Multimodal profiling of lung granulomas reveals cellular correlates of tuberculosis control

Authors: Hannah P. Gideon[1, 2]*, Travis K. Hughes[3,4,5]*, Marc H. Wadsworth II[3,4,5,6], Ang Andy Tu[7], Todd M. Gierahn[7], Joshua M. Peters[4,7], Forrest F. Hopkins [4,8], Jun-Rong Wei [4,8], Conner Kummerlowe[9], Nicole L. Grant[1], Kievershen Nargan[10], JiaYao Phuah[1], H. Jacob Borish[1], Pauline Maiello[1], Alexander G. White[1], Caylin G. Winchell[1,2,11], Sarah K. Nyquist[3,4,5,9,12], Sharie Keanne C. Ganchua[1], Amy Myers[1], Kush V Patel[1], Cassaundra L. Ameel[1], Catherine T. Cochran[1], Samira Ibrahim[3,4,5], , Jaime A Tomko[1], Lonnie James Frye[1], Jacob M. Rosenberg[4,8,13], Angela Shih[13], Michael Chao[4,8], Charles A. Scanga[1,2], Jose Ordovas-Montanes[4,5], Bonnie Berger[12], Joshua T. Mattila[2,14], Rajhmun Madansein[15], J. Christopher Love[4,16,17], Philana Ling Lin[2,18], Alasdair Leslie[10,19,20], Samuel M. Behar[21], Bryan Bryson[4,7], JoAnne L Flynn[1,2#], Sarah M. Fortune[4,5,8#], Alex K. Shalek[3,4,5,6,17#]

*These first authors contributed equally to this work.

#These last authors contributed equally to this work.

Affiliations:

[1]Department of Microbiology and Molecular Genetics, University of Pittsburgh School of Medicine, Pittsburgh PA USA

[2]Center for Vaccine Research, University of Pittsburgh, Pittsburgh PA USA

[3]Institute for Medical Engineering & Science, Massachusetts Institute of Technology, Cambridge, MA

[4]Ragon Institute of MGH, MIT, and Harvard, Cambridge, MA

[5]Broad Institute of MIT and Harvard, Cambridge, MA

[6]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA

[7]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA

[8]Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA

[9]Program in Computational and Systems Biology, Massachusetts Institute of Technology, Cambridge, MA, USA.

[10]Africa Health Research Institute, Durban, South Africa

[11]Division of Pulmonary, Allergy and Critical Care Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

[12]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, USA.

[13]Division of Infectious Diseases, Massachusetts General Hospital, Boston, MA

[14]Department of Infectious Diseases and Microbiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA

[15]Department of Cardiothoracic Surgery, University of KwaZulu Natal, Durban, South Africa

[16]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States

[17]The Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States
[18]Department of Pediatrics, University of Pittsburgh School of Medicine, UPMC Children's Hospital of Pittsburgh, Pittsburgh PA USA
[19]School of Laboratory Medicine and Medical Sciences, University of KwaZulu-Natal, Durban, South Africa.
[20]Department of Infection and Immunity, University College London, London, United Kingdom
[21]Department of Microbiology and Physiological Systems, University of Massachusetts Medical School, Worcester, Massachusetts, USA

Corresponding authors: JoAnne L. Flynn: joanne@pitt.edu; Sarah M. Fortune: sfortune@hsph.harvard.edu; Alex K. Shalek: shalek@mit.edu

Abstract: *Mycobacterium tuberculosis* lung infection results in a complex multicellular structure, the granuloma. In some granulomas, immune activity promotes bacterial clearance; in others, bacteria persist and grow. We identified correlates of bacterial control in cynomolgus macaque granulomas by co-registering longitudinal PET-CT imaging, single-cell RNA-sequencing, and measures of bacterial clearance. Bacterial control associates with the dynamics of granuloma formation and cellular composition. Early granulomas have limited capacity for bacterial restriction and are characterized by Type 2 immune features—plasma cells, mast cells, and IL-4/13 signaling. Granulomas that emerge after the onset of adaptive responses exhibit superior bacterial killing and are enriched for hybrid Type1-Type17 and certain cytotoxic T cells—host immune targets that can be leveraged to develop new vaccine and therapeutic strategies for TB.

One-Sentence Summary: Bacterial control in TB lung granulomas correlates with the timing of granuloma formation and the cellular immune environment.

Main Text:

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (Mtb), remains a major global health threat (*1*). Mtb infection is characterized by the formation of granulomas predominantly in the lungs and lymph nodes (*2-5*). These spatially organized structures, composed of a mixture of immune and non-immune cells (*4-12*), are key sites of host-pathogen interactions which can either restrict or facilitate bacterial survival **(Figure S1A)**. Understanding the cellular and molecular features in granulomas that are associated with bacterial restriction, versus failure to control infection, is critical for the development of next-generation treatments and vaccines for TB. Delineating such protective responses in humans has been challenging given the limited accessibility of affected lung tissue and difficulty determining the true extent of bacterial control. The cynomolgus macaque model of Mtb infection, which recapitulates the diversity of human infection outcomes and granuloma pathologies, has been a transformative advance in the field, enabling detailed studies of the features of immunologic success and failure in Mtb granulomas (*4, 7, 13*).

A spectrum of granuloma types, organizations and cellular compositions have been described in both humans and NHP (*4, 7, 13-15*). Studies of Mtb infection in NHP have demonstrated that individual granulomas are dynamic (*2, 16, 17*), changing in response to evolving interactions between bacteria and diverse host cell types (*4, 5, 8-10, 18*). The bacterial burden in individual granulomas is highest early in infection and then decreases due to increased bacterial killing as the immune response matures, even in animals that ultimately develop active TB (**Figure S1B-C)** (*2, 19, 20*). Strikingly, however, individual granulomas within a single host follow independent trajectories with respect to inflammation, cellular composition, reactivation risk, and ability to kill Mtb (*2, 16, 17, 21-24*). We and others have profiled immune responses among individual cell types in macaque lung granulomas, including those of T cells (*21, 25-29*), macrophages (*8*), B cells (*9, 30*), and neutrophils (*6, 31*), and also examined the instructive roles of cytokines, including IFN-γ, IL-2, TNF, IL-17 and IL-10 (*21, 32, 33*). While these analyses have enabled key insights into how specific canonical cell types and effector molecules relate to bacterial burden, they have been relatively narrow and directed in focus and have not revealed how the integrated actions of diverse cell types within individual granulomas influence control.

The recent emergence of high-throughput single-cell genomic profiling methods affords transformative opportunities to define the cell types, phenotypic states and intercellular circuits that comprise granulomas and inform their dynamics (*34*). Here, we developed and applied a multifactorial profiling pipeline—integrating longitudinal PET-CT imaging, single-cell RNA-sequencing (scRNA-seq)-based immunophenotyping, and molecular measures of bacterial killing—to identify features of TB lung granulomas that correlate with bacterial clearance in cynomolgus macaques (**Figure 1A**). Collectively, our data provide a holistic view of the TB lung granuloma cellular microenvironments in which Mtb is controlled or persists, and these are influenced host immune state in which granulomas arise, suggesting several novel therapeutic and prophylactic targets for future investigation. In general, this should include a brief (1-2 paragraph) introduction, followed by a statement of the specific scope of the study, followed

100 by results and then interpretations. Please avoid statements of future work, claims of priority, and repetition of conclusions at the end.

Results

Timing of granuloma establishment is associated with bacterial burden

We sought to define the complex cellular ecosystems of granulomas that manifest different degrees of bacterial control
105 in NHP. Four cynomolgus macaques were infected with a low dose of Mtb (<10 CFU; Erdman strain) and followed for 10 weeks (**Figure 1A**). Progression of Mtb infection and individual granuloma dynamics were monitored at 4, 8, and 10 weeks post infection (p.i.) using PET-CT imaging of FDG avidity as a proxy for inflammation **(Figure S1D-E, Table S1)** (*16, 35*). At necropsy, individual PET-CT identified lung granulomas were excised and dissociated to obtain a single-cell suspension; viable bacterial burden (CFU, colony forming units – i.e., culturable live bacterial
110 burden) and cumulative (live + dead) bacterial load (chromosomal equivalents, CEQ) were measured to define the extent of bacterial growth and killing in each granuloma (*2, 36*) (**Methods**).

Twenty-six granulomas from these four animals were randomly selected at the time of necropsy for scRNA-seq analysis. By serial PET-CT imaging, 15/26 granulomas were detected at 4 weeks p.i., while 11/26 granulomas were
115 observed only at 10-week p.i. or necropsy, hereafter referred to as early and late granulomas, respectively (**Figure S1E**, **Table S1**). Among the 26, there was a range of granuloma-level bacterial burdens, from sterile (0 CFU/granuloma) to high (4.6 $\log_{10}$ CFU/granuloma) (**Figure 1B-C**). There was a significant difference in CFU between early and late granulomas (3.6 vs 2 $\log_{10}$ CFU/granuloma, $p<0.0001$, Mann Whitney U test; **Figure 1C**), suggesting a relationship between bacterial burden and time of granuloma detection by PET-CT. To validate this
120 finding, we further evaluated bacterial burden between early and late granulomas in a total of 10 animals at 10 weeks p.i. (**Figure S1F-G**) and again found that the median CFU/granuloma per animal was significantly lower in late granulomas as compared to early ones.

To determine whether low CFU in late granulomas reflected reduced bacterial growth or increased bacterial killing,
125 we evaluated cumulative bacterial burden. We observed no significant difference in CEQ values between late (4.2 $\log_{10}$ CEQ/granuloma) and early granulomas (4.7 $\log_{10}$ CEQ/granuloma, $p=0.07$, Mann Whitney U test), indicating that the granulomas supported roughly similar cumulative Mtb growth over the course of infection (**Figure 1D**). However, the extent of bacterial killing, calculated as the ratio of CFU to CEQ, was significantly higher in late (-2.1 $\log_{10}$ CFU/CEQ per granuloma) as compared to early granulomas (-1.2 $\log_{10}$ CFU/CEQ per granuloma, $p=0.01$, Mann
130 Whitney U test) (**Figure 1E**), suggesting that the late granulomas have greater capacity to kill Mtb.

Broad cellular composition of TB lung Granulomas

To identify cellular and molecular factors associated with increased Mtb killing in an unbiased fashion, we loaded a single-cell suspension from each of the 26 granulomas onto a Seq-Well array (*37*) under Biosafety Level 3 conditions,
135 and then processed and sequenced as previously described (*37*). After aligning the data to the *Macaca fascicularis* (cynomolgus macaque) genome and performing robust quality controls and granuloma-specific technical corrections, we retained 109,584 high-quality single-cell transcriptomes for downstream analysis (**Figure S2; Table S2**).

Among these, we resolved 13 general cell types (**Figures 2A,B** and **S3A-G**) through dimensionality reduction, Louvain clustering, and examination of canonical linage defining genes and reference signatures from the Tabula Muris (*38*), Mouse Cell Atlas (*39*) and SaVanT database (*40*) (**Figure S3 A-G, Table S3**). These 13 encompass groups of lymphocytes, including B cells (defined by expression of *MS4A1*, *CD79B*, & *BANK1*), T and NK cells (T/NK; *GNLY, TRAC, CD3D,* & *GZMH*) and plasma cells (*IGHG1* & *JCHAIN)*); myeloid cells, including conventional dendritic cells (cDCs; *CLEC9A, CST3,* & *CPVL*), plasmacytoid dendritic cells (pDCs; *LILRA4* and *IRF8*), and macrophages (*APOC1, LYZ,* and *APOE*); mast cells (*CPA3* & *TPSAB1*); neutrophils (*CCL2, CXCL8,* & *CSF3R*); erythroid cells (*HBA1* & *HBB*); stromal cells, including endothelial cells (*RNASE1, EPAS1,* & *FCN3*) and fibroblasts (*COL3A1, COL1A1,* & *DCN)*; Type-1 pneumocytes (*AGER*); and, Type-2 pneumocytes (*SFTPC, SFTPB,* and *SFTPA1*) (**Figure 2A & B, Figure S3G** and **Table S3 & S4**). For each of the 13 cell types, we also performed further within cell-type subclustering; in these analyses, we only detected substructure among the T/NK and macrophage clusters (detailed below).

Cell types associated with timing of granuloma formation and control

We compared cellular frequencies between early and late granulomas, given the significant association between bacterial burden and the timing of granuloma detection. Our data reveal multiple cell types that are both significantly enriched in early granulomas and/or associated with increased bacterial burden, including B cells (relative cell abundance vs CFU, p=0.4, non-parametric Spearman's rho correlation test; early vs late, p=0.04, Mann Whitney U test), plasma cells (p<0.0001; p=0.001), mast cells (p=0.0024; p=0.001), and endothelial cells (p=0.001; p=0.01) (**Figure 2C-D, Table S5**). T/NK cells were abundant in late granulomas and were associated with bacterial clearance (p=0.0055; p=0.01) (**Figure 2C-D, Table S5**). To control for inter-subject variability, each of the cellular associations with granuloma dynamics and bacterial control was examined both across all animals and lesions, and through a directed analysis of the granulomas from the NHP host for which we had captured a broad representation of early and late lesions (4017) (**Figure S3H**). We further confirmed these trends by performing deconvolution on bulk RNA-sequencing of 12 additional granulomas (6 early and 6 late lesions) from separate macaques (**Figure S4A**).

Early granulomas are characterized by Type 2 immune features

The general presence of B and plasma cells is appreciated in TB granulomas, especially in well-defined lymphoid follicles (*9*). In mice and NHP, B cells have been described not only as protective regulators of the immune response to Mtb infection but also as markers of active TB in human studies (*41-43*). The increased B cell fractional abundances detected here may reflect elevated antigen levels or an attempt to limit a pathologic immune response. In contrast to our studies in individual granulomas, a recent study on lung tissue from Mtb infected macaques reported increased abundance of B cells in those that appear to be latently infected compared to those with active pulmonary TB (*89*). Plasma cells also have been noted histologically in NHP and human granulomas, and antibody features similarly have been implicated in Mtb control but also as a biomarker of active disease (*44*). Among the plasma cells in our scRNA-seq dataset, the vast majority express either IGHA or IGHG constant chains (**Figure S4B)**. This suggests that IgA and IgG are the dominant antibody classes induced by Mtb in the granuloma microenvironment.

In contrast, the presence and function of mast cells in Mtb lung granulomas has not been previously described. Therefore, to validate this observation, we performed immunohistochemistry on NHP and human granuloma sections using Tryptase and C-kit/CD117 markers (**Figure S4D & E**). This confirmed the presence of mast cells within both NHP and human granulomas, and further revealed that they primarily localize to the outer regions of NHP granulomas, including the lymphocyte cuff (**Figure S4D**), and can be found within and around human granulomas (**Figure SE**). In our data, mast cells are distinguished by their expression of IL-4 and IL-13 (**Figure S4B**), which we also recently observed in a study of human nasal polyposis, a type 2 inflammatory disease associated with dramatic epithelial remodeling (*45*). Taken together, these data suggest early granulomas represent a Type 2 immune enriched environment with poor capacity for bacterial clearance.

T and NK functional subclusters as mediators of protection

Of the 13 broad cell types, only the T/NK cell subcluster is associated with late granulomas and superior bacterial control (p=0.01 Mann Whitney U; p=0.007 Spearman rho, respectively) (**Figure 2C-D**). To further assess functional diversity within the 41,622 cells that comprise the T and NK cell cluster and their association with timing of granuloma detection and bacterial control, we performed additional subclustering analyses. This revealed 13 T/NK cell subclusters which we annotated based upon expression of lineage defining markers, known cytotoxic, regulatory and proliferation genes (**Figure 3A,C** and **S5, Tables 1 and S6)** and TCR constant gene (*TRAC*, *TRBC* and *TRDC*) expression (**Figure 3B).** The process of annotation revealed that most subclusters did not correspond neatly to canonical T and NK cell subsets, consistent with recent studies in other systems (*46*). Where possible, we annotated each based on known T cell markers and literature-derived genes of interest; we note that these genes are parts of broader transcriptional signatures that appear to reflect dominant cellular response states superimposed on cell lineage-associated gene expression programs. Among the 13 T/NK cell subclusters, 7 were significantly associated with the timing of granuloma detection; 6 with late granulomas, and one with early lesions (**Figure 3D, Table S5**).

A prominent role for Type1-Type 17 T cells in bacterial control

One T/NK cell subcluster is the most abundant cell type across all granulomas (8.8%) (**Table S4**) and the strongest correlate with late granulomas and bacterial burden (p=0.0005 Mann Whitney U; p=0.001 Spearman rho) (**Figure 3D**; **Table S4 & S5**). This subcluster, designated Type1-Type17 (T1-T17), is enriched for expression of classical Th1-associated genes, including *IFNG* and *TNF* (*47*), as well as transcription factors associated with Th17 differentiation (*48*), including *RORA* (*49*), *RORC* (*50*), *RBPJ* (*51*) and *BHLHE40* (*52-54*). While we also detect additional features of T17 cells, including *CCR6* (*55*) and *IL23R (56)*, we do not observe expression of either *IL17A* or *IL17F* (**Figure 4A; Table S6)**. Collectively, this hybrid gene expression state is consistent with previously described expression programs for Th1* or ex-Th17 cells, which are believed to be precursors to tissue resident memory (*57*). Previous studies have revealed a prominent role for CD4 Th1 and Th17 cytokines in control of Mtb infection, including IFN-γ, TNF, and IL-17 (*58-66*), and studies in NHP granulomas suggest an association between T1 and T17 cytokine expression and bacterial burden (*21*). In addition, in murine models, BHLHE40 is required for control of Mtb infection, as a repressor of IL-10 production (*54*). While Th1* and ex-Th17 subsets have been described primarily as CD4 T

cells (*21, 58, 67, 68*), our T1-T17 sub-cluster is characterized by the expression of both *CD4* and *CD8A/B* transcripts **(Figures 3C and 4C, Figure S5D-E)**.

To resolve whether this subcluster was a mixture of T1 and T17 cells or a *bona fide* hybrid state, we further sub-clustered the 9,234 T1-T17 cells. This revealed 4 distinct subpopulations, each of which represents a unique hybrid T1-T17 state **(Figure 4B, Table S7**): T1-T17 subpopulation 1 is distinguished by expression of *CD4* and markers of activation and motility, including *IL7R*, *CD6*, *TXNIP*, *PDE4D*, *ZFP36L2*, *ITGB1*, *CCR6 and CXCR3* (**Figure 4B,C; Tables 1 and S4**), making it most akin to ex-Th17 cells; T1-T17 subpopulation 2 is characterized by increased relative expression of cytotoxic effector molecules and both *CD8A* and *CD8B*; T1-T17 subpopulation 3, which includes cells expressing *CD8A/B* or *CD4*, is characterized by cytokine gene expression (*IFNG, TNF, LTA*, and *LTB*) and markers of an inhibitory cell state (*CTLA4*, *GADD45B* and *SLA*); T1-T17 subpopulation 4 is very low in abundance and characterized by heat shock and DNA damage associated transcripts (*DNAJB1* and *HSPH1*). Late granulomas had higher proportions of T1-T17 subpopulation 1 (p=0.03; p=0.055) and T1-T17 subpopulation 2 (p=0.02; p=0.02) compared to early granulomas. Surprisingly, T1-T17 subpopulation 3 is not correlated with time of detection or bacterial burden, despite expressing elevated levels of *IFNG* and *TNF* (**Figure 4E**, **Table S5**), cytokines generally considered as critical mediators of control in Mtb infection (*64, 65*).

CD4 and CD8 subclusters associated with granuloma formation dynamics

Among the remaining 12 T/NK cell subclusters, 6 are enriched for both CD4 and CD8 expression. Of these, 4 significantly associated with late lesions (**Figure 3B,D & S5D-E**). The most abundant of these (8.3% of granuloma cells, p=0.04 Mann Whitney U; p=0.03 Spearman rho, respectively; **Figure 3D, Table S4, S5**), we annotated as stem-like T cells based on elevated expression of markers of naïve and memory T cells (*TCF7*, *CCR7, IL7R,* and *TXNIP*) and activation or memory state (*CD69* and *ITGB1*) (**Figure 3B**). These cells may represent a "stem-like" population of T cells, which have been described as an early differentiating memory phenotype, distinct from naïve T cells, that are long-lived and possess a unique ability to proliferate and self-renew (*69-71*). The second, regulatory T cells (1.2%; p=0.03; p=0.1; **Figure 3D, Table S4, S5**), is defined by elevated expression of canonical Treg markers (*FOXP3, CTLA4, TIGIT,* and *IL1RL1*) and *GATA3*, a Th2 lineage-defining transcription factor that has been observed in a subset of tissue-resident Tregs (**Figure 3B)**. The third, Metallothionein expressing T cells (0.05%; p=0.003; p=0.03; **Figure 3D, Table S4, S5**), is defined by metallothionein genes such as *MT1* and *MT2* **(Figure 3B,D)** which play a role in negative regulation of Type 1 regulatory (Tr1) CD4+ cells (*72*). The fourth is a proliferating T cell subcluster (2.4%; p=0.002; p=0.03 **Figure 3D, Table S4, S5**), characterized by high expression of transcripts associated with cellular proliferation (*MKI67*, *STMN1*, and *TOP2A*) **(Figure 3B, Table S4, S5)**, consistent with published data that T cell proliferation occurs within NHP and human granulomas (*9, 21, 25, 30, 73, 74*).

The remaining 2 CD4/CD8 subclusters, both unassociated with the timing of granuloma formation, are interferon responsive T cells and SRRM2-T cells **(Figure 3D, Table S4, S5)**. The interferon cluster (0.4%) is enriched for CD4 expression and Type-I interferon inducible molecules (*MX1*, *ISG15*, *IFIT3*, *IFI6*, *IFIT1*, *RSAD2*, and *MX2*) *(75),* and may represent activated CD4+ T cells. The SRRM2-T cells (0.6%) are characterized by enrichment of genes associated

with nuclear speckles and splicing factors such as *PNISR* and *SRRM2* **(Figure 3B, Table S4, S5)**, the latter of which has been associated with alternate splicing in Parkinson disease (*76*) and has a critical role in organization of 3D genome (*77*).

255

Bacterial control is associated with specific cytotoxic T cell states

The remaining 6 T/NK subclusters are broadly defined by expression of cytotoxic genes (designated Cytotoxic 1-6; C1-6) including granzymes (*GZMA, GZMB, GZMH, GZMK* and *GZMM*), granulysin (*GNLY)*, and/or perforin (*PRF1*) (**Figure 3B, Table 1**). We confirmed expression of multiple granzymes among CD8$\alpha\beta$ T cells in Mtb granulomas by
260 flow cytometry (**Figure S6**) from separate animals. Two of these cytotoxic subclusters are associated with the timing of granuloma formation and bacterial control.

