Single-cell and spatial profiling reveals a role for tuberculosis-induced myofibroblasts in the immunopathology of infected lungs

Ian M. Mbano[1,2‡], Nuo Liu[3,4,5,11‡], Marc H. Wardsworth II[3,4,5], Mark J. Chambers[1,2], Thabo Mpotje[1,2], Osaretin E. Asowata[1,2], Sarah Nyquist[3,4,6,7], Kievershen Nargan[1], Duran Ramsuran[1], Farina Karim[1,2], Travis K. Hughes[4,5,8,9], Joshua D. Bromley[3,4,5,10], Robert Krause[1,2], Threnesan Naidoo[1,2], Liku B. Tezera[11,13], Michaela T. Reichmann[11,12], Sharie Keanne Ganchua[17], Henrik N. Kløverpris[1,2,14], Kaylesh J. Dullabh[2], Rajhmun Madansein[2], Sergio Triana[3,4,5], Adrie J. C. Steyn[1,2,15], Bonnie Berger[8,16], Mohlopheni J Marakalala[1,2,13], Sarah M. Fortune[4,19], JoAnne L. Flynn[17,18,], Paul Elkington[11,12,†], Alex K. Shalek[3,4,5,†], Alasdair Leslie[1,2,13,†,*]

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

[2]School of Laboratory Medicine and Medical Sciences, College of Health Sciences, University of KwaZulu-Natal, Durban, South Africa

[3]Institute for Medical Engineering & Science, Department of Chemistry, and Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[4]Ragon Institute of MGH, MIT, and Harvard, Cambridge, MA 02139, USA

[5]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

[6]Program in Computational & Systems Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[7]Computer Science & Artificial Intelligence Lab, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[8]Program in Health Sciences & Technology, Harvard Medical School & Massachusetts Institute of Technology, Boston, MA 02115, USA

[9]Program in Immunology, Harvard Medical School, Boston, MA 02115, USA

[10]Microbiology Graduate Program, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[11]NIHR Biomedical Research Centre, School of Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, United Kingdom

[12]Institute for Life Sciences, University of Southampton, Southampton, United Kingdom

[13]Department of Infection and Immunity, University College London, UK

[14]Department of Immunology and Microbiology, University of Copenhagen, Denmark

[15]Department of Microbiology, Centre for AIDS Research and Free Radical Biology, University of Alabama at Birmingham, Birmingham, United States

[16]Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

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

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

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

‡ Contributed equally

† Senior authors

*Correspondence: Alasdair Leslie, al.leslie@ahri.org

Abstract

Tuberculosis (TB) remains a global health challenge, causing ~1.3 million deaths annually. Pulmonary TB often leads to lung destruction and fibrosis, yet the cellular drivers of human TB immunopathology remain poorly defined due to limited access to relevant tissues and differences from animal models. We performed single-cell RNA-sequencing and spatial transcriptomics on lung tissues from TB-infected and TB-negative individuals in a highly endemic South African region. We identified 30 distinct immune, parenchymal, and stromal cell subsets, with several linked to TB pathology and corroborated through immunohistochemistry, flow cytometry, and an independent human lymph node granuloma cohort. Among these, we found that fibroblasts were a major driver of intercellular interaction in both active TB granuloma and TB-diseased lung tissue. In particular, the *MMP1*$^{+}$*CXCL5*$^{+}$ fibroblast subset, which expressed elevated levels of a myofibroblast-like gene signature, was associated with severe disease, as well as higher bacterial burden in non-human primate granulomas. Network analyses revealed crosstalk between *MMP1*$^{+}$*CXCL5*$^{+}$ fibroblasts and *SPP1*$^{+}$ macrophages within the granuloma cuff, which has been reported in other disease contexts, and may play an important role in the immunopathology of TB. Overall, our findings highlight previously unappreciated cell populations and interactions that may be targetable in host-directed TB therapies.

INTRODUCTION

Tuberculosis (TB), caused by infection with *Mycobacterium tuberculosis* (*M.tb*), remains a global epidemic, with approximately 10.6 million new cases and 1.3 million deaths annually *(1, 2)*. The development of highly effective anti-TB drugs and programmatic improvements led to global cure rates of approximately 85% in drug-susceptible TB from 1995 to 2015, as well as reduced mortality rates *(3)*. Unfortunately, however, mortality remains persistently high *(4)*, highlighting the need for improved interventions.

M.tb infection occurs primarily in the lung where interactions between host cells and the pathogen typically result in the formation of a granuloma – an aggregation of infected myeloid cells, usually surrounded by an inner ring of macrophages and an outer cuff of lymphoid cells. This specialized immunological niche is highly heterogeneous in its overall cellular makeup, with the composition of each lesion independently influencing bacterial growth and disease progression *(5, 6)*. In progressive TB, extensive lung extracellular matrix (ECM) remodeling via both matrix destruction and fibrosis leads to the formation of lung cavities that facilitate transmission *(7)*. This ECM remodeling also increases the risk of post tuberculosis lung disease (PTLD), resulting in high rates of recurrent TB infection and mortality – even after successful eradication of initial infection *(8, 9)*. Although some features of the immunopathology of TB infection that lead to PTLD are known, including granuloma formation, cytokine production, hypoxia-inducible factors, and production of matrix

metalloproteinases (MMPs), the exact mechanisms remains unclear *(8)*. While animal models of TB disease—from zebrafish to non-human primates—have provided valuable insights into aspects of these processes, they do not fully recapitulate human pathology *(10)*. Critically, these models generally reflect primary infection, often without the cavities and the extensive ECM remodeling observed in human TB, and fail to capture the development of chronic secondary TB disease that arises in humans *(11)*. Consequently, the key features and cellular drivers of immunopathology in human TB remain poorly understood.

The advent of high-throughput single-cell RNA-sequencing (scRNA-seq) has transformed our ability to analyze the cellular makeup of complex tissues and phenotypic changes associated with disease *(12)*. For example, application of this technology to study idiopathic pulmonary fibrosis (IPF) – a lung disease characterized by dysregulated ECM turnover – identified aberrant basal-like cells, peribronchiolar endothelial cells, *SPP1*$^{+}$ macrophages, and myofibroblasts as key drivers of pulmonary tissue remodeling, suggesting new strategies to combat the disease *(13)*. Similar characterization of the cell types and states involved in the immunopathogenesis of human TB and PTLD could potentially help uncover effective targets for host directed therapies (HDT) *(14)*.

Here, we applied scRNA-seq and spatial transcriptomics to human TB-diseased lung tissues and TB-negative controls to examine the cellular and molecular features of TB lung disease and investigate mediators of immunopathology. Overall, we identified depletion of most macrophage subsets and an enrichment of fibroblast and neutrophil subsets in TB-diseased lungs, consistent with altered fibrotic and pro-inflammatory activity. We validated these observations with bulk RNA-sequencing data of lymph node TB granuloma from a well-characterized cohort of treatment-naïve, culture-confirmed TB patients *(15)*. To further contextualize specific disease-associated cell subsets, we integrated our data with those from the Human Lung Cell Atlas (HLCA) and Non-Human Primate (NHP) lung TB granulomas *(16, 17)*. This enabled us to uncover a putative central role for fibroblast subsets – including a *MMP1*$^{+}$*CXCL5*$^{+}$ fibroblast cluster expressing a myofibroblast-like gene module – in TB immunopathology, where we further evidenced via flow cytometry and immunohistochemistry. Through cell network analyses on the single-cell data, we found that these cells appear to coordinate their activities with macrophages, including a *SPP1*$^{+}$ subset not previously implicated in TB biology that was observed to be coresident in immunohistochemical stainings of human lung TB granuloma. Moreover, these two subsets were co-inducible by a standard skin challenge of TB patients with *M.tb* derived antigen (Tuberculin), and analyses of spatial transcriptomics data from an independent cohort of TB patients showed colocalization of both this myofibroblast signature and the *SPP1*$^{+}$ macrophage signal to lung TB granuloma cuffs. Overall, our data reveal key cellular subsets and pathways that could inform next-generation HDTs and provide an essential reference for the community.

RESULTS

Cellular composition of human TB-infected lung tissue

Fresh, TB-diseased human lung tissue pieces were obtained from 9 participants (7 HIV^+ TB; 2 HIV^- TB) enrolled in the African Health Research Institute (AHRI) lung cohort study (**Fig. 1A**). All participants underwent TB treatment after initial diagnosis but had subsequent lung resection surgery to treat complications consistent with PTLD, including hemoptysis and bronchiectasis (**Table S1**). As a control, TB-negative lung samples were obtained from the healthy tissue margins of four surgically resected lung tumors (1 HIV^+ cancer control; 3 HIV^- cancer control). All participants, irrespective of TB status, received prophylactic anti-TB treatment prior to surgery. Tissue pieces were washed thoroughly and homogenized into a single-cell suspension via mechanical and enzymatic digestion using an optimized protocol in BSL3 containment *(18)*. Lung cells were then processed and sequenced following the Seq-Well S^3 protocol as described previously to obtain our scRNA-seq dataset *(19)*. An additional 30 samples were obtained from different participants in the same TB lung cohort and profiled using the 10x Visium Spatial Gene Expression platform (**Fig. 1B, Data File 5**). As above, all participants received TB-drug treatment prior to surgery. Fresh tissue pieces were removed from resected lung tissue and preserved using standard formalin-fixed paraffin embedded (FFPE) procedures, followed by Visium v2 chemistry protocols with paired hematoxylin and eosin (H&E) staining to generate reference images (**Fig. S1A**; **Methods**) *(20)*. 21 samples (10 HIV^+, 11 HIV^-) were derived from subjects with active microbiologically confirmed TB, termed "current TB". The remaining samples (5 HIV^+, 4 HIV^-) termed "post-TB", were obtained from individuals in whom bacterial load was no longer detectable from bronchoalveolar lavage (BAL) TB culture. This spatial dataset contains both samples with TB lung granulomas and samples with inducible bronchus-associated lymphoid tissues (iBALTs) or lung draining lymph nodes (LNs), which are considered as less severe pathological states. For each granuloma sample, pathological grading and manual annotation of the granuloma structures on the H&E image were performed by an expert TB pathologist to enable better disease contextualization (**Fig. S1B**, **Data File S1**).

After quality control of the scRNA-seq data, we retained 19,632 high-quality single-cell profiles from the homogenized lung tissues. Neighborhood-based clustering revealed 16 canonical cell types. Further sub-clustering of high abundance populations resulted in a total of 30 phenotypically distinct immune, parenchymal, and stromal subsets (**Fig. 1A**, **Data Set 1A-E, Fig. S2A-F, Data Set 2A-B; Methods**). The fractional representation of cells per participant and clinical characteristic varied between clusters reflecting biological heterogeneity between patients, TB disease states, and potentially anatomical sampling location, though our data are limited with respect to the latter (**Fig. 2A**, **Table S1**). Most cells derived from HIV^- TB samples, and while most of the clusters contained cells from the majority of patients, we observed substantial inter-patient variability in cell numbers (**Fig. 2B, Table S2**). Canonical cell type markers and genes differentially expressed between clusters were examined for manual annotation (**Fig. 2C; Methods**). Notably, we found large populations of neutrophils, which are captured by Seq-Well S^3 but often underrepresented by other scRNA-seq technologies due to their fragility *(19)*. Overall, observed clusters closely mirrored those seen in a scRNA-seq characterization of lung tissue from idiopathic pulmonary fibrosis (IPF) patients and healthy donors *(21)*.

Next, we looked for evidence of differential abundance by comparing the fractional representation of each cell type per donor between the TB-diseased and TB-negative lung groups, irrespective of HIV status (**Fig. 2D**). There were pronounced shifts in the frequency of most cell types between the TB-diseased and TB-negative groups, including an expansion of neutrophils in the TB-diseased group, consistent with several human studies linking neutrophil recruitment with TB lung pathology *(22–24)*. We also found an increased frequency of mast and plasma B cells in TB-diseased tissue, supporting findings from recent scRNA-seq studies of NHP models where both cell populations were expanded in TB granuloma with higher bacterial burden *(16)*. In addition to these immune cell populations, fibroblasts were enriched in TB-diseased lung tissue. Conversely, in TB-diseased lung samples, we found a decrease in the proportions of macrophages, the cell type targeted by and primarily responsible for killing bacilli, and CD8 T cells, thought to contribute to *M.tb* control *(25, 26)*. Although we detected significant changes between the abundance of these cell subsets using a Fisher's exact test, given high inter-patient variability and limited sample numbers, we were underpowered to determine significance using a Dirichlet-multinomial regression or a Wilcoxon test *(27)*.

Specific innate immune cell subclusters are associated with TB-diseased human lung tissue.

Given the limited participant numbers that compose our scRNA-seq dataset, we leveraged this data to impute cell type abundances in both the current- and post-TB lung samples in our spatial transcriptomics cohort. This allowed us to better understand the phenotypic shifts associated with TB disease, and select relevant single-cell subclusters for further characterization. This strategy drove us to focus on neutrophils, macrophages, monocytes, and fibroblasts, whose abundances also showed the most dramatic shifts between the TB and control samples in the scRNA-seq data **(Fig 2D)**.

Neutrophil subclusters

Neutrophils play a crucial role in the innate immune system and are quickly recruited as a first line of defense against bacterial infections. They are suggested to have immunoregulatory functions in TB granulomas in NHP models *(28)*; however, their role in the immunopathogenesis of human TB has been contentious and less well understood *(29)*.

Neutrophils were highly enriched in TB-diseased lung tissue in our single-cell dataset (**Fig. 2D**) and associated with three distinct subclusters (termed "Pro-inflammatory Neutrophils", "Activated Neutrophils", and "Heat-Shock (*HSP*) Neutrophils"). Both Activated Neutrophil and Pro-inflammatory Neutrophil subclusters expressed markers genes associated with IFN-γ and TNF-α signaling – critical responses linked to inflammation and immune activation in TB disease *(30–32)* (**Fig. S2A-D, Data File S2**). Activated Neutrophils were annotated by their high expression of neutrophil activation markers, including *IL1RN,* and *IL1B* and *IL8,* inflammatory cytokines involved in neutrophil recruitment (**Fig. S2B-D**) *(33, 34)*. They also expressed *GBP1 and GBP5,* genes involved in a previously-described blood neutrophil transcriptional signature used to diagnose pulmonary TB *(35)*. Pro-inflammatory Neutrophils, in contrast, highly expressed high levels of *MMP9*, *CST7* and *LDHA* (**Fig. S2B-D**). MMP9 is a proteinase involved in the degradation of extracellular matrix that

is strongly associated with TB granuloma *(36)*; CST7 (Cystatin F) is a neutrophil marker of acute inflammation *(37)*; and, *LDHA* encodes lactate dehydrogenase, which enhances neutrophil migration and activity, and is highly elevated in hypoxic lung TB granuloma in animals *(38, 39)*. Pro-inflammatory Neutrophils also highly expressed, *FKBP5* and *CEBPD*, both implicated in an immunometabolic network predictive of TB progression *(40)*; and, *VEGFA*, *PLAUR*, *TPM4*, and *CD44*, which are involved in neutrophil recruitment and lymphangiogenesis during inflammation *(41–44)* (**Data File S2**). The remaining small subcluster of neutrophils, marked by high expression of heat-shock protein genes ("*HSP* Neutrophils"), was also elevated in TB diseased lungs, which is notable given that heat shock proteins expression by neutrophils can trigger proinflammatory response in macrophages *(45, 46)* (**Fig. S2A-E**).

Given the small samples size and high HIV prevalence in our scRNA-seq dataset, we examined neutrophils in the spatial cohort to understand the link between neutrophils and TB disease, running cell type deconvolution using the scRNA-seq cohort as reference and imputing individual cell type abundances (**Methods**). Within granuloma structures, neutrophil abundance was significantly higher in sample from current-TB infections than those from post-TB infections, consistent with the recruitment of this cell type to the granuloma during active disease (**Fig. S2F**). Interestingly, however, neutrophils were more abundant in non-granuloma tissues (e.g. iBALT, LN) compared to granuloma, though the difference was less pronounced in current-TB. This may reflect the involvement of neutrophils in tissue remodeling and chronic inflammation associated with both active TB and post TB lung disease *(49)*.

To further test the association between neutrophil subsets and TB disease, we quantified the expression of each subset's top markers genes in an independent bulk RNA-seq dataset generated from laser-captured human lymph node (LN) TB granuloma, in which all patients were HIV negative (**Methods; Data and materials availability**) *(15)*. Importantly, these LN were excised prior to TB therapy initiation and contained viable *M.tb* bacilli. We found that 7 of the top 10 unique marker genes associated with the Activated Neutrophils were significantly upregulated in LN granuloma compared to non-infected LN controls (**Data File S3**), as well as the Pro-inflammatory markers *CEBPD* and *LDHA* (**Data File S3**). We note that although LNs are common sites of extrapulmonary TB, LN granulomas have functional and structural differences from those found in the lung, which may contribute to differences in expression levels of these marker genes *(47, 48)*.

Monocyte and Macrophage subclusters

Macrophages are necessary to control TB disease, but also provide a niche for bacterial growth and survival *(50)*. In addition, they have been implicated in pulmonary remodeling, with reported roles in both promoting and inhibiting pathology *(51)*. Tissue infiltrating monocytes, meanwhile, provide a source for macrophage differentiation and are key players in inflammatory response and bacterial persistence *(52)*. Hence, understanding the functional differences in monocytes and macrophages between the TB-diseased and TB-negative controls could provide insights into understanding TB immunopathology. In aggregate, macrophages were significantly decreased in TB-diseased lung tissue, and monocytes were decreased albeit not significantly (**Fig. 2D**).

Sub-clustering of 8,313 macrophages/monocytes single-cell transcriptomes generated ten distinct subclusters, which we annotated manually based on marker genes (**Fig. 3A-C, Data Set 2A**). Alveolar Macrophages (*INHBA*$^{+}$*FABP4*$^{+}$*MARCO*$^{+}$) and a subcluster we termed "Heat-Shock (*HSP*) Macrophages" were significantly reduced in TB diseased lung tissue compared to TB-negative lungs. The former may reflect the loss of normal lung alveolar structure observed in TB diseased lung tissue. Upregulated proteins in *HSP* Macrophages included those encoding for Hsp70 family proteins (e.g., *HSPA1A, HSPA1B, HSPA6*, and *HSPA8*), which are known to modulate NF-κB mediated release of pro-inflammatory cytokines from alveolar macrophages in pulmonary TB *(53, 54)*. In addition, when tested separately, all monocyte subsets were significantly reduced in TB disease, possibly due to rapid transition to macrophage phenotypes in the proinflammatory environment of the diseased lung *(55)*. The remaining three macrophage subsets (defined by *ARL4C/EMP1*; *LGMN/SEPP1*; and *SPP1/CHI3L1*) were higher in TB diseased lung, but subtly so (**Fig. 3D**). Therefore, to explore the potential skewing of macrophage subsets further, we performed cluster-free differential abundance testing using Milo, which models cellular states as overlapping neighborhoods on k-nearest neighbor graphs representing the similarities between single-cell profiles *(56)*. This analysis highlighted the underrepresentation of alveolar macrophages in TB diseased lung tissue, as this was the only subcluster with its phenotypic neighborhoods depleted among TB disease samples (**Data set 2C-D**). In contrast, although present at low frequency, *ARL4C*$^{+}$*EMP1*$^{+}$, *LGMN*$^{+}$*SEPP1*$^{+}$ and *SPP1*$^{+}$*CHI3L1*$^{+}$ Macrophages were all significantly associated with TB diseased lung. Marker genes enriched in *ARL4C*$^{+}$*EMP1*$^{+}$ Macrophages did not obviously associate with published functional annotations, but included *GPR138* which favors *M.tb* replication in macrophages *(57)*. *LGMN*$^{+}$*SEPP1*$^{+}$ Macrophages were enriched for lipid metabolism activities, while *SPP1* encodes for osteopontin, a known macrophage attractant, which has been associated with granulomatous diseases and is upregulated in *M.tb.* infection *(58, 59)* (**Data File S2**).

In our spatial cohort, we observed a higher abundance of both macrophages and monocytes in current-TB compared to post-TB, consistent with continuous recruitment of myeloid cells during active disease (**Fig. 3E left, Fig. 3F left**). As with neutrophils, macrophages were more abundant in granuloma from current-TB samples compared to post-TB. Additionally, in current-TB macrophages were more abundant in granuloma compared to non-granuloma tissue while the opposite was true for monocytes, which may be explained by maturation into macrophages within this environment. (**Fig. 3E right**). HIV co-infection was associated with increases in both macrophage and monocyte abundance in current-TB samples, and increases in monocyte abundance in post-TB samples, suggesting a potential effect of HIV co-infection on myeloid populations during both active TB and PTLD (**Fig. 3F**). In active TB samples, HIV infection may lead to more macrophages to compensate for the loss of CD4^{+} T cells, which are important for adaptive immune responses against *M.*tb *(60)*. The decrease in macrophages associated with HIV-coinfection in post-TB samples, however, might reflect impaired monocyte differentiation, or persistent depletion of macrophage precursors *(61, 62)*.

To further contextualize monocyte/macrophage subclusters in human TB granuloma, we similarly assessed expression of subcluster marker genes in the LN dataset described above. We found the strongest signal for

the *SPP1*$^{+}$*CHI3L1*$^{+}$ Macrophages, where 5/10 of the top markers of the subcluster were significantly upregulated in human LN TB, over 5-fold in the case of *SPP1* and *FN1* (**Data File S3**). Bulk gene expression deconvolution of this data supported significant increase in the frequency of several populations in untreated TB granuloma compared to control LN, including the *SPP1*$^{+}$*CHI3L1*$^{+}$ Macrophage (**Fig. S3A**).

Finally, to investigate these myeloid subsets under control condition, we examined the expression of subcluster defining genes within single-cell data from TB-granuloma isolated from experimentally infected NHPs *(16)*. In this dataset, we identified overlapping gene signatures between most of the subclusters observed in our current study, including the *SPP1*$^{+}$*CHI3L1*$^{+}$ Macrophage population (*LILRB4, MMP9, PKM, MYOF, CHI3L1*) (**Data File S4**) *(16)*. Collectively, these data identify diverse myeloid subsets in TB diseased lung tissue and support a putative role for *SPP1*$^{+}$*CHI3L1*$^{+}$ Macrophages in TB immunopathology in humans and NHPs.

Single-cell analysis identifies TB-associated fibroblast populations.

Despite playing a prominent role in tissue remodeling in other lung diseases, there is limited understanding of how fibroblasts contribute to granuloma formation, immunopathology and protective TB immunity *(5, 63)*. In our spatial transcriptomics samples, fibroblast abundance was estimated to be higher in granuloma samples than iBALT/LN samples for both TB conditions, suggesting fibroblast involvement in long-term tissue remodeling and granuloma formation (**Fig. S3B**). Holding granuloma status or HIV status constant, we observed a higher abundance of fibroblasts in post-TB samples relative to current-TB samples, consistent with the role fibroblasts play in long-term tissue damage in post-TB lung disease (**Fig. S3B-C**). Further sub-clustering of the 1,627 fibroblasts in the scRNA-seq dataset revealed five distinct subclusters (**Fig. 4A; Fig. S3D-E**): *IL6*$^{+}$ *CCL2*$^{+}$ Fibroblasts, *SERPINE2*$^{+}$*COL1A1*$^{+}$ Fibroblasts, Heat-Shock (*HSP*) Fibroblasts, *COMP*$^{+}$*CILP*$^{+}$ Fibroblasts, and *MMP1*$^{+}$*CXCL5*$^{+}$ Fibroblasts (**Fig. 4B; Fig. S4A).** We note that most of the fibroblasts we recovered came from TB-diseased patients, a trend consistent with fibrotic change due to TB damage in the lung *(64)*. One of the subclusters, *MMP1*$^{+}$*CXCL5*$^{+}$ Fibroblasts, almost solely consists of cells from HIV^{-} TB patients, whereas the others were mostly occupied by cells from HIV^{+} TB patients, suggesting potentially different phenotypes for fibroblasts in HIV/TB coinfected patients versus patients with TB alone (**Fig. 4A; Fig. S3E**). This subcluster also had the strongest phenotypic shifts among fibroblast populations in a Milo analysis (**Data Set 2C-D**).

To better understand the phenotypic properties of these five subclusters, we further contextualized them against the existing literature by mapping them onto a trained reference model for lung stromal cell annotation *(65)* (**Fig. 4C**). The majority mapped strongly to adventitial fibroblasts, which are associated with pulmonary vascular remodeling in response to stress, including hypoxia and infection *(66)*. Although canonically associated with vascular beds, adventitial fibroblasts become highly migratory and invasive in response to activating signals, notably including Osteopontin (SPP1), and have been shown to drive tissue remodeling by inducing a proinflammatory/profibrotic phenotype in macrophages through IL-6 signaling, *(67, 68)*. The *MMP1*$^{+}$*CXCL5*$^{+}$ Fibroblast cluster, however, mapped primarily to the myofibroblast phenotype, followed by

the lipofibroblast phenotype. Myofibroblasts are involved in wound healing after tissue injury and can differentiate from recruited fibroblasts under mechanical stress, through the influence of cytokines like TGF-β, and epithelial-to-mesenchymal transition (EMT) *(69, 70)*. In addition, lipofibroblasts can differentiate into myofibroblasts during fibrosis *(71)*. Consistent with this, overrepresentation analysis (ORA) showed that the *MMP1*$^{+}$*CXCL5*$^{+}$ fibroblast markers were enriched among genes associated with EMT and myoblast differentiation (**Data File S2**).

To test the association between *MMP1*$^{+}$*CXCL5*$^{+}$ fibroblast markers and TB, we again examined the LN dataset and found that 5/10 top unique marker genes in the *MMP1*$^{+}$*CXCL5*$^{+}$ Fibroblast subcluster were upregulated in the LN TB data, including *MMP1, CA12, TDO2, POSTN,* and *COL12A1* (**Data File S3**). Interestingly, *CA12* plays a role in many biological processes, including preventing calcification *(72)*, an essential process in granuloma resolution *(73)*. In addition, this gene was found to co-express with *MMP1* and *CXCL5* in a subset of Cancer Associated Fibroblasts associated with poor clinical outcome *(74)*. We also observed a significant increase in the imputed frequency of *MMP1*$^{+}$*CXCL5*$^{+}$ Fibroblast in TB LN granuloma compared to control LNs via deconvolution of bulk RNA-seq profiles (**Fig. S3A**). Together these data suggest TB is associated with skewing of lung fibroblasts to phenotypes that overlap with known disease processes in the infected lung.

To confirm the presence of the *MMP1*$^{+}$*CXCL5*$^{+}$ phenotype via an orthogonal method, we stained sections of human lung from the same surgical cohort that contained distinct TB granuloma (5μm sections from two patients) for associated gene products of the *MMP1*$^{+}$*CXCL5*$^{+}$ subcluster: COL1, TDO2, MMP1, MMP3 and CA12, together with PI-15 and CTHRC-1, which were also significantly upregulated in this subcluster (**Fig. S4B**; **Data File S2**). COL1, a general fibroblast marker was expressed across lung tissue; the MMPs, which are secreted to facilitate ECM breakdown, has less strict localization; TDO2, CA12, PI-15 and CTHRC-1, meanwhile, were expressed higher in the granuloma compared to the surrounding tissue. These data support the presence of *MMP1*$^{+}$*CXCL5*$^{+}$ fibroblasts in TB diseased human lung and their localization with TB granuloma. It is worth noting, however, that some differences in fibroblast populations we observe between TB conditions may be exacerbated by the limited number of control lung samples and difficulties associated with extracting stromal cells from fresh tissues during single-cell isolation *(75)*.

Reference mapping to Human Lung Cell Atlas reveals distinct activities between TB-diseased and control fibroblasts

Given the limited recovery of fibroblasts from TB-negative controls, we next explored how the fibroblast subsets detected in TB diseased lung tissue relate to lung fibroblasts in published datasets. For this, we used the data from the Human Lung Cell Atlas (HLCA), which integrates 49 datasets from the human respiratory system, encompassing 2.4 million single cells, to generate consensus cell type annotations *(17)*. Using the HLCA as a reference, we confirmed the heterogenous immune and non-immune cell types present in our lung tissue samples (**Fig. 4D**). Via label transfer, we independently re-annotated our fibroblasts, observing high consistency with our original annotations (>95% fibroblasts were re-annotated as fibroblast/myofibroblast;

Methods). Combining cells mapped to fibroblasts/myofibroblasts in our data and the HLCA, we performed differential expression (DE) analysis between all TB-negative cells (mostly consisting of healthy cells from the HLCA reference) and our TB-diseased cells (**Fig. 4E**). We then ran gene set enrichment analysis (GSEA) on the resulting top DE genes using the MSigDB Hallmark database (**Fig. 4F**). This confirmed upregulation of EMT processes, thought to directly contribute to the fibroblast/myofibroblasts pool during fibrosis, in TB-diseased fibroblasts. In addition, oxidative phosphorylation was highly upregulated, consistent with alteration of metabolic activity in fibrotic lung disease *(76)*. Several enriched terms are related to inflammatory process, including TNF-α signaling, TGF-β signaling, IL2/IL6 signaling, were also observed, suggesting an overall elevated inflammatory response in the TB fibroblasts.

