

Supplementary Information:

Figure Legends

Supp. Figure 1: Characterizing how ontogeny and intra-lineage diversity affect macrophage responses to Mtb infection using scRNA-seq. **(a)** Visualization of Cell Annotations Using the ImmGen Reference. The plot visualizes the annotations of cells from the scRNA-seq timepoint dataset, leveraging the ImmGen reference dataset to assign cell type labels. The SingleR method¹¹⁵ has been employed to match the gene expression profile of each cell in the scRNA-seq dataset with those from the ImmGen reference, a comprehensive database containing finely annotated immune cell types. The heatmap displays the matching scores for each cell in the scRNA-seq dataset against potential cell types from ImmGen. Each column in the heatmap represents a cell from the dataset, while the rows correspond to the potential cell type labels from ImmGen. The cells are grouped based on their inherent clusters from the scRNA-seq dataset. The color intensity in the heatmap indicates the degree of match between the cell's gene expression profile and the reference profile of the cell type in ImmGen. Yellow shades suggest a stronger match, indicating higher confidence in the assigned cell type label. **(b)** DotPlot illustrating genes involved in monocyte-to-macrophage differentiation processes. Each dot's size corresponds to the percentage of cells within a cluster expressing a given gene, while the color intensity indicates the average expression level, ranging from red (low expression) to blue (high expression). Clusters include different subsets of tissue-resident macrophages and monocyte populations. Source data are provided as a Source Data file.

Supp. Figure 2: Analysis of macrophage responses and pathway regulation in Mtb infection across genetic backgrounds and timepoints. **(a)** Flow cytometry analysis of the Mtb smyc::mCherry/hsp α ::GFP reporter strain at 2 wpi in Rag1^{-/-} and IFN- γ ^{-/-} KO mice. **(b)** CD11b Protein Expression in macrophages, stratified by infection status and timepoint. Data are presented as violin plots with an overlaid box plot. The center line in the box plot represents the median expression level of CD11b, the bounds of the box represent the interquartile range (IQR) from the 25th to the 75th percentile, and the whiskers extend to 1.5 times the IQR from the quartiles. Statistical significance of pairwise comparisons between infection states at each timepoint was assessed using two-sided t-tests. P-values were adjusted for multiple comparisons utilizing the Benjamini-Hochberg method; annotations on the plot reflect these adjusted p-values. **(c-d)** Pathways upregulated at late (left) and early (right) timepoints. Generalized additive models (GAMs) were fitted to assess temporal changes in pathway expression levels. An analysis of variance (ANOVA) was performed on these models to derive p-values, which were subsequently adjusted for multiple comparisons using the Benjamini-Hochberg method. The red line represents the fitted GAM-predicted trend over time for each pathway, with specific macrophage clusters labeled as appropriate. Exact adjusted p-values are provided. Each cell cluster is represented by distinct colored points. Average pathway expression levels are on the y-axis, with inferred time from early to late on the x-axis. **(e)** Mtb mRNA Reads Distribution Across Timepoints. This figure displays the distribution of the recovered Mtb mRNA reads at distinct post-infection timepoints. Each panel represents a timepoint, with labels on top. Cells are colored based on the presence of Mtb reads. Source data are provided as a Source Data file.

Supp. Figure 3: Differential dynamics of alveolar and interstitial macrophage subsets in response to infection over time. **(a)** Flow cytometry analysis of infected AM and IM at 2 and 4 wpi. Infected macrophages were first identified based on expression of Mertk, CD64, and

mCherry. Within this infected population, cells were further differentiated into AM (SiglecF⁺) and IM (SiglecF⁻). The proportion of each subset at 2 and 4 wpi is shown. Data from n=5 independent animals.

Supp. Figure 4: Multimodal analysis of immune responses and bacterial burden in BCG-vaccinated versus unvaccinated mice. **(a)** Flow cytometry analysis of the Mtb mCherry signal in BCG-treated and unvaccinated mice at 2 wpi. BCG-vaccinated mice show a clear reduction in the percentage of mCherry⁺ cells, indicating lower bacterial burden. **(b)** The figure illustrates the distribution of cells containing hsp_x-high and low bacteria across two conditions: BCG-vaccinated and unvaccinated (Control). On the X-axis, the vaccination status is represented, while the Y-axis displays the cell counts. Within each vaccination status, there are bars differentiating between Gfp-high (hsp_x-high) cells, colored in pink, and Gfp-low (hsp_x-low) cells, colored in green. The exact cell counts are labeled above each bar. **(c)** Visualization of BCG-specific scWGCNA-derived Module 2 (463 genes) expression levels mapped onto UMAP embeddings. **(d)** This heatmap showcases the expression of selected pro-inflammatory genes in BCG-vaccinated versus unvaccinated cells. Genes are listed on the Y-axis and cells are grouped by their identity on the X-axis. The color gradient—ranging from dark blue (downregulated) through white (average) to yellow (upregulated)—reflects z-scaled gene expression levels. **(e)** Violin and box plots showing Small GTPase gene expression levels in BCG vs. Control conditions. Within each violin plot, the density of the shading illustrates the distribution of gene expression levels, with internal white boxplots providing median and interquartile ranges. Gene expression differences between the conditions were statistically analyzed using the Wilcoxon rank-sum test. Adjusted p-values, calculated using the False Discovery Rate (FDR) method to control for multiple testing, are displayed within each facet. Effect size measured as Cliff's Delta is also displayed directly on each plot, providing a standardized measure of the magnitude of expression differences. Source data are provided as a Source Data file.

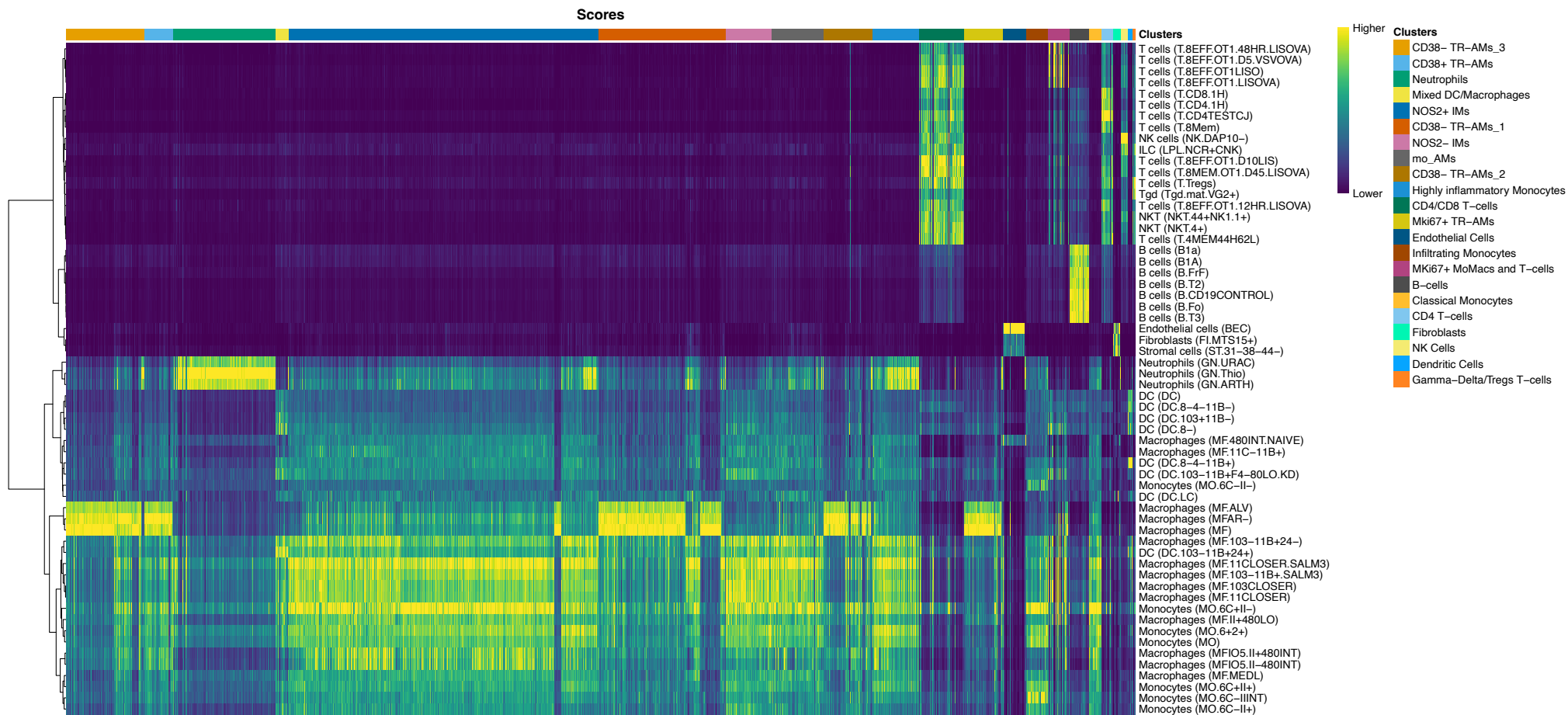
Supp. Figure 5: Multi-panel display illustrating immune profiling in BCG-vaccinated versus unvaccinated mice. **(a)** Horizontal bar chart showing top gene set enrichment results for the scWGCNA "Module 2" in BCG-vaccinated macrophages. Gene sets categories are ordered by significance with p-value logarithms on the X-axis. Pathway enrichment analysis was conducted using the G::Profiler and the two-sided g:SCS method, as described¹¹⁸. **(b)** Flow cytometry analysis of CD38 staining in infected cells from BCG-vaccinated and control mice at 2 weeks post-infection (wpi). **(c)** Flow cytometry analysis of the Mtb mCherry signal at 4 wpi reveals fewer mCherry⁺ cells in BCG-vaccinated mice, indicating lower bacterial burden. **(d)** Visualization of CD64 staining within the scRNA-seq dataset, with a color gradient from light gray (low expression) to dark red (high expression). **(e)** Flow cytometry analysis of the relative proportions of infected macrophages, when gating by Mertk and CD64 at 2wpi. In BCG-vaccinated mice, the vast majority of infected macrophages are AMs when using this gating strategy, because of the loss of expression of CD64 in monocyte-derived cells. **(f)** Flow cytometry analysis of the relative proportions of infected macrophages, when gating only by SiglecF at 2 wpi. In BCG-vaccinated mice, the vast majority of infected cells are of monocyte-origin, in line with the findings from our scRNA-seq dataset. A consistent increase in the percentage of infected AMs in BCG-vaccinated mice compared to the unvaccinated counterparts, underscores their improved ability to restrict Mtb growth. **(g)** Flow cytometry analysis of the relative proportions of infected macrophages, when gating only by SiglecF at 4 wpi. In BCG-vaccinated mice, the majority of infected cells are of monocyte-origin. A consistent increase in the percentage of infected AMs in BCG-vaccinated mice compared to the unvaccinated counterparts, underscores their improved ability to restrict Mtb growth. For

experiments in panels (e-g), a high infection dose of 5×10^3 CFU was used, consistent with protocols from Mata et al. Source data are provided as a Source Data file.

Supp. Figure 6: Mapping of chromatin accessibility and regulatory landscapes in Naive mice using single-cell ATAC-seq. **(a)** Transcription Start Site (TSS) Enrichment Analysis of scATAC-seq Data from Naïve mice. The plot showcases the normalized insertion profile of the single-cell ATAC-seq dataset relative to the TSS. The x-axis represents the distance from the TSS, and the y-axis indicates the normalized insertion profile. A pronounced peak at or near the center is indicative of significant enrichment at transcription start sites, suggesting a high-quality ATAC-seq dataset. **(b)** This UMAP representation highlights the chromatin accessibility patterns of two marker genes, "*Mafb*" and "*Ly6a*", in scATAC-seq data. Each point on the UMAP plot corresponds to a cell, and its color intensity reflects the gene accessibility score for the specified marker genes. Yellow shades indicate higher accessibility, suggesting a potential for higher gene activity. **(c)** Visualization of open chromatin regions near key genes "*Nos2*", "*Cd38*", "*Slc7a11*", "*Ccl5*", "*Ptgs2*," "*Cxcl2*," and "*Il1b*" within the single-cell ATAC-seq dataset. The browser-like display is grouped by distinct scATAC-seq clusters, allowing for a comