Multimodal profiling of lung granulomas reveals cellular correlates of tuberculosis control

Authors: Hannah P. Gideon[1, 2]*, Travis K. Hughes[3,4,5]*, Constantine N. Tzouanas[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], Jia Yao 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

24 [7]Department of Biological Engineering, Massachusetts Institute of Technology,
25 Cambridge, MA
26 [8]Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of
27 Public Health, Boston, Massachusetts, USA
28 [9]Program in Computational and Systems Biology, Massachusetts Institute of Technology,
29 Cambridge, MA, USA
30 [10]Africa Health Research Institute, Durban, South Africa
31 [11]Division of Pulmonary, Allergy and Critical Care Medicine, University of Pittsburgh
32 School of Medicine, Pittsburgh, PA, USA
33 [12]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of
34 Technology, Cambridge, MA, USA
35 [13]Division of Infectious Diseases, Massachusetts General Hospital, Boston, MA
36 [14]Department of Infectious Diseases and Microbiology, Graduate School of Public
37 Health, University of Pittsburgh, Pittsburgh, PA
38 [15]Department of Cardiothoracic Surgery, University of KwaZulu Natal, Durban, South
39 Africa
40 [16]Department of Chemical Engineering, Massachusetts Institute of Technology,
41 Cambridge, Massachusetts, United States
42 [17]The Koch Institute for Integrative Cancer Research, Massachusetts Institute of
43 Technology, Cambridge, Massachusetts, United States
44 [18]Department of Pediatrics, University of Pittsburgh School of Medicine, UPMC
45 Children's Hospital of Pittsburgh, Pittsburgh PA USA
46 [19]School of Laboratory Medicine and Medical Sciences, University of KwaZulu-Natal,
47 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 lung granulomas by co-registering longitudinal PET-CT imaging, single-cell RNA-sequencing, and measures of bacterial clearance. We find that bacterial persistence occurs in granulomas enriched for mast, endothelial, fibroblast and plasma cells, signaling amongst themselves via Type II immunity and wound healing pathways. In contrast, these interactions are largely absent in granulomas that drive bacterial control, which are often those that form later in the course of infection; these restrictive lesions are characterized by cellular ecosystems enriched for Type1-Type17, stem-like, and cytotoxic T cells engaged in pro-inflammatory signaling networks that involve diverse myeloid and non-immune cell populations. There is also a temporal aspect to bacterial control, in that granulomas that arise later in infection (in the context of an established immune response) share the functional characteristics of restrictive granulomas and are more capable of killing Mtb. Taken together, our results define the complex multicellular ecosystems underlying (lack of) granuloma resolution and highlight host immune targets that can be leveraged to develop new vaccine and therapeutic strategies for TB.

72

73 **One-Sentence Summary:** Bacterial control in TB lung granulomas correlates with distinct

74 cellular immune microenvironments and time of formation after infection.

Introduction

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (Mtb), remains a major global health threat (WHO, 2019). Mtb infection is characterized by the formation of granulomas predominantly in the lungs and lymph nodes (Flynn, 2010; Lin et al., 2014b; Russell et al., 2010; Ulrichs and Kaufmann, 2006). These spatially organized structures, composed of a mixture of immune and non-immune cells (Ehlers and Schaible, 2012; Flynn, 2010; Gideon et al., 2019; Lin et al., 2006; Mattila et al., 2013; Pagan and Ramakrishnan, 2014; Phuah et al., 2012; Reece and Kaufmann, 2012; Ulrichs and Kaufmann, 2006), are key sites of host-pathogen interactions which can either restrict or facilitate bacterial survival **(Fig 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 (Canetti, 1955; Flynn, 2010; Lin *et al.*, 2006).

A spectrum of granuloma types, organization and cellular composition has been described in both humans and non-human primates (NHP) (Canetti, 1955; Flynn, 2010; Hunter, 2011; 2016; Lin *et al.*, 2006). Studies of Mtb infection in NHP have demonstrated that individual granulomas are dynamic (Coleman et al., 2014b; Lin et al., 2013; Lin *et al.*, 2014b), changing in response to evolving interactions between bacteria and diverse host cell types (Ehlers and Schaible, 2012; Flynn, 2010; Flynn et al., 2003; Mattila *et al.*, 2013; Phuah *et al.*, 2012; Ulrichs and Kaufmann,

99 2006). The bacterial burden in individual granulomas is highest early in infection and then
100 decreases due to increased bacterial killing as the immune response matures, even in animals that
101 ultimately develop active TB (**Fig S1B-C)** (Cadena et al., 2016; Lin *et al.*, 2014b; Maiello et al.,
102 2018). Strikingly, however, individual granulomas within a single host follow independent
103 trajectories with respect to inflammation, cellular composition, reactivation risk, and ability to kill
104 Mtb (Coleman *et al.*, 2014b; Gideon et al., 2015; Lenaerts et al., 2015; Lin *et al.*, 2013; Lin *et al.*,
105 2014b; Malherbe et al., 2016; Martin et al., 2017). We and others have profiled immune responses
106 among individual cell types in macaque lung granulomas, including those of T cells (Diedrich et
107 al., 2020; Foreman et al., 2016; Gideon *et al.*, 2015; Lin et al., 2012; Mattila et al., 2011; Wong et
108 al., 2018), macrophages (Mattila *et al.*, 2013), B cells (Phuah et al., 2016; Phuah *et al.*, 2012), and
109 neutrophils (Gideon *et al.*, 2019; Mattila et al., 2015), and also examined the instructive roles of
110 cytokines, including IFN-γ, IL-2, TNF, IL-17 and IL-10 (Gideon *et al.*, 2015; Lin et al., 2010;
111 Wong et al., 2020). While these analyses have enabled key insights into how specific canonical
112 cell types and effector molecules relate to bacterial burden, they have been relatively narrow and
113 directed in focus, and have not revealed how the integrated actions of diverse cell types within
114 individual granulomas influence control.
115

116 The emergence of high-throughput single-cell genomic profiling methods affords transformative
117 opportunities to define the cell types, phenotypic states and intercellular circuits that comprise
118 granulomas and inform their dynamics (Prakadan et al., 2017). Here, we developed and applied a
119 multifactorial profiling pipeline—integrating longitudinal PET-CT imaging, single-cell RNA-
120 sequencing (scRNA-seq)-based immunophenotyping, molecular measures of bacterial killing with
121 immunohistochemistry and flow cytometry—to identify features of TB lung granulomas that
122 correlate with bacterial clearance in cynomolgus macaques (**Fig 1A**). Leveraging it, we define the

general cellular compositions and specific cell states associated with bacterial persistence or control. We further uncover TB-associated intercellular signaling networks and how they differ across granulomas that have different levels of bacterial clearance, identifying distinct participating cell types and pathways implicated in bacterial persistence or control. Collectively, our data define the cellular environments and holistic interaction networks within TB lung granulomas in which Mtb is controlled or alternatively survives and multiplies, nominating novel therapeutic and prophylactic targets for future investigation.

130 **Results**

131 We sought to define the complex cellular ecosystems of granulomas that manifest different degrees
132 of bacterial control in NHP. Four cynomolgus macaques were infected with a low dose of Mtb
133 (<10 CFU; Erdman strain) and followed for 10 weeks (**Fig 1A**). 10 weeks post-infection was
134 chosen as a pivotal time point in which bacterial killing can be identified in some but not all
135 granulomas (**Fig S1B**), providing the potential to examine a range of bacterial burdens across
136 granulomas in our analyses. Progression of Mtb infection and individual granuloma dynamics were
137 monitored at 4, 8, and 10 weeks post infection (p.i.) using PET-CT imaging of FDG avidity as a
138 proxy for inflammation **(Fig S1D-E, Table S1)** (Coleman *et al.*, 2014b; White et al., 2017). At
139 necropsy, individual PET-CT identified lung granulomas were excised and dissociated to obtain a
140 single-cell suspension; viable bacterial burden (CFU, colony forming units – i.e., culturable live
141 bacterial burden) and cumulative (live + dead) bacterial load (chromosomal equivalents, CEQ)
142 were measured to define the extent of bacterial growth and killing in each granuloma (Lin *et al.*,
143 2014b; Munoz-Elias et al., 2005).

144

145 Twenty-six granulomas from these four animals were randomly selected at the time of necropsy
146 for scRNA-seq analysis. Among the 26, there was a range of granuloma-level bacterial burdens,
147 from sterile (0 CFU/granuloma) to high (4.6 $\log_{10}$ CFU/granuloma) (**Fig 1B-C**; **Table S1**). The
148 granulomas were binned based on bacterial burden (low, n=13 and high, n=13). There was a
149 significant difference in CFU between low and high CFU granulomas (median 2.2 (low) vs 3.6
150 (high) $\log_{10}$ CFU/granuloma, $p<0.0001$, Mann Whitney U test) (**Fig 1C).** To determine whether
151 low CFU reflected reduced bacterial growth or increased bacterial killing, we assessed the total
152 number of bacterial genomes (CEQ), where we have previously shown that the genomes of dead
153 bacteria are not readily cleared and that CEQ provides a measure of cumulative bacterial load

154 (Munoz-Elias *et al.*, 2005). We observed no significant difference in CEQ values between low and
155 high burden granulomas, indicating that the granulomas supported roughly similar cumulative Mtb
156 growth over the course of infection (**Fig 1D**). However, the extent of bacterial killing, calculated
157 as the ratio of CFU to CEQ, was significantly higher in the low bacterial burden granulomas (**Fig**
158 **1E**), indicating that the lower CFU reflected greater killing rather than more limited bacterial
159 growth.
160
161 We then sought to identify granuloma features correlated with the degree of bacterial control. *Post-*
162 *hoc* analysis of serial PET CT imaging data revealed a strong association between the apparent
163 timing of lesion formation and the extent of bacterial control. All the high bacterial burden
164 granulomas were detected at the 4-week scan, while most (11/13) of the low bacterial burden
165 granulomas were first detected at the final pre-necropsy scan (10 weeks) (**Fig S1E, Fig 1F-G**).
166 Consistent with these data, we further evaluated bacterial burden between early and late appearing
167 granulomas in a total of 10 animals at 10 weeks p.i. (**Fig S1G-H**) and again found that the median
168 CFU/granuloma per animal was significantly lower in late granulomas as compared to early ones.
169 We considered the model that late lesions have lower CFU because the bacterial population has
170 simply not had sufficient time to expand. However, since the cumulative bacterial burdens (CEQ)
171 in early and late lesions were not significantly different (**Fig 1H**), the data are consistent with more
172 bacterial killing in late appearing granulomas (-2.1 $\log_{10}$ CFU/CEQ per granuloma) as compared
173 to early appearing ones (-1.2 $\log_{10}$ CFU/CEQ per granuloma, p=0.01, Mann Whitney U test)
174 (**Figure 1I**). Late appearing granulomas could be due to differences in the timing of lesion
175 formation, most likely due to a dissemination event from an early granuloma, such that bacterial
176 replication occurred in the context of an activated immune response (Martin *et al.*, 2017) or

177 differences in the characteristics of the initial inflammatory response such that late appearing
178 granulomas were not detectable by imaging until later in infection.
179
180 ***Cellular composition of TB lung granulomas***
181 To identify cellular and molecular factors associated with increased Mtb killing in an unbiased
182 fashion, we loaded a single-cell suspension from each of the 26 granulomas onto a Seq-Well array
183 (Gierahn et al., 2017) under Biosafety Level 3 conditions, and then processed and sequenced as
184 previously described (Gierahn *et al.*, 2017). After aligning the data to the *Macaca fascicularis*
185 (cynomolgus macaque) genome and performing robust quality controls and granuloma-specific
186 technical corrections, we retained 109,584 high-quality single-cell transcriptomes for downstream
187 analysis (**Fig S2; Table S2**).
188
189 Among these, we resolved 13 general cell types (**Fig 2A,B** and **Fig S3A-G**) through dimensionality
190 reduction, Louvain clustering, and examination of canonical lineage defining genes and reference
191 signatures from the Tabula Muris (Tabula Muris et al., 2018), Mouse Cell Atlas (Han et al., 2018)
192 and SaVanT database (Lopez et al., 2017) (**Fig S3 A-G, Table S3**). These 13 encompass groups
193 of lymphocytes, including B cells (defined by expression of *MS4A1*, *CD79B*, & *BANK1*), T and
194 NK cells (T/NK; *GNLY, TRAC, CD3D,* & *GZMH*) and plasma cells (*IGHG1* & *JCHAIN)*); myeloid
195 cells, including conventional dendritic cells (cDCs; *CLEC9A, CST3,* & *CPVL*), plasmacytoid
196 dendritic cells (pDCs; *LILRA4* and *IRF8*), and macrophages (*APOC1, LYZ,* and *APOE*); mast cells
197 (*CPA3* & *TPSAB1*); neutrophils (*CCL2, CXCL8,* & *CSF3R*); erythroid cells (*HBA1* & *HBB*);
198 stromal cells, including endothelial cells (*RNASE1, EPAS1,* & *FCN3*) and fibroblasts (*COL3A1,*
199 *COL1A1,* & *DCN)*; Type-1 pneumocytes (*AGER*); and, Type-2 pneumocytes (*SFTPC, SFTPB,*
200 and *SFTPA1*) (**Fig 2A & B, Fig S3G** and **Table S3 & S4**). For each of the 13 cell types, we also

201 performed further within cell-type sub-clustering; in these analyses, we only detected substructure
202 among the T/NK and macrophage clusters (detailed below, **Methods**).
203
204 ***Cell types associated with timing of granuloma formation and control***
205 To investigate the relationship between cell type composition and bacterial burden, we quantified
206 the correlation between cellular frequency and CFU across all granulomas. We considered cellular
207 frequencies in granulomas as a function of viable bacterial burden. Our data reveal multiple cell
208 types that are significantly enriched in higher bacterial burden (early appearing) granulomas,
209 including plasma cells (relative cell abundance vs CFU, $p<0.0001$, non-parametric Spearman's rho
210 correlation test), mast cells ($p=0.002$), endothelial cells ($p=0.001$) and fibroblasts ($p=0.011$) (**Fig**
211 **2C, Table S5**). By contrast, T/NK cells were more abundant in lower bacterial burden (late
212 appearing) granulomas ($p=0.0055$) (**Fig 2C, Table S5**). Cynomolgus macaques are variable in
213 their infection outcomes (**Fig 1B**), so to control for inter-subject variability, each of the cellular
214 associations between granuloma dynamics and bacterial control was examined both across all
215 animals and lesions, and through a directed analysis of the granulomas from a single NHP host
216 (4017) (**Fig S3H**). We further confirmed these trends by performing deconvolution on bulk RNA-
217 sequencing data of 12 additional granulomas (6 high CFU (early) and 6 low (late) bacterial burden
218 granulomas) from separate macaques (**Fig S4A**).
219
220 ***High bacterial burden granulomas are characterized by fibrosis and Type II immune features***
221 The presence and function of mast cells in Mtb lung granulomas has not been previously described.
222 Therefore, to validate this observation, we performed immunohistochemistry on NHP and human
223 granuloma sections using Tryptase and C-kit/CD117 markers (**Fig S4D & E**). This confirmed the
224 presence of mast cells within both NHP and human granulomas, and further revealed that they

225 primarily localize to the outer regions of NHP granulomas, including the lymphocyte cuff (**Fig**
226 **S4D**), and can be found within and around human granulomas (**Fig S4E**). In our data, mast cells
227 are distinguished by their expression of *IL4* and *IL13* (**Fig S4B**), which we also recently observed
228 in a study of human nasal polyposis, a Type II inflammatory disease associated with dramatic
229 epithelial remodeling (Ordovas-Montanes et al., 2018). Mast cells are also marked by expression
230 of *ALOX5A* and *ALOX5AP*, which encode the system to synthesize the anti-inflammatory lipoxin
231 LXA4; the balance between LXA4 and the pro-inflammatory lipoxin LTB_4 has been strongly
232 implicated in the progression of TB disease in humans (Tobin et al., 2012; Tobin et al., 2010).
233
234 Plasma cells are also abundant in high burden lesions, consistent with previous findings (Jacobs et
235 al., 2016; Phuah *et al.*, 2012). Recruitment of mast cells can be characteristic of allergic Type II
236 immune responses mediated by IgE (Kanagaratham et al., 2020), but mast cell function is also
237 regulated by IgG, which is much more abundant in the circulation and tissues. Among the plasma
238 cells in our scRNA-seq dataset, the vast majority express either *IGHG* or *IGHA* (Collins and
239 Jackson, 2013) constant chains (**Fig S4B, C)**, suggesting that IgG and IgA are the dominant
240 antibody classes induced by Mtb infection in the granuloma microenvironment. Taken together,
241 these data suggest that granulomas with failed bacterial clearance are characterized by a Type II
242 immune environment, but the antibody features are not consistent with a canonical allergic
243 response.
244
245 ***T and NK functional subclusters as mediators of protection***
246 Of the 13 broad cell types, only the T/NK cell subcluster is associated with more robust bacterial
247 control in granulomas (p=0.0055, non-parametric Spearman's rho correlation test) (**Fig 2C**). To
248 further assess functional diversity within the 41,622 cells that comprise the T and NK cell cluster

249 and their association with bacterial burden, we performed additional sub-clustering analyses. This
250 revealed 13 T/NK cell subclusters which we annotated based upon expression of lineage defining
251 markers, known cytotoxic, regulatory and proliferation genes (**Fig 3A, C** and **S5, Tables 1 and**
252 **S6)** and TCR constant gene (*TRAC*, *TRBC*, and *TRDC*) expression (**Fig 3B).** The process of
253 annotation revealed that most subclusters did not correspond neatly to canonical T and NK cell
254 subsets, consistent with recent studies in other systems (Rath et al., 2020). Where possible, we
255 annotated each based on known T cell markers and literature-derived genes of interest; we note
256 that these genes are parts of broader transcriptional signatures that appear to reflect dominant
257 cellular response states superimposed on cell lineage-associated gene expression programs.
258 Among the 13 T/NK cell subclusters, 6 were significantly negatively associated with bacterial
259 burden (**Fig 3D, Table S5**).
260

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

262 One T/NK cell subcluster represented the most abundant cell type identified across all granulomas
263 (8.8%) (**Table S4**) and the strongest correlate with bacterial control (p=0.001, non-parametric
264 Spearman's rho correlation test) (**Fig 3D**; **Table S4 & S5**). This subcluster, which we designated
265 Type1-Type17 (T1-T17) (**Fig 3C**), is enriched for expression of classical Th1-associated genes,
266 including *IFNG* and *TNF* (Raphael et al., 2015), as well as transcription factors associated with
267 Th17 differentiation (Yosef et al., 2013), including *RORA* (Yang et al., 2008), *RORC* (Ivanov et
268 al., 2006), *RBPJ* (Meyer Zu Horste et al., 2016), and *BHLHE40* (Huynh et al., 2018; Lin et al.,
269 2016; Lin et al., 2014a). While we also detected additional features of T17 cells, including *CCR6*
270 (Hirota et al., 2007) and *IL23R (Kobayashi et al., 2008)*, we did not observe expression of either
271 *IL17A* or *IL17F* (**Fig 4A; Table S6-7)**. Collectively, this hybrid gene expression state is consistent
272 with previously described expression programs for Th1* or ex-Th17 cells, which are believed to

273 be precursors to tissue resident memory cells (Amezcua Vesely et al., 2019). Previous studies have
274 revealed a prominent role for CD4 Th1 and Th17 cytokines in control of Mtb infection, including
275 IFN-γ, TNF, and IL-17 (Algood et al., 2005; Green et al., 2013; Khader et al., 2007; Khader and
276 Gopal, 2010; Lin et al., 2007; Lyadova and Panteleev, 2015; Millington et al., 2007; O'Garra et
277 al., 2013; Scriba et al., 2017), and studies in NHP granulomas suggest an association between T1
278 and T17 cytokine expression and bacterial burden (Gideon *et al.*, 2015). In addition, in murine
279 models, BHLHE40 is required for control of Mtb infection, as a repressor of IL-10 production
280 (Huynh *et al.*, 2018). Notably, while Th1* and ex-Th17 subsets are described primarily as CD4 T
281 cells (Darrah et al., 2020; Gideon *et al.*, 2015; Lyadova and Panteleev, 2015; Mpande et al., 2018),
282 our T1-T17 sub-cluster is characterized by the expression of both *CD4* and *CD8A/B* transcripts
283 **(Fig 3C and 4C, Fig S5D-E)**.
284

285 To better resolve the identities of the cells in this cluster, we further sub-clustered the 9,234 T1-
286 T17 cells. This revealed 4 distinct subpopulations, each of which expressed T1-T17 cluster
287 markers (*RORA, RORC, IL23R*, and *BHLHE40*) but were further distinguished by markers of cell
288 type and state **(Fig 4B, Table S7**): T1-T17 subpopulation 1 is distinguished by expression of *CD4*
289 and markers of activation and motility, including *IL7R*, *CD6*, *TXNIP*, *PDE4D*, *ZFP36L2*, *ITGB1*,
290 *CCR6*, and *CXCR3* (**Fig 4B,C; Tables 1 and S7**), making it most akin to ex-Th17 cells; T1-T17
291 subpopulation 2 is characterized by increased relative expression of both *CD8A* and *CD8B* and
292 cytotoxic effector molecules; T1-T17 subpopulation 3, which includes cells expressing either
293 *CD8A/B* or *CD4*, is characterized by cytokine gene expression (*IFNG, TNF, LTA*, and *LTB*) and
294 markers of an inhibitory cell state (*CTLA4*, *GADD45B*, and *SLA*); T1-T17 subpopulation 4 is very
295 low in abundance and characterized by heat shock and DNA damage associated transcripts
296 (*DNAJB1* and *HSPH1*). There was a trend towards negative association between bacterial burden

297 and higher abundance of T1-T17 subpopulation 1 (p=0.055, non-parametric Spearman's rho
298 correlation test) and a significant negative association between bacterial burden and abundance of
299 T1-T17 subpopulation 2 (p=0.02). Surprisingly, T1-T17 subpopulation 3 was not correlated with
300 bacterial burden, despite expressing elevated levels of *IFNG* and *TNF* (**Fig 4E**, **Table S5**),
301 cytokines generally considered as critical mediators of control in Mtb infection (O'Garra *et al.*,
302 2013; Scriba *et al.*, 2017).
303
304 ***CD4 and CD8 subclusters associated with low bacterial burden***
305 Among the remaining 12 T/NK cell subclusters, 6 are enriched for both CD4 and CD8 expression
306 (**Fig 3A-C**, **Fig S5D&E**, **Table 1, S6**). Of these, 5 are significantly associated with more robust
307 bacterial control (**Figure 3D & S5D-E**). We annotated the most abundant of these as stem-like T
308 cells (8.3% of granuloma cells, p=0.03 non-parametric Spearman's rho correlation test, **Fig 3D,**
309 **Table S5)** based on elevated expression of markers of naïve and memory T cells (*TCF7*, *CCR7,*
310 *IL7R,* and *TXNIP*) and activation or memory state (*CD69* and *ITGB1*) (**Fig 3C, Table S6**). These
311 cells may represent a "stem-like" population of T cells, which has been described as an early
312 differentiating memory phenotype, distinct from naïve T cells, that are long-lived and possess a
313 unique ability to proliferate and self-renew (Ahmed et al., 2016; Caccamo et al., 2018; Gattinoni
314 et al., 2011). The second CD4/CD8 subcluster associated with control contains proliferating T cells
315 (2.4%; p=0.03; **Fig 3D, Table S5**) and is characterized by high expression of transcripts associated
316 with cellular proliferation (*MKI67*, *STMN1*, and *TOP2A*) **(Fig 3C, Table S6)**, consistent with
317 published data that T cell proliferation occurs within NHP and human granulomas (Gideon *et al.*,
318 2015; McCaffrey et al., 2020; Ohtani, 2013; Phuah *et al.*, 2016; Phuah *et al.*, 2012; Wong *et al.*,
319 2018). The third is a very small population of Metallothionein expressing T cells (0.05%; p=0.03;
320 **Fig 3D, Table S5**), defined by metallothionein genes, such as *MT1* and *MT2* **(Fig 3C, Table S6),**

321 which play a role in negative regulation of Type 1 regulatory (Tr1) CD4+ cells (Wu et al., 2013).
322 The fourth, SRRM2-T cells (0.6%, p=0.007), is characterized by enrichment of genes associated
323 with nuclear speckles and splicing factors such as *PNISR* and *SRRM2* **(Figure 3C&D, Table S5-**
324 **6)**, the latter of which has been associated with alternate splicing in Parkinson disease (Shehadeh
325 et al., 2010) and has a critical role in the structural organization of the genome (Hu et al., 2019).
326
327 The remaining two CD4/CD8 subclusters are not associated with bacterial control. Interestingly,
328 one is regulatory T cells (1.2%), defined by elevated expression of canonical Treg markers
329 (*FOXP3, CTLA4, TIGIT,* and *IL1RL1*) and *GATA3*, a Th2 lineage-defining transcription factor that
330 has been observed in a subset of tissue-resident Tregs (**Fig 3C&D, Table S5-6)**. The final
331 subcluster is interferon responsive T cells (0.4%), which are enriched for Type-I interferon
332 inducible molecules (*MX1*, *ISG15*, *IFIT3*, *IFI6*, *IFIT1*, *RSAD2*, and *MX2*) *(Szabo et al., 2019)* **(Fig**
333 **3C-D, Table S5-6).**
334
335 ***Bacterial control is associated with a specific cytotoxic T cell population***
336 The remaining 6 T/NK subclusters are broadly defined by expression of *CD8A* and/or *CD8B* and
337 cytotoxic genes, including granzymes (*GZMA, GZMB, GZMH, GZMK*, and *GZMM*), granulysin
338 (*GNLY)*, and/or perforin (*PRF1*) (designated Cytotoxic 1-6, **Fig 3C, Table 1**). We confirmed
339 expression of multiple granzymes among CD8$\alpha\beta$ T cells in Mtb granulomas by flow cytometry
340 (**Fig S6**) from animals in other ongoing studies.
341
342 Low bacterial burden granulomas are associated with a higher proportion of cells from cytotoxic
343 subcluster C4 (3.8% of granuloma cells; p=0.02, non-parametric Spearman's rho correlation test)
344 (**Fig 3D; Table S5**). C4 expresses both *CD8A* and *CD8B* and *TCRA* and *TCRB*, but not *TCRD,*

345 indicating that it is composed primarily of conventional CD8$\alpha\beta$ T cells (**Fig 3B&C, S5D)**. C4 is
346 further enriched for genes associated with cytotoxic effector functions (*PRF1, GZMH*, *GZMB,* and
347 *GZMM)*, motility, migration and tissue residency (*CX3CR1, TGFBR3*, and *S100A10*), and
348 regulators of cell state (*AHNAK*, *KLF3*, and *ZEB2*; **Fig 3C, Table S6**).
349
350 The remaining 5 cytotoxic subclusters did not associate with bacterial control. Cytotoxic
351 subclusters C1-3 are enriched for the expression of *CD8A* but not *CD8B* and elevated *TCRD*,
352 implying that these cells possess innate cytotoxic function (**Fig 3B-C)**. C1 is further characterized
353 by high expression of cytotoxic effector genes—*GNLY* and *PRF1; GZMH*, *GZMA* and *GZMB;* as
354 well as *KLRD1, KLRC1*, *KLRC2*, and *NKG7*—which suggests that subcluster 1 contains a greater
355 proportion of highly cytotoxic innate CD8+ T cells (possibly NKT cells), $\gamma\delta$ T cells, and NK cells
356 (**Fig 3B-C, Table 1, S6**). C2 is also enriched for NK receptors and CD8 T cell activation markers
357 in addition to a trio of transcription factors (*EGR1, EGR2*, and *DUSP2*) described to distinguish
358 peripheral tolerant CD8 T cells (Schietinger et al., 2012) **(Fig 3B-C, Table 1, S6)**. C3 appears to
359 be more selectively enriched for NK cells with elevated expression of cytotoxic and NK cell
360 markers and low expression of *CD3D* and *CD3G.* C5, which like C4 expresses both *CD8A* and
361 *CD8B* and *TCRA* and *TCRB*, but not *TCRD*, is distinguished by elevated expression of *GZMK* (**Fig**
362 **3C)**; granzyme K expressing CD8 cells have been recently described as a hallmark of immune
363 dysfunction in inflammation (Mogilenko et al., 2021). C6 was not detected in sufficient frequency
364 (<0.3%) to draw meaningful conclusions. The functional complexity of these 6 subclusters, along
365 with the common and distinct responses they represent, suggests a significant and
366 underappreciated role for cytotoxic cells in TB granulomas.
367
368

369 ***Macrophage heterogeneity in Mtb granulomas***

370 While macrophages are responsible for much of the bacterial killing within granulomas, we did
371 not observe any association between overall macrophage abundance and bacterial burden (**Fig 2**
372 **and S7**). Yet, like the T/NK cell cluster, the macrophage cluster had discernable substructure based
373 on unbiased gene expression analyses. Among the 27,670 macrophages, we identified 9
374 subclusters (**Table S8**), none of which were independently associated with bacterial control with
375 the exception of Mac 4 (0.07%), a very small subpopulation of macrophages expressing *INSIG1*
376 and *EREG* ($p<0.0001$) (**Fig S7E, Table S8**)).

377

378 ***Cellular ecology of pulmonary TB granulomas***

379 Given demonstrable differences in cellular composition across the bacterial burden spectrum, we
380 wondered whether specific cell types significantly co-occur in TB lung granulomas and
381 collectively influence control. We calculated the pairwise Pearson correlation matrix between all
382 major cell types, subclusters, and subpopulations across the 26 granulomas (**Fig 5A**). Using
383 hierarchical clustering of this pairwise correlation matrix, we defined 5 groups of cell types whose
384 collective abundances are associated across granulomas (**Fig 5A**, **Table S9**). Of these, Group 2
385 (shown in red), which includes mast cells, plasma cells, macrophage subcluster 4 and certain
386 stromal populations, is significantly expanded in high bacterial burden granulomas (Mann-
387 Whitney U Test, $p=3*10^{-4}$; **Fig 5B, Table S10, S11**). Group 3 (shown in blue) is significantly more
388 abundant in low bacterial burden granulomas ($p=0.026$; **Fig 5B**, **Table S10, S11**) and consists of
389 many T cell subclusters/subpopulations, including Stem-like, Cytotoxic subclusters C2, C4, & C6,
390 Metallothionein, Proliferating, SRRM2+, T1-T17 subpopulations 1,3 and 4, as well as a single
391 macrophage subset, Mac7. This macrophage subset is distinguished in part, by expression of the
392 immunomodulatory genes *IDO* and *CHIT* (encoding chitotriosidase), which is abundantly

393 produced by lipid-laden macrophages in other conditions such as Gaucher's disease, Niemenn-
394 Pick disease, and atherosclerosis (Barone et al., 2007; Yap et al., 2020).
395
396 ***Distinct cellular ecosystems associate with granuloma-level bacterial burden***
397 To further explore how specific cellular compositions might constitute distinct tissue niches that
398 support different levels of bacterial control, we examined putative cell-cell interactions within each
399 granuloma. For each potential interacting cell-type pair, we constructed edge weights for receptor-
400 ligand combinations, adjusting to account for differences in the abundance of the sender cell type,
401 relative ligand/receptor expression, and the percent of receptor positive cells (**Methods**).
402
403 To obtain an initial view of cell-cell signaling across granulomas of different burden, we examined
404 the extent and strengths of interactions across cell type groups. High bacterial burden lesions are
405 dominated by signals sent by Group 2 cell types (i.e., mast, fibroblast, endothelial, plasma, type I
406 pneumocyte, and macrophage subset 4); these cell types display the highest counts of high burden-
407 linked interactions as well as those most strengthened in high burden granulomas ($p < 2.2E-16$,
408 binomial test against null of all Groups having equal interaction likelihoods) (**Fig 5C-D**). In
409 contrast, interactions in low burden granulomas more evenly involve Groups 1, 3, 4 and 5, with
410 Group 3 showing the strongest enrichment for signaling activity strengthened in low burden
411 granulomas ($p = 1.2E-4$, binomial test against null of all Groups having equal interaction
412 likelihoods; $p = 0.008$, binomial test against null of equal interaction likelihoods among non-Group
413 2 cells) (**Fig 5E-F**). These contrasting patterns of intercellular communication suggest distinct
414 signaling architectures underlying different degree of bacterial control, with Group 2 cells
415 dominating activity within high-burden lesions, compared to coordinated signaling across Groups
416 in low-burden cases.

417

418 We further examined shifts in intercellular interaction network topology by more comprehensively
419 quantifying the sender and receiver activity associated with different levels of bacterial burden.
420 This more directed investigation revealed significantly different patterns of intercellular signaling
421 between high and low burden granulomas ($p < 2.2E-16$, Pearson's chi-squared test). Subsetting all
422 interactions to those strengthened in high burden granulomas, we find that Group 2 cell types are
423 the key source of intercellular signals (i.e., senders in 67% of interactions strengthened in high
424 burden granulomas) (**Fig 5G**). High burden lesions also exhibit strong intra-Group 2 signaling,
425 with 58% of signals received by Group 2 cell types originating from Group 2 cell types themselves.
426 This suggests that high burden lesions are driven by self-reinforcing interactions amongst Group
427 2 cell types (e.g., between mast cells, plasma cells, fibroblasts, and endothelial cells). In contrast,
428 when subsetting to interactions strengthened in low burden granulomas, we find only sparse
429 contributions from Group 2 cell types (**Fig 5H**); instead, low burden granulomas are characterized
430 by a more even distribution of signals stemming and terminating in Group 1, 3, 4, and 5 cell types,
431 suggestive of a coordinated immune response involving multiple cellular subsets (e.g., the T cell
432 and macrophage subsets present in these Groups).

433

434 We next examined which specific axes of intercellular communication, and among whom,
435 associate with varying levels of bacterial control. In looking more directly at the signals underlying
436 the interaction networks that associate with burden, we find dramatic differences in intercellular
437 crosstalk involving both canonical and non-canonical immune mediators that may impair or
438 facilitate bacterial control. Among the ligands whose interactions are most strengthened in high
439 burden granulomas, we identify genes implicated in fibrosis (e.g, *FGF1, PDGFB, CTGF, FGF7,*
440 *IL34*), vascular remodeling (*VEGFB, VEGFC*, *ANGPTL4*) and TGFβ signaling (*TGFB3, BMP6*),

441 suggestive of a wound healing response (**Fig 5I**) (Joshi et al., 2020; Padela et al., 2008). In addition,
442 we observe evidence of intercellular communication via genes implicated in Type II immunity
443 (*CCL11, CCL13, CD5L, IL4, IL5, IL13, IL24*) and allergy-linked inflammation (*CCL19*) (Nakano
444 et al., 2019). We note that these specific ligands are largely produced and received by Group 2 cell
445 types (with only sparse contributions from Groups 3-5). Collectively, this supports a model where
446 intra-Group 2 signaling drives a self-reinforcing high burden microenvironment via wound
447 healing-like responses and associated Type II immune activity (**Fig 5J**). This interpretation is
448 further supported by an enrichment of pathways such as FGF, VEGFR, and PI3K signaling, as
449 well as organogenesis and tissue remodeling processes (**Fig S8A**).
450
451 In contrast, low burden granulomas exhibit cell-cell interactions consistent with Type I immune
452 responses (*CCL3, CXCL9/10/11, DLL1, IFNG, IL18*) and Th17 chemoattraction (*CXCL16,*
453 *CCL20*) and successful immune mobilization and activation (Li et al., 2013; Lim et al., 2008;
454 Touzot et al., 2014). Ligands specifically associated with low burden granulomas include co-
455 stimulatory molecules important in immune activation (*CD40LG, CD48, CD70, CD80, CD86*),
456 those involved in lymphocyte adhesion (*CD58*), and antimicrobial peptides (*DEFB1, SLPI*) (**Fig.**
457 **5I**) (Tateosian et al., 2012). Various antimicrobial peptides have been implicated previously in
458 direct control of MTB infection (Fabri et al., 2011; Liu et al., 2006); whether intercellular
459 communication is an essential or auxiliary role remains to be determined. Importantly, production
460 of low burden-linked ligands is distributed across the cell types of Groups 1, 3, 4, and 5, but not
461 Group 2 (**Fig. 5J**); signaling occurs between multiple T and macrophage cell subsets, suggesting
462 that successful Mtb control requires coordinated interactions across diverse innate and adaptive
463 immune cell types. Supporting this interpretation, gene set enrichment analyses on ligands and
464 receptors whose interactions are strengthened in low burden granulomas revealed enrichment for

465 processes including T cell activation and differentiation and signaling associated with pro-
466 inflammatory cytokines (e.g., TNF) (**Fig S8A**). Likewise enriched in interactions associated with
467 low burden granulomas are additional metabolic processes such as fatty acid metabolism and heat
468 generation, which have been individually studied extensively in Mtb but here connect to broader
469 signaling interactions associated with Mtb control.
470
471 Beyond ascribing a simple binary role to each cell type, our cell-cell interaction analyses also
472 indicate context-dependent roles for particular cell types and ligands. For instance, with respect to
473 cell types, the macrophage-dominated Group 1 is not statistically correlated with granuloma
474 control in our compositional analyses (**Fig 5A**), but participates in the second most interactions in
475 both high and low burden granulomas (**Fig 5B-C,E,G-H**). The idea of dual roles for Group 1 cells
476 is borne out by examination of the ligands produced by Group 1 cell types in high (e.g., *PDGFB,*
477 *CD5L, TNFSF13*) and low burden (e.g., *CXCL9/10/11, CD86, IL18, CCL20*) microenvironments
478 (**Fig 5I-J**). Similarly, we observe that some individual ligands participate in interactions
479 strengthened in both high and low burden granulomas, suggesting pleiotropic effects for these
480 molecules. As one specific example, IL-1's effects on Mtb control vary based on disease stage and
481 model (Juffermans et al., 2000; Law et al., 1996; Mayer-Barber et al., 2014); based on our analyses,
482 *IL1A* and *IL1B* each mediate interactions associated with both high and low bacterial burden, but
483 are derived from different sender cell populations in the two instances. Thus, our intercellular
484 interaction analyses uncover axes of cellular plasticity and ligand pleiotropy across granuloma
485 microenvironments, important for improved understanding and therapeutic modulation of Mtb.
486
487
488

489 **Discussion**

490 Within an individual with Mtb infection, distinct granulomas can achieve sterilizing immunity,
491 immune standoff, or frank immune failure (Flynn, 2006; 2010; Lin *et al.*, 2014b; Lin et al., 2009).
492 In NHPs, which most closely recapitulate human Mtb infection and disease (Coleman et al.,
493 2014a), this heterogeneity provides an opportunity to define the cellular and molecular factors that
494 correlate with bacterial control to identify potential host-directed prevention and cure strategies
495 for TB. While a spectrum of granuloma-level bacterial control has been appreciated previously,
496 the immune correlates of bacterial control within granulomas have not been mapped
497 comprehensively. By coupling advanced serial imaging, scRNA-seq, and molecular measures of
498 bacterial growth and killing, the present study provides new insights into the immunologic control
499 and temporal evolution of granulomas in Mtb infection: we discover and define how the timing of
500 granuloma appearance correlates with distinct microenvironmental signaling networks formed
501 through host responses and shapes eventual bacterial persistence or control. Overall, our data
502 substantiate a model where the state of the surrounding host cellular ecosystem informs a
503 granuloma's infection trajectory, leading to long-term, stable states which either permit or restrict
504 bacterial survival.

505

506 We find that high CFU, early granulomas are characterized by significantly higher proportions of
507 mast cells and plasma cells, as well as a central group of cell types (further including fibroblasts,
508 and endothelial cells) that exhibits extensive self-directed signaling exchanges. While mast cells
509 have been described in granulomatous conditions, such as TB lymphadenitis (Taweevisit and
510 Poumsuk, 2007), leprosy skin lesions (Bagwan et al., 2004), and liver granulomas (Celasun et al.,
511 1992), and may orchestrate immune cross talk in TB (Garcia-Rodriguez et al., 2017), this is the
512 first description of direct correlation with failure of Mtb control in TB granulomas. Structurally,

513 we find mast cells inter-digited in the lymphocyte cuff of TB granulomas, physically well
514 positioned to play significant regulatory roles.
515
516 The mast cells in high-burden granulomas are major producers of Type II cytokines, especially
517 *IL4, IL5,* and *IL13*, which are important down-modulators of lymphocyte and macrophage
518 antimicrobial activity, including inhibiting the cytolytic functions of CD8+ T cells (Kienzle et al.,
519 2005; Wijesundara et al., 2013). However, IL4 and IL13 have broader functions in the context of
520 wound healing. Indeed, the cellular interactions in high burden granulomas reveal both specific
521 signaling molecules (e.g., FGF1 from Type 1 pneumocytes, PDGFB from endothelial cells,
522 ANGPTL4 from plasma and mast cells, among others) and broad pathways (e.g., FGF and VEGF,
523 among others) that reflect fibrosis, metabolic remodeling, and angiogenesis. Collectively, these
524 data suggest a cascade of interactions in early appearing granulomas with failed control, whereby
525 an initially permissive environment is reinforced by a tissue remodeling response that seeks to
526 limit and wall off pathologic activity, thereby allowing for persistence of both Mtb and the Type
527 II/wound healing microenvironment itself. While more detailed studies on the roles of wound
528 healing responses and tissue remodeling in TB are indicated, these features may represent critical
529 targets for host-directed therapies that not only need to enhance restrictive adaptive immune
530 responses but also address the maladaptive features of microenvironments permissive to
531 granuloma persistence.
532
533 Indeed, it is striking how strongly the timing of granuloma appearance (as identified by PET-CT
534 imaging) correlates with the formation of distinct classes of complex yet stable cellular
535 communities and their accompanying levels of bacterial control. We note that granulomas
536 identified late by PET-CT imaging may either be formed later—for example through

537 dissemination (Martin *et al.*, 2017) —or take more time to reach the threshold to be identified by
538 PET-CT scans (limit of detection ≥1mm) because of more efficient immune control or differences
539 in the quality of the inflammatory response (Cronan et al., 2021). Regardless of the exact
540 mechanism, late appearing granulomas are characterized by the enrichment of multiple T and NK
541 cell subsets, as well as extensive pro-inflammatory, pro-activating, and pro-migratory signaling
542 networks predominated by T cell subsets, which may exclude or prevent the establishment of self-
543 reinforcing Type II signaling. Moreover, our measures of cumulative bacterial burden (CEQ)
544 indicate that late granulomas have lower bacterial burden because of greater bacterial killing
545 (CFU/CEQ), linking these adaptive immune features to true sterilizing immunity.
546
547 The strongest cellular correlate of bacterial control was a subcluster of cells with transcriptional
548 features of both Type 1 and Type 17 T cells that was expanded in granulomas with bacterial
549 control. Aspects of these data are consistent with recent observations that granulomas established
550 in immune primed environments—e.g., existing Mtb infection (Cadena et al., 2018) or intravenous
551 or intrabronchial BCG vaccination—are characterized by Th1/17 expression patterns that are
552 associated with protection (Darrah *et al.*, 2020; Dijkman et al., 2019); however, we extend these
553 findings, defining appreciable substructure among the T1-T17 subcluster of relevance to control.
554 The *CD4* T1-T17 subpopulation (subpopulation 1) is most consistent with published descriptions
555 of Th1/17 cells (e.g., Th1* or ex-Th17) (Amezcua Vesely *et al.*, 2019). These cells may represent
556 precursors to long lived tissue memory, which has been shown to play a crucial protective role in
557 autoimmunity, bacterial control, and memory immune responses to pathogens (Amezcua Vesely
558 *et al.*, 2019; Liang et al., 2015; van Hamburg and Tas, 2018; Wacleche et al., 2016), including Mtb
559 infection. A recent study using flow cytometry and immunohistochemistry in Mtb infected rhesus
560 macaques support an association of Th1 (IFNγ+) and Th17 (IL-17+) cells in lung tissue with latent

561 infection (Shanmugasundaram et al., 2020); in contrast, another study using scRNA-seq reported
562 activated CD4 and CD8 T cells including Th1 and Th17 in the lung tissue of macaques with
563 pulmonary TB (Esaulova et al., 2021). The *CD8* subsets within the T1/T17 subcluster
564 (subpopulations 2 & 3), meanwhile, have not been described previously. The former of these is
565 strongly associated with bacterial control and may represent a novel immunologic paradigm that
566 can be exploited for vaccine development. Subpopulation 3 intriguingly, expresses elevated *TNF*
567 and *IFNG* but does not associate with bacterial restriction; further profiling will be necessary to
568 establish the significance of this subset and its relation to previously appreciated Type 1 and Type
569 17 features of control (Algood *et al.*, 2005; Gideon *et al.*, 2015; Green *et al.*, 2013; Khader *et al.*,
570 2007; Khader and Gopal, 2010; Lin *et al.*, 2007; Lyadova and Panteleev, 2015; Millington *et al.*,
571 2007; O'Garra *et al.*, 2013; Scriba *et al.*, 2017).
572
573 Our data also revealed an interesting CD4 and CD8 expressing T cell subcluster associated with
574 low burden granulomas that resembles stem-like T cells (Ahmed *et al.*, 2016; Caccamo *et al.*,
575 2018; Cartwright et al., 2016; Fuertes Marraco et al., 2015; Gattinoni *et al.*, 2011; Mateus et al.,
576 2015; Todryk, 2018). We hypothesize that these cells may be a source of T cell renewal in
577 granulomas and may differentiate into the various functional subsets we observe within them. It is
578 possible, however, that these represent memory T cells that are not specific for Mtb antigens but
579 migrate to the granuloma in response to inflammation and/or chemokine gradients. Indeed, flow-
580 cytometry based studies support that a majority of T cells in granulomas do not respond to Mtb
581 antigens by making cytokines and do not display hallmarks of exhaustion (Gideon *et al.*, 2015;
582 Sakai et al., 2016; Wong *et al.*, 2018). These stem-like T cells warrant additional study, as they
583 associate with control of Mtb in granulomas and, if antigen specific, could be explored as a
584 potential vaccine target.

585

586 Although both CD4 and CD8 T cells have been implicated in control of Mtb infection, the
587 cytotoxic function of lymphocytes in Mtb infection has been relatively understudied, with
588 emphasis placed instead on macrophage activating cytokines, such as IFN-γ and TNF. However,
589 we also find previously unappreciated complexity among granuloma cytotoxic cells of relevance
590 to bacterial control. In accordance with another recent study (Rath *et al.*, 2020), our 6 cytotoxic
591 T/NK subclusters do not align neatly with canonical markers of cellular identity that would define
592 them as classical CD8$\alpha\beta$ or CD4 T cells, NK, NK T cells, or $\gamma\delta$ T cells, but instead appear to be
593 variable mixtures of innate and adaptive cell types with common transcriptional programming. Of
594 these, cytotoxic subcluster 4, which is enriched in CD8$\alpha\beta$ T cells and defined by expression of
595 several granzymes and perforin, likely represents cytotoxic effector T cells that target infected
596 cells for apoptosis and is associated with low burden granulomas. A recent study on lung tissue
597 from Mtb infected macaques also found evidence of cytotoxic molecule expression associated with
598 controlled infection (Esaulova *et al.*, 2021). These findings reveal the importance of cytotoxic
599 innate and adaptive lymphocytes in temporal control of Mtb in granulomas, and a potential role
600 for in future prevention and cure strategies.

601

602 Complementing analyses that characterize individual cell states associated with Mtb control, our
603 cell-cell interaction analyses support connections between control, timing of granuloma
604 appearance, and primed immune responses. Robust control of Mtb at the granuloma level
605 correlated with interactions between particular subsets of T cells and macrophages and was
606 mediated via specific proinflammatory cytokines (e.g., CCL3, IFNγ), T cell chemoattractants (e.g.,
607 CXCL9/10/11/16, CCL20), and co-stimulatory molecules (e.g., CD40LG, CD80, CD86). The pro-
608 inflammatory T cell-mediated signaling in late-appearing, low bacterial burden granulomas stands

609 in contrast to early type II immune and would healing activities in high bacterial burden
610 granulomas, highlighting key cell types and interdependencies behind integrated, holistic host
611 responses.
612
613 Importantly, our analyses reveal not just sets of biological pathways utilized in the host cells of
614 high vs. low burden granulomas, but also assign roles to the specific cell types that drive these
615 signaling patterns. In particular, the strong internal signaling among Group 2 cell types and
616 comparatively weaker cross-talk to other groups in early lesions may drive establishment of a
617 cellular ecosystem dominated by Type II immune and wound healing responses that preclude
618 effective T cell engagement and conversion to a more restrictive state. By comparison, in late-
619 appearing lesions, primed T cell populations, in concert with different innate populations, may use
620 a variety of pro-inflammatory and pro-activation interactions to control Mtb growth or
621 dissemination; a similar phenomenon might explain how infection with Mtb can protect against
622 subsequent reinfection (Lin *et al.*, 2014b), even in the presence of ongoing original infection, by
623 locally recruiting adaptive responses that can act before self-reinforcing Group 2 responses work
624 to limit pathology. Future work will be necessary to determine the relative importance of each
625 adaptive response for control. More broadly, we will need to define the relative stability of these
626 two broad cellular microenvironments and how host perturbations—whether vaccination,
627 therapies or coinfections—impact their balance.
628
629 In addition to identifying cellular populations that relatively exclusively associate with high or low
630 burden granulomas, we find cellular plasticity among some cell populations which appear capable
631 of producing ligands linked to either bacterial persistence or control. For example, the Group 1
632 macrophage populations vary their interaction patterns, perhaps based on the signals they receive

633 from their microenvironment. These responses may, in turn, help mold the phenotypes of
634 surrounding cell types via an immunologic feedback loop (e.g., contributing to persistence of
635 wound healing and Type II immune signaling, or to effective immune recruitment and activation
636 for granuloma clearance). Indeed, we find that individual ligands (IL15, TNFSF13, IL1A and
637 IL1B) can also exhibit pleiotropic effects and participate in interactions enriched in either high or
638 low burden granulomas. Such pleiotropic ligand effects may arise from differing spatial contexts
639 around sender cells (e.g., whether TNFSF13 is secreted from lymphocyte cuff-localized mast cells
640 vs. from macrophage populations closer to the granuloma core), or from combinatorial interactions
641 with other ligands whose presence varies with the distinct microenvironmental ecosystems of high
642 vs. low burden granulomas. These results may help reconcile contrasting findings on ligands' roles
643 differing by disease stage and model (e.g., IL-1), but also inform the selection of targets for
644 therapies seeking to unwind deleterious microenvironments or reinforce adaptive responses
645 (Juffermans *et al.*, 2000; Law *et al.*, 1996; Mayer-Barber *et al.*, 2014).
646
647 Importantly, we note that the contrasting microenvironments revealed through our analyses can
648 occur within the same individual. This suggests the importance of rationally designing new classes
649 of host directed TB preventions and cures that seek to destabilize one set of interactions while
650 reinforcing the other. Indeed, the current standard of care for Mtb calls for multiple antibiotics to
651 be administered for months and has largely remained unchanged for decades (Keshavjee and
652 Farmer, 2012). Our work now defines the complexities of cellular ecosystems encapsulated as
653 granulomas (e.g., reprogramming of plastic tissue-resident cells, recruitment of non-resident
654 immune cells, etc.). Knowledge of intercellular networks underlying granuloma stability will spur
655 future research efforts that identify and manipulate linchpins that serve as key nodes in limiting or
656 enhancing the efficacy of therapeutic and prophylactic measures. For instance, the ligands and

657 receptors implicated in interactions strengthened in high burden granulomas are also enriched for
658 targets of several vitamin A derivatives, including alitretinoin, beta-carotene, and retinol (**Fig**
659 **S8B**), consistent with vitamin A's known activity in promoting wound healing. We note, however,
660 that vitamin A deficiency is a strong risk factor for progression to TB disease (Albana et al., 2017).
661 These observations may be consistent with wound healing responses that create conditions in
662 which bacteria cannot be eradicated but can be contained, speaking to the complexity of
663 intervening in Mtb pathology. In contrast, low burden-linked signaling molecules are also enriched
664 for targets of immunomodulatory drugs used to treat dermatoses and keratoses (e.g.,
665 fludroxycortide, imiquimod) (**Fig S8B**), aligning with a model where successful immune
666 activation circumvents the need for wound healing responses. To most effectively target these
667 complex granuloma ecosystems, we will need new computational methods that can pinpoint the
668 relative importance of different molecular targets and cell types to granuloma stability and
669 determine the most promising points of intervention to destabilize and modulate a densely
670 interacting multicellular community toward adaptive states.
671
672 To fully optimize these host directed therapies, additional work on intercellular communication
673 will be necessary since certain classes of regulatory and effector interactions are not fully captured
674 in this type of analysis. For example, as part of our cell-cell interaction analyses, we found strong
675 enrichment for the expression of distinct neuro-hormonal modulators by Group 2 (e.g., *NRG1,*
676 *RLN3, NTS*) and Group 3 cells (e.g., *NRG2, UCN3*) but did not capture any potential neural
677 interactors in our scRNA-seq dataset, limiting our ability to discern fully how they associate with,
678 and might be leveraged to achieve, control. Nonetheless, ligands and receptors implicated in low-
679 burden interactions are enriched for targets of several neuropsychiatric agents, including
680 spiperone, scopolamine and serotonin, where serotonin reuptake inhibitors have already been

