Title: Mycobacterial exposure remodels alveolar macrophages and the early innate response to *Mycobacterium tuberculosis* infection

Authors:

Dat Mai[1], Ana Jahn[1], Tara Murray[1], Michael Morikubo[1], Johannes Nemeth[1,3], Kevin Urdahl[1], Alan H. Diercks[1], Alan Aderem[1], Alissa C. Rothchild[2,*]

Affiliations:

[1] Center for Global Infectious Disease Research, Seattle Children's Research Institute, Seattle, WA 98109

[2] Department of Veterinary and Animal Sciences, University of Massachusetts Amherst, Amherst, MA 01003

[3] Current address: University Hospital Zurich, University of Zurich, Division of Infectious Diseases and Hospital Epidemiology, Raemistrasse 100, 8091 Zürich, Switzerland.

*Corresponding author: arothchild@umass.edu

Keywords:

innate immunity, alveolar macrophages, *Mycobacterium tuberculosis*, lung, vaccination, bacterial infection, host-pathogen interaction

Abstract

As innate sentinels in the lung, alveolar macrophages (AMs) play a critical role during *Mycobacterium tuberculosis* (Mtb) infection as the first cells to encounter bacteria. We previously showed that AMs initially respond to Mtb infection *in vivo* by mounting a cell-protective, rather than pro-inflammatory response, yet whether the AM response could be modified by environmental factors was unknown. Here, we characterize how previous exposure to mycobacteria, either through subcutaneous vaccination with *Mycobacterium bovis* (scBCG) or through a contained Mtb infection (coMtb), impacts the initial response by AMs and early innate response in the lung. We find that both scBCG and coMtb accelerate early innate cell activation and recruitment and generate a stronger pro-inflammatory AM response to Mtb *in vivo*. AMs from scBCG vaccinated mice mount a robust interferon response, while AMs from coMtb mice produce a broader and more diverse inflammatory response. Using single-cell RNA-sequencing, we identify exposure-induced changes to airway-resident cells, with scBCG and coMtb enriching for different AM subpopulations. *Ex vivo* stimulation assays reveal that AMs from scBCG and coMtb mice switch on an interferon-dependent response, which is absent in AMs from unexposed mice. Overall, our studies reveal significant, durable, and cell-intrinsic modifications to AMs following exposure to *mycobacterium,* and comparison of scBCG and coMtb models reveals that AMs can be reprogrammed into more than one state. These findings highlight the plasticity of innate responses in the airway and underscore the unexplored potential of targeting AMs through vaccination or host-directed therapy to augment host responses.

Introduction

Mycobacterium tuberculosis (Mtb), the causative agent of Tuberculosis (TB), leads to more than 1.5 million deaths each year and, for the first time since 2005, the number of TB deaths worldwide is increasing[1, 2]. These trends highlight the urgent need for new vaccination strategies. Traditionally, vaccine design has focused on generating a rapid, robust, and effective adaptive immune response. However, recent studies suggest that the innate immune system can undergo long-term changes in the form of trained immunity[3], which affect the outcome of infection and may be important components of an effective TB vaccine[4, 5]. Initial studies focused primarily on central trained immunity, long-term changes to hematopoietic stem cells that lead to functional changes in short-lived innate cell compartments (i.e., monocytes, NK cells, dendritic cells)[3]. In the context of TB, exposure to BCG and Mtb were found to have opposing effects on central myelopoiesis: BCG vaccination augments myelopoiesis while Mtb infection limits it[6, 7]. More recent studies have examined innate training in tissue-resident macrophages and demonstrated that these cells can respond to remote injury and inflammation[8], undergo long-term changes[3], and display altered responses to bacteria after pulmonary viral infection[9, 10, 11].

Lung resident alveolar macrophages (AMs) are the first cells to become infected with inhaled Mtb and engage a cell-protective response, mediated by the transcription factor Nrf2, that impedes their ability to control bacterial growth[12, 13]. In this study, we examined how prior mycobacterial exposure reprograms AMs and alters the overall innate response in the lung to aerosol challenge with Mtb. We chose to compare the effects of subcutaneous BCG vaccination (scBCG) with those arising from a contained Mtb-infection model (coMtb) as these represent two of the most common mycobacterial exposures in humans, and both exposures are associated with changes to innate immunity. In mice with an established contained but persistent lymph node Mtb infection, generated by intradermal inoculation, the AM response to inhaled Mtb is more inflammatory and bacterial growth over the course of disease is restricted compared to naïve mice[14, 15]. CoMtb also protects mice against heterologous challenges,

including infection with *Listeria monocytogenes* and expansion of B16 melanoma cells, suggesting substantial remodeling of innate immune responses[15]. BCG, a live-attenuated TB vaccine derived from *M. bovis,* provides protection against disseminated pediatric disease but has lower efficiency against adult pulmonary disease[16, 17]. In addition to enhancement of Mtb-specific adaptive responses, based on shared antigens, BCG vaccination also leads to changes in hematopoiesis and epigenetic reprogramming of myeloid cells in the bone marrow[7], early monocyte recruitment and Mtb dissemination[18], and innate activation of dendritic cells critical for T cell priming[19]. Intranasal BCG vaccination protects against *Streptococcus pneumoniae* and induces long term activation of AMs[20]. BCG vaccination is also associated with trained immunity effects in humans[21, 22, 23], including well-described reductions in all-cause neonatal mortality and protection against bladder cancer[3, 24].

Here, we show that while both coMtb and scBCG protect against low dose Mtb aerosol challenge, they remodel the innate response in different ways. In AMs, scBCG enables a largely interferon response to infection, while coMtb promotes a broader pro-inflammatory response. Prior mycobacteria exposure also generates altered immune cellularity in the lung prior to aerosol challenge and significant changes in the early dynamics of the overall innate response. The observation that the innate response to Mtb, including intrinsic effects of AMs, can be reprogrammed in multiple directions highlights the plasticity of tissue resident cells and suggests opportunities to modulate their function through vaccination or host-directed therapies.

Results

Prior mycobacterial exposure accelerates activation and innate cell recruitment that is associated with protection from Mtb infection

We first determined the earliest stage of infection when the immune response was altered by prior exposure to mycobacteria. Mice were vaccinated with scBCG ($1x10^6$ *M. bovis* BCG, Pasteur strain) or treated with coMtb ($1x10^4$ H37Rv, delivered intradermally in the ear),

rested for 8 weeks, and then challenged with low-dose (~100 CFU) H37Rv aerosol infection. We measured both the cellularity and activation of innate immune cells in the lung at 10, 12 and 14 days following infection, the earliest timepoints when innate cells are known to be recruited[12, 13, 25]. We observed a significant increase in MHC II Median Fluorescence Intensity (MFI) as early as day 10 for AMs from coMtb mice and day 12 for AMs from scBCG mice compared to controls (**Fig 1A**, **S1**). There were no significant differences in MHC II expression prior to challenge on day 0 (**Fig 1A**). There were significant increases in the numbers of monocyte-derived macrophages (MDM), neutrophils (PMN), dendritic cells, $Ly6C^+$ $CD11b^+$ monocytes, and $Ly6C^+$ $CD11b^-$ monocytes by day 10 in coMtb mice compared to controls, with further increases by days 12 and 14 (**Fig 1B, Fig S1**). scBCG elicited similar increases in these populations starting at day 10, but the increases were not as robust or rapid as those observed in the coMtb. Significant differences between scBCG and coMtb groups were found at d10, d12, d14 in MDM, d14 in PMN, d12, d14 in dendritic cells, and d14 in $Ly6C^+$ $CD11b^+$ monocytes (**Fig 1B,** designated with +).

In addition to early changes in innate cell activation and recruitment, we observed earlier recruitment of activated $CD44^+$ $CD4^+$ and $CD8^+$ T cells in both coMtb and scBCG mice starting at day 10 as well as TB antigen-specific T cells, ESAT6-$tetramer^+$ $CD4^+$ T cells and TB10.4-$tetramer^+$ $CD8^+$ T cells in coMtb mice starting at day 10 compared to controls and to scBCG mice (**Fig 1C, Fig S1**). The differences in the recruitment of ESAT6-$tetramer^+$ $CD4^+$ T cells between scBCG and coMtb were expected, as the ESAT6 antigen is expressed by H37Rv but not BCG.

We also evaluated whether these cell recruitment differences correlated with changes in bacterial burden. We found that both modalities generated a significant reduction in bacterial burden in the lung, spleen, and lung-draining lymph node (LN) at day 14 and at day 28, as has been previously reported[15, 18, 26] (**Fig S2A-D)**. Interestingly, we also observed a significant reduction in bacterial burden in the lung by day 12 in coMtb but not scBCG mice and significant

reduction in CFU in the LN in both models compared to controls, at a time when we first detected innate cell recruitment to the lungs of coMtb mice (**Fig S2B**). The majority of control mice had undetectable bacteria in spleen and LN at day 10 and there was no reduction in bacterial burden in the lung at that time. Our results demonstrate that prior mycobacteria exposure leads to accelerated innate activation and innate cell recruitment within the first two weeks of infection, with coMtb generating a faster and more robust response compared to scBCG. These early immune changes are associated with early reductions in bacterial load in the lung, while reductions in bacterial burden in the LN and spleen suggest delays in bacterial dissemination.

Mycobacterium exposure alters the *in vivo* alveolar macrophage response to Mtb infection

To examine the earliest response to Mtb, we measured the gene expression profiles of Mtb-infected AMs 24 hours following aerosol challenge with high dose mEmerald-H37Rv (depositions: 4667, 4800) in scBCG-vaccinated mice and compared these measurements to previously generated profiles of AMs from control (unexposed) mice[12] and coMtb mice[15] (**Table S1**). Mtb-infection induced distinct expression changes for coMtb, scBCG, and control AMs as measured by Principal Component Analysis (**Fig 2A**) and the majority of up-regulated Differentially Expressed Genes (DEG) (fold change > 2, FDR < 0.05) were unique to each condition (control: 151 unique genes/257 total genes, scBCG: 222/289, coMtb: 156/229) (**Fig 2B**). The divergence in the responses of Mtb-infected AMs from each of the 3 conditions was also reflected in the diversity in the Top 20 Canonical Pathways identified by Ingenuity Pathway Analysis (**Fig S3**).

To identify trends between groups, we performed Gene Set Enrichment Analysis using a set of 50 Hallmark Pathways. As we've shown previously, Mtb-infected AMs from control mice at 24 hours had strong enrichment for "Xenobiotic Metabolism" and "Reactive Oxygen Species"

pathways, indicative of the Nrf2-associated cell-protective response (**Fig 2C**). While these two pathways were not among most enriched pathways in the exposed groups, Mtb-infected AMs from all groups upregulated genes associated with the cell-protective Nrf2-driven response[12] (**Fig 2D**). Expression profiles for Mtb-infected AMs from scBCG mice showed the strongest enrichment for "Interferon Alpha Response" and "Interferon Gamma Response" pathways, which contain many shared genes (**Fig 2C**). The strength of the interferon response in these AMs was further highlighted by examining gene expression changes in a set of Interferon Stimulated Genes (ISGs) identified from macrophages responding to IFNα (fold change > 2, p-value < 0.01)[27] (**Fig 2D**). Expression profiles for Mtb-infected AMs from coMtb mice showed a weaker enrichment for interferon response pathways with fewer up-regulated ISGs compared to scBCG, and instead showed enrichment across a number of inflammatory pathways including "IL6 JAK STAT3 signaling" in comparison to the other groups (**Fig 2C, D**).

In summary, mycobacteria exposures alter the initial *in vivo* response of AMs to Mtb infection 24 hours after challenge and remodel the AM response in different ways. AMs from scBCG vaccinated animals mount a robust interferon associated response, while AMs from coMtb mice generate a more expansive inflammatory program across multiple pathways in response to infection.