Late granulomas contained higher proportion of cells in subcluster C4 which is associated with bacterial control (3.8% of granuloma cells; p=0.04; Mann Whitney U; p=0.02) (**Figure 3D; Table S4, S5**). C4 expresses both *CD8A* and
265 *CD8B* and *TCRA* and *TCRB*, but not *TCRD,* indicating that it is composed primarily of conventional CD8$\alpha\beta$ T cells (**Figure 3B,C, S5D)**. C4 is further enriched for genes associated with cytotoxic effector functions (*PRF1, GZMH*, *GZMB,* and *GZMM)*, motility, migration and tissue residency (*CX3CR1, TGFBR3*, and *S100A10*), and regulators of cell state, such as *AHNAK*, *KLF3*, and *ZEB2* (**Figure 3C, Table S6**). Conversely, C5, which also expresses both *CD8A* and *CD8B* and *TCRA* and *TCRB*, but not *TCRD*, was the only T/NK subcluster associated with early granulomas and
270 failed control (p=0.047; p=0.7 ns; **Figure 3D, Table S4, S5**). C5 is distinguished by elevated expression of *GZMK* (**Figure 3B)**; interestingly, granzyme K expressing CD8 cells have been recently described as a hallmark of immune dysfunction in inflammation (*78*).

The remaining 4 cytotoxic subclusters did not associate with the timing of granuloma formation or bacterial burden.
275 While C6 was not detected in sufficient frequency (<0.3%) to draw meaningful conclusions, C1-3 were abundant. All three are enriched for the expression of *CD8A* but not *CD8B* and elevated TCRD, implying that these cells possess innate cytotoxic function (**Figure 3B-C)**. C1 is further characterized by high expression of all three classes of cytotoxic effectors genes—*GNLY*, *PRF1 and GZMH*, *GZMA GZMB*—as well as *KLRD1, KLRC1*, *KLRC2*, *NKG7*, which suggests that subcluster 1 contains a greater proportion of highly cytotoxic innate CD8+ T cells (possibly NKT cells),
280 $\gamma\delta$ T cells, and NK cells (**Figure 3B, Table 1, S6**). C2 is also enriched for NK receptors and CD8 T cell activation markers in addition to a trio of transcription factors (*EGR1, EGR2* and *DUSP2*) described to distinguish peripherally tolerant CD8 T cells (*79*)**-(Figure 3B, Table 1, S6)**. C3 appears to be more selectively enriched for NK cells with elevated expression of cytotoxic and NK cell markers and low expression of *CD3D* and *CD3G.* The functional complexity of these 6 subclusters, along with the common and distinct responses they represent, suggest a significant
285 and underappreciated role for cytotoxic cells in TB granulomas.

Macrophage heterogeneity in Mtb granulomas

While macrophages are responsible for much of the bacterial killing within granulomas, we did not observe any association between overall macrophage abundance and the timing of granuloma detection or burden (**Figures 2 and**

290 **S7**). Yet, like the T/NK cell cluster, the macrophage cluster had discernable substructure based on unbiased gene expression analyses. Among the 27,670 macrophages, we identified 9 subclusters (**Table S8**), 2 of which were significantly enriched among early granulomas . Most notably, in early granulomas, we uncovered expansion of alveolar macrophages defined by *MCEMP1*, *MRC1* and *PPARG* (Mac 5; 2.3% of granuloma cells; p=0.0008) (**Figure S7E, Table S8**). These data are consistent with several recent studies demonstrating that alveolar macrophages are a
295 more permissive niche for Mtb than recruited monocyte derived macrophages (*80*). In early lesions, we also found increased frequency of a subpopulation of macrophages expressing *INSIG1* and *EREG* (Mac 4; p<0.0001), but this association may be driven by two outlier granulomas and will require further work to establish its potential biologic significance (**Figure S7E**).

300 The remaining 7 macrophage subclusters were not associated with timing of granuloma formation or bacterial burden. Among these, we identified 2 subclusters of monocytes defined by expression of *VCAN* (Mac 2; 5.6%) and *CD16* (Mac 9; 0.5%), respectively. We also identified: a population of inflammatory macrophages (Mac 7; 0.9%), expressing multiple chemokines including *CXCL9*, *CXCL10*, and *CXCL11*; a subcluster of proliferating macrophages, defined by *MKI67* and *TOP2A* (Mac 8; 0.5%); a subcluster of recruited monocytes defined by expression of *FUCA1* and *LGMN,*
305 *lysosomal proteins involved in fucose metabolism and antigen processing* (81) (Mac 3; 2.3%); a subcluster defined by expression of *ATP13A3* (Mac 6; 1.0%), a component of the polyamine transport system linked to pulmonary hypertension ; and, a subcluster (Mac 1; 6.8%) distinguished by persistent expression of ambient RNA contaminants and non-macrophage lineage-defining gene expression (**Methods**), which possibly represents an efferocytotic macrophage population.

310

Cellular ecology of pulmonary TB granulomas

Given demonstrable differences in cellular composition between early and late granulomas (and bacterial burden), we assessed whether specific cell types co-occurred in TB lung granulomas more than would be expected by chance to collectively influence control. We calculated the pairwise Pearson correlation matrix between all major cell types and
315 subclusters across 26 granulomas (**Figure 5A**; **Methods**). Using hierarchical clustering of this pairwise correlation matrix, we defined 5 groups of cell types whose collective abundances are associated across granulomas (**Figure 5A**, **Table S9; Methods**). Of these, group 2 (shown in red), which includes mast cells, plasma cells, macrophage subcluster 4 and certain stromal populations, is significantly expanded in early granulomas. Group 4 (shown in navy blue), which consists of T cell subclusters T1-T17, Stem-like, Cytotoxic C2, C4, & C6, Metallothionein and SRRM2+, is
320 significantly higher in late granulomas (**Figure 5B**, **Table S10,S11**).

Cell-Cell Interactions Correlate with Granuloma-level Bacterial Burden

To further explore how the distinct cellular ecosystems observed in early and late lesions may successfully or unsuccessfully coordinate bacterial control, we examined putative cell-cell interactions within each. We focused our
325 analysis on signaling from the groups of cells uniquely enriched within early and late lesions. For each potential interacting cell-type pair, we constructed edge weights for receptor-ligand combinations, adjusting to account for

differences in the abundance of the sender cell type, relative ligand/receptor expression, and the percent of receptor positive cells (**Methods**).

330 Collectively, our interaction analysis revealed dramatic potential differences between early and late-appearing lesions that may reinforce reduced bacterial control in the former and facilitate bacterial clearance in the latter. Specifically, in early-appearing granulomas, we infer dense communication between and from the cellular subsets of group 2; this is dominated by mast cells signaling to plasma cells, T/NK cells and multiple macrophage populations, including Mac 2 and Mac 5 (**Figure 5C, Table S12**), via canonical type-2 cytokines, including IL-4/13, which may play a role in the
335 macrophage epithelialization associated with granuloma formation (*82*), as well as active endothelial crosstalk. In late-appearing granulomas, on the other hand, we observe augmented T1-T17 signaling to other T/NK populations and various macrophage populations (**Figure 5D, Table S12**). However, the majority of cells in group 4 also have the capacity drive bacterial control through cytolytic effector activity with limited intercellular coordination.

340 **Discussion**

Within an individual with Mtb infection, distinct granulomas can achieve sterilizing immunity, immune standoff, or frank immune failure (*2, 4, 83, 84*). In NHPs, which most closely recapitulate human Mtb infection and disease (*85*), this heterogeneity provides an opportunity to define the cellular and molecular factors that correlate with bacterial control to identify potential host-directed prevention and cure strategies for TB. While a spectrum of granuloma-level
345 bacterial control has been appreciated previously, the relationship between the timing of granuloma formation and bacterial control has not been fully explored, nor have the correlates of bacterial control been mapped comprehensively. By coupling advanced serial imaging, scRNA-seq, and molecular measures of bacterial growth and killing, the present study provides new insights into the temporal evolution of granulomas and immunologic control in Mtb infection.

350

Overall, our data highlight the importance of early host-pathogen interactions in control of Mtb. Granulomas that are identified late by PET-CT imaging may either be formed later, likely through dissemination (*22*), or take more time to reach the threshold to be identified by PET-CT scans (limit of detection $\geq$1mm) because of more efficient immune control or a reduction in the Type 2 signaling necessary to drive granuloma formation (*82*). Regardless of the exact
355 mechanism, late granuloma evolution appears to occur in the context of a primed adaptive immune response, characterized by multiple T and NK cell subclusters. Moreover, our measures of cumulative bacterial burden (CEQ) indicate that late granulomas have lower bacterial burden because of greater bacterial killing (CFU/CEQ), linking these adaptive immune features to true sterilizing immunity. Consistent with previous observations, our findings reinforce a critical role for T cells in the control of Mtb infection. Nevertheless, given the substantial increase in
360 resolution, our data paint a more nuanced picture, highlighting several key T/NK subsets—including those defined by specific hybrid Type1-Type 17, cytotoxic, and stem-like memory signatures—that may play a critical role in bacterial control at the local granuloma level and be actionable.

The strongest correlate of control was a subcluster of cells with transcriptional features of both Type 1 and Type 17 T
365 cells that was expanded in late granulomas. Aspects of these data are consistent with recent observations that granulomas established in immune primed environments—e.g., existing Mtb infection (*86*) or intravenous or intrabronchial BCG vaccination—are characterized by Th1/17 expression patterns that are associated with protection (*67, 87*); however, we extend these findings, defining appreciable substructure among the T1-T17 subcluster of relevance to control. The *CD4* T1-T17 subpopulation (subpopulation 1) is most consistent with published descriptions
370 of Th1/17 cells (e.g., Th1* or ex-Th17) (*57*). These cells may represent precursors to long lived tissue memory, which has been shown to play a crucial protective role in autoimmunity, bacterial control and memory immune responses to pathogens (*57, 90-92*), including Mtb infection. A recent study using flow cytometry and immunohistochemistry in Mtb infected rhesus macaques support an association of Th1 (IFNγ+) and Th17 (IL-17+) cells in lung tissue with latent infection (*88*); in contrast, another study using scRNAseq reported activated CD4 and CD8 T cells including
375 Th1 and Th17 in the lung tissue of macaques with pulmonary TB (*89*). The *CD8* subsets within the T1/T17 subcluster (subpopulations 2 & 3), meanwhile, have not been described previously. The former of these is strongly associated with bacterial control and may represent a novel immunologic paradigm that can be exploited for vaccine development. Subpopulation 3 intriguingly, expresses elevated *TNF* and *IFNG* but does not associated with bacterial restriction; further profiling will be necessary to establish the significance of this subset and its relation to previously
380 appreciated Type 1 and Type 17 features of control (*21, 58-66*).

Our data also revealed an interesting CD4 and CD8 expressing T cell subcluster associated with late granulomas that resembles stem-like T cells (*69-71, 93-96*). We hypothesize that these cells may be a source of T cell renewal in granulomas, and may differentiate into the various functional subsets we observe within them. It is possible, however,
385 that these represent memory T cells that are not specific for Mtb antigens, but migrate to the granuloma in response to inflammation and/or chemokine gradients. Indeed, flow-cytometry based studies support that a majority of T cells in granulomas do not respond to Mtb antigens by making cytokines and do not display hallmarks of exhaustion (*21, 25, 97*). These stem-like T cells warrant additional study, as they associate with control of Mtb in granulomas and, if antigen specific, could be explored as a potential vaccine target.

390

Although both CD4 and CD8 T cells have been implicated in control of Mtb infection, the cytotoxic function of lymphocytes in Mtb infection has been relatively understudied, with emphasis placed instead on macrophage activating cytokines, such as IFN-γ and TNF. Here, we describe previously unappreciated complexity among granuloma cytotoxic cells of relevance to bacterial control. In accordance with another recent study (*46*), our 6 T/NK
395 subclusters do not align neatly with canonical markers of cellular identity that would define them as classical CD8$\alpha\beta$ or CD4 T cells, NK, NK T cells, or $\gamma\delta$ T cells, but instead appear to be variable mixtures of innate and adaptive cell types with common transcriptional programming. Of these, cytotoxic subcluster 4, which is enriched in CD8$\alpha\beta$ T cells and defined by expression of several granzymes and perforin, likely represents cytotoxic effector T cells that target infected cells for apoptosis and is associated with late granulomas. Cytotoxic C5, characterized by Granzyme
400 K expression (*98*) which has been associated with T cell dysfunction in the setting of immune-aging, are enriched in early granulomas and associated with higher bacterial burden. A recent study on lung tissue (but not granulomas) from

Mtb infected macaques also found evidence of cytotoxic molecule expression associated with controlled infection (*89*). These findings reveal the importance of cytotoxic innate and adaptive lymphocytes in temporal control of Mtb in granulomas, and a potential role for in future prevention and cure strategies.

Equally importantly, our data reveal an unexpected strong association between Type 2 immune features and granulomas that appear early in infection and fail to control. The high CFU, early granulomas are characterized by significantly higher proportions of B cells, plasma and mast cells. In other conditions such as rheumatoid arthritis, mast cells are known to induce differentiation of B cells to specific IgA secreting plasma cells (*99*). The expression of IGHA and IGHG and presence of plasma cells in granulomas support the notion that antibodies may play a prominent role in Mtb infection, perhaps with different effects as a function of antibody quality (*44, 100, 101*). The localization of mast cells—major producers of Type 2 cytokines, such as IL-4 and IL-13 which have been shown to modulate CD8 T cell function, inhibiting cytotoxic activity (*102, 103*)—in and around the lymphocyte cuff suggests potential regulatory interactions, further reinforced by our cell-cell interaction modeling. While mast cells have been described in granulomatous conditions, such as TB lymphadenitis (*104*), leprosy skin lesions (*105*), and liver granulomas (*106*), and may orchestrate immune cross talk in TB (*107*), this is the first description of direct correlation with failure of Mtb control in TB granulomas. However, it is possible that these cell types are a result of, rather than a cause of, higher bacterial burden and their regulatory features may reflect an attempt at curbing pathologic immune activation. While more detailed studies on the roles of mast cells in TB are indicated, this observation provides exciting new avenues to explore the immune architecture of failed immunity in TB lung granulomas and suggests novel intervention strategies.

There are several limitations to this study. The granuloma is an inherently heterogenous environment and includes necrotic debris, requiring robust technical correction and quality control; this results in an analysis of only high-quality cells. Since only a fraction of cells from each granuloma are analyzed, proportions may not reflect the true composition of cells within a granuloma and may be skewed toward lymphocytes, highlighting the importance of orthogonal validations. In bulk RNA-sequencing analysis of a distinct set of early and late granulomas, we observed generally similar trends in cell-type composition. Further, we focused primarily on cell types, sub-clusters, and subpopulations that were correlated with time of granuloma appearance and control. While macrophages are clearly an important component of the immune response in TB granulomas, the heterogeneity of the myeloid populations requires further in-depth evaluation with additional samples and time points to appreciate which functions and cell types are associated with bacterial restriction or permissiveness, and how. Relatedly, the granuloma landscape investigated here is from a single, albeit pivotal, time point. It is likely that expression of certain genes that arise early in infection and then are downregulated as infection progresses will be missed, as will some populations critical to guiding overall granuloma outcome. More generally, matched analyses of earlier and later time point post-infection, along with analysis of lung tissue and granulomas from vaccinated or reinfected and protected animals will provide a more complete picture of the temporal control of Mtb in granulomas and is the subject of future work.

In summary, our data represent the first scRNA-seq investigation of the cellular and molecular features that dynamically associate with natural control of Mtb in pulmonary granulomas. Beyond recapitulating canonical correlates, our analysis defines nuanced actionable innate and adaptive functional cell states, and sheds light on essential dynamics among host-pathogen interactions (*108*). Collectively, our data substantiate a model where Mtb burden within early forming lesions is dictated by the interplay among restrictive, inflammatory innate-like and permissive, protective type-2 (wound healing) responses seeking to balance bacterial control with the maintenance of essential tissue functionality, respectively. In those lesions forming later in infection, this balance can be tipped by the emergence of adaptive T1-T17 and cytotoxic responses which are capable of controlling local disease, given sufficient access. Such a framework is consistent with previous observations of natural (*86*) or induced (*67*) control, and nominates several putative axes of intra- and intercellular signaling that may prove therapeutically or prophylactically valuable, as well as intellectual links to other inflammatory and infectious diseases that affect epithelial barrier tissues (*45, 109*).

Reference

1. WHO, Global Tuberculosis Report. (2019).
2. P. L. Lin *et al.*, Sterilization of granulomas is common in active and latent tuberculosis despite within-host variability in bacterial killing. *Nat Med* **20**, 75-79 (2014).
3. D. G. Russell, C. E. Barry, 3rd, J. L. Flynn, Tuberculosis: what we don't know can, and does, hurt us. *Science* **328**, 852-856 (2010).
4. J. L. Flynn, Klein, E.,, Pulmonary Tuberculosis in Monkeys. *A Color Atlas of Comparative Pathology of Pulmonary Tuberculosis*, 83-105 (2010).
5. T. Ulrichs, S. H. Kaufmann, New insights into the function of granulomas in human tuberculosis. *J Pathol* **208**, 261-269 (2006).
6. H. P. Gideon, J. Phuah, B. A. Junecko, J. T. Mattila, Neutrophils express pro- and anti-inflammatory cytokines in granulomas from Mycobacterium tuberculosis-infected cynomolgus macaques. *Mucosal Immunol* **12**, 1370-1381 (2019).
7. P. L. Lin *et al.*, Early events in Mycobacterium tuberculosis infection in cynomolgus macaques. *Infect Immun* **74**, 3790-3803 (2006).
8. J. T. Mattila *et al.*, Microenvironments in tuberculous granulomas are delineated by distinct populations of macrophage subsets and expression of nitric oxide synthase and arginase isoforms. *J Immunol* **191**, 773-784 (2013).
9. J. Y. Phuah, J. T. Mattila, P. L. Lin, J. L. Flynn, Activated B cells in the granulomas of nonhuman primates infected with Mycobacterium tuberculosis. *Am J Pathol* **181**, 508-514 (2012).
10. S. Ehlers, U. E. Schaible, The granuloma in tuberculosis: dynamics of a host-pathogen collusion. *Front Immunol* **3**, 411 (2012).
11. S. T. Reece, S. H. Kaufmann, Floating between the poles of pathology and protection: can we pin down the granuloma in tuberculosis? *Curr Opin Microbiol* **15**, 63-70 (2012).
12. A. J. Pagan, L. Ramakrishnan, Immunity and Immunopathology in the Tuberculous Granuloma. *Cold Spring Harb Perspect Med* **5**, (2014).
13. G. Canetti, The tubercle bacillus in the pulmonary lesion of man: histobacteriology and its bearing on the therapy of pulmonary tuberculosis. *Springer*, (1955).
14. R. L. Hunter, Tuberculosis as a three-act play: A new paradigm for the pathogenesis of pulmonary tuberculosis. *Tuberculosis (Edinb)* **97**, 8-17 (2016).
15. R. L. Hunter, Pathology of post primary tuberculosis of the lung: an illustrated critical review. *Tuberculosis (Edinb)* **91**, 497-509 (2011).
16. M. T. Coleman *et al.*, Early Changes by (18)Fluorodeoxyglucose positron emission tomography coregistered with computed tomography predict outcome after Mycobacterium tuberculosis infection in cynomolgus macaques. *Infect Immun* **82**, 2400-2404 (2014).
17. P. L. Lin *et al.*, Radiologic Responses in Cynomolgus Macaques for Assessing Tuberculosis Chemotherapy Regimens. *Antimicrob Agents Chemother* **57**, 4237-4244 (2013).

18. J. L. Flynn *et al.*, Non-human primates: a model for tuberculosis research. *Tuberculosis (Edinb)* **83**, 116-118 (2003).
19. A. M. Cadena, J. L. Flynn, S. M. Fortune, The Importance of First Impressions: Early Events in Mycobacterium tuberculosis Infection Influence Outcome. *mBio* **7**, e00342-00316 (2016).
20. P. Maiello *et al.*, Rhesus Macaques Are More Susceptible to Progressive Tuberculosis than Cynomolgus Macaques: a Quantitative Comparison. *Infect Immun* **86**, (2018).
21. H. P. Gideon *et al.*, Variability in tuberculosis granuloma T cell responses exists, but a balance of pro- and anti-inflammatory cytokines is associated with sterilization. *PLoS Pathog* **11**, e1004603 (2015).
22. C. J. Martin *et al.*, Digitally Barcoding Mycobacterium tuberculosis Reveals In Vivo Infection Dynamics in the Macaque Model of Tuberculosis. *mBio* **8**, (2017).
23. A. Lenaerts, C. E. Barry, 3rd, V. Dartois, Heterogeneity in tuberculosis pathology, microenvironments and therapeutic responses. *Immunol Rev* **264**, 288-307 (2015).
24. S. T. Malherbe *et al.*, Persisting positron emission tomography lesion activity and Mycobacterium tuberculosis mRNA after tuberculosis cure. *Nat Med* **22**, 1094-1100 (2016).
25. E. A. Wong *et al.*, Low Levels of T Cell Exhaustion in Tuberculous Lung Granulomas. *Infect Immun* **86**, (2018).
26. T. W. Foreman *et al.*, CD4+ T-cell-independent mechanisms suppress reactivation of latent tuberculosis in a macaque model of HIV coinfection. *Proc Natl Acad Sci U S A* **113**, E5636-5644 (2016).
27. P. L. Lin *et al.*, CD4 T cell depletion exacerbates acute Mycobacterium tuberculosis while reactivation of latent infection is dependent on severity of tissue depletion in cynomolgus macaques. *AIDS research and human retroviruses* **28**, 1693-1702 (2012).
28. J. T. Mattila, C. R. Diedrich, P. L. Lin, J. Phuah, J. L. Flynn, Simian immunodeficiency virus-induced changes in T cell cytokine responses in cynomolgus macaques with latent Mycobacterium tuberculosis infection are associated with timing of reactivation. *J Immunol* **186**, 3527-3537 (2011).
29. C. R. Diedrich *et al.*, SIV and Mycobacterium tuberculosis synergy within the granuloma accelerates the reactivation pattern of latent tuberculosis. *bioRxiv*, (2020).
30. J. Phuah *et al.*, Effects of B Cell Depletion on Early Mycobacterium tuberculosis Infection in Cynomolgus Macaques. *Infect Immun* **84**, 1301-1311 (2016).
31. J. T. Mattila, P. Maiello, T. Sun, L. E. Via, J. L. Flynn, Granzyme B-expressing neutrophils correlate with bacterial load in granulomas from Mycobacterium tuberculosis-infected cynomolgus macaques. *Cell Microbiol* **17**, 1085-1097 (2015).
32. P. L. Lin *et al.*, Tumor necrosis factor neutralization results in disseminated disease in acute and latent Mycobacterium tuberculosis infection with normal granuloma structure in a cynomolgus macaque model. *Arthritis Rheum* **62**, 340-350 (2010).
33. E. A. Wong *et al.*, IL-10 Impairs Local Immune Response in Lung Granulomas and Lymph Nodes during Early Mycobacterium tuberculosis Infection. *J Immunol* **204**, 644-659 (2020).
34. S. M. Prakadan, A. K. Shalek, D. A. Weitz, Scaling by shrinking: empowering single-cell 'omics' with microfluidic devices. *Nat Rev Genet* **18**, 345-361 (2017).
35. A. G. White *et al.*, Analysis of 18FDG PET/CT Imaging as a Tool for Studying Mycobacterium tuberculosis Infection and Treatment in Non-human Primates. *J Vis Exp*, (2017).
36. E. J. Munoz-Elias *et al.*, Replication dynamics of Mycobacterium tuberculosis in chronically infected mice. *Infect Immun* **73**, 546-551 (2005).
37. T. M. Gierahn *et al.*, Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. *Nat Methods* **14**, 395-398 (2017).
38. C. Tabula Muris *et al.*, Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature* **562**, 367-372 (2018).
39. X. Han *et al.*, Mapping the Mouse Cell Atlas by Microwell-Seq. *Cell* **173**, 1307 (2018).
40. D. Lopez *et al.*, SaVanT: a web-based tool for the sample-level visualization of molecular signatures in gene expression profiles. *BMC Genomics* **18**, 824 (2017).
41. A. G. Loxton, Bcells and their regulatory functions during Tuberculosis: Latency and active disease. *Mol Immunol* **111**, 145-151 (2019).
42. K. P. Lyashchenko, H. M. Vordermeier, W. R. Waters, Memory B cells and tuberculosis. *Vet Immunol Immunopathol* **221**, 110016 (2020).
43. M. Rao *et al.*, B in TB: B Cells as Mediators of Clinically Relevant Immune Responses in Tuberculosis. *Clin Infect Dis* **61Suppl 3**, S225-234 (2015).
44. A. J. Jacobs, J. Mongkolsapaya, G. R. Screaton, H. McShane, R. J. Wilkinson, Antibodies and tuberculosis. *Tuberculosis (Edinb)* **101**, 102-113 (2016).

45. J. Ordovas-Montanes *et al.*, Allergic inflammatory memory in human respiratory epithelial progenitor cells. *Nature* **560**, 649-654 (2018).
46. J. A. Rath *et al.*, Single-cell transcriptomics identifies multiple pathways underlying antitumor function of TCR- and CD8alphabeta-engineered human CD4(+) T cells. *Sci Adv* **6**, eaaz7809 (2020).
47. I. Raphael, S. Nalawade, T. N. Eagar, T. G. Forsthuber, T cell subsets and their signature cytokines in autoimmune and inflammatory diseases. *Cytokine* **74**, 5-17 (2015).
48. N. Yosef *et al.*, Dynamic regulatory network controlling TH17 cell differentiation. *Nature* **496**, 461-468 (2013).
49. X. O. Yang *et al.*, Molecular antagonism and plasticity of regulatory and inflammatory T cell programs. *Immunity* **29**, 44-56 (2008).
50. Ivanov, II *et al.*, The orphan nuclear receptor RORgammat directs the differentiation program of proinflammatory IL-17+ T helper cells. *Cell* **126**, 1121-1133 (2006).
51. G. Meyer Zu Horste *et al.*, RBPJ Controls Development of Pathogenic Th17 Cells by Regulating IL-23 Receptor Expression. *Cell Rep* **16**, 392-404 (2016).
52. C. C. Lin *et al.*, IL-1-induced Bhlhe40 identifies pathogenic T helper cells in a model of autoimmune neuroinflammation. *J Exp Med* **213**, 251-271 (2016).
53. C. C. Lin *et al.*, Bhlhe40 controls cytokine production by T cells and is essential for pathogenicity in autoimmune neuroinflammation. *Nat Commun* **5**, 3551 (2014).
54. J. P. Huynh *et al.*, Bhlhe40 is an essential repressor of IL-10 during Mycobacterium tuberculosis infection. *J Exp Med* **215**, 1823-1838 (2018).
55. K. Hirota *et al.*, Preferential recruitment of CCR6-expressing Th17 cells to inflamed joints via CCL20 in rheumatoid arthritis and its animal model. *J Exp Med* **204**, 2803-2812 (2007).
56. T. Kobayashi *et al.*, IL23 differentially regulates the Th1/Th17 balance in ulcerative colitis and Crohn's disease. *Gut* **57**, 1682-1689 (2008).
57. M. C. Amezcua Vesely *et al.*, Effector TH17 Cells Give Rise to Long-Lived TRM Cells that Are Essential for an Immediate Response against Bacterial Infection. *Cell* **178**, 1176-1188 e1115 (2019).
58. I. V. Lyadova, A. V. Panteleev, Th1 and Th17 Cells in Tuberculosis: Protection, Pathology, and Biomarkers. *Mediators Inflamm* **2015**, 854507 (2015).
59. S. A. Khader, R. Gopal, IL-17 in protective immunity to intracellular pathogens. *Virulence* **1**, 423-427 (2010).
60. S. A. Khader *et al.*, IL-23 and IL-17 in the establishment of protective pulmonary CD4+ T cell responses after vaccination and during Mycobacterium tuberculosis challenge. *Nat Immunol* **8**, 369-377 (2007).
61. H. M. Algood, P. L. Lin, J. L. Flynn, Tumor necrosis factor and chemokine interactions in the formation and maintenance of granulomas in tuberculosis. *Clin Infect Dis* **41 Suppl 3**, S189-193 (2005).
62. P. L. Lin, H. L. Plessner, N. N. Voitenok, J. L. Flynn, Tumor necrosis factor and tuberculosis. *J Investig Dermatol Symp Proc* **12**, 22-25 (2007).
63. K. A. Millington *et al.*, Dynamic relationship between IFN-gamma and IL-2 profile of Mycobacterium tuberculosis-specific T cells and antigen load. *J Immunol* **178**, 5217-5226 (2007).
64. A. O'Garra *et al.*, The immune response in tuberculosis. *Annu Rev Immunol* **31**, 475-527 (2013).
65. T. J. Scriba, A. K. Coussens, H. A. Fletcher, Human Immunology of Tuberculosis. *Microbiol Spectr* **5**, (2017).
66. A. M. Green, R. Difazio, J. L. Flynn, IFN-gamma from CD4 T cells is essential for host survival and enhances CD8 T cell function during Mycobacterium tuberculosis infection. *J Immunol* **190**, 270-277 (2013).
67. P. A. Darrah *et al.*, Prevention of tuberculosis in macaques after intravenous BCG immunization. *Nature* **577**, 95-102 (2020).
68. C. A. M. Mpande *et al.*, Functional, Antigen-Specific Stem Cell Memory (TSCM) CD4(+) T Cells Are Induced by Human Mycobacterium tuberculosis Infection. *Front Immunol* **9**, 324 (2018).
69. R. Ahmed *et al.*, Human Stem Cell-like Memory T Cells Are Maintained in a State of Dynamic Flux. *Cell Rep* **17**, 2811-2818 (2016).
70. N. Caccamo, S. A. Joosten, T. H. M. Ottenhoff, F. Dieli, Atypical Human Effector/Memory CD4(+) T Cells With a Naive-Like Phenotype. *Front Immunol* **9**, 2832 (2018).
71. L. Gattinoni *et al.*, A human memory T cell subset with stem cell-like properties. *Nat Med* **17**, 1290-1297 (2011).
72. C. Wu *et al.*, Metallothioneins negatively regulate IL-27-induced type 1 regulatory T-cell differentiation. *Proc Natl Acad Sci U S A* **110**, 7802-7807 (2013).

73. H. Ohtani, Granuloma cells in chronic inflammation express CD205 (DEC205) antigen and harbor proliferating T lymphocytes: similarity to antigen-presenting cells. *Pathol Int* **63**, 85-93 (2013).
74. E. F. McCaffrey *et al.*, Multiplexed imaging of human tuberculosis granulomas uncovers immunoregulatory features conserved across tissue and blood. *bioRxiv*, 2020.2006.2008.140426 (2020).
75. P. A. Szabo *et al.*, Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat Commun* **10**, 4706 (2019).
76. L. A. Shehadeh *et al.*, SRRM2, a potential blood biomarker revealing high alternative splicing in Parkinson's disease. *PLoS One* **5**, e9104 (2010).
77. S. Hu, P. Lv, Z. Yan, B. Wen, Disruption of nuclear speckles reduces chromatin interactions in active compartments. *Epigenetics Chromatin* **12**, 43 (2019).
78. D. A. Mogilenko *et al.*, Comprehensive Profiling of an Aging Immune System Reveals Clonal GZMK(+) CD8(+) T Cells as Conserved Hallmark of Inflammaging. *Immunity* **54**, 99-115 e112 (2021).
79. A. Schietinger, J. J. Delrow, R. S. Basom, J. N. Blattman, P. D. Greenberg, Rescued tolerant CD8 T cells are preprogrammed to reestablish the tolerant state. *Science* **335**, 723-727 (2012).
80. S. B. Cohen *et al.*, Alveolar Macrophages Provide an Early Mycobacterium tuberculosis Niche and Initiate Dissemination. *Cell Host Microbe* **24**, 439-446 e434 (2018).
81. L. Xu *et al.*, Downregulation of alpha-l-fucosidase 1 suppresses glioma progression by enhancing autophagy and inhibiting macrophage infiltration. *Cancer Sci* **111**, 2284-2296 (2020).
82. M. R. Cronan *et al.*, A non-canonical type 2 immune response coordinates tuberculous granuloma formation and epithelialization. *Cell* **184**, 1757-1774 e1714 (2021).
83. J. L. Flynn, Lessons from experimental Mycobacterium tuberculosis infections. *Microbes Infect* **8**, 1179-1188 (2006).
84. P. L. Lin *et al.*, Quantitative comparison of active and latent tuberculosis in the cynomolgus macaque model. *Infect Immun* **77**, 4631-4642 (2009).
85. M. T. Coleman *et al.*, PET/CT imaging reveals a therapeutic response to oxazolidinones in macaques and humans with tuberculosis. *Sci Transl Med* **6**, 265ra167 (2014).
86. A. M. Cadena *et al.*, Concurrent infection with Mycobacterium tuberculosis confers robust protection against secondary infection in macaques. *PLoS Pathog* **14**, e1007305 (2018).
87. K. Dijkman *et al.*, Prevention of tuberculosis infection and disease by local BCG in repeatedly exposed rhesus macaques. *Nat Med* **25**, 255-262 (2019).
88. U. Shanmugasundaram *et al.*, Pulmonary Mycobacterium tuberculosis control associates with CXCR3- and CCR6-expressing antigen-specific Th1 and Th17 cell recruitment. *JCI Insight* **5**, (2020).
89. E. Esaulova *et al.*, The immune landscape in tuberculosis reveals populations linked to disease and latency. *Cell Host Microbe* **29**, 165-178 e168 (2021).
90. Y. Liang, H. F. Pan, D. Q. Ye, Tc17 Cells in Immunity and Systemic Autoimmunity. *Int Rev Immunol* **34**, 318-331 (2015).
91. J. P. van Hamburg, S. W. Tas, Molecular mechanisms underpinning T helper 17 cell heterogeneity and functions in rheumatoid arthritis. *J Autoimmun* **87**, 69-81 (2018).
92. V. S. Wacleche *et al.*, New insights into the heterogeneity of Th17 subsets contributing to HIV-1 persistence during antiretroviral therapy. *Retrovirology* **13**, 59 (2016).
93. E. K. Cartwright *et al.*, Initiation of Antiretroviral Therapy Restores CD4+ T Memory Stem Cell Homeostasis in Simian Immunodeficiency Virus-Infected Macaques. *J Virol* **90**, 6699-6708 (2016).
94. S. A. Fuertes Marraco, C. Soneson, M. Delorenzi, D. E. Speiser, Genome-wide RNA profiling of long-lasting stem cell-like memory CD8 T cells induced by Yellow Fever vaccination in humans. *Genom Data* **5**, 297-301 (2015).
95. J. Mateus *et al.*, Low frequency of circulating CD8+ T stem cell memory cells in chronic chagasic patients with severe forms of the disease. *PLoS Negl Trop Dis* **9**, e3432 (2015).
96. S. M. Todryk, T Cell Memory to Vaccination. *Vaccines (Basel)* **6**, (2018).
97. S. Sakai *et al.*, CD4 T Cell-Derived IFN-gamma Plays a Minimal Role in Control of Pulmonary Mycobacterium tuberculosis Infection and Must Be Actively Repressed by PD-1 to Prevent Lethal Disease. *PLoS Pathog* **12**, e1005667 (2016).
98. Y. Guo, J. Chen, L. Shi, Z. Fan, Valosin-containing protein cleavage by granzyme K accelerates an endoplasmic reticulum stress leading to caspase-independent cytotoxicity of target tumor cells. *J Immunol* **185**, 5348-5359 (2010).
99. S. Merluzzi *et al.*, Mast cells enhance proliferation of B lymphocytes and drive their differentiation toward IgA-secreting plasma cells. *Blood* **115**, 2810-2817 (2010).
100. L. L. Lu *et al.*, A Functional Role for Antibodies in Tuberculosis. *Cell* **167**, 433-443 e414 (2016).

101. J. M. Achkar, J. Chan, A. Casadevall, B cells and antibodies in the defense against Mycobacterium tuberculosis infection. *Immunol Rev* **264**, 167-181 (2015).
102. N. Kienzle *et al.*, Progressive differentiation and commitment of CD8+ T cells to a poorly cytolytic CD8low phenotype in the presence of IL-4. *J Immunol* **174**, 2021-2029 (2005).
103. D. K. Wijesundara, D. C. Tscharke, R. J. Jackson, C. Ranasinghe, Reduced interleukin-4 receptor alpha expression on CD8+ T cells correlates with higher quality anti-viral immunity. *PLoS One* **8**, e55788 (2013).
104. M. Taweevisit, U. Poumsuk, High mast cell density associated with granulomatous formation in tuberculous lymphadenitis. *Southeast Asian J Trop Med Public Health* **38**, 115-119 (2007).
105. I. N. Bagwan, M. M. Khandekar, P. Kadam, M. V. Jadhav, S. D. Deshmukh, A study of mast cells in granulomatous lesions of skin, with special emphasis on leprosy. *Indian J Lepr* **76**, 31-37 (2004).
106. B. Celasun, J. Crow, P. J. Scheuer, Mast cells in granulomatous liver disease. *Pathol Res Pract* **188**, 97-100 (1992).
107. K. M. Garcia-Rodriguez, A. Goenka, M. T. Alonso-Rasgado, R. Hernandez-Pando, S. Bulfone-Paus, The Role of Mast Cells in Tuberculosis: Orchestrating Innate Immune Crosstalk? *Front Immunol* **8**, 1290 (2017).
108. A. Iwasaki, R. Medzhitov, Control of adaptive immunity by the innate immune system. *Nat Immunol* **16**, 343-353 (2015).
109. T. K. Hughes *et al.*, Second-Strand Synthesis-Based Massively Parallel scRNA-Seq Reveals Cellular States and Molecular Features of Human Inflammatory Skin Pathologies. *Immunity* **53**, 878-894 e877 (2020).
110. J. Schindelin *et al.*, Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682 (2012).
111. E. Z. Macosko *et al.*, Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202-1214 (2015).
112. M. D. Young, and Behjati, S., SoupX removes ambient RNA contamination from droplet based single cell RNA sequencing data. *BioRxiv*, (2018).
113. A. T. L. Lun *et al.*, EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol* **20**, 63 (2019).
114. C. S. McGinnis, L. M. Murrow, Z. J. Gartner, DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst* **8**, 329-337 e324 (2019).
115. F. A. Wolf, P. Angerer, F. J. Theis, SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* **19**, 15 (2018).
116. A. Liberzon *et al.*, Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739-1740 (2011).
117. L. Varemo, J. Nielsen, I. Nookaew, Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. *Nucleic Acids Res* **41**, 4378-4391 (2013).
118. X. Guo *et al.*, Global characterization of T cells in non-small-cell lung cancer by single-cell sequencing. *Nat Med* **24**, 978-985 (2018).
119. R. Zilionis *et al.*, Single-Cell Transcriptomics of Human and Mouse Lung Cancers Reveals Conserved Myeloid Populations across Individuals and Species. *Immunity* **50**, 1317-1334 e1310 (2019).
120. A. M. Newman *et al.*, Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* **12**, 453-457 (2015).

Acknowledgments: We are grateful to the research and veterinary technicians: Chelsea Chedrick, Carolyn Bigbee, Nicholas Schindler, Mark Rogers, Tara Rutledge, Chelsea Causgrove and Brianne Stein in the Flynn lab who assisted with this work, as well as helpful discussions with members of the Flynn, Scanga, Mattila, Lin and Shalek laboratories. We also thank the efforts of the University of Pittsburgh Division of Laboratory Animal Research technicians for husbandry of the animals.

Funding:
Bill and Melinda Gates Foundation (OP1139972: AL, SMB, SMF, JLF, AKS; OPP1202327: AKS)

Searle Scholars Program (AKS)

The Beckman Young Investigator Program (AKS),

Sloan Fellowship in Chemistry (AKS)

NIH (5U24AI118672, BAA-NIAID-NIHAI201700104) (AKS).

American Lung Association RG571577(HPG), F30-AI143160 (TKH), NIH T32A1065380 (NLG), NSF GRFP grant (SKN 1122374) Wellcome Trust Fellowship award 210662/Z/16/Z (AL), Koch Institute Support (core) grant P30-CA14051 from the National Cancer Institute (CL), NIH CFAR P30 AI060354(BB) NIH R01A1022553 (BB), JR (T32 A1007387).

Author contributions:

Conceptualization: JLF, SMF, AKS

Data Curation: HPG, TKH, FFH, PM, AGW, NLG, AL

Formal Analysis: HPG,TKH, NLF, FFH, AGW,

Methodology: HPG,TKH, MHW, AAT, TMG, FFH, CK, PM, AGW, SKN, HJB, BB, JCL

Investigation: HPG,TKH, MHW, AAT, TG, FFH, JW, CK, JMP, PM, AGW, SKN, HJB, SKCG, AM, KVP, CLA, CTC, JAT, LJF, HJB, PLL, SI, JYP, JMR, AS, JOM

Visualization: HPG, TKH

Validation: HPG, TKH, NLG, KN, CGW, SI

Resources: JLF, SMF, AKS, JCL. RM, AL

Funding acquisition: JLF, SMF, AKS

Project administration: CAS

Supervision: JLF, SMF, AKS, SMB, BDB, AL, JCL, BB

Writing – original draft: HPG, TKH, SMB, JLF, SMF, AKS

Writing – review & editing: HPG, TKH, MC, SMB, JLF, SMF, AKS

Competing interests: A.K.S. reports compensation for consulting and/or SAB membership from Merck, Honeycomb Biotechnologies, Cellarity, Repertoire Immune Medicines, Third Rock Ventures, Ochre Bio, and Dahlia Biosciences.

CL: shareholder and consultant Honeycomb biotechnologies

TKH: shareholder and consultant nference, inc.

Data and materials availability

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Alex K. Shalek (shalek@mit.edu).

Materials Availability

The study did not generate new unique reagents.

Data and Code Availability

Raw and processed data will be available on the gene expression omnibus. Additional code is available upon request from the lead contact.

Figures

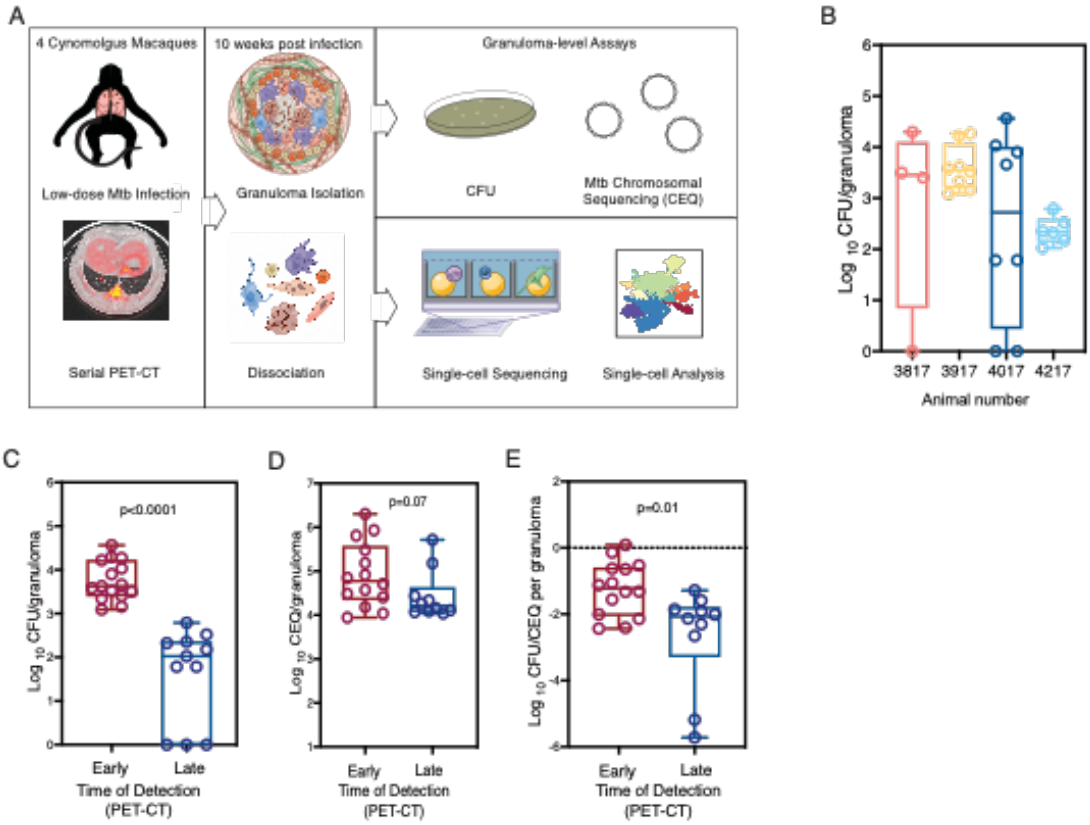


Figure 1. Study design, experimental set up, characteristics of animals over the course of Mtb infection and granuloma bacterial burden.