681 identified in screens for host-acting compounds that improve macrophage control of Mtb,
682 supporting potential for their further investigation (Heemskerk et al., 2021; Stanley et al., 2014) .
683 Equally critically, a significant proportion of cell types in Group 3 expressed cytolytic effector
684 genes that can directly drive bacterial control, suggesting a potential therapeutic role for IL15
685 super-agonists in clinical development that can drive expansion of cytotoxic populations.
686
687 It should be noted that granulomas are inherently heterogenous and include necrotic debris,
688 requiring robust technical correction and quality control; this results in an analysis of only high-
689 quality cells. Since only a fraction of cells from each granuloma are analyzed, proportions may not
690 reflect the true composition of cells within a granuloma and may be skewed toward lymphocytes,
691 highlighting the importance of orthogonal validations. In bulk RNA-sequencing analysis of a
692 distinct set of early and late granulomas, we observe generally similar trends in cell-type
693 composition, supporting our conclusions. Relatedly, the transcriptomic granuloma landscape
694 investigated here is from a single (albeit pivotal) time point, while including granulomas across a
695 spectrum of growth trajectories. It is likely that expression of certain genes that arise early in
696 infection and then are downregulated as infection progresses will be missed, as will some
697 populations critical to guiding overall granuloma outcome. More generally, matched analyses of
698 earlier and later time point post-infection, along with analysis of lung tissue and granulomas from
699 vaccinated or reinfected and protected animals, will provide a more complete picture of the
700 temporal control of Mtb in granulomas and is the subject of future work.
701
702 In summary, our data represent the first scRNA-seq investigation of the cellular and molecular
703 features that dynamically associate with natural control of Mtb in pulmonary granulomas. Beyond
704 recapitulating canonical correlates, our analysis defines nuanced actionable innate and adaptive

705 functional cell states, and sheds light on essential dynamics among host-pathogen interactions
706 (Iwasaki and Medzhitov, 2015). Collectively, our data substantiate a model where high Mtb burden
707 within granulomas is dictated at a local level by Type II immune and tissue-protective (wound
708 healing) responses that seek to maintain essential tissue functionality, at the expense of creating a
709 niche for bacterial persistence. In granulomas that form later in infection and therefore in the
710 context of an adaptive immune response, this balance is tipped towards bacterial control by the
711 emergence of adaptive T1-T17 and cytotoxic responses, with interactions involving innate immune
712 cell types enabling sufficient infiltration and activation of these T cell subsets. As a result,
713 successful immune coordination across cell types in late-forming granulomas may obviate the self-
714 reinforcing Type II/wound healing response that would otherwise exclude immune effector
715 functions needed for Mtb control. We also identify cell types and ligands that participate in both
716 high and low burden granulomas, potentially indictive of phenotypic plasticity and pleiotropic
717 effects that may both be molded by and (in turn) reinforce distinct, pathology-associated
718 granuloma microenvironments. Such a framework is consistent with previous observations of
719 natural (Cadena *et al.*, 2018) or induced (Darrah *et al.*, 2020) control, and supports the need to
720 look to new combinatorial host-directed paradigms for the development of novel efficacious
721 therapeutic and prophylactic measures. Moving beyond the perspective of individual molecular
722 targets, our work highlights the importance of the complexities of divergent host cellular
723 ecosystems in driving Mtb persistence or control. By defining and nominating several putative
724 axes of intra- and intercellular signaling associated with contrasting Mtb outcomes, our work
725 provides a foundation for enabling effective manipulation of the properties and states of complex
726 cellular ecosystems, therapeutically-relevant destabilization of pathologic molecular environments
727 to enable adaptive immune access, and fundamental connections to other inflammatory and

728 infectious diseases that affect epithelial barrier tissues (Hughes et al., 2020; Ordovas-Montanes *et*
729 *al.*, 2018).
730
731
732 **Reference**

733 Ahmed, R., Roger, L., Costa Del Amo, P., Miners, K.L., Jones, R.E., Boelen, L., Fali, T.,
734 Elemans, M., Zhang, Y., Appay, V., et al. (2016). Human Stem Cell-like Memory T Cells Are
735 Maintained in a State of Dynamic Flux. Cell Rep *17*, 2811-2818. 10.1016/j.celrep.2016.11.037.
736 Albana, O., Franke, M.F., Huang, C.-C., Galea, J.T., Calderon, R., Zhang, Z., Becerra, M.C.,
737 Smith, E.R., Ronnenberg, A.G., Contreras, C., et al. (2017). Impact of Vitamin A and
738 Carotenoids on the Risk of Tuberculosis Progression. Clinical Infectious Diseases *65*, 900-909.
739 10.1093/cid/cix476.
740 Algood, H.M., Lin, P.L., and Flynn, J.L. (2005). Tumor necrosis factor and chemokine
741 interactions in the formation and maintenance of granulomas in tuberculosis. Clin Infect Dis *41*
742 *Suppl 3*, S189-193. 10.1086/429994.
743 Amezcua Vesely, M.C., Pallis, P., Bielecki, P., Low, J.S., Zhao, J., Harman, C.C.D., Kroehling,
744 L., Jackson, R., Bailis, W., Licona-Limon, P., et al. (2019). Effector TH17 Cells Give Rise to
745 Long-Lived TRM Cells that Are Essential for an Immediate Response against Bacterial
746 Infection. Cell *178*, 1176-1188 e1115. 10.1016/j.cell.2019.07.032.
747 Bagwan, I.N., Khandekar, M.M., Kadam, P., Jadhav, M.V., and Deshmukh, S.D. (2004). A study
748 of mast cells in granulomatous lesions of skin, with special emphasis on leprosy. Indian J Lepr
749 *76*, 31-37.
750 Barone, R., Sotgiu, S., and Musumeci, S. (2007). Plasma chitotriosidase in health and
751 physiology. Clinical Laboratory *53*, 321-333.
752 Caccamo, N., Joosten, S.A., Ottenhoff, T.H.M., and Dieli, F. (2018). Atypical Human
753 Effector/Memory CD4(+) T Cells With a Naive-Like Phenotype. Front Immunol *9*, 2832.
754 10.3389/fimmu.2018.02832.
755 Cadena, A.M., Flynn, J.L., and Fortune, S.M. (2016). The Importance of First Impressions: Early
756 Events in Mycobacterium tuberculosis Infection Influence Outcome. mBio *7*, e00342-00316.
757 10.1128/mBio.00342-16.
758 Cadena, A.M., Hopkins, F.F., Maiello, P., Carey, A.F., Wong, E.A., Martin, C.J., Gideon, H.P.,
759 DiFazio, R.M., Andersen, P., Lin, P.L., et al. (2018). Concurrent infection with Mycobacterium
760 tuberculosis confers robust protection against secondary infection in macaques. PLoS Pathog *14*,
761 e1007305. 10.1371/journal.ppat.1007305.
762 Canetti, G. (1955). The tubercle bacillus in the pulmonary lesion of man: histobacteriology and
763 its bearing on the therapy of pulmonary tuberculosis. Springer.
764 Cartwright, E.K., Palesch, D., Mavigner, M., Paiardini, M., Chahroudi, A., and Silvestri, G.
765 (2016). Initiation of Antiretroviral Therapy Restores CD4+ T Memory Stem Cell Homeostasis in
766 Simian Immunodeficiency Virus-Infected Macaques. J Virol *90*, 6699-6708. 10.1128/JVI.00492-
767 16.
768 Celasun, B., Crow, J., and Scheuer, P.J. (1992). Mast cells in granulomatous liver disease. Pathol
769 Res Pract *188*, 97-100. 10.1016/S0344-0338(11)81163-5.

770 Coleman, M.T., Chen, R.Y., Lee, M., Lin, P.L., Dodd, L.E., Maiello, P., Via, L.E., Kim, Y.,
771 Marriner, G., Dartois, V., et al. (2014a). PET/CT imaging reveals a therapeutic response to
772 oxazolidinones in macaques and humans with tuberculosis. Sci Transl Med *6*, 265ra167.
773 10.1126/scitranslmed.3009500.
774 Coleman, M.T., Maiello, P., Tomko, J., Frye, L.J., Fillmore, D., Janssen, C., Klein, E., and Lin,
775 P.L. (2014b). Early Changes by (18)Fluorodeoxyglucose positron emission tomography
776 coregistered with computed tomography predict outcome after Mycobacterium tuberculosis
777 infection in cynomolgus macaques. Infect Immun *82*, 2400-2404. 10.1128/IAI.01599-13.
778 Collins, A.M., and Jackson, K.J. (2013). A Temporal Model of Human IgE and IgG Antibody
779 Function. Front Immunol *4*, 235. 10.3389/fimmu.2013.00235.
780 Cronan, M.R., Hughes, E.J., Brewer, W.J., Viswanathan, G., Hunt, E.G., Singh, B., Mehra, S.,
781 Oehlers, S.H., Gregory, S.G., Kaushal, D., and Tobin, D.M. (2021). A non-canonical type 2
782 immune response coordinates tuberculous granuloma formation and epithelialization. Cell *184*,
783 1757-1774 e1714. 10.1016/j.cell.2021.02.046.
784 Darrah, P.A., Zeppa, J.J., Maiello, P., Hackney, J.A., Wadsworth, M.H., 2nd, Hughes, T.K.,
785 Pokkali, S., Swanson, P.A., 2nd, Grant, N.L., Rodgers, M.A., et al. (2020). Prevention of
786 tuberculosis in macaques after intravenous BCG immunization. Nature *577*, 95-102.
787 10.1038/s41586-019-1817-8.
788 Diedrich, C.R., Rutledge, T., Maiello, P., Baranowski, T.M., White, A.G., Borish, H.J., Karell,
789 P., Hopkins, F., Brown, J., Fortune, S.M., et al. (2020). SIV and Mycobacterium tuberculosis
790 synergy within the granuloma accelerates the reactivation pattern of latent tuberculosis. bioRxiv.
791 Dijkman, K., Sombroek, C.C., Vervenne, R.A.W., Hofman, S.O., Boot, C., Remarque, E.J.,
792 Kocken, C.H.M., Ottenhoff, T.H.M., Kondova, I., Khayum, M.A., et al. (2019). Prevention of
793 tuberculosis infection and disease by local BCG in repeatedly exposed rhesus macaques. Nat
794 Med *25*, 255-262. 10.1038/s41591-018-0319-9.
795 Ehlers, S., and Schaible, U.E. (2012). The granuloma in tuberculosis: dynamics of a host-
796 pathogen collusion. Front Immunol *3*, 411. 10.3389/fimmu.2012.00411.
797 Esaulova, E., Das, S., Singh, D.K., Choreno-Parra, J.A., Swain, A., Arthur, L., Rangel-Moreno,
798 J., Ahmed, M., Singh, B., Gupta, A., et al. (2021). The immune landscape in tuberculosis reveals
799 populations linked to disease and latency. Cell Host Microbe *29*, 165-178 e168.
800 10.1016/j.chom.2020.11.013.
801 Fabri, M., Stenger, S., Shin, D.-M., Yuk, J.-M., Liu, P.T., Tealegeno, S., Lee, H.-M., Krutzik,
802 S.R., Schenk, M., Sieling, P.A., et al. (2011). Vitamin D is required for IFN-gamma-mediated
803 antimicrobial activity of human macrophages. Science Translational Medicine *3*.
804 10.1126/scitranslmed.3003045.
805 Flynn, J.L. (2006). Lessons from experimental Mycobacterium tuberculosis infections. Microbes
806 Infect *8*, 1179-1188. 10.1016/j.micinf.2005.10.033.
807 Flynn, J.L., Capuano, S.V., Croix, D., Pawar, S., Myers, A., Zinovik, A., and Klein, E. (2003).
808 Non-human primates: a model for tuberculosis research. Tuberculosis (Edinb) *83*, 116-118.
809 10.1016/s1472-9792(02)00059-8.
810 Flynn, J.L., Klein, E., (2010). Pulmonary Tuberculosis in Monkeys. A Color Atlas of
811 Comparative Pathology of Pulmonary Tuberculosis, 83-105.
812 Foreman, T.W., Mehra, S., LoBato, D.N., Malek, A., Alvarez, X., Golden, N.A., Bucsan, A.N.,
813 Didier, P.J., Doyle-Meyers, L.A., Russell-Lodrigue, K.E., et al. (2016). CD4+ T-cell-
814 independent mechanisms suppress reactivation of latent tuberculosis in a macaque model of HIV
815 coinfection. Proc Natl Acad Sci U S A *113*, E5636-5644. 10.1073/pnas.1611987113.

816 Fuertes Marraco, S.A., Soneson, C., Delorenzi, M., and Speiser, D.E. (2015). Genome-wide
817 RNA profiling of long-lasting stem cell-like memory CD8 T cells induced by Yellow Fever
818 vaccination in humans. Genom Data *5*, 297-301. 10.1016/j.gdata.2015.06.024.
819 Garcia-Rodriguez, K.M., Goenka, A., Alonso-Rasgado, M.T., Hernandez-Pando, R., and
820 Bulfone-Paus, S. (2017). The Role of Mast Cells in Tuberculosis: Orchestrating Innate Immune
821 Crosstalk? Front Immunol *8*, 1290. 10.3389/fimmu.2017.01290.
822 Gattinoni, L., Lugli, E., Ji, Y., Pos, Z., Paulos, C.M., Quigley, M.F., Almeida, J.R., Gostick, E.,
823 Yu, Z., Carpenito, C., et al. (2011). A human memory T cell subset with stem cell-like
824 properties. Nat Med *17*, 1290-1297. 10.1038/nm.2446.
825 Gideon, H.P., Phuah, J., Junecko, B.A., and Mattila, J.T. (2019). Neutrophils express pro- and
826 anti-inflammatory cytokines in granulomas from Mycobacterium tuberculosis-infected
827 cynomolgus macaques. Mucosal Immunol *12*, 1370-1381. 10.1038/s41385-019-0195-8.
828 Gideon, H.P., Phuah, J., Myers, A.J., Bryson, B.D., Rodgers, M.A., Coleman, M.T., Maiello, P.,
829 Rutledge, T., Marino, S., Fortune, S.M., et al. (2015). Variability in tuberculosis granuloma T
830 cell responses exists, but a balance of pro- and anti-inflammatory cytokines is associated with
831 sterilization. PLoS Pathog *11*, e1004603. 10.1371/journal.ppat.1004603.
832 Gierahn, T.M., Wadsworth, M.H., 2nd, Hughes, T.K., Bryson, B.D., Butler, A., Satija, R.,
833 Fortune, S., Love, J.C., and Shalek, A.K. (2017). Seq-Well: portable, low-cost RNA sequencing
834 of single cells at high throughput. Nat Methods *14*, 395-398. 10.1038/nmeth.4179.
835 Green, A.M., Difazio, R., and Flynn, J.L. (2013). IFN-gamma from CD4 T cells is essential for
836 host survival and enhances CD8 T cell function during Mycobacterium tuberculosis infection. J
837 Immunol *190*, 270-277. 10.4049/jimmunol.1200061.
838 Guo, X., Zhang, Y., Zheng, L., Zheng, C., Song, J., Zhang, Q., Kang, B., Liu, Z., Jin, L., Xing,
839 R., et al. (2018). Global characterization of T cells in non-small-cell lung cancer by single-cell
840 sequencing. Nat Med *24*, 978-985. 10.1038/s41591-018-0045-3.
841 Han, X., Wang, R., Zhou, Y., Fei, L., Sun, H., Lai, S., Saadatpour, A., Zhou, Z., Chen, H., Ye,
842 F., et al. (2018). Mapping the Mouse Cell Atlas by Microwell-Seq. Cell *173*, 1307.
843 10.1016/j.cell.2018.05.012.
844 Heemskerk, M.T., Korbee, C.J., Esselink, J.J., Carvalho Dos Santo, C., van Veen, S., Gordijn,
845 I.F., Vrieling, F., Engele, C.G., Dijkman, K., Wilson, L., et al. (2021). Repurposing
846 diphenylbutylpiperidine-class antipsychotic drugs for host-directed therapy of Mycobacterium
847 tuberculosis and Salmonella enterica infections. Scientific Reports *11*. 10.1038/s41598-021-
848 98980-z.
849 Hirota, K., Yoshitomi, H., Hashimoto, M., Maeda, S., Teradaira, S., Sugimoto, N., Yamaguchi,
850 T., Nomura, T., Ito, H., Nakamura, T., et al. (2007). Preferential recruitment of CCR6-expressing
851 Th17 cells to inflamed joints via CCL20 in rheumatoid arthritis and its animal model. J Exp Med
852 *204*, 2803-2812. 10.1084/jem.20071397.
853 Hu, S., Lv, P., Yan, Z., and Wen, B. (2019). Disruption of nuclear speckles reduces chromatin
854 interactions in active compartments. Epigenetics Chromatin *12*, 43. 10.1186/s13072-019-0289-2.
855 Hughes, T.K., Wadsworth, M.H., 2nd, Gierahn, T.M., Do, T., Weiss, D., Andrade, P.R., Ma, F.,
856 de Andrade Silva, B.J., Shao, S., Tsoi, L.C., et al. (2020). Second-Strand Synthesis-Based
857 Massively Parallel scRNA-Seq Reveals Cellular States and Molecular Features of Human
858 Inflammatory Skin Pathologies. Immunity *53*, 878-894 e877. 10.1016/j.immuni.2020.09.015.
859 Hunter, R.L. (2011). Pathology of post primary tuberculosis of the lung: an illustrated critical
860 review. Tuberculosis (Edinb) *91*, 497-509. 10.1016/j.tube.2011.03.007.
861 Hunter, R.L. (2016). Tuberculosis as a three-act play: A new paradigm for the pathogenesis of
862 pulmonary tuberculosis. Tuberculosis (Edinb) *97*, 8-17. 10.1016/j.tube.2015.11.010.

863 Huynh, J.P., Lin, C.C., Kimmey, J.M., Jarjour, N.N., Schwarzkopf, E.A., Bradstreet, T.R.,
864 Shchukina, I., Shpynov, O., Weaver, C.T., Taneja, R., et al. (2018). Bhlhe40 is an essential
865 repressor of IL-10 during Mycobacterium tuberculosis infection. J Exp Med *215*, 1823-1838.
866 10.1084/jem.20171704.
867 Ivanov, II, McKenzie, B.S., Zhou, L., Tadokoro, C.E., Lepelley, A., Lafaille, J.J., Cua, D.J., and
868 Littman, D.R. (2006). The orphan nuclear receptor RORgammat directs the differentiation
869 program of proinflammatory IL-17+ T helper cells. Cell *126*, 1121-1133.
870 10.1016/j.cell.2006.07.035.
871 Iwasaki, A., and Medzhitov, R. (2015). Control of adaptive immunity by the innate immune
872 system. Nat Immunol *16*, 343-353. 10.1038/ni.3123.
873 Jacobs, A.J., Mongkolsapaya, J., Screaton, G.R., McShane, H., and Wilkinson, R.J. (2016).
874 Antibodies and tuberculosis. Tuberculosis (Edinb) *101*, 102-113. 10.1016/j.tube.2016.08.001.
875 Joshi, N., Watanabe, S., Verma, R., Jablonski, R.P., Chen, C.Y., Cheresh, P., Markov, N.S.,
876 Reyfman, P.A., McQuattie-Pimentel, A.C., Sichizya, L., et al. (2020). A spatially restricted
877 fibrotic niche in pulmonary fibrosis is sustained by M-CSF/M-CSFR signalling in monocyte-
878 derived alveolar macrophages. European Respiratory Journal *55*. 10.1183/13993003.00646-
879 2019.
880 Juffermans, N.P., Florquin, S., Camoglio, L., Verbon, A., Kolk, A.H., Speelman, P., van
881 Deventer, S.J.H., and van der Poll, T. (2000). Interleukin-1 Signaling Is Essential for Host
882 Defense during Murine Pulmonary Tuberculosis. The Journal of Infectious Diseases *182*, 902-
883 908. doi.org/10.1086/315771.
884 Kanagaratham, C., El Ansari, Y.S., Lewis, O.L., and Oettgen, H.C. (2020). IgE and IgG
885 Antibodies as Regulators of Mast Cell and Basophil Functions in Food Allergy. Front Immunol
886 *11*, 603050. 10.3389/fimmu.2020.603050.
887 Keshavjee, S., and Farmer, P.E. (2012). Tuberculosis, Drug Resistance, and the History of
888 Modern Medicine. New England Journal of Medicine *367*, 931-936. 10.1056/NEJMra1205429.
889 Khader, S.A., Bell, G.K., Pearl, J.E., Fountain, J.J., Rangel-Moreno, J., Cilley, G.E., Shen, F.,
890 Eaton, S.M., Gaffen, S.L., Swain, S.L., et al. (2007). IL-23 and IL-17 in the establishment of
891 protective pulmonary CD4+ T cell responses after vaccination and during Mycobacterium
892 tuberculosis challenge. Nat Immunol *8*, 369-377. 10.1038/ni1449.
893 Khader, S.A., and Gopal, R. (2010). IL-17 in protective immunity to intracellular pathogens.
894 Virulence *1*, 423-427. 10.4161/viru.1.5.12862.
895 Kienzle, N., Olver, S., Buttigieg, K., Groves, P., Janas, M.L., Baz, A., and Kelso, A. (2005).
896 Progressive differentiation and commitment of CD8+ T cells to a poorly cytolytic CD8low
897 phenotype in the presence of IL-4. J Immunol *174*, 2021-2029. 10.4049/jimmunol.174.4.2021.
898 Kobayashi, T., Okamoto, S., Hisamatsu, T., Kamada, N., Chinen, H., Saito, R., Kitazume, M.T.,
899 Nakazawa, A., Sugita, A., Koganei, K., et al. (2008). IL23 differentially regulates the Th1/Th17
900 balance in ulcerative colitis and Crohn's disease. Gut *57*, 1682-1689. 10.1136/gut.2007.135053.
901 Law, K., Weiden, M., Harkin, T., Tchou-Wong, K., Chi, C., and Rom, W.N. (1996). Increased
902 release of interleukin-1 beta, interleukin-6, and tumor necrosis factor-alpha by bronchoalveolar
903 cells lavaged from involved sites in pulmonary tuberculosis. American Journal of Respiratory
904 and Critical Care Medicine *153*, 799-804. 10.1164/ajrccm.153.2.8564135.
905 Lenaerts, A., Barry, C.E., 3rd, and Dartois, V. (2015). Heterogeneity in tuberculosis pathology,
906 microenvironments and therapeutic responses. Immunol Rev *264*, 288-307. 10.1111/imr.12252.
907 Li, W., Laumonnier, Y., Syrovets, T., and Simmet, T. (2013). Recruitment of CCR6-expressing
908 Th17 cells by CCL20 secreted from plasmin-stimulated macrophages. Acta Biochim Biophys
909 Sin *45*, 593-600. 10.1093/abbs/gmt049.

910 Liang, Y., Pan, H.F., and Ye, D.Q. (2015). Tc17 Cells in Immunity and Systemic Autoimmunity.
911 Int Rev Immunol *34*, 318-331. 10.3109/08830185.2014.954698.
912 Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdottir, H., Tamayo, P., and Mesirov,
913 J.P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics *27*, 1739-1740.
914 10.1093/bioinformatics/btr260.
915 Lim, H.W., Lee, J., Hillsamer, P., and Kim, C.H. (2008). Human Th17 Cells Share Major
916 Trafficking Receptors with Both Polarized Effector T Cells and FOXP3+ Regulatory T Cells.
917 The Journal of Immunology *180*, 122-129. 10.4049/jimmunol.180.1.122.
918 Lin, C.C., Bradstreet, T.R., Schwarzkopf, E.A., Jarjour, N.N., Chou, C., Archambault, A.S., Sim,
919 J., Zinselmeyer, B.H., Carrero, J.A., Wu, G.F., et al. (2016). IL-1-induced Bhlhe40 identifies
920 pathogenic T helper cells in a model of autoimmune neuroinflammation. J Exp Med *213*, 251-
921 271. 10.1084/jem.20150568.
922 Lin, C.C., Bradstreet, T.R., Schwarzkopf, E.A., Sim, J., Carrero, J.A., Chou, C., Cook, L.E.,
923 Egawa, T., Taneja, R., Murphy, T.L., et al. (2014a). Bhlhe40 controls cytokine production by T
924 cells and is essential for pathogenicity in autoimmune neuroinflammation. Nat Commun *5*, 3551.
925 10.1038/ncomms4551.
926 Lin, P.L., Coleman, T., Carney, J.P., Lopresti, B.J., Tomko, J., Fillmore, D., Dartois, V., Scanga,
927 C., Frye, L.J., Janssen, C., et al. (2013). Radiologic Responses in Cynomolgus Macaques for
928 Assessing Tuberculosis Chemotherapy Regimens. Antimicrob Agents Chemother *57*, 4237-
929 4244. 10.1128/AAC.00277-13.
930 Lin, P.L., Ford, C.B., Coleman, M.T., Myers, A.J., Gawande, R., Ioerger, T., Sacchettini, J.,
931 Fortune, S.M., and Flynn, J.L. (2014b). Sterilization of granulomas is common in active and
932 latent tuberculosis despite within-host variability in bacterial killing. Nat Med *20*, 75-79.
933 10.1038/nm.3412.
934 Lin, P.L., Myers, A., Smith, L., Bigbee, C., Bigbee, M., Fuhrman, C., Grieser, H., Chiosea, I.,
935 Voitenek, N.N., Capuano, S.V., et al. (2010). Tumor necrosis factor neutralization results in
936 disseminated disease in acute and latent Mycobacterium tuberculosis infection with normal
937 granuloma structure in a cynomolgus macaque model. Arthritis Rheum *62*, 340-350.
938 10.1002/art.27271.
939 Lin, P.L., Pawar, S., Myers, A., Pegu, A., Fuhrman, C., Reinhart, T.A., Capuano, S.V., Klein, E.,
940 and Flynn, J.L. (2006). Early events in Mycobacterium tuberculosis infection in cynomolgus
941 macaques. Infect Immun *74*, 3790-3803. 10.1128/IAI.00064-06.
942 Lin, P.L., Plessner, H.L., Voitenok, N.N., and Flynn, J.L. (2007). Tumor necrosis factor and
943 tuberculosis. J Investig Dermatol Symp Proc *12*, 22-25. 10.1038/sj.jidsymp.5650027.
944 Lin, P.L., Rodgers, M., Smith, L., Bigbee, M., Myers, A., Bigbee, C., Chiosea, I., Capuano, S.V.,
945 Fuhrman, C., Klein, E., and Flynn, J.L. (2009). Quantitative comparison of active and latent
946 tuberculosis in the cynomolgus macaque model. Infect Immun *77*, 4631-4642.
947 10.1128/IAI.00592-09.
948 Lin, P.L., Rutledge, T., Green, A.M., Bigbee, M., Fuhrman, C., Klein, E., and Flynn, J.L. (2012).
949 CD4 T cell depletion exacerbates acute Mycobacterium tuberculosis while reactivation of latent
950 infection is dependent on severity of tissue depletion in cynomolgus macaques. AIDS research
951 and human retroviruses *28*, 1693-1702. 10.1089/AID.2012.0028.
952 Liu, P.T., Stenger, S., Li, H., Wenzel, L., Tan, B.H., Krutzik, S.R., Ochoa, M.T., Schauber, J.,
953 Wu, K., Meinken, C., et al. (2006). Toll-like receptor triggering of a vitamin D-mediated human
954 antimicrobial response. Science *311*, 1770-1773. 10.1126/science.1123933.
955 Lopez, D., Montoya, D., Ambrose, M., Lam, L., Briscoe, L., Adams, C., Modlin, R.L., and
956 Pellegrini, M. (2017). SaVanT: a web-based tool for the sample-level visualization of molecular
957 signatures in gene expression profiles. BMC Genomics *18*, 824. 10.1186/s12864-017-4167-7.

958 Lun, A.T.L., Riesenfeld, S., Andrews, T., Dao, T.P., Gomes, T., participants in the 1st Human
959 Cell Atlas, J., and Marioni, J.C. (2019). EmptyDrops: distinguishing cells from empty droplets in
960 droplet-based single-cell RNA sequencing data. Genome Biol *20*, 63. 10.1186/s13059-019-1662-
961 y.
962 Lyadova, I.V., and Panteleev, A.V. (2015). Th1 and Th17 Cells in Tuberculosis: Protection,
963 Pathology, and Biomarkers. Mediators Inflamm *2015*, 854507. 10.1155/2015/854507.
964 Macosko, E.Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas,
965 A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression
966 Profiling of Individual Cells Using Nanoliter Droplets. Cell *161*, 1202-1214.
967 10.1016/j.cell.2015.05.002.
968 Maiello, P., DiFazio, R.M., Cadena, A.M., Rodgers, M.A., Lin, P.L., Scanga, C.A., and Flynn,
969 J.L. (2018). Rhesus Macaques Are More Susceptible to Progressive Tuberculosis than
970 Cynomolgus Macaques: a Quantitative Comparison. Infect Immun *86*. 10.1128/IAI.00505-17.
971 Malherbe, S.T., Shenai, S., Ronacher, K., Loxton, A.G., Dolganov, G., Kriel, M., Van, T., Chen,
972 R.Y., Warwick, J., Via, L.E., et al. (2016). Persisting positron emission tomography lesion
973 activity and Mycobacterium tuberculosis mRNA after tuberculosis cure. Nat Med *22*, 1094-1100.
974 10.1038/nm.4177.
975 Martin, C.J., Cadena, A.M., Leung, V.W., Lin, P.L., Maiello, P., Hicks, N., Chase, M.R., Flynn,
976 J.L., and Fortune, S.M. (2017). Digitally Barcoding Mycobacterium tuberculosis Reveals In
977 Vivo Infection Dynamics in the Macaque Model of Tuberculosis. mBio *8*. 10.1128/mBio.00312-
978 17.
979 Mateus, J., Lasso, P., Pavia, P., Rosas, F., Roa, N., Valencia-Hernandez, C.A., Gonzalez, J.M.,
980 Puerta, C.J., and Cuellar, A. (2015). Low frequency of circulating CD8+ T stem cell memory
981 cells in chronic chagasic patients with severe forms of the disease. PLoS Negl Trop Dis *9*, e3432.
982 10.1371/journal.pntd.0003432.
983 Mattila, J.T., Diedrich, C.R., Lin, P.L., Phuah, J., and Flynn, J.L. (2011). Simian
984 immunodeficiency virus-induced changes in T cell cytokine responses in cynomolgus macaques
985 with latent Mycobacterium tuberculosis infection are associated with timing of reactivation. J
986 Immunol *186*, 3527-3537. 10.4049/jimmunol.1003773.
987 Mattila, J.T., Maiello, P., Sun, T., Via, L.E., and Flynn, J.L. (2015). Granzyme B-expressing
988 neutrophils correlate with bacterial load in granulomas from Mycobacterium tuberculosis-
989 infected cynomolgus macaques. Cell Microbiol *17*, 1085-1097. 10.1111/cmi.12428.
990 Mattila, J.T., Ojo, O.O., Kepka-Lenhart, D., Marino, S., Kim, J.H., Eum, S.Y., Via, L.E., Barry,
991 C.E., 3rd, Klein, E., Kirschner, D.E., et al. (2013). Microenvironments in tuberculous
992 granulomas are delineated by distinct populations of macrophage subsets and expression of nitric
993 oxide synthase and arginase isoforms. J Immunol *191*, 773-784. 10.4049/jimmunol.1300113.
994 Mayer-Barber, K.D., Andrade, B.B., Oland, S.D., Amaral, E.P., Barber, D.L., Gonzales, J.,
995 Derrick, S.C., Shi, R., Kumar, N.P., Wei, W., et al. (2014). Host-directed therapy of tuberculosis
996 based on interleukin-1 and type I interferon crosstalk. Nature *511*, 99-103.
997 McCaffrey, E.F., Donato, M., Keren, L., Chen, Z., Fitzpatrick, M., Jojic, V., Delmastro, A.,
998 Greenwald, N.F., Baranski, A., Graf, W., et al. (2020). Multiplexed imaging of human
999 tuberculosis granulomas uncovers immunoregulatory features conserved across tissue and blood.
1000 bioRxiv, 2020.2006.2008.140426. 10.1101/2020.06.08.140426.
1001 McGinnis, C.S., Murrow, L.M., and Gartner, Z.J. (2019). DoubletFinder: Doublet Detection in
1002 Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. Cell Syst *8*, 329-337
1003 e324. 10.1016/j.cels.2019.03.003.

1004 Meyer Zu Horste, G., Wu, C., Wang, C., Cong, L., Pawlak, M., Lee, Y., Elyaman, W., Xiao, S.,
1005 Regev, A., and Kuchroo, V.K. (2016). RBPJ Controls Development of Pathogenic Th17 Cells by
1006 Regulating IL-23 Receptor Expression. Cell Rep *16*, 392-404. 10.1016/j.celrep.2016.05.088.
1007 Millington, K.A., Innes, J.A., Hackforth, S., Hinks, T.S., Deeks, J.J., Dosanjh, D.P., Guyot-
1008 Revol, V., Gunatheesan, R., Klenerman, P., and Lalvani, A. (2007). Dynamic relationship
1009 between IFN-gamma and IL-2 profile of Mycobacterium tuberculosis-specific T cells and
1010 antigen load. J Immunol *178*, 5217-5226. 10.4049/jimmunol.178.8.5217.
1011 Mogilenko, D.A., Shpynov, O., Andhey, P.S., Arthur, L., Swain, A., Esaulova, E., Brioschi, S.,
1012 Shchukina, I., Kerndl, M., Bambouskova, M., et al. (2021). Comprehensive Profiling of an
1013 Aging Immune System Reveals Clonal GZMK(+) CD8(+) T Cells as Conserved Hallmark of
1014 Inflammaging. Immunity *54*, 99-115 e112. 10.1016/j.immuni.2020.11.005.
1015 Mpande, C.A.M., Dintwe, O.B., Musvosvi, M., Mabwe, S., Bilek, N., Hatherill, M., Nemes, E.,
1016 Scriba, T.J., and Team, S.C.I. (2018). Functional, Antigen-Specific Stem Cell Memory (TSCM)
1017 CD4(+) T Cells Are Induced by Human Mycobacterium tuberculosis Infection. Front Immunol
1018 *9*, 324. 10.3389/fimmu.2018.00324.
1019 Munoz-Elias, E.J., Timm, J., Botha, T., Chan, W.T., Gomez, J.E., and McKinney, J.D. (2005).
1020 Replication dynamics of Mycobacterium tuberculosis in chronically infected mice. Infect Immun
1021 *73*, 546-551. 10.1128/IAI.73.1.546-551.2005.
1022 Nakano, H., Lyons-Cohen, M., Nakano, K., Karcz, T., Thomas, S.Y., Whitehead, G., and Cook,
1023 D.N. (2019). Type 1 alveolar epithelial cell-derived chemokine CCL19 promotes airway
1024 inflammation. The Journal of Immunology *202*.
1025 Newman, A.M., Liu, C.L., Green, M.R., Gentles, A.J., Feng, W., Xu, Y., Hoang, C.D., Diehn,
1026 M., and Alizadeh, A.A. (2015). Robust enumeration of cell subsets from tissue expression
1027 profiles. Nat Methods *12*, 453-457. 10.1038/nmeth.3337.
1028 O'Garra, A., Redford, P.S., McNab, F.W., Bloom, C.I., Wilkinson, R.J., and Berry, M.P. (2013).
1029 The immune response in tuberculosis. Annu Rev Immunol *31*, 475-527. 10.1146/annurev-
1030 immunol-032712-095939.
1031 Ohtani, H. (2013). Granuloma cells in chronic inflammation express CD205 (DEC205) antigen
1032 and harbor proliferating T lymphocytes: similarity to antigen-presenting cells. Pathol Int *63*, 85-
1033 93. 10.1111/pin.12036.
1034 Ordovas-Montanes, J., Dwyer, D.F., Nyquist, S.K., Buchheit, K.M., Vukovic, M., Deb, C.,
1035 Wadsworth, M.H., 2nd, Hughes, T.K., Kazer, S.W., Yoshimoto, E., et al. (2018). Allergic
1036 inflammatory memory in human respiratory epithelial progenitor cells. Nature *560*, 649-654.
1037 10.1038/s41586-018-0449-8.
1038 Padela, S., Yi, M., Cabacungan, J., Shek, S., Belcastro, R., Masood, A., Jankov, R.P., and
1039 Tanswell, A.K. (2008). A Critical Role for Fibroblast Growth Factor-7 during Early Alveolar
1040 Formation in the Neonatal Rat. Pediatric Research *63*, 232-238.
1041 Pagan, A.J., and Ramakrishnan, L. (2014). Immunity and Immunopathology in the Tuberculous
1042 Granuloma. Cold Spring Harb Perspect Med *5*. 10.1101/cshperspect.a018499.
1043 Phuah, J., Wong, E.A., Gideon, H.P., Maiello, P., Coleman, M.T., Hendricks, M.R., Ruden, R.,
1044 Cirrincione, L.R., Chan, J., Lin, P.L., and Flynn, J.L. (2016). Effects of B Cell Depletion on
1045 Early Mycobacterium tuberculosis Infection in Cynomolgus Macaques. Infect Immun *84*, 1301-
1046 1311. 10.1128/IAI.00083-16.
1047 Phuah, J.Y., Mattila, J.T., Lin, P.L., and Flynn, J.L. (2012). Activated B cells in the granulomas
1048 of nonhuman primates infected with Mycobacterium tuberculosis. Am J Pathol *181*, 508-514.
1049 10.1016/j.ajpath.2012.05.009.
1050 Prakadan, S.M., Shalek, A.K., and Weitz, D.A. (2017). Scaling by shrinking: empowering
1051 single-cell 'omics' with microfluidic devices. Nat Rev Genet *18*, 345-361. 10.1038/nrg.2017.15.

1052 Raphael, I., Nalawade, S., Eagar, T.N., and Forsthuber, T.G. (2015). T cell subsets and their
1053 signature cytokines in autoimmune and inflammatory diseases. Cytokine *74*, 5-17.
1054 10.1016/j.cyto.2014.09.011.
1055 Rath, J.A., Bajwa, G., Carreres, B., Hoyer, E., Gruber, I., Martinez-Paniagua, M.A., Yu, Y.R.,
1056 Nouraee, N., Sadeghi, F., Wu, M., et al. (2020). Single-cell transcriptomics identifies multiple
1057 pathways underlying antitumor function of TCR- and CD8alphabeta-engineered human CD4(+)
1058 T cells. Sci Adv *6*, eaaz7809. 10.1126/sciadv.aaz7809.
1059 Reece, S.T., and Kaufmann, S.H. (2012). Floating between the poles of pathology and
1060 protection: can we pin down the granuloma in tuberculosis? Curr Opin Microbiol *15*, 63-70.
1061 10.1016/j.mib.2011.10.006.
1062 Russell, D.G., Barry, C.E., 3rd, and Flynn, J.L. (2010). Tuberculosis: what we don't know can,
1063 and does, hurt us. Science *328*, 852-856. 10.1126/science.1184784.
1064 Sakai, S., Kauffman, K.D., Sallin, M.A., Sharpe, A.H., Young, H.A., Ganusov, V.V., and
1065 Barber, D.L. (2016). CD4 T Cell-Derived IFN-gamma Plays a Minimal Role in Control of
1066 Pulmonary Mycobacterium tuberculosis Infection and Must Be Actively Repressed by PD-1 to
1067 Prevent Lethal Disease. PLoS Pathog *12*, e1005667. 10.1371/journal.ppat.1005667.
1068 Schietinger, A., Delrow, J.J., Basom, R.S., Blattman, J.N., and Greenberg, P.D. (2012). Rescued
1069 tolerant CD8 T cells are preprogrammed to reestablish the tolerant state. Science *335*, 723-727.
1070 10.1126/science.1214277.
1071 Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch,
1072 S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for
1073 biological-image analysis. Nat Methods *9*, 676-682. 10.1038/nmeth.2019.
1074 Scriba, T.J., Coussens, A.K., and Fletcher, H.A. (2017). Human Immunology of Tuberculosis.
1075 Microbiol Spectr *5*. 10.1128/microbiolspec.TBTB2-0016-2016.
1076 Shanmugasundaram, U., Bucsan, A.N., Ganatra, S.R., Ibegbu, C., Quezada, M., Blair, R.V.,
1077 Alvarez, X., Velu, V., Kaushal, D., and Rengarajan, J. (2020). Pulmonary Mycobacterium
1078 tuberculosis control associates with CXCR3- and CCR6-expressing antigen-specific Th1 and
1079 Th17 cell recruitment. JCI Insight *5*. 10.1172/jci.insight.137858.
1080 Shehadeh, L.A., Yu, K., Wang, L., Guevara, A., Singer, C., Vance, J., and Papapetropoulos, S.
1081 (2010). SRRM2, a potential blood biomarker revealing high alternative splicing in Parkinson's
1082 disease. PLoS One *5*, e9104. 10.1371/journal.pone.0009104.
1083 Stanley, S.A., Barczak, A.K., Silvis, M.R., Luo, S.S., Sogi, K., Vokes, M., Bray, M.-A.,
1084 Carpenter, A.E., Moore, C.B., Siddiqi, N., et al. (2014). Identification of Host-Targeted Small
1085 Molecules That Restrict Intracellular Mycobacterium tuberculosis Growth. PLos Pathogens *10*.
1086 10.1371/journal.ppat.1003946.
1087 Szabo, P.A., Levitin, H.M., Miron, M., Snyder, M.E., Senda, T., Yuan, J., Cheng, Y.L., Bush,
1088 E.C., Dogra, P., Thapa, P., et al. (2019). Single-cell transcriptomics of human T cells reveals
1089 tissue and activation signatures in health and disease. Nat Commun *10*, 4706. 10.1038/s41467-
1090 019-12464-3.
1091 Tabula Muris, C., Overall, c., Logistical, c., Organ, c., processing, Library, p., sequencing,
1092 Computational data, a., Cell type, a., Writing, g., et al. (2018). Single-cell transcriptomics of 20
1093 mouse organs creates a Tabula Muris. Nature *562*, 367-372. 10.1038/s41586-018-0590-4.
1094 Tateosian, N., Pasquinelli, V., del Pino, R.H., Garcia, V., and Chuluyan, E. (2012). The negative
1095 feedback mechanism between SLPI and IFNγ is impaired in tuberculosis patients. The Journal of
1096 Immunology *188*.
1097 Taweevisit, M., and Poumsuk, U. (2007). High mast cell density associated with granulomatous
1098 formation in tuberculous lymphadenitis. Southeast Asian J Trop Med Public Health *38*, 115-119.

1099 Tobin, D.M., Roca, F.J., Oh, S.F., McFarland, R., Vickery, T.W., Ray, J.P., Ko, D.C., Zou, Y.,
1100 Bang, N.D., Chau, T.T.H., et al. (2012). Host genotype-specific therapies can optimize the
1101 inflammatory response to mycobacterial infections. Cell *148*, 434-446.
1102 10.1016/j.cell.2011.12.023.
1103 Tobin, D.M., Vary, J.C., Ray, J.P., Walsh, G.S., Dunstan, S.J., Bang, N.D., Hagge, D.A.,
1104 Khadge, S., King, M.-C., Hawn, T.R., et al. (2010). The lta4h Locus Modulates Susceptibility to
1105 Mycobacterial Infection in Zebrafish and Humans. Cell *140*, 717-730.
1106 10.1016/j.cell.2010.02.013.
1107 Todryk, S.M. (2018). T Cell Memory to Vaccination. Vaccines (Basel) *6*.
1108 10.3390/vaccines6040084.
1109 Touzot, M., Grandclaudon, M., Cappuccio, A., Satoh, T., Martinez-Cingolani, C., Servant, N.,
1110 Manel, N., and Soumelis, V. (2014). Combinatorial flexibility of cytokine function during human
1111 T helper cell differentiation. Nature Communications *5:3987*. 10.1038/ncomms4987.
1112 Ulrichs, T., and Kaufmann, S.H. (2006). New insights into the function of granulomas in human
1113 tuberculosis. J Pathol *208*, 261-269. 10.1002/path.1906.
1114 van Hamburg, J.P., and Tas, S.W. (2018). Molecular mechanisms underpinning T helper 17 cell
1115 heterogeneity and functions in rheumatoid arthritis. J Autoimmun *87*, 69-81.
1116 10.1016/j.jaut.2017.12.006.
1117 Varemo, L., Nielsen, J., and Nookaew, I. (2013). Enriching the gene set analysis of genome-wide
1118 data by incorporating directionality of gene expression and combining statistical hypotheses and
1119 methods. Nucleic Acids Res *41*, 4378-4391. 10.1093/nar/gkt111.
1120 Wacleche, V.S., Goulet, J.P., Gosselin, A., Monteiro, P., Soudeyns, H., Fromentin, R., Jenabian,
1121 M.A., Vartanian, S., Deeks, S.G., Chomont, N., et al. (2016). New insights into the heterogeneity
1122 of Th17 subsets contributing to HIV-1 persistence during antiretroviral therapy. Retrovirology
1123 *13*, 59. 10.1186/s12977-016-0293-6.
1124 White, A.G., Maiello, P., Coleman, M.T., Tomko, J.A., Frye, L.J., Scanga, C.A., Lin, P.L., and
1125 Flynn, J.L. (2017). Analysis of 18FDG PET/CT Imaging as a Tool for Studying Mycobacterium
1126 tuberculosis Infection and Treatment in Non-human Primates. J Vis Exp. 10.3791/56375.
1127 WHO (2019). Global Tuberculosis Report.
1128 Wijesundara, D.K., Tscharke, D.C., Jackson, R.J., and Ranasinghe, C. (2013). Reduced
1129 interleukin-4 receptor alpha expression on CD8+ T cells correlates with higher quality anti-viral
1130 immunity. PLoS One *8*, e55788. 10.1371/journal.pone.0055788.
1131 Wolf, F.A., Angerer, P., and Theis, F.J. (2018). SCANPY: large-scale single-cell gene
1132 expression data analysis. Genome Biol *19*, 15. 10.1186/s13059-017-1382-0.
1133 Wong, E.A., Evans, S., Kraus, C.R., Engelman, K.D., Maiello, P., Flores, W.J., Cadena, A.M.,
1134 Klein, E., Thomas, K., White, A.G., et al. (2020). IL-10 Impairs Local Immune Response in
1135 Lung Granulomas and Lymph Nodes during Early Mycobacterium tuberculosis Infection. J
1136 Immunol *204*, 644-659. 10.4049/jimmunol.1901211.
1137 Wong, E.A., Joslyn, L., Grant, N.L., Klein, E., Lin, P.L., Kirschner, D.E., and Flynn, J.L. (2018).
1138 Low Levels of T Cell Exhaustion in Tuberculous Lung Granulomas. Infect Immun *86*.
1139 10.1128/IAI.00426-18.
1140 Wu, C., Pot, C., Apetoh, L., Thalhamer, T., Zhu, B., Murugaiyan, G., Xiao, S., Lee, Y.,
1141 Rangachari, M., Yosef, N., and Kuchroo, V.K. (2013). Metallothioneins negatively regulate IL-
1142 27-induced type 1 regulatory T-cell differentiation. Proc Natl Acad Sci U S A *110*, 7802-7807.
1143 10.1073/pnas.1211776110.
1144 Yang, X.O., Nurieva, R., Martinez, G.J., Kang, H.S., Chung, Y., Pappu, B.P., Shah, B., Chang,
1145 S.H., Schluns, K.S., Watowich, S.S., et al. (2008). Molecular antagonism and plasticity of

1146 regulatory and inflammatory T cell programs. Immunity *29*, 44-56.
1147 10.1016/j.immuni.2008.05.007.
1148 Yap, J., McCurdy, S., Alcala, M., Irei, J., Garo, J., Regan, W., Lee, B.-H., Kitamoto, S., and
1149 Boisvert, W.A. (2020). Expression of Chitotriosidase in Macrophages Modulates Atherosclerotic
1150 Plaque Formation in Hyperlipidemic Mice. Frontiers in Physiology *11*.
1151 10.3389/fphys.2020.00714.
1152 Yosef, N., Shalek, A.K., Gaublomme, J.T., Jin, H., Lee, Y., Awasthi, A., Wu, C., Karwacz, K.,
1153 Xiao, S., Jorgolli, M., et al. (2013). Dynamic regulatory network controlling TH17 cell
1154 differentiation. Nature *496*, 461-468. 10.1038/nature11981.
1155 Young, M.D., and Behjati, S. (2018). SoupX removes ambient RNA contamination from droplet
1156 based single cell RNA sequencing data. BioRxiv.
1157 Zilionis, R., Engblom, C., Pfirschke, C., Savova, V., Zemmour, D., Saatcioglu, H.D., Krishnan,
1158 I., Maroni, G., Meyerovitz, C.V., Kerwin, C.M., et al. (2019). Single-Cell Transcriptomics of
1159 Human and Mouse Lung Cancers Reveals Conserved Myeloid Populations across Individuals
1160 and Species. Immunity *50*, 1317-1334 e1310. 10.1016/j.immuni.2019.03.009.

1161

1162 **Acknowledgments:** We are grateful to the research and veterinary technicians: Chelsea
1163 Chedrick, Carolyn Bigbee, Nicholas Schindler, Mark Rogers, Tara Rutledge, Chelsea Causgrove
1164 and Brianne Stein in the Flynn lab who assisted with this work, as well as helpful discussions
1165 with members of the Flynn, Scanga, Mattila, Lin and Shalek laboratories. We also thank the
1166 efforts of the University of Pittsburgh Division of Laboratory Animal Research technicians for
1167 husbandry of the animals.

1168

1169 **Funding:**

1170 Bill and Melinda Gates Foundation (OP1139972: AL, SMB, SMF, JLF, AKS; OPP1202327:
1171 AKS)

1172 Searle Scholars Program (AKS)

1173 The Beckman Young Investigator Program (AKS)

1174 Sloan Fellowship in Chemistry (AKS)

1175 NIH (5U24AI118672, BAA-NIAID-NIHAI201700104) (AKS)

1176 American Lung Association RG571577(HPG)

1177 F30-AI143160 (TKH)

1178 NIH T32A1065380 (NLG)

1179 NSF GRFP grant (CNT, SKN 1122374)

1180 Fannie and John Hertz Foundation Fellowship (CNT)

1181 Wellcome Trust Fellowship award 210662/Z/16/Z (AL)

1182 Koch Institute Support (core) grant P30-CA14051 from the National Cancer Institute (CL)

1183 NIH CFAR P30 AI060354 (BB)

1184 NIH R01A1022553 (BB)

1185 T32 A1007387 (JR)

1186 NIH K12 (CW)

1187

1188 **Author contributions:**

1189 Conceptualization: JLF, SMF, AKS

1190 Data Curation: HPG, TKH, FFH, PM, AGW, NLG, AL

1191 Formal Analysis: HPG, TKH, CNT, NLF, FFH, AGW,

1192 Methodology: HPG, TKH, MHW, CNT, AAT, TMG, FFH, CK, PM, AGW, SKN, HJB, BB, JCL

1193 Investigation: HPG, TKH, MHW, CNT, AAT, TG, FFH, JW, CK, JMP, PM, AGW, SKN, HJB,
1194 SKCG, AM, KVP, CLA, CTC, JAT, LJF, HJB, PLL, SI, JYP, JMR, AS, JOM

1195 Visualization: HPG, TKH, CNT

1196 Validation: HPG, TKH, NLG, KN, CGW, SI

1197 Resources: JLF, SMF, AKS, JCL. RM, AL

1198 Funding acquisition: JLF, SMF, AKS

1199 Project administration: CAS

1200 Supervision: JLF, SMF, AKS, SMB, BDB, AL, JCL, BB

1201 Writing – original draft: HPG, TKH, SMB, JLF, SMF, AKS

1202 Writing – review & editing: HPG, TKH, CNT, MC, SMB, JLF, SMF, AKS

1203

1204 **Competing interests: A.K.S.** reports compensation for consulting and/or SAB membership from
1205 Merck, Honeycomb Biotechnologies, Cellarity, Repertoire Immune Medicines, Third Rock
1206 Ventures, Hovione, Relation Therapeutics, FL82, Empress Therapeutics, Ochre Bio, and Dahlia
1207 Biosciences.

1208 **CL:** shareholder and consultant Honeycomb biotechnologies

1209 **TKH:** shareholder and consultant nference, inc.

1210

1211 **Data and materials availability**

1212 ***Lead Contact***

1213 Further information and requests for resources and reagents should be directed to and will be
1214 fulfilled by the Lead Contact, Alex K. Shalek (shalek@mit.edu).

1215 ***Materials Availability***

1216 The study did not generate new unique reagents.

1217 ***Data and Code Availability***

1218 Raw and processed data will be available on the Gene Expression Omnibus and can be accessed
1219 at https://singlecell.broadinstitute.org/single_cell/study/SCP257/cellular-ecology-of-m-
1220 tuberculosis-granulomas?scpbr=the-alexandria-project#study-visualize. Additional code is
1221 available upon request from the lead contact.

1222 **Materials and Methods**

1223 ***Ethics Statement***

1224 All experimental manipulations, protocols, and care of the animals were approved by the University of
1225 Pittsburgh School of Medicine Institutional Animal Care and Use Committee (IACUC). The protocol assurance
1226 number for our IACUC is D16-00118. Our specific protocol approval numbers for this project are 18124275 and IM-
1227 18124275-1. The IACUC adheres to national guidelines established in the Animal Welfare Act (7 U.S.C. Sections
1228 2131 - 2159) and the Guide for the Care and Use of Laboratory Animals (8th Edition) as mandated by the U.S. Public
1229 Health Service Policy.

1230 All macaques used in this study were housed at the University of Pittsburgh in rooms with autonomously
1231 controlled temperature, humidity, and lighting. Animals were singly housed in caging at least 2 square meters apart
1232 that allowed visual and tactile contact with neighboring conspecifics. The macaques were fed twice daily with biscuits
1233 formulated for nonhuman primates, supplemented at least 4 days/week with large pieces of fresh fruits or vegetables.
1234 Animals had access to water *ad libitem*. Because our macaques were singly housed due to the infectious nature of
1235 these studies, an enhanced enrichment plan was designed and overseen by our nonhuman primate enrichment
1236 specialist. This plan has three components. First, species-specific behaviors are encouraged. All animals have access
1237 to toys and other manipulata, some of which will be filled with food treats (e.g. frozen fruit, peanut butter, etc.). These
1238 are rotated on a regular basis. Puzzle feeders foraging boards, and cardboard tubes containing small food items also
1239 are placed in the cage to stimulate foraging behaviors. Adjustable mirrors accessible to the animals stimulate
1240 interaction between animals. Second, routine interaction between humans and macaques are encouraged. These
1241 interactions occur daily and consist mainly of small food objects offered as enrichment and adhere to established
1242 safety protocols. Animal caretakers are encouraged to interact with the animals (by talking or with facial expressions)
1243 while performing tasks in the housing area. Routine procedures (e.g. feeding, cage cleaning, etc) are done on a strict
1244 schedule to allow the animals to acclimate to a routine daily schedule. Third, all macaques are provided with a variety
1245 of visual and auditory stimulation. Housing areas contain either radios or TV/video equipment that play cartoons or
1246 other formats designed for children for at least 3 hours each day. The videos and radios are rotated between animal
1247 rooms so that the same enrichment is not played repetitively for the same group of animals.

1248 All animals are checked at least twice daily to assess appetite, attitude, activity level, hydration status, etc.
1249 Following *M. tuberculosis* infection, the animals are monitored closely for evidence of disease (e.g., anorexia, weight
1250 loss, tachypnea, dyspnea, coughing). Physical exams, including weights, are performed on a regular basis. Animals

1251 are sedated prior to all veterinary procedures (e.g. blood draws, etc.) using ketamine or other approved drugs. Regular
1252 PET/CT imaging is conducted on most of our macaques following infection and has proved very useful for monitoring
1253 disease progression. Our veterinary technicians monitor animals especially closely for any signs of pain or distress. If
1254 any are noted, appropriate supportive care (e.g. dietary supplementation, rehydration) and clinical treatments
1255 (analgesics) are given. Any animal considered to have advanced disease or intractable pain or distress from any cause
1256 is sedated with ketamine and then humanely euthanatized using sodium pentobarbital.
1257
1258 ***Research Animals***
1259 Four Cynomolgus macaques (*Macaca fascicularis*), >4 years of age, (Valley Biosystems, Sacramento, CA)
1260 were housed within a Biosafety Level 3 (BSL-3) primate facility as previously described and as above. Animals were
1261 infected with low dose (~10 colony-forming units (CFUs)) *M tuberculosis* (Erdman strain) via bronchoscopic
1262 instillation. Infection was confirmed by PET-CT scan at 4 weeks and monitored with clinical and radiographic
1263 examinations until 10 weeks post infection.
1264
1265 ***Serial PET-CT Imaging***
1266 Animals underwent PET-CT scans after Mtb infection at 4 weeks, 8 weeks and pre necropsy (i.e. 10 weeks
1267 post-infection) as previously described (White *et al.*, 2017). Briefly, animals were sedated, intubated and imaged by
1268 2-deoxy-2-^{18}F-D-deoxyglucose (FDG) PET imaging (microPET Focus 220 preclinical PET scanner, Seimens Medical
1269 Solutions, USA, Malvern, PA) and Cretom CT scanner (Neurologica Corp, Danvers, MA, USA) within biosafety level
1270 3 facility. The total lung FDG avidity was analyzed using Osirix viewer, an open-source PACS workstation and
1271 DICOM viewer (Pixmeo, Bernex, Switzerland). The whole lung was segmented on CT by using the growing region
1272 algorithm on the Osirix viewer to create a ROI of normal lung (Hounsfield units < 200). The closing tool was used to
1273 include individual nodules and other pulmonary disease. The ROI was transferred to the co-registered PET scan and
1274 manually edited to ensure all pulmonary disease was included. Voxels outside the ROI were set to zero and voxels
1275 with an SUV greater than or equal to normal lung (SUV > 2.3) were isolated. Finally, the "Export ROIs" plug-in was
1276 then used to export the data from these isolated ROIs to a spreadsheet where the total SUV per voxel were summed
1277 to represent the total lung FDG activity. Total FDG activity in lungs was used to estimate thoracic bacterial burden
1278 prior to reinfection (**Figure 1C**), as previously published(Coleman *et al.*, 2014b; White *et al.*, 2017). Granulomas were
1279 individually characterized by their date of establishment (scan date), size (mm), and relative metabolic activity as a

1280 proxy for inflammation ([^{18}F]-FDG standard uptake normalized to muscle [SUVR])(Coleman *et al.*, 2014b; White *et*
1281 *al.*, 2017). Granulomas greater than 1mm are detected by CT scan.
1282
1283 ***Necropsy***
1284 Necropsy was performed as previously described (Gideon *et al.*, 2015; Lin *et al.*, 2013; Lin *et al.*, 2009;
1285 Maiello *et al.*, 2018). Briefly, an ^{18}F-FDG PET-CT scan was performed on every animal 1-3 days prior to necropsy to
1286 measure disease progression and identify individual granulomas. At necropsy, monkeys were maximally bled and
1287 humanely sacrificed using pentobarbital and phenytoin (Beuthanasia; Schering-Plough, Kenilworth, NJ). Individual
1288 granulomas previously identified by PET-CT and those that were not seen on imaging from lung and mediastinal
1289 lymph nodes were excised for histological analysis, bacterial burden, and other immunological studies. TB specific
1290 gross pathologic lesions and overall gross pathologic disease burden was quantified using a previously published
1291 method (Maiello *et al.*, 2018). The size of each granuloma was measured by pre-necropsy scans and at necropsy.
1292 Granulomas were enzymatically dissociated using the Gentlemacs dissociator system (Miltenyi Biotec Inc) to obtain
1293 single cell suspension and used to enumerate bacterial burden and applied on a Seq-Well device for single cell RNA-
1294 sequencing (scRNA-seq).
1295
1296 **Bacterial burden**
1297 200 µl of each granuloma homogenate were plated in serial dilutions onto 7H11 medium, and the CFU of *M.*
1298 *tuberculosis* growth were enumerated 21 days later to determine the number of bacilli in each granuloma (Gideon *et*
1299 *al.*, 2015). As a quantitative measure of overall bacterial burden, a CFU score was derived from the summation of the
1300 log-transformed CFU/gram of each sample at the time of necropsy.
1301
1302 **Chromosomal equivalents, CEQ**
1303 DNA extraction and qPCR was performed with modifications as described previously ((Lin *et al.*, 2014b)).
1304 Briefly, frozen aliquots of homogenates were thawed and volumes recorded throughout the extraction process.
1305 Samples were transferred to tubes containing 150 µl of 0.1mm zirconia-silica beads (Biospec Products) before adding
1306 600µl of Tris-EDTA buffer, pH 8.0. Three hundred microliters of phenol/chloroform/isoamyl alcohol (25:24:1, Sigma-
1307 Aldrich) at 70°C were subsequently added and the samples incubated at room temperature for 10 minutes. The samples
1308 were then vortexed, the aqueous layer separated and supplemented with 50 µl 5M NaCl and a second phenol

1309 chloroform extraction performed on the extracted aqueous layer. DNA was precipitated with the addition of one
1310 volume of 100% isopropanol and one-tenth volume of 3M sodium acetate and incubating at -20°C overnight. The
1311 DNA pellet was washed with 70% ethanol, dried and resuspended in nuclease-free water. Mtb genomes were then
1312 quantified using Taqman Universal Master Mix II (Life Technologies) and previously published *sigF* primer-probe
1313 combination (Lin *et al.*, 2014b). Each sample was amplified in triplicate using an ABI Systems 7900HT machine.
1314 Chromosomal equivalents (CEQ) were quantified by comparing the samples with a standard curve derived from serial
1315 dilution of Mtb genomes prepared from liquid culture. Our detection limit for the standard curve was 10 copies per
1316 reaction. When we calculated the number of genomes for the whole granuloma, our detection limit was 1,000 copies
1317 per granuloma. Of the 26 granulomas analysed, 2 granulomas failed at the CEQ quantification and they were
1318 eliminated from CEQ and CFU/CEQ analysis.
1319
1320 **Immunohistochemistry analysis**
1321 Granulomas from macaques were harvested at 10 or 11 weeks post Mtb infection from other published
1322 (Phuah *et al.*, 2016) and unpublished studies at the University of Pittsburgh. Following formalin fixation and paraffin
1323 embedding, 5 µm sections were placed on slides for staining. Slides were deparaffinized in xylenes, hydrated in a
1324 series of graded ethanol dips, and then antigen retrieval was performed by boiling the slides in a pressure cooker
1325 containing antigen retrieval citrate buffer for slides stained with c-kit and tryptase or Tris-EDTA buffer (Mattila *et al.*,
1326 2013) for slides stained with CD11c, CD20, and CD3. Sections were cooled to room temperature and washed with 1X
1327 PBS then stained overnight at 4°C in a humidified chamber using anti-human c-kit , anti-mast cell tryptase antibodies,
1328 or rabbit-anti-CD3 and mouse anti-CD11c antibodies as previously described (Phuah *et al.*, 2016). For the c-kit and
1329 tryptase stained slides, the tissue sections were washed three times using 1X PBS and then incubated with anti-mouse
1330 IgG1 AF546 to label the anti-c-kit antibodies for 1 hour at room temperature in a humidified chamber. Tryptase
1331 staining was performed overnight at 4°C with anti-tryptase antibodies that were labeled with an Alexa Fluor 488 anti-
1332 rabbit IgG Zenon labeling kit. For the CD3, C11c, and CD20 stained sections, the CD3 and CD11c antibodies were
1333 labeled with donkey anti-rabbit IgG Alexa Fluor 647 and anti-mouse IgG Alexa Fluor 488-conjugated secondaries
1334 purchased Jackson ImmunoResearch Laboratories (West Grove, PA) or ThermoFisher, respectively. After the
1335 secondary antibodies were removed with PBS washes, CD20 was stained with rabbit anti-CD20 that was labeled with
1336 Alex Fluor 546 anti-rabbit IgG Zenon labeling kit. For both staining panels, the sections were washed again in 1X
1337 PBS and coverslips were applied using ProLong Gold Antifade Mountant with DAPI. For the slides stained with CD3,

1338 CD11c, and CD20, individual image channels were acquired with an Olympus FluoView 500 laser scanning confocal
1339 microscope (Olympus, Life Sciences Waltham, MA) maintained by the University of Pittsburgh's Center for Biologic
1340 Imaging and combined and pseudocolored with the FIJI build of ImageJ (Schindelin et al., 2012). Images of c-kit and
1341 tryptase-stained slides were acquired with a Nikon e1000 epifluorescence microscope (Nikon Instruments, Melville,
1342 NY) operated by the NIS-Elements AR software package (Nikon).

1343 Human granulomas were identified from sections of lung tissue obtained at subjects undergoing partial lung
1344 resection for clinical indications at King Dinzulu Hospital and Inksosi Albert Luthili central Hospital in Durban, South
1345 Africa. Gross pathology was assessed by Haematoxylin and Eosin (H&E) staining. Briefly, samples of lung were fixed
1346 in 10% neutral buffered formalin and processed routinely in a vacuum filtration processor using a xylene-free method
1347 with isopropanol as the main substitute fixative. Tissue sections were embedded in paraffin wax. Sections cut at 4 µm
1348 using a microtome, heated at 56°C for 15 min, dewaxed through two changes of xylene and rehydrated through
1349 descending grades of alcohol to water and stained with Haematoxylin & Eosin (H&E, 5 minute incubation with each
1350 stain). Slides were dehydrated in ascending grades of alcohol, cleared in xylene, and mounted with a mixture of
1351 distyrene, plasticizer, and xylene (DPX). For immunohistochemistry, 4 µm sections and were mounted on charged
1352 slides and heated at 56°C for 15 min. Mounted sections were dewaxed in xylene followed by rinsing in 100% ethanol
1353 and 1 change of SVR (95%). Slides were then washed under running water for 2 min followed by antigen retrieval via
1354 Heat Induced Epitope Retrieval (HIER) in Tris-sodium chloride (pH 6.0) for 30 minutes. Slides were then cooled for
1355 15 min and rinsed under running water for 2 min. Endogenous peroxide activity was blocked using 3% hydrogen
1356 peroxide for 10 min at room temperature (RT). Slides were then washed in phosphate-buffered saline with 1% Tween
1357 (PBST) and blocked with protein block (Novolink) for 5 min at RT. Sections were incubated with primary antibodies
1358 for CD117 (A4502-CD117,c-kit, DAKO, 1:500), followed by washing and incubation with post primary (Novolink)
1359 for 30 minutes at RT. Slides were washed with PBST followed by incubation with the polymer (Novolink) for 30 min
1360 at RT. Slides were then washed and stained with DAB for 5 min, washed under running water and counterstained with
1361 hematoxylin for 2 min. Slides were rinsed under running water, blued in 3% ammoniated water for 30 s, washed under
1362 water, dehydrated and mounted in DPX.

1363

1364 **Flow cytometry**

1365 Granulomas harvested from other Mtb infected NHPs were used in the flow cytometery analysis and processed as
1366 previously published(Gideon *et al.*, 2015). Cells were counted and stained for viability using fixable viability dye

1367 (Zombie NIR, Biolegend) and other surface and intracellular markers using the standard protocols. Surface markers
1368 include: CD3 (SP34-2, BD), CD4 (L200, BD), CD8a (RPA-T8, BD), CD8b (2ST8.5H7, BD), TCR gd (5A6.E9,
1369 ThermoFisher), CD16 (3G8, BD), NKG2A (Z199, Beckman Coulter and intracellular markers include: Granzyme B
1370 (GB11, BD), Granzyme A (CB9, BD) and Granzyme K (G3H69, BD). Samples were acquired on a Cytek Aurora
1371 spectral cytometer (5 laser configuration) and unmixed using SpectroFlo software (Cytek). Final analysis was
1372 performed in FlowJo (v10, FlowJo)

1373 **Single-cell RNA-Sequencing (scRNA-seq)**

1374 High-throughput scRNA-seq was performed using the Seq-Well platform as previously described (Gierahn
1375 *et al.*, 2017). Briefly, total cell counts from single-cell suspension of granuloma homogenate were enumerated and
1376 ~15,000-30,000 cells were applied to the surface of a Seq-Well device loaded with capture beads in the BSL-3 facility
1377 at University of Pittsburgh. Following cell loading, Seq-Well devices were reversibly sealed with a polycarbonate
1378 membrane and incubated at 37°C for 30 minutes. After membrane sealing, Seq-Well devices were submerged in lysis
1379 buffer (5 M guanidine thiocyanate, 10 mM EDTA, 0.1% β-mercaptoethanol, 0.1% Sarkosyl) and rocked for 30
1380 minutes. Following cell lysis, arrays were rocked for 40 minutes in 2 M NaCl to promote hybridization of mRNA to
1381 bead-bound capture oligos. Beads were removed from arrays by centrifugation and reverse transcription was
1382 performed at 52°C for 2 hours. Following reverse transcription, arrays were washed with TE-SDS (TE Buffer + 0.1%
1383 SDS) and twice with TE-Tween (TE Buffer + 0.01% Tween20). Following ExoI digestion, PCR amplification was
1384 performed to generate whole-transcriptome amplification (WTA) libraries. Specifically, a total of 2,000 beads were
1385 amplified in each PCR reaction using 16 cycles as previously described (Gierahn *et al.*, 2017). Following PCR
1386 amplification, SPRI purification was performed at 0.6x and 0.8x volumetric ratios and eluted samples were quantified
1387 using a Qubit. Sequencing libraries were prepared by tagmentation of 800 pg of cDNA input using Illumina Nextera
1388 XT reagents. Tagmented libraries were purified using 0.6x and 0.8x volumetric SPRI ratios and final library
1389 concentrations were determined using a Qubit. Library size distributions were established using an Agilent
1390 TapeStation with D1000 High Sensitivity ScreenTapes (Agilent, Inc., USA).

1391

1392 **Bulk RNA Sequencing**

1393 Bulk RNA sequencing was performed using cells obtained from a total of 12 granulomas from a separate set of animals
1394 infected with Mtb for 10 weeks. Initially, granulomas were enzymatically dissociated and cells from each granuloma
1395 were placed in 100 uL of lysis buffer. RNA was then extracted from whole lysates using RNEasy kits (Qiagen, Inc.)

1396 and combined with mRNA capture beads. Reverse transcription, whole transcriptome amplification, tagmentation and
1397 sequencing were performed as described above. Within each bulk RNA sequencing sample, expression values were
1398 summarized across bead barcodes to obtain an aggregate expression profile for each population.
1399
1400 **Sequencing and Alignment**
1401 Libraries for each sample were sequenced on a NextSeq550 75 Cycle High Output sequencing kit (Illumina
1402 Inc., Sunnyvale, CA, USA). For each library, 20 bases were sequenced in read 1, which contains information for cell
1403 barcode (12 bp) and unique molecular identifier (UMI, 8bp), while 50 bases were obtained for each read 2 sequence.
1404 Cell barcode and UMI tagging of transcript reads was performed using DropSeqTools v1.12 (Macosko et al., 2015).
1405 Barcode and UMI-tagged sequencing reads were aligned to the Macaca fascicularis v5 genome
1406 (https://useast.ensembl.org/Macaca_fascicularis/Info/Index) using the STAR aligner. Aligned reads were then
1407 collapsed by barcode and UMI sequences to generate digital gene expression matrices with 10,000 barcodes for each
1408 array.
1409
1410 **QUANTIFICATION AND STATISTICAL ANALYSIS**
1411 **Data Processing and Quality Control**
1412 Initially, after examining a range of cell inclusion thresholds, a combined dataset of 169,830 barcodes was
1413 generated by applying a cutoff of 500 genes and 750 transcripts (UMIs). We visualized cells from each array using t-
1414 SNE across 30 principal components and performed Louvain clustering in Seurat. For many arrays, large clusters of
1415 cell barcodes were identified that were not marked by distinct cell-type defining gene expression. Instead, these cells
1416 were marked by distributed, low-level expression of genes presumed to originate from other cell types (e.g. *HBB* from
1417 erythrocytes, *JCHAIN* from plasma cells, and *CPA3* from mast cells). To understand the identity of these barcodes
1418 more fully, sequencing quality metrics were initially examined, and non-descript clusters did not significantly differ
1419 in the total number of aligned reads, detected genes, UMIs/cell, or mitochondrial percentage.
1420
1421 To more fully understand the identity of these clusters, multiple modeling approaches were pursued **(Figure S2)**:
1422
1423 1. Initially, low-quality clusters were modelled as array-specific doublets. Here, models were constructed in which
1424 pseudo-doublets/multiplets (n=2, 5, 10, 15, or 20 cells) were created from random sampling of the remaining cell

1425 type clusters. However, in these models, there was not significant overlap between the generated pseudo-
1426 multiplets and the clusters with non-distinct gene expression patterns.
1427
1428 2. Random cells were created by binomial sampling a pseudo-population average expression vector generated by
1429 summation of expression profiles across all cell type clusters not suspected to be derived from ambient
1430 contamination. In these models, direct overlap was not observed between the simulated mixed population and
1431 those clusters with non-distinct gene expression patterns.
1432
1433 3. Finally, we examined whether these clusters might represent deep sampling of ambient contamination or cellular
1434 debris by generating a "contamination" scoring scheme. First, to identify the clusters within each array, 30
1435 principal components were calculated (this was observed to consistently capture the majority of variation in each
1436 array), and Louvain clustering (resolution = 1.25) was performed using all significant principal components
1437 (JackStraw Empirical P-value < 0.05). Next, within each array, cluster-specific "contamination" scores were
1438 generated that consisted of 3 components:
1439
1440 a. **A measure of array-specific background contamination by cluster ("soup expression").** For each
1441 array, a background expression profile was generated based on low-UMI barcodes (See ***Correction for***
1442 ***Residual Background Contamination*** below for full details). A set of "soup"-defining genes was
1443 identified at a range of thresholds for soup-defining gene expression (0.01, 0.005, 0.001, and 0.0005;), a
1444 value that represents the proportional contribution of a given gene to the cumulative soup expression
1445 profile for each array. Array-specific, background-contamination scores were generated for the set of
1446 soup-defining transcripts using the AddModuleScore function in Seurat. Clusters with
1447 ambiguous/overlapping expression of lineage-defining gene expression signatures (Erythrocytes: *HBB*,
1448 Plasma cells: *JCHAIN*, Mast cells: *CPA3*, etc.) were observed to be significantly enriched for soup-
1449 defining gene expression. Finally, to calculate "contamination' scores, expression scores for soup genes
1450 at a threshold of 0.001 were generated to calculate the average soup-profile score for each cluster within
1451 each array.
1452

1453 b. **An estimate of biological signal ("biological signal")**. Here, the average log-fold change for the top 5
1454 genes enriched within each cluster was calculated. For clusters dominated by ambient RNA, lower fold
1455 change enrichments for their biological signature genes were observed relative to clusters characterized
1456 by expression of canonical cluster-defining genes. In cases where the highest average log-fold change
1457 values within a cluster were below the "return threshold" in Seurat, we set the value to the default return
1458 threshold of 0.25.

1459
1460

1461 c. **A measure of co-expression of lineage-defining genes ("soup linage coexpression")**. 5 genes were
1462 manually selected that were recurrently over-represented in clusters suspected to arise from ambient
1463 contamination and cellular debris. Specifically, the following genes were selected: *HBB* (An erythrocyte-
1464 defining gene), *JCHAIN* (A plasma cell defining gene), *COL3A1* (A fibroblast defining gene), *SFTPC*
1465 (A type 2 pneumocyte defining genes), and *CPA3* (A mast cell defining gene). For each cell barcode,
1466 the number of these five genes with non-zero expression was calculated as a measure of lineage-defining
1467 co-expression. Within each cluster, the average co-expression of these genes was calculated and one was
1468 subtracted from this average to allow for endogenous expression of 1 lineage-defining gene. This
1469 parameter was specifically added to avoid exclusion of *bona fide* cell clusters with high-background
1470 contamination (presumably due to low endogenous RNA content) and low biological signal (e.g., naïve
1471 T cells). Here, cell populations that scored high for markers of a single lineage yet had higher soup-
1472 expression scores presented with lower rates of co-expression of these soup and lineage defining
1473 transcripts relative to clusters which did not, likely representing ambient RNA and debris.

1474

1475 Using these three values, cluster-specific background "contamination" scores were calculated for each array in 2 ways:

1476

1477 $$Contamination\ Score\ 1 = \frac{(Soup\ Expression)\ x\ (Soup\ Lineage\ Coexpression)}{Biological\ Signal}$$

1478

1479 $$Contamination\ Score\ 2 = \frac{(Scaled - Soup\ Expression)\ x\ (Soup\ Lineage\ Coexpression)}{Biological\ Signal}$$

1480

1481 These two "contamination" scores quantify both the (1) absolute and (2) relative soup-profile contamination in
1482 subsequent cluster classification.

1483 Next, for each array, clustering was performed to identify clusters with array-specific ambient contamination
1484 and debris. Specifically, hierarchical clustering was performed using a total of 7 variables to identify clusters defined
1485 by ambient contamination: the 2 contamination scores (shown above), three scaled soup scores (soup gene thresholds:
1486 0.01, 0.05 and 0.001), the average log-fold change for the top 5 cluster genes, and soup/lineage gene co-expression.
1487 For each array, the hierarchical clustering tree was cut at the first branch point to identify clusters with a signature of
1488 ambient contamination. In total, 41 array-specific clusters, comprising 56,590 barcodes from 21 out of 32 total arrays,
1489 were identified as characterized by ambient RNA contamination and cellular debris and removed them in all
1490 subsequent analyses.