***Mycobacterium* exposure modifies the baseline phenotype of alveolar macrophages in the airway**

Although scBCG and coMtb mice potentiate significantly altered AM responses to Mtb infection *in vivo*, these effects are not widely evident in the transcriptomes prior to infection as measured by population-level RNA-sequencing (**Fig S4**). However, we posited that remodeling effects were likely not homogenous across the entire AM population and that small heterogenous changes to baseline profiles might be detectable using a single cell approach. We therefore analyzed pooled BAL samples taken from 10 age- and sex-matched mice from each

of the three conditions (control, scBCG, coMtb) eight weeks following mycobacteria exposure by single cell RNA-sequencing (scRNAseq). Gross cellularity was unaffected by mycobacterial exposure as measured by flow cytometry analysis of common lineage markers with AMs the dominant hematopoietic cell type (57.4-85.8% of $CD45^+$ live cells), followed by lymphocytes (5.26-22.7% of $CD45^+$ live cells) with small contributions from other innate cell populations (**Fig S5**).

Six samples, with an average of 2,709 cells per sample (range: 2,117-4,232), were analyzed together for a total of 17,788 genes detected. The most prominent expression cluster mapped to an AM profile, with smaller clusters mapping to T and B lymphocytes, dendritic cells, and neutrophils (**Fig 3A**). All cells that mapped to a macrophage profile were extracted and reclustered into 11 macrophage subclusters (**Fig 3B, C**). All but one of the macrophage subclusters expressed AM lineage markers (*Siglecf, Mertk, Fcgr1* (CD64), *Lyz2* (LysM), and *Itgax* (CD11c) and had low expression of *Itgam* (CD11b) (**Fig 3D**). The one exception was cluster 6 that showed high *Itgam* and *Lyz2* expression and lower *Siglecf* expression, likely representing a small monocyte-derived macrophage population in the airway.

To interpret the various expression subclusters, we identified the genes that most distinguished each cluster from the others (**Fig S6, Table S2**). As has been reported by other groups[28, 29], a small proportion of the AMs formed two clusters (4, 10) with high expression of cell cycle genes (i.e., *Top2a, Mki67*), indicative of cell proliferation (**Fig 3E, Table S2**). Cluster 0 formed the most abundant macrophage cluster with high expression of lipid metabolism genes (i.e., *Abcg1, Fabp1*) and with a trend of slightly decreased relative frequency in scBCG and coMtb samples compared to controls (**Fig 3F, Table S2**). Cluster 2 was significantly increased in relative frequency for scBCG samples compared to coMtb ($p = 0.032$, *One-way ANOVA with Tukey post-test*) and associated with oxidative stress response genes (*Hmox1, Gclm*) **(Fig 3G, Table S2**). Cluster 7 was the only cluster with an increase in relative frequency trending for both scBCG and coMtb ($p = 0.076$, *One-way ANOVA*). AMs in this cluster had high expression of

Interferon Stimulated Genes (**Fig 3H, Table S2**). Cluster 3 had significantly higher relative frequency for coMtb samples compared to control and scBCG samples (p = 0.021, 0.039, respectively, *One-way ANOVA with Tukey post-test*) and was distinguished by expression of the macrophage-associated transcription factors (*Cebpb*, *Zeb2*, *Bhlhe40*)[30, 31], hemoglobin metabolism (*Hba-a1, Hba-a2, Hbb-bs*), mitochondrial oxidative phosphorylation (*mt-Co1, mt-Cytb, mt-Nd2*), chromatin remodeling (*Ankrd11, Baz1a*), and immune signaling including the CARD9 complex (*Malt1, Spag9, Bcl10, Prkcd*) (**Fig 3I, S7, Table S2**). This expression profile closely matches a subcluster of AMs previously described by Pisu et al, as an "interstitial macrophage-like" AM population (labeled "AM_2") that expanded in relative frequency in lung samples 3 weeks following low-dose H37Rv infection[28]. The changes in the baseline phenotypes revealed by scRNAseq correspond to differences observed in the AM *in vivo* response (**Fig 2**), with scBCG driving AMs towards an interferon response and coMtb shifting AMs towards a pro-inflammatory response with more diverse qualities.

Interestingly, Cluster 2 (higher relative frequency in scBCG) and Cluster 3 (higher relative frequency in coMtb) represent divergent endpoints of a pseudotime plot generated by performing a trajectory inference analysis, regardless of whether the starting point is the most abundant cluster in the control group (Cluster 0) (**Fig 3J, *top***) or the cluster of proliferating cells (Cluster 4) (**Fig 3J, *bottom***). This result suggests that scBCG and coMtb drive AM phenotypes in different directions and may indicate the possibility of driving different flavors of an innate tissue-resident response, rather than flipping an "on/off" switch.

***Mycobacterium* exposure modifies T cell populations in the airway**

While AMs are the dominant immune cell type in the airway, other cell populations make up an average of 18.4% of the cells within the BAL in controls (range: 10.4-26.3%) and 31.3% in exposed groups (range: 14.0-48.8%). To determine how mycobacteria exposure influenced other cells in the airway, we focused on T cells and dendritic cells (DCs) which have the two

highest relative frequencies after AMs (**Fig 4A, B**). As expected, the overall frequency of T cells in BAL was increased following exposure (**Fig 4B**). To examine the alterations in T cell cellularity in detail, we combined and reclustered the two original T cell clusters into 7, which were manually annotated using the closest Immgen profiles and the expression of key lineage specific markers (**Fig 4A-C, Fig S8**). We focused specifically on the 5 most abundant T cell subclusters (Clusters 0-4). While we observed subtle shifts in the relative frequency of each group, none reached statistical significance. Cluster 0, the most abundant cluster, had an expression profile most consistent with γδ T cells, including expression of *Cd3e* with low to nil *Cd4* and *Cd8a* and some expression of *Zbtb16* (PLZF) and *Tmem176a*, an ion channel regulated by RORγt and reported to be expressed by lung γδ T cells[32, 33] (**Fig 4D-F, S8**). Cluster 1, consistent with a profile for effector $CD4^+$ T cells, had slightly higher relative frequency in scBCG samples (**Fig 4D-F, S8**). Cluster 2 had a profile associated with naïve $CD8^+$ T cells, with minimal differences in relative frequency between groups (**Fig 4D-F, S8**). Cluster 3, with high relative frequency in just one of the two coMtb samples, had a profile consistent with effector memory/resident memory $CD8^+$ T cells ($T_{EM/RM}$). (**Fig 4D-F, Fig S7**). Lastly, Cluster 4, with slightly higher relative frequency for both scBCG and coMtb samples compared to controls, had a profile consistent with NK cells. Overall, *mycobacterium* exposure increases the total frequency of T cells in the airway, but leads to only subtle shifts in relative frequencies of T cell subclusters in the airway, mostly notably a slight increase in effector T cell or memory T cell clusters and a decrease in naïve T cells, trends that have been documented previously[18, 26].

***Mycobacterium* exposure modifies the dendritic cell airway landscape**

Similar re-clustering of DCs yielded 2 major clusters and 1 minor cluster (**Fig 4G**). BAL from coMtb mice had a slightly higher relative frequency of Cluster 0, notable for expression of *Clec9a, Itgae* (CD103), consistent with an expression profile of lung $CD103^+$ cDCs[34]. Cluster 0

also had high expression of antigen presenting molecules such as *H2-Ab1* and *H2-DMa* (**Fig 4I**). Cluster 1 had slightly higher relative frequency in control BAL and expressed *Batf3, Ccr7*, and *Fscn1*. Cluster 2 had a mixed phenotype with expression of genes from both major clusters. All of the clusters had high *Irf8* expression and low expression of *Xcr1*, *Irf4*, and *Itgam* (CD11b), (**Fig 4I**). Antigen presentation genes were more highly expressed by Cluster 0, while *Ccr7,* which is known to be important for migration of DC out of the lung into the draining lymph node during Mtb infection[35], was more highly expressed by Cluster 1, hinting at a potential division of labor for T cell priming. Overall, scRNAseq analysis shows that mycobacteria exposure alters the airway landscape and may impact how airway resident T cells and DCs interact with AMs and with each other following subsequent aerosol infection.

Cell-intrinsic remodeling following *mycobacterium* exposure licenses an alveolar macrophage interferon response *in vitro*

In vivo transcriptional analysis of the AM response within 24 hours of aerosol challenge demonstrates that the very earliest immune response to Mtb is altered by previous *mycobacterium* exposure. However, a trade-off to measuring responses *in vivo* is the inability to discern whether observed changes are cell-intrinsic or dependent on the changed tissue environment. Therefore, to determine whether mycobacteria exposure induces cell-intrinsic changes to AM responses or whether AMs simply respond differently due to a changed environment (i.e., the presence of greater number of antigen-specific T cells), we isolated AMs from control, scBCG, or coMtb mice, stimulated them *ex vivo* with LPS (10 ng/ml), Pam3Cys (10 ng/ml), or H37Rv, and measured their transcriptional profile 6 hours later. AMs from both coMtb and scBCG mice showed remarkably different responses than AMs from control mice to both LPS and H37Rv stimulation, and minimal differences in response to Pam3Cys stimulation (**Fig 5A, S9, Table S3**). Performing Gene Set Enrichment Analysis, we found that the greatest changes for LPS and H37Rv responses were associated with “Interferon Gamma Response”,

"Interferon Alpha Response", "TNFa signaling via NF-kB", and "Inflammatory Response" pathways (**Fig 5B)**. To assess whether the cell-intrinsic changes observed were long-lasting, we compared the responses of AMs at 8 or 23 weeks following scBCG vaccination by RT-qPCR. Increases in gene expression were as robust or even enhanced 23 weeks following exposure compared to 8 weeks, suggesting that exposure-induced changes to AMs are relatively long-lived (**Fig S10**).

We observed that many of the genes whose response to H37Rv stimulation was altered by mycobacteria exposure could be categorized as ISGs[27] (**Fig 5C, S9**). We previously defined "Type I IFN dependent" (352 genes) or "Type I IFN independent" (339 genes) portions of the response of bone-marrow-derived macrophages to H37Rv stimulation using cells from IFNAR$^{-/-}$ mice[36]. Interestingly, expression of H37Rv-induced IFN dependent genes was minimally induced by H37Rv in control AMs but strongly induced in AMs from *mycobacterium* exposed mice (**Fig 5D, *left***). In contrast, expression of H37Rv-induced IFN independent genes was modestly upregulated in control AMs and only slightly altered by *mycobacterium* exposure (**Fig 5D, *right***). These results demonstrate that previous *mycobacterium* exposure leads to cell-intrinsic changes in AMs that are most evident following secondary stimulation and can be long-lasting and that one of the most significant of these is the licensing of a more robust Type I Interferon response.

Discussion

Here, we describe remodeling of AMs, long-lived airway-resident innate immune cells, following two types of *mycobacterium* exposure, scBCG vaccination and coMtb, a model of contained H37Rv infection. We find that prior *mycobacterium* exposure generates faster innate cell activation and immune cell recruitment to the lung following Mtb aerosol infection, immune modifications that are detectable within the first 10-14 days of infection. We observe that the AM response within the first 24 hours of infection is substantially altered in mice previously exposed

to mycobacteria and that scBCG and coMtb lead to qualitatively different AM responses: AMs from BCG vaccinated mice mount a robust and dominant interferon response, while AMs from coMtb mice generate a broader inflammatory response that includes but is not dominated by interferon pathways. Profiling the airway landscape by scRNAseq, we find sub-populations of AMs that are differentially enriched following scBCG or coMtb exposures. These sub-populations, defined by expression of ISGs, oxidative stress response genes, or interstitial macrophage-like profiles, reveal how prior exposure may generate different flavors of innate immune responses in the airway. Subtle shifts in the relative frequency and phenotype of T cells and DCs within the airway following *mycobacterium* exposure highlight potential areas where innate and adaptive interactions may be substantially altered during the early stages of infection, complementing previous descriptions of accelerated immune responses following vaccination[18, 25, 26]. *Ex vivo* stimulations of AMs from either scBCG or coMtb mice demonstrate cell-intrinsic effects, with the most robust expression changes occurring in genes that are IFN-dependent, suggesting that prior *mycobacterium* exposure licenses AMs to mount an interferon response that is otherwise lacking in control animals.

AMs are the first cells to be productively infected in the lung following aerosol Mtb infection[12, 13], but their role during subsequent stages of infection and their contribution to protection mediated by vaccination or concomitant immunity is not fully understood. We previously showed that AMs initially respond to Mtb infection with a cell-protective, Nrf2-driven program that is detrimental to early host control[12]. Others have shown that depletion of AMs or strategies that “skip” the AM stage including directly injecting antigen-primed DCs or activating DCs accelerates the immune response and reduces bacterial burden[19, 37, 38]. However, strategies that target AMs to make them better early responders have not been well studied[39]. We demonstrate here that following mycobacteria exposure, whether from scBCG or coMtb, the AM response to Mtb becomes more pro-inflammatory, including an up-regulation of ISGs. This coincides with an accelerated innate response and early bacterial control. Typically, interferon

signatures are associated with active TB or TB disease progression in both humans and non-human primates[40, 41, 42]. In addition, type I IFN has known negative consequences for infection. Host perturbations such as treatment with poly I:C or viral co-infection that induce type I IFN lead to worsened disease[43, 44], type I IFN has been shown to block production of IL-1β in myeloid cells during Mtb infection[45], and type I IFN drives mitochondrial stress and metabolic dysfunction in Mtb infected macrophages[36]. However, there are also examples where AM production of type I IFN is critical for host protection due to the ability to activate other innate cells in the lung during acute viral infection[46, 47]. Determining whether early type I IFN production by AMs helps or hurts in the acceleration of the host response during Mtb will require further study.

What signals are required to induce long-term remodeling in AMs? A number of recent studies investigated how prior viral infection alters AM function, uncovering either enhanced AM antimicrobial phenotypes[9, 10, 11] or impaired responses[48, 49] following exposure. Other studies have discovered long-lasting changes to AMs following intranasal immunization of either adenoviral-based or inactivated whole cell vaccines[20, 50, 51]. Several reports have identified T cell-derived IFNγ as critical for altering AM function, although the immunological outcome varies substantially based on the context. In one study, T cell-derived IFNγ following adenoviral infection leads to AM activation, innate training and protection from *S. pneumoniae*[10], while in another study influenza-induced T cell-derived IFNγ leads to AM dysfunction and impaired clearance of *S. pneumoniae*[49]. Importantly, a study of BAL samples from 88 SARS-CoV-2 patients identified AMs and T cell-derived IFNγ as part of a positive feedback loop in the airway that generates more progressive disease and lung pathology[52]. IFNγ is a likely candidate driving the effects of *mycobacterium* exposure on AMs described in this study. IFNγ was shown to be important for the generation of trained immunity in bone marrow-derived myeloid cells following BCG vaccination[6, 7]. While pulmonary H37Rv infection is associated with induction of type I IFN

and a reduction in myeloid training[6], we previously found that coMtb leads to low-level systemic cytokinemia, including IFNγ production, and using WT:Ifngr1$^{-/-}$ mixed bone marrow chimeras, we showed that IFNγ signaling was responsible for monocyte and AM activation following establishment of coMtb[15].

By providing a side-by-side comparison of multiple remodeled states of AMs following different initial exposures (scBCG vs coMtb), our study highlights the plasticity of AM phenotypes and the impact of the local and/or systemic environments. Heterogeneity in myeloid reprogramming is a feature that has also been demonstrated in human monocytes[53]. While scBCG and coMtb provide models of mycobacterial exposure that are highly relevant to human biology and host protection derived from vaccination or concomitant immunity, they also have drawbacks. One significant limitation is that perturbations of the likely mechanisms for scBCG or coMtb-mediated AM reprogramming alter the initial exposures themselves. For example, containment in the coMtb model is lost following T cell depletion or anti-IFNγ blockade[14]. For this reason, we have not been able to dissect how different signals derived from scBCG versus coMtb push AMs towards different phenotypic states. Instead, we envision future studies using other models that can examine the specific effects of individual cytokines etc. on AM remodeling.

There are many other remaining questions. We do not yet know whether the altered AM phenotypes require constant exposure from the changed environment or whether the cell-intrinsic nature is stable in the absence of environmental cues. These would require complex cell transfer experiments that are beyond the scope of this study. Additionally, we do not yet know the durability of these changes and whether they are mediated by epigenetic effects. Our longest experiment showed retention of changes to AMs after 23 weeks. Whether the effects derived from *mycobacterium* exposure require ongoing bacterial replication is also unclear. In Nemeth et al, we showed that antibiotic treatment lessened the protection mediated by coMtb,

suggesting that ongoing replication is a key part of the effect[15]. However, we rarely find live bacteria in the lungs of scBCG vaccinated mice after 8 weeks yet AMs from these mice show robust alterations, suggesting that the effects can be mediated in the absence of ongoing bacterial replication. We performed several of these studies with intravenous BCG vaccination (ivBCG), which in the mouse model leads to much more bacterial dissemination and ongoing replication (data not shown). While we saw similar changes to AMs in the ivBCG model, these were not dramatically different than those of scBCG vaccination, despite major differences in bacterial replication and far greater T cell recruitment to the airway.

There is still much unknown about the signals that drive reprogramming of tissue-resident innate cells. Ideally, vaccines would be designed to leverage these signals and generate the most effective interactions between innate and adaptive responses. Identifying the ways that AMs are reprogrammed by inflammatory signals and the effects of their changed phenotypes on the early stages of infection will help to improve future vaccines or host-directed therapies.

Materials and Methods

Mice

C57BL/6 mice were purchased from Jackson Laboratories (Bar Harbor, ME). Mice were housed and maintained in specific pathogen-free conditions at Seattle Children's Research Institute and experiments were performed in compliance with the Institutional Animal Care and Use Committee. 6-12 week old male and female mice were used for all experiments, except for RNA-sequencing, which used only female mice for uniformity. Mice infected with Mtb were housed in a Biosafety Level 3 facility in an Animal Biohazard Containment Suite.

Mycobacteria exposure models: BCG immunization and establishment of coMtb

BCG-Pasteur was cultured in Middlebrook 7H9 broth at 37°C to an OD of 0.1–0.3. Bacteria was diluted in PBS and 1 x 10^6 CFU in 200 ml was injected SC. Intradermal infections to establish coMtb were performed as formerly described[14], with some modifications as detailed previously[15]. Briefly, 10,000 CFU of Mtb (H37Rv) in logarithmic phase growth were injected intradermally into the ear in 10 µL PBS using a 10 µL Hamilton Syringe, following anesthesia with ketamine/xylazine.

M. tuberculosis Aerosol Infections and Lung Mononuclear Cell Isolation

Aerosol infections were performed with wildtype H37Rv, including some transformed with an mEmerald reporter pMV261 plasmid, generously provided by Dr. Chris Sassetti and Christina Baer (University of Massachusetts Medical School, Worcester, MA). For both standard (~100 CFU) and high dose (~2,000-4,000 CFU) infections, mice were enclosed in an aerosol infection chamber (Glas-Col) and frozen stocks of bacteria were thawed and placed inside the associated nebulizer. To determine the infectious dose, three mice in each infection were sacrificed one day later and lung homogenates were plated onto 7H10 plates for CFU enumeration. High dose challenge and sorting of Mtb-infected AM was performed 4 weeks following scBCG vaccination and 2 weeks following coMtb vaccination as previously described[54]. All other analysis was performed 8 weeks following mycobacterium exposures.

Lung Single Cell Suspensions

At each time point, lungs were removed, and single-cell suspensions of lung mononuclear cells were prepared by Liberase Blendzyme 3 (70 ug/ml, Roche) digestion containing DNaseI (30 µg/ml; Sigma-Aldrich) for 30 mins at 37°C and mechanical disruption using a gentleMACS dissociator (Miltenyi Biotec), followed by filtering through a 70 µM cell strainer. Cells were resuspended in FACS buffer (PBS, 1% FBS, and 0.1% NaN_3) prior to staining for flow cytometry. For bacterial enumeration, lungs were processed in 0.05% Tween-80 in PBS using a

gentleMACS dissociator (Miltenyi Biotec) and were plated onto 7H10 plates for CFU enumeration.

Alveolar Macrophage Isolation

Bronchoalveolar lavage was performed by exposing the trachea of euthanized mice, puncturing the trachea with Vannas Micro Scissors (VWR) and injecting 1 mL PBS using a 20G-1" IV catheter (McKesson) connected to a 1 mL syringe. The PBS was flushed into the lung and then aspirated three times and the recovered fluid was placed in a 15mL tube on ice. The wash was repeated 3 additional times. Cells were filtered and spun down. For antibody staining, cells were suspended in FACS buffer. For cell culture, cells were plated at a density of 5 x 10^4 cells/well (96-well plate) in complete RPMI (RPMI plus FBS (10%, VWR), L-glutamine (2mM, Invitrogen), and Penicillin-Streptomycin (100 U/ml; Invitrogen) and allowed to adhere overnight in a 37°C humidified incubator (5% CO_2). Media with antibiotics were washed out prior to infection with *M. tuberculosis*.

Cell Sorting and Flow Cytometry

Fc receptors were blocked with anti-CD16/32 (2.4G2, BD Pharmingen). Cell viability was assessed using Zombie Violet dye (Biolegend). Cells were suspended in 1X PBS (pH 7.4) containing 0.01% NaN_3 and 1% fetal bovine serum (i.e., FACS buffer). Surface staining, performed at 4 degrees for 20 minutes, included antibodies specific for murine: Siglec F (E50-2440, BD Pharmingen), CD11b (M1/70), CD64 (X54-5/7.1), CD45 (104), CD3 (17A2, eBiosciences), CD19 (1D3, eBiosciences), CD11c (N418), I-A/I-E (M5/114.15.2), Ly6G (1A8), and Ly6C (HK1.4) (reagents from Biolegend unless otherwise noted). MHC class II tetramers ESAT-6 (I-A(b) 4–17, sequence: QQWNFAGIEAAASA) and MHC class I tetramers TB10.4 (H-2K(b) 4-11, sequence: IMYNYPAM) were obtained from the National Institutes of Health Tetramer Core Facility. Cell sorting was performed on a FACS Aria (BD Biosciences). Sorted

cells were collected in complete media, spun down, resuspended in TRIzol, and frozen at -80°C overnight prior to RNA isolation. Samples for flow cytometry were fixed in 2% paraformaldehyde solution in PBS and analyzed using a LSRII flow cytometer (BD Biosciences) and FlowJo software (Tree Star, Inc.).

Bulk RNA-sequencing and Analysis

RNA isolation was performed using TRIzol (Invitrogen), two sequential chloroform extractions, Glycoblue carrier (Thermo Fisher), isopropanol precipitation, and washes with 75% ethanol. RNA was quantified with the Bioanalyzer RNA 6000 Pico Kit (Agilent). cDNA libraries were constructed using the SMARTer Stranded Total RNA-Seq Kit (v2) - Pico Input Mammalian (Clontech) following the manufacturer's instructions. Libraries were amplified and then sequenced on an Illumina NextSeq (2 x 76, paired-end (sorted BAL cells) or 2 x 151, paired-end (ex vivo stimulation samples)). Stranded paired-end reads were preprocessed: The first three nucleotides of R2 were removed as described in the SMARTer Stranded Total RNA-Seq Kit - Pico Input Mammalian User Manual (v2: 063017) and read ends consisting of more than 66% of the same nucleotide were removed). The remaining read pairs were aligned to the mouse genome (mm10) + Mtb H37Rv genome using the gsnap aligner[55] (v. 2018-07-04) allowing for novel splicing. Concordantly mapping read pairs (~20 million / sample) that aligned uniquely were assigned to exons using the subRead program and gene definitions from Ensembl Mus_Musculus GRCm38.78 coding and non-coding genes. Genes with low expression were filtered using the "filterByExpr" function in the edgeR package[56]. Differential expression was calculated using the "edgeR" package[56] from bioconductor.org. False discovery rate was computed with the Benjamini-Hochberg algorithm. Hierarchical clusterings were performed in R using 'TSclust' and 'hclust' libraries. Heat map and scatterplot visualizations were generated in R using the 'heatmap.2' and 'ggplot2' libraries, respectively.

