigma-Aldrich) after a PBS wash. For the comparison of mouse CTSZ between uninfected B6 mice and *Tip5*<sup>S</sup> CC strains (CC033 and CC038), whole lungs were collected from male mice (8 weeks of age) into 1mL of Trizol reagent. Samples were homogenized with sterile beads at 4.5 m/s for 30s using the FastPrep-24 Homogenizer (MP Biomedicals). For samples in Trizol, protein was precipitated for 15min using 9 volumes of 100% methanol at room temperature. The protein precipitate was centrifuged at 3000rpm for 5min, dried for 5min, and washed in an equal volume of 90% methanol. The protein precipitates were then centrifuged for 1min at 3000rpm, dried for 10min, resuspended in 1mL of RIPA buffer and 1X Protease Inhibitor Cocktail, and heated for 5-10min at 95°C. Equal volumes of each sample were combined with Laemmli Sample Buffer (BioRad) and 2-Mercaptoethanol (BioRad) and heated at 95°C for 5min. SDS-PAGE was performed using BioRad Western Blotting kit along with Precision Plus Protein All Blue Prestained Protein Standards (BioRad). Protein

was separated using a 4-20% Mini-PROTEAN TGX Stain-Free Protein Gel (BioRad) and transferred to a polyvinylidene fluoride (PVDF) membrane using a semi-dry transfer protocol on a Trans-Blot Turbo Transfer System (BioRad). Membrane was blocked using EveryBlot Blocking Buffer (BioRad). Primary staining was performed at 4°C overnight using Human/Mouse/Rat Cathepsin X/Z/P Antibody (R&D Systems; AF934; 1:2000 dilution in EveryBlot Blocking Buffer). For B6 and *Tip5<sup>S</sup>* CC mice, 0.1% Tween 20 was added to the blocking buffer and primary staining also included Vinculin (E1E9V) XP Rabbit mAb (Cell Signaling; #13901; 1:5000 dilution in EveryBlot Blocking Buffer + 0.1% Tween 20). For *Ctsz<sup>+/+</sup>* and *Ctsz<sup>-/-</sup>* mice, secondary staining was performed at room temperature for 60min using Donkey anti-Goat 680 (LI-COR; 1:20000 dilution in EveryBlot Blocking Buffer + 0.1% SDS). Blot was washed in TBS-T between blocking and antibody stains, and fluorescence was measured using a LI-COR Odyssey. Secondary staining for B6 and *Tip5<sup>S</sup>* CC mice was performed at room temperature for 60min using HRP-conjugated Rabbit Anti-Goat IgG (Proteintech; SA00001-4; 1:5000 dilution in EveryBlot Blocking Buffer + 0.1% Tween 20) and HRP-conjugated Goat Anti-Rabbit IgG (Proteintech; SA00001-2; 1:5000 dilution in EveryBlot Blocking Buffer + 0.1% Tween 20). Blot was washed in PBS-T (0.1% Tween 20). Chemiluminescence was developed using SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Fisher Scientific) and imaged using a ChemiDoc Plus Imaging System (BioRad). Quantification of the blot was performed with ImageLab software (version 6.1).

### *Human CTSZ Analysis*

We queried the summary statistics from a published genome-wide association study (GWAS) of TB severity in cases from Kampala, Uganda [61]. Briefly, two independent cohorts of culture-confirmed adult TB cases (n=149, n=179) [84] were included in a GWAS. TB severity was quantified using the Bandim TBscore, which enumerates TB symptoms (e.g., cough, hemoptysis, dyspnea) and clinical signs (e.g., anemia, low body mass index, high body temperature) [62,85]. SNPs within *CTSZ* were identified using a 5kb flanking region around the *CTSZ* start and end positions reported in Ensembl (GRCh38). Pairwise linkage disequilibrium (LD) for these SNPs was evaluated as the squared inter-variant allele count correlations ( $R^2$ ) using PLINK (version 1.90) in the larger of the two cohorts (n=179). An LD plot was generated from these pairwise LD measures using the R package LDheatmap (version 1.0-5) [86]. The model used to estimate allele effects accounted for sex, HIV status, RNA-Seq batch, and genotypic principal components 1 and 2.

SNP eQTL assessment was performed for the four significant SNPs indicated in **Table 1**. A linear mixed effect model was developed in the R package kinma [87] to compare baseline, media, and *Mtb*-induced *CTSZ* expression against each SNP genotype. The eQTL model accounted for sex, HIV status, RNA-Seq batch, and genotypic principal components 1 and 2. *CTSZ* expression as  $\log_2$ (counts per million) was obtained from RNA-Seq data normalized using voom [88]. RNA-Seq data used for these analyses originated from a previously published dataset of CD14<sup>+</sup> monocytes isolated from individuals within the Uganda cohort [63]. Monocytes were subjected to 6-hour media or *Mtb* stimulation and transcriptionally assayed.



558

## 559 *Statistical Analysis & Data Visualization*

560 Hypothesis testing was performed using R statistical software (version 4.3.1). Statistical  
561 tests used for hypothesis testing are noted in figure legends. Shapiro-Wilks tests were  
562 used to assess normality in phenotype data prior to parametric hypothesis testing, and  
563 log<sub>10</sub>-transformation was applied for normalization where appropriate. Kaplan-Meier  
564 survival curves were calculated using the R package survminer (version 0.5.0). A  
565 visualization of mouse chromosome 2 was generated using the R package karyoploteR  
566 (version 1.16.0) from the GRCm38/mm10 mouse genome build. Heatmaps in **Figure 1A**  
567 and **Figure 2C** were generated using the R packages ComplexHeatmap (version 2.21.2)  
568 and heatmaply (version 1.5.0), respectively. Optimization and sparse partial least squares  
569 discriminant analysis (sPLS-DA) on time course infection cohorts were performed on  
570 timepoint data using the R package mixOmics (version 6.24.0).

571

## 572 **Supplemental Table Legend**

573 **Table S1: Complete list of 81 CTSZ SNPs present in Ugandan household contact**  
574 **study cohorts and their associations with TB severity.** TB severity was evaluated by  
575 Bandim TBScore. Summary statistics for the CTSZ variants shown are based on a meta-  
576 analysis of two independent cohorts of culture-confirmed adult TB cases (described in  
577 McHenry et al., 2023 [61]). Each cohort utilized a linear regression model that controlled  
578 for HIV status, sex, and one principal component. Unadjusted p-values are reported.  
579 Abbreviations: CHR, chromosome; BP, base pair from GRCh38 build; MAF, minor allele  
580 frequency.