1491

1492 **Correction for Residual Background Contamination**

1493 After removal of cell barcodes that were derived from background contamination and extracellular debris,
1494 additional correction for ambient RNA contamination was performed among remaining cell barcodes on an array-by-
1495 array basis. Among filtered cell barcodes, array-specific, ambient RNA contamination was observed to be marked by
1496 ectopic expression of cell-type defining genes (e.g. widespread expression of *JCHAIN*, *HBB*, and *CPA3* etc.).
1497 Specifically, this contamination was observed to vary in relation to the overall distribution of cell types recovered
1498 from each array. To correct for residual ambient contamination within each array, SoupX (Young, 2018) was used to:
1499 (1) generate array-specific profiles of background contamination, (2) estimate per-cell contamination fractions, and
1500 (3) generate corrected background-corrected UMI counts matrices. To generate background expression profiles,
1501 counts matrices containing up to 50,000 barcodes were generated to assemble a collection of low-UMI cell barcodes
1502 that presumably represent extracellular mRNA. For each array, a UMI threshold for background expression was
1503 determined using EmptyDrops (Lun et al., 2019) to estimate the likelihood distribution that low-UMI barcodes
1504 represent cells rather than ambient contamination. Using an array-specific UMI-threshold (Range: 20-100 UMIs), a
1505 composite background profile was created for each array. To estimate the per-cell contamination fraction, a set of
1506 lineage-defining genes was first identified with bimodal expression patterns across cells (i.e., lineage defining genes
1507 with leaky expression). For each array, this set of soup-defining, lineage genes was used to estimate contamination
1508 fraction for cell types with known endogenous expression. Finally, the composite soup profile was subtracted from
1509 each the transcriptional profile of each cell based on the estimated contamination fraction. For each array, individual

1510 transcripts most likely to be contamination were removed from each single-cell based on the estimated contamination
1511 fraction. Specifically, individual transcripts were sequentially removed from each single-cell transcriptome until the
1512 probability of subsequent transcripts being soup-derived was less than 0.5 to generate a background-corrected counts
1513 matrix for each array **(Table S2b)**.
1514
1515 **Separation of Doublets**
1516 Within each array, doublet identification and separation were performed using DoubletFinder. To account
1517 for differences in cell loading densities and expected cell doublet frequencies, array-specific estimates of the expected
1518 number of doublets were generated **(Table S2a).** For example, for a total of 20,000 cells applied to a Seq-Well device
1519 containing 85,000 wells (lambda = 20,000), an expected doublet rate of >2.37% (since not all of the array's surface
1520 area contains wells) was calculated. For each array, pseudo-doublets were generated using DoubletFinder (McGinnis
1521 et al., 2019). Here, the pK parameter estimate was separately optimized for each array by performing a parameter
1522 sweep in which we selected the pK value with the maximum bimodality coefficient, while a pN = 0.25 was maintained
1523 across all arrays based on published recommendations (McGinnis *et al.*, 2019). Cells were identified as doublets based
1524 on their rank order in the distribution of the proportion of artificial nearest neighbors (pANN). Specifically, the pANN
1525 value for the cell at the expected doublet percentile was identified and the corresponding pANN value was used as a
1526 threshold to remove additional cells in the event of ties. In total, we excluded 3,656 cells as doublets (**Table S2a,c**).
1527
1528 **Integrated Cell Type Classification**
1529 Following the aforementioned quality filtering, a combined dataset of 109,584 cells was used in downstream
1530 analysis **(Table S2d)**. An initial dimensionality reduction was performed on these cells by selecting 1580 variable
1531 genes, performing principal component analysis (PCA), UMAP dimensionality reduction and Louvain clustering
1532 using Scanpy (Wolf et al., 2018). To identify broad cell types, we examined cluster assignments at multiple levels of
1533 clustering resolution (Resolutions: 0.5 to 2.25). We selected a cluster resolution of 1.00 because this was the resolution
1534 beyond which branching did not result in discovery of clusters that represent distinct cell lineages (e.g., division of
1535 Type 1 and Type 2 pneumocytes) (**Table S5**). To define major cell populations, extensive comparisons to existing
1536 signatures of lung parenchyma and immune cell populations were performed using data from the Tabula Muris (Tabula
1537 Muris *et al.*, 2018) and Mouse Cell Atlas (Han *et al.*, 2018) studies. Specifically, lung scRNA-seq data from both
1538 studies were collected and used to calculate enriched gene expression signatures for each lung cell type cluster using

1539 a Wilcox rank-sum test. For each cluster, the top 20 genes **(Table S3)** were selected as a cluster-specific expression
1540 signature and then used them to score all cells in the granuloma dataset. The average signature score within each
1541 cluster was calculated and the distribution of signature score was examined within each granuloma cell type, and
1542 significance was determined via permutation testing.
1543
1544 **Cell Type Assignment of Proliferating Cells**
1545 Among our top-level clusters was one defined by markers of cellular proliferation (*MKI67*, *TOP2A*, and
1546 *CDK1*). To identify the underlying cell type identity for these cells, a separate dimensionality reduction and clustering
1547 was performed among 3,123 cells defined by this proliferation signature. UMAP dimensionality reduction and
1548 Louvain clustering was running at multiple clustering resolutions (0.4-0.8), and a resolution of 0.70 was selected as
1549 the value beyond which no additional major cell type clusters were observed (**Figure S3E**). For each of the major cell
1550 types identified in the global clustering analysis, we generated a gene signature using the top 20 enriched genes and
1551 scored the proliferating cells clusters using the AddModuleScore function in Seurat. We then examined the distribution
1552 of cell-type signature scores across each of the sub-clusters of proliferating cells and re-assigned clusters based on
1553 enrichment of lineage-specific gene expression. Here, we assessed the significance of the cluster scores using a
1554 permutation test. More specifically, 1,000 permutations were performed in which the proliferating clusters were down-
1555 sampled to have the same number of cells. Cluster assignments of the cells were randomized and the average generic
1556 cell type signature score was calculated for each randomized cluster. The significance of a cell type score for each
1557 proliferating cluster was determined by comparing the observed average signature score to the random null
1558 distribution. Through this approach, distinct clusters of proliferating B cells, macrophages, neutrophils, plasma cells,
1559 and T cells were identified and re-assigned to their respective cell types.
1560
1561 **Filtering of Soup-Defining Transcripts**
1562 To avoid artifacts from ambient RNA contamination and cellular debris in sub-clustering of T cells and
1563 macrophages, genes that were observed to be soup-defining for any array were excluded. Specifically, a set of 210
1564 soup defining genes was identified that comprised 0.001 of total soup expression in any array. The threshold of 0.001
1565 was selected to maximize the cumulative fraction of soup expression with the least number of genes to avoid removing
1566 underlying biology. Here, this threshold value represents cumulative fraction of soup expression accounted for by a
1567 given gene for each array. In a further effort to avoid removing cell type specific biology, any genes with average log-

1568 fold changes greater than 1.00 in T cells and macrophages compared to all other generic cell types were retained. In
1569 total, 204 and 180 genes were removed prior to sub-clustering analysis of T cells and macrophages, respectively.
1570
1571 **Sub-clustering of Granuloma unified T and NK cells**
1572 Across the complete set of 44,766 T and NK cells, Louvain clustering was initially performed at a range of
1573 resolution of values (0.30 – 0.75) to examine the relationships between cluster membership. In this analysis, a cluster
1574 was observed to be defined by persistent expression of contaminating transcripts derived from macrophage and mast
1575 cells (Cluster 4 - Louvain Resolution 0.60). To confirm that these cells did not represent persistent doublets, all T cells
1576 were scored by expression of the top 20 cluster defining T cells and similar signature scores between the contaminated
1577 cell population were observed. Additional sub-clustering within the "contaminated" T cell cluster was performed to
1578 understand whether residual contamination obscured additional T cell biology; this failed to reveal additional T cell
1579 clusters not identified among the remaining non-contaminated populations. Since this contamination cluster was not
1580 observed to obscure a novel T cell phenotype, this population was excluded from downstream analysis. Following
1581 removal of the cluster of T cells defined by residual contamination, dimensionality reduction and clustering at multiple
1582 clustering resolutions (Louvain resolution: 0.25 – 0.75) were performed. In this final analysis, a total of 12 T cell
1583 populations were identified at a clustering resolution of 0.75. Finally, additional sub-clustering was performed within
1584 the population of 2,377 γδ and cytotoxic T cells, including dimensionality reduction and clustering at multiple
1585 resolutions (0.30 – 0.75). Here, 2 primary populations of cells were identified: sub-cluster 2, a population of cytotoxic
1586 cells enriched for expression of *TRDC* and sub-cluster 3, a population of XCL1+ NK cells. Differential expression
1587 analysis was performed to determine differences in gene expression between these clusters upon which the
1588 classification of these cells was based.
1589 Additional sub-clustering analysis was performed within the T1-T17 population through repeated variable
1590 gene identification, dimensionality reduction and Louvain clustering (Resolution = 0.55), and 4 distinct sub-
1591 populations were discovered. Differential expression analysis was performed within the 9,234 T1-T17 cells using a
1592 Wilcox test in Seurat to identify sub-cluster defining gene signatures.
1593
1594 **Annotation of T /NK subclusters**

1595 T cell populations were classified using a combination of manual curation and comparison to literature-derived
1596 sequences. Granuloma T cell populations were compared to publicly available T cell population and scRNA-seq
1597 signatures. Specifically, comparisons were performed in the following ways:
1598
1599 1. For each T cell cluster, cluster-defining genes were compared to publicly available databases of immune
1600 signatures, including IPA, GeneGO, MSigDb (Liberzon et al., 2011) and SaVant (Lopez *et al.*, 2017). This
1601 was performed by comparing the set of T cell cluster-defining genes (Adjusted p-value < 0.001 and log-FC
1602 > 0.2) to the signatures in GSEA and the SaVant data using Piano (Lopez *et al.*, 2017; Varemo et al., 2013).
1603 Specifically, significance was assessed using a hypergeometic test to examine the likelihood of the observed
1604 frequency of enriched genes. Among cluster-defining genes for each T/NK cell sub-cluster, comparisons
1605 were performed within each GSEA collection C1-7 (https://www.gsea-
1606 msigdb.org/gsea/msigdb/collections.jsp) and to the SaVant database. Expression signatures were also
1607 compared to MSigDB signatures using GSEA. Here, pseudo-bulk expression signatures were generated for
1608 each T/NK sub-population as the average gene expression across all cells within each cluster. These average
1609 expression values were used to perform GSEA for each cluster in which the expression values were compared
1610 to all other clusters using 1,000 permutations.
1611
1612 2. Each T cell cluster was compared to literature-derived signatures of T cells from another scRNA-seq study.
1613 Here, cell signature scores were generated in Seurat using the AddModuleScore function using gene
1614 expression signatures obtained from human lung cancer (Guo et al., 2018). To determine the significance of
1615 these score, 1,000 permutations were performed in which T cell cluster identity was randomly re-assigned to
1616 generate a null distribution of module scores.
1617
1618 3. Finally, extensive manual curation was performed based on literature evidence. For each cell population, an
1619 extensive literature search was performed to support classification of T cell sub-populations based on patterns
1620 of enriched gene expression. For example, regulatory T cells were identified on the basis of expression of
1621 known regulatory T cell markers (*FOXP3, IKZF1*, and *TNFSF18/GITR*). However, in many cases, surface
1622 markers used to define canonical T cell populations were not detected in the scRNA-seq data.
1623

1624 Next, expression of *TRAC* and *TRBC* or *TRDC* was evaluated within T cells in the scRNA-seq data and the
1625 frequency of cells expressing either *TRAC/TRBC* (yellow) or *TRDC* (green) within each of the 13 clusters was
1626 calculated. While *TRAC/TRBC* expression was observed in all 13 subclusters, *TRDC* expression was observed mainly
1627 in subclusters 1-3 compared to subclusters 4-13. Finally, cluster-specific expression of *CD4* and *CD8A* and *CD8B*
1628 were examined as the proportion of cells with non-zero expression of *CD4*, *CD8A/B* or *CD4&CD8 (A/B)*.
1629

1630 **Sub-clustering of Granuloma Macrophages**

1631 Across 27,670 macrophages, dimensionality reduction and Louvain clustering at multiple clustering
1632 resolutions was performed. In initial clustering, a cluster defined by contaminating transcripts derived from other cell
1633 types (including mast cells (*KIT* and *CLU*), T cells (*CD3D*), and plasma cells (*JCHAIN*)) and soup-defining gene
1634 expression was identified. By comparing the distribution of macrophage-defining gene expression in this cluster to
1635 other clusters, this cluster was observed to have enriched signature scores relative other clusters. The enrichment of
1636 macrophage expression signatures was examined to determine the population of macrophages that have a core
1637 macrophage expression program. While this population of macrophages is primarily soup-defining gene expression,
1638 this cluster was not excluded due to the possibility that this represents an efferocytotic macrophage population.
1639

1640 **Classification of Macrophage Populations**

1641 Identities of the macrophage clusters were established through a combination of manual curation and comparison
1642 to published gene expression signatures from both population and scRNA-seq studies. More specifically:
1643

1644 1. For each macrophage cluster, similar comparison to databases of immune signatures including MSigDb and
1645 SaVanT were performed (See ***Identification of T cell Populations***).
1646

1647 2. A series of gene expression signatures were generated from published scRNA-Seq studies of macrophage
1648 states. For example, a recently published atlas of myeloid states in lung (Zilionis et al., 2019) was used to
1649 score granuloma macrophages. Further, a list of myeloid expression signatures was generated using lung
1650 myeloid cells from the Mouse Cell Atlas (Han *et al.*, 2018). For each study, signatures for the top 20 cluster-
1651 defining genes were selected to generate gene expression signatures **(Table S5).** Signature scores were
1652 generated for each cell using the AddModuleScore function in Seurat.

1653

1654 3. Finally, in cases where an existing description of a macrophage population was not discovered, extensive
1655 literature searches were performed to contextualize possible identities of macrophage populations.

1656
1657 **Deconvolution of Bulk RNA-Sequencing Data**

1658 Population deconvolution was performed using CiberSort (Newman et al., 2015) using reference populations
1659 generated from random sampling of a quarter of the single cells within each of the 13 generic cell types identified in
1660 our single-cell analysis.

1661

1662 **Co-variation in Granuloma Composition**

1663 We calculated correlations in cell-type proportions to identify underlying structure in the co-occurrence of
1664 cell types across all granulomas. Specifically, we calculated Pearson correlation coefficients for all pair-wise cell-type
1665 combinations (**NB** we also performed each analysis using Spearman correlation coefficients and obtained similar
1666 results). For each pairwise combination of cell types, we calculated permutation p-values by randomly re-assigning
1667 cell type labels to generate a set of background correlation values (**Table S9,S10**)

1668 We then performed hierarchical clustering to identify clusters of correlated cell-types across granulomas,
1669 calculating the proportional composition of correlated cell-type clusters within each lesion. For each of the 5 clusters
1670 identified through hierarchical clustering, we calculated permutation p-values to examine average correlation values.
1671 To understand the relationship between identified cell-type clusters and granuloma-level bacterial burden, we
1672 examined the abundance of correlated cell types by grouping lesions by timing of granuloma formation.

1673

1674 **Cell-Communication Analysis**

1675 To examine cell-cell interactions, we first generated a curated list of receptor-ligand pairs through a
1676 combination of publicly-available databases and literature review. Within each granuloma, we generated edge weights
1677 between cell types for a given receptor ligand pair by multiplying the average receptor expression in Cell Type 1 by
1678 the average ligand expression in Cell Type 2. Edge weights were constructed for all receptor-ligand pairs and pairwise-
1679 cell type combinations within granulomas individually. Within each granuloma, we performed a total of 1,000
1680 permutations for each receptor-ligand pair in which cell-type identifiers were randomly resorted and the resulting edge

1681 weight was recorded. For each receptor-ligand pair, the significance of the observed value was calculated from a z-
1682 score comparison of the observed value relative the permuted values.

1683 We further performed adjustment of receptor-ligand edge weights at multiple levels. (1) To account for
1684 differences in the relative abundance of 'sender' cell types, we multiplied receptor-ligand edge weights by the
1685 proportion of all 'sender' cells within a granuloma. In effect, this generates a pool of 'sender' cell derived ligand that
1686 is available to act upon cell types bearing appropriate receptors. (2) To identify the most likely receiver cells, we
1687 weighted receptor-ligand edge-weights by the proportion of total receptor expression within the receiving cell subset
1688 cluster relative to the average receptor expression across all cells in the granuloma. In this scheme, receptors with
1689 more uniform expression across the entire granuloma will be down-weighted to reflect non-autonomous sinks of
1690 extracellular ligands, while receptors predominantly expressed by a single cell subset will be up-weighted. (3) Finally,
1691 we adjusted receptor-ligand edge weights to account for the percent of cells within the receiver cell subset expressing
1692 a given receptor by multiplying our receptor-ligand edge weights by the proportion of all 'receiver' cells expressing
1693 the receptor within a the receiver cell subset.

1694 To identify axes of intercellular communication with differential weights across granulomas, we performed
1695 student's t-tests of receptor-ligand edge weights between (A) high-burden and low-burden lesions, and (B) original
1696 and late-blooming lesions. We filtered results based on the following criteria: (1) the average permutation p-values
1697 for the receptor-ligand pair within high or low-burden lesions < 0.05, (2) p-value from student's t-test in (A) or (B)
1698 above < 0.05, and (3) fold-change of the adjusted receptor edge-weight > 0 in the (A) or (B) comparisons. The "dplyr"
1699 package in R was used to filter the resulting cell-cell interaction database to count significant interactions across cell
1700 type groups and granuloma burdens, identify cell type groups contributing to the top 10% of ligands most strengthened
1701 in either high or low burden granulomas, identify ligands most associated with high or low burden granulomas, and
1702 identify cell type specificity of these ligands. The "circlize" package in R was used to generate circus plots of the
1703 topology of signaling networks across high and low burden granulomas.

1704

1705 **Statistical methods**

1706 Non-parametric Spearman's rho was calculated for correlation analysis for evaluating the degree of
1707 relationship between cellular abundance and bacterial burden. Non parametric t-test was used when comparing two
1708 groups (Mann- Whitney U). P values, or where appropriate adjusted or permutation p values, ≤ 0.05 were considered

1709 significant. Statistical analysis was performed using GraphPad Prism v8 (GraphPad software, San Diego, CA), JMP

1710 Pro v12 and R base statistics.

1711 **Table legends:**

1712 **Table 1:** T/NK subclusters characteristics and annotations

1713 **Table S1:** Granuloma CFU, CEQ, CFU/CEQ; PET-CT: SUV-R, Size and Time of detection

1714 **Table S2a:** Seq-Well array loading densities and doublet rate

1715 **Table S2b:** Technical correction data: SoupX

1716 **Table S2c:** Doublet removal Metadata

1717 **Table S2d:** Cell level metadata

1718 **Tablet S3:** Canonical cell type enrichment gene list: 13 cell type clusters

1719 **Table S4:** Cell type composition: percentage of assigned granuloma cells. A) canonical cell type clusters,

1720 b)macrophage subclusters, c) T/NK subclusters and d) T1T17 subpopulation

1721 **Table S5:** Correlation (Spearman's rho) with bacterial burden and difference between in percentage of

1722 cells in early high burden and late low burden granulomas (Mann Whitney U): A) canonical cell type

1723 clusters, b) T/NK subclusters and C) T1T17 subpopulation

1724 **Table S6:** T/NK subclustering: enrichment gene list :13 T/NK subclusters

1725 **Table S7**: Type1-Type-17 subpopulation enrichment

1726 **Tablet S8:** Macrophage subcluster enrichment: 9 subclusters

1727 **Table S9**: Cellular ecology

1728 **Table S10**: Cellular ecology correlation permutations (Spearman and Pearson)

1729 **Table S11:** Association of cell group abundance with bacterial burden: (1) All: CFU low vs high, (2)

1730 timing of granuloma detection (Early vs late)

1731 **Table S12:** Interaction analysis: receptor-ligand senders and receivers in early and late granulomas.

Fig 1

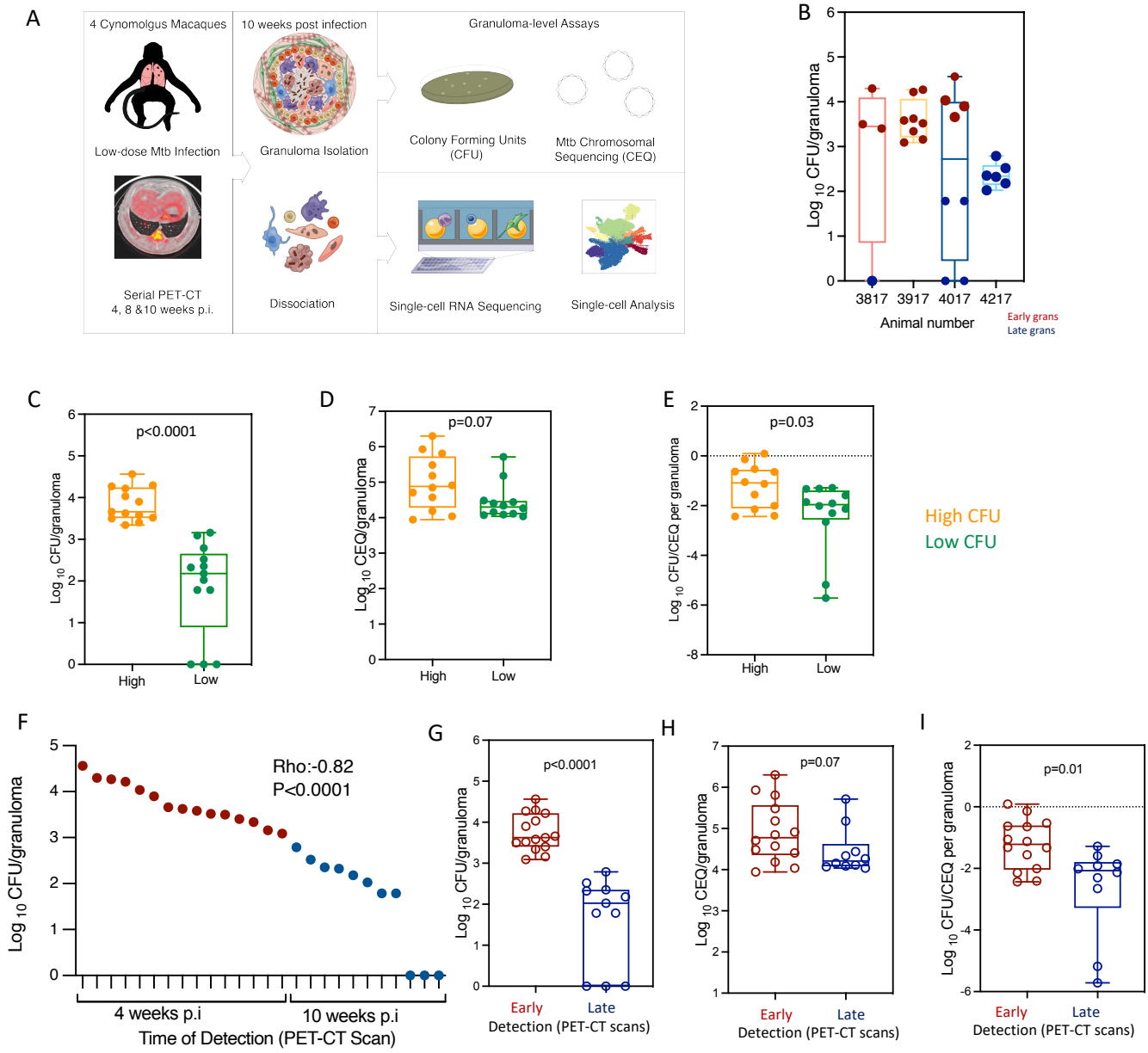


Figure 1. Study design, experimental set up, characteristics of animals over the course of Mtb infection and granuloma bacterial burden

(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) /(G)** CFU $\log_{10}$ per granuloma (total live bacteria); **(D)/ (H)** CEQ $\log_{10}$ per granuloma (Chromosomal equivalents, CEQ, live + dead Mtb) organized by time of detection; **(E)/(I)** 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. **(C-E)** organized by bacterial burden: low (Green); high (orange). **(F)** Individual Granuloma bacterial burden (log CFU) ploted with time of detection by PET-CT scans: 4 weeks post infection (early) or 10 weeks post infection (late). The granulomas in X axis is arranged in order of bacterial burden and time of detection. (**F-I**) time of detection by PET-CT scan (Table S1): early granulomas (maroon), late granulomas (blue). Each symbol is a granuloma. Box plot showing median, IQR and range. Mann Whitney U for panels E-G.

Fig 2

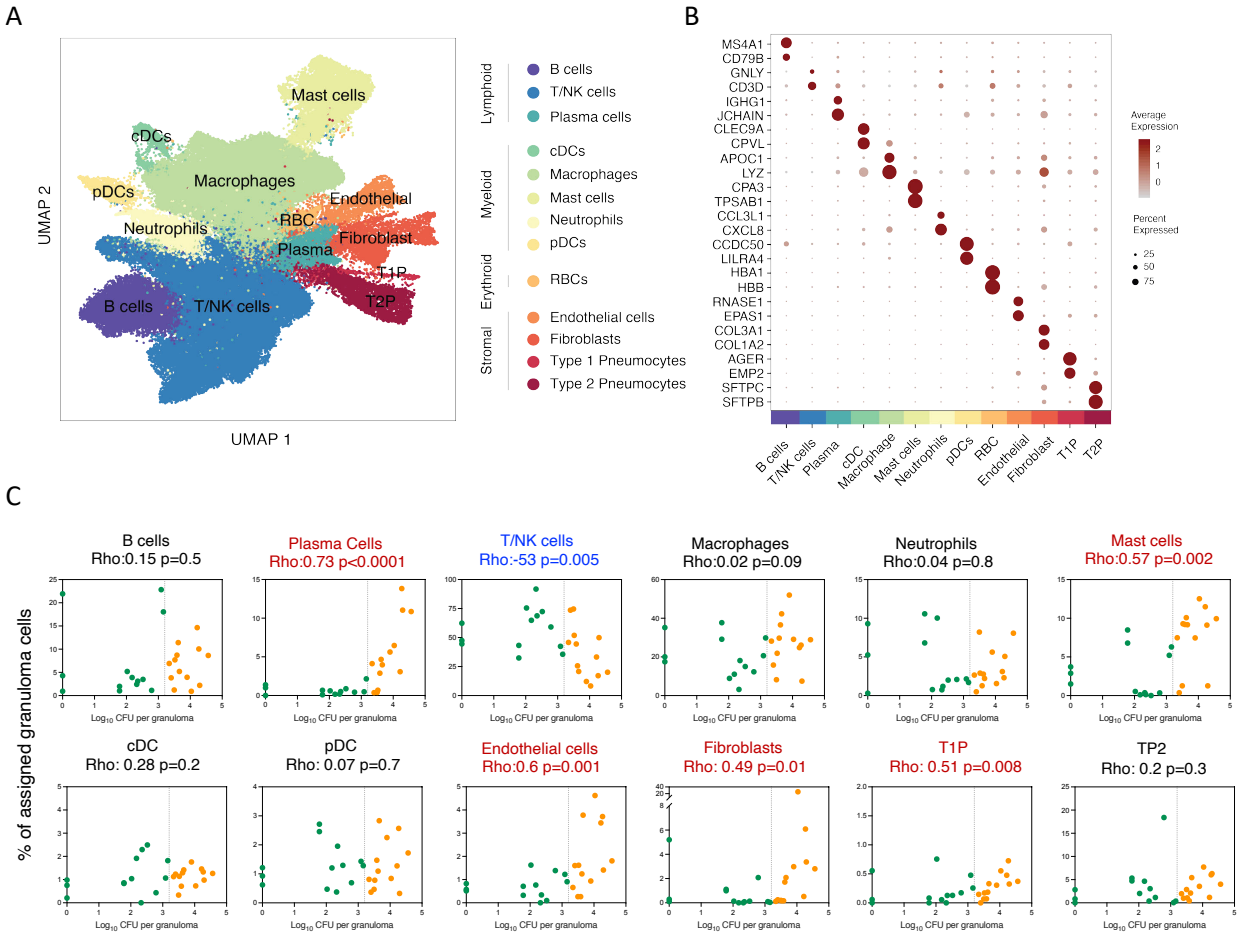


Figure 2 Analysis of single-cell sequencing of tuberculosis lung granulomas

(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 canonical cell types with bacterial burden of individual granulomas (Log_{10} CFU per granuloma) using non-parametric Spearman's rho correlation test. Color indicated binned granuloma bacterial burden: low (green) and high (orange).