(A) Study design: Cynomolgus macaques (n=4) were infected with a low-dose inoculum of Mtb (Erdman strain) and serial PET-CT scans were performed at 4, 8, and 10 weeks post-infection with the final scan used as a map for lesion identification at necropsy. Individual granulomas were excised and homogenized. CFU and CEQ assays were performed on all granulomas (top right) and 26 individual granulomas across 4 animals were randomly selected at necropsy for Seq-Well assays (bottom right). **(B)** Distribution of CFU per granuloma sampled for Seq-Well assay for each animal. Each dot is an individual granuloma. **(C)** CFU $\log_{10}$ per granuloma (total live bacteria) organized by time of detection by PET-CT scan (Table S1): early granulomas (maroon), late granulomas (navy blue). Each symbol is a granuloma. Box plot showing median, IQR and range. Mann Whitney U for panels E-G. **(D)** CEQ $\log_{10}$ per granuloma (Chromosomal equivalents, CEQ, live + dead Mtb) organized by time of detection. **(E)** Ratio between CFU (viable bacteria) and CEQ (total bacterial burden) i.e., relative bacterial survival. Lower ratio (negative values) corresponds to increased killing and higher ratio corresponds to increased Mtb survival.

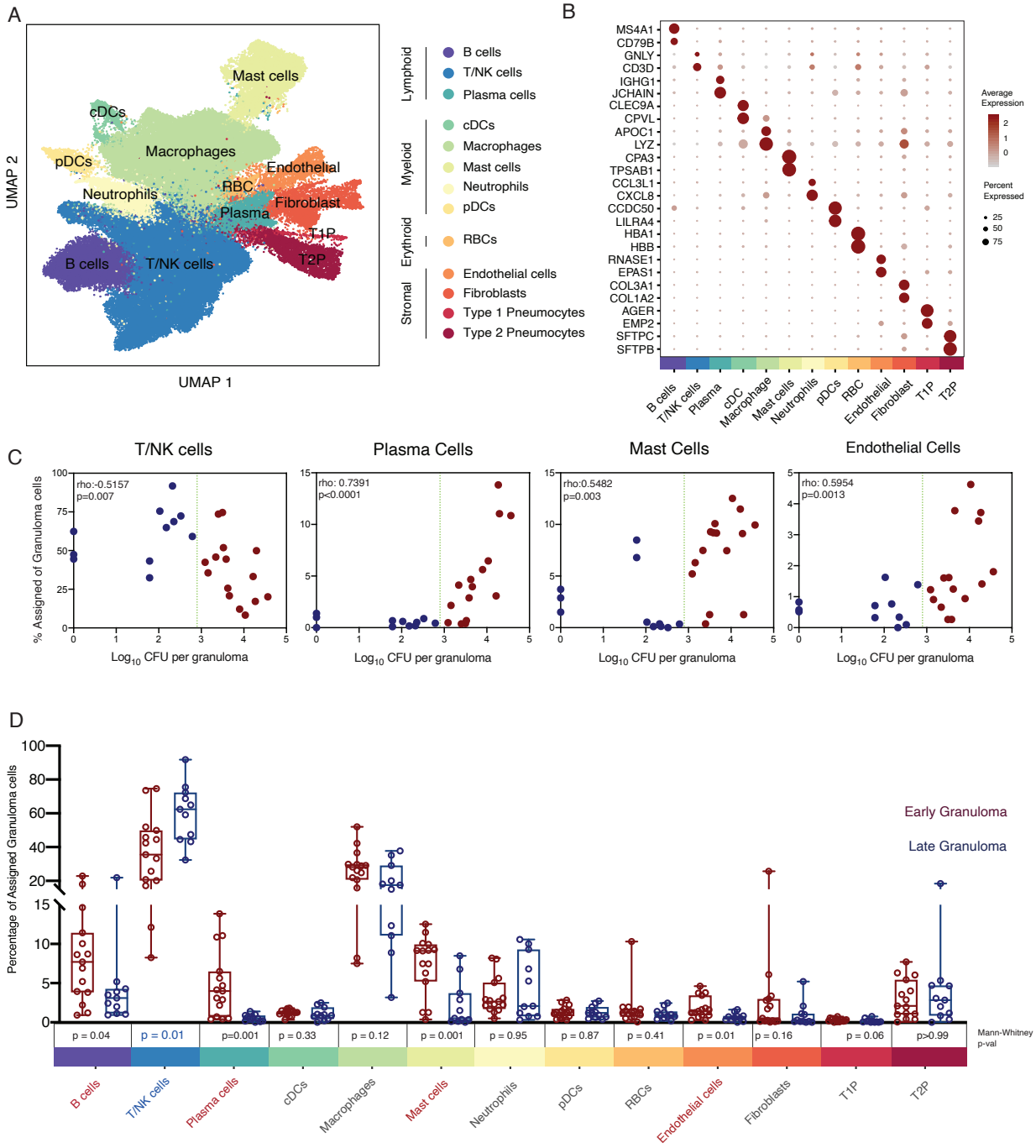


Figure 2 Analysis of single-cell sequencing of tuberculosis lung granulomas

(A) UMAP plot of 109,584 cells from 26 granulomas colored by identities of 13 generic cell types. **(B)** Expression levels of cluster defining genes enriched across 13 generic cell types. Color intensity corresponds to the level of gene expression, while the size of dots represents the percent of cells with non-zero expression in each cluster. **(C)** Significant correlations between proportion of T/NK cells, mast cells, plasma cells and endothelial cells with bacterial burden of individual granulomas (Log_{10} CFU per granuloma) using non-parametric Spearman's rho correlation test. **(D)** Difference in granuloma proportional composition of cell type clusters and CFU between early (maroon) and late

granulomas (navy blue). Statistics: Mann Whitney U. p values are presented in boxes. Box plot showing median, IQR and range; each dot represents a granuloma.

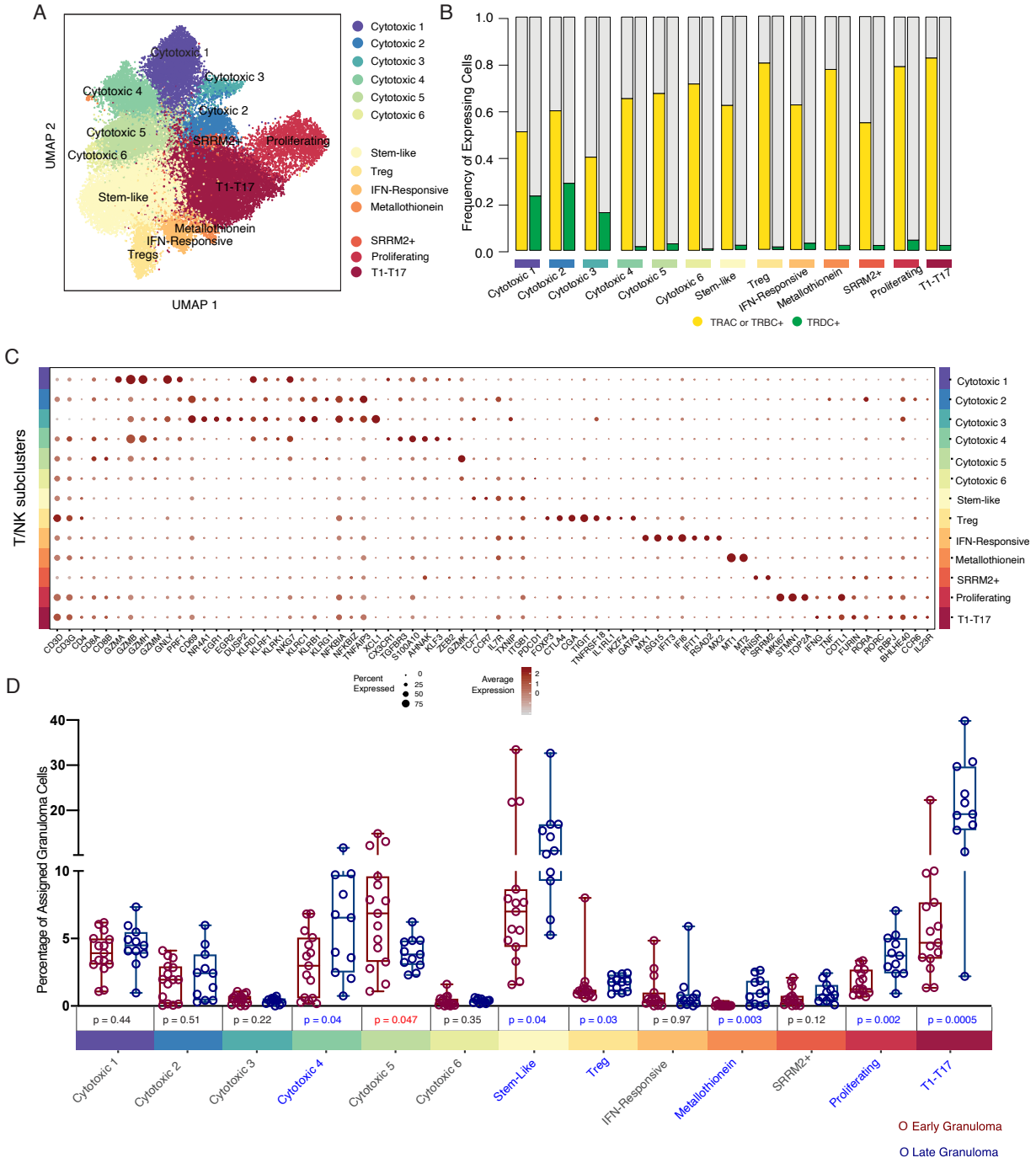
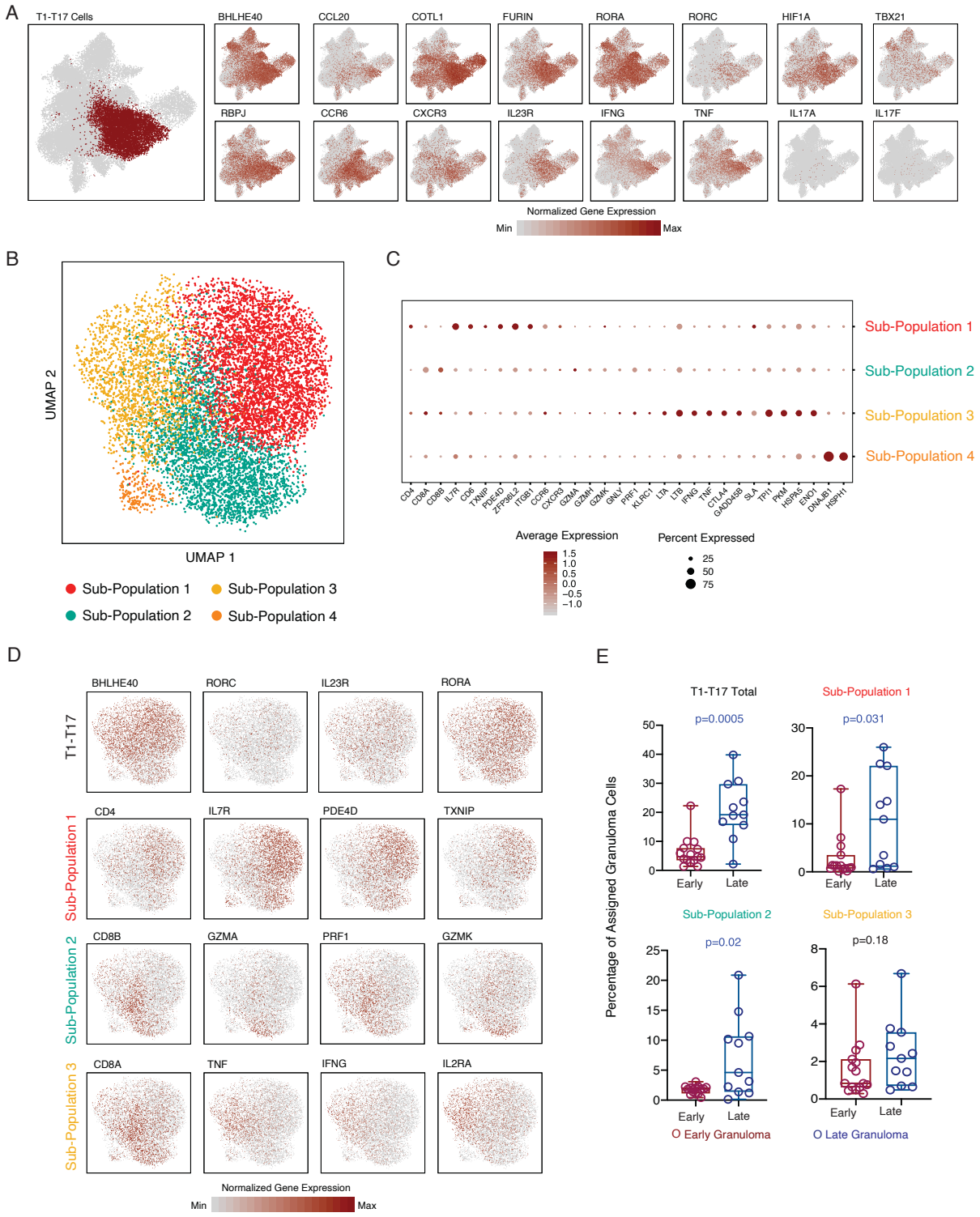


Figure 3.Diversity in the unified T and NK cell cluster and relationship to granuloma-level bacterial burden.

(A) Subclustering 41,222 cells in the unified T/NK cell cluster, colored by subclusters. Subclusters are numbered based the expression patterns. **(B)** Frequency of expression of TCR genes *TRAC*, *TRBC1* or *TRBC2* (yellow) and *TRDC* (green) across 13 T/NK cell subclusters. **(C)** Expression levels of T/NK cell cluster-defining genes. Color intensity corresponds to the level of gene expression and the size of dots represents the percent of cells with non-zero expression in each cluster. Y-axis identifies subclusters. **(D)** Difference in assigned proportion of T/NK cells

and subclusters (1-13) between early (maroon) and late granulomas (navy blue). Statistics: Mann Whitney U. p values are presented in boxes. Box plot showing median, IQR and range; each dot represents a granuloma.



805 **Figure 4 Phenotypic Diversity in T1-T17 cells.**

(A) T1-T17 subcluster overlaid on unified T/NK cell cluster (left) and colored by normalized expression values for T1-T17 subcluster-defining genes (bold outlined boxes) and non-enriched canonical Type1 and Type 17 genes (right). **(B)** Subclustering of 9,234 T1-T17 cells resulting in 4 phenotypic sub-populations. **(C)** Cluster defining genes for T1-T17 subpopulation 1, 2, 3 and 4. Color intensity corresponds to the level of gene expression and the size of dots represents the percent of cells with non-zero expression in each cluster. **(D)** Subclustering of T1-T17 cells colored by normalized gene expression values for selected subcluster (top row) and sub-population defining genes. **(E)** Difference in T1-T17 sub-populations proportion between early (maroon) and late granulomas (navy blue). Statistics: Mann Whitney U.. Box plot showing median, IQR and range; each dot represents a granuloma.

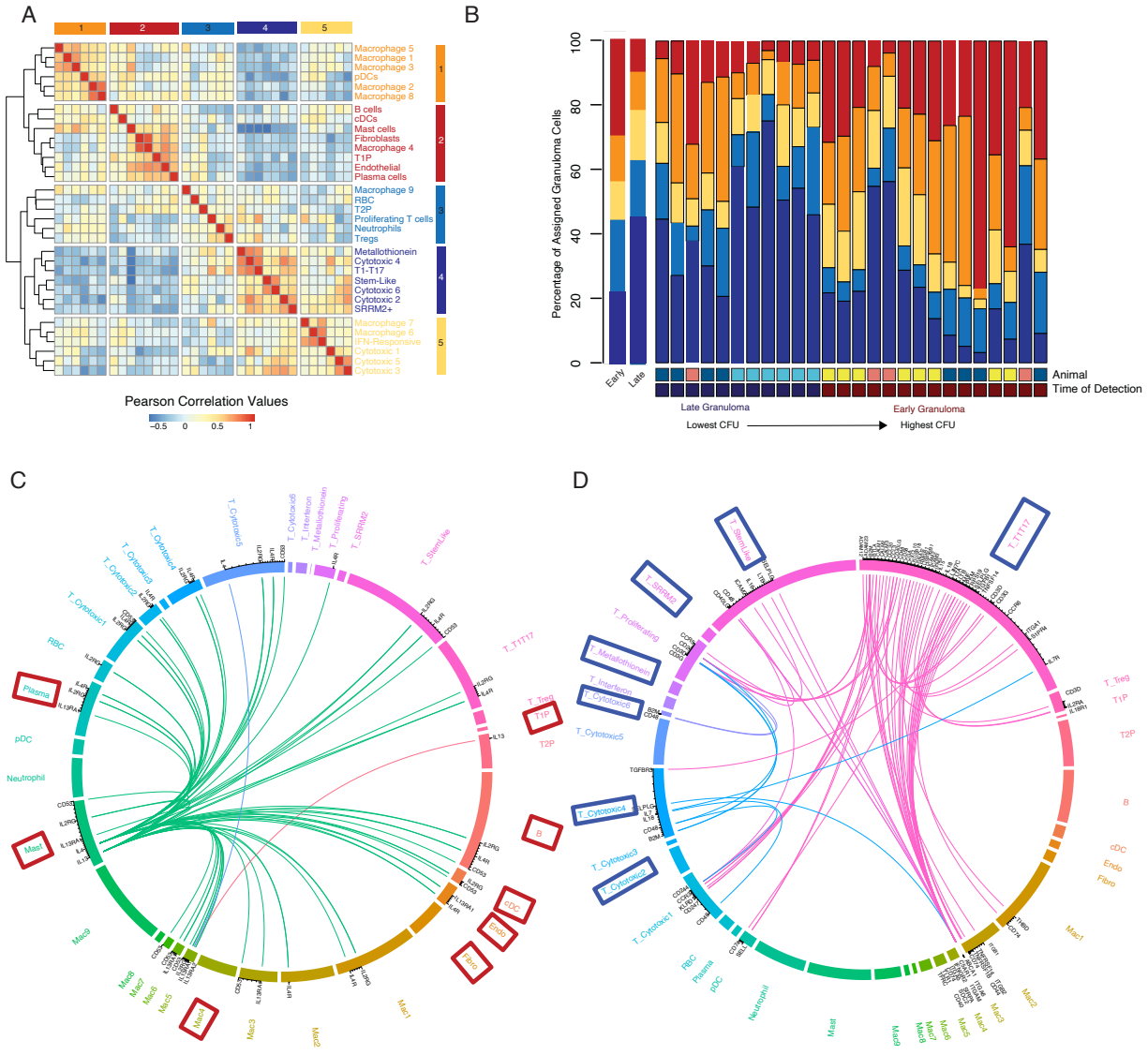


Figure 5 Cellular ecosystem in TB lung granulomas.

(A) Pairwise Pearson correlation values proportions of canonical cell types and T/NK and macrophage subclusters across 26 granulomas. Hierarchical clustering of correlation coefficients identified 5 groups (indicated by color and number) of cell types with correlated abundance in granulomas. **(B)** Relationship between the distribution of correlated cell-types and bacterial burden and timing of granuloma establishment (left), and across all 26 granulomas ordered from lowest CFU (left) to highest CFU (right) and timing of granuloma detection in PET-CT.

Colored boxes which granulomas came from which animal by salmon boxes (3817), yellow boxes (3917), Navy blue boxes (4017) and 4217 boxes (light blue) and time of detection is indicated by dark blue boxes (10 weeks) and maroon boxes (4 weeks). **(C)** Circos plot showing receptor-ligand edge weights for interactions between various cell types in early granulomas. Group 2 cell types highlighted with red boxes.**(D)** Circos plot showing receptor-ligand edge weights for interactions between various cell types in late granulomas. Group 4 cell types are highlighted with blue boxes.

Table 1. T/NK subclusters characteristics and annotation

Subclusters	CD4/CD8 and TCR	Markers	Inference
Cytotoxic C1	CD8A, TCRA, TCRB, TCRD	*GNLY*, *PRF1, GZMH*, *GZMA GZMB, KLRD1, KLRC1*, *KLRC2*, *NKG7*	innate CD8+ T cells (possibly NKT cells), gd T cells, and natural killer cells (NK)tri-cytotoxic potential
Cytotoxic C2	CD8A, TCRA, TCRB, TCRD	*PRF1,KLRC1, KLRB1, KLRG1, CD69, NR4A1, EGR1, EGR2 and DUSP2. NFKBIA* (IkB), *NFKBIZ* and *TNFAIP3*,	Gamma delta T cells, Peripheral tolerance, Activation, Inhibition of NFkB signaling
Cytotoxic C3	CD8A, TCRA, TCRB, TCRD	*XCL1, EGR1, NR4A1*	XCL-1+ NK cells
Cytotoxic C4	CD8A, TCRA, TCRB, CD4 (low)	*PRF1, GZMH*, *GZMB, GZMM, CXCR1, TGFBR3, S100A10, AHNAK*, *KLF3*, and *ZEB2. KLRD1, KLRF1, KLRK1* and *NKG*	CD8 T cells, Early differentiation, Regulation, Apoptosis, Tissue residency
Cytotoxic C5	CD8A, TCRA, TCRB,	*GZMK, CRTAM, PIK3R1, GZMM, EOMES, KLRG1*	Granzyme K expressing CD8 T cells
Cytotoxic C6	CD8A, TCRA, TCRB,	GZMK, FTH1	
Stem-like	CD4, CD8A/B	*PLK2, TCF7*, *CCR7, IL7R, TXNIP, CD69* and *ITGB1*	Stem-like, early memory, self-renewal
T reg	CD4, CD8A/B	*FOXP3, CTLA4, CGA, TIGIT, TNFRSF18, IL1RL1, IKZF4, GARTA3*	Regulatory T cells
Interferon	CD4, CD8A/B	*MX1*, *ISG15*, *IFIT3*, *IFI6*, *IFIT1*, *RSAD2*, and *MX2*	Interferon inducible T cells
Metallothionin	CD4, CD8A/B	*MT1* and *MT2*	Metallothionin+ T cells
SRRM2	CD4, CD8A/B	*PNISR* and *SRRM2*	SRRM2+ T cells
Proliferation	CD4, CD8A/B	*MKI67*, *STMN1*, and *TOP2A*	Proliferating T cells
T1T17	CD4, CD8A/B, TCRA,TCRB,	**Transcription factors** *RORA, RORC, RBPJ, BHLEHE40, FURIN, COTL1, CCL20* **Surface receptors:***CCR6, IL23R,CXCR3* **Cytokine:** IFNG, TNF	T1T17 hybrid tissue-resident effector and effector-memory T cells