Identification of fibroblast gene modules associated to bacterial burden within TB granuloma

Having compared our fibroblast subclusters against public lung stromal datasets, we next examined how the cell states within these subclusters might shift with TB disease in an experimentally controlled setting. Given a paucity of TB-associated gene signatures, we opted to pursue unbiased gene module identification on the entire fibroblast population, applying a tool for weighted gene co-expression network analysis in high-dimensional single cell transcriptomics data (hdWGCNA) *(77)*. This yielded 7 gene modules with varying degrees of expression across the 5 clusters and disease states (named Fibroblast-M1-7; **Fig. 5A; Data File S5**). The Fibroblast-M1 module was highly enriched in the *MMP1*$^{+}$*CXCL5*$^{+}$ Fibroblast cluster (**Fig. 5B**). Top hub genes in this module included: *MMP1, CA12, CXCL5, CXCL13, TDO2, PDPN,* and *FAP*, showing a high degree of overlap with cluster markers for *MMP1*$^{+}$*CXCL5*$^{+}$ Fibroblast (**Fig. 4B**; **Fig. S4C-D**). There was also a clear overlap between *SERPINE2*$^{+}$*COL1A1*$^{+}$ Fibroblasts and Fibroblast-M2 module (**Fig. 4B, Fig. S4D**). Over-representation analysis showed the M1 module was highly enriched for known biological processes associated with immune cell migration and chemotaxis (“myeloid leukocyte migration”, ”granulocyte chemotaxis”) and control of ECM structure (“external encapsulating structure”, “collagen fibril organization”, “extracellular matrix disassembly”), as well as myofibroblast-related signatures *(69)* (“response to wounding”, “muscle tissue development”, “myoblast differentiation”, “response to mechanical stimulus”) (**Fig. 5C; Data File S2**). Hence, we refer to this M1 module, enriched in the *MMP1*$^{+}$*CXCL5*$^{+}$ Fibroblast population, as the “human TB-myofibroblast” module.

To investigate these modules in the context of defined pulmonary TB granuloma, we evaluated the expression of each module in fibroblasts from a well-controlled SIV-uninfected NHP TB granuloma dataset from Gideon *et al*. mentioned above *(16)*. This study collected data on positron emission tomography (PET)-tracked granulomas isolated at 4 weeks and 10 weeks post-infection, with each granuloma individually resected, homogenized and subjected to scRNA-seq, as well as quantification of total and viable *M.tb*. Interestingly, our human TB-myofibroblast module (M1)**,** M2 and M3 modules were significantly elevated in the week 4 granulomas, which contained higher *M.tb* burdens, compared to those at week 10 (**Fig. 5D, Fig. S4E**). These data suggest that the human TB-myofibroblast phenotype, in addition to other diverse fibroblasts, is likely present in untreated early TB lung granuloma, and that their frequency is associated with bacterial burden. Next, to further localize human TB-myofibroblast phenotype relative to granuloma, we evaluated the

expression of this module in fibroblasts in an independent dataset of SIV-uninfected NHP TB lung TB dataset which included single-cell data from uninvolved lung tissue *(78, 79)*. From this dataset, lung granulomas from 4 weeks p.i. (granuloma data published in Bromley *et al.*) were compared against uninvolved lung samples from the same experimental condition *(78)* (uninvolved lung data unpublished, **Fig. 5E; Table S3; Methods**). Evaluating our hdWGCNA modules, we observed that the human TB-myofibroblast module was upregulated in the granuloma compared to uninvolved lung tissues, confirming that this phenotype is associated with granuloma-specific structural or cellular processes that reflect a local response to *M.tb*.

Confirmation of the myofibroblast-like phenotype in different TB disease contexts

Taken together, our findings suggest a previously underappreciated role for fibroblasts – including a myofibroblast-like, *MMP1*$^{+}$*CXCL5*$^{+}$ subcluster – in TB immunopathology. However, whilst this population was detected in 5/9 TB diseased samples, the majority of cells derived from a single donor. Therefore, we quantified this fibroblast subset in additional patients undergoing surgery for post-TB lung complications by flow cytometry (**Table S4**). For this, we gated on non-hematopoietic cells (CD45^{-}), lacking expression of CD234a (duffy antigen), CD31 (endothelial cells), EPCAM (epithelial cells) and CD34 (progenitors), but expressing the fibroblast marker CD90, as well as PDPN and FAP, which are both canonical myofibroblast markers and hubs genes in our human TB-myofibroblast module (**Fig. 6A**). These data confirmed the presence of PDPN^{+}FAP^{+} fibroblasts in 5/5 TB diseased lung samples. In addition, by examining tissue samples from regions of lung tissue with varying degree of disease pathology, as determined by operating surgeon, we found that PDPN^{+}FAP^{+} fibroblasts were elevated in the most diseased lung pieces *(p*-value=$8*10^{-4}$, Friedman test; **Fig. 6B**).

To determine whether a *M.tb* stimulation *in vivo* could induce these TB-associated fibroblast gene signatures, we evaluated the expression of our TB-myofibroblast gene module in a previously published bulk transcriptomics dataset from a standardized tuberculin skin test (TST) challenge (**Fig. 6C; Methods**) *(80)*. In this study HIV^{-} participants with active pulmonary TB or "latent TB" (i.e. individuals with T-cell memory to *M.tb* but no evidence of TB disease) received a standard TST (intradermal injection of purified *M.tb* proteins) challenge or saline contol. The TST site was biopsied 48 hours later and processed for bulk RNA-sequencing. Consistent with our observations, the human TB-myofibroblast signature was induced in response to the standardized mycobacterial antigen stimulation *in vivo* compared to saline controls, where no inflammatory response is expected. This signal was amplified in the context of active TB disease compared to "latent TB". This implies systematic inflammation from active *M.tb.* infection may prime the differentiation of a pathological fibroblast cell state. Interestingly, genes associated with the *SPP1*$^{+}$*CHI3L1*$^{+}$ Macrophages subset were similarly induced by TST, supporting the hypothesis that *M.tb* stimulation induces myofibroblast-like phenotype and *SPP1*$^{+}$ macrophages in humans, especially in the context of active TB disease (**Fig. 6C**).

Cell-cell interaction analysis reveals fibroblasts dominate cellular crosstalk in TB-diseased lung.

To identify putative intercellular interactions regulating differentially expressed genes between TB-diseased and control lung niches, we used MultiNichenet *(81)* (**Fig. 7A-B; Methods**). This indicated that fibroblasts were both the dominant sender and receiver cell type in TB-diseased lungs (**Fig. 7C**). Interestingly, a significant proportion of fibroblasts expressed ligand and receptor pairs consistent with autocrine signaling. In contrast, top interactions in the TB-negative condition were more diverse, with pronounced T cell and myeloid involvement.

We performed additional analysis with LIANA to drill down into the specific cellular subclusters contributing to niche crosstalk in TB infected and control lungs *(82)* (**Data Set 3A-B, Data Set 4A**). This integrated ligand-receptor analysis framework leverages multiple resources and methods to generate aggregated inference on samples from each condition. Our results suggested a dominant role for *COMP*$^{+}$*CILP*$^{+}$, *IL6*$^{+}$*CCL2*$^{+}$*, MMP1*$^{+}$*CXCL5*$^{+}$*, and SERPINE2*$^{+}$*COL1A1*$^{+}$ Fibroblasts, but not *HSP* Fibroblasts, in TB-diseased conditions. This analysis also implied that AT1 cell sender signaling is upregulated in the TB-diseased lung, although these cells were significantly depleted in TB-diseased lungs (**Fig. 2D, Data Set 4B**). AT1 cells normally serve as the interface of oxygen exchange in the alveoli, but we found they have high expression of collagen in the TB-diseased lungs, which broadly targets other cell types (**Data Set 4B**). In TB-negative diseased lungs, only *HSP* Fibroblasts were predicted to contribute to signaling (**Data Set 3B**). It is important to note, however, that the lack of fibroblasts in TB-negative lungs may influence this analysis.

To understand broader signaling patterns, we grouped sender and receiver cell types based on similarities of their signaling patterns (**Methods**). Within TB-diseased samples, we observed more distinctive patterns among sender cell types than receivers, with senders roughly grouped by cell type (**Fig. 7D**). The opposite was observed in TB-negative lung (**Data set 4C**). In TB, most of the fibroblast sender subclusters (*MMP1*$^{+}$*CXCL5*$^{+}$*, COMP*$^{+}$*CILP*$^{+}$, *IL6*$^{+}$*CCL2*$^{+}$) grouped together and with other non-immune cells (endothelial cells, AT1, and club cells; Sender Group 1). Quantification of net cell signaling flux – defined as the product of a sender population's relative abundance and its average expression of a given signal – highlighted that, despite making up small proportion of the entire dataset, *MMP1*$^{+}$ *CXCL5*$^{+}$ Fibroblasts were among the most prolific signal senders in the TB-diseased condition (**Fig. 8A**).

Next, to quantify ligand-driven changes in cellular crosstalk during TB infection, we calculated the difference in ligand interaction strengths between TB-diseased and TB-negative lung samples (**Fig. 8B; Methods**). TB sender Group 1 secreted most of the top upregulated ligands in TB-diseased lung (binomial test p-value<0.01, **Data Set 5A, left**). Top senders of all upregulated interactions in TB-diseased lungs were *COMP*$^{+}$*CLIP*$^{+}$ Fibroblasts followed by *MMP1*$^{+}$*CXCL5*$^{+}$ Fibroblasts (**Data Set 5A, right**). In contrast, control sender group 2, which consists mostly of monocytes and macrophages, exhibited the greatest signaling flux in control lung (binomial test p-value<0.01, **Data Set 5B**). Notably, *MMP1*$^{+}$*CXCL5*$^{+}$ Fibroblasts expressed most of the top flux ligands (9/30) with increased overall interaction strength in TB, supporting a central role in TB disease (**Fig. 8B**). The top 5 ligands from the *MMP1*$^{+}$*CXCL5*$^{+}$ Fibroblasts, ranked by average interaction strength across receptors, were CXCL13, CXCL6, DSG2, GREM1, and NTN1 (**Fig. 8C**). CXCL13 may act in both autocrine and paracrine modes, signaling to B cells via CXCR5, a homing marker for activated lymphocyte to lymphoid tissues and on B cells in NHP lung TB granuloma where it regulates host-pathogen interactions

(83). CXCL6 appears to function similarly, signaling to neutrophils and self, consistent with known functions in inducing fibroblast matrix expression, neutrophil recruitment and activation *(84, 85)*. DSG2 (desmoglein) is known to induce pro-proliferative activity in dermal fibroblasts *(86)*, and is highly upregulated in zebrafish and human granuloma *(87)*. GREM1, part of the TGF-β superfamily, contributes to pulmonary fibrosis during the early stages of disease *(88)*. NTN1 (netrin-1), meanwhile, supports endothelial survival and regulates angiogenesis, an important process for dissemination of the pathogen *(89, 90)*.

We also examined other top ligands sent by this subcluster in TB-diseased lungs specifically. Our analyses suggested a prominent role for the *MMP1*$^{+}$*CXCL5*$^{+}$ subcluster in coordinating fibrosis and inflammation through expression of collagen proteins, *MMP1* and cytokines (**Fig. 8D**). Notably, MMP1 itself acts as a ligand for ITGA2, a receptor expressed on epithelial cells, endothelial cells, and *IL6*$^{+}$*CCL2*$^{+}$ Fibroblasts, mirroring the AT1-fibroblast interactions aforementioned. CXCL5 interacts with CXCR2 on *CD16*$^{+}$ monocytes and neutrophils, a key axis for recruitment these cells during TB infection and likely fueling granuloma formation *(91–93)*. *MMP1*$^{+}$ *CXCL5*$^{+}$ Fibroblasts also secrete numerous ECM-related ligands: for instance, collagen molecules *COL1A1, COL1A2, COL6A3, COL3A1,* and *COL5A2* signal to epithelial cells, monocytes, in addition to other fibroblasts via an autocrine loop; ECM proteins *POSTN, FBN-1* and *DCN* signal with other non-immune cell types; and, *MXRA5*, a matrix remodeling protein like *MMP1*, communicates to AT1/AT2 cells via *AGER*. Notably, many of these ECM related ligands are highly upregulated in LN TB granuloma (**Data Set 5C)**. Collectively, our analyses suggest aberrant lung remodeling may be driven by fibroblast and AT1 communication, leading to the fibrosis typical of TB – a trend not necessarily reflected from cell type abundance changes

Cell-cell interaction analysis underscores the relevance of *SPP1*$^{+}$ macrophages in human TB

The main receiver of fibroblasts signaling were fibroblasts and macrophages (**Fig. 7C**). Amongst macrophage subclusters, the top receivers for TB-upregulated fibroblast signals were the *ARL4C*$^{+}$*EMP1*$^{+}$ and *SPP1*$^{+}$*CHI3L1*$^{+}$ cells (**Fig. S5A**). Conversely, *SPP1*$^{+}$*CHI3L1*$^{+}$ subcluster mostly signal to fibroblasts, followed by macrophages (**Fig. S5B**), suggesting a potentially important role in fibroblast-macrophage crosstalk. *SPP1*$^{+}$ macrophages have been identified in the lungs of individuals with COVID-19, IPF, and lung carcinoma, and in bronchoalveolar lavage fluids from TB and latent TB patients *(17, 94)*. In tumors, *SPP1*$^{+}$ macrophages are highly immunosuppressive and associated with poor outcomes, and they have been shown to orchestrate fibroblast activation during fibrosis, driving myofibroblast activation in heart and kidney injuries *(95–97)*. Comparing against other known markers, we noted that our *SPP1*$^{+}$*CHI3L1*$^{+}$ Macrophages appeared congruent with *SPP1*$^{+}$ macrophages described in other disease contexts (**Fig. S5C-D**). Cell-cell interaction analysis showed that fibroblasts were the major receiver of *SPP1*$^{+}$*CHI3L1*$^{+}$ Macrophage signals (binomial test *p*-value = $8.7*10^{-14}$), and nominated *SPP1* and *FN1* as the major ligand genes driving cross talk with fibroblasts (**Fig. S5E**).

To further confirm the presence of *SPP1*$^{+}$ macrophages in human lung TB granulomas, we performed immunohistochemical staining of tissues from two independent donors. We observed abundant total

macrophages ($CD68^+$) in both the granuloma and surrounding lung tissue, and localization of alveolar macrophages ($CD68^+CD206^+$) in the non-granulomatous lung tissue, where alveolar sacs were still visible (**Fig. S5F left**). In stark contrast, $CD68^+SPP1^+$ macrophages localized to the inner cellular periphery immediately bordering the necrotic core of the granuloma and were largely absent from surrounding lung tissue. Quantification of SPP1 expression shows significant difference between the inner cellular layer and the other regions and to a larger degree than CD68 (**Fig. S5F, right)**. Notably, CTHRC-1, a marker for *MMP1*$^+$*CXCL5*$^+$ Fibroblasts that was localized to granuloma at the protein level, has been suggested to play a role in cross talk with *SPP1*$^+$ macrophages *(98)* (**Fig. S4B**). These lines of evidence support the direct interaction between *SPP1*$^+$ macrophages and myofibroblast-like phenotype in human TB granuloma implied by the single-cell data.

Spatial transcriptomics confirms myofibroblast-like phenotype in independent human cohort

Finally, to confirm our observations from human TB lung and LN granulomas, we investigate cells within the Visium dataset for expression of the myofibroblast-like module (**Table S5, Fig. 1B, Fig. S1A**). Consistent with our other data, the human TB-myofibroblast signature was detected in both current and post-TB lesions, and was particularly highly expressed around granuloma structures (**Fig. 9A**). In addition, we found that both HIV^+ and HIV^- samples displayed clear human TB-myofibroblast signature expression, suggesting it not limited to TB mono-infected individuals, as potentially suggested by our single-cell data (**Fig. 4A**). Indeed, in both current- and post-TB samples, HIV was associated with a higher TB-myofibroblast signature expression (**Fig. 9B**). This may be because HIV impairs CD4+ T cells and macrophage-driven repair, and increases TGF-β release, keeping myofibroblast-like cells chronically active *(99–101)*. In HIV^- samples, current-TB was associated with elevated expression of the myofibroblast-like module, but the opposite was true in HIV^+ samples, likely due to persistent systematic immune activation from HIV. These observations suggest that both pathogens can exacerbate expression of this disease associated module. Within each disease condition, we found granuloma samples had higher human TB-myofibroblast signature expression, with the exception of HIV^+ post-TB group, where expression was highest in the iBALT sample (**Data Set 6A).** However, only one iBALT sample was available for this condition which limits our confidence in the observation. Nevertheless, these data confirm that the human TB-myofibroblast phenotype is localized to human TB lung granuloma in both active TB and post TB lung disease, irrespective of concurrent HIV infection.

For each of the granuloma samples, annotations on granuloma borders (“granuloma cuff”) were designated in paired H&E staining images by a trained histopathologist and used to examine the spatial distribution of gene signatures (**Fig. S1B**). This analysis confirmed that the human TB-myofibroblast signature was strongly expressed in the granuloma cuffs compared to surrounding regions, with a slightly higher presence outside the granuloma compared to the granuloma core (**Fig. 9C**). Interestingly, examine the other fibroblast modules revealed distinct spatial orientation relative to the granuloma (**Data set 6B**). Like *MMP1*$^+$ *CXCL5*$^+$ Fibroblasts*,* *COMP*$^+$*CLIP*$^+$ and *SERPINE2*$^+$*COL1A*$^+$ Fibroblast displayed a similar pattern of enrichment around the granuloma cuff, whereas *IL6*$^+$*CCL2*$^+$ Fibroblast and *HSP* Fibroblast exhibited greater enrichment outside the

granuloma. *MMP1*$^{+}$ *CXCL5*$^{+}$ Fibroblast, however, showed the largest difference for marker expression between the Visium spots on granuloma cuff and those inside/outside the cuff.

SPP1^{+}*CHI3L1*$^{+}$ Macrophage marker expression was similarly enriched on the granuloma cuff, supporting the co-localization of myofibroblast-like phenotype and *SPP1*$^{+}$*CHI3L1*$^{+}$ Macrophages at this site (**Fig. 9C**). To confirm this relationship, we looked at the correlation between all macrophage subset markers with the human TB-myofibroblast signature across all samples and found the strongest correlation with *SPP1*$^{+}$*CHI3L1*$^{+}$ Macrophages compared to the other macrophage subsets (**Fig. 9D**). Finally, we conducted a ligand-receptor analysis to identify spatially co-expressed ligand-receptor pairs using the same database as our analysis on scRNA-seq data (**Methods**). This identified the same L-R pairs as the top pairs in both samples, including, for example, SPP1-CD44. This interaction was nominated as the top L-R pair in several samples, specifically highlighted around the granuloma cuffs and in our scRNA-seq data (**Fig. 9E**, **Fig. S5E**).

Taken together, our scRNA-seq and spatial transcriptomics data support the robustness and generalizability of the human TB-myofibroblast signature, and confirm its co-localization and cross talk with *SPP1*$^{+}$ macrophages in human TB lung granuloma.

DISCUSSION

TB is a global pandemic, and transformative interventions are hindered by an incomplete understanding of its pathogenic processes, including the extensive lung remodeling in pulmonary TB that drives transmission, mortality, and a high burden of PTLD following successful treatment *(102)*. Several sequencing studies have highlighted a central role for ECM remodeling of human lung in TB, but none have resolved the contributions of individual cell-type *(103)*. Additionally, an emerging issue in TB research is that findings from the circulation – the compartment mostly frequently studied – often fail to reflect processes in diseased tissue *(104)*. To address these gaps, we analyzed scRNA-seq data generated from lung tissue freshly resected to treat complication arising from TB disease, and systematically cross-referenced our findings with public datasets from *M.tb*-infected NHPs, the HCLA, LN TB granulomas, and TST challenge, as well as additional immunohistochemical, flow-cytometric, and spatial transcriptomic data from the same cohort to identify TB-specific changes at the cellular level. Collectively, our lung datasets provides a key resource defining the cellular subsets present in TB diseased lung and dissecting immunopathogenic mechanisms. Our data demonstrate substantial heterogeneity amongst key innate immune populations, such as macrophages and neutrophils, in infected lung tissue. We find that several of these subsets correlate with a recent single-cell analysis of *M.tb.*-infected NHPs *(16)*, a study not limited by tissue availability or complicated by comorbidities such as HIV, further strengthening our observations. In addition, our data highlight a possible central role for diverse fibroblast subsets with TB diseased lung tissue and with TB granuloma, particularly an underappreciated *MMP1*$^{+}$*CXCL5*$^{+}$ fibroblast population that colocalizes with *SPP1*$^{+}$ macrophages at the granuloma cuff. We hypothesize that interaction between these cells, which express a myofibroblast-like gene module, and *SPP1*$^{+}$ macrophages may play an important role in human TB granuloma development and post TB-lung disease, potentially aggravating granuloma progression and lung fibrosis. Further examining these

putative interactions could more deeply inform our understanding of granuloma biology and suggest promising targets for novel TB host-directed therapies.

Previously, limited knowledge on matrix turnover mechanisms has hindered development of clinical strategies for managing PTLD *(109)*; here, our study identifies potential cell targets, including heterogenous fibroblast subsets such as those expressing a myofibroblast-like gene module. Lung myofibroblasts are thought to arise from a variety routes, ranging from differentiation of tissue-resident fibroblasts, epithelial to mesenchymal transition (EMT) *(110)*, endothelial to mesenchymal transition (EndoMT) *(111)*, and bone marrow-derived progenitors such as fibrocytes *(112)*. The myofibroblast-like cells showed in this study express genes observed in immune fibroblasts (lineage$^-$, CD34$^-$, CD90$^+$, FAP$^+$, PDPN$^+$) *(113)*. These cells are critical for the formation of tertiary lymphoid structures (TLS), which arise in response to sustained inflammation *(114)* and are commonly observed in TB infected lung tissue *(115)*. Additionally, matrix remodeling driven by skewed fibroblast populations can profoundly impact the cellular niche. Changes in ECM composition can further perpetuate fibroblast reprogramming and ECM remodeling, as seen in escalating MMP1 expression *(116)*. These findings help guide interpretation of our cell-cell interaction analyses, highlighting significant roles for ECM-related molecules.

Post-primary human TB is often paucibacillary *(117)*, and it remains puzzling how profound lung destruction is generated under such conditions. The data presented here may support a model in which fibroblasts-ECM interactions exacerbate and perpetuate lung destruction in human TB, and highlight the emerging immune regulatory role of fibroblasts *(118)*. Of note, a phase II clinical trial in patients with pulmonary TB found that 2 weeks doxycycline, an MMP inhibitor, led to significant changes in the peripheral transcriptome at 8 weeks *(119)*, demonstrating how a matrix-modulating host-directed therapy may influence the immunological trajectory of disease. Overall, our single-cell and spatial transcriptomics analyses highlight a previously overlooked role for myofibroblast-like phenotype as likely key player in orchestrating the immune response and regulating immunopathology in TB.

Anti-inflammatory macrophages are generally enriched in TB-diseased tissue during chronic TB infection, potentially limiting immunopathology but also creating a favorable niche for *M.tb.* replication *(105)*. Here, we found most macrophage populations were skewed in TB diseased lung tissue compared to TB-negative tissues, with a similar trend between post-TB to current-TB spatial samples. Of particular interest are *SPP1*$^+$ macrophages which were elevated in TB-diseased lung tissue and strongly associated with the granuloma cuff in our spatial transcriptomics and histology data. This population has not been characterized in TB lung granuloma, but are emerging as important players in tumors, IPF-diseased lung tissue, and other fibrotic conditions *(106, 107)*. The presence of *SPP1*$^+$ macrophages in TB granuloma was further supported by granuloma RNA-seq data from human LNs and experimentally infected NHPs. Furthermore, the *SPP1*$^+$ macrophage markers were upregulated following TST challenge, which was amplified by concurrently active TB disease, linking their induction to *M.tb* exposure. Moreover, we found evidence of cross talk between *SPP1*$^+$ macrophages and the human TB-myofibroblast phenotype, a previously underappreciated but

potentially important interaction in TB. This putative interaction is supported by histological and spatial transcriptomics data indicating both *SPP1*$^{+}$*CHI3L1*$^{+}$ Macrophages and *MMP1*$^{+}$ *CXCL5*$^{+}$ Fibroblasts are tightly associated with the granuloma cuff. In IPF, *SPP1*$^{+}$ macrophages are highly expanded in fibrotic lesions and crosstalk with myofibroblasts to drive fibrotic changes *(106)*; in colorectal cancer, there are direct interactions between *SPP1*$^{+}$ macrophages and *FAP*$^{+}$ fibroblasts expressing high levels of *MMP1*/*3 (107).* In addition, mechanistic work in murine models showed *SPP1*$^{+}$ macrophages can directly activate myofibroblasts via SPP1 and FN1 *(97)*, both of which are implicated in the *SPP1*+ macrophage-fibroblast crosstalk we found in TB lung tissues. This interaction was also linked to an immune-suppressive, pro-tumorigenic microenvironment through active ECM deposition – resembling granuloma formation in tuberculosis *(108)*. Thus, we hypothesize the *SPP1*$^{+}$ macrophages-myofibroblast axis likely play an important role in TB granuloma biology.

While our study provides much needed information on TB-diseased human lungs, several limitations should be acknowledged. Our cohort size is modest, and substantial variability between patients and sampling location exists in both the primary resections used in the single-cell analysis and flow cytometry experiments. We attempted to address these challenges by obtaining additional samples for spatial transcriptomics and by integrating our analyses with data from relevant public datasets. However, we are still likely to have missed some biological features underlying TB pathology. In addition, further work is needed to dissect the mechanistic role of the myofibroblast-like phenotype and the interaction of the cells that express it interactions with *SPP1*$^{+}$ macrophages in TB immunopathology. Possible avenues include co-culture systems, conditioned media assays, or recruitment assays to determine whether and how these fibroblasts influence, and are influenced by, macrophage behaviors, as well as whether chemotactic interactions exist. *Ex vivo* stimulation experiments with TB antigens on isolated fibroblasts or macrophages could help establish whether TB-specific cues directly drive differentiation towards these disease phenotypes. Genetic approaches, such as targeted knockout of key genes in *MMP1*$^{+}$*CXCL5*$^{+}$ Fibroblasts or genome-wide CRISPR screens in fibroblasts within animal models of TB, could clarify causal relationships between these cells and TB pathogenesis and tissue remodeling. Beyond identifying causality, studying earlier time points in TB infection will be necessary to understand disease progression and the origins of TB complications. Ultimately, an integrated spatial, temporal, single-cell resolution disease map may be required to fully understand pulmonary reprogramming due to TB and guide optimal treatment strategies that maximizes bacterial clearance while minimizing or restoring post-TB lung damage.

In sum, our study demonstrates the power of single-cell profiling to help identify, and spatial transcriptomics to contextualize, potential drivers of immunopathology underlying lung remodeling in TB disease. Our analysis highlights specific macrophage and fibroblast populations, as well as ECM related processes, as promising targets for novel host-directed therapies that could complement or offer alternatives to standard antibiotic regimens.

MATERIALS AND METHODS

Human study ethics and participants

Human lung tissue was obtained from patients undergoing surgery due to TB sequelae including but not limited to haemoptysis, cavitation, bronchiectasis, shrunken or collapsed lung, at the Department of Cardiothoracic Surgery at King Dinizulu Hospital in Durban, KwaZulu Natal and Inkosi Albert Luthuli Central Hospital in KwaZulu-Natal. All samples were collected with approval from the Biomedical Research Ethics Committee and written informed consents obtained from all subjects (BREC no 019/13).

Human lung tissue preparation

Single-cell RNA-seq samples: The lung tissue was processed within 3 hours of receipt as described[17]. Briefly, a piece of the lung tissue was cut for histology and placed in 4% paraformaldehyde (PFA). The remaining piece of tissue was dissected into small pieces (5x5x5mm) and infiltrated with a collagenase (Sigma-Aldrich), DNase 1 (Sigma-Aldrich) in RPMI (Sigma-Aldrich) with 10% FBS (Hyclone) for 30 minutes. Mechanical digestion at room temperature was performed using the Gentle MACS (Miltenyi Biotec) followed by agitation at 37°C for 30 minutes. The mechanical digestion and agitation were repeated once more, followed by filtration of the resulting cellular suspension using the 70mm (Corning) and 40mm (Corning) strainer, followed by the lysis of red blood cells. Cells were then stained with trypan blue (Thermo Fischer) and enumerated using an automated cell counter (BioRad) or a manual counter (Kova).

Spatial transcriptomics (Visium) samples: A section of lung was cut and transferred to 10% buffered formalin to fix for 24 hours, then transferred to 70% ethanol until wax embedding. The sample was then processed in a vacuum filtration processor using a xylene-free method and isopropanol as the main substitute fixative. The tissues were embedded in paraffin wax. Tissue sections (5 µm) of specimens of good quality, as determined by trained histotechnologist, were mounted on charged slides, air dried for 30 min then at 42 °C for 3 hours in a desiccator and stored in a desiccator at room temperature until use.

Non-human primate (NHP) study ethics and research animals

The macaques used for generating the scRNA-seq data were part of the study published by Ganchua *et al.* and the same ethical and maintenance procedures were followed *(79)*.

NHP infections and disease tracking by PET-CT

Five cynomolgus macaques (Macaca fascicularis, aged between 5.3–9.1 years), obtained from Valley Biosystems (Sacramento, California) were part of a previously published study as the “immune naïve” control group *(78, 79)*. They only received a low dose infection (7 CFU) with *M. tuberculosis* strain Erdman and were necropsied 4 weeks post infection. PET-CT was performed just prior to necropsy and results were analyzed using OsiriX viewer as previously described, with a detection limit of 1 mm *(120)*. The infection dose was

determined by colony counts after plating an aliquot of the infection inoculum on 7H11 agar plates, which were incubating for 3 weeks at 37°C/5% CO2.

Necropsy protocols

Procedures carried out during necropsy have been previously described *(79)*. Briefly, 1–3 days before necropsy, PET-CT scans were taken to pinpoint the location and metabolic activity (FDG activity) of granulomas. These scans served as a guide during necropsy for precise identification and excision of these samples. On the day of necropsy, macaques were sacrificed humanely by infection of sodium pentobarbital and terminally bled. Individual granulomas and uninvolved lung tissue were all excised and homogenized separately into single cell suspensions. Homogenates were aliquoted for plating on 7H11 agar for bacterial burden, freezing for DNA extraction, and staining for flow cytometry analysis. Any remaining samples were frozen for future use.