466 **Gene Set Enrichment Analysis (GSEA)**

467 Input data for GSEA consisted of lists, ranked by -log(p-value), comparing RNAseq expression
468 measures of target samples and naïve controls including directionality of fold-change. Mouse
469 orthologs of human Hallmark genes were defined using a list provided by Molecular Signatures
470 Database (MSigDB)[57]. GSEA software was used to calculate enrichment of ranked lists in each
471 of the respective hallmark gene lists, as described previously[58]. A nominal p-value for each ES
472 is calculated based on the null distribution of 1,000 random permutations. To correct for multiple
473 hypothesis testing, a normalized enrichment score (NES) is calculated that corrects the ES
474 based on the null distribution. A false-discovery rate (FDR) is calculated for each NES. Leading
475 edge subsets are defined as the genes in a particular gene set that are part of the ranked list at
476 or before the running sum reaches its maximum value.

477

478 **Ingenuity Pathway Analysis (IPA)**

479 IPA (QIAGEN) was used to identify enriched pathways for differentially expressed genes
480 between naïve and Mtb-infected AMs (cut-off values: FDR < 0.01, |FC| > 2). The top 20
481 canonical pathways with enrichment score p-value < 0.05 with greater than 10 gene members
482 are reported.

483

484 **Single cell RNA-sequencing**

485 BAL from 10 mice per condition was pooled for each sample, with two independent replicates
486 per condition. Samples were prepared for methanol fixation following protocol “CG000136 Rev.
487 D” from 10X Genomics[59]. Briefly, samples were filtered with 70 μm filters and red blood cells
488 were lysed with ACK lysis buffer. Samples were resuspended in 1 mL ice-cold DPBS using a
489 wide-bore tip and transferred to a 1.5 mL low-bind Eppendorf tube. Samples were centrifuged at
490 700 × g for 5 minutes at 4°C. Supernatant was carefully removed with a p1000 pipette, and the

cell pellet was washed two more times with DPBS, counted, and resuspended in 200 µL ice-cold DPBS/1 × 10^6 cells. 800 µL of ice-cold methanol was added drop-wise for a final concentration of 80% methanol. Samples were incubated at -20°C for 30 minutes and then stored at -80°C for up to 6 weeks prior to rehydration. For rehydration, frozen samples were equilibrated to 4°C, centrifuged at 1,000 × g for 10 minutes at 4°C, and resuspended in 50 µL of Wash-Resuspension Buffer (0.04% BSA + 1mM DTT + 0.2U/µL Protector RNAase Inhibitor in 3× SSC buffer) to achieve ~1,000 cells/µL (assuming 75% sample loss).

Single cell RNA-sequencing Analysis

Libraries were prepared using the Next GEM Single Cell 3□ Reagent Kits v3.1 (Dual Index) (10X Genomics) following the manufacturer's instructions. Raw sequencing data were aligned to the mouse genome (mm10) and UMI counts determined using the Cell Ranger pipeline (10X Genomics). Data processing, integration, and analysis was performed with Seurat v.3[60]. Droplets containing less than 200 detected genes, more than 4000 detected genes (doublet discrimination), or more than 5% mitochondrial were discarded. Genes expressed by less than 3 cells across all samples were removed. Unbiased annotation of clusters using the Immgen database[61] as a reference was performed with "SingleR" package[62]. Pseudotime analysis was performed using the "SeuratWrappers" and "Monocle3" R packages[63]. Data visualization was performed with the "Seurat", "tidyverse", "cowplot", and "viridis" R packages.

Alveolar Macrophage *Ex Vivo* Stimulation

AMs were isolated by bronchoalveolar lavage and pooled from 5 mice per group. Cells were plated at a density of 5 x 10^4 cells/well (96-well plate) in complete RPMI (RPMI plus FBS (10%, VWR), L-glutamine (2mM, Invitrogen), and Penicillin-Streptomycin (100 U/ml; Invitrogen) and allowed to adhere overnight in a 37°C humidified incubator (5% CO_2). Media with antibiotics and

non-adherent cells were washed out prior to stimulation. AM were stimulated with LPS (LPS from Salmonella Minnesota, List Biologicals, #R595, 10 ng/ml), Pam3Cys (Pam3CSK4, EMC Microcollections, GmbH, 10 ng/ml), or H37Rv (MOI 25:1). H37Rv was prepared by culturing from frozen stock in 7H9 media at 37^{o}C for 48 hours to O.D. of 0.1-0.3. The final concentration was calculated based on strain titer and bacteria was added to macrophages for two hours. Cultures were then washed three times to remove extracellular bacteria. Cell cultures were washed once in PBS after 6 hours to remove dead cells and collected in TRIzol for RNA isolation via chloroform/isopropanol extraction.

Filtering for IFN dependent and independent gene sets

"IFN dependent" and "IFN independent" gene sets were generated from data from Olson et al[36], using the following filters starting from a total of 1,233 genes up-regulated in H37Rv-stimulated WT BMDM with average CPM >1, $\log_2$ fold change > 1 and FDR < 0.01:

"IFN dependent" = H37Rv-stimulated IFNAR$^{-/-}$ BMDM: $\log_2$ fold change < 1 AND H37Rv-stimulated WT vs IFNAR$^{-/-}$: $\log_2$ fold change > 2 = **352 genes**

"IFN independent" = H37Rv-stimulated IFNAR$^{-/-}$ BMDM: $\log_2$ fold change > 1, FDR < 0.01 AND H37Rv-stimulated WT vs IFNAR$^{-/-}$: $\log_2$ fold change < 2 = **339 genes**

qRT-PCR

Quantitative PCR reactions were carried out using TaqMan primer probes (ABI) and TaqMan Fast Universal PCR Master Mix (ThermoFisher) in a CFX384 Touch Real-Time PCR Detection System (BioRad). Data were normalized by the level of EF1a expression in individual samples.

Statistical Analyses

RNA-sequencing was analyzed using the edgeR package from Bioconductor.org and the false discovery rate was computed using the Benjamini-Hochberg algorithm. All other data are

presented as mean ± SEM and analyzed by one-way ANOVA (95% confidence interval) with Tukey post-test (for comparison of multiple conditions). Statistical analysis and graphical representation of data was performed using either GraphPad Prism v6.0 software or R. PCA plots generated using "Prcomp" and "Biplot" packages. Venn diagrams and gene set intersection analysis was performed using Intervene[64]. p-values, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Acknowledgements: We thank the staff at Seattle Children's Research Institute vivarium for animal care, Pamela Troisch and the Next Gen Sequencing core at the Institute for Systems Biology, and members of the Aderem, Urdahl, and Rothchild labs for helpful discussions.

Funding: This work was supported by National Institute of Allergy and Infectious Disease of the National Institute of Health under Awards U19AI135976 (A.A.), R01AI032972 (A.A.), 75N93019C00070-P00006-9999-1 (A.A., K.U., A.C.R.)**,** and R21AI163809 (A.C.R.).

Author contributions: D.M., A.C.R, J.N., A.H.D., K.U., and A.A. designed the experiments. A.C.R., D.M., A.J., T.M. conducted the experiments. A.H.D., A.C.R., M.M. performed computational analyses. A.C.R., A.H.D. wrote the paper.

Competing interests: The authors declare no competing interests.

Data and materials availability: Raw and processed RNA-sequencing data can be accessed from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database under accession number GSE212205 [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE212205]. (*Submission currently private.)*

Figure 1: Prior mycobacteria exposure leads to faster activation and innate cell recruitment following aerosol Mtb challenge. Control, scBCG, and coMtb mice, 8 weeks following exposure, challenged with standard low-dose H37Rv. Lungs collected on day 10, 12,

568 and 14 post-infection. A) AM MHC II MFI. B) Total numbers of MDMs, PMN, DC, Ly6C$^+$ CD11b$^+$
569 and LyC6$^+$ CD11b$^+$ monocytes. C) Total numbers of CD44$^+$ CD4$^+$ T cells, ESAT6-tetramer$^+$
570 CD4$^+$ T cells, CD44$^+$ CD8$^+$ T cells, and TB10.4-tetramer$^+$ CD8$^+$ T cells. Mean +/- SEM, 5 mice
571 per group, representative of 3 independent experiments. One-way ANOVA with Tukey post-test.
572 * p< 0.05, **p< 0.01, ***p < 0.001. B, C) *, **, and *** scBCG or coMtb vs control; +, ++ scBCG
573 vs coMtb.
574
575 **Figure 2: Mycobacterium exposure alters the alveolar macrophage transcriptional**
576 **response to Mtb infection in vivo.** Gene expression of naive and Mtb-infected AMs 24 hours
577 following high-dose mEmerald-H37Rv infection in mice previously exposed with scBCG or
578 coMtb or controls (controls- reported in Rothchild et al, 2019[12]; CMTB- reported in Nemeth et al,
579 2020[15]). A) Principal Component Analysis using DEG (|fold change| > 2, FDR < 0.05) in Mtb-
580 infected AMs. B) Venn Diagram and Intersection plot of overlap in up-regulated DEG between
581 the 3 conditions. C) Gene Set Enrichment Analysis of 50 Hallmark Pathways. D) Heatmap of
582 131 Nrf2-associated DEG at 24 hours in Mtb-infected AM (left), Interferon Stimulated Genes,
583 derived from macrophage response to IFNα (fold change >2, p-value < 0.01) Mostafavi et al,
584 2016[27] (middle), IL6 JAK STAT3 hallmark pathway (right). Compiled from 4 independent
585 experiments per condition for control, 2 independent experiments per condition for scBCG and
586 coMtb.
587
588 **Figure 3: *Mycobacterium* exposure modifies the alveolar macrophage phenotype in the**
589 **airway pre-challenge**. Single-cell RNA-sequencing of BAL samples from control, scBCG, and
590 coMtb mice. A) Compiled scRNAseq data for all BAL samples, highlighted by major clusters,
591 annotated based on closest Immgen sample match. B) Highlighting of the two clusters used for
592 macrophage recluster analysis. C) The 12 clusters generated by the macrophage recluster
593 analysis, separated by condition. D) Expression of major macrophage-specific markers: *Siglecf,*

Mertk, Fcgr1, Lyz2, Itgam (CD11b), and *Itgax* (CD11c). E-I) Relative frequency of each macrophage sub-cluster by condition. Expression level of representative genes that are distinguished by that particular cluster compared to other clusters. J) Pseudotime analysis using Monocle3 with starting node at the largest cluster in control, Cluster 0 (*top*) and at the cluster of proliferating cells, Cluster 4,9 (*bottom*). Data is compiled from two independent experiments with 3 conditions each for a total of 6 scRNA-seq BAL samples. One-way ANOVA with Tukey post-test, * $p < 0.05$.

Figure 4: Mycobacterium exposure modifies airway T cell and dendritic cell profiles. Single-cell RNA-sequencing of BAL samples from control, scBCG, and coMtb mice. A) Compiled scRNAseq data for all BAL samples, with major T cell and dendritic cell clusters highlighted. B) Relative frequency of T cells, DC #1, and DC #2 clusters for each condition. C-F) T cell subcluster analysis. C) T cell subclusters compiled (left) and split by condition (right). Annotations below were made following Immgen profile matches and manual marker inspection. D) Relative frequency of Clusters 0-4 for each condition. E) UMAP gene expression plot for general T cell markers. F) UMAP gene expression plot cluster-specific markers split by condition. G-I) Dendritic cell subcluster analysis. G) Dendritic cell subcluster, colored by each of 3 different clusters, compiled (top) and split by conditions (bottom). H) Relative frequency of Clusters 0-2 for each condition. I) Violin plots for cluster-specific markers and genes of interest. Data is compiled from two independent experiments with 3 conditions each for a total of 6 scRNA-seq BAL samples.

Figure 5: Alveolar macrophage remodeling following *mycobacterium* exposure licenses an interferon response upon re-stimulation *ex vivo*. AMs were stimulated for 6 hours with Pam3Cys (10 ng/ml), LPS (10 ng/ml), and H37Rv (MOI 10:1). A) Scatterplots for $\log_2$ fold change for stimulated versus unstimulated AMs for each background (control, scBCG, coMtb).

Differentially expressed genes (DEG) are highlighted for one or both conditions (|Fold change| > 2, FDR < 0.05 for Pam3Cys and LPS; |Fold change| > 2, FDR < 0.2 for H37Rv). B) Gene Set Enrichment Analysis results for 50 HALLMARK pathways. Pathways shown have NES > 1.5 and FDR < 0.05 for at least one of the conditions. C) Scatterplots for log2 fold change for stimulated versus unstimulated AMs for each background. Genes highlighted are Interferon Stimulated Genes. D) Heatmaps depicting log_2 fold change for AM (control, scBCG, coMtb) following 6 hour H37Rv stimulation for IFN-dependent genes (352 total) and IFN-independent genes (339 total) based on WT vs $IFNAR^{-/-}$ BMDM bulk RNA-seq dataset (Olson et al, 2021).

Figure S1 (related to Figure 2): Flow cytometry gating schemes. Gating strategies for myeloid (A) and T cell (B) analysis.

Figure S2 (related to Figure 1): Mycobacterium exposure provides protection against standard dose H37Rv challenge. A) Lung, spleen, and lung-draining lymph node (LN) CFU in control mice at deposition, day 10, 12, 14, and 28. B-E) Summary plots of CFU change (log) in lung, spleen, and LN following low-dose infection with H37Rv at day 10 (B), day 12 (C), day 14 (D), and day 28 (E). *p < 0.05, **p < 0.01, ***p < 0.001. One-way ANOVA with Tukey post-test. Data compiled from 2-3 independent experiments per condition, with 5 mice per group for each experiment.

Figure S3 (related to Figure 2): Top 20 Canonical Pathways by Ingenuity Pathway Analysis for up-regulated genes by Mtb-infected alveolar macrophages. IPA analysis for Mtb-infected AMs from control, scBCG, and coMtb mice 24 hours following high dose mEmerald-H37Rv infection. Data representative of 3 independent experiments per condition.

Figure S4 (related to Figure 3): Transcriptional changes to naive alveolar macrophages following mycobacterium exposure by bulk RNA-sequencing. Volcano plots depicting changes in baseline gene expression of naive AMs from scBCG (A) and coMtb (B) mice compared to naive AMs from control mice. Significantly changed genes (FDR < 0.05, |FC| > 2) highlighted and labeled. Compiled from 2 independent experiments for each condition.

Figure S5 (related to Figure 3): Flow analysis of BAL samples prepared for 10X single-cell RNA-sequencing. Percentage of each population (AM, lymphocytes, eosinophils, MDM, other $CD45^+$) out of $CD45^+$ ZV^-. AM = Siglec F^+ $CD64^+$, Eosinophils = Siglec F^+ $CD64^-$, lymphocytes = $CD3/CD19^+$, MDM = Siglec F^- $CD64^+$, other $CD45^+$ = $CD3^-$ $CD19^-$ Siglec F^- $CD64^-$. Note: One of the two coMtb samples analyzed by flow cytometry did not have an accompanying 10X sample. The second coMtb 10X sample was processed separately without flow analysis.

Figure S6 (related to Figure 3): **Top 10 genes differentially expressed for each of 11 macrophage sub-clusters** Heatmap of genes that are most differentially expressed for each of 11 clusters with all other clusters. Genes filtered with log fold change threshold of > 0.25 and minimum percentage expression of 25% of cells. All genes but one (Gsto1) had an adjusted p-value of < 1.0x10-5 . *Five genes (Fabp4, Fabp5, Stmn1, Mki67, Cbr2) met this criterion for more than one cluster, grouped with the more abundant cluster. Data is compiled from two independent experiments, 3 conditions each, for a total of 6 scRNA-seq BAL samples.

Figure S7 (related to Figure 4): UMAP gene expression plots for genes associated with macrophage sub-cluster 3 and found in AM_2 (Pisu et al). Genes associated with mitochondrial oxidative phosphorylation (*mt-Co1, mt-Cytb, mt-Nd2*), chromatin remodeling (*Ankrd11, Baz1*), macrophage-associated transcription factors (*Cebpb, Zeb2, Bhlhe40, Hif1a*), CARD9 signaling (*Malt1, Bcl10, Prkcd*), hemaglobin metabolism (*Hba-a1, Hba-a2, Hbb-bs*).

Data is compiled from two independent experiments with 4 conditions each for a total of 8 scRNA-seq BAL samples.

Figure S8 (related to Figure 4): UMAP gene expression plots of cluster and lineage marker genes of interest for T cell subclusters. Data is compiled from two independent experiments with 4 conditions each for a total of 8 scRNA-seq BAL samples.

Figure S9 (related to Figure 5): Volcano plots of alveolar macrophage *ex vivo* stimulations. AMs were stimulated for 6 hours with Pam3Cys (10 ng/ml), LPS (10 ng/ml), or H37Rv (MOI 10:1). Volcano plots depict fold change (log2) and P-value (-log10) for each stimulation condition for each of the three groups (control scBCG, coMtb) compared to the respective unstimulated control. DEG (p-value < 0.001; |fold change| > 2) highlighted and labeled, space permitting. Compiled from 3 independent experiments.

Figure S10 (related to Figure 5): Cell-intrinsic changes in alveolar macrophage response is retained 23 weeks following vaccination. Gene expression of *Mx1, Cxcl10, Il1b, Cxcl2, Irf7,* and *Il6* as measured by qPCR in AMs isolated by BAL from mice 8 and 23 weeks following scBCG vaccination and from age-matched controls, with and without LPS (10 ng/ml) stimulation. Data is representative of technical AM duplicates from a single experiment

Table S1: RNA-Sequencing data for alveolar macrophages 24 hours following high dose H37Rv-mEmerald challenge from scBCG mice

Table S2: Top differentially expressed genes for individual clusters for macrophage, T cell, and dendritic cell sub-cluster analysis

697 **Table S3: RNA-Sequencing data for *ex vivo* stimulated alveolar macrophages**

References

1. World Health Organization. Global tuberculosis report 2021. Geneva. Licence: CC BY-NC-SA 3.0 IGO. (2021).

2. Dheda, K. *et al.* The intersecting pandemics of tuberculosis and COVID-19: population-level and patient-level impact, clinical presentation, and corrective interventions. *Lancet Respir Med* (2022).

3. Netea, M.G. *et al.* Defining trained immunity and its role in health and disease. *Nat Rev Immunol* **20**, 375-388 (2020).

4. Sherwood, E.R. *et al.* Innate Immune Memory and the Host Response to Infection. *J Immunol* **208**, 785-792 (2022).

5. Khader, S.A. *et al.* Targeting innate immunity for tuberculosis vaccination. *J Clin Invest* **129**, 3482-3491 (2019).

6. Khan, N. *et al.* M. tuberculosis Reprograms Hematopoietic Stem Cells to Limit Myelopoiesis and Impair Trained Immunity. *Cell* **183**, 752-770 e722 (2020).

7. Kaufmann, E. *et al.* BCG Educates Hematopoietic Stem Cells to Generate Protective Innate Immunity against Tuberculosis. *Cell* **172**, 176-190 e119 (2018).

8. Hoyer, F.F. *et al.* Tissue-Specific Macrophage Responses to Remote Injury Impact the Outcome of Subsequent Local Immune Challenge. *Immunity* **51**, 899-914 e897 (2019).

9. Aegerter, H. *et al.* Influenza-induced monocyte-derived alveolar macrophages confer prolonged antibacterial protection. *Nat Immunol* **21**, 145-157 (2020).

10. Yao, Y. *et al.* Induction of Autonomous Memory Alveolar Macrophages Requires T Cell Help and Is Critical to Trained Immunity. *Cell* **175**, 1634-1650 e1617 (2018).

11. Zhu, B. *et al.* Uncoupling of macrophage inflammation from self-renewal modulates host recovery from respiratory viral infection. *Immunity* **54**, 1200-1218 e1209 (2021).

12. Rothchild, A.C. *et al.* Alveolar macrophages generate a noncanonical NRF2-driven transcriptional response to Mycobacterium tuberculosis in vivo. *Sci Immunol* **4** (2019).

13. Cohen, S.B. *et al.* Alveolar Macrophages Provide an Early Mycobacterium tuberculosis Niche and Initiate Dissemination. *Cell Host Microbe* **24**, 439-446 e434 (2018).

14. Kupz, A., Zedler, U., Staber, M. & Kaufmann, S.H. A Mouse Model of Latent Tuberculosis Infection to Study Intervention Strategies to Prevent Reactivation. *PLoS One* **11**, e0158849 (2016).

15. Nemeth, J. *et al.* Contained Mycobacterium tuberculosis infection induces concomitant and heterologous protection. *PLoS Pathog* **16**, e1008655 (2020).

16. Mangtani, P. *et al.* Protection by BCG vaccine against tuberculosis: a systematic review of randomized controlled trials. *Clin Infect Dis* **58**, 470-480 (2014).

17. Trunz, B.B., Fine, P. & Dye, C. Effect of BCG vaccination on childhood tuberculous meningitis and miliary tuberculosis worldwide: a meta-analysis and assessment of cost-effectiveness. *Lancet* **367**, 1173-1180 (2006).

18. Delahaye, J.L. *et al.* Cutting Edge: Bacillus Calmette-Guerin-Induced T Cells Shape Mycobacterium tuberculosis Infection before Reducing the Bacterial Burden. *J Immunol* **203**, 807-812 (2019).

19. Das, S. *et al.* Lung Epithelial Signaling Mediates Early Vaccine-Induced CD4(+) T Cell Activation and Mycobacterium tuberculosis Control. *mBio* **12**, e0146821 (2021).

20. Mata, E. *et al.* Pulmonary BCG induces lung-resident macrophage activation and confers long-term protection against tuberculosis. *Sci Immunol* **6**, eabc2934 (2021).

21. Arts, R.J.W. *et al.* BCG Vaccination Protects against Experimental Viral Infection in Humans through the Induction of Cytokines Associated with Trained Immunity. *Cell Host Microbe* **23**, 89-100 e105 (2018).

22. Kleinnijenhuis, J. *et al.* BCG-induced trained immunity in NK cells: Role for non-specific protection to infection. *Clin Immunol* **155**, 213-219 (2014).

23. Koeken, V. *et al.* The effect of BCG vaccination on alveolar macrophages obtained from induced sputum from healthy volunteers. *Cytokine* **133**, 155135 (2020).

24. Soto, J.A. *et al.* BCG vaccination induces cross-protective immunity against pathogenic microorganisms. *Trends Immunol* **43**, 322-335 (2022).

25. Wolf, A.J. *et al.* Mycobacterium tuberculosis infects dendritic cells with high frequency and impairs their function in vivo. *J Immunol* **179**, 2509-2519 (2007).

26. Mollenkopf, H.J., Kursar, M. & Kaufmann, S.H. Immune response to postprimary tuberculosis in mice: Mycobacterium tuberculosis and Miycobacterium bovis bacille Calmette-Guerin induce equal protection. *J Infect Dis* **190**, 588-597 (2004).

27. Mostafavi, S. *et al.* Parsing the Interferon Transcriptional Network and Its Disease Associations. *Cell* **164**, 564-578 (2016).

28. Pisu, D. *et al.* Single cell analysis of M. tuberculosis phenotype and macrophage lineages in the infected lung. *J Exp Med* **218** (2021).

29. Travaglini, K.J. *et al.* A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* **587**, 619-625 (2020).

30. Scott, C.L. *et al.* The Transcription Factor ZEB2 Is Required to Maintain the Tissue-Specific Identities of Macrophages. *Immunity* **49**, 312-325 e315 (2018).

31. Cain, D.W. *et al.* Identification of a tissue-specific, C/EBPbeta-dependent pathway of differentiation for murine peritoneal macrophages. *J Immunol* **191**, 4665-4675 (2013).

32. Ciofani, M. *et al.* A validated regulatory network for Th17 cell specification. *Cell* **151**, 289-303 (2012).

33. Edwards, S.C. *et al.* Single-cell analysis uncovers 1 differential regulation of lung γδ T cell subsets by the co-inhibitory molecules, PD-1 and TIM-3. *bioRxiv* **2021.07.04.451035;** (2021).

34. Guilliams, M. *et al.* Alveolar macrophages develop from fetal monocytes that differentiate into long-lived cells in the first week of life via GM-CSF. *J Exp Med* **210**, 1977-1992 (2013).

35. Olmos, S., Stukes, S. & Ernst, J.D. Ectopic activation of Mycobacterium tuberculosis-specific CD4+ T cells in lungs of CCR7-/- mice. *J Immunol* **184**, 895-901 (2010).

36. Olson, G.S. *et al.* Type I interferon decreases macrophage energy metabolism during mycobacterial infection. *Cell Rep* **35**, 109195 (2021).

37. Huang, L., Nazarova, E.V., Tan, S., Liu, Y. & Russell, D.G. Growth of Mycobacterium tuberculosis in vivo segregates with host macrophage metabolism and ontogeny. *J Exp Med* **215**, 1135-1152 (2018).

38. Griffiths, K.L. *et al.* Targeting dendritic cells to accelerate T-cell activation overcomes a bottleneck in tuberculosis vaccine efficacy. *Nat Commun* **7**, 13894 (2016).

39. Lim, P.N., Cervantes, M.M., Pham, L.K. & Rothchild, A.C. Alveolar macrophages: novel therapeutic targets for respiratory diseases. *Expert Rev Mol Med* **23**, e18 (2021).

40. Berry, M.P. *et al.* An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature* **466**, 973-977 (2010).

41. Zak, D.E. *et al.* A blood RNA signature for tuberculosis disease risk: a prospective cohort study. *Lancet* **387**, 2312-2322 (2016).

42. Esaulova, E. *et al.* The immune landscape in tuberculosis reveals populations linked to disease and latency. *Cell Host Microbe* **29**, 165-178 e168 (2021).

43. Antonelli, L.R. *et al.* Intranasal Poly-IC treatment exacerbates tuberculosis in mice through the pulmonary recruitment of a pathogen-permissive monocyte/macrophage population. *J Clin Invest* **120**, 1674-1682 (2010).

44. Redford, P.S. *et al.* Influenza A virus impairs control of Mycobacterium tuberculosis coinfection through a type I interferon receptor-dependent pathway. *J Infect Dis* **209**, 270-274 (2014).

45. Mayer-Barber, K.D. *et al.* Innate and adaptive interferons suppress IL-1alpha and IL-1beta production by distinct pulmonary myeloid subsets during Mycobacterium tuberculosis infection. *Immunity* **35**, 1023-1034 (2011).

46. Goritzka, M. *et al.* Alveolar macrophage-derived type I interferons orchestrate innate immunity to RSV through recruitment of antiviral monocytes. *J Exp Med* **212**, 699-714 (2015).

47. Kumagai, Y. *et al.* Alveolar macrophages are the primary interferon-alpha producer in pulmonary infection with RNA viruses. *Immunity* **27**, 240-252 (2007).

48. Correa-Macedo, W. *et al.* Alveolar macrophages from persons living with HIV show impaired epigenetic response to Mycobacterium tuberculosis. *J Clin Invest* **131** (2021).

49. Verma, A.K., Bansal, S., Bauer, C., Muralidharan, A. & Sun, K. Influenza Infection Induces Alveolar Macrophage Dysfunction and Thereby Enables Noninvasive Streptococcus pneumoniae to Cause Deadly Pneumonia. *J Immunol* **205**, 1601-1607 (2020).

50. D'Agostino, M.R. *et al.* Airway Macrophages Mediate Mucosal Vaccine-Induced Trained Innate Immunity against Mycobacterium tuberculosis in Early Stages of Infection. *J Immunol* **205**, 2750-2762 (2020).

51. Gu, H. *et al.* Vaccination induces rapid protection against bacterial pneumonia via training alveolar macrophage in mice. *Elife* **10** (2021).

52. Grant, R.A. *et al.* Circuits between infected macrophages and T cells in SARS-CoV-2 pneumonia. *Nature* **590**, 635-641 (2021).

53. Zhang, B. *et al.* Single-cell RNA sequencing reveals induction of distinct trained-immunity programs in human monocytes. *J Clin Invest* **132** (2022).

54. Rothchild, A.C., Mai, D., Aderem, A. & Diercks, A.H. Flow Cytometry Analysis and Fluorescence-activated Cell Sorting of Myeloid Cells from Lung and Bronchoalveolar Lavage Samples from Mycobacterium tuberculosis-infected Mice. *Bio Protoc* **10** (2020).

55. Wu, T.D. & Nacu, S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* **26**, 873-881 (2010).

56. Robinson, M.D., McCarthy, D.J. & Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010).

57. Liberzon, A. *et al.* The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* **1**, 417-425 (2015).

58. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550 (2005).

59. Chen, J. *et al.* PBMC fixation and processing for Chromium single-cell RNA sequencing. *J Transl Med* **16**, 198 (2018).

60. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902 e1821 (2019).

61. Heng, T.S., Painter, M.W. & Immunological Genome Project, C. The Immunological Genome Project: networks of gene expression in immune cells. *Nat Immunol* **9**, 1091-1094 (2008).

62. Aran, D. *et al.* Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat Immunol* **20**, 163-172 (2019).

63. Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496-502 (2019).

64. Khan, A. & Mathelier, A. Intervene: a tool for intersection and visualization of multiple gene or genomic region sets. *BMC Bioinformatics* **18**, 287 (2017).

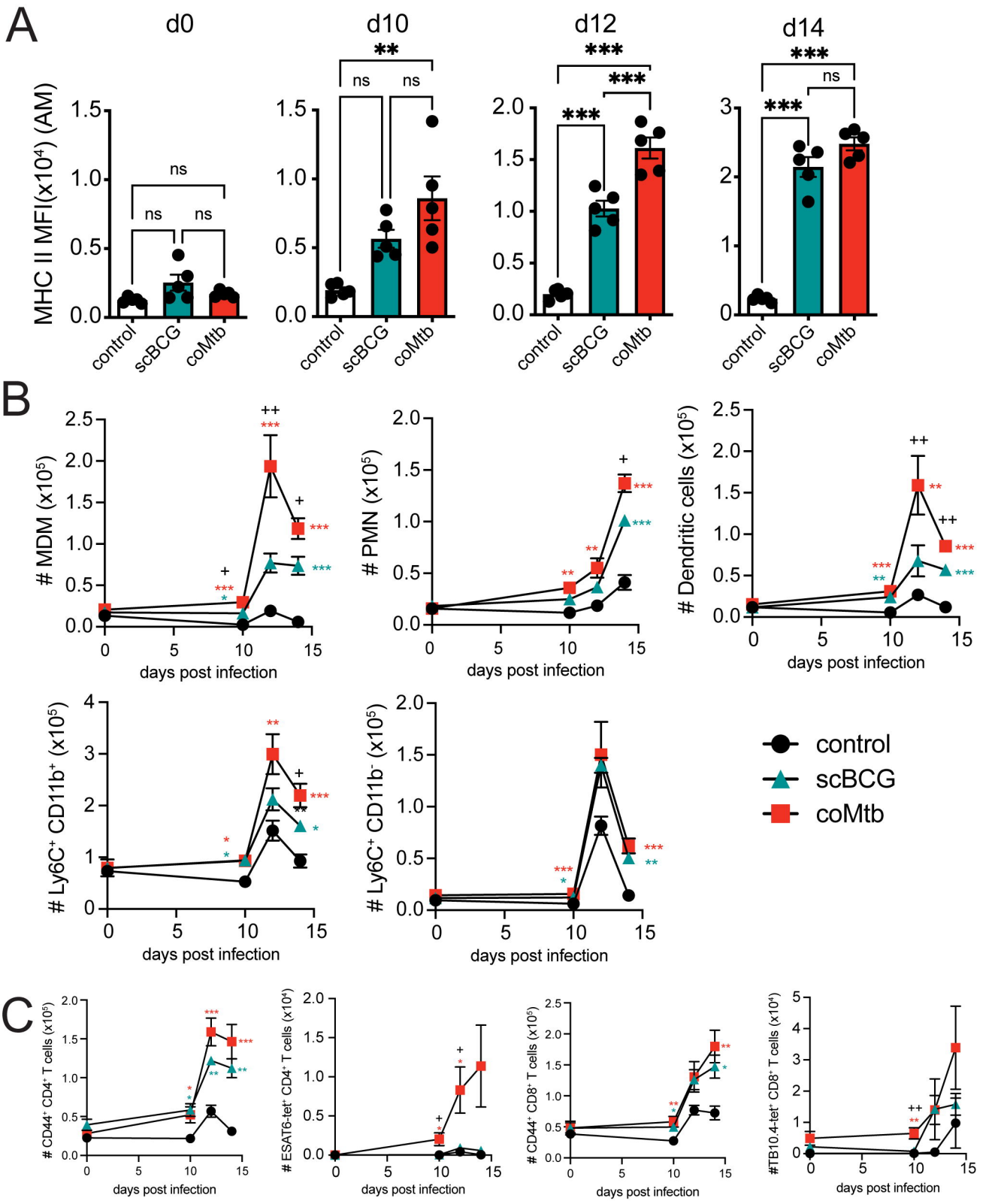


Figure 1: Prior mycobacteria exposure leads to faster activation and innate cell recruitment following aerosol Mtb challenge. Control, scBCG, and coMtb mice, 8 weeks following exposure, challenged with standard low-dose H37Rv. Lungs collected on day 10, 12, and 14 post-infection. A) AM MHC II MFI. B) Total numbers of MDMs, PMN, DC, Ly6C$^+$ CD11b$^+$ and LyC6$^+$ CD11b$^-$ monocytes. C) Total numbers of CD44$^+$ CD4$^+$ T cells, ESAT6-tetramer$^+$ CD4$^+$ T cells, CD44$^+$ CD8$^+$ T cells, and TB10.4-tetramer$^+$ CD8$^+$ T cells. Mean +/- SEM, 5 mice per group, representative of 3 independent experiments. One-way ANOVA with Tukey post-test. * p< 0.05, **p< 0.01, ***p < 0.001. B, C) *, **, and *** scBCG or coMtb vs control; +, ++ scBCG vs coMtb.

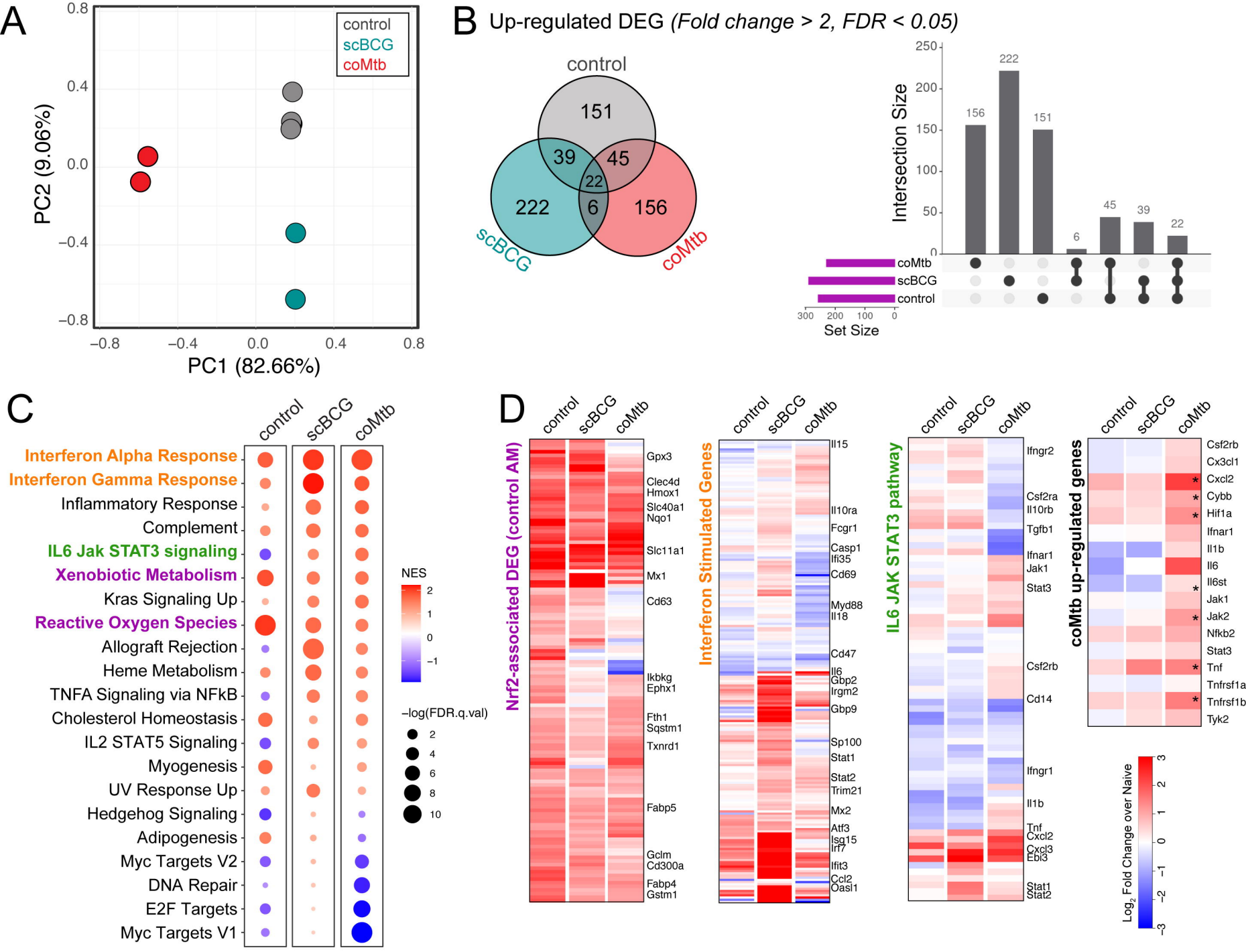


Figure 2: *Mycobacterium* exposure alters the alveolar macrophage transcriptional response to Mtb infection *in vivo*. Gene expression of naive and Mtb-infected AMs 24 hours following high-dose mEmerald-H37Rv infection in mice previously exposed with scBCG or coMtb or controls (*controls- reported in Rothchild et al, 2019; CMTB- reported in Nemeth et al, 2020*). A) Principal Component Analysis using DEG (|fold change| > 2, FDR < 0.05) in Mtb-infected AMs. B) Venn Diagram and Intersection plot of overlap in up-regulated DEG between the 3 conditions. C) Gene Set Enrichment Analysis of 50 Hallmark Pathways. D) Heatmap of 131 Nrf2-associated DEG at 24 hours in Mtb-infected AM (*left*), Interferon Stimulated Genes, derived from macrophage response to IFNα (fold change >2, p-value < 0.01) (Mostafavi et al, 2016) (*middle-left*), IL6 JAK STAT3 hallmark pathway (*middle-right*) and selected coMtb signature genes (*right, *FDR < 0.05, FC>2*). Compiled from 4 independent experiments per condition for control, 2 independent experiments per condition for scBCG and coMtb.

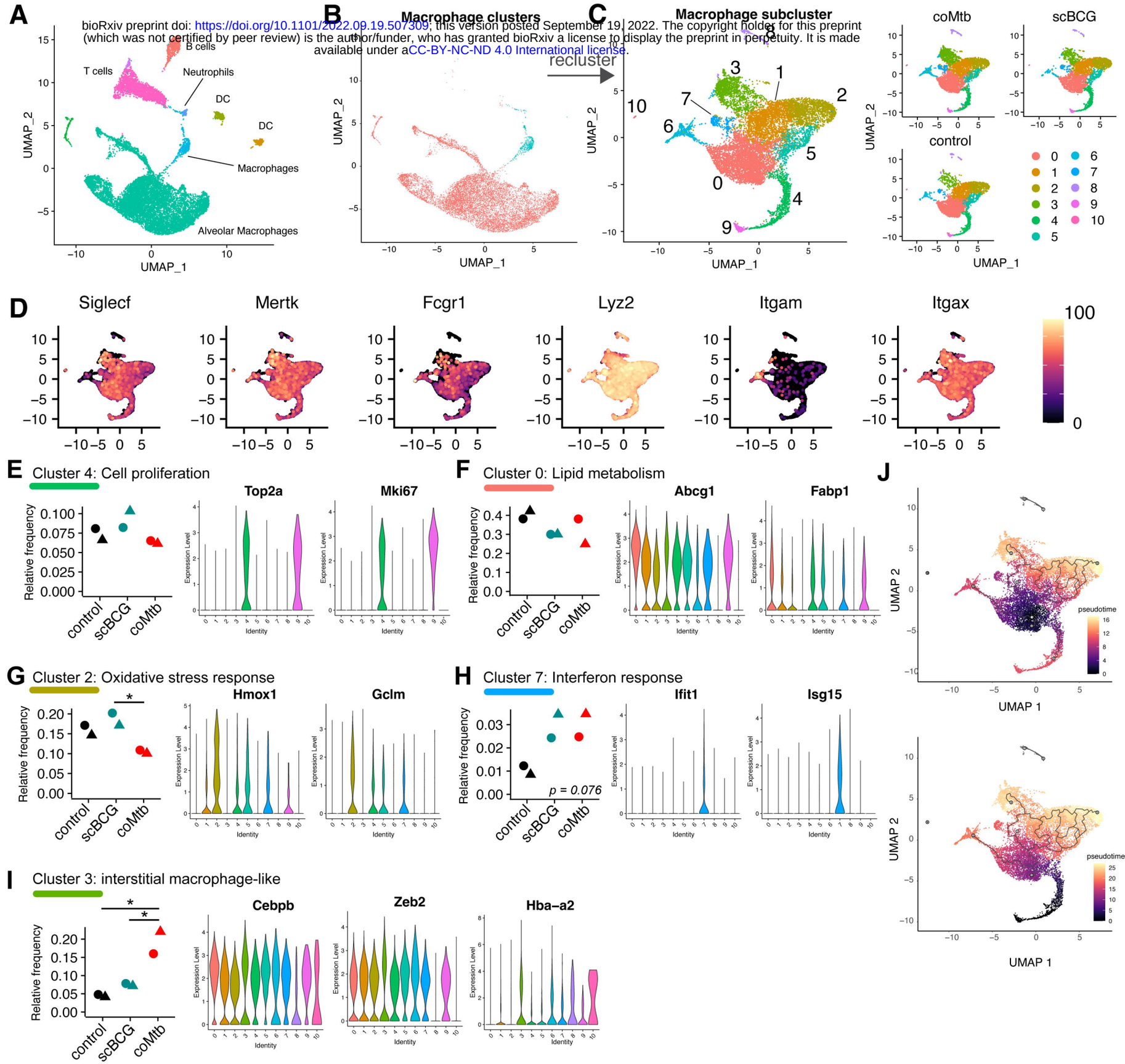


Figure 3: *Mycobacterium* exposure modifies the alveolar macrophage phenotype in the airway pre-challenge. Single-cell RNA-sequencing of BAL samples from control, scBCG, and coMtb mice. A) Compiled scRNAseq data for all BAL samples, highlighted by major clusters, annotated based on closest Immgen sample match. B) Highlighting of the two clusters used for macrophage recluster analysis. C) The 12 clusters generated by the macrophage recluster analysis, separated by condition. D) Expression of major macrophage-specific markers: Siglecf, Mertk, Fcgr1, Lyz2, Itgam (CD11b), and Itgax (CD11c). E-I) Relative frequency of each macrophage sub-cluster by condition. Expression level of representative genes that are distinguished by that particular cluster compared to other clusters. J) Pseudotime analysis using Monocle3 with starting node at the largest cluster in control, Cluster 0 (*top*) and at the cluster of proliferating cells, Cluster 4,9 (*bottom*). Data is compiled from two independent experiments with 3 conditions each for a total of 6 scRNA-seq BAL samples. One-way ANOVA with Tukey post-test, * p< 0.05.