Fig 3

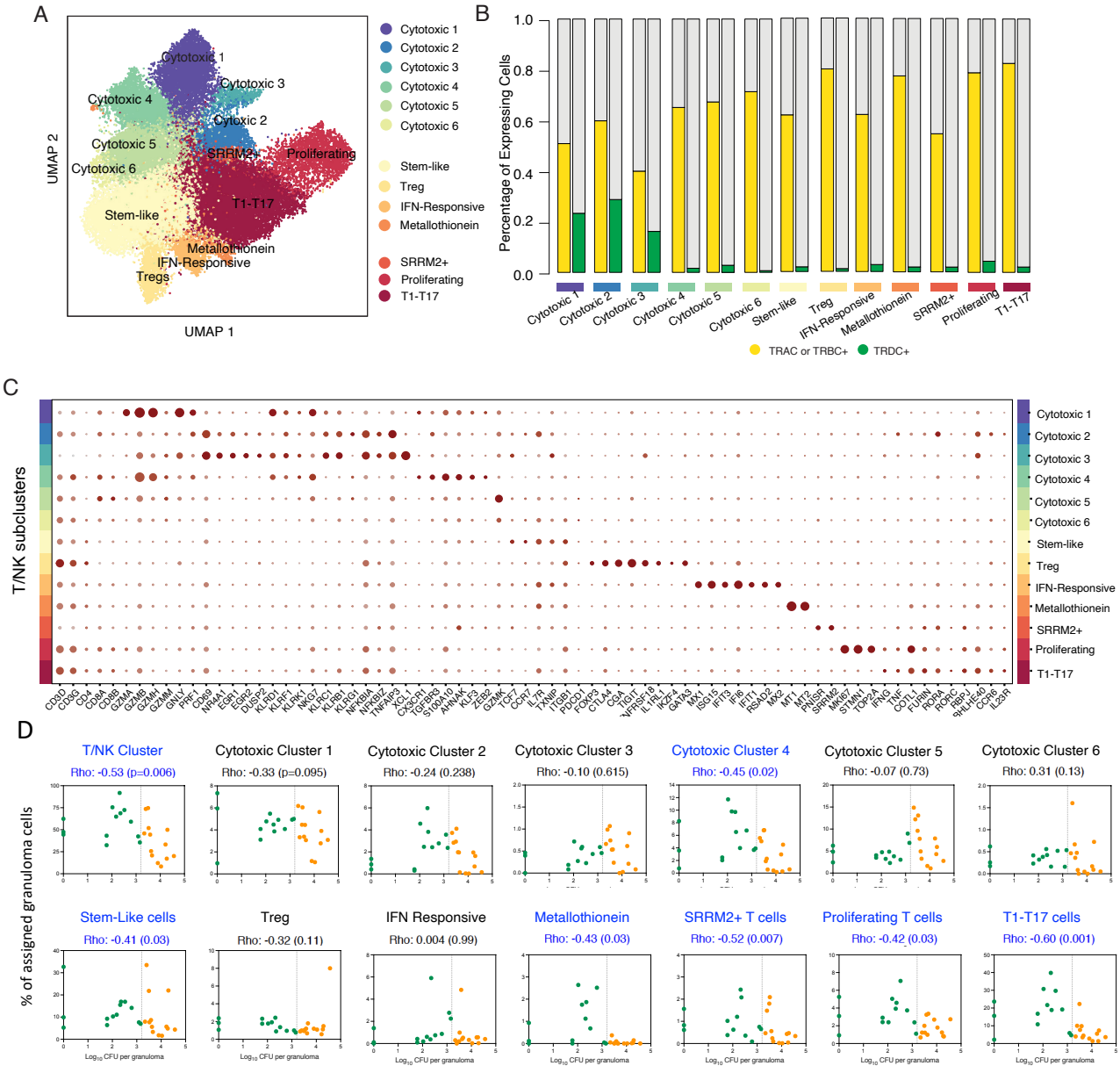


Figure 3.Diversity in the unified T and NK cell cluster and relationship to granuloma-level bacterial burden

(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)** Significant correlations between proportion of T/NK subclusters with bacterial burden of individual granulomas (Log_{10} CFU per granuloma) using non-parametric Spearman's rho correlation test. Color indicated binned granuloma bacterial burden: low (green) and high (orange).

Fig 4

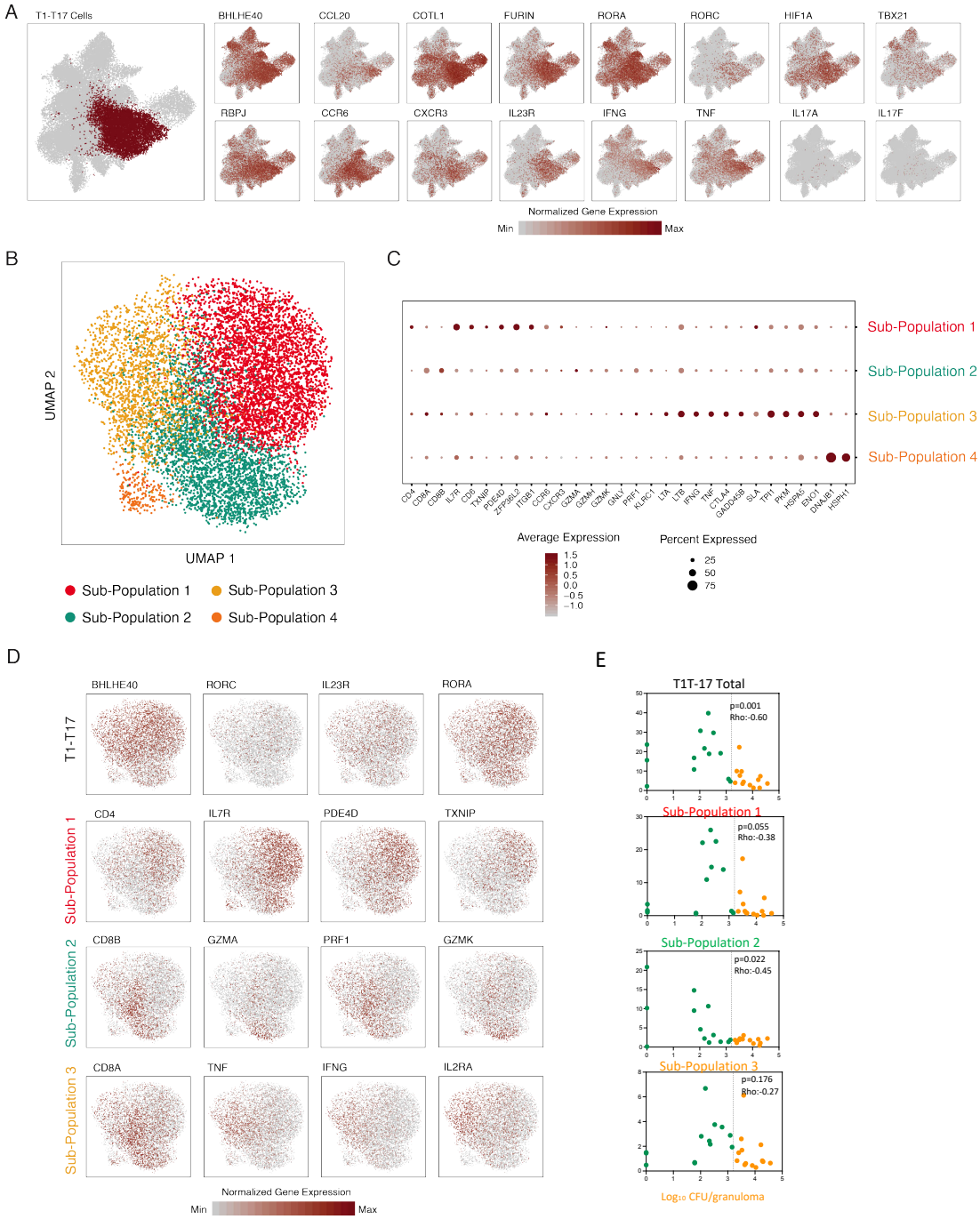


Figure 4 Phenotypic Diversity in T1-T17 cells

(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)** Significant correlations between proportion of T1-T17 subcluster and subpopulations with bacterial burden of individual granulomas (Log_{10} CFU per granuloma) using non-parametric Spearman's rho correlation test.

Fig 5

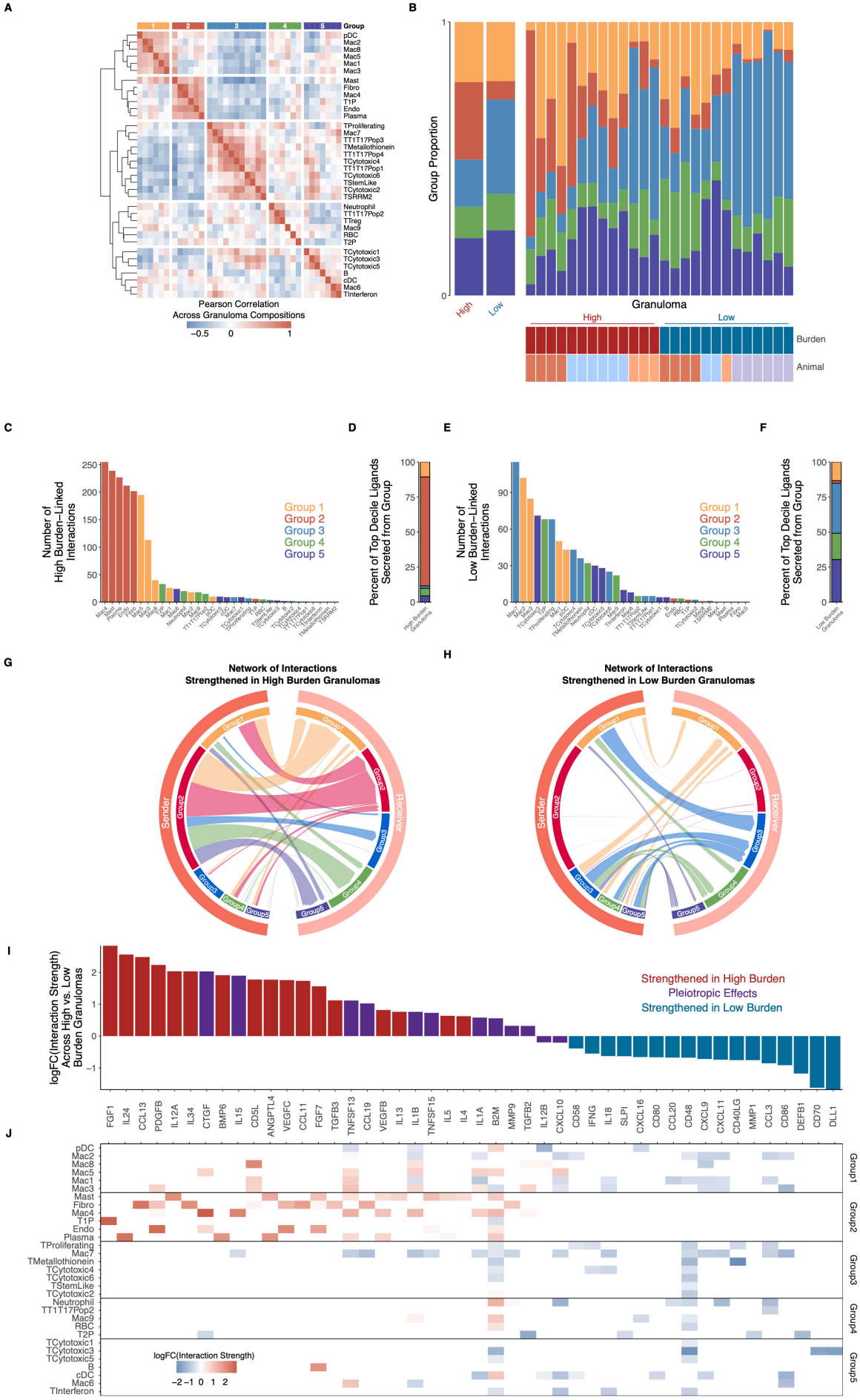
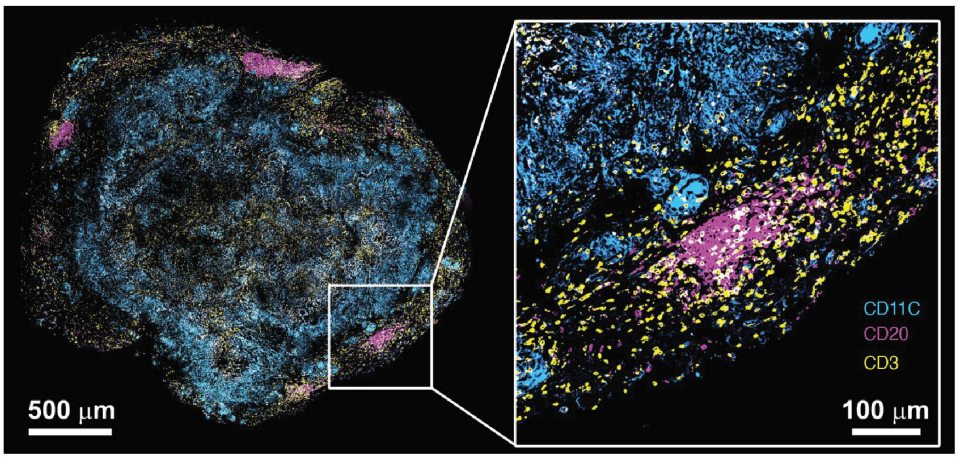

A
Group
pDC
Mac2
Mac8
Mac5
Mac1
Mac3
Mast
Fibro
Mac4
T1P
Endo
Plasma
TProliferating
Mac7
TT1T17Pop3
TMetallothionein
TT1T17Pop4
TCytotoxic4
TT1T17Pop1
TCytotoxic6
TStemLike
TCytotoxic2
TSRRM2
Neutrophil
TT1T17Pop2
TTreg
Mac9
RBC
T2P
TCytotoxic1
TCytotoxic3
TCytotoxic5
B
cDC
Mac6
TInterferon
Pearson Correlation
Across Granuloma Compositions
−0.5
0
1
B
Group Proportion
High
Low
Granuloma
Burden
Animal
C
Number of High Burden−Linked Interactions
Group 1
Group 2
Group 3
Group 4
Group 5
D
Percent of Top Decile Ligands Secreted from Group
E
Number of Low Burden−Linked Interactions
F
Percent of Top Decile Ligands Secreted from Group
G
Network of Interactions
Strengthened in High Burden Granulomas
Sender
Receiver
H
Network of Interactions
Strengthened in Low Burden Granulomas
I
logFC(Interaction Strength) Across High vs. Low Burden Granulomas
Strengthened in High Burden
Pleiotropic Effects
Strengthened in Low Burden
FGF1
IL24
CCL13
PDGFB
IL12A
IL34
CTGF
BMP6
IL15
CD5L
ANGPTL4
VEGFC
CCL11
FGF7
TGFB3
TNFSF13
CCL19
VEGFB
IL13
IL1B
TNFSF15
IL5
IL4
IL1A
B2M
MMP9
TGFB2
IL12B
CXCL10
CD58
IFNG
IL18
SLPI
CXCL16
CD80
CCL20
CD48
CXCL9
CXCL11
CD40LG
MMP1
CCL3
CD86
DEFB1
CD70
DLL1
J
Group1
Group2
Group3
Group4
Group5
logFC(Interaction Strength)
−2 −1 0 1 2

Figure 5 Cellular ecosystem in TB lung granulomas

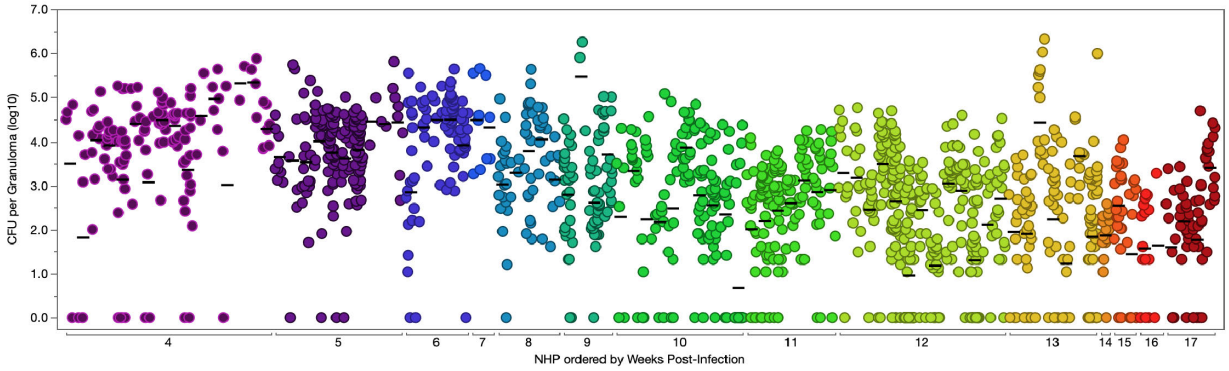
(A) Pairwise Pearson correlation values proportions of canonical cell types and T/NK and macrophage subclusters across 26 granulomas. Hierarchical clustering of correlation coefficients identified 5 groups (indicated by color and number) of cell types with correlated abundance in granulomas. **(B)** Composition of each granuloma by cell type group. Left bar graph shows all high burden and all low burden granulomas grouped together, with right bar graph split by individual granuloma. **(C)** Number of interactions strengthened in high burden granulomas, organized by sender cell clusters (i.e., cell cluster producing the ligand). **(D)** Representation of each cell type group as sender cell population among the 10% of ligands most strengthened in high burden granulomas. **(E)** Number of interactions strengthened in low burden granulomas, organized by sender cell clusters. **(F)** Representation of each cell type group as sender cell population among the 10% of ligands most strengthened in low burden granulomas. **(G)** Network of interactions across cell type groups, subsetted to only highlight interactions strengthened in high burden granulomas. Widths of arcs are proportional to number of interactions between cell type groups, and widths are on same scale as for subfigure F. n = 2,586 statistically significant interactions, 1,715 of which were strengthened in high burden granulomas. **(H)** Network of interactions across cell type groups, subsetted to only highlight interactions strengthened in low burden granulomas. Widths of arcs are proportional to number of interactions between cell type groups, and widths are on same scale as for subfigure E. n = 2,586 statistically significant interactions, 871 of which were strengthened in high burden granulomas. **(I)** Overall high-vs-low granuloma burden fold-change of interactions strengths of key ligands, averaged across all statistically significant interactions. **(J)** Cell cluster-specific interaction strength fold changes of each ligand, averaged across all statistically significant interactions where each cell cluster was the sender population.

Fig S1

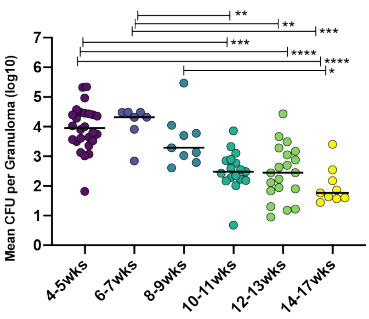
A



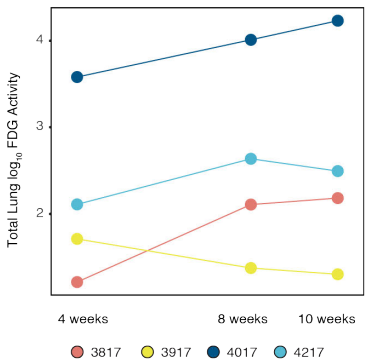
B



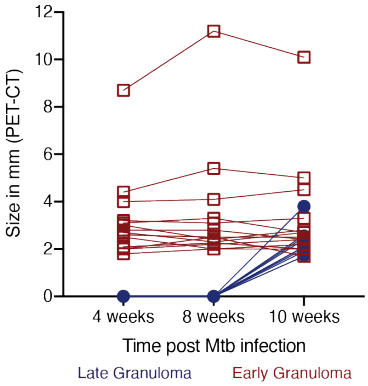
C



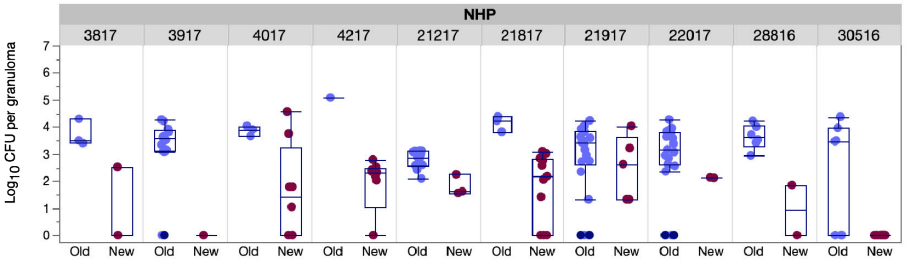
D



E



F



G

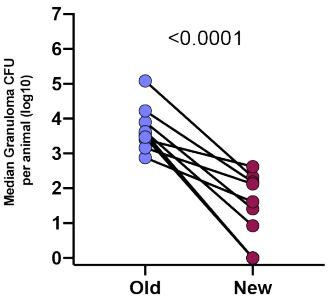


Figure S1. Granuloma architecture and CFU per granuloma decreases over time

(**A**) Architecture of macaque TB lung granuloma, where lymphocytes and macrophages are present in distinct regions. Immunohistochemistry and confocal microscopy were performed on a granuloma from an animal at 11 weeks post-Mtb infection to visualize localization of CD11c+ macrophages (cyan), CD3+ T cells (yellow), and CD20+ B cells (magenta). (B) Each column depicts the CFU for all granulomas of an individual macaque (N=88 macaques), ranging from 4 weeks to 17 weeks post-infection. Each dot represents a granuloma. Lines are at means (per animal) and different colors represent weeks post-infection. (C) CFU per granuloma decreases significantly starting at 10-11 weeks post-infection. Each dot represents the mean CFU per granuloma of an individual animal, with the x-axis indicating weeks post-infection at which necropsy was performed. Lines are at medians. Differences between time points were tested using Kruskal-Wallis test with Dunn's multiple comparison adjustment. (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.) (D) Total lung FDG activity (in log scale) measured by PET scans of each animal at 4, 8 and 10-weeks post-Mtb infection showing trajectories of lung inflammation. (E) Size of each granuloma measured by CT scans at 4, 8 and 10 weeks post-mtb infection. Early granulomas are those identified at 4 weeks post infection (in maroon) and late granulomas are those identified at 10 weeks post infection (in dark blue). (F) CFU per granuloma is shown for early detection (blue) and late detection (red) within each animal. Box plots lines represent the median, IQR and range Each dot represents a granuloma. (G) CFU is significantly lower in new granulomas within animals. Each dot (and line) represents the median CFU per granuloma of each animal. Statistics: paired t-test

Fig S2

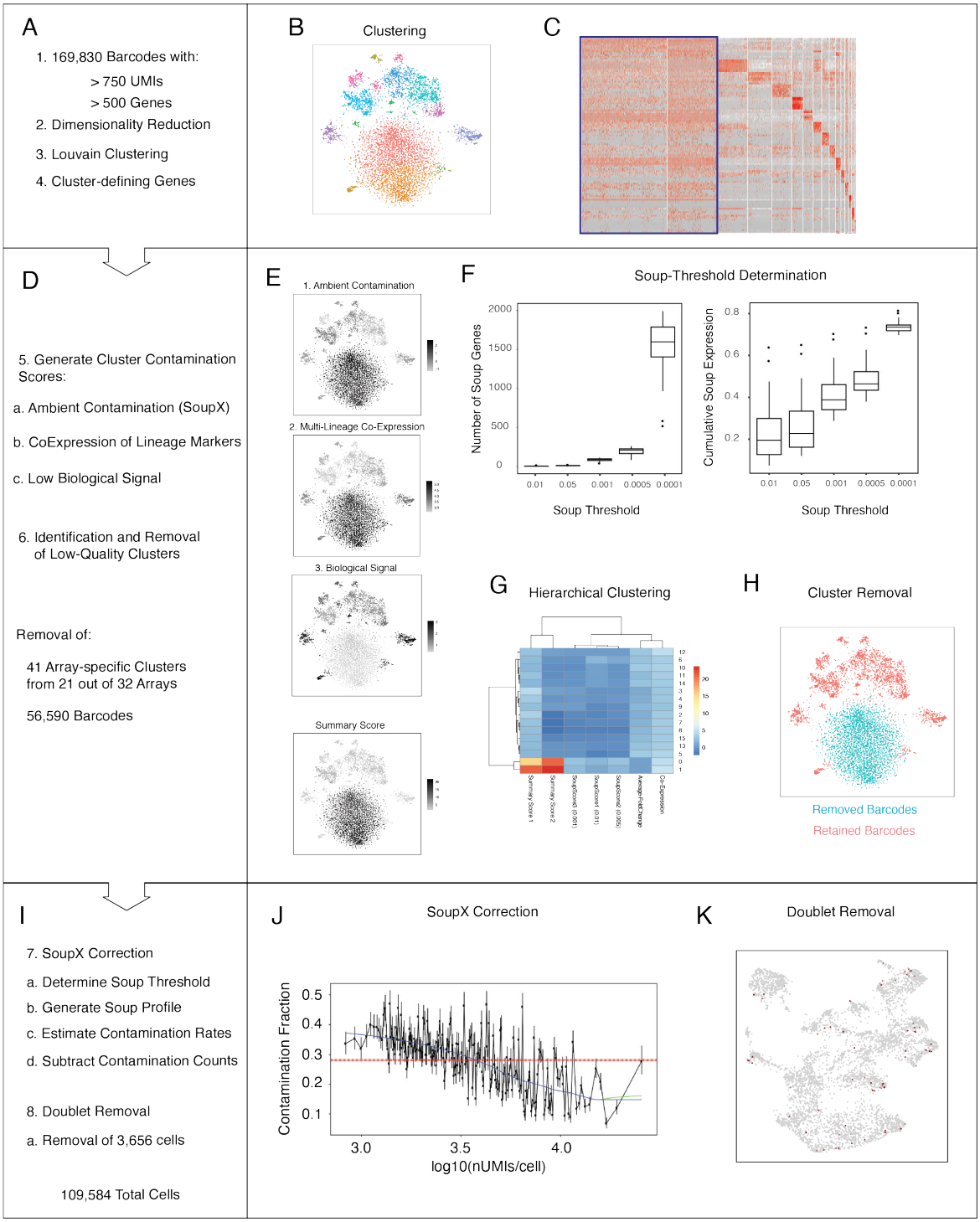


Figure S2: Sequencing, alignment and QC pipeline (see methods)

(A, D, I) Array-specific processing pipeline. **(B)** Array specific Louvain clustering (Resolution = 1.25). **(C)** Cluster-defining gene expression was determined within each array. **(E)** Overview of Cluster-Specific Summary Score. **(F)** Estimation of soup-thresholds for correction of ambient RNA contamination. Left: Relationship between soup-thresholds (x-axis) the number of soup defining genes detected for each array (y-axis). Right: Relationship between soup-thresholds (x-axis) and the cumulative proportion of soup-defining gene expression (y-axis). **(G)** Hierarchical clustering results used to identify and remove clusters defined by ambient contamination from each array. **(H)** t-SNE plot showing removal of clusters characterized as ambient RNA. **(J)** Estimation of array-specific contamination rates using SoupX. **(K)** Identification and removal of array-specific doublets.

Fig S3

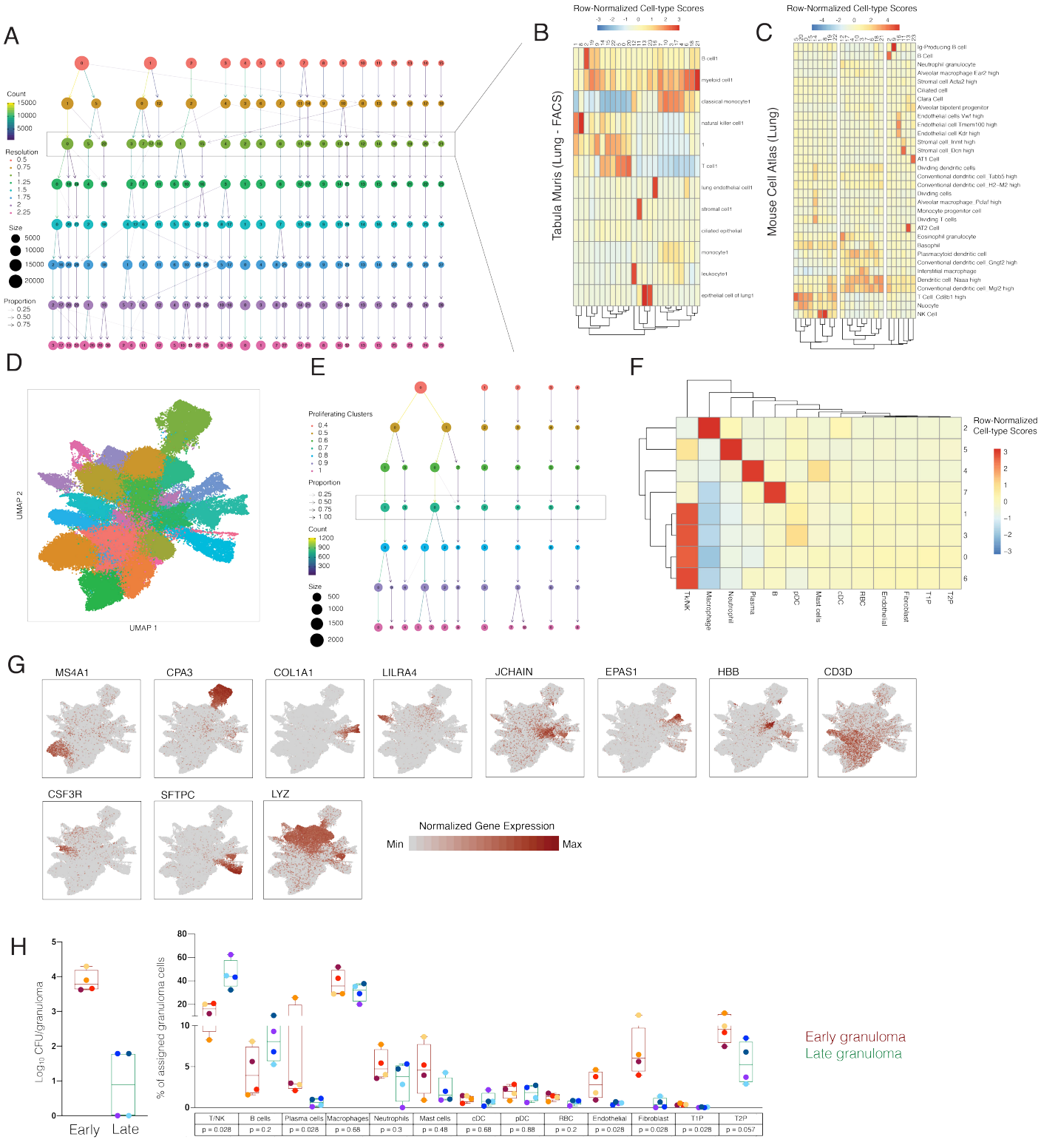

A
Count
15000
10000
5000
Resolution
0.5
0.75
1
1.25
1.5
1.75
2
2.25
Size
5000
10000
15000
20000
Proportion
0.25
0.50
0.75
B
Row-Normalized Cell-type Scores
Tabula Muris (Lung - FACS)
B cell1
myeloid cell1
classical monocyte1
natural killer cell1
T cell1
lung endothelial cell1
stromal cell1
ciliated epithelial
monocyte1
leukocyte1
epithelial cell of lung1
C
Row-Normalized Cell-type Scores
Mouse Cell Atlas (Lung)
Ig Producing B cell
B Cell
Neutrophil granulocyte
Alveolar macrophage Ear2 high
Stromal cell Acta2 high
Ciliated cell
Clara Cell
Alveolar bipotent progenitor
Endothelial cells Vwf high
Endothelial cell Tmem100 high
Endothelial cell Kdr high
Stromal cell Irxnl high
Stromal cell Dcn high
AT1 Cell
Dividing dendritic cells
Conventional dendritic cell Tubb5 high
Conventional dendritic cell H2-M2 high
Dividing cells
Alveolar macrophage Pclaf high
Monocyte progenitor cell
Dividing T cells
AT2 Cell
Eosinophil granulocyte
Basophil
Plasmacytoid dendritic cell
Conventional dendritic cell Gngt2 high
Interstitial macrophage
Dendritic cell Naaa high
Conventional dendritic cell Mgl2 high
T Cell Cd8b1 high
Nuocyte
NK Cell
D
UMAP 2
UMAP 1
E
Proliferating Clusters
0.4
0.5
0.6
0.7
0.8
0.9
1
Proportion
0.25
0.50
0.75
1.00
Count
1200
900
600
300
Size
500
1000
1500
2000
F
Row-Normalized Cell-type Scores
T/NK
Macrophage
Neutrophil
Plasma
B
pDC
Mast cells
cDC
RBC
Endothelial
Fibroblast
T1P
T2P
G
MS4A1
CPA3
COL1A1
LILRA4
JCHAIN
EPAS1
HBB
CD3D
CSF3R
SFTPC
LYZ
Normalized Gene Expression
Min
Max
H
Log10 CFU/granuloma
Early
Late
% of assigned granuloma cells
T/NK p = 0.028
B cells p = 0.2
Plasma cells p = 0.028
Macrophages p = 0.68
Neutrophils p = 0.3
Mast cells p = 0.48
cDC p = 0.68
pDC p = 0.88
RBC p = 0.2
Endothelial p = 0.028
Fibroblast p = 0.028
T1P p = 0.028
T2P p = 0.057
Early granuloma
Late granuloma

Figure S3: Identification of Canonical Cell Types.

(A) Waterfall plot showing stability of cell-type clustersat multiple clustering resolutions. Boxed row (resolution=1.00) selected for downstream analysis.

(B, C) Distribution of lung cell-type signatures obtained from the Tabula muris (B) and Mouse cell (C) atlas. **(D**) UMAP plot of 109,584 cells colored by Louvain clusters (resolution = 1.00). **(E)** Waterfall plot showing the stability of sub-clustering analysis of 3,123 cells with a proliferating gene signature. **(F)** Distribution of canonical cell type signatures across subclusters of proliferating cells. **(G)** Expression levels of cluster-defining genes overlaid on UMAP plot in panel 2A. (**H) Left:** CFU per granuloma based on the timing of detection by PET CT scan in one animal : 4017. **Right:** Difference in granuloma proportional composition of cell type clusters between early (maroon box plot) and late granulomas (green) within an animal (4017). Each granuloma is coloured. Statistics: Mann Whitney U. p values are presented in boxes. Box plot showing median, IQR and range; each dot represents a granuloma.

Fig S4

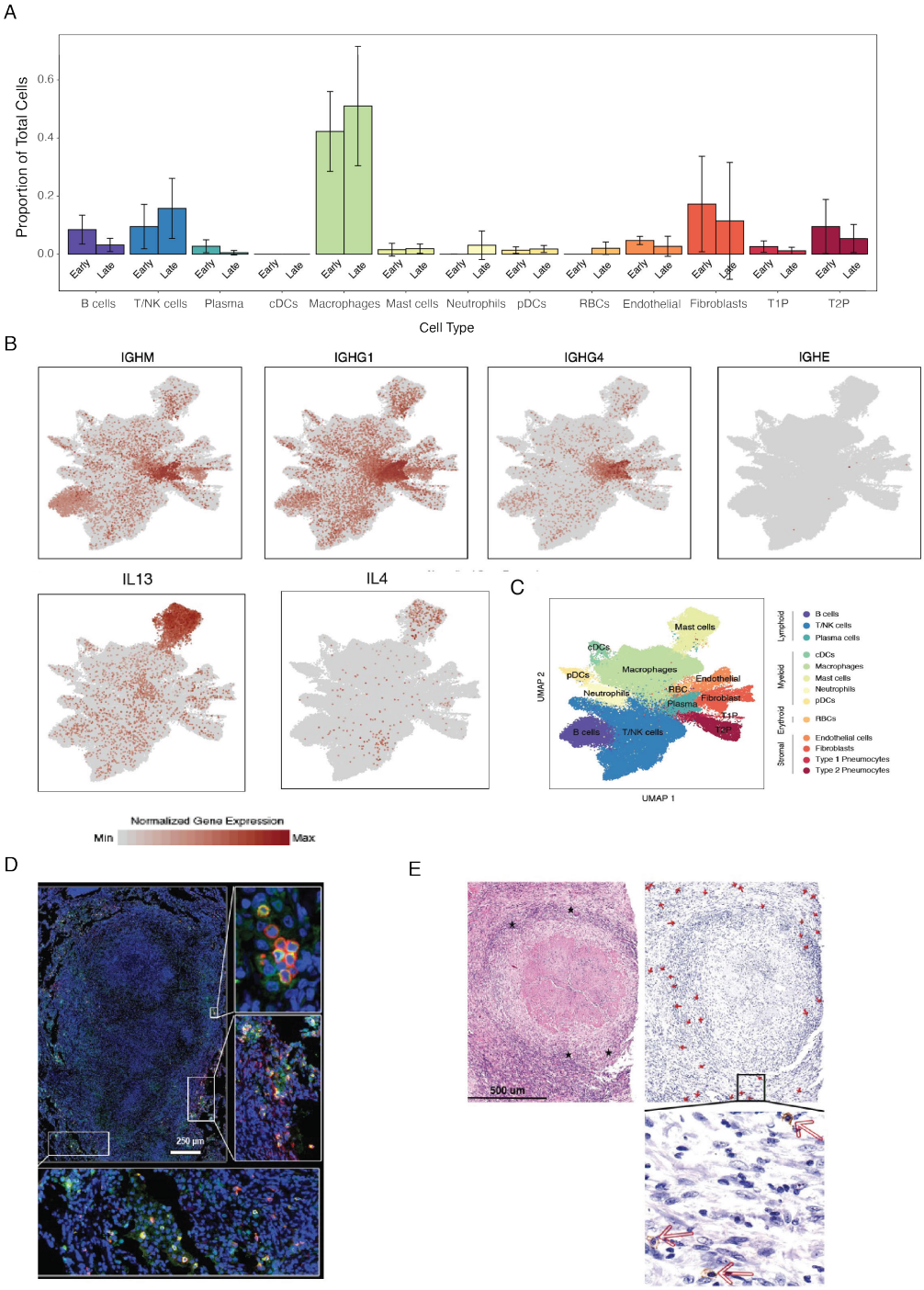


Figure S4. Cell type confirmation and Expression of selected functional transcripts. (A) Proportion of cell types in granulomas from bulk sequencing of 6 early and 6 late granulomas to confirm the trend seen in scRNAseq. **(B)** UMAP plot of 109,584 cells from 26 granulomas colored by identities of 13 generic cell types. **(C)** Expression levels of select functional genes overlaid on UMAP plot of 109,584 cells. **(D)** Detection of mast cells in a 10-week NHP granuloma using immunohistochemistry, staining for tryptase (green) and c-kit (CD117)(red). **(E)** Detection of mast cells in a human lung granuloma. Hematoxylin and eosin stain and immunohistochemistry with multinucleated giant cells (stars, (top left) and c-kit (CD117) staining (indicated by arrows, top and bottom right).

Fig S5

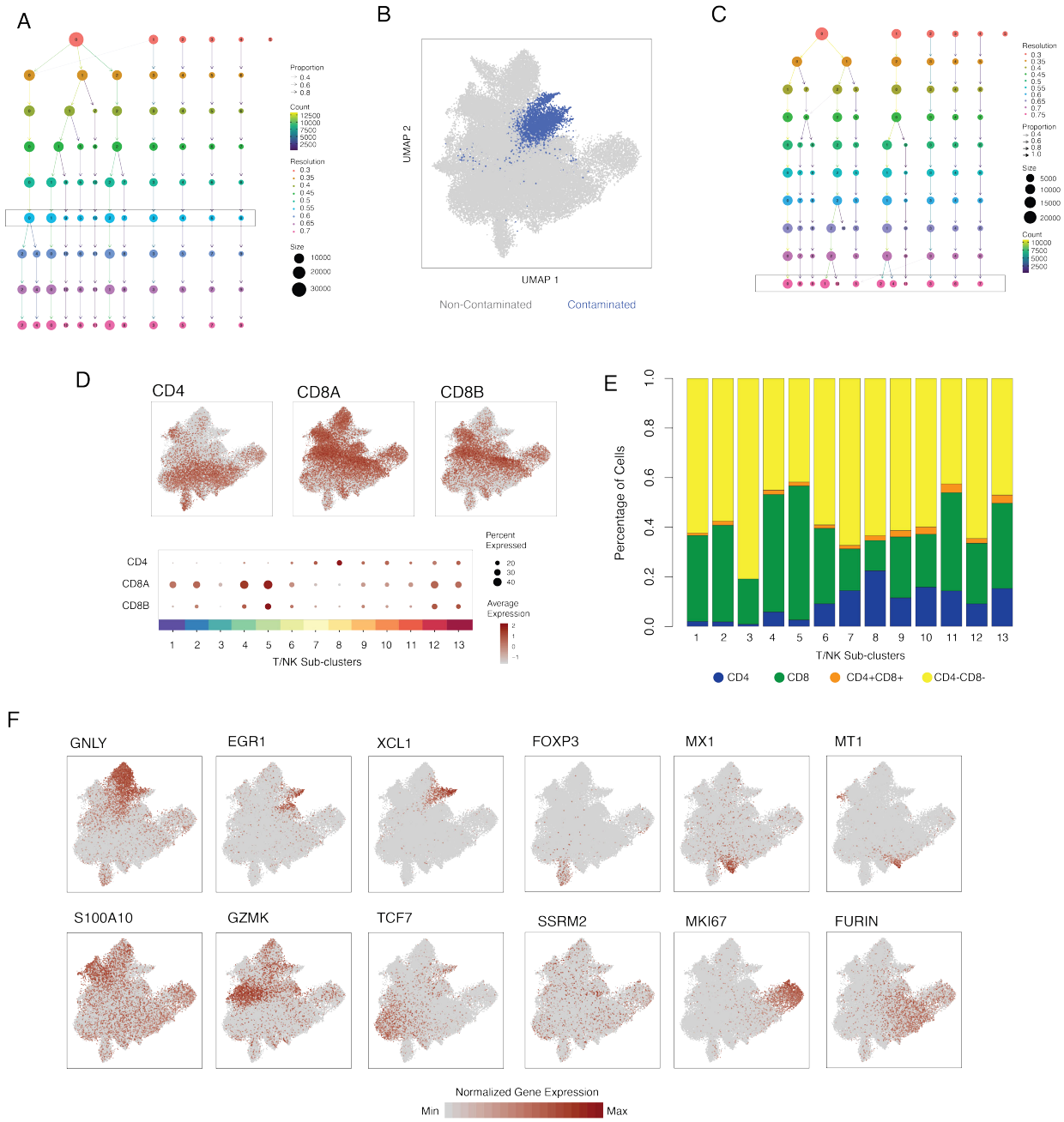
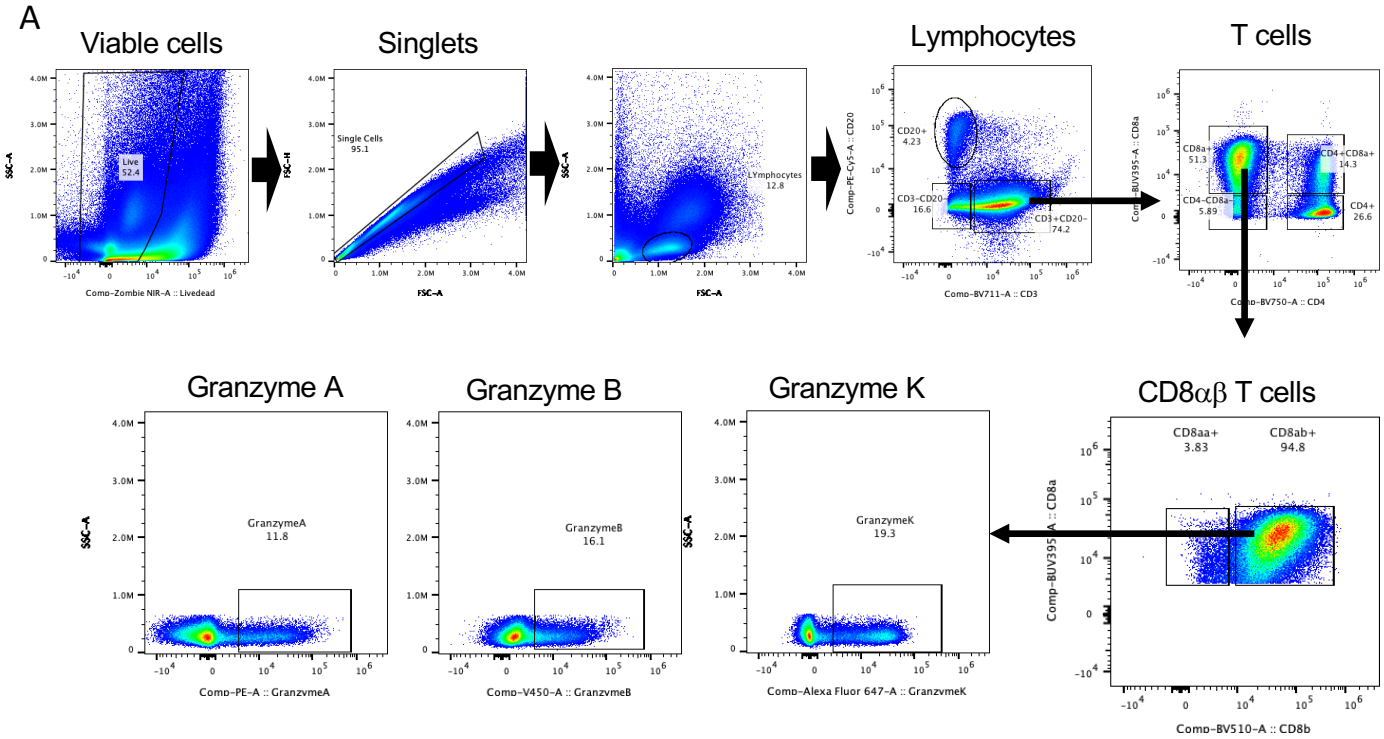


Figure S5. Sub-clustering and phenotypic identification of T/NK cell populations.

(A) Waterfall plot showing the stability of T/NK cell sub-clustering. Boxed row (resolution=0.55) selected for downstream analysis. **(B)** UMAP plot of 44,766 T/NK cells with a sub-cluster of 3,544 T/NK cells defined by residual contamination highlighted (blue). **(C)** Waterfall plot showing the stability of T/NK cell sub-clustering following removal of contaminated T cell sub-cluster. Boxed row (resolution=0.75) selected for downstream analysis. **(D)** T/NK subclustering UMAP overlaid with normalized gene expression for CD4, CD8A, and CD8B (top). Expression of these genes across 13 sub-clusters (bottom) where color intensity corresponds to level of gene expression and size of dots represents the percent of cells with non-zero expression in each cluster. **(E)** Frequency of expression of *CD4* (blue), *CD8A* and/ *CD8B* (green), *CD4* and *CD8A/B* (orange) or no expression of *CD4/CD8A/B* (yellow) across 13 T/NK cell subclusters. **(F)** UMAP plots overlaid with normalized expression levels for selected T/NK cell subcluster-defining genes.

Fig S6



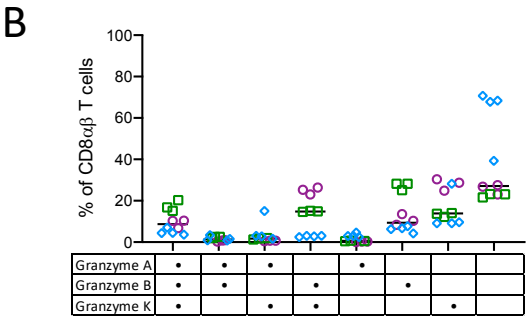


Figure S6: Flow cytometry confirmation of cytotoxic molecules in TB granulomas

(A) Gating Tree showing identification CD8ab T cells in lung granuloma samples and population of Granzyme A, Granzyme B and Granzyme K + CD8abT cells. **(B)** Frequency of CD8ab T cells in lung granulomas making one or more (two , three) types of Granzymes (A, B or K). Each symbol is a granuloma and each colour identifies an animal. This data supports different types of granzyme producing cytotoxic cells identified in scRNAseq.

Fig S7

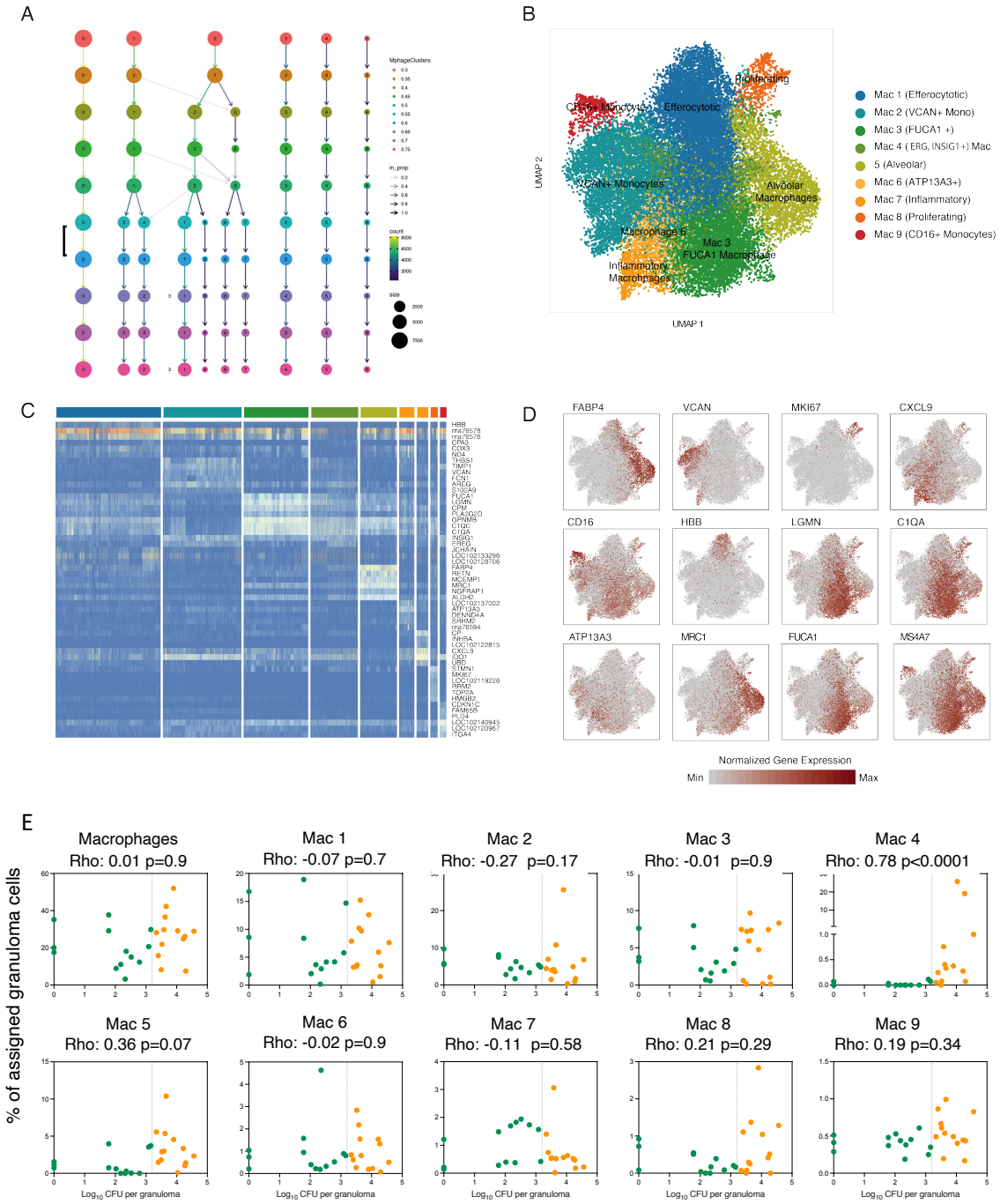


Figure S7 Macrophage heterogeneity in Mtb granulomas

(A) Waterfall plot showing the stability of macrophage sub-clusters. Boxed row (resolution=0.55) selected for downstream analysis. **(B)** UMAP plot of 27,670 macrophage cluster colored by phenotypes. **(C)** Cluster-defining genes across macrophage subclusters. **(D)** Macrophage subcluster-defining genes overlaid on macrophage plot in panel B. **(E)** Significant correlations between proportion of Macrophage subclusters with bacterial burden of individual granulomas (Log_{10} CFU per granuloma) using non-parametric Spearman's rho correlation test. Color indicated binned granuloma bacterial burden: low (green) and high (orange).

Fig S8

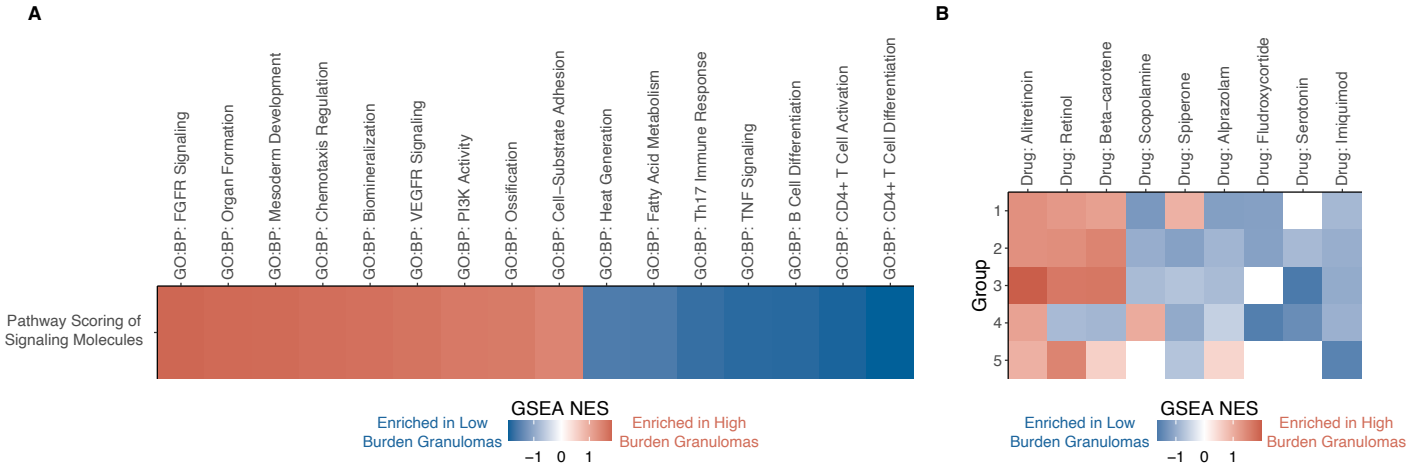


Figure S8 Transcriptomic pathways associated with granuloma burden.

(A) Pathways enriched in signaling molecules associated with high vs. low granuloma burden. Signaling molecules were ranked according to their log(fold-change in high vs. low burden granulomas) as input to GSEA. **(B)** Drugs with targets enriched in signaling molecules associated with each cell type group.