Human lung single-cell RNA-sequencing (scRNA-seq) with Seq-Well S^3

Seq-Well S^3 was implemented as described[18]. Briefly, the single-cell suspension was diluted to 15,000 cells in 200µL of RPMI (Sigma-Aldrich) plus 10% FBS (Hyclone) and loaded onto a polymethylsiloxane (PDMS) array pre-treated with the same solution for 15 minutes. The cells were allowed to settle into the microwells by gravity and the array was washed with PBS (Sigma-Aldrich) and sealed with a plasma functionalized polycarbonate membrane (Sterlitech). The arrays were then sealed followed by incubation at 37˚C for 40 minutes, followed by a 20-minute incubation in lysis buffer containing guanidium thiocyanate (Sigma-Aldrich), EDTA (Thermo Fischer), 1% betamercaptoethanol (Sigma-Aldrich) and sarkosyl (Sigma-Aldrich) at room temperature. The arrays were then transferred to a hybridization buffer containing NaCl (Thermo Fischer), $MgCl_2$ (Sigma), PBS (Thermo Fischer) and polyethylene glycol (SigmaAldrich) and were gently shaken at 60rpm for 40 minutes. The capture beads hybridized with released mRNA from the lysed cells were collected from the array by a series of three wash steps with wash buffer containing NaCl (Thermo Fischer), $MgCl_2$(Sigma), Tris-HCl (Thermo Fischer) and Water (Inqaba Biotech), with centrifugation at 2500g for 5 minutes each iteration. The beads were resuspended in a master mix for reverse transcriptase containing Maxima H Minus Reverse Transcriptase, Maxima Buffer, dNTPs, RNAse inhibitor, a template switch oligonucleotide and PEG for 30 minutes at room temperature and overnight with endto-end mixing at 52˚C. This was followed by the standard exonuclease digestion and denaturation of complementary DNA (cDNA) hybridized to the bead by 5-minute incubation in NaOH (Sigma-Alrich) and washed with a solution containing Tris-HCl, EDTA and Tween-20 (Thermo Fischer). The beads were resuspended in a master mix containing Klenow Fragment (NEB), dNTPs, PEG and the dN-SMRT oligonucleotide, incubating for 45 minutes at 38˚C. PCR was performed as described in the protocol and the product was subjected to 2 rounds of AMPure XP SPRI (Agencourt) bead cleanup at 0.6x and 0.8x volumetric ratios sequentially. The library size was analyzed using an Agilent Tape station hsD5000 kit, ensuring that the expected product had an average size of ~1000bp and the absence of primer dimers especially below 200bp. The Qubit High Sensitivity DNA kit was used to quantify the libraries and they were prepared for Illumina sequencing using the Nextera XT DNA Sample Preparation kit. A total of 900pg of the different libraries were added the tagmentation reaction. The

amplified product was purified with the AMPure XP SPRI beads and the libraries were pooled for loading onto the NovaSeq 6000 using paired end read structure with custom read 1 primer: read 1:20 bases, read 2: 50 bases, read 1 index: 8 bases.

Spatial transcriptomics with Visium and paired H&E staining

Tissue slides were baked at 60 °C for 2 hours and dewaxed using two xylene changes and rehydrated with descending grades of alcohol to water. They were then Hematoxylin and Eosin (H&E) stained and imaged as the reference image and the same slide was then processed as per Visium v2 chemistry protocol following manufacturer's recommendations (Visium Spatial Gene Expression for FFPE – Deparaffinization, H&E Staining, Imaging & Decrosslinking, document CG000409 RevD, 10x Genomics, (Sep 2023); Visium Spatial Gene Expression for FFPE Imaging Guidelines, document CG000436 RevB, 10x Genomics, (Sept 2023); Visium Spatial Gene Expression Reagent Kits for FFPE User Guide, document GC000407 Rev E, 10x Genomics, (Sept 2023)). The sequencing results were processed through the SpaceRanger software following manual alignment of the fiduciary frames using the 10x Loupe browser.

NHP sample scRNA-seq with Seq-Well S^3

ScRNA-seq was performed on both uninvolved lung tissues and granuloma tissues using the Seq-Well S^3 platform as described in Bromley *et al.,* where and the granuloma data was previously published *(78).*

NHP single-cell data alignment and analysis

The transcript reads were aligned as described in Bromley *et al.(78).* Briefly, transcript reads were tagged for cell barcode and UMI using *DropSeqTools* v1.12, then aligned to the Macaca fascicularis v5 genome (https://useast.ensembl.org/Macaca_fascicularis/Info/Index) through the *Dropseq-tools* pipeline on the Terra platform (app.terra.bio) *(121)*. Aligned reads were collapsed by barcode and UMI sequences to generate digital gene expression matrices for each array, covering 10,000 barcodes. For each sample, gene expression matrices with >=10,000 barcodes were processed through *CellBender* to estimate ambient RNA fraction. The "*remove-background*" function in *CellBender* was applied with default settings. Next, the matrices "corrected" by *CellBender* were analyzed with *Scrublet*, with default parameters to detect potential doublets. Any transcriptome with a *doublet_score* >0.30 were removed from downstream analyses.

After that, the gene expression matrices for each sample were merged and processed in *Scanpy* (version 1.8.2). Transcriptomes were filtered using following the following criteria: min_genes > 300, min_counts>500, mitochondrial_threshold=0.05, and genes expressed in at least 10 cells. Gene expression counts were normalized using default *Scanpy* parameters (i.e., log2(TP10K+1)). Coarse-level cell type clustering and iterative sub-clustering were used to annotate cell types and further detect low-quality transcriptomes (e.g., doublets). Cell types were identified using canonical markers, and only fibroblast cells were included in the analysis presented in this study.

Human lung single-cell data analysis and cell type identification

The raw sequencing reads from the NovaSeq run were aligned to the hg19 genome assembly and processed in accordance with the Drop-Seq Computational Protocol v2.0 (https://github.com/broadinstitute/Drop-seq). The output (cell by gene matrix) was then loaded to the Seurat R package v3.1.0 (https://satijalab.org/seurat/), transformed to $\log_e$(UMI + 1) followed by scaling by a factor of 10000. The overall quality was assessed by the distribution of reads, transcripts and genes per cell (percentage of mitochondrial genes <5, nFeature_RNA<2500, nFeature_RNA>200, nCount_RNA>200). SCTransform by Seurat was called to perform normalization of the gene counts, selecting top 3000 highly variable genes, and scaling normalized gene counts. Principle component analysis was run on the selected highly variable genes to give the top 50 PCs. A custom elbow-based method was used to find the smallest number of PCs (*n_pcs*) where the eigengap between two adjacent PCs drops below 20-percentile of all eigengaps amongst top 50 PCs. Uniform Manifold Approximation and Projejection (UMAP) was calculated using the *RunUMAP* function and neighborhood graph was calculated by *FindNeighbors*, both using *reduction='pca'* and selecting top *n_pcs* as input dimensions. Unsupervised Louvain clustering using the FindClusters was used to identify transcriptionally similar cells with parameters *assay='integrated'*, dims.use=n_pcs, *k.param=ceiling(0.5*sqrt(#cells))* and we performed a resolution scan for the best clustering resolution from 0.2 to 2 while optimizing for silhouette score. Cell type annotation were done by cross-referencing canonical cluster defining genes with well curated lists, online databases such SaVant T (http://newpathways.mcdb.ucla.edu/savant-dev/) and GSEA/MsigDB (https://www.gsea-msigdb. org/gsea/msigdb/index.jsp). Doublet clusters where multiple canonical markers were expressed are identified and removed and the entire dataset are re-processed starting from the SCtransform step. Final differentially expressed (DE) gene for each of the major clusters were found by calling *FindAllMarkers* from *Seurat* using default setting and adjusted p-value cutoff < 0.05 and top DE genes were found by ranking log fold-change values from high to low. Heatmap of DE genes were plotted using Seurat function *DoHeatmap* and dotplot was achieved using function DotPlot.

Sub-clustering for major cell groups (macrophage/monocytes, neutrophiles, epithelial cells, fibroblasts) were performed similarly to the entire dataset after subsetting to the specific cell types. Marker genes for each subcluster was found by calling *FindAllMarkers* from *Seurat* using default setting and significant genes (adjusted p-value<0.05) are visualized with custom volcano plots.

Comparison with human lymph node data was done for the top 10 DE genes in each cellular subcluster and checked over the TB vs. control differential testing result from the human lymph node granuloma study.

Clustering analysis on cell subtypes

Proportion of cell subtypes in each patient was calculated, and Pearson's correlations between every pair of broad level cell type is calculated. For each pair of cell types, we ran permutation test by randomly re-assigning cell type labels to generate a set of background correlation values, and p values are calculated as the percentage the permutated correlation values exceeds the original observation. Hierarchical clustering on

the cell types are done by feeding in the pairwise correlation into Python function *linkage* with method='average', metric='correlation', we then use function *fcluster* with a defined k to call cluster from the returned linkage result with criterion='maxclust'. We grid searched through k from 2 to 29 (one less than the number of cell types) and determined the optimal cluster number by computing the silhouette score from each clustering result with function *silhouette_score* and a precomputed correlation distance. This allowed us to select k=12 which resulted in the highest silhouette score. For each of the 12 clusters identified through hierarchical clustering, we further calculated permutation test p values to examine average correlation values within and outside of each cluster and annotate those that has within-group p value <0.05.

Differential abundance testing

Statistical differences in the cell type abundance between TB diseased and TB-negative lungs were tested by two-sided Fisher's exact text and adjusted for multiple testing correction by Holm's method.

Cluster-free differential abundance testing is done using *milopy* in python. Neighborhoods are constructed over the entire dataset using $k = ceil(0.5 \times \sqrt{n})$ where $ceil$ rounds up to the nearest integer and *n* is the number of cells. Neighborhood are made with $prop = 1$. Function *DA_nhoods* was called with $design = \sim HIV + TB$ to account for the effect of HIV status. For interpretation, we only kept neighborhoods with neighborhood annotation fraction > 0.5, and label them with the majority cells' annotation. Due to the small sample size, we opted to use P-value instead of the spatial FDR devised in *milopy* for significance.

Bulk RNA-seq profile deconvolution and comparison

For comparing the marker genes in each subcluster with DE genes in bulk RNA-seq on human lymph node TB granuloma samples, we first selected top 50 DE genes in each subcluster. Note that some of the DE genes in a broad cell type may overlap with the DE genes in another, since the differential analysis was done within each broad cell type. Hence, we remove the genes that are shared between cell types, re-ranked the remaining DE genes by log-fold change, and took the top 10 DE genes to compare with the bulk DE results.

For deconvolution of the human lymph node TB granuloma and control samples, we applied tool MuSiC (1.0.0) separately on TB and control samples, using annotated data in our study as single cell reference. We kept all the cell types for deconvolution except alveolar macrophages, which should only exist in lungs. Other parameters are kept as default.

We applied a standard two-sided T-test to compare the difference between inferred cell type proportions between TB and control lymph node samples, with Bonferroni correction for multiple testing.

Fibroblast label transfer and gene signature finding

Travaglini *et al.* stromal cell type calling: Top 20 markers for each stromal cell population by Travaglini *at el.* 2020 Supplementary Table S4, and was found by filtering on P-value<0.05, and sorted by average log

fold-change. *AddModuleScore* from Seurat was used to calculate module score of these markers and "Travaglini.fib.subtype" was called based on which cell type gives the maximum module score, where "ambiguous" was assigned if no score gives a positive value. Proportion of Travaglini.fib.subtype was calculated in each fibroblast cluster given this new cell annotation.

HLCA label transfer: Human Lung Cell Atlas (HLCA) label transfer onto our dataset was achieved following their tutorial (https://github.com/theislab/scarches/blob/hlca_tutorial_improvements/notebooks/hlca_map_classify.ipynb). Briefly, label transfer was done using asArches on the raw counts of the entire dataset on the genes that are part of the reference model. Annotation level 3 data was used in this paper. Celled called as "fibroblast" or "myofibroblast" are considered together as fibroblast population which are highly consistent with our manual annotation (>95% true positive rate). For better comparison, we only included HLCA fibroblasts (and myofibroblast) with tissue source annotation "lung parenchyma". Differential gene expression analysis was performed between all TB-negative controls (from both HLCA and our study) and TB-diseased lungs (only from our study) on log normalized counts. *GSEA* was run in R using genesets from MSigDB (accessed using *msigdbr*) on DE genes passing filter for Benjamin-Hochberg adjusted P-value<0.05.

Gene module finding with hdWGCNA: single cell version of WGCNA was run using tool hdWGCNA following tutorial (https://smorabit.github.io/hdWGCNA/articles/basic_tutorial.html). Briefly, gene_select = "variable" was used to set the variable gene selection approach using *SetupForWGCNA* . To avoid sparsity in the single-cell data, we first constructed metacells that aggregates the expression profile based on neighborhood information. Metacells were constructed through *MetacellsByGroup* call with parameters k = 10, max_shared = 5, min_cells = 20, group.by uses the categories for fibroblast subcluster and disease status (TB, HIVTB, HIV control and cancer control) and ident.group is also set to be the subcluster. SetDatExpr was called with "SCT" assay and "data" slot for scaled expression. *TestSoftPowers* function was called with networkType = "signed". The rest follows the default analysis workflow. Top genes in each module ranked by eigen-based connectivity (kME) are visualzed by running *PlotkMEs*. Feature plot of module eigengenes (MEs) for each module was plotted by running ModuleFeaturePlot with features='MEs'. *ModuleCorrelogram* function was used to visualize the correlation between each module based on their MEs, and *VlnPlot* from Seurat was used to visualize the difference of module MEs between subclusters.

Lymph node granuloma laser capture microdissection RNA-seq study

Formalin-fixed paraffin-embedded (FFPE) clinical samples from twenty-four adult patients undergoing mediastinal or neck lymph node biopsy were selected (seven TB, ten sarcoidosis, seven normal), and the first analysis has previously been reported (PMID24798354; 10.1136/bcr-2013-202127, GEO accession code GSE174443). The patients were treatment naïve and had no significant comorbidities, were HIV negative and non-smokers. Sections of 10μm thickness were cut, floated in RNase-free water, mounted on to

polyethylene naphthalate (PEN) membrane glass slides and dried at 37°C overnight. Sections were dewaxed with xylene immersion followed by xylene removal with 100% EtOH. Laser capture microdissection was used to isolate granulomas or similar area of control normal tissue. Each sample underwent total RNA extraction and sequenced using Ion Torrent sequencing. Raw sequencing data was aligned using *kallisto* software and annotated to gene level by *ensembldb* and sleuth programs to ensure similar results were found. Inter-sample normalization was performed using TMM normalization (*edgeR*).

Evaluation of differential genes in lymph node granuloma dataset

Genes identified from each cluster during single cell sequencing analysis were searched within the bulk RNAseq dataset of granulomas isolated by laser capture microdissection (GEO accession code GSE174443), where differential gene expression analysis was performed using *limma* with its *voomWithQualityWeights* function (version 3.38.3, R) with Benjamini-Hochberg FDR of less than 0.05. Filter values were optimized to yield the highest number of differentially expressed genes across the study cohort. GraphPad Prism 9 was used to plot the average gene expression of 7 control and 7 TB lymph nodes, with box-and-whisker values generated using one-tailed unpaired T test.

Evaluate gene module in non-human primate dataset

Gene modules found from above are taken to be evaluated in NHP data. Top 50 hub genes are ranked by eigengene-based connectivity (kME) and used to score on fibroblasts from the NHP dataset using function *score_genes* from python package *scanpy*. Two sided Mann-Whitney U test without correction was used to compare module usage between different conditions.

Evaluate gene modules in human TST challenge dataset

Top 50 hub genes from the Fibroblast-M1 module from hdWGCNA is taken as the human TB-myofibroblast module as described above, along with differentially expressed marker genes from *SPP1*$^{+}$*CHI3L1*$^{+}$ Macrophages (**Data File S2**), they are used to score on the bulk RNA-seq data which has been preprocessed following methods in Pollara *et al.* followed by calculating geometric means of all the genes in set *(80)*. Two sided Mann-Whitney U test without correction was used to compare module usage between different conditions.

Cell-cell interaction analyses

MultiNicheNet: analysis was run using package *multinichenetr* following tutorials on https://github.com/saeyslab/multinichenetr. Briefly, recommended ligand-receptor network and ligand-target matrix was downloaded from https://zenodo.org/record/7074291/files and a SingleCellExperiment object was constructed from the RNA assay from the Seurat object. Analysis was defined for senders and receivers as all broad level cell types shown in Figure 1. We performed genome-wide differential expression analysis of

receiver and sender cell types to define DE genes between the conditions of interest (TB-negative and TB-diseased lungs). Empirical P-values were calculated after DE calculation using function *get_empirical_pvals*. Then, we predicted NicheNet ligand activities and NicheNet ligand-target links based on calculated differential expression results using function *get_ligand_activities_targets_DEgenes* with parameters logFC_threshold = 0.50, p_val_threshold = 0.05, fraction_cutoff = 0.05, p_val_adj = FALSE, top_n_target = 250. We sse the information collected above to prioritize all sender-ligand-receiver-receptor pairs using function *generate_prioritization_tables* with prioritizing weights: "de_ligand" = 1,"de_receptor" = 1, "activity_scaled" = 2, "exprs_ligand" = 2, "exprs_receptor" = 2, "frac_exprs_ligand_receptor" = 1, "abund_sender" = 0, "abund_receiver" = 0, and fraction_cutoff=0.05, grouping_tbl consists of sample ID and TB status. Circoplot visualizations of top 20 ligand-receptor pairs in each TB status group were done on prioritization table outputs. Summary heatmap was done over top 200 interactions for enrichment of interactions between cell types.

Given the requirement to perform genome-wide differential expression analysis to identify DE genes between TB conditions, we couldn't apply the same MultiNicheNet framework to all subclusters given some subclusters don't have enough power to detect. Hence, we switched to LIANA for an unbiased cell-cell communication analysis at the subcluster level.

LIANA: liana analysis was first independently run on both TB-diseased data and healthy control data using function *liana_wrap* followed by *liana_aggregate* from the *liana* package in R using default parameters on RNA assay from Seurat. We kept only keep interactions concordant between methods by filtering for interactions with aggregate_rank <= 0.01. Top 20 $MMP1^{+}$ $CXCL5^{+}$ Fibroblast specific signaling in TB was extracted, where interaction specificities are extracted from natmi.edge_specificity values and expression magnitudes are from sca.LRscore value between interactions. Senders/receivers are ordered by hierarchical clustering based on Pearson's correlation of sca.LRscore values.

We summarize the sender-receiver interaction frequencies from the filtered interactions in each TB status group and calculated the difference between the two frequency matrices. Lastly, we normalized by the largest absolute value of differences for plotting the interaction difference heatmap. To visualize interactions strengthened in TB-diseased group and TB-negative group, we defined the edge weight of interactions by the natmi.edge_specificity from LIANA output and edge_FC as the fold change between the TB group and control group with a pseudo edge weight of 10^-6 if control group is 0. We counted the number of interactions between sender-receiver groups involved in interactions of edge_FC>1, defined as "poslogFC.cellcell.count" and similarly the number of interactions between sender-receiver groups involved in interactions of edge_FC<1, defined as "neglogFC.cellcell.count". We clustered sender and receiver in TB-upregulated interactions (summarized in "poslogFC.cellcell.count") and TB-downregulated inetractions (summarized in "neglogFC.cellcell.count") based on Pearson's correlation of interaction count similarities between senders and receivers respectively. Hierarchical clustering was done using *pheatmap* followed by inspecting tree clusters and calling groups using *cutree*. For circos plots of TB-upregulated sender-receiver pairs and TB-downregulated sender-receiver pairs, we only selected for pairs with interaction counts exceeding 80-percentile of all pairs in the particular condition, using function *chordDiagram* from R package *circlize*.

For visualizing the interactions between fibroblast and macrophages and *SPP1*$^{+}$ macrophage signaling, we visualized the count of interactions with aggregate_rank <= 0.01 and edge_FC>1 and used *chordDiagram* to plot.

Ligand interaction strength calculation: mean TB edge is defined as the mean of interaction edge weight in TB-diseased group for specific sender and ligand combination and mean_CTR_edge is defined as the mean of interaction edge weight in TB-negative group. weighted_mean_TB is defined as mean_TB_edge weighted by the count of interaction involving that ligand in each sender group, and similarly for weighted_mean_CTR. Finally, weighted_mean_FC (e.g., interaction strength change) is defined as weighted_mean_TB/weighted_mean_CTR. Top 30 and bottom 30 ligands by the interaction strength are chosen to be visualized in Fig. 5D.

We also calculate an unweighted mean_FC = mean TB edge/ mean_CTR_edge for the interaction strength change in each sender cell type, and we consider an interaction involving a ligand as positive if the log10(mean_FC) is positive and negative if the log10(mean_FC) is negative, which reflect whether the interaction is stronger or weaker in TB-diseased vs. TB-negative group. Positive and negative interaction counts are tallied for each ligand and a ligand is thought to be dominantly “positive” (colored red in barplot) if positive interaction count is 50% higher than negative interaction count and “negative” (colored blue) if negative interaction count is 50% higher than positive interaction count. Mean_FC and log10(mean_FC) between TB-diseased and TB-negative samples are used to indicate ligand activity importance in each sender cell type, sender cell types are grouped according to clustering for TB-diseased senders in circo plot . Grouping of sender cell types in the ligand interaction strength analysis was the same as before. Top 5 ligands in *MMP1*$^{+}$*CXCL5*$^{+}$ Fibroblast by mean TB edge metric was visualized for their proposed targets and number of possible receptor interactions on each cell type.

For summarizing top 10% of ligands in each TB condition, we calculate the mean of edge_FC for each source/ligand combination. The mean edge_FC is then sorted by descending order, where the top 10% and bottom 10% are visualized as top ligands upregulated in TB (**Data set 4B**, **C**). For the barplot of number of interactions upregulated in each TB condition, we filter for all interactions with edge_FC<1 or edge_FC>1 and count the number of interactions by sender cell type. We use the same cluster groupings/colors for the senders as for the circus plot in Fig. 4D.

Sender signaling co-occurrence analysis: we first filter out sender-ligand combinations that are upregulated in TB (edge_FC>1). Then, for each cell type of interest, the normalized RNA count for the upregulated ligands in this cell type is retrieved for all the TB-diseased patients. The ligand expression in each cell is then weighted by log10(mean_FC) so ligands with larger degree of change are weighted higher for their expressions. Then, patient averages of all the weighted ligand expressions are calculated and

summed to arrive at a final patient-sender activity score. Pearson's correlation is computed across each pair of cell type's sender activity scores in 9 patients.

Interaction flux analysis: In this analysis, we define the flux of interaction in the direction from sender to receiver cell types. First, we calculate the mean of edge_FC between all LIANA inferred significant interactions (aggregate_rank <= 0.01) for each sender-receiver pair. Then, for each pair of sender-receiver, the flux of interaction is calculated by multiplying the sender cell count. The total flux of a given sender cell type is then the sum of flux to all receiver cell types.

Fluorescent immunohistochemistry staining

*Sample preparation***:** Multiplex fluorescent immunohistochemistry staining of macrophage markers was performed on lung tissue sections using the Opal™ 6-Plex Manual Detection Kit 50 Slides (AKOYA, USA) as directed by the manufacturers. Multiplex fluorescent immunohistochemistry staining of fibroblast markers was performed on lung tissue sections using the Opal™ 4-Color Manual IHC Kit 50 Slides (PerkinElmer, USA) as directed by the manufacturers. For both, lung tissue samples fixed in 4% formalin were paraffin-embedded. Four mm sections were cut on X-tra adhesive precleaned micro slides (Leica, Germany), allowed to dry for a minimum of 24 hours and the slides were baked at 60^{o}C overnight.

*Deparaffinization, rehydration, and antigen Retrieval***:** The combined process of deparaffinization, rehydration and antigen retrieval of the tissue sections was done using 1x Envision Target Retrieval Solution, High PH (Dako, USA) in the PT-Link Pre-Treatment instrument (Dako, USA). Thereafter, slides were incubated for 1 minute in distilled water and equilibrated in EnVision^{TM} FLEX Wash Buffer (Dako, USA) at room temperature for 10 minutes (2x 5 minutes using fresh buffer for each period) for macrophage markers staining and 5 minutes for fibroblast markers staining. Then, the macrophage slides were incubated in EnVision^{TM} FLEX Peroxidase blocking solution (Dako, USA) for 10 minutes and fibroblast slides were incubated in Peroxidase blocking solution (PerkinElmer) for 10 minutes, both then washed in wash buffer (Dako, USA) as before immediately at room temperature.

Background reduction: The macrophage slides were incubated in blocking buffer (0.05g BSA + 10% goat serum dissolved in EnVision^{TM} FLEX Wash Buffer) for 20 minutes. The fibroblast slides were incubated in Bloxall blocking solution (PerkinElmer) for 10 minutes.

Antibody staining: The macrophage slides were incubated in primary antibody-1 for 45 minutes, fibroblast slides for 30 minutes, at room temperature, then washed for 5 minutes in wash buffer. The macrophage slides were then incubated in Secondary Opal Polymer Horseradish Peroxidase (HRP) Mouse and Rabbit (AKOYA, USA) for 20 minutes and fibroblast slides were incubated in Secondary Opal Polymer Horseradish Peroxidase (HRP) Mouse and Rabbit (PerkinElmer) for 30 minutes. Then, the slides were washed twice in wash buffer as before, drained and the sections were incubated in Opal Polymer Fluorophore (macrophage slides: AKOYA, USA; fibroblast slides: PerkinElmer) working solution for signal amplification at room temperature for 10 minutes in the dark. The slides were then washed for 10 minutes (2x 5 minutes using fresh buffer for each time) for macrophage slides, 5 min for fibroblast slides, in wash buffer at room temperature.

Antibody stripping: Afterwards, the antigen retrieval via microwave treatment was done by placing the slides in a slide jar with pre-warmed buffer AR6 (macrophage slides: AKOYA, USA; fibroblast slides: PerkinElmer). The jar was loosely covered and placed in a microwave for 2 minutes at 100% power (High setting), 10 minutes at 50% (Medium setting) power and 5 minutes at 20% (low setting) power. Slides were cooled down in the dark by placing the slide jar on ice for 20 minutes and the slides were rinsed in distilled water, followed by incubation in the wash buffer for 10 (2x 5 minutes) minutes for macrophage slides, 5 minutes for fibroblast slides, to equilibrate slides. The microwave step re-exposes the antigen on the tissue and allows the introduction of the next primary antibody. For the detection of the next target (primary antibody 2), the protocol was restarted at the blocking step using blocking buffer (macrophage slides: 0.05g BSA + 10% goat serum dissolved in EnVision^{TM} FLEX Wash Buffer; fibroblast slides: Bloxall blocking solution from PerkinElmer) for 10 minutes. After the third target was detected (primary antibody 3), a working solution of DAPI (macrophage slides: AKOYA, USA; fibroblast slides: PerkinElmer) was applied to the sections as the nuclear counterstain for 5 minutes in a humidity chamber. The slides were washed in wash buffer for 5 minutes, then in distilled water for 5 minutes and drained. Then, the sections were coverslip with Fluorescence Mounting Medium (Agilent Technologies, Inc.) and the edges of the coverslip were sealed with nail varnish. Slides were stored in a humidity chamber at 4^{o}C until images are acquired.

*Antibodies and fluorophores***:** For macrophage slides, the unconjugated primary antibodies used are Anti-CD68 (conc. clone: Ab213363, Abcam), Anti-CD206 (clone: Abcam), Anti-Osteopontin (clone: ab302942, Abcam). The primary antibodies were diluted in antibody diluent (PerkinElmer) as recommended by the antibody manufacturer, and the Opal fluorophores diluted in amplification diluent (PerkinElmer). The fluorophores used for signal generation in this study are FITC, Tetramethylrhodamine (TRITC), and Cy5. For fibroblast slides, the unconjugated primary antibodies used are Anti-Collagen I (clone: ab34710, Abcam), Anti- Anti-CTHRC1 (clone: ab85739, Abcam), Anti-TDO2 (clone: OT14G2, Thermo Fisher Scientific), Anti-PI15 (clone: PA5-52312, Thermo Fisher Scientific), Anti-ACTA2 (clone: 1A4, LSBio). The primary antibodies were diluted in antibody diluent (PerkinElmer) as recommended by the antibody manufacturer, and the Opal fluorophores diluted in amplification diluent (PerkinElmer). The fluorophores used for signal generation in this study are FITC, Texas-Red and Cy5.
*Imaging***:** For macrophage slides, the images were acquired on Hamamatsu NanoZoomer S60, (Japan) and analysed with NDP.view2 (version 2.9.29) imaging software (TissueGnostics). For fibroblast slides, the images were acquired on a Zeiss Axio Observer Z1 inverted microscope (Olympus) and analysed with TissueFAXS imaging software (TissueGnostics).

Quantification: for macrophage slides, Using QuPath software (version 0.5.0-x64), TB granulomas were segmented into three distinct layers:

1. An **outer cellular layer** primarily composed of lymphocytes,
2. An **inner cellular layer** predominantly consisting of myeloid cells (mainly macrophages),
3. A **necrotic core** characterized by cellular debris and dead cells.

To assess the expression levels of SPP1 (Green) and CD68 (Yellow), we divided the granuloma radially similar to pie-cutting, into ten regions, which further divides each granuloma layer into ten sub-regions. Each sub-region is defined and analyzed as a separate Region of Interest (ROI) where they are numbered clockwise from 1 to 10 so ROI1 from the necrotic core is immediately adjacent to ROI1 from the inner cellular layer, etc. This enabled a more precise evaluation of spatial differences in protein expression. Expression intensity was measured at 5 µm per pixel to capture variability in protein distribution across the granuloma architecture.

Mean intensity for each ROI was used to calculate the statistical significance between the groups using two-sided Mann-Whitney U test without correction for SPP1. The ratio of mean intensity between inner cellular layer and outer cellular layer was calculated between the ROI1 and ROI1, ROI2 and ROI2, etc. The same was done for the ratio of mean intensity between inner cellular layer and the necrotic core.

Flow Cytometry

Lung pieces collected after removal from *M.tb* infected patients were used in flow cytometry analysis after processing as from scRNA-seq (**Table S4**). Cells were counted and stained with antibody cocktail for 30 minutes at RT and in the dark followed by 2x wash steps with PBS and resuspension of stained cells in FACSLyse. The surface markers used were: CD45 (CD45-APC, BioLegend, cat#304012), CD34 (CD34-FITC, BioLegend, cat#324226), EpCAM (EpCAM-BV650, BioLegend, cat#324204), CD11b (CD11b-PeCy7, BD, cat#557743), CD31 (CD31-BV605, BioLegend, cat#303121), VCAM1 (VCAM-PE, BioLegend, cat#305805), ICAM1 (ICAM-APC-Cy7, BioLegend, cat#353121), podoplanin (PDPN-PERCPefluor710, ThermoFischer, cat#46-9381-42) and CD235a (CD235a-PECF594, BioLegend, cat#349119). Viability was determined using the Invitrogen Live/Dead Aqua Fluorescent reactive dye on the HV500 channel. Samples were acquired on the BDFACS Aria Fusion flow cytometer. Analysis of samples was subsequently carried out using FlowJo (v10, FlowJo).

The Friedman test was used to assess significant changes in the fibroblast subset of interest across different lung resection severities.

Human lung tissue spatial transcriptomics data analysis

Filtered 10x spatial data from each sample processed by *Space Ranger* was read into an anndata object using the function "*read*_visium" from python package *scanpy*, along with the corresponding high-resolution image of the H&E stain. No filtering on spots or genes was done to preserve the maximum amount of information as the nature of these data are intrinsically sparse. Data was log-normalized with standard *scanpy* workflow. Top 50 hub genes from the hdWGCNA Fibroblast-M1 module were used to score for human TB-myofibroblast signature on each Visium spot using *score_genes* from *scanpy*. The top 20 differentially expressed markesr from the macrophage subsets (**Data File S2**) were used to calculate DE marker scores in a similar fashion. Spearman's correlation and its significance between the human TB-myofibroblast

signature and macrophage subset markers were calculated using the function *spearmanr* from python package *scipy*. A two-sided Mann-Whitney U test without correction was used to compare module usage between different conditions and Spearman's correlation was used to compare human TB-myofibroblast and different macrophage subsets.

Deconvolution of spatial transcriptomic data

Since Visium v2 chemistry has spot size of diameter = 55μm (generally larger than one cell), we estimated the cell type abundance of each spot using python package *cell2location*, a Bayesian model that estimates the combination and abundance of cell types that could give rise to the mRNA counts in each spatial location. We first learned reference signatures of each broad-level cell type from the original scRNA-seq cohort generated using Seq-Well S^3, then decomposed the Visium multi-cell RNA counts into these reference signatures, establishing a spatial mapping of cell types. For training the reference signatures, we used patient ID as categorical_covariate_*keys* and sequencing batch as *batch_key, num_samples=1,000, batch_size=2,500, max_epochs=250,* with the rest set to default. For the posterior estimating, we created and trained the model with hyperparameters: *N_cells_per_location*=10, *detection_alpha*=20, *max_epoches*=15,000, with the rest set to default. For each boxplot comparing cell type abundance, 5% quantile of the estimated posterior was used to represent cell type abundance at each Visium spot, which represents the value of abundance the model has high confidence in.

Annotation of granuloma structures on H&E images

Granuloma structures were manually annotated using ImageJ by experts in TB lung pathology. A band of ~20-pixel width was then drawn outside the selection area to approximate the granuloma cuff. The spots from the Visium data are categorized to be "in", "on" and "out" of the granuloma border based on the spot's corresponding position relative to this segmentation band.

Spatial Ligand-Receptor Analysis

Each sample was log-normalized with the *scanpy* package. The python version of *LIANA* package was then used to impute spatial ligand-receptor interactions. First, spatial neighborhoods were calculated using the *spatial_neighbors* with *bandwidth=10, cutoff=0.1, kernel='gaussian', set_diag=True*. Then bivariate scores for potential ligand-receptor pairs is calculated using the function *bivariate* using with parameters *resource_name='consensus', local_name='cosine'* (spatially-weighted cosine similarity for local score), *global_name='morans'* (bivariate Moran's R for global score), *n_perms=100*, *nz_prop=0.05*, and default settings for the rest. Top L-R pairs are selected by sorting for descending Moran's R as it describes global co-expression.

List of Supplementary Materials

Materials and Methods
Figs. S1 to S11
Tables S1 to S5

Data files S1 to S5

Reference 1

References and Notes

1. Global tuberculosis report 2023 (available at https://www.who.int/publications-detail-redirect/9789240083851).

2. C. Carranza, S. Pedraza-Sanchez, E. de Oyarzabal-Mendez, M. Torres, Diagnosis for Latent Tuberculosis Infection: New Alternatives. *Front. Immunol.* **11**, 2006 (2020).

3. Global tuberculosis report 2020 (2024) (available at https://www.who.int/publications/i/item/9789240013131).

4. M. Pai, T. Kasaeva, S. Swaminathan, Covid-19's Devastating Effect on Tuberculosis Care - A Path to Recovery. *N. Engl. J. Med.* **386**, 1490–1493 (2022).

5. E. F. McCaffrey, M. Donato, L. Keren, Z. Chen, A. Delmastro, M. B. Fitzpatrick, S. Gupta, N. F. Greenwald, A. Baranski, W. Graf, R. Kumar, M. Bosse, C. C. Fullaway, P. K. Ramdial, E. Forgó, V. Jojic, D. Van Valen, S. Mehra, S. A. Khader, S. C. Bendall, M. van de Rijn, D. Kalman, D. Kaushal, R. L. Hunter, N. Banaei, A. J. C. Steyn, P. Khatri, M. Angelo, The immunoregulatory landscape of human tuberculosis granulomas. *Nat. Immunol.* **23**, 318–329 (2022).

6. J. M. Davis, L. Ramakrishnan, The Role of the Granuloma in Expansion and Dissemination of Early Tuberculous Infection. *Cell* **136**, 37–49 (2009).

7. S. Ravimohan, H. Kornfeld, D. Weissman, G. P. Bisson, Tuberculosis and lung damage: from epidemiology to pathophysiology. *Eur. Respir. Rev. Off. J. Eur. Respir. Soc.* **27** (2018), doi:10.1183/16000617.0077-2017.

8. B. W. Allwood, A. Byrne, J. Meghji, A. Rachow, M. M. van der Zalm, O. D. Schoch, Post-Tuberculosis Lung Disease: Clinical Review of an Under-Recognised Global Challenge. *Respiration* **100**, 751–763 (2021).

9. K. Romanowski, B. Baumann, C. A. Basham, F. Ahmad Khan, G. J. Fox, J. C. Johnston, Long-term all-cause mortality in people treated for tuberculosis: a systematic review and meta-analysis. *Lancet Infect. Dis.* **19**, 1129–1137 (2019).

10. K. L. Fonseca, P. N. S. Rodrigues, I. A. S. Olsson, M. Saraiva, Experimental study of tuberculosis: From animal models to complex cell systems and organoids. *PLoS Pathog.* **13**, e1006421 (2017).

11. R. L. Hunter, The Pathogenesis of Tuberculosis: The Early Infiltrate of Post-primary (Adult Pulmonary) Tuberculosis: A Distinct Disease Entity. *Front. Immunol.* **9**, 2108 (2018).

12. Y. Cui, Y. Zheng, X. Liu, L. Yan, X. Fan, J. Yong, Y. Hu, J. Dong, Q. Li, X. Wu, S. Gao, J. Li, L. Wen, J. Qiao, F. Tang, Single-Cell Transcriptome Analysis Maps the Developmental Track of the Human Heart. *Cell Rep.* **26**, 1934-1950.e5 (2019).

13. T. S. Adams, J. C. Schupp, S. Poli, E. A. Ayaub, N. Neumark, F. Ahangari, S. G. Chu, B. A. Raby, G. DeIuliis, M. Januszyk, Q. Duan, H. A. Arnett, A. Siddiqui, G. R. Washko, R. Homer, X. Yan, I. O. Rosas, N. Kaminski, Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Sci. Adv.* **6**, eaba1983 (2020).

14. T. R. Hawn, A. I. Matheson, S. N. Maley, O. Vandal, Host-directed therapeutics for tuberculosis: can we harness the host? *Microbiol. Mol. Biol. Rev. MMBR* **77**, 608–627 (2013).

15. M. T. Reichmann, L. B. Tezera, A. F. Vallejo, M. Vukmirovic, R. Xiao, J. Reynolds, S. Jogai, S. Wilson, B. Marshall, M. G. Jones, A. Leslie, J. M. D'Armiento, N. Kaminski, M. E. Polak, P. Elkington, Integrated transcriptomic analysis of human tuberculosis granulomas and a biomimetic model identifies therapeutic targets. *J. Clin. Invest.* **131** (2021), doi:10.1172/JCI148136.

16. H. P. Gideon, T. K. Hughes, C. N. Tzouanas, M. H. 2nd Wadsworth, A. A. Tu, T. M. Gierahn, J. M. Peters, F. F. Hopkins, J.-R. Wei, C. Kummerlowe, N. L. Grant, K. Nargan, J. Y. Phuah, H. J. Borish, P. Maiello, A. G. White, C. G. Winchell, S. K. Nyquist, S. K. C. Ganchua, A. Myers, K. V. Patel, C. L. Ameel, C. T. Cochran, S. Ibrahim, J. A. Tomko, L. J. Frye, J. M. Rosenberg, A. Shih, M. Chao, E. Klein, C. A. Scanga, J. Ordovas-Montanes, B. Berger, J. T. Mattila, R. Madansein, J. C. Love, P. L. Lin, A. Leslie, S. M. Behar, B. Bryson, J. L. Flynn, S. M. Fortune, A. K. Shalek, Multimodal profiling of lung granulomas in macaques reveals cellular correlates of tuberculosis control. *Immunity* **55**, 827-846.e10 (2022).

17. L. Sikkema, C. Ramírez-Suástegui, D. C. Strobl, T. E. Gillett, L. Zappia, E. Madissoon, N. S. Markov, L. E. Zaragosi, Y. Ji, M. Ansari, M. J. Arguel, L. Apperloo, M. Banchero, C. Bécavin, M. Berg, E. Chichelnitskiy, M. i. Chung, A. Collin, A. C. A. Gay, J. Gote-Schniering, B. Hooshiar Kashani, K. Inecik, M. Jain, T. S. Kapellos, T. M. Kole, S. Leroy, C. H. Mayr, A. J. Oliver, M. von Papen, L. Peter, C. J. Taylor, T. Walzthoeni, C. Xu, L. T. Bui, C. De Donno, L. Dony, A. Faiz, M. Guo, A. J. Gutierrez, L. Heumos, N. Huang, I. L. Ibarra, N. D. Jackson, P. Kadur Lakshminarasimha Murthy, M. Lotfollahi, T. Tabib, C. Talavera-López, K. J. Travaglini, A. Wilbrey-Clark, K. B. Worlock, M. Yoshida, Y. Chen, J. S. Hagood, A. Agami, P. Horvath, J. Lundeberg, C. H. Marquette, G. Pryhuber, C. Samakovlis, X. Sun, L. B. Ware, K. Zhang, M. van den Berge, Y. Bossé, T. J. Desai, O. Eickelberg, N. Kaminski, M. A. Krasnow, R. Lafyatis, M. Z. Nikolic, J. E. Powell, J. Rajagopal, M. Rojas, O. Rozenblatt-Rosen, M. A. Seibold, D. Sheppard, D. P. Shepherd, D. D. Sin, W. Timens, A. M. Tsankov, J. Whitsett, Y. Xu, N. E. Banovich, P. Barbry, T. E. Duong, C. S. Falk, K. B. Meyer, J. A. Kropski, D. Pe'er, H. B. Schiller, P. R. Tata, J. L. Schultze, S. A. Teichmann, A. V. Misharin, M. C. Nawijn, M. D. Luecken, F. J. Theis, An integrated cell atlas of the lung in health and disease. *Nat. Med.* **29** (2023), doi:10.1038/s41591-023-02327-2.

18. A. Ardain, R. Domingo-Gonzalez, S. Das, S. W. Kazer, N. C. Howard, A. Singh, M. Ahmed, S. Nhamoyebonde, J. Rangel-Moreno, P. Ogongo, L. Lu, D. Ramsuran, M. de la Luz Garcia-Hernandez, T. K Ulland, M. Darby, E. Park, F. Karim, L. Melocchi, R. Madansein, K. J. Dullabh, M. Dunlap, N. Marin-Agudelo, T. Ebihara, T. Ndung'u, D. Kaushal, A. S. Pym, J. K. Kolls, A. Steyn, J. Zúñiga, W. Horsnell, W. M. Yokoyama, A. K. Shalek, H. N. Kløverpris, M. Colonna, A. Leslie, S. A. Khader, Group 3 innate lymphoid cells mediate early protective immunity against tuberculosis. *Nature* **570**, 528–532 (2019).

19. T. K. Hughes, M. H. Wadsworth, T. M. Gierahn, T. Do, D. Weiss, P. R. Andrade, F. Ma, B. J. de Andrade Silva, S. Shao, L. C. Tsoi, J. Ordovas-Montanes, J. E. Gudjonsson, R. L. Modlin, J. C. Love, A. K. Shalek, Second-Strand Synthesis-Based Massively Parallel scRNA-Seq Reveals Cellular States and Molecular Features of Human Inflammatory Skin Pathologies. *Immunity* **53**, 878-894.e7 (2020).

20. Visium Spatial Gene Expression Reagent Kits for FFPE User Guide*10x Genomics* (available at https://www.10xgenomics.com/support/spatial-gene-expression-ffpe/documentation/steps/library-construction/visium-spatial-gene-expression-reagent-kits-for-ffpe-user-guide).

21. P. A. Reyfman, J. M. Walter, N. Joshi, K. R. Anekalla, A. C. McQuattie-Pimentel, S. Chiu, R. Fernandez, M. Akbarpour, C.-I. Chen, Z. Ren, R. Verma, H. Abdala-Valencia, K. Nam, M. Chi, S. Han, F. J. Gonzalez-Gonzalez, S. Soberanes, S. Watanabe, K. J. N. Williams, A. S. Flozak, T. T. Nicholson, V. K. Morgan, D. R. Winter, M. Hinchcliff, C. L. Hrusch, R. D. Guzy, C. A. Bonham, A. I. Sperling, R. Bag, R. B. Hamanaka, G. M. Mutlu, A. V. Yeldandi, S. A. Marshall, A. Shilatifard, L. A. N. Amaral, H. Perlman, J. I. Sznajder, A. C. Argento, C. T. Gillespie, J. Dematte, M. Jain, B. D. Singer, K. M. Ridge, A. P. Lam, A. Bharat, S. M. Bhorade, C. J. Gottardi, G. R. S. Budinger, A. V. Misharin, Single-Cell Transcriptomic

Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis. *Am. J. Respir. Crit. Care Med.* **199**, 1517–1536 (2019).

22. C. N. Muefong, J. S. Sutherland, Neutrophils in Tuberculosis-Associated Inflammation and Lung Pathology. *Front. Immunol.* **11**, 962 (2020).

23. J. S. Sutherland, D. J. Jeffries, S. Donkor, B. Walther, P. C. Hill, I. M. O. Adetifa, R. A. Adegbola, M. O. C. Ota, High granulocyte/lymphocyte ratio and paucity of NKT cells defines TB disease in a TB-endemic setting. *Tuberc. Edinb. Scotl.* **89**, 398–404 (2009).

24. G. R. Leisching, Susceptibility to Tuberculosis Is Associated With PI3K-Dependent Increased Mobilization of Neutrophils. *Front. Immunol.* **9**, 1669 (2018).

25. C. J. Queval, R. Brosch, R. Simeone, The macrophage: A disputed fortress in the battle against Mycobacterium tuberculosis. *Front. Microbiol.* **8** (2017), doi:10.3389/fmicb.2017.02284.

26. C. G. Winchell, S. K. Nyquist, M. C. Chao, P. Maiello, A. J. Myers, F. Hopkins, M. Chase, H. P. Gideon, K. V. Patel, J. D. Bromley, A. W. Simonson, R. Floyd-O'Sullivan, M. Wadsworth, J. M. Rosenberg, R. Uddin, T. Hughes, R. J. Kelly, J. Griffo, J. Tomko, E. Klein, B. Berger, C. A. Scanga, J. Mattila, S. M. Fortune, A. K. Shalek, P. L. Lin, J. A. L. Flynn, CD8+ lymphocytes are critical for early control of tuberculosis in macaques. *J. Exp. Med.* **220** (2023), doi:10.1084/JEM.20230707.

27. C. S. Smillie, M. Biton, J. Ordovas-Montanes, K. M. Sullivan, G. Burgin, D. B. Graham, R. H. Herbst, N. Rogel, M. Slyper, J. Waldman, M. Sud, E. Andrews, G. Velonias, A. L. Haber, K. Jagadeesh, S. Vickovic, J. Yao, C. Stevens, D. Dionne, L. T. Nguyen, A.-C. Villani, M. Hofree, E. A. Creasey, H. Huang, O. Rozenblatt-Rosen, J. J. Garber, H. Khalili, A. N. Desch, M. J. Daly, A. N. Ananthakrishnan, A. K. Shalek, R. J. Xavier, A. Regev, Intra- and Inter-cellular Rewiring of the Human Colon during Ulcerative Colitis. *Cell* **178**, 714-730.e22 (2019).

28. H. P. Gideon, J. Phuah, B. A. Junecko, J. T. Mattila, Neutrophils express pro- and anti-inflammatory cytokines in granulomas from Mycobacterium tuberculosis-infected cynomolgus macaques. *Mucosal Immunol.* **12**, 1370–1381 (2019).

29. E. Gaffney, D. Murphy, A. Walsh, S. Connolly, S. A. Basdeo, J. Keane, J. J. Phelan, Defining the role of neutrophils in the lung during infection: Implications for tuberculosis disease. *Front. Immunol.* **13** (2022), doi:10.3389/fimmu.2022.984293.

30. P. Chandra, S. J. Grigsby, J. A. Philips, Immune evasion and provocation by Mycobacterium tuberculosis. *Nat. Rev. Microbiol.* **20**, 750–766 (2022).

31. The Formation and Function of Granulomas | Annual Reviews (available at https://www.annualreviews.org/content/journals/10.1146/annurev-immunol-032712-100022).

32. A. D. Luster, R. Alon, U. H. von Andrian, Immune cell migration in inflammation: present and future therapeutic targets. *Nat. Immunol.* **6**, 1182–1190 (2005).

33. L. R. Prince, L. Allen, E. C. Jones, P. G. Hellewell, S. K. Dower, M. K. B. Whyte, I. Sabroe, The Role of Interleukin-1β in Direct and Toll-Like Receptor 4-Mediated Neutrophil Activation and Survival. *Am. J. Pathol.* **165**, 1819–1826 (2004).

34. S. de Oliveira, C. C. Reyes-Aldasoro, S. Candel, S. A. Renshaw, V. Mulero, Â. Calado, Cxcl8 (Interleukin-8) mediates neutrophil recruitment and behavior in the zebrafish inflammatory response. *J. Immunol. Baltim. Md 1950* **190**, 4349–4359 (2013).

35. D. E. Zak, A. Penn-Nicholson, T. J. Scriba, E. Thompson, S. Suliman, L. M. Amon, H. Mahomed, M. Erasmus, W. Whatney, G. D. Hussey, D. Abrahams, F. Kafaar, T. Hawkridge, S. Verver, E. J. Hughes, M. Ota, J. Sutherland, R. Howe, H. M. Dockrell, W. H. Boom, B. Thiel, T. H. M. Ottenhoff, H. Mayanja-Kizza, A. C. Crampin, K. Downing, M. Hatherill, J. Valvo, S. Shankar, S. K. Parida, S. H. E. Kaufmann, G. Walzl, A. Aderem, W. A. Hanekom, A blood RNA signature for tuberculosis disease risk: a prospective cohort study. *Lancet Lond. Engl.* **387**, 2312–2322 (2016).

36. U. K. Rohlwink, N. F. Walker, A. A. Ordonez, Y. J. Li, E. W. Tucker, P. T. Elkington, R. J. Wilkinson, K. A. Wilkinson, Matrix Metalloproteinases in Pulmonary and Central Nervous System Tuberculosis-A Review. *Int. J. Mol. Sci.* **20** (2019), doi:10.3390/ijms20061350.

37. S. S. Y. Huang, M. Toufiq, L. R. Saraiva, N. Van Panhuys, D. Chaussabel, M. Garand, Transcriptome and literature mining highlight the differential expression of erlin1 in immune cells during sepsis. *Biology* **10** (2021), doi:10.3390/biology10080755.

38. G. Krishnamoorthy, P. Kaiser, U. Abu Abed, J. Weiner, P. Moura-Alves, V. Brinkmann, S. H. E. Kaufmann, FX11 limits Mycobacterium tuberculosis growth and potentiates bactericidal activity of isoniazid through host-directed activity. *Dis. Model. Mech.* **13**, dmm041954 (2020).

39. C. S. Chowdhury, E. Wareham, J. Xu, S. Kumar, M. Kofron, S. Lakshmikanthan, M. Chrzanowska, M. D. Filippi, Rap1b-loss increases neutrophil lactate dehydrogenase activity to enhance neutrophil migration and acute inflammation in vivo. *Front. Immunol.* **13** (2022), doi:10.3389/fimmu.2022.1061544.

40. F. J. Duffy, J. Weiner, S. Hansen, D. L. Tabb, S. Suliman, E. Thompson, J. Maertzdorf, S. Shankar, G. Tromp, S. Parida, D. Dover, M. K. Axthelm, J. S. Sutherland, H. M. Dockrell, T. H. M. Ottenhoff, T. J. Scriba, L. J. Picker, G. Walzl, S. H. E. Kaufmann, D. E. Zak, R. Golinski, M. Jacobson, G. McEwen, G. F. Black, G. Van Der Spuy, K. Stanley, M. Kriel, N. DuPlessis, N. Nene, A. G. Loxton, N. N. Chegou, M. Fisher, H. Mahomed, J. Hughes, K. Downing, A. Penn-Nicholson, H. Mulenga, B. Abel, M. Bowmaker, B. Kagina, W. Kwong, C. W. Hanekom, M. R. Klein, M. C. Haks, K. L. F ranken, A. Geluk, K. E. Van Meijgaarden, S. A. Joosten, D. Van Baarle, F. Miedema, W. H. Boom, B. Thiel, J. Sadoff, D. Sizemore, S. Ramachandran, L. Barker, M. Brennan, F. Weichold, S. Muller, L. Geiter, G. Schoolnik, G. Dolganov, T. Van, H. Mayanja-Kizza, M. Joloba, S. Zalwango, M. Nsereko, B. Okwera, H. Kisingo, S. Smith, P. Gorak-Stolinska, Y. G. Hur, M. Lalor, J. S. Lee, A. C. Crampin, N. French, B. Ngwira, A. B. Smith, K. Watkins, L. Ambrose, F. Simukonda, H. Mvula, F. Chilongo, J. Saul, K. Branson, D. Kassa, A. Abebe, T. Mesele, B. Tegbaru, R. Howe, A. Mihret, A. Aseffa, Y. Bekele, R. Iwnetu, M. Tafesse, L. Yamuah, M. Ota, J. Sutherland, P. Hill, R. Adegbola, T. Corrah, M. Antonio, T. Togun, I. Adetifa, S. Donkor, P. Andersen, I. Rosenkrands, M. Doherty, K. Weldingh, Immunometabolic signatures predict risk of progression to active tuberculosis and disease outcome. *Front. Immunol.* **10** (2019), doi:10.3389/fimmu.2019.00527.

41. K. W. Tan, S. Z. Chong, F. H. S. Wong, M. Evrard, S. M.-L. Tan, J. Keeble, D. M. Kemeny, L. G. Ng, J.-P. Abastado, V. Angeli, Neutrophils contribute to inflammatory lymphangiogenesis by increasing VEGF-A bioavailability and secreting VEGF-D. *Blood* **122**, 3666–3677 (2013).

42. W. Adams, T. Espicha, J. Estipona, Getting Your Neutrophil: Neutrophil Transepithelial Migration in the Lung. *Infect. Immun.* **89**, 10.1128/iai.00659-20 (2021).

43. Y. He, K. B. V. Døssing, M. Rossing, F. O. Bagger, A. Kjaer, uPAR (PLAUR) Marks Two Intra-Tumoral Subtypes of Glioblastoma: Insights from Single-Cell RNA Sequencing. *Int. J. Mol. Sci.* **25**, 1998 (2024).

44. X. Zhou, X. Zhu, J. Yao, X. Wang, N. Wang, Comprehensive analysis of clinical prognosis and molecular immune characterization of tropomyosin 4 in pancreatic cancer. *Invest. New Drugs* **39**, 1469–1483 (2021).

45. K. D. Kauffman, M. A. Sallin, S. Sakai, O. Kamenyeva, J. Kabat, D. Weiner, M. Sutphin, D. Schimel, L. Via, C. E. 3rd Barry, T. Wilder-Kofie, I. Moore, R. Moore, D. L. Barber, Defective positioning in granulomas but not lung-homing limits CD4 T-cell interactions with Mycobacterium tuberculosis-infected macrophages in rhesus macaques. *Mucosal Immunol.* **11**, 462–473 (2018).

46. L. Zheng, M. He, M. Long, R. Blomgran, O. Stendahl, Pathogen-Induced Apoptotic Neutrophils Express Heat Shock Proteins and Elicit Activation of Human Macrophages1. *J. Immunol.* **173**, 6319–6326 (2004).

47. S. K. C. Ganchua, A. G. White, E. C. Klein, J. L. Flynn, Lymph nodes-The neglected battlefield in tuberculosis. *PLoS Pathog.* **16**, e1008632 (2020).

48. D. Mekonnen, A. Derbie, A. Mihret, S. A. Yimer, T. Tønjum, B. Gelaw, E. Nibret, A. Munshae, S. J. Waddell, A. Aseffa, Lipid droplets and the transcriptome of Mycobacterium tuberculosis from direct sputa: a literature review. *Lipids Health Dis.* **20**, 129 (2021).

49. A. P. Santos, L. S. Rodrigues, N. Rother, F. C. de Q. Mello, C. Magis-Escurra, The role of neutrophil response in lung damage and post-tuberculosis lung disease: a translational narrative review. *Front. Immunol.* **16**, 1528074 (2025).

50. E. Guirado, L. S. Schlesinger, G. Kaplan, Macrophages in tuberculosis: friend or foe. *Semin. Immunopathol.* **35**, 563–583 (2013).

51. A. Kishore, M. Petrek, Roles of Macrophage Polarization and Macrophage-Derived miRNAs in Pulmonary Fibrosis. *Front. Immunol.* **12**, 678457 (2021).

52. P. Sampath, K. Moideen, U. D. Ranganathan, R. Bethunaickan, Monocyte Subsets: Phenotypes and Function in Tuberculosis Infection. *Front. Immunol.* **9**, 1726 (2018).

53. C.-H. Wang, P.-C. Chou, F.-T. Chung, H.-C. Lin, K.-H. Huang, H.-P. Kuo, Heat shock protein70 is implicated in modulating NF-κB activation in alveolar macrophages of patients with active pulmonary tuberculosis. *Sci. Rep.* **7**, 1214 (2017).

54. J. Radons, The human HSP70 family of chaperones: where do we stand? *Cell Stress Chaperones* **21**, 379–404 (2016).

55. G. Desalegn, O. Pabst, Inflammation triggers immediate rather than progressive changes in monocyte differentiation in the small intestine. *Nat. Commun.* **10**, 3229 (2019).

56. E. Dann, N. C. Henderson, S. A. Teichmann, M. D. Morgan, J. C. Marioni, Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nat. Biotechnol.* **40**, 245–253 (2022).

57. J. Tang, Y. Shi, L. Zhan, C. Qin, Downregulation of GPR183 on infection restricts the early infection and intracellular replication of mycobacterium tuberculosis in macrophage. *Microb. Pathog.* **145**, 104234 (2020).

58. D. Wang, X. Tong, L. Wang, S. Zhang, J. Huang, L. Zhang, H. Fan, The association between osteopontin and tuberculosis: A systematic review and meta-analysis. *PLOS ONE* **15**, e0242702 (2020).

59. G. J. Nau, P. Guilfoile, G. L. Chupp, J. S. Berman, S. J. Kim, H. Kornfeld, R. A. Young, A chemoattractant cytokine associated with granulomas in tuberculosis and silicosis. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 6414–6419 (1997).

60. J. D. Bromley, S. K. C. Ganchua, S. K. Nyquist, P. Maiello, M. Chao, H. J. Borish, M. Rodgers, J. Tomko, K. Kracinovsky, D. Mugahid, S. Nguyen, D. Wang, J. M. Rosenberg, E. C. Klein, H. P. Gideon, R. Floyd-O'Sullivan, B. Berger, C. A. Scanga, P. L. Lin, S. M. Fortune, A. K. Shalek, J. L. Flynn, CD4+ T cells are homeostatic regulators during Mtb reinfection. *bioRxiv* , 2023.12.20.572669 (2023).

61. J. H. Campbell, A. C. Hearps, G. E. Martin, K. C. Williams, S. M. Crowe, The importance of monocytes and macrophages in HIV pathogenesis, treatment, and cure. *AIDS Lond. Engl.* **28**, 2175–2187 (2014).

62. T. H. Burdo, A. Lackner, K. C. Williams, Monocyte/macrophages and their role in HIV neuropathogenesis. *Immunol. Rev.* **254**, 102–113 (2013).

63. H.-S. Lee, H.-S. Hua, C.-H. Wang, M.-C. Yu, B.-C. Chen, C.-H. Lin, Mycobacterium tuberculosis induces connective tissue growth factor expression through the TLR2-JNK-AP-1 pathway in human lung fibroblasts. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **33**, 12554–12564 (2019).

64. X. Gai, B. Allwood, Y. Sun, Post-tuberculosis lung disease and chronic obstructive pulmonary disease. *Chin. Med. J. (Engl.)* **136**, 1923–1928 (2023).

65. K. J. Travaglini, A. N. Nabhan, L. Penland, R. Sinha, A. Gillich, R. V. Sit, S. Chang, S. D. Conley, Y. Mori, J. Seita, G. J. Berry, J. B. Shrager, R. J. Metzger, C. S. Kuo, N. Neff, I. L. Weissman, S. R. Quake, M. A. Krasnow, A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* **587**, 619–625 (2020).

66. K. R. Stenmark, E. Nozik-Grayck, E. Gerasimovskaya, A. Anwar, M. Li, S. Riddle, M. Frid, The adventitia: Essential role in pulmonary vascular remodeling. *Compr. Physiol.* **1**, 141–161 (2011).

67. A. Anwar, M. Li, M. G. Frid, B. Kumar, E. V. Gerasimovskaya, S. R. Riddle, B. Alexandre McKeon, R. Thukaram, B. O. Meyrick, M. A. Fini, K. R. Stenmark, Osteopontin is an endogenous modulator of the constitutively activated phenotype of pulmonary adventitial fibroblasts in hypoxic pulmonary hypertension. *Am. J. Physiol. - Lung Cell. Mol. Physiol.* **303** (2012), doi:10.1152/ajplung.00050.2012.

68. K. C. El Kasmi, S. C. Pugliese, S. R. Riddle, J. M. Poth, A. L. Anderson, M. G. Frid, M. Li, S. S. Pullamsetti, R. Savai, M. A. Nagel, M. A. Fini, B. B. Graham, R. M. Tuder, J. E. Friedman, H. K. Eltzschig, R. J. Sokol, K. R. Stenmark, Adventitial Fibroblasts Induce a Distinct Proinflammatory/Profibrotic Macrophage Phenotype in Pulmonary Hypertension. *J. Immunol.* **193** (2014), doi:10.4049/jimmunol.1303048.

69. H. E. Talbott, S. Mascharak, M. Griffin, D. C. Wan, M. T. Longaker, Wound healing, fibroblast heterogeneity, and fibrosis. *Cell Stem Cell* **29**, 1161–1180 (2022).

70. M. Li, F. Luan, Y. Zhao, H. Hao, Y. Zhou, W. Han, X. Fu, Epithelial-mesenchymal transition: An emerging target in tissue fibrosis. *Exp. Biol. Med.* **241** (2016), doi:10.1177/1535370215597194.

71. E. El Agha, A. Moiseenko, V. Kheirollahi, S. De Langhe, S. Crnkovic, G. Kwapiszewska, M. Szibor, D. Kosanovic, F. Schwind, R. T. Schermuly, I. Henneke, B. MacKenzie, J. Quantius, S. Herold, A. Ntokou, K. Ahlbrecht, T. Braun, R. E. Morty, A. Günther, W. Seeger, S. Bellusci, Two-Way Conversion between Lipogenic and Myogenic Fibroblastic Phenotypes Marks the Progression and Resolution of Lung Fibrosis. *Cell Stem Cell* **20**, 261-273.e3 (2017).

72. X. Zhao, P. Shen, H. Li, Y. Yang, J. Guo, S. Chen, Y. Ma, J. Sheng, S. Shen, G. Liu, X. Fang, Carbonic Anhydrase 12 Protects Endplate Cartilage From Degeneration Regulated by IGF-1/PI3K/CREB Signaling Pathway. *Front. Cell Dev. Biol.* **8**, 595969 (2020).

73. P. L. Lin, C. B. Ford, M. T. Coleman, A. J. Myers, R. Gawande, T. Ioerger, J. Sacchettini, S. M. Fortune, J. L. Flynn, Sterilization of granulomas is common in active and latent tuberculosis despite within-host variability in bacterial killing. *Nat. Med.* **20**, 75–79 (2014).

74. P. Qin, H. Chen, Y. Wang, L. Huang, K. Huang, G. Xiao, C. Han, J. Hu, D. Lin, X. Wan, Y. Zheng, Y. Liu, G. Li, H. Yang, S. Ye, M. Luo, Y. Fu, H. Xu, L. Wen, Z. Guo, X. Shen, Z. Li, C. Wang, X. Chen, L. Wang, L. Sun, D. Ren, L. Wu, J. Wang, S. Liu, H. Lin, Cancer-associated fibroblasts undergoing

neoadjuvant chemotherapy suppress rectal cancer revealed by single-cell and spatial transcriptomics. *Cell Rep. Med.* **4** (2023), doi:10.1016/J.XCRM.2023.101231.

75. M. Guilliams, J. Bonnardel, B. Haest, B. Vanderborght, C. Wagner, A. Remmerie, A. Bujko, L. Martens, T. Thoné, R. Browaeys, F. F. De Ponti, B. Vanneste, C. Zwicker, F. R. Svedberg, T. Vanhalewyn, A. Gonçalves, S. Lippens, B. Devriendt, E. Cox, G. Ferrero, V. Wittamer, A. Willaert, S. J. F. Kaptein, J. Neyts, K. Dallmeier, P. Geldhof, S. Casaert, B. Deplancke, P. Ten Dijke, A. Hoorens, A. Vanlander, F. Berrevoet, Y. Van Nieuwenhove, Y. Saeys, W. Saelens, H. Van Vlierberghe, L. Devisscher, C. L. Scott, Spatial proteogenomics reveals distinct and evolutionarily conserved hepatic macrophage niches. *Cell* **185**, 379-396.e38 (2022).

76. J. Geng, Y. Liu, H. Dai, C. Wang, Fatty Acid Metabolism and Idiopathic Pulmonary Fibrosis. *Front. Physiol.* **12** (2022), doi:10.3389/fphys.2021.794629.

77. S. Morabito, F. Reese, N. Rahimzadeh, E. Miyoshi, V. Swarup, hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. *Cell Rep. Methods* **3**, 100498 (2023).

78. J. D. Bromley, S. K. C. Ganchua, S. K. Nyquist, P. Maiello, M. Chao, H. J. Borish, M. Rodgers, J. Tomko, K. Kracinovsky, D. Mugahid, S. Nguyen, Q. D. Wang, J. M. Rosenberg, E. C. Klein, H. P. Gideon, R. Floyd-O'Sullivan, B. Berger, C. A. Scanga, P. L. Lin, S. M. Fortune, A. K. Shalek, J. L. Flynn, CD4+ T cells re-wire granuloma cellularity and regulatory networks to promote immunomodulation following Mtb reinfection. *Immunity* **57**, 2380-2398.e6 (2024).

79. S. K. Ganchua, P. Maiello, M. Chao, F. Hopkins, D. Mugahid, P. L. Lin, S. M. Fortune, J. L. Flynn, Antibiotic treatment modestly reduces protection against Mycobacterium tuberculosis reinfection in macaques. *Infect. Immun.* **92**, e0053523 (2024).

80. G. Pollara, C. T. Turner, J. Rosenheim, A. Chandran, L. C. K. Bell, A. Khan, A. Patel, L. F. Peralta, A. Folino, A. Akarca, C. Venturini, T. Baker, S. Ecker, F. L. M. Ricciardolo, T. Marafioti, C. Ugarte-Gil, D. A. J. Moore, B. M. Chain, G. S. Tomlinson, M. Noursadeghi, Exaggerated IL-17A activity in human in vivo recall responses discriminates active tuberculosis from latent infection and cured disease. *Sci. Transl. Med.* **13**, eabg7673 (2021).

81. R. Browaeys, W. Saelens, Y. Saeys, NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat. Methods* **17**, 159–162 (2020).

82. D. Dimitrov, D. Türei, M. Garrido-Rodriguez, P. L. Burmedi, J. S. Nagai, C. Boys, R. O. Ramirez Flores, H. Kim, B. Szalai, I. G. Costa, A. Valdeolivas, A. Dugourd, J. Saez-Rodriguez, Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. *Nat. Commun.* **13** (2022), doi:10.1038/s41467-022-30755-0.

83. A. G. Loxton, Bcells and their regulatory functions during Tuberculosis: Latency and active disease. *Mol. Immunol.* **111**, 145–151 (2019).

84. P. Mittal, R. Romero, J. P. Kusanovic, S. S. Edwin, F. Gotsch, S. Mazaki-tovi, J. Espinoza, O. Erez, C. L. Nhan-Chang, N. G. Than, E. Vaisbuch, S. S. Hassan, CXCL6 (Granulocyte Chemotactic Protein-2): a novel chemokine involved in the innate immune response of the amniotic cavity. *Am. J. Reprod. Immunol. N. Y. N 1989* **60**, 246–246 (2008).

85. H. Bahudhanapati, J. Tan, R. M. Apel, B. Seeliger, X. Li, T.-Y. Chen, D. Sullivan, J. Sembrat, M. Rojas, T. Tabib, E. Valenzi, R. Lafyatis, C. Jawale, P. Biswas, J. Tedrow, T. Adams, N. Kaminski, W. A. Wuyts, J. F. McDyer, J. K. Alder, Y. Zhang, M. Nouraie, A. Prasse, D. J. Kass, High CXCL6 drives matrix expression and correlate with markers of poor outcome in IPF. *bioRxiv* , 2021.06.22.449424-2021.06.22.449424 (2021).

86. A. M. Overmiller, J. A. Pierluissi, P. J. Wermuth, S. Sauma, U. Martinez-Outschoorn, M. Tuluc, A. Luginbuhl, J. Curry, L. A. Harshyne, J. K. Wahl, A. P. South, M. G. Mahoney, Desmoglein 2 modulates extracellular vesicle release from squamous cell carcinoma keratinocytes. *FASEB J.* **31**, 3412–3424 (2017).

87. M. R. Cronan, R. W. Beerman, A. F. Rosenberg, J. W. Saelens, M. G. Johnson, S. H. Oehlers, D. M. Sisk, K. L. Jurcic Smith, N. A. Medvitz, S. E. Miller, L. A. Trinh, S. E. Fraser, J. F. Madden, J. Turner, J. E. Stout, S. Lee, D. M. Tobin, Macrophage Epithelial Reprogramming Underlies Mycobacterial Granuloma Formation and Promotes Infection. *Immunity* **45**, 861–876 (2016).

88. X. Shi, J. Wang, X. Zhang, S. Yang, W. Luo, S. Wang, J. Huang, M. Chen, Y. Cheng, J. Chao, GREM1/PPP2R3A expression in heterogeneous fibroblasts initiates pulmonary fibrosis. *Cell Biosci.* **12**, 123–123 (2022).

89. M. Castets, M. M. Coissieux, C. Delloye-Bourgeois, L. Bernard, J. G. Delcros, A. Bernet, V. Laudet, P. Mehlen, Inhibition of Endothelial Cell Apoptosis by Netrin-1 during Angiogenesis. *Dev. Cell* **16** (2009), doi:10.1016/j.devcel.2009.02.006.

90. H. Polena, F. Boudou, S. Tilleul, N. Dubois-Colas, C. Lecointe, N. Rakotosamimanana, M. Pelizzola, S. F. Andriamandimby, V. Raharimanga, P. Charles, J. L. Herrmann, P. Ricciardi-Castagnoli, V. Rasolofo, B. Gicquel, L. Tailleux, Mycobacterium tuberculosis exploits the formation of new blood vessels for its dissemination. *Sci. Rep.* **6** (2016), doi:10.1038/srep33162.

91. G. Nouailles, A. Dorhoi, M. Koch, J. Zerrahn, J. 3rd Weiner, K. C. Faé, F. Arrey, S. Kuhlmann, S. Bandermann, D. Loewe, H.-J. Mollenkopf, A. Vogelzang, C. Meyer-Schwesinger, H.-W. Mittrücker, G. McEwen, S. H. E. Kaufmann, CXCL5-secreting pulmonary epithelial cells drive destructive neutrophilic inflammation in tuberculosis. *J. Clin. Invest.* **124**, 1268–1282 (2014).

92. K. V. Sawant, K. M. Poluri, A. K. Dutta, K. M. Sepuru, A. Troshkina, R. P. Garofalo, K. Rajarathnam, Chemokine CXCL1 mediated neutrophil recruitment: Role of glycosaminoglycan interactions. *Sci. Rep.* **6**, 33123 (2016).

93. N. V. Serbina, T. Jia, T. M. Hohl, E. G. Pamer, Monocyte-Mediated Defense Against Microbial Pathogens. *Annu. Rev. Immunol.* **26**, 421–452 (2008).

94. Q. Yang, F. Qi, T. Ye, J. Li, G. Xu, X. He, G. Deng, P. Zhang, M. Liao, K. Qiao, Z. Zhang, The interaction of macrophages and CD8 T cells in bronchoalveolar lavage fluid is associated with latent tuberculosis infection. *Emerg. Microbes Infect.* **12**, 2239940 (2023).

95. W. Gao, D. Liu, H. Sun, Z. Shao, P. Shi, T. Li, S. Yin, T. Zhu, SPP1 is a prognostic related biomarker and correlated with tumor-infiltrating immune cells in ovarian cancer. *BMC Cancer* **22**, 1367 (2022).

96. E. Matsubara, Y. Komohara, S. Esumi, Y. Shinchi, S. Ishizuka, R. Mito, C. Pan, H. Yano, D. Kobayashi, Y. Fujiwara, K. Ikeda, T. Sakagami, M. Suzuki, SPP1 Derived from Macrophages Is Associated with a Worse Clinical Course and Chemo-Resistance in Lung Adenocarcinoma. *Cancers* **14**, 4374 (2022).

97. K. Hoeft, G. J. L. Schaefer, H. Kim, D. Schumacher, T. Bleckwehl, Q. Long, B. M. Klinkhammer, F. Peisker, L. Koch, J. Nagai, M. Halder, S. Ziegler, E. Liehn, C. Kuppe, J. Kranz, S. Menzel, I. Costa, A. Wahida, P. Boor, R. K. Schneider, S. Hayat, R. Kramann, Platelet-instructed SPP1+ macrophages drive myofibroblast activation in fibrosis in a CXCL4-dependent manner. *Cell Rep.* **42** (2023), doi:10.1016/j.celrep.2023.112131.

98. J. Liu, Y. Huang, Y. Gong, Q. Liu, J. Lin, J. Liu, M. Liu, J. Huang, W. Pu, Y. Ma, Y. Zhang, H. Li, X. Shi, Y. Zhang, J. Wang, Y. Zhu, Q. Wang, K. Wei, J. Wang, Y. Sha, J. Wang, W. Wu, CTHRC1+ fibroblasts are stimulated by macrophage-secreted SPP1 to induce excessive collagen deposition in keloids. *Clin. Transl. Med.* **12**, e1115 (2022).

99. TGF-β: the master regulator of fibrosis | Nature Reviews Nephrology (available at https://www.nature.com/articles/nrneph.2016.48#Abs3).

100. J. Joseph, W. Daley, D. Lawrence, E. Lorenzo, P. Perrin, V. R. Rao, S.-Y. Tsai, V. Varthakavi, Role of macrophages in HIV pathogenesis and cure: NIH perspectives. *J. Leukoc. Biol.* **112**, 1233–1243 (2022).

101. A. J. Theron, R. Anderson, T. M. Rossouw, H. C. Steel, The Role of Transforming Growth Factor Beta-1 in the Progression of HIV/AIDS and Development of Non-AIDS-Defining Fibrotic Disorders. *Front. Immunol.* **8**, 1461 (2017).

102. K. Dheda, C. E. 3rd Barry, G. Maartens, Tuberculosis. *Lancet Lond. Engl.* **387**, 1211–1226 (2016).

103. P. Elkington, M. E. Polak, M. T. Reichmann, A. Leslie, Understanding the tuberculosis granuloma: the matrix revolutions. *Trends Mol. Med.* **28**, 143–154 (2022).

104. P. Ogongo, A. J. Steyn, F. Karim, K. J. Dullabh, I. Awala, R. Madansein, A. Leslie, S. M. Behar, Differential skewing of donor-unrestricted and γδ T cell repertoires in tuberculosis-infected human lungs. *J. Clin. Invest.* **130**, 214–230 (2020).

105. D. Shim, H. Kim, S. J. Shin, Mycobacterium tuberculosis Infection-Driven Foamy Macrophages and Their Implications in Tuberculosis Control as Targets for Host-Directed Therapy. *Front. Immunol.* **11**, 910 (2020).

106. C. Morse, T. Tabib, J. Sembrat, K. L. Buschur, H. T. Bittar, E. Valenzi, Y. Jiang, D. J. Kass, K. Gibson, W. Chen, A. Mora, P. V. Benos, M. Rojas, R. Lafyatis, Proliferating SPP1/MERTK-expressing macrophages in idiopathic pulmonary fibrosis. *Eur. Respir. J.* **54** (2019), doi:10.1183/13993003.02441-2018.

107. J. Qi, H. Sun, Y. Zhang, Z. Wang, Z. Xun, Z. Li, X. Ding, R. Bao, L. Hong, W. Jia, F. Fang, H. Liu, L. Chen, J. Zhong, D. Zou, L. Liu, L. Han, F. Ginhoux, Y. Liu, Y. Ye, B. Su, Single-cell and spatial analysis reveal interaction of FAP(+) fibroblasts and SPP1(+) macrophages in colorectal cancer. *Nat. Commun.* **13**, 1742 (2022).

108. CTHRC1+ fibroblasts and SPP1+ macrophages synergistically contribute to pro-tumorigenic tumor microenvironment in pancreatic ductal adenocarcinoma | Scientific Reports (available at https://www.nature.com/articles/s41598-024-68109-z).

109. S. C. van Kampen, A. Wanner, M. Edwards, A. D. Harries, B. J. Kirenga, J. Chakaya, R. Jones, International research and guidelines on post-tuberculosis chronic lung disorders: a systematic scoping review. *BMJ Glob. Health* **3**, e000745 (2018).

110. B. C. Willis, R. M. duBois, Z. Borok, Epithelial origin of myofibroblasts during fibrosis in the lung. *Proc. Am. Thorac. Soc.* **3**, 377–382 (2006).

111. S. Piera-Velazquez, Z. Li, S. A. Jimenez, Role of endothelial-mesenchymal transition (EndoMT) in the pathogenesis of fibrotic disorders. *Am. J. Pathol.* **179**, 1074–1080 (2011).

112. L. Mori, A. Bellini, M. A. Stacey, M. Schmidt, S. Mattoli, Fibrocytes contribute to the myofibroblast population in wounded skin and originate from the bone marrow. *Exp. Cell Res.* **304**, 81–90 (2005).

113. S. Nayar, J. Campos, C. G. Smith, V. Iannizzotto, D. H. Gardner, F. Mourcin, D. Roulois, J. Turner, M. Sylvestre, S. Asam, B. Glaysher, S. J. Bowman, D. T. Fearon, A. Filer, K. Tarte, S. A. Luther, B. A. Fisher, C. D. Buckley, M. C. Coles, F. Barone, Immunofibroblasts are pivotal drivers of tertiary lymphoid structure formation and local pathology. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 13490–13497 (2019).

114. C. Gago da Graça, L. G. M. van Baarsen, R. E. Mebius, Tertiary Lymphoid Structures: Diversity in Their Development, Composition, and Role. *J. Immunol. Baltim. Md 1950* **206**, 273–281 (2021).

115. A. J. Sawyer, E. Patrick, J. Edwards, J. S. Wilmott, T. Fielder, Q. Yang, D. L. Barber, J. D. Ernst, W. J. Britton, U. Palendira, X. Chen, C. G. Feng, Spatial mapping reveals granuloma diversity and histopathological superstructure in human tuberculosis. *J. Exp. Med.* **220** (2023), doi:10.1084/jem.20221392.

116. M. A. Cole, T. Quan, J. J. Voorhees, G. J. Fisher, Extracellular matrix regulation of fibroblast function: redefining our perspective on skin aging. *J. Cell Commun. Signal.* **12**, 35–43 (2018).

117. R. Hunter, J. Actor, The pathogenesis of post-primary tuberculosis. A game changer for vaccine development. *Tuberc. Edinb. Scotl.* **116S**, S114–S117 (2019).

118. S. Davidson, M. Coles, T. Thomas, G. Kollias, B. Ludewig, S. Turley, M. Brenner, C. D. Buckley, Fibroblasts as immune regulators in infection, inflammation and cancer. *Nat. Rev. Immunol.* **21**, 704–717 (2021).

119. *Doxycycline host-directed therapy in human pulmonary tuberculosis.* (United States, 2021).

120. A. G. White, P. Maiello, M. T. Coleman, J. A. Tomko, L. J. Frye, C. A. Scanga, P. L. Lin, J. L. Flynn, Analysis of 18FDG PET/CT Imaging as a Tool for Studying Mycobacterium tuberculosis Infection and Treatment in Non-human Primates. *J. Vis. Exp. JoVE* , 56375 (2017).

121. E. Z. Macosko, A. Basu, R. Satija, J. Nemesh, K. Shekhar, M. Goldman, I. Tirosh, A. R. Bialas, N. Kamitaki, E. M. Martersteck, J. J. Trombetta, D. A. Weitz, J. R. Sanes, A. K. Shalek, A. Regev, S. A. McCarroll, Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).

Acknowledgments

We extend our appreciation to the technical and veterinary teams of the Flynn, Shalek, and Fortune laboratories for their dedication and assistance in generating the NHP dataset.

Funding

Wellcome Trust grant 210662/Z/18/Z (A.L.)

Bill and Melinda Gates Foundation grant OPP1137006 (A.L., A.K.S)

Medical Research Council (MRC) MR/W025728/1: MR CF2021-2\126 (P.E)

NIH contract IMPAc-TB; 75N93019C00071 (S.M.F., J.L.F., A.K.S.).

Author Contributions

I.M.M, M.W.H, P.E, A.K.S and A.L designed the study. I.M.M collected lung samples and performed single cell sequencing experiments. M.H.W, J.D.B and S.N sequenced the single-cell libraries. N.L. and I.M.M performed the computational analysis with supervision from with supervision from S.N, A.K.S and A.L. S.T. provided insights on single-cell analysis. O.E.A performed immunohistochemistry fluorescent staining. K.N and D.S were responsible for haematoxylin and eosin staining and slide scanning. F.K coordinated the lung study. M.J.C and R.K performed flow cytometry. M.T.R analyzed the LCMD lymph node data. L.T performed the cellular stimulation experiments. S.K.G., P.M., E.C.K., J.L.F, S.M.F. generated and provided insights on the NHP infection data. I.M.M, N.L., P.E, A.K.S and A.L wrote the manuscript with input from all the authors.

Declaration of Interest

A.K.S reports compensation for consulting and/or SAB membership from Honeycomb Biotechnologies, Cellarity, Ochre Bio, Relation Therapeutics, JnJ, IntrECate Biotherapeutics, Passkey Therapeutics, Quotient Therapeutics, Parabilis Biomedicines, Bio-Rad Laboratories, and Dahlia Biosciences. M.H.W is an employee of Pfizer. T.K.H is an employee at Nference. T.K.H , M.H.W and A.K.S are coinventors on a provisional patent filled by MIT relating to the second strand synthesis and Seq-Well technology. S.M.F. reports compensation for board of directors' membership from Oxford Nanopore unrelated to this work. S.T. was supported in part by a Pew Latin American Fellowship Program in Biomedical Sciences. The authors have no additional financial interests.

Data and materials availability

The raw and analyzed scRNA-seqs data from this study will be deposited on GEO and Broad Institute Single Cell Portal. The 4-week post-infection NHP were previously used in another study by Ganchua *et al* and Bromley *et al.(78, 79)*. The related scRNA-seq dataset will be publicly available for download and visualization via the Alexandria and Broad Institute Single Cell Portal, and fairdomhub upon publication. The other 4-week p.i and 10-week p.i NHP granuloma dataset from Gideon *et al.* can be accessed from GEO with accession number GSE200151. and The Human Lung Cell Atlas can be accessed at https://data.humancellatlas.org/hca-bio-networks/lung/atlases/lung-v1-0. The human TB lymph node bulk RNA-seq data from Reichmann *et al.* can be accessed on GEO with accession number GSE174443. Any additional information required to reanalyze the data from this study is available from the lead contact upon request.

Figures

Figure 1: Overview of the single-cell and spatial data generated from TB-diseased and control lungs. (**A**) Schematic showing the experimental flow for the isolation of cells from human lung tissues, generation of single-cell libraries using Seq-Well S^3. 4 TB-negative and 9 TB-positive lung samples were processed through single-cell RNA-seq. Shown adjacent to the process flow is a low-dimensional embedding (UMAP) of the 19,632 cells passing quality control annotated with high-level cell types (middle) or detailed cell subtype (right). (**B**) 10x Visium platform workflow for spatial transcriptomics profiling on formalin-fixed paraffin-embedded tissue (FFPE) samples from TB-diseased lung resections. 21 of these samples come from current-TB patients with detectable *M.tb*, 9 came from post-TB patient with where bacteria are no longer detected in bronchoalveolar lavage (BAL) TB culture after infection. Samples contain either granulomas, inducible bronchus-associated lymphoid tissues (iBALTs), or lung lymph nodes (LNs), representing different pathological states.

Figure 2: Overview of tissue heterogeneity and cell type abundance in the single-cell dataset.

(**A**) Cell type proportions by disease status (left) and patient (right, n = 7 HIV^+TB^+; n = 2 TB^+; n =1 HIV^+, n = 3 Cancer Control). (B) Low dimensional embedding (UMAP) of all single-cell RNA-seq data colored by patient HIV status (left) and TB status (right). (**C**) Dot plot showing expression levels of top 2 DE genes in each of the broad level cell types. (**D**) Two-sided Fisher's exact test for abundance of major cell types between samples from patients with previous TB diagnosis and samples from control patients. Holm's method was applied to adjust p-values for multiple-testing correction. Statistical annotations: p-value < 0.05 (*), p-value < 0.01 (**), p-value < 0.001 (***).

Figure 3: Single-cell transcriptomic reveals heterogeneity within monocyte and macrophage populations with disease specific difference. (**A**) Monocyte/macrophage (n=8,318) sub-clustering reveals 10 subclusters (left), also colored by patient ID (middle) and disease condition (right). (**B**) Heatmap of subtype top 10 DE genes in each of the monocyte/macrophage subcluster. (**C**) Expression of marker genes in monocyte/macrophage subclusters by disease conditions. (**D**) Two-sided Fisher's exact test on abundance of detailed macrophage (left), monocyte (right) subclusters between TB conditions. Holm's method was applied to adjust p-values for multiple-testing correction. Statistical annotations: p-value < 0.05 (*), p-value < 0.01 (**),p-value < 0.001 (***), fold-change> 1 (Δ), fold-change >2 ($\Delta\Delta$), fold-change <1 (∇). **(E)** *Cell2loc* imputed macrophage (left) and monocyte (right) abundance distribution on the Visium dataset grouped by TB and granuloma status (**Methods**). The 5% quantile of the estimated posterior distribution of cell abundance at each Visium spot is displayed, representing the value of cell abundance that the model has high confidence in. Two-sided Mann-Whitney U test without correction were used for statistical testing. Statistical annotations: p-value < 0.0001 (****). **(F)** Similar to **(E)** but group by TB status and HIV status.

Figure 4: Fibroblast exhibit TB specific phenotypes. (**A**) Fibroblast (n=1,627) sub-clustering reveals 5 subclusters (left), also colored by patient ID (middle) and disease condition (right). (**B**) Heatmap of subtype top 10 DE genes in each of the fibroblast subcluster. Columns (cells) are annotated by fibroblast subclusters and sample source disease status. (**C**) Comparing annotation against literature stromal annotation from Travaglini *et al*. *(65)*. Left: original fibroblast UMAP as seen in (**A**) colored by mapped cell types in Travaglini *et al*. Right: barplot showing distributions of mapped cell type in each original subcluster. ASM: airway smooth muscle; VSM: vascular smooth muscle; MyoF: myofibroblast; FibM: fibromyocyte; AdvF: adventitial fibroblast; AlvF: alveolar fibroblast; LipF: lipofibroblast; Peri: pericyte; Meso: mesothelial. (**D**) Reference mapping to the Human Lung Cell Atlas (HLCA). Query (all cells in this study, n=19,632) vs reference cells (n= 584,944) on integrated UMAP with transferred label from HLCA to query cells. (**E**) Query (all fibroblasts in this study that was mapped to fibroblast/myofibroblast in label transfer, n=1,601) and reference lung fibroblast cells (n=17,500) from HLCA colored by annotation (either 'Fibroblast' or 'Myofibroblast') and TB conditions. (**F**) Gene-set enrichment analysis (GSEA) on DE genes between TB fibroblasts and TB-negative fibroblasts on HLCA-integrated data.

Figure 5: Fibroblast weighted gene co-expression network analysis (hdWGCNA). (A) High dimensional weighted gene co-expression network analysis (hdWGCNA) for gene module detection in fibroblasts of this study. UMAPs are colored by eigengene of each of the 7 modules. **(B)** Evaluation of M1 module expression (MEs) in fibroblast subclusters. Bonferroni adjusted P computed from two-sided Wilcoxon test are shown. **(C)** Over-representation analysis (ORA) by enricher on all assigned M1 module genes using MSigDB Gene-ontology Biological Processes (GOBP) gene set database. (**D**) Top: bacterial burden of NHP lung granulomas from Gideon *et al. (16)* grouped by the time point. Bottom: evaluation of human TB-myofibroblast module expression in NHP TB fibroblasts on 4-week and 10-week samples. Two-sided Mann-Whitney U test without correction was used. Statistical annotations: p-value < 0.05 (*), p-value < 0.01 (**), p-value < 0.001 (***). **(E)** Evaluation of human TB-myofibroblast module expression in fibroblasts from granuloma vs uninvolved lungs in an independent non-human primate study with 4-week post-infection (p.i) macaques *(78)* (**Methods**). Two-sided Mann-Whitney U test without correct was used. Statistical annotations: p-value < 0.05 (*), p-value < 0.01 (**), p-value < 0.001 (***).

Figure 6: Evidence of *MMP1*$^{+}$ *CXCL5*$^{+}$ Fibroblast populations in TB-diseased human lungs. (**A**) Representative flow cytometry plot showing the isolation strategy of the PDPN^{+}FAP^{+} fibroblast population from the CD45-EPCAM- cell fraction. (**B**) Cumulative data on frequency of PDPN^{+}FAP^{+}CD90^{+} fibroblasts as a fraction of live lung cells from 5 separate lung resections. Three separate sections were taken from each TB-diseased lung, corresponding to the most diseased and least diseased tissues areas and an intermediate lung piece, according to the expert opinion of the operating surgeon. Friedman test was used to ascertain statistical significance between proportion of PDPN^{+}FAP^{+} fibroblast between severity groups. (**C**) Expression of human TB-myofibroblast signature and *SPP1*$^{+}$*CHI3L1*$^{+}$ marker genes in the tuberculin skin test (TST) challenge site *in vivo* model. Active TB TST (n=48): biopsies from participants with microbiologically confirmed pulmonary TB disease within the first month of treatment who underwent TST; latent TB TST (n=191): biopsies from participants lacking clinical and radiological evidence of active TB disease but with a positive peripheral blood IFN-γ release assay; saline (n=34): biopsies from participants that received saline under the skin instead of tuberculin. Each dot corresponds to a sample, horizontal lines represent median values. Two-sided Mann-Whitney U test without correct was used. Statistical annotations: p-value < 0.0001 (****).

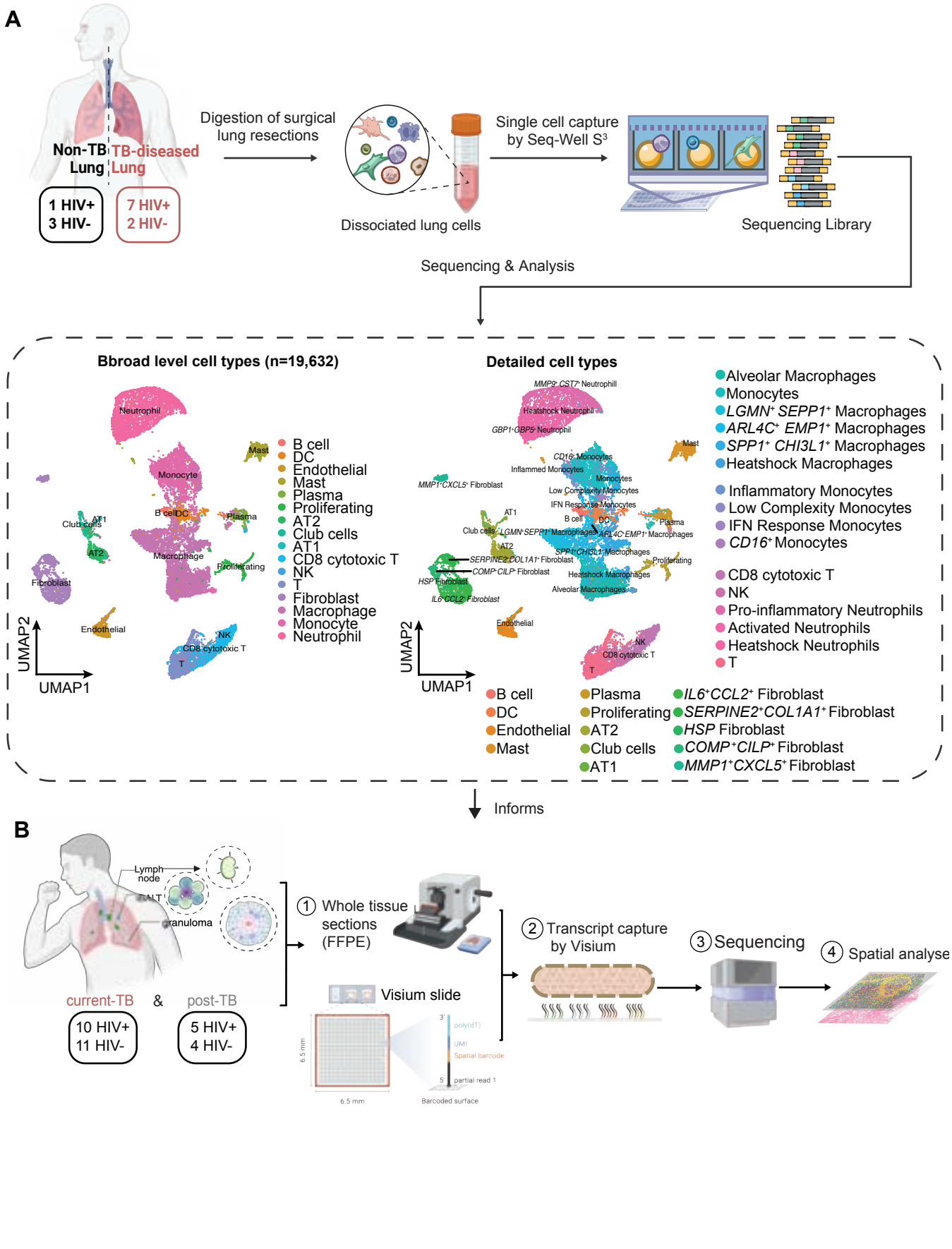
Figure 7: Cell-cell interaction analysis reveals key discrepancies between TB-diseased and control lung niches. (**A**) Top 20 ligand-receptor (L-R) pairs from MultiNichenet analysis highlighting putative interaction pairs with upregulated interactions in TB-negative lung compared to TB-diseased lung. **(B)** Top 20 ligand-receptor (L-R) pairs from MultiNichenet analysis highlighting putative interaction pairs differentially communicating in TB-diseased lungs. (**C**) Summary of top 200 interactions in TB-diseased and TB-negative/control lungs, by the number of interactions between each cell pair. Cartoons on the right of each heatmap shows the suggested major modes interactions in each condition. (**D**) Circos plots of significant interaction pairs in TB-diseased lungs from LIANA where sender and receiver cell types in each condition are clustered to reflect similar patterns.

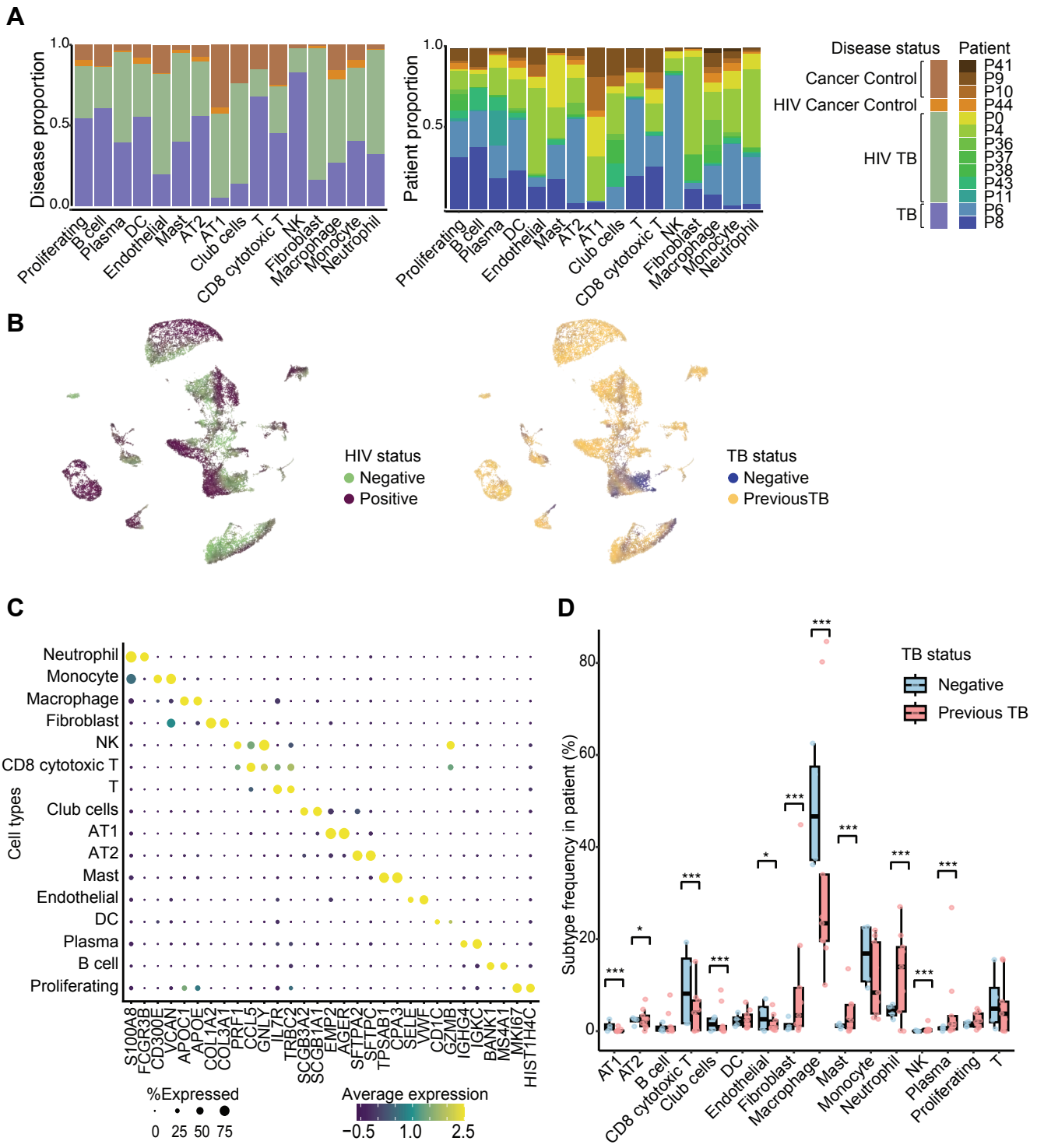
Figure 8: Global interaction analysis identifies key players in cellular communication within TB-diseased lung tissues

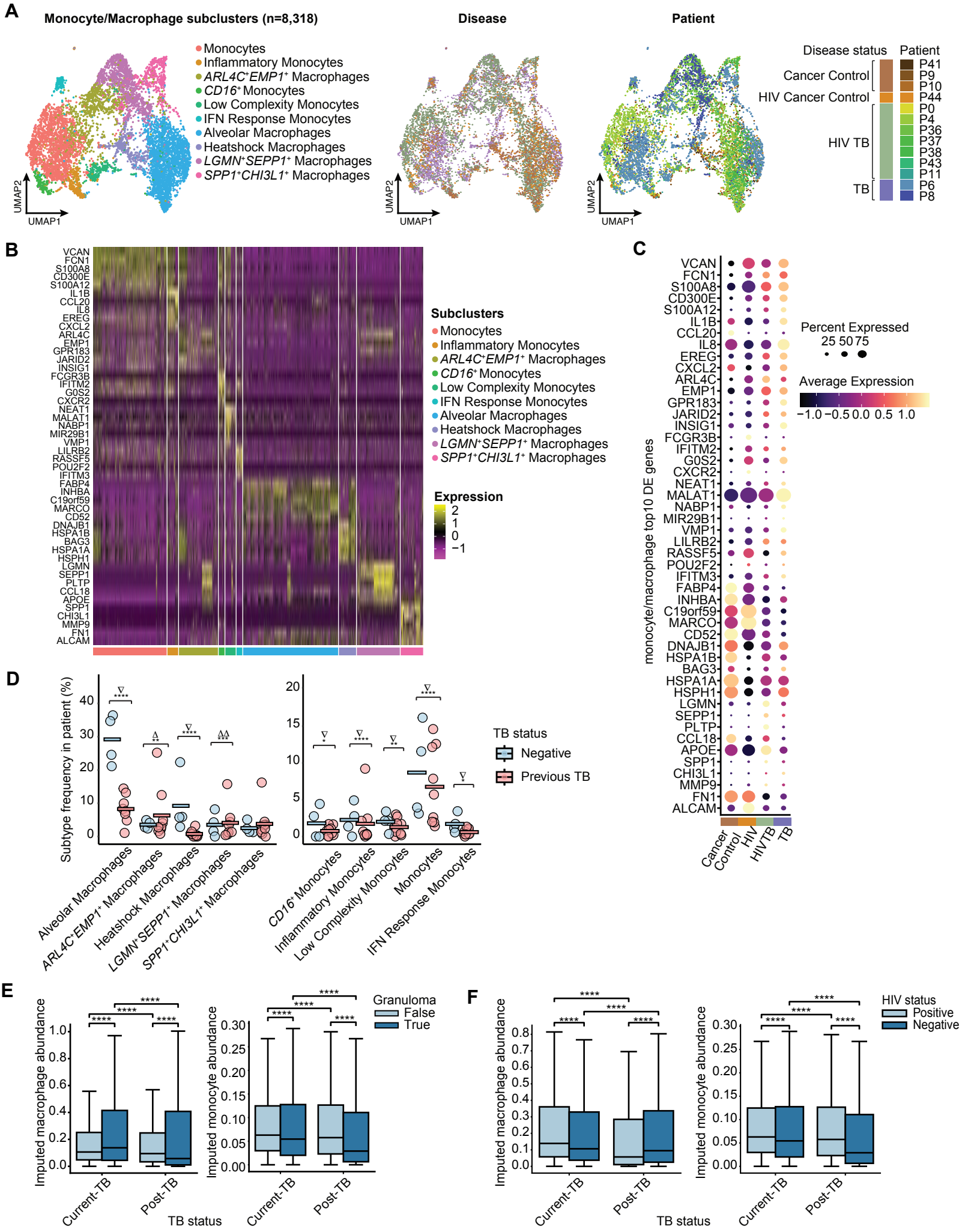
(A) Heatmap visualization of interaction flux analysis. Rows represent sender cell types; columns represent receiver cell types. Each entry represents the potential flux of interaction from sender cell to receiver cell, whereas the total flux of each sender cell type is summarized on the left. Sender cell types are sorted based on descending order of total flux (**Methods**). (**B**) Top: Bar plot showing top 30 and bottom 30 ligands by log fold-change of interaction strength between TB and control lungs across all sender cell types, bottom: log fold-change of interaction strength between TB and control lungs in each sender cell type (**Methods**). (**C**) Dot plot of top 5 ligand by ligand activity in TB-diseased lungs secreted by *MMP1*$^{+}$*CXCL5*$^{+}$ Fibroblasts and their receivers (**Methods**). (**D**) L-R interactions with *MMP1*$^{+}$*CXCL5*$^{+}$ Fibroblasts in the TB-diseased lungs, rows (L-R pair) and columns (target cell types) are hierarchically clustered by correlation distance (**Methods**).

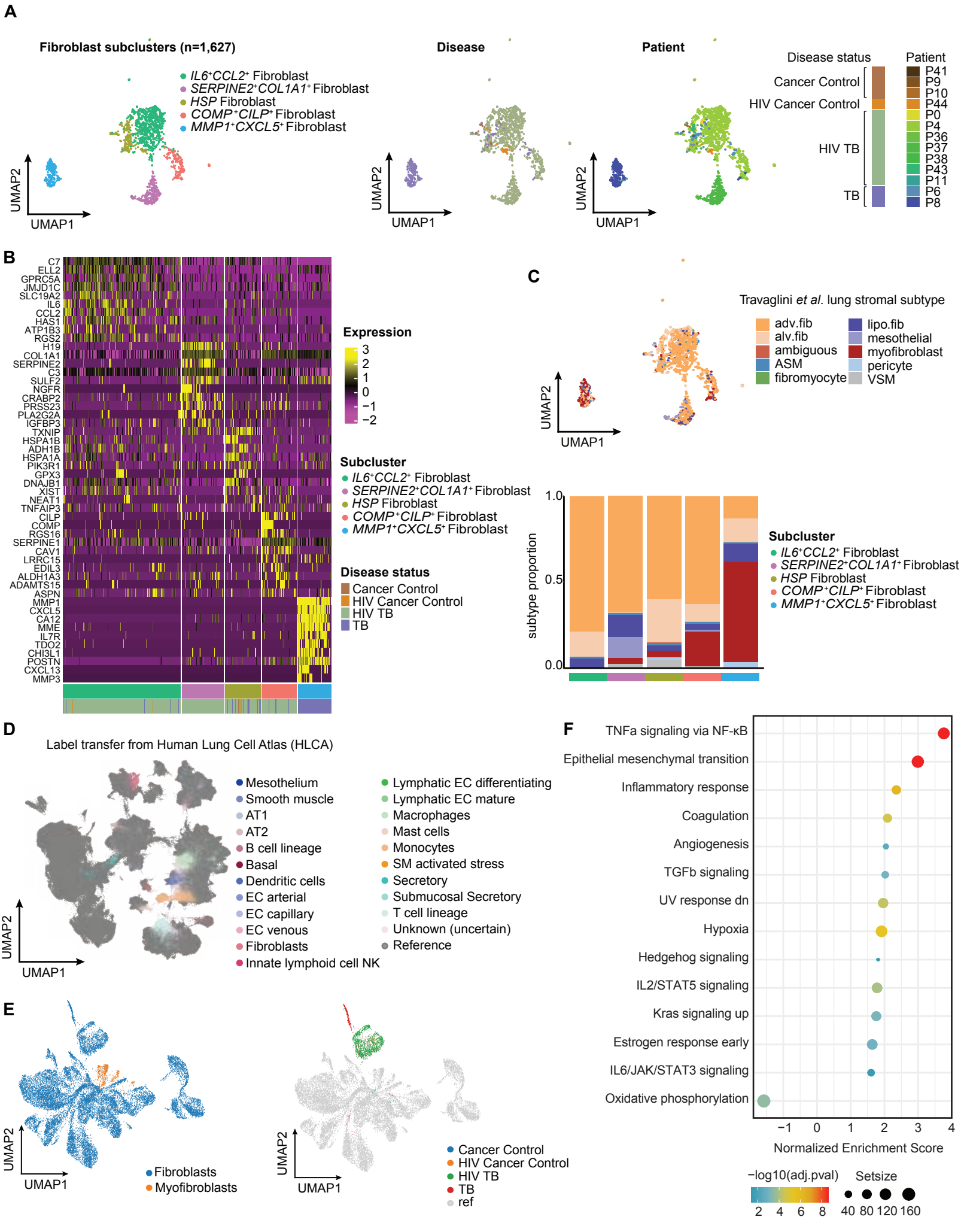
Figure 9: Spatial transcriptomics analysis on post- and current-TB lung resections

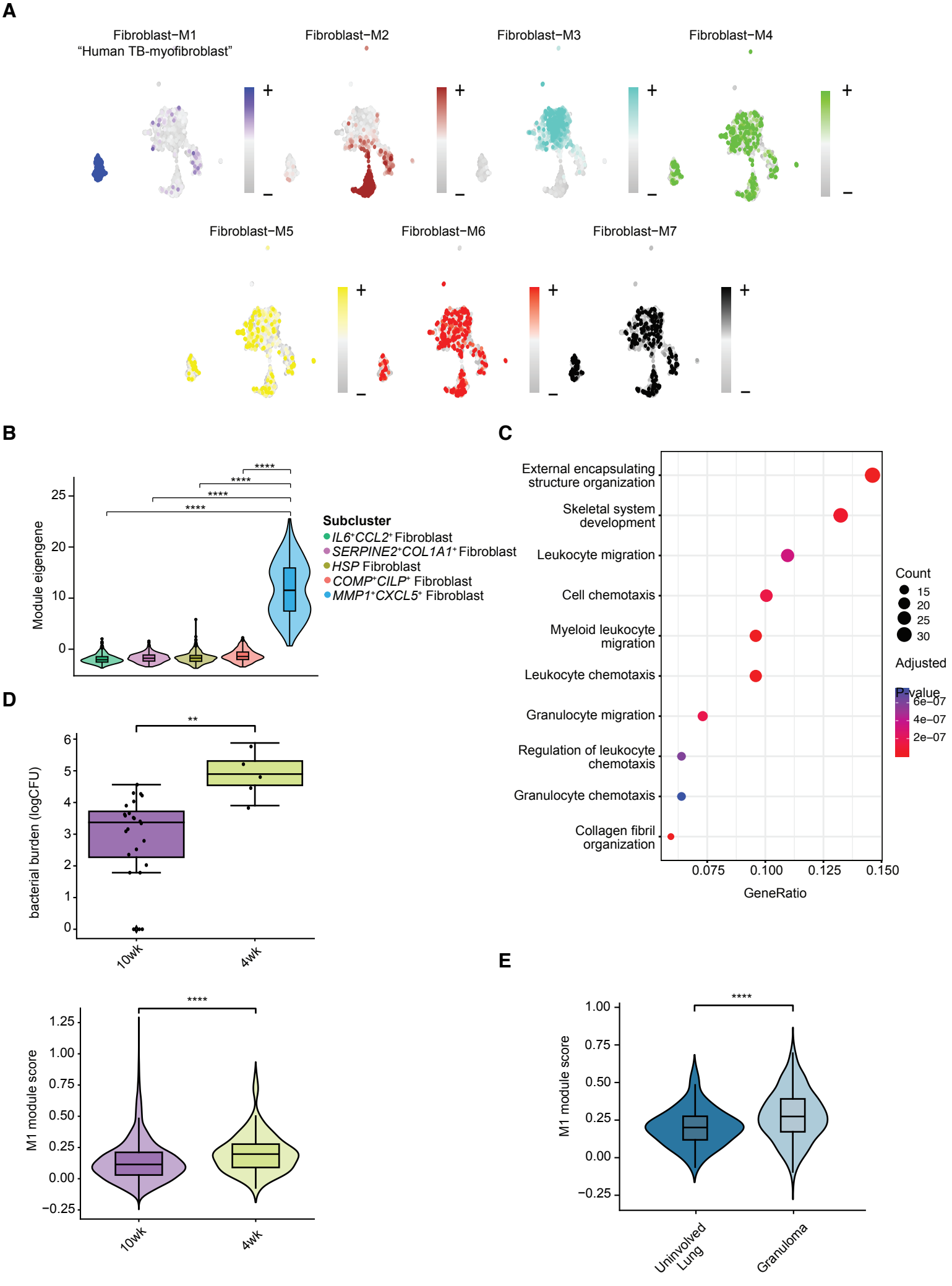
(A) Heatmap showing the expression of human TB-myofibroblast gene signature and *SPP1*$^{+}$*CHI3L1*+ macrophage markers on selective tissue slides from patients who are post-TB (top) or current-TB (bottom), alongside paired H&E staining. **(B)** Distribution of human TB-myofibroblast signature expression on the spatial cohort. HIV statuses are shown in different shades of blue for positive or negative. Two-sided Mann-Whitney U test without correction was used for statistical testing. Statistical annotation: *p*-value<0.0001 (****). **(C)** Distribution of *SPP1*$^{+}$*CHI3L1*$^{+}$ macrophage markers and human TB-myofibroblast signature on the spatial data across all Visium spots. Left two panels: manual segmentation of the granuloma structure was done to allow separation of the Visium slide into three different regions: in granuloma, on granuloma border (cuff), outside of granuloma (**Methods**). Right two panels: the same as left panels with the exception that “on border”=True means on granuloma cuff and False means the rest. Two-sided Mann-Whitney U test without correction was used for statistical testing. Statistical annotation: *p*-value<0.0001 (****). **(D)** Correlation between human TB-myofibroblast signature and all macrophage subpopulations’ markers. Each circle represents a Visium sample. Boxplot of the Pearson’s r distribution is shown for each macrophage subtype. Mann-Whitney U test without correction were used for statistical testing. Statistical annotation: *p*-value<0.0001 (****). **(E)** Spatially informed ligand-receptor (L-R) analysis using *LIANA*+ on Visium samples. Examples are shown where SPP1(L)-CD44(R) interactions are being nominated as top L-R pairs. H&E overlaid with pathology annotation for granuloma structures are shown next to heatmap of L-R interaction scores, which are calculated at each Visium spot using spatially-weighted Cosine similarity (**Methods**)

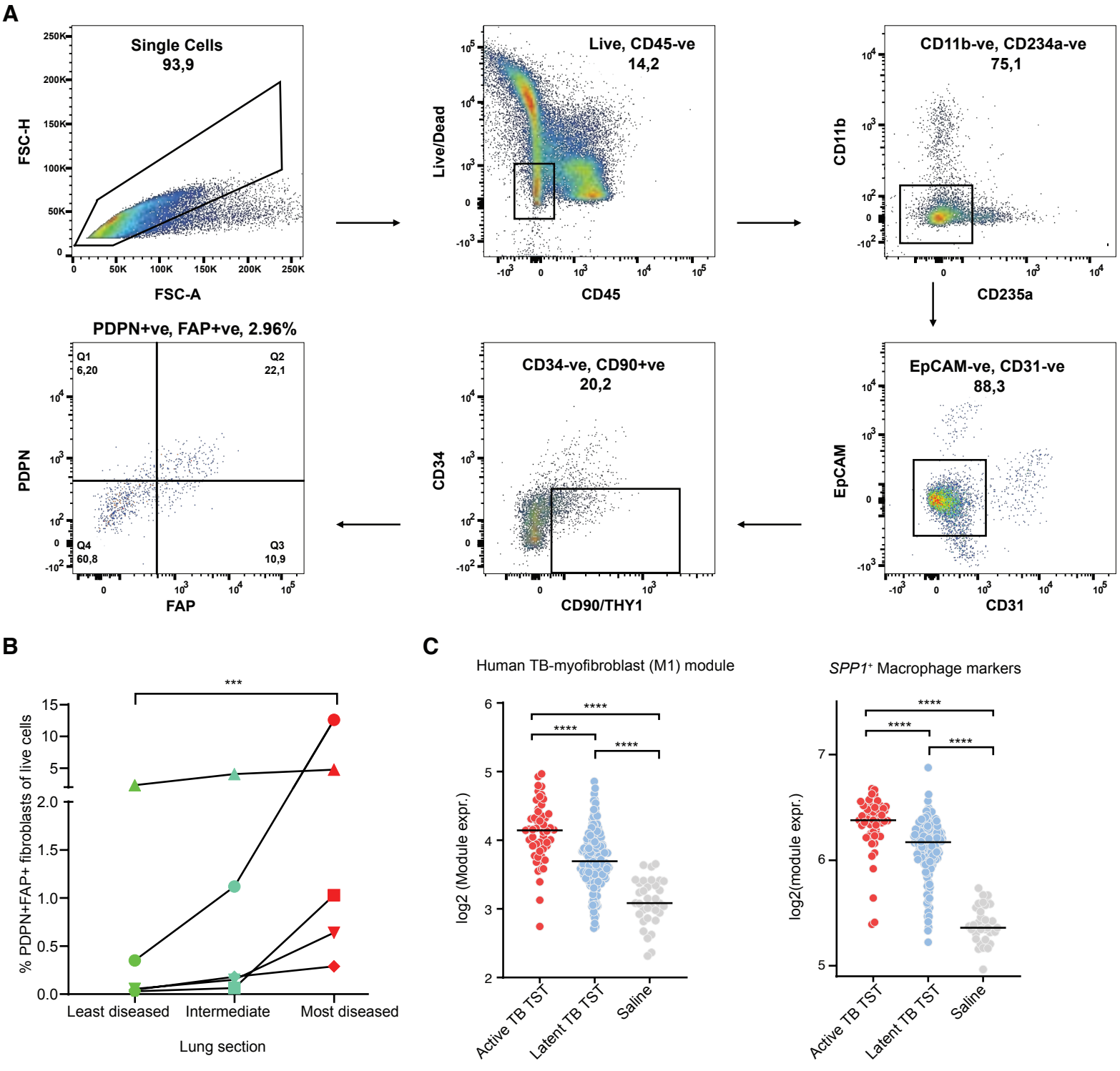

A
Non-TB Lung
TB-diseased Lung
1 HIV+
3 HIV-
7 HIV+
2 HIV-
Digestion of surgical lung resections
Dissociated lung cells
Single cell capture by Seq-Well S3
Sequencing Library
Sequencing & Analysis
Bbroad level cell types (n=19,632)
Neutrophil
Monocyte
B cell
DC
Plasma
Macrophage
Proliferating
AT2
Club cells
AT1
Fibroblast
Endothelial
NK
CD8 cytotoxic T
T
Mast
UMAP2
UMAP1
B cell
DC
Endothelial
Mast
Plasma
Proliferating
AT2
Club cells
AT1
Detailed cell types
MMP9+ CST7+ Neutrophil
Heatshock Neutrophil
GBP1+GBP5+ Neutrophil
CD16+ Monocytes
Inflammed Monocytes
Monocytes
MMP1+CXCL5+ Fibroblast
Low Complexity Monocytes
IFN Response Monocytes
LGMN+SEPP1+ Macrophages
ARL4C+ EMP1+ Macrophages
SPP1+ CHI3L1+ Macrophages
SERPINE2+COL1A1+ Fibroblast
COMP+CILP+ Fibroblast
HSP Fibroblast
IL6+CCL2+ Fibroblast
Heatshock Macrophages
Alveolar Macrophages
Alveolar Macrophages
Monocytes
LGMN+ SEPP1+ Macrophages
ARL4C+ EMP1+ Macrophages
SPP1+ CHI3L1+ Macrophages
Heatshock Macrophages
Inflammatory Monocytes
Low Complexity Monocytes
IFN Response Monocytes
CD16+ Monocytes
CD8 cytotoxic T
NK
Pro-inflammatory Neutrophils
Activated Neutrophils
Heatshock Neutrophils
T
IL6+CCL2+ Fibroblast
SERPINE2+COL1A1+ Fibroblast
HSP Fibroblast
COMP+CILP+ Fibroblast
MMP1+CXCL5+ Fibroblast
Informs
B
Lymph node
Granuloma
current-TB
&
post-TB
10 HIV+
11 HIV-
5 HIV+
4 HIV-
1 Whole tissue sections (FFPE)
Visium slide
poly(dT)
UMI
Spatial barcode
partial read 1
6.5 mm
Barcoded surface
2 Transcript capture by Visium
3 Sequencing
4 Spatial analyse

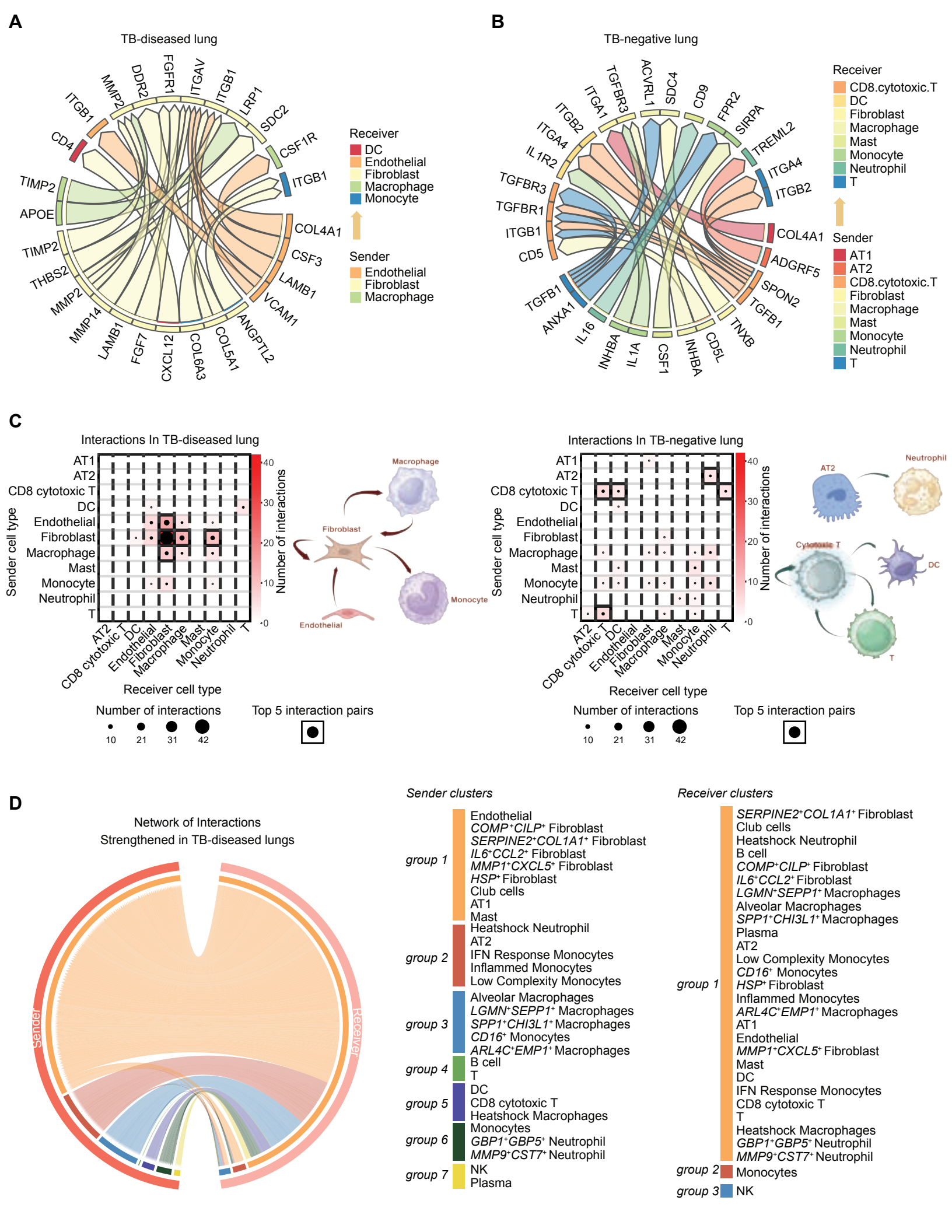

A
Disease proportion
Patient proportion
1.0
0.5
0.0
Proliferating
B cell
Plasma
DC
Endothelial
Mast
AT2
AT1
Club cells
T
CD8 cytotoxic T
NK
Fibroblast
Macrophage
Monocyte
Neutrophil
Disease status
Patient
Cancer Control
HIV Cancer Control
HIV TB
TB
P41
P9
P10
P44
P0
P4
P36
P37
P38
P43
P11
P6
P8
B
HIV status
Negative
Positive
TB status
Negative
PreviousTB
C
Cell types
S100A8
FCGR3B
CD300E
VCAN
APOC1
APOE
COL1A2
COL3A1
PRF1
CCL5
GNLY
IL7R
TRBC2
SCGB3A2
SCGB1A1
EMP2
AGER
SFTPA2
SFTPC
TPSAB1
CPA3
SELE
VWF
CD1C
GZMB
IGHG4
IGKC
BANK1
MS4A1
MKI67
HIST1H4C
%Expressed
0 25 50 75
Average expression
−0.5 1.0 2.5
D
Subtype frequency in patient (%)
TB status
Negative
Previous TB
***
*

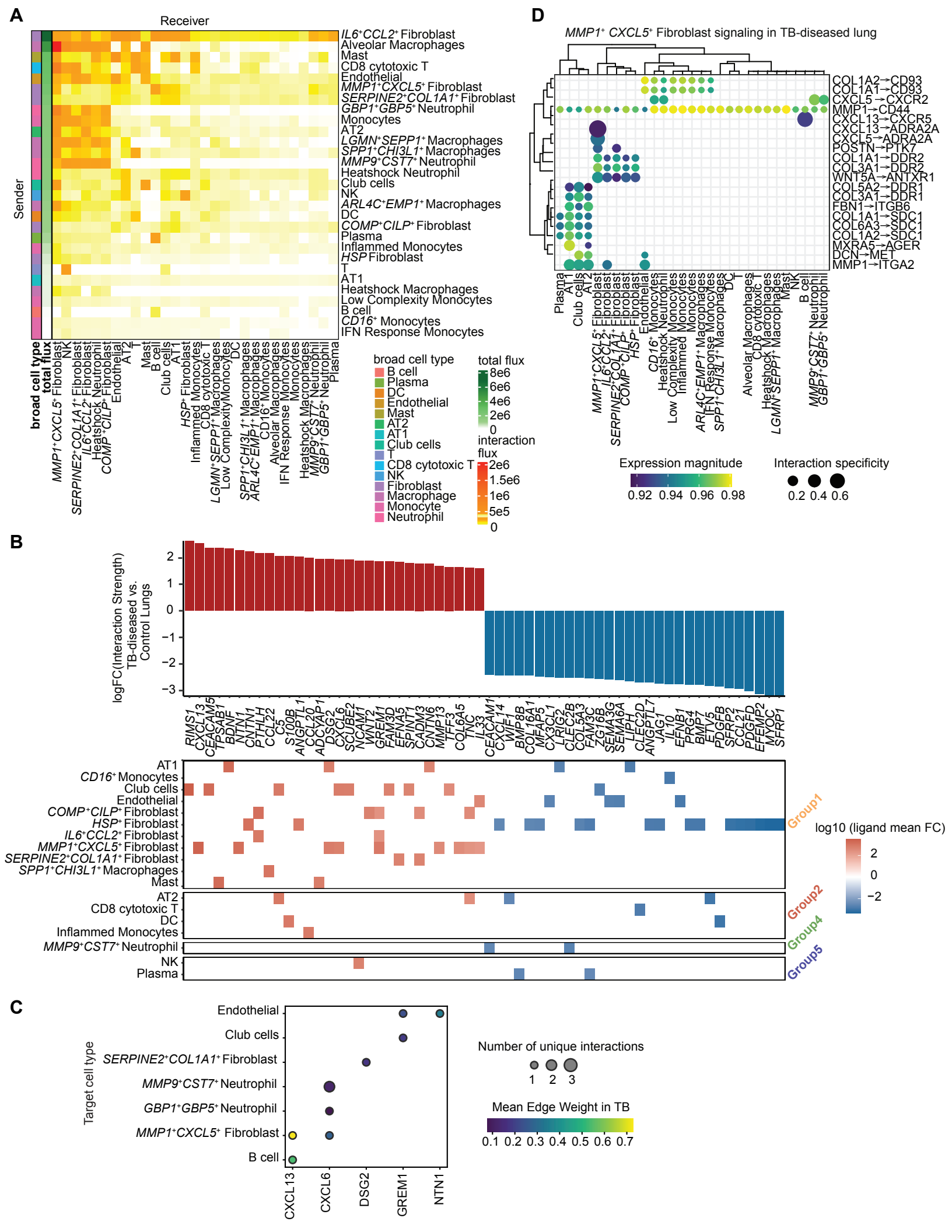

A
Monocyte/Macrophage subclusters (n=8,318)
Monocytes
Inflammatory Monocytes
ARL4C+EMP1+ Macrophages
CD16+ Monocytes
Low Complexity Monocytes
IFN Response Monocytes
Alveolar Macrophages
Heatshock Macrophages
LGMN+SEPP1+ Macrophages
SPP1+CHI3L1+ Macrophages
UMAP1
UMAP2
Disease
Patient
Disease status
Patient
Cancer Control
HIV Cancer Control
HIV TB
TB
P41
P9
P10
P44
P0
P4
P36
P37
P38
P43
P11
P6
P8
B
VCAN
FCN1
S100A8
CD300E
S100A12
IL1B
CCL20
IL8
EREG
CXCL2
ARL4C
EMP1
GPR183
JARID2
INSIG1
FCGR3B
IFITM2
G0S2
CXCR2
NEAT1
MALAT1
NABP1
MIR29B1
VMP1
LILRB2
RASSF5
POU2F2
IFITM3
FABP4
INHBA
C19orf59
MARCO
CD52
DNAJB1
HSPA1B
BAG3
HSPA1A
HSPH1
LGMN
SEPP1
PLTP
CCL18
APOE
SPP1
CHI3L1
MMP9
FN1
ALCAM
Subclusters
Monocytes
Inflammatory Monocytes
ARL4C+EMP1+ Macrophages
CD16+ Monocytes
Low Complexity Monocytes
IFN Response Monocytes
Alveolar Macrophages
Heatshock Macrophages
LGMN+SEPP1+ Macrophages
SPP1+CHI3L1+ Macrophages
Expression
2
1
0
−1
C
monocyte/macrophage top10 DE genes
Percent Expressed
25 50 75
Average Expression
−1.0 −0.5 0.0 0.5 1.0
Cancer Control
HIV
HIVTB
TB
D
Subtype frequency in patient (%)
TB status
Negative
Previous TB
Alveolar Macrophages
ARL4C+EMP1+ Macrophages
Heatshock Macrophages
LGMN+SEPP1+ Macrophages
SPP1+CHI3L1+ Macrophages
CD16+ Monocytes
Inflammatory Monocytes
Low Complexity Monocytes
Monocytes
IFN Response Monocytes
E
Imputed macrophage abundance
Imputed monocyte abundance
Granuloma
False
True
Current-TB
Post-TB
TB status
F
Imputed macrophage abundance
Imputed monocyte abundance
HIV status
Positive
Negative
Current-TB
Post-TB
TB status

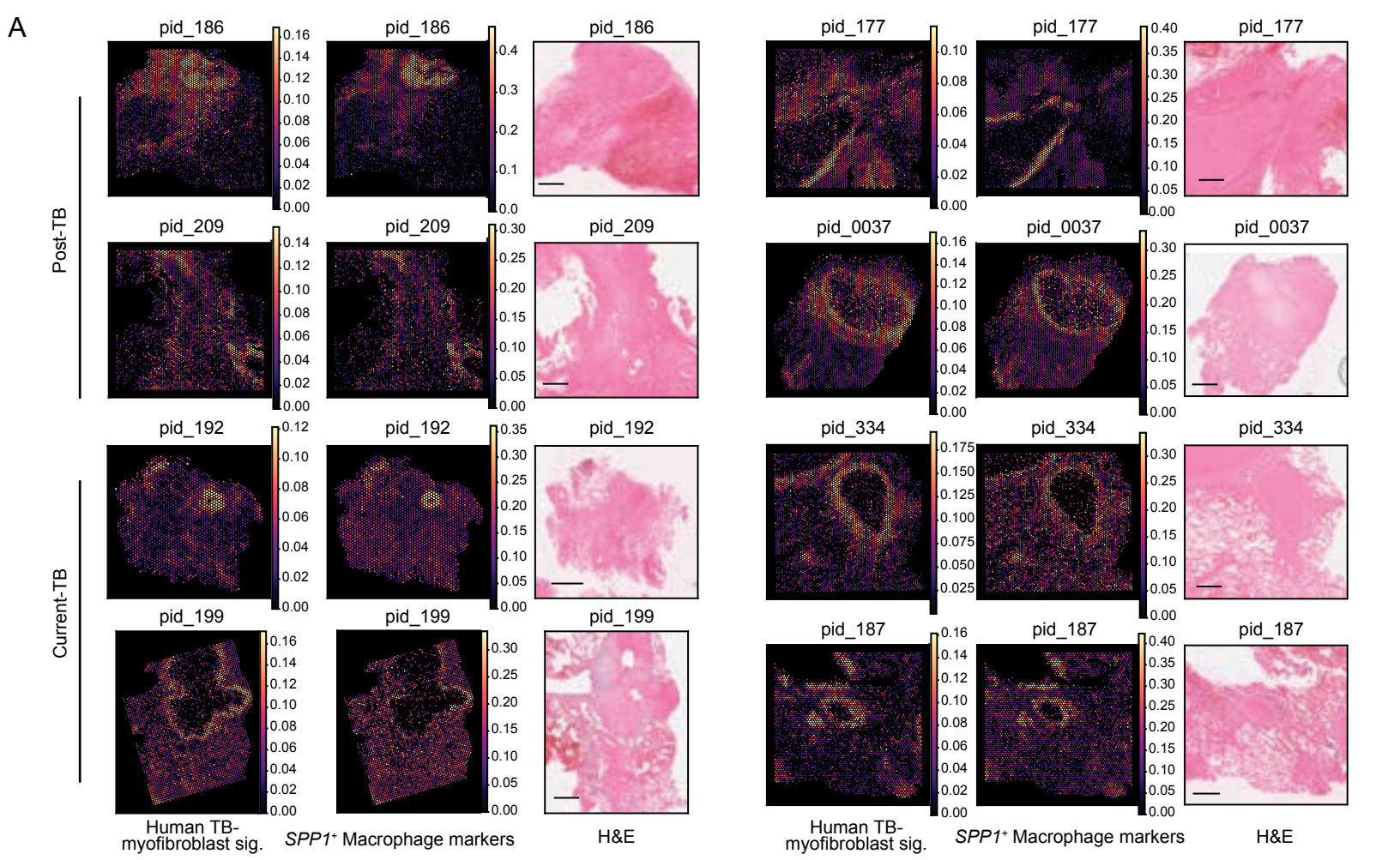

A
Fibroblast subclusters (n=1,627)
IL6+CCL2+ Fibroblast
SERPINE2+COL1A1+ Fibroblast
HSP Fibroblast
COMP+CILP+ Fibroblast
MMP1+CXCL5+ Fibroblast
Disease
Patient
Disease status
Cancer Control
HIV Cancer Control
HIV TB
TB
Patient
P41
P9
P10
P44
P0
P4
P36
P37
P38
P43
P11
P6
P8
UMAP1
UMAP2
B
C7
ELL2
GPRC5A
JMJD1C
SLC19A2
IL6
CCL2
HAS1
ATP1B3
RGS2
H19
COL1A1
SERPINE2
C3
SULF2
NGFR
CRABP2
PRSS23
PLA2G2A
IGFBP3
TXNIP
HSPA1B
ADH1B
HSPA1A
PIK3R1
GPX3
DNAJB1
XIST
NEAT1
TNFAIP3
CILP
COMP
RGS16
SERPINE1
CAV1
LRRC15
EDIL3
ALDH1A3
ADAMTS15
ASPN
MMP1
CXCL5
CA12
MME
IL7R
TDO2
CHI3L1
POSTN
CXCL13
MMP3
Expression
3
2
1
0
−1
−2
Subcluster
IL6+CCL2+ Fibroblast
SERPINE2+COL1A1+ Fibroblast
HSP Fibroblast
COMP+CILP+ Fibroblast
MMP1+CXCL5+ Fibroblast
Disease status
Cancer Control
HIV Cancer Control
HIV TB
TB
C
Travaglini et al. lung stromal subtype
adv.fib
alv.fib
ambiguous
ASM
fibromyocyte
lipo.fib
mesothelial
myofibroblast
pericyte
VSM
UMAP1
UMAP2
subtype proportion
1.0
0.5
0.0
Subcluster
IL6+CCL2+ Fibroblast
SERPINE2+COL1A1+ Fibroblast
HSP Fibroblast
COMP+CILP+ Fibroblast
MMP1+CXCL5+ Fibroblast
D
Label transfer from Human Lung Cell Atlas (HLCA)
Mesothelium
Smooth muscle
AT1
AT2
B cell lineage
Basal
Dendritic cells
EC arterial
EC capillary
EC venous
Fibroblasts
Innate lymphoid cell NK
Lymphatic EC differentiating
Lymphatic EC mature
Macrophages
Mast cells
Monocytes
SM activated stress
Secretory
Submucosal Secretory
T cell lineage
Unknown (uncertain)
Reference
UMAP1
UMAP2
E
Fibroblasts
Myofibroblasts
Cancer Control
HIV Cancer Control
HIV TB
TB
ref
F
TNFa signaling via NF-κB
Epithelial mesenchymal transition
Inflammatory response
Coagulation
Angiogenesis
TGFb signaling
UV response dn
Hypoxia
Hedgehog signaling
IL2/STAT5 signaling
Kras signaling up
Estrogen response early
IL6/JAK/STAT3 signaling
Oxidative phosphorylation
−1
0
1
2
3
4
Normalized Enrichment Score
−log10(adj.pval)
2
4
6
8
Setsize
40
80
120
160

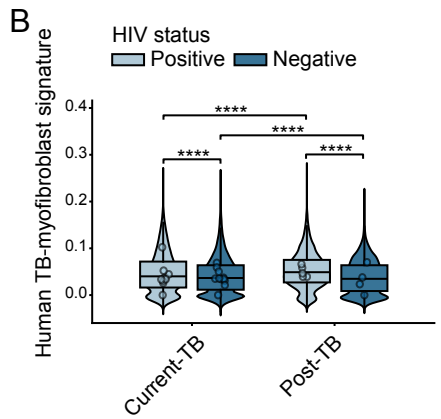

A
Fibroblast−M1
“Human TB-myofibroblast”
Fibroblast−M2
Fibroblast−M3
Fibroblast−M4
Fibroblast−M5
Fibroblast−M6
Fibroblast−M7
B
Module eigengene
Subcluster
IL6+CCL2+ Fibroblast
SERPINE2+COL1A1+ Fibroblast
HSP Fibroblast
COMP+CILP+ Fibroblast
MMP1+CXCL5+ Fibroblast
C
External encapsulating structure organization
Skeletal system development
Leukocyte migration
Cell chemotaxis
Myeloid leukocyte migration
Leukocyte chemotaxis
Granulocyte migration
Regulation of leukocyte chemotaxis
Granulocyte chemotaxis
Collagen fibril organization
GeneRatio
Count
Adjusted P-value
D
bacterial burden (logCFU)
10wk
4wk
M1 module score
E
Uninvolved Lung
Granuloma

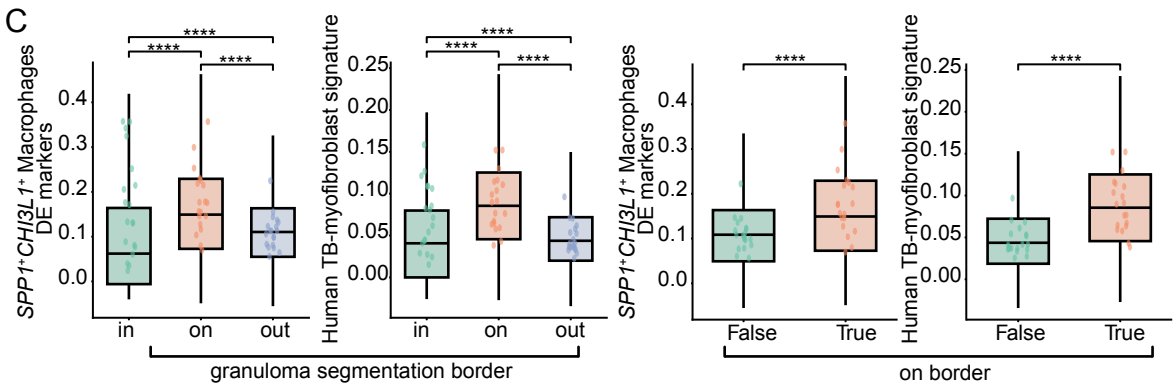

A
Single Cells
93,9
FSC-H
FSC-A
Live, CD45-ve
14,2
Live/Dead
CD45
CD11b-ve, CD234a-ve
75,1
CD11b
CD235a
EpCAM-ve, CD31-ve
88,3
EpCAM
CD31
CD34-ve, CD90+ve
20,2
CD34
CD90/THY1
PDPN+ve, FAP+ve, 2.96%
Q1
6,20
Q2
22,1
Q4
60,8
Q3
10,9
PDPN
FAP
B
% PDPN+FAP+ fibroblasts of live cells
***
Least diseased
Intermediate
Most diseased
Lung section
C
Human TB-myofibroblast (M1) module
log2 (Module expr.)
****
Active TB TST
Latent TB TST
Saline
SPP1+ Macrophage markers
log2(module expr.)

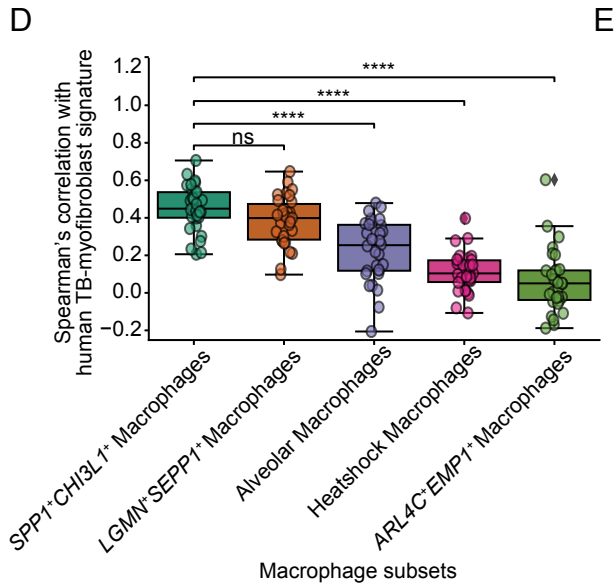

A
TB-diseased lung
ITGB1
MMP2
DDR2
FGFR1
ITGAV
ITGB1
LRP1
SDC2
CSF1R
ITGB1
CD4
TIMP2
APOE
TIMP2
THBS2
MMP2
MMP14
LAMB1
FGF7
CXCL12
COL6A3
COL5A1
ANGPTL2
VCAM1
LAMB1
CSF3
COL4A1
Receiver
DC
Endothelial
Fibroblast
Macrophage
Monocyte
Sender
Endothelial
Fibroblast
Macrophage
B
TB-negative lung
TGFBR3
ACVRL1
SDC4
CD9
FPR2
SIRPA
TREML2
ITGA4
ITGB2
ITGA1
ITGB2
ITGA4
IL1R2
TGFBR3
TGFBR1
ITGB1
CD5
TGFB1
ANXA1
IL16
INHBA
IL1A
CSF1
INHBA
CD5L
TNXB
TGFB1
SPON2
ADGRF5
COL4A1
Receiver
CD8.cytotoxic.T
DC
Fibroblast
Macrophage
Mast
Monocyte
Neutrophil
T
Sender
AT1
AT2
CD8.cytotoxic.T
Fibroblast
Macrophage
Mast
Monocyte
Neutrophil
T
C
Interactions In TB-diseased lung
Sender cell type
AT1
AT2
CD8 cytotoxic T
DC
Endothelial
Fibroblast
Macrophage
Mast
Monocyte
Neutrophil
T
Number of interactions
40
30
20
10
0
Receiver cell type
Macrophage
Fibroblast
Monocyte
Endothelial
Number of interactions
10
21
31
42
Top 5 interaction pairs
Interactions In TB-negative lung
Neutrophil
AT2
Cytotoxic T
DC
T
D
Network of Interactions
Strengthened in TB-diseased lungs
Sender
Receiver
Sender clusters
group 1
Endothelial
COMP+CILP+ Fibroblast
SERPINE2+COL1A1+ Fibroblast
IL6+CCL2+ Fibroblast
MMP1+CXCL5+ Fibroblast
HSP+ Fibroblast
Club cells
AT1
Mast
group 2
Heatshock Neutrophil
AT2
IFN Response Monocytes
Inflammed Monocytes
Low Complexity Monocytes
group 3
Alveolar Macrophages
LGMN+SEPP1+ Macrophages
SPP1+CHI3L1+ Macrophages
CD16+ Monocytes
ARL4C+EMP1+ Macrophages
group 4
B cell
T
group 5
DC
CD8 cytotoxic T
Heatshock Macrophages
group 6
Monocytes
GBP1+GBP5+ Neutrophil
MMP9+CST7+ Neutrophil
group 7
NK
Plasma
Receiver clusters
group 1
SERPINE2+COL1A1+ Fibroblast
Club cells
Heatshock Neutrophil
B cell
COMP+CILP+ Fibroblast
IL6+CCL2+ Fibroblast
LGMN+SEPP1+ Macrophages
Alveolar Macrophages
SPP1+CHI3L1+ Macrophages
Plasma
AT2
Low Complexity Monocytes
CD16+ Monocytes
HSP+ Fibroblast
Inflammed Monocytes
ARL4C+EMP1+ Macrophages
AT1
Endothelial
MMP1+CXCL5+ Fibroblast
Mast
DC
IFN Response Monocytes
CD8 cytotoxic T
T
Heatshock Macrophages
GBP1+GBP5+ Neutrophil
MMP9+CST7+ Neutrophil
group 2
Monocytes
group 3
NK

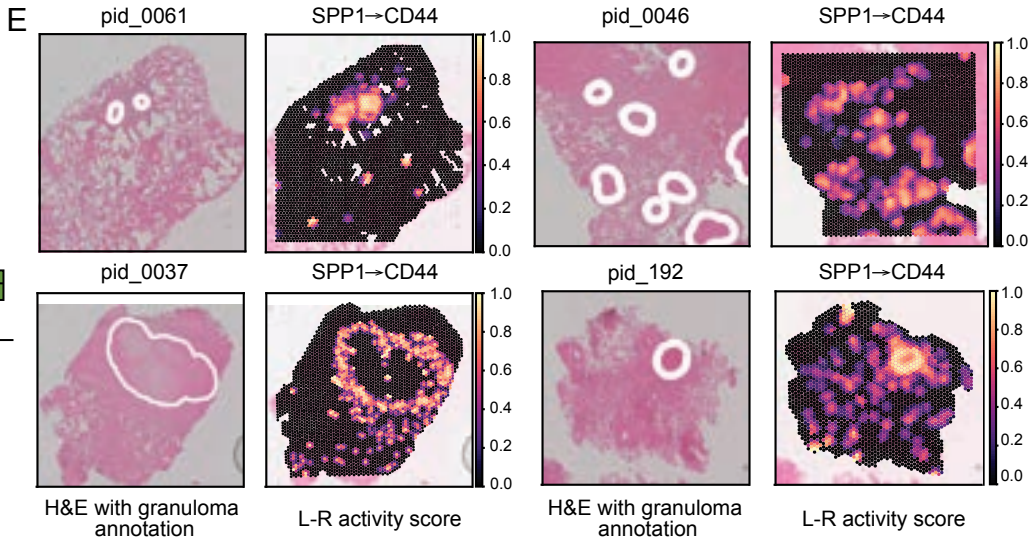

A
Receiver
Sender
IL6+CCL2+ Fibroblast
Alveolar Macrophages
Mast
CD8 cytotoxic T
Endothelial
MMP1+CXCL5+ Fibroblast
SERPINE2+COL1A1+ Fibroblast
GBP1+GBP5+ Neutrophil
Monocytes
AT2
LGMN+SEPP1+ Macrophages
SPP1+CHI3L1+ Macrophages
MMP9+CST7+ Neutrophil
Heatshock Neutrophil
Club cells
NK
ARL4C+EMP1+ Macrophages
DC
COMP+CILP+ Fibroblast
Plasma
Inflammed Monocytes
HSP Fibroblast
T
AT1
Heatshock Macrophages
Low Complexity Monocytes
B cell
CD16+ Monocytes
IFN Response Monocytes
broad cell type
total flux
broad cell type
B cell
Plasma
DC
Endothelial
Mast
AT2
AT1
Club cells
T
CD8 cytotoxic T
NK
Fibroblast
Macrophage
Monocyte
Neutrophil
total flux
8e6
6e6
4e6
2e6
0
interaction flux
2e6
1.5e6
1e6
5e5
0
D
MMP1+ CXCL5+ Fibroblast signaling in TB-diseased lung
COL1A2→CD93
COL1A1→CD93
CXCL5→CXCR2
MMP1→CD44
CXCL13→CXCR5
CXCL13→ADRA2A
CXCL5→ADRA2A
POSTN→PTK7
COL1A1→DDR2
COL3A1→DDR2
WNT5A→ANTXR1
COL5A2→DDR1
COL3A1→DDR1
FBN1→ITGB6
COL1A1→SDC1
COL6A3→SDC1
COL1A2→SDC1
MXRA5→AGER
DCN→MET
MMP1→ITGA2
Expression magnitude
0.92 0.94 0.96 0.98
Interaction specificity
0.2 0.4 0.6
B
logFC(Interaction Strength)
TB-diseased vs. Control Lungs
2
1
0
−1
−2
−3
RIMS1 CXCL13 CEACAM5 TPSAB1 BDNF NTN1 CNTN1 PTHLH CCL22 C5 S100B ANGPTL1 CCL20 ADCYAP1 DSG2 CXCL6 SCUBE2 NCAM1 WNT2 GREM1 FAM3D EFNA5 SPINT1 CADM3 CNTN6 MMP13 TFF3 COL6A5 TNC IL33 CEACAM1 CXCL14 WIF1 BMP8B COL16A1 MFAP5 CX3CL1 LRIG2 CLEC2B COL5A3 FAM3C ZG16B SEMA3G SEMA6A LIPH CLEC2D ANGPTL7 JAG1 IL10 EFNB1 PRG4 BMP7 ETV5 PDGFB SFRP2 CCL21 PDGFD EFEMP2 MYOC SFRP1
AT1
CD16+ Monocytes
Club cells
Endothelial
COMP+CILP+ Fibroblast
HSP+ Fibroblast
IL6+CCL2+ Fibroblast
MMP1+CXCL5+ Fibroblast
SERPINE2+COL1A1+ Fibroblast
SPP1+CHI3L1+ Macrophages
Mast
AT2
CD8 cytotoxic T
DC
Inflammed Monocytes
MMP9+CST7+ Neutrophil
NK
Plasma
Group1
Group2
Group4
Group5
log10 (ligand mean FC)
2
0
−2
C
Target cell type
Endothelial
Club cells
SERPINE2+COL1A1+ Fibroblast
MMP9+CST7+ Neutrophil
GBP1+GBP5+ Neutrophil
MMP1+CXCL5+ Fibroblast
B cell
CXCL13
CXCL6
DSG2
GREM1
NTN1
Number of unique interactions
1 2 3
Mean Edge Weight in TB
0.1 0.2 0.3 0.4 0.5 0.6 0.7

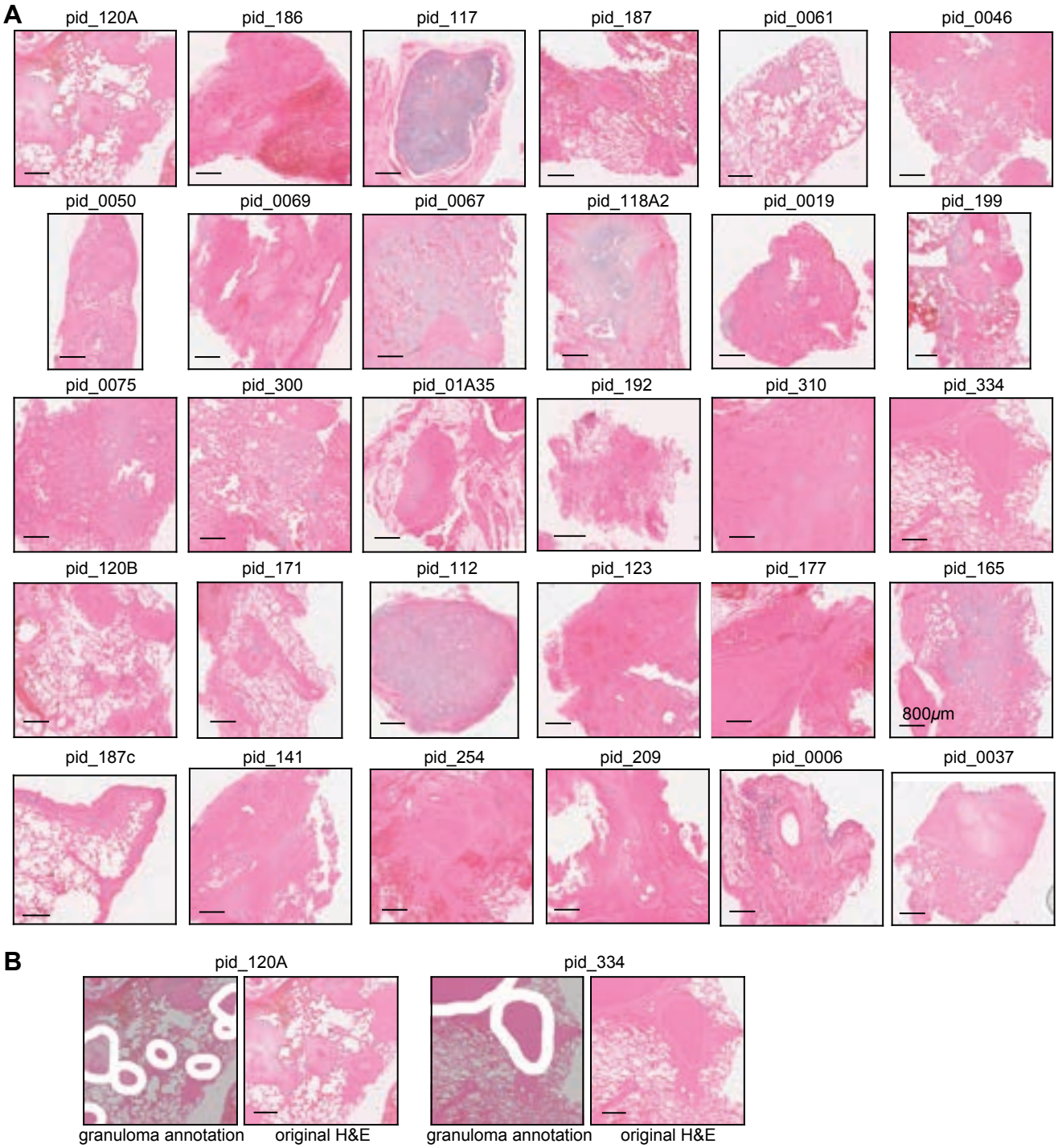

A
Post-TB
Current-TB
pid_186
pid_209
pid_192
pid_199
pid_177
pid_0037
pid_334
pid_187
Human TB-myofibroblast sig.
SPP1+ Macrophage markers
H&E

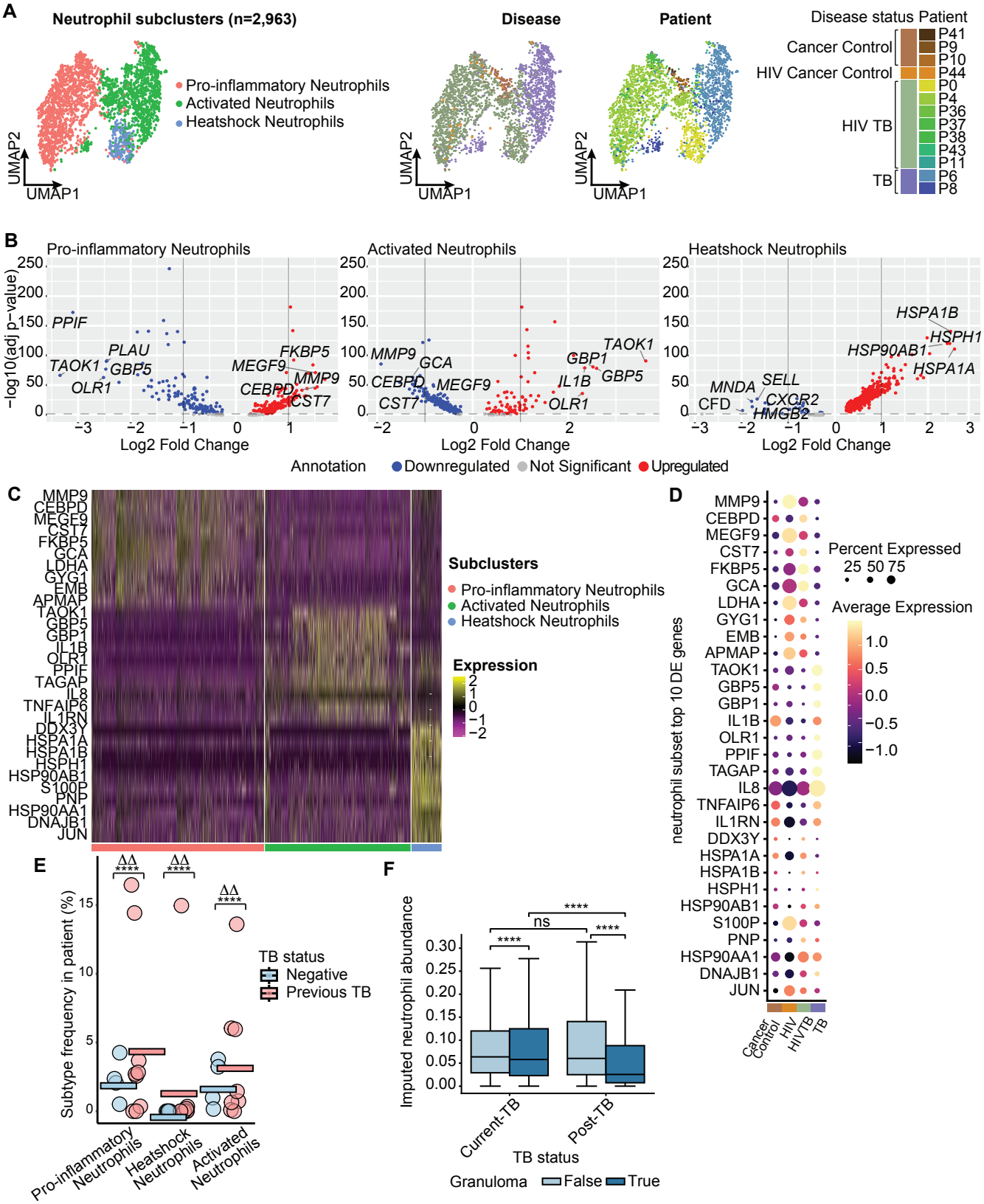

B
HIV status
Positive
Negative
Human TB-myofibroblast signature
Current-TB
Post-TB

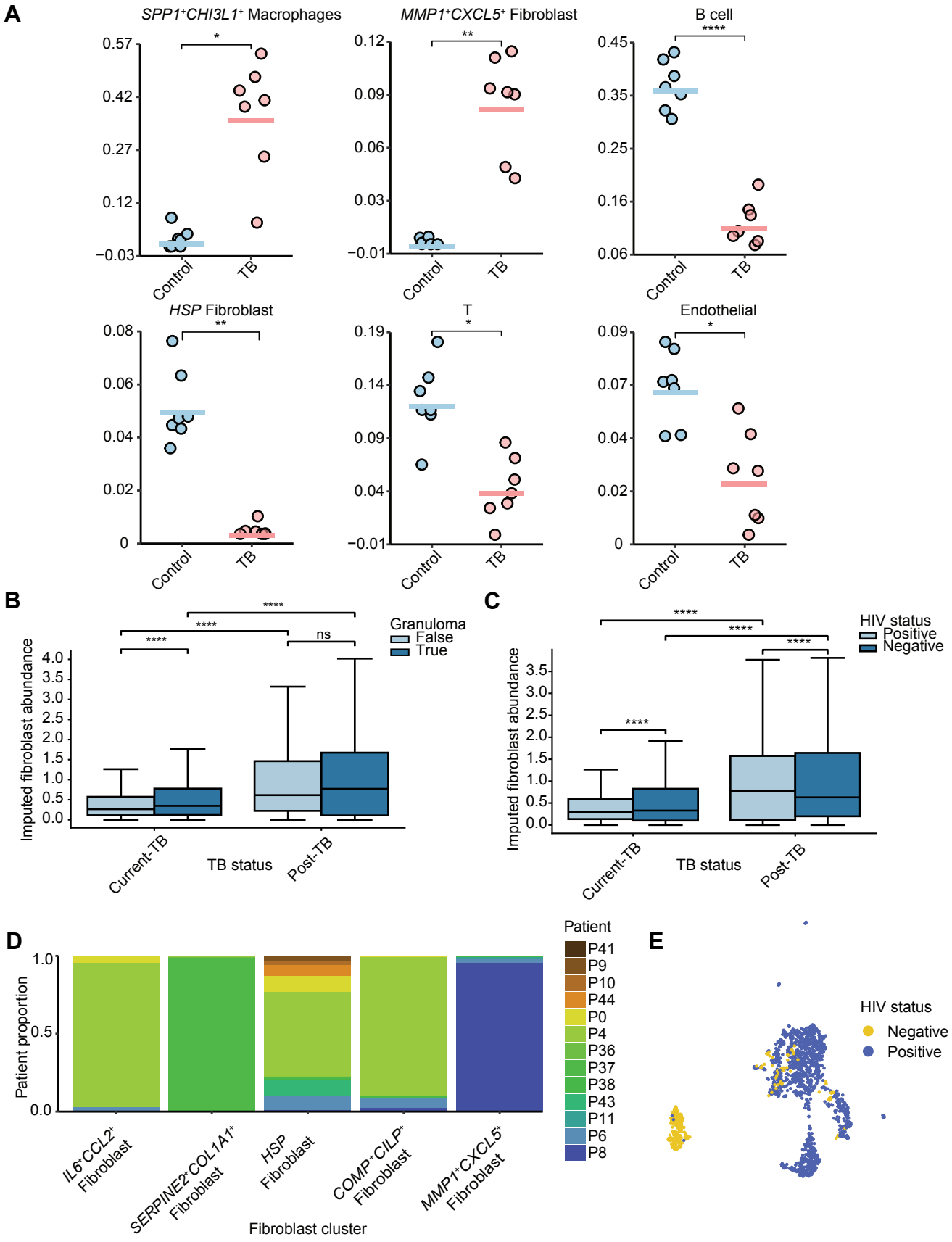

C
SPP1+CHI3L1+ Macrophages DE markers
Human TB-myofibroblast signature
in
on
out
granuloma segmentation border
False
True
on border

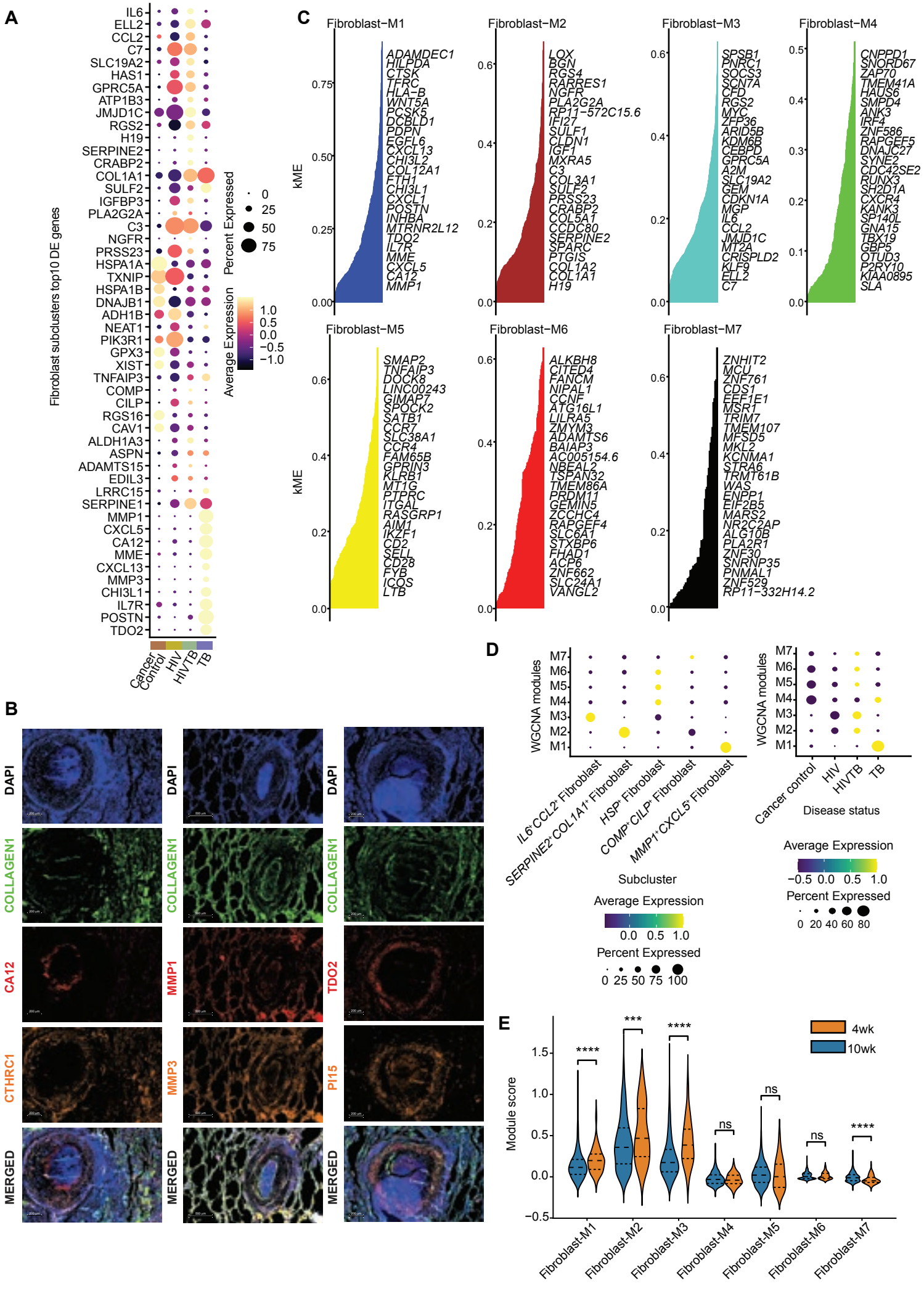

D
Spearman's correlation with human TB-myofibroblast signature
ns
SPP1+CHI3L1+ Macrophages
LGMN+SEPP1+ Macrophages
Alveolar Macrophages
Heatshock Macrophages
ARL4C+EMP1+ Macrophages
Macrophage subsets

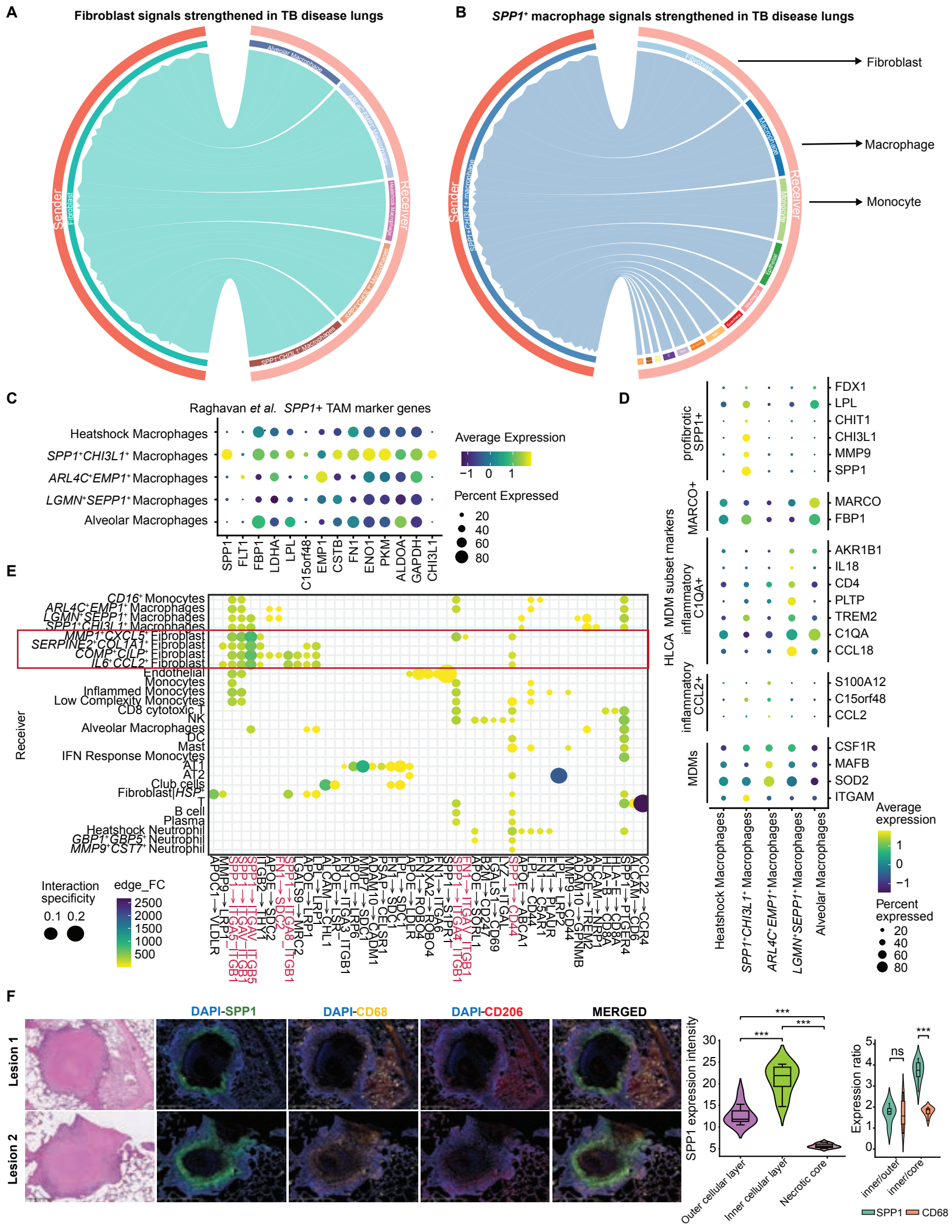

E
pid_0061
pid_0046
pid_0037
pid_192
SPP1→CD44
H&E with granuloma annotation
L-R activity score


A
pid_120A
pid_186
pid_117
pid_187
pid_0061
pid_0046
pid_0050
pid_0069
pid_0067
pid_118A2
pid_0019
pid_199
pid_0075
pid_300
pid_01A35
pid_192
pid_310
pid_334
pid_120B
pid_171
pid_112
pid_123
pid_177
pid_165
800μm
pid_187c
pid_141
pid_254
pid_209
pid_0006
pid_0037
B
pid_120A
pid_334
granuloma annotation
original H&E
granuloma annotation
original H&E


A
Neutrophil subclusters (n=2,963)
Pro-inflammatory Neutrophils
Activated Neutrophils
Heatshock Neutrophils
UMAP2
UMAP1
Disease
Patient
Disease status Patient
Cancer Control
HIV Cancer Control
HIV TB
TB
P41
P9
P10
P44
P0
P4
P36
P37
P38
P43
P11
P6
P8
B
Pro-inflammatory Neutrophils
Activated Neutrophils
Heatshock Neutrophils
−log10(adj p−value)
Log2 Fold Change
PPIF
PLAU
TAOK1
GBP5
OLR1
FKBP5
MEGF9
MMP9
CEBPD
CST7
GCA
TAOK1
GBP1
GBP5
IL1B
OLR1
HSPA1B
HSPH1
HSP90AB1
HSPA1A
MNDA
SELL
CFD
CXCR2
HMGB2
Annotation
Downregulated
Not Significant
Upregulated
C
MMP9
CEBPD
MEGF9
CST7
FKBP5
GCA
LDHA
GYG1
EMB
APMAP
TAOK1
GBP5
GBP1
IL1B
OLR1
PPIF
TAGAP
IL8
TNFAIP6
IL1RN
DDX3Y
HSPA1A
HSPA1B
HSPH1
HSP90AB1
S100P
PNP
HSP90AA1
DNAJB1
JUN
Subclusters
Pro-inflammatory Neutrophils
Activated Neutrophils
Heatshock Neutrophils
Expression
D
neutrophil subset top 10 DE genes
Percent Expressed
25 50 75
Average Expression
Cancer Control
HIV
HIV TB
TB
E
Subtype frequency in patient (%)
TB status
Negative
Previous TB
Pro-inflammatory Neutrophils
Heatshock Neutrophils
Activated Neutrophils
F
Imputed neutrophil abundance
ns
Current-TB
Post-TB
TB status
Granuloma
False
True


A
SPP1+CHI3L1+ Macrophages
MMP1+CXCL5+ Fibroblast
B cell
HSP Fibroblast
T
Endothelial
Control
TB
B
Granuloma
False
True
Imputed fibroblast abundance
Current-TB
Post-TB
TB status
ns
C
HIV status
Positive
Negative
D
Patient
P41
P9
P10
P44
P0
P4
P36
P37
P38
P43
P11
P6
P8
Patient proportion
IL6+CCL2+ Fibroblast
SERPINE2+COL1A1+ Fibroblast
HSP Fibroblast
COMP+CILP+ Fibroblast
MMP1+CXCL5+ Fibroblast
Fibroblast cluster
E
HIV status
Negative
Positive


A
Fibroblast subclusters top10 DE genes
IL6
ELL2
CCL2
C7
SLC19A2
HAS1
GPRC5A
ATP1B3
JMJD1C
RGS2
H19
SERPINE2
CRABP2
COL1A1
SULF2
IGFBP3
PLA2G2A
C3
NGFR
PRSS23
HSPA1A
TXNIP
HSPA1B
DNAJB1
ADH1B
NEAT1
PIK3R1
GPX3
XIST
TNFAIP3
COMP
CILP
RGS16
CAV1
ALDH1A3
ASPN
ADAMTS15
EDIL3
LRRC15
SERPINE1
MMP1
CXCL5
CA12
MME
CXCL13
MMP3
CHI3L1
IL7R
POSTN
TDO2
Cancer Control
HIV
HIVTB
TB
Percent Expressed
0
25
50
75
Average Expression
1.0
0.5
0.0
−0.5
−1.0
C
Fibroblast−M1
ADAMDEC1
HILPDA
CTSK
TFRC
HLA−B
WNT5A
PCSK5
DCBLD1
PDPN
EGFL6
CXCL13
CHI3L2
COL12A1
FTH1
CHI3L1
CXCL1
POSTN
INHBA
MTRNR2L12
TDO2
IL7R
MME
CXCL5
CA12
MMP1
kME
Fibroblast−M2
LOX
BGN
RGS4
RARRES1
NGFR
PLA2G2A
RP11−572C15.6
IFI27
SULF1
CLDN1
IGF1
MXRA5
C3
COL3A1
SULF2
PRSS23
CRABP2
COL5A1
CCDC80
SERPINE2
SPARC
PTGIS
COL1A2
COL1A1
H19
Fibroblast−M3
SPSB1
PNRC1
SOCS3
SCN7A
CFD
RGS2
MYC
ZFP36
ARID5B
KDM6B
CEBPD
GPRC5A
A2M
SLC19A2
GEM
CDKN1A
MGP
IL6
CCL2
JMJD1C
MT2A
CRISPLD2
KLF9
ELL2
C7
Fibroblast−M4
CNPPD1
SNORD67
ZAP70
TMEM41A
HAUS6
SMPD4
ANK3
IRF4
ZNF586
RAPGEF5
DNAJC27
SYNE2
CDC42SE2
RUNX3
SH2D1A
CXCR4
KANK3
SP140L
GNA15
TBX19
GBP5
OTUD3
P2RY10
KIAA0895
SLA
Fibroblast−M5
SMAP2
TNFAIP3
DOCK8
LINC00243
GIMAP7
SPOCK2
SATB1
CCR7
SLC38A1
CCR4
FAM65B
GPRIN3
KLRB1
MT1G
PTPRC
ITGAL
RASGRP1
AIM1
IKZF1
CD2
SELL
CD28
FYB
ICOS
LTB
Fibroblast−M6
ALKBH8
CITED4
FANCM
NIPAL1
CCNF
ATG16L1
LILRA5
ZMYM3
ADAMTS6
BAIAP3
AC005154.6
NBEAL2
TSPAN32
TMEM86A
PRDM11
GEMIN5
ZCCHC4
RAPGEF4
SLC6A1
STXBP6
FHAD1
ACP6
ZNF662
SLC24A1
VANGL2
Fibroblast−M7
ZNHIT2
MCU
ZNF761
CDS1
EEF1E1
MSR1
TRIM7
TMEM107
MFSD5
MKL2
KCNMA1
STRA6
TRMT161B
WAS
ENPP1
EIF2B5
MARS2
NR2C2AP
ALG10B
PLA2R1
ZNF30
SNRNP35
PNMAL1
ZNF529
RP11−332H14.2
B
DAPI
COLLAGEN1
CA12
MMP1
TDO2
CTHRC1
MMP3
PI15
MERGED
D
WGCNA modules
M7
M6
M5
M4
M3
M2
M1
IL6+CCL2+ Fibroblast
SERPINE2+COL1A1+ Fibroblast
HSP+ Fibroblast
COMP+CILP+ Fibroblast
MMP1+CXCL5+ Fibroblast
Subcluster
Average Expression
0.0 0.5 1.0
Percent Expressed
0 25 50 75 100
Cancer control
HIV
HIVTB
TB
Disease status
Average Expression
−0.5 0.0 0.5 1.0
Percent Expressed
0 20 40 60 80
E
Module score
4wk
10wk
****
***
****
ns
ns
ns
****
Fibroblast-M1
Fibroblast-M2
Fibroblast-M3
Fibroblast-M4
Fibroblast-M5
Fibroblast-M6
Fibroblast-M7


A
Fibroblast signals strengthened in TB disease lungs
Sender
Receiver
Fibroblast
Alveolar Macrophage
ARL4C+EMP1+ Macrophages
Heatshock Macrophage
SPP1+CHI3L1+ Macrophages
B
SPP1+ macrophage signals strengthened in TB disease lungs
Fibroblast
Macrophage
Monocyte
C
Raghavan et al. SPP1+ TAM marker genes
Heatshock Macrophages
SPP1+CHI3L1+ Macrophages
ARL4C+EMP1+ Macrophages
LGMN+SEPP1+ Macrophages
Alveolar Macrophages
SPP1 FLT1 FBP1 LDHA LPL C15orf48 EMP1 CSTB FN1 ENO1 PKM ALDOA GAPDH CHI3L1
Average Expression
−1 0 1
Percent Expressed
20 40 60 80
D
profibrotic SPP1+
FDX1 LPL CHIT1 CHI3L1 MMP9 SPP1
MARCO+
MARCO FBP1
inflammatory C1QA+
AKR1B1 IL18 CD4 PLTP TREM2 C1QA CCL18
inflammatory CCL2+
S100A12 C15orf48 CCL2
MDMs
CSF1R MAFB SOD2 ITGAM
HLCA MDM subset markers
Heatshock Macrophages
SPP1+CHI3L1+ Macrophages
ARL4C+EMP1+ Macrophages
LGMN+SEPP1+ Macrophages
Alveolar Macrophages
Average expression
1 0 −1
Percent expressed
20 40 60 80
E
CD16+ Monocytes
ARL4C+EMP1+ Macrophages
LGMN+SEPP1+ Macrophages
SPP1+CHI3L1+ Macrophages
MMP1+CXCL5+ Fibroblast
SERPINE2+COL1A1+ Fibroblast
COMP+CILP+ Fibroblast
IL6+CCL2+ Fibroblast
Endothelial
Monocytes
Inflammed Monocytes
Low Complexity Monocytes
CD8 cytotoxic T
NK
Alveolar Macrophages
DC
Mast
IFN Response Monocytes
AT1
AT2
Club cells
Fibroblast|HSP+
T
B cell
Plasma
Heatshock Neutrophil
GBP1+GBP5+ Neutrophil
MMP9+CST7+ Neutrophil
Receiver
Interaction specificity
0.1 0.2
edge_FC
2500 2000 1500 1000 500
F
DAPI-SPP1
DAPI-CD68
DAPI-CD206
MERGED
Lesion 1
Lesion 2
SPP1 expression intensity
***
***
***
Outer cellular layer
Inner cellular layer
Necrotic core
Expression ratio
ns
***
inner/outer
inner/core
SPP1 CD68