581

## 582 Literature Cited

- 583 1. Sabin S, Herbig A, Vågane ÅJ, Ahlström T, Bozovic G, Arcini C, et al. A  
584 seventeenth-century *Mycobacterium tuberculosis* genome supports a Neolithic  
585 emergence of the *Mycobacterium tuberculosis* complex. *Genome Biol.*  
586 2020;21:201. doi:10.1186/s13059-020-02112-1
- 587 2. Ernst JD. The immunological life cycle of tuberculosis. *Nat Rev Immunol.*  
588 2012;12:581–591. doi:10.1038/nri3259
- 589 3. Horton KC, Richards AS, Emery JC, Esmail H, Houben RMGJ. Reevaluating  
590 progression and pathways following *Mycobacterium tuberculosis* infection within  
591 the spectrum of tuberculosis. *Proc Natl Acad Sci U S A.* 2023;120:e2221186120.  
592 doi:10.1073/pnas.2221186120
- 593 4. Kallmann FJ, Reisner D. Twin Studies on Genetic Variations in Resistance to  
594 Tuberculosis. *J Hered.* 1943;34:269–276.  
595 doi:10.1093/oxfordjournals.jhered.a105302
- 596 5. Simonds B. Tuberculosis in Twins. London: Pitman; 1963.
- 597 6. Comstock GW. Tuberculosis in twins: A re-analysis of the Proffit survey. *Am Rev*  
598 *Respir Dis.* 1978;117:621–624. doi:10.1164/arrd.1978.117.4.621
- 599 7. Bellamy R, Ruwende C, Corrah T, McAdam KPWJ, Whittle HC, Hill AVS.  
600 Variations in the NRAMP1 Gene and Susceptibility to Tuberculosis in West  
601 Africans. *N Engl J Med.* 1998;338:640–644. doi:10.1056/NEJM199803053381002
- 602 8. Goldfeld AE, Delgado JC, Thim S, Bozon MV, Ugliarolo AM, Turbay D, et al.  
603 Association of an HLA-DQ Allele With Clinical Tuberculosis. *J Am Med Assoc.*  
604 1998;279:226–228. doi:10.1001/jama.279.3.226
- 605 9. Bellamy R, Ruwende C, Corrah T, McAdam KPWJ, Thursz M, Whittle HC, et al.  
606 Tuberculosis and Chronic Hepatitis B Virus Infection in Africans and Variation in  
607 the Vitamin D Receptor Gene. *J Infect Dis.* 1999;179:721–724. doi:10.1086/314614
- 608 10. Wilkinson RJ, Llewelyn M, Toossi Z, Patel P, Pasvol G, Lalvani A, et al. Influence  
609 of vitamin D deficiency and vitamin D receptor polymorphisms on tuberculosis  
610 among Gujarati Asians in west London: A case-control study. *Lancet.*  
611 2000;355:618–621. doi:10.1016/S0140-6736(99)02301-6
- 612 11. da Cruz HLA, da Silva RC, Segat L, de Carvalho MSZ de MG, Brandão LAC,  
613 Guimarães RL, et al. *MBL2* gene polymorphisms and susceptibility to tuberculosis  
614 in a northeastern Brazilian population. *Infect Genet Evol.* 2013;19:323–329.  
615 doi:10.1016/j.meegid.2013.03.002

12. Mhmoud N, Fahal A, van de Sande WJW. Association of IL-10 and CCL5 single nucleotide polymorphisms with tuberculosis in the Sudanese population. *Trop Med Int Health*. 2013;18:1119–1127. doi:10.1111/tmi.12141
13. Chen M, Liang Y, Li W, Wang M, Hu L, Abuaku BK, et al. Impact of *MBL* and *MASP-2* gene polymorphism and its interaction on susceptibility to tuberculosis. *BMC Infect Dis*. 2015;15:1–6. doi:10.1186/s12879-015-0879-y
14. Nägler DK, Ménard R. Human cathepsin X: A novel cysteine protease of the papain family with a very short proregion and unique insertions. *FEBS Lett*. 1998;434:135–139. doi:10.1016/S0014-5793(98)00964-8
15. Santamaría I, Velasco G, Pendás AM, Fueyo A, López-Otín C. Cathepsin Z, a Novel Human Cysteine Proteinase with a Short Propeptide Domain and a Unique Chromosomal Location. *J Biol Chem*. 1998;273:16816–16823. doi:10.1074/jbc.273.27.16816
16. Nägler DK, Zhang R, Tam W, Sulea T, Purisima EO, Ménard R. Human Cathepsin X: A Cysteine Protease with Unique Carboxypeptidase Activity. *Biochemistry*. 1999;38:12648–12654. doi:10.1021/bi991371z
17. Sivaraman J, Nägler DK, Zhang R, Ménard R, Cygler M. Crystal Structure of Human Procathepsin X: A Cysteine Protease with the Proregion Covalently Linked to the Active Site Cysteine. *J Mol Biol*. 2000;295:939–951. doi:10.1006/jmbi.1999.3410
18. Gunčar G, Klemenčič I, Turk B, Turk V, Karaoglanovic-Carmona A, Juliano L, et al. Crystal structure of cathepsin X: a flip–flop of the ring of His23 allows carboxy-monopeptidase and carboxy-dipeptidase activity of the protease. *Structure*. 2000;8:305–313. doi:10.1016/S0969-2126(00)00108-8
19. Deussing J, von Olshausen I, Peters C. Murine and human cathepsin Z: cDNA-cloning, characterization of the genes and chromosomal localization. *Biochim Biophys Acta*. 2000;1491:93–106. doi:10.1016/S0167-4781(00)00021-X
20. Therrien C, Lachance P, Sulea T, Purisima EO, Qi H, Ziomek E, et al. Cathepsins X and B Can Be Differentiated through Their Respective Mono- and Dipeptidyl Carboxypeptidase Activities. *Biochemistry*. 2001;40:2702–2711. doi:10.1021/bi002460a
21. Devanathan G, Turnbull JL, Ziomek E, Purisima EO, Ménard R, Sulea T. Carboxy-monopeptidase substrate specificity of human cathepsin X. *Biochem Biophys Res Commun*. 2005;329:445–452. doi:10.1016/j.bbrc.2005.01.150
22. Dolenc I, Štefe I, Turk D, Taler-Verčič A, Turk B, Turk V, et al. Human cathepsin X/Z is a biologically active homodimer. *Biochim Biophys Acta*. 2021;1869:140567. doi:10.1016/j.bbapap.2020.140567

23. Cooke GS, Campbell SJ, Bennett S, Lienhardt C, McAdam KPWJ, Sirugo G, et al. Mapping of a Novel Susceptibility Locus Suggests a Role for *MC3R* and *CTSZ* in Human Tuberculosis. *Am J Respir Crit Care Med*. 2008;178:203–207. doi:10.1164/rccm.200710-1554OC
24. Adams LA, Möller M, Nebel A, Schreiber S, van der Merwe L, van Helden PD, et al. Polymorphisms in *MC3R* promoter and *CTSZ* 3'UTR are associated with tuberculosis susceptibility. *Eur J Hum Genet*. 2011;19:676–681. doi:10.1038/ejhg.2011.1
25. Stein CM, Zalwango S, Malone LL, Won S, Mayanja-Kizza H, Mugerwa RD, et al. Genome Scan of *M. tuberculosis* Infection and Disease in Ugandans. *PLoS ONE*. 2008;3:e4094. doi:10.1371/journal.pone.0004094
26. Baker AR, Zalwango S, Malone LL, Igo RP, Qiu F, Nsereko M, et al. Genetic susceptibility to tuberculosis associated with cathepsin Z haplotype in a Ugandan household contact study. *Hum Immunol*. 2011;72:426–430. doi:10.1016/j.humimm.2011.02.016
27. Journet A, Chapel A, Kieffer S, Louwagie M, Luche S, Garin J. Towards a human repertoire of monocytic lysosomal proteins. *Electrophoresis*. 2000;21:3411–3419. doi:10.1002/1522-2683(20001001)21:16<3411::AID-ELPS3411>3.0.CO;2-M
28. Garin J, Diez R, Kieffer S, Dermine J-F, Duclos S, Gagnon E, et al. The Phagosome Proteome: Insight into Phagosome Functions. *J Cell Biol*. 2001;152:165–180. doi:10.1083/jcb.152.1.165
29. Kos J, Jevnikar Z, Obermajer N. The role of cathepsin X in cell signaling. *Cell Adhes Migr*. 2009;3:164–166. doi:10.4161/cam.3.2.7403
30. Schmiedel BJ, Singh D, Madrigal A, Valdovino-Gonzalez AG, White BM, Zapardiel-Gonzalo J, et al. Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. *Cell*. 2018;175:1701-1715.e16. doi:10.1016/j.cell.2018.10.022
31. Obermajer N, Švajger U, Bogyo M, Jeras M, Kos J. Maturation of dendritic cells depends on proteolytic cleavage by cathepsin X. *J Leukoc Biol*. 2008;84:1306–1315. doi:10.1189/jlb.0508285
32. Jevnikar Z, Obermajer N, Bogyo M, Kos J. The role of cathepsin X in the migration and invasiveness of T lymphocytes. *J Cell Sci*. 2008;121:2652–2661. doi:10.1242/jcs.023721
33. Obermajer N, Repnik U, Jevnikar Z, Turk B, Kreft M, Kos J. Cysteine protease cathepsin X modulates immune response via activation of  $\beta_2$  integrins. *Immunology*. 2008;124:76–88. doi:10.1111/j.1365-2567.2007.02740.x

34. Pires D, Marques J, Pombo JP, Carmo N, Bettencourt P, Neyrolles O, et al. Role of Cathepsins in *Mycobacterium tuberculosis* Survival in Human Macrophages. *Sci Rep*. 2016;6:32247. doi:10.1038/srep32247
35. Lewis MS, Danelishvili L, Rose SJ, Bermudez LE. MAV\_4644 Interaction with the Host Cathepsin Z Protects *Mycobacterium avium* subsp. *hominissuis* from Rapid Macrophage Killing. *Microorganisms*. 2019;7:144. doi:10.3390/microorganisms7050144
36. Saul MC, Philip VM, Reinholdt LG, Center for Systems Neurogenetics of Addiction, Chesler EJ. High-Diversity Mouse Populations for Complex Traits. *Trends Genet*. 2019;35:501–514. doi:10.1016/j.tig.2019.04.003
37. Villar-Hernández R, Ghodousi A, Konstantynovska O, Duarte R, Lange C, Raviglione M. Tuberculosis: current challenges and beyond. *Breathe*. 2023;19:220166. doi:10.1183/20734735.0166-2022
38. Medina E, North RJ. Evidence Inconsistent with a Role for the *Bcg* Gene (*Nramp1*) in Resistance of Mice to Infection with Virulent *Mycobacterium tuberculosis*. *J Exp Med*. 1996;183:1045–1051. doi:10.1084/jem.183.3.1045
39. Medina E, North RJ. Resistance ranking of some common inbred mouse strains to *Mycobacterium tuberculosis* and relationship to major histocompatibility complex haplotype and *Nramp1* genotype. *Immunology*. 1998;93:270–274. doi:10.1046/j.1365-2567.1998.00419.x
40. Green EL. *Genetics and Probability in Animal Breeding Experiments*. 1st ed. London: Red Globe Press; 1981. Available: <https://doi.org/10.1007/978-1-349-04904-2>
41. Taylor BA, Heiniger HJ, Meier H. Genetic Analysis of Resistance to Cadmium-Induced Testicular Damage in Mice. *Proc Soc Exp Biol Med*. 1973;143:629–633. doi:10.3181/00379727-143-37380
42. Taylor BA, Wnek C, Kotlus BS, Roemer N, MacTaggart T, Phillips SJ. Genotyping new BXD recombinant inbred mouse strains and comparison of BXD and consensus maps. *Mamm Genome*. 1999;10:335–348. doi:10.1007/s003359900998
43. Peirce JL, Lu L, Gu J, Silver LM, Williams RW. A new set of BXD recombinant inbred lines from advanced intercross populations in mice. *BMC Genet*. 2004;5:1–17. doi:10.1186/1471-2156-5-7
44. Ashbrook DG, Arends D, Prins P, Mulligan MK, Roy S, Williams EG, et al. A platform for experimental precision medicine: The extended BXD mouse family. *Cell Syst*. 2021;12:235-247.e9. doi:10.1016/j.cels.2020.12.002

- 723 45. The Complex Trait Consortium. The Collaborative Cross, a community resource for  
724 the genetic analysis of complex traits. *Nat Genet.* 2004;36:1133–1137.  
725 doi:10.1038/ng1104-1133
- 726 46. Collaborative Cross Consortium. The Genome Architecture of the Collaborative  
727 Cross Mouse Genetic Reference Population. *Genetics.* 2012;190:389–401.  
728 doi:10.1534/genetics.111.132639
- 729 47. Srivastava A, Morgan AP, Najarian ML, Sarsani VK, Sigmon JS, Shorter JR, et al.  
730 Genomes of the Mouse Collaborative Cross. *Genetics.* 2017;206:537–556.  
731 doi:10.1534/genetics.116.198838
- 732 48. Meade RK, Smith CM. Immunological roads diverged: mapping tuberculosis  
733 outcomes in mice. *Trends Microbiol.* 2025;33:15–33. doi:10.1016/j.tim.2024.06.007
- 734 49. Meade RK, Long JE, Jinich A, Rhee KY, Ashbrook DG, Williams RW, et al.  
735 Genome-wide screen identifies host loci that modulate *Mycobacterium tuberculosis*  
736 fitness in immunodivergent mice. *G3 Genes|Genomes|Genetics.* 2023;13:jkad147.  
737 doi:10.1093/g3journal/jkad147
- 738 50. Smith CM, Baker RE, Proulx MK, Mishra BB, Long JE, Park SW, et al. Host-  
739 pathogen genetic interactions underlie tuberculosis susceptibility in genetically  
740 diverse mice. *eLife.* 2022;11. doi:10.7554/eLife.74419
- 741 51. Koyuncu D, Niazi MKK, Tavorara T, Abeijon C, Ginese ML, Liao Y, et al. CXCL1: A  
742 new diagnostic biomarker for human tuberculosis discovered using Diversity  
743 Outbred mice. *PLoS Pathog.* 2021;17:e1009773. doi:10.1371/journal.ppat.1009773
- 744 52. Ahmed M, Thirunavukkarasu S, Rosa BA, Thomas KA, Das S, Rangel-Moreno J, et  
745 al. Immune correlates of tuberculosis disease and risk translate across species. *Sci*  
746 *Transl Med.* 2020;12:eaay0233. doi:10.1126/scitranslmed.aay0233
- 747 53. Moreira-Teixeira L, Tabone O, Graham CM, Singhanian A, Stavropoulos E, Redford  
748 PS, et al. Mouse transcriptome reveals potential signatures of protection and  
749 pathogenesis in human tuberculosis. *Nat Immunol.* 2020;21:464–476.  
750 doi:10.1038/s41590-020-0610-z
- 751 54. Zak DE, Penn-Nicholson A, Scriba TJ, Thompson E, Suliman S, Amon LM, et al. A  
752 blood RNA signature for tuberculosis disease risk: a prospective cohort study.  
753 *Lancet.* 2016;387:2312–2322. doi:10.1016/S0140-6736(15)01316-1
- 754 55. Adams DJ, Doran AG, Lilue J, Keane TM. The Mouse Genomes Project: a  
755 repository of inbred laboratory mouse strain genomes. *Mamm Genome.*  
756 2015;26:403–412. doi:10.1007/s00335-015-9579-6
- 757 56. Desvignes L, Wolf AJ, Ernst JD. Dynamic Roles of Type I and Type II IFNs in Early  
758 Infection with *Mycobacterium tuberculosis*. *J Immunol.* 2012;188:6205–6215.  
759 doi:10.4049/jimmunol.1200255



57. Wolf AJ, Desvignes L, Linas B, Banaiee N, Tamura T, Takatsu K, et al. Initiation of the adaptive immune response to *Mycobacterium tuberculosis* depends on antigen production in the local lymph node, not the lungs. *J Exp Med*. 2007;205:105–115. doi:10.1084/jem.20071367
58. 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;188:1137–1146. doi:10.1164/rccm.201304-0803OC
59. Cronan MR, Hughes EJ, Brewer WJ, Viswanathan G, Hunt EG, Singh B, et al. A non-canonical type 2 immune response coordinates tuberculous granuloma formation and epithelialization. *Cell*. 2021;184:1757-1774.e14. doi:10.1016/j.cell.2021.02.046
60. Gideon HP, Hughes TK, Tzouanas CN, Wadsworth MH, Tu AA, Gierahn TM, et al. Multimodal profiling of lung granulomas in macaques reveals cellular correlates of tuberculosis control. *Immunity*. 2022;55:827-846.e10. doi:10.1016/j.immuni.2022.04.004
61. McHenry ML, Simmons J, Hong H, Malone LL, Mayanja-Kizza H, Bush WS, et al. Tuberculosis severity associates with variants and eQTLs related to vascular biology and infection-induced inflammation. *PLoS Genet*. 2023;19:e1010387. doi:10.1371/journal.pgen.1010387
62. Rudolf F. The Bandim TBscore – reliability, further development, and evaluation of potential uses. *Glob Health Action*. 2014;7:24303. doi:10.3402/gha.v7.24303
63. Simmons JD, Dill-McFarland KA, Stein CM, Van PT, Chihota V, Ntshiq T, et al. Monocyte Transcriptional Responses to *Mycobacterium tuberculosis* Associate with Resistance to Tuberculin Skin Test and Interferon Gamma Release Assay Conversion. *mSphere*. 2022;7:e00159-22. doi:10.1128/msphere.00159-22
64. Cronan MR. In the Thick of It: Formation of the Tuberculous Granuloma and Its Effects on Host and Therapeutic Responses. *Front Immunol*. 2022;13:820134. doi:10.3389/fimmu.2022.820134
65. Obermajer N, Premzl A, Zavašnik Bergant T, Turk B, Kos J. Carboxypeptidase cathepsin X mediates  $\beta_2$ -integrin-dependent adhesion of differentiated U-937 cells. *Exp Cell Res*. 2006;312:2515–2527. doi:10.1016/j.yexcr.2006.04.019
66. Keller C, Hoffmann R, Lang R, Brandau S, Hermann C, Ehlers S. Genetically Determined Susceptibility to Tuberculosis in Mice Causally Involves Accelerated and Enhanced Recruitment of Granulocytes. *Infect Immun*. 2006;74:4295–4309. doi:10.1128/iai.00057-06
67. Augenstreich J, Haanappel E, Sayes F, Simeone R, Guillet V, Mazeres S, et al. Phthiocerol Dimycocerosates From *Mycobacterium tuberculosis* Increase the



- 798 Membrane Activity of Bacterial Effectors and Host Receptors. *Front Cell Infect*  
799 *Microbiol.* 2020;10:420. doi:10.3389/fcimb.2020.00420
- 800 68. Krueger S, Kalinski T, Hundertmark T, Wex T, Küster D, Peitz U, et al. Up-  
801 regulation of cathepsin X in *Helicobacter pylori* gastritis and gastric cancer. *J*  
802 *Pathol.* 2005;207:32–42. doi:10.1002/path.1820
- 803 69. Selkig J, Li N, Hausmann A, Mangan MSJ, Zietek M, Mateus A, et al.  
804 Spatiotemporal proteomics uncovers cathepsin-dependent macrophage cell death  
805 during *Salmonella* infection. *Nat Microbiol.* 2020;5:1119–1133.  
806 doi:10.1038/s41564-020-0736-7
- 807 70. Wendt W, Zhu X-R, Lübbert H, Stichel CC. Differential expression of cathepsin X in  
808 aging and pathological central nervous system of mice. *Exp Neurol.* 2007;204:525–  
809 540. doi:10.1016/j.expneurol.2007.01.007
- 810 71. Kraus S, Bunsen T, Schuster S, Cichoń MA, Tacke M, Reinheckel T, et al. Cellular  
811 senescence induced by cathepsin X downregulation. *Eur J Cell Biol.* 2011;90:678–  
812 686. doi:10.1016/j.ejcb.2011.03.008
- 813 72. Allan ERO, Campden RI, Ewanchuk BW, Tailor P, Balce DR, McKenna NT, et al. A  
814 role for cathepsin Z in neuroinflammation provides mechanistic support for an  
815 epigenetic risk factor in multiple sclerosis. *J Neuroinflammation.* 2017;14:1–11.  
816 doi:10.1186/s12974-017-0874-x
- 817 73. Nishida N, Aiba Y, Hitomi Y, Kawashima M, Kojima K, Kawai Y, et al. *NELFCD* and  
818 *CTSZ* loci are associated with jaundice-stage progression in primary biliary  
819 cholangitis in the Japanese population. *Sci Rep.* 2018;8:8071. doi:10.1038/s41598-  
820 018-26369-6
- 821 74. Aiba Y, Harada K, Ito M, Suematsu T, Aishima S, Hitomi Y, et al. Increased  
822 expression and altered localization of cathepsin Z are associated with progression  
823 to jaundice stage in primary biliary cholangitis. *Sci Rep.* 2018;8:11808.  
824 doi:10.1038/s41598-018-30146-w
- 825 75. Dera AA, Ranganath L, Barraclough R, Vinjamuri S, Hamill S, Barraclough DL.  
826 Cathepsin Z as a novel potential biomarker for osteoporosis. *Sci Rep.* 2019;9:9752.  
827 doi:10.1038/s41598-019-46068-0
- 828 76. Hafner A, Glavan G, Obermajer N, Živin M, Schliebs R, Kos J. Neuroprotective role  
829 of  $\gamma$ -enolase in microglia in a mouse model of Alzheimer's disease is regulated by  
830 cathepsin X. *Aging Cell.* 2013;12:604–614. doi:10.1111/accel.12093
- 831 77. Olson OC, Joyce JA. Cysteine cathepsin proteases: regulators of cancer  
832 progression and therapeutic response. *Nat Rev Cancer.* 2015;15:712–729.  
833 doi:10.1038/nrc4027

78. Li J, Zhou X, Li L, Ji L, Li J, Qu Y, et al. The association between *CTSZ* methylation in peripheral blood and breast cancer in Chinese women. *Front Oncol.* 2023;13:1148635. doi:10.3389/fonc.2023.1148635
79. Vizin T, Christensen I, Nielsen H, Kos J. Cathepsin X in serum from patients with colorectal cancer: relation to prognosis. *Radiol Oncol.* 2012;46:207–212. doi:10.2478/v10019-012-0040-0
80. Nägler DK, Krüger S, Kellner A, Ziomek E, Menard R, Buhtz P, et al. Up-regulation of cathepsin X in prostate cancer and prostatic intraepithelial neoplasia. *Prostate.* 2004;60:109–119. doi:10.1002/pros.20046
81. Wang J, Chen L, Li Y, Guan X-Y. Overexpression of Cathepsin Z Contributes to Tumor Metastasis by Inducing Epithelial-Mesenchymal Transition in Hepatocellular Carcinoma. *PLoS ONE.* 2011;6:e24967. doi:10.1371/journal.pone.0024967
82. Sevenich L, Schurigt U, Sachse K, Gajda M, Werner F, Müller S, et al. Synergistic antitumor effects of combined cathepsin B and cathepsin Z deficiencies on breast cancer progression and metastasis in mice. *Proc Natl Acad Sci U S A.* 2010;107:2497–2502. doi:10.1073/pnas.0907240107
83. Wang J, Sha J, Strong E, Chopra AK, Lee S. FDA-Approved Amoxapine Effectively Promotes Macrophage Control of Mycobacteria by Inducing Autophagy. *Microbiol Spectr.* 2022;10:e02509-22. doi:10.1128/spectrum.02509-22
84. Stein CM, Zalwango S, Malone LL, Thiel B, Mupere E, Nsereko M, et al. Resistance and Susceptibility to *Mycobacterium tuberculosis* Infection and Disease in Tuberculosis Households in Kampala, Uganda. *Am J Epidemiol.* 2018;187:1477–1489. doi:10.1093/aje/kwx380
85. Wejse C, Gustafson P, Nielsen J, Gomes VF, Aaby P, Andersen PL, et al. TBscore: Signs and symptoms from tuberculosis patients in a low-resource setting have predictive value and may be used to assess clinical course. *Scand J Infect Dis.* 2008;40:111–120. doi:10.1080/00365540701558698
86. Shin J-H, Blay S, McNeney B, Graham J. LDheatmap: An R Function for Graphical Display of Pairwise Linkage Disequilibria Between Single Nucleotide Polymorphisms. *J Stat Softw.* 2006;16:1–9. doi:10.18637/jss.v016.c03
87. Dill-McFarland KA, Mitchell K, Batchu S, Segnitz RM, Benson B, Janczyk T, et al. Kimma: flexible linear mixed effects modeling with kinship covariance for RNA-seq data. *Bioinformatics.* 2023;39:btad279. doi:10.1093/bioinformatics/btad279
88. Law CW, Chen Y, Shi W, Smyth GK. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* 2014;15:R29. doi:10.1186/gb-2014-15-2-r29