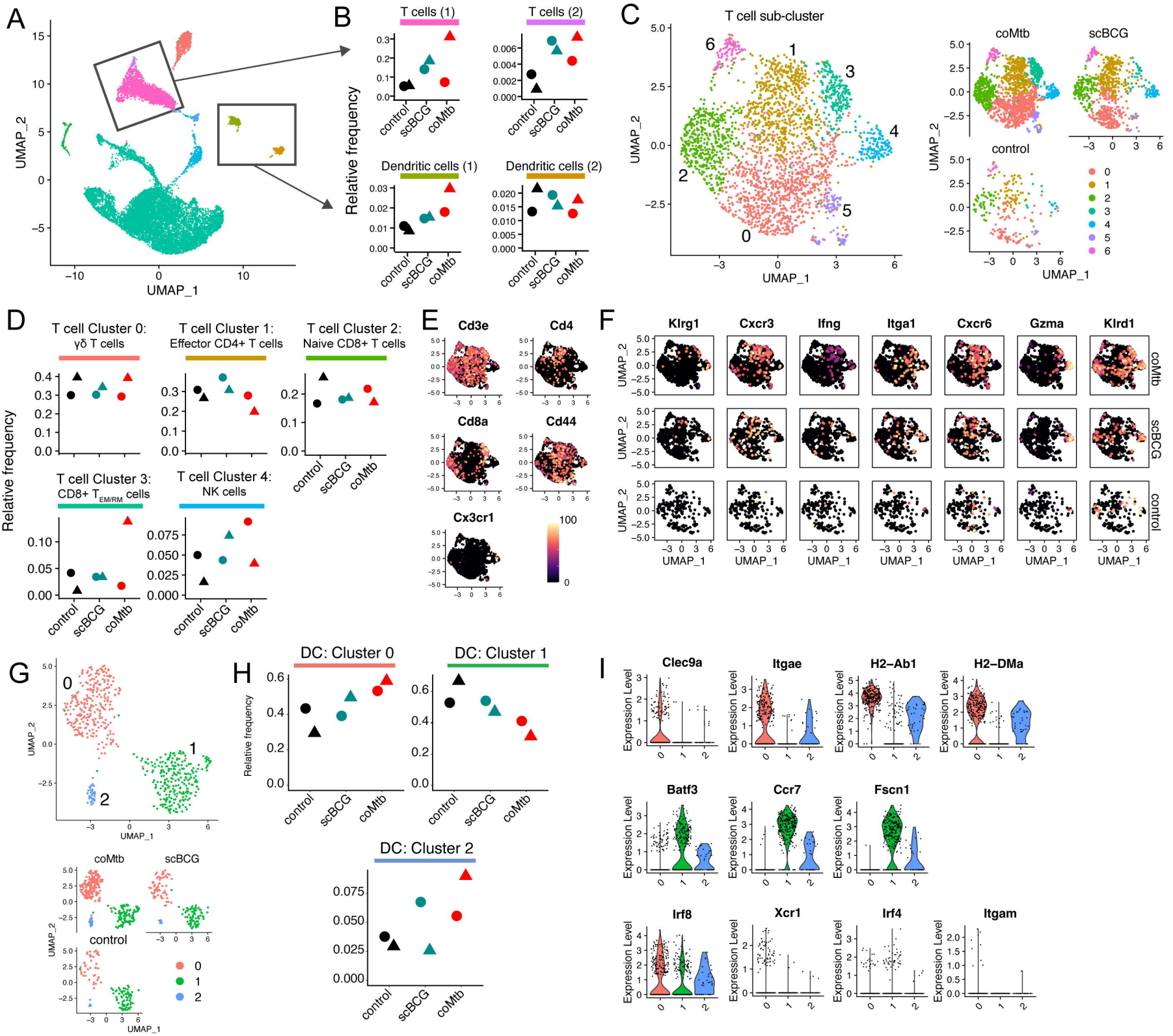


Figure 4: *Mycobacterium* exposure modifies airway T cell and dendritic cell profiles. Single-cell RNA-sequencing of BAL samples from control, scBCG, and coMtb mice. A) Compiled scRNAseq data for all BAL samples, with major T cell and dendritic cell clusters highlighted. B) Relative frequency of T cells, DC #1, and DC #2 clusters for each condition. C-F) T cell subcluster analysis. C) T cell subclusters compiled (*left*) and split by condition (*right*). Annotations below were made following Immgen profile matches and manual marker inspection. D) Relative frequency of Clusters 0-4 for each condition. E) UMAP gene expression plot for general T cell markers. F) UMAP gene expression plot cluster-specific markers split by condition. G-I) Dendritic cell subcluster analysis. G) Dendritic cell subcluster, colored by each of 3 different clusters, compiled (*top*) and split by conditions (*bottom*). H) Relative frequency of Clusters 0-2 for each condition. I) Violin plots for cluster-specific markers and genes of interest. Data is compiled from two independent experiments with 3 conditions each for a total of 6 scRNA-seq BAL samples.

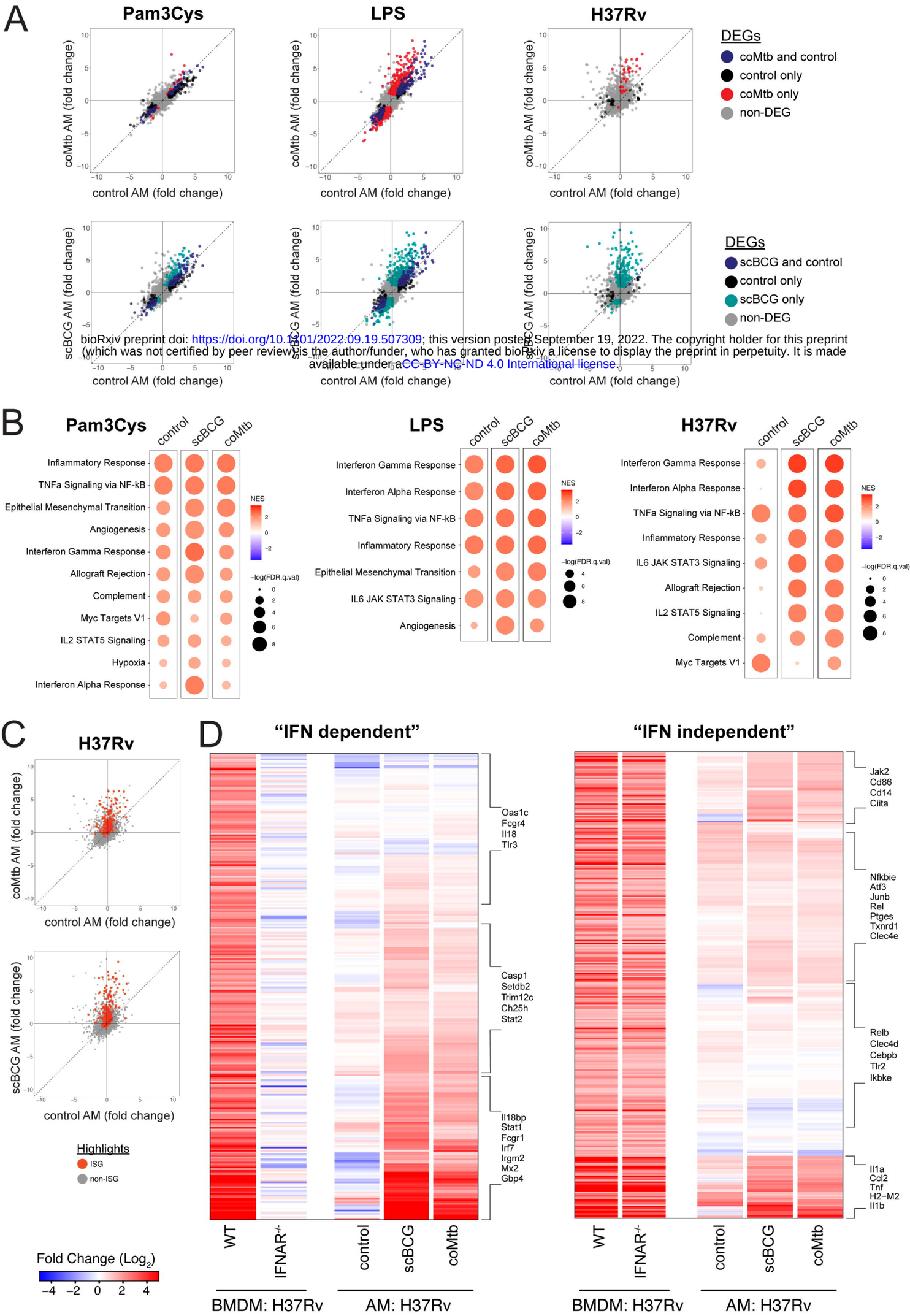
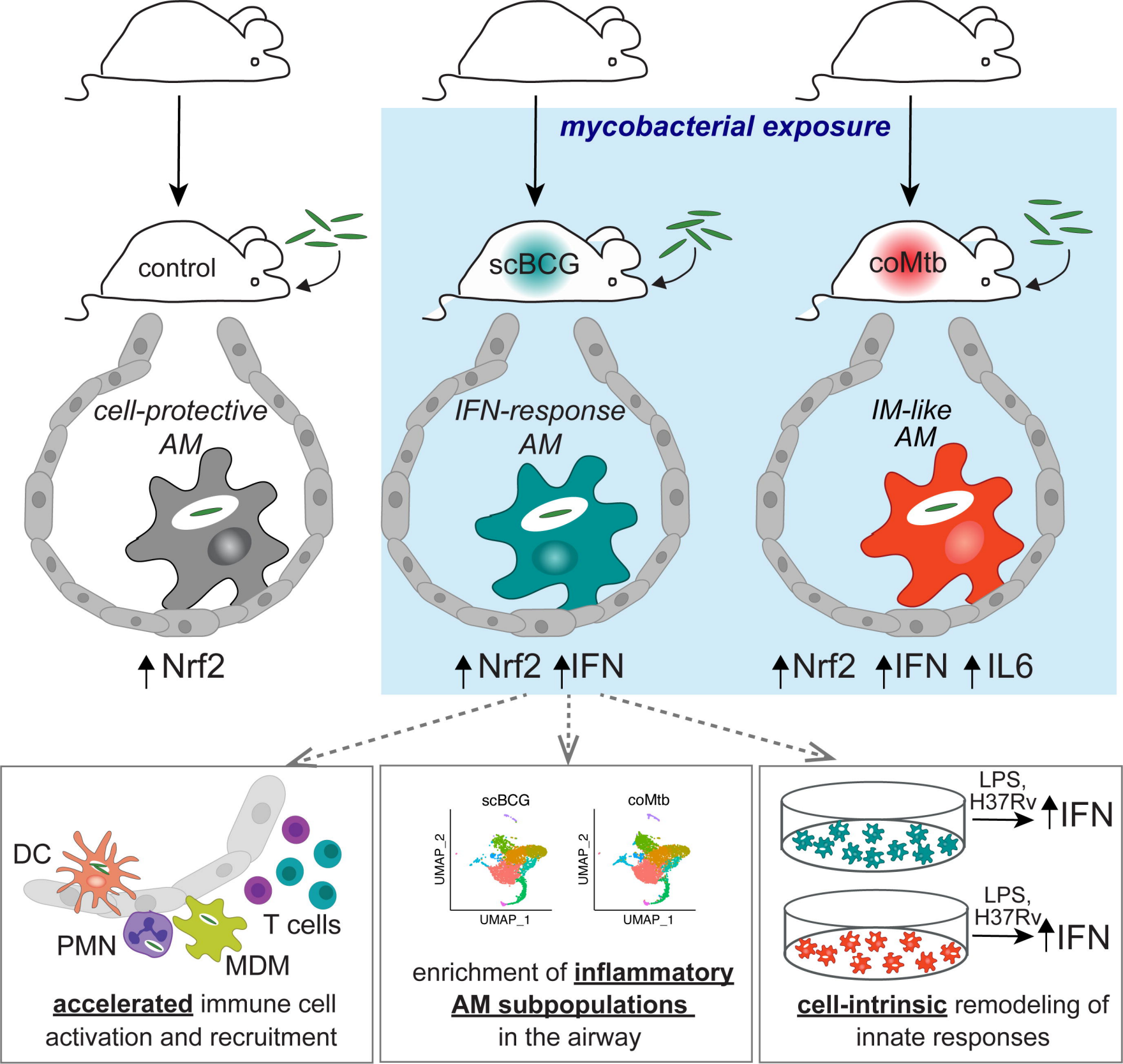


Figure 5: Alveolar macrophage remodeling following *mycobacterium* exposure licenses an interferon response upon re-stimulation *ex vivo*. AMs were stimulated for 6 hours with Pam3Cys (10 ng/ml), LPS (10 ng/ml), and H37Rv (MOI 10:1). A) Scatterplots for $\log_2$ fold change for stimulated versus unstimulated AMs for each background (control, scBCG, coMtb). Differentially expressed genes (DEG) are highlighted for one or both conditions (|Fold change| > 2, FDR < 0.05 for Pam3Cys and LPS; |Fold change| > 2, FDR < 0.2 for H37Rv). B) Gene Set Enrichment Analysis results for 50 HALLMARK pathways. Pathways shown have NES > 1.5 and FDR < 0.05 for at least one of the conditions. C) Scatter-plots for $\log_2$ fold change for stimulated versus unstimulated AMs for each background. Genes highlighted are Interferon Stimulated Genes. D) Heatmaps depicting $\log_2$ fold change for AM (control, scBCG, coMtb) following 6 hour H37Rv stimulation for IFN-dependent genes (352 total) and IFN-independent genes (339 total) based on WT vs IFNAR$^{-/-}$ BMDM bulk RNA-seq dataset (Olson et al, 2021).


mycobacterial exposure
control
scBCG
coMtb
cell-protective AM
IFN-response AM
IM-like AM
↑Nrf2
↑Nrf2 ↑IFN
↑Nrf2 ↑IFN ↑IL6
DC
PMN
MDM
T cells
accelerated immune cell activation and recruitment
scBCG
coMtb
UMAP_2
UMAP_1
enrichment of inflammatory AM subpopulations in the airway
LPS, H37Rv
↑IFN
LPS, H37Rv
↑IFN
cell-intrinsic remodeling of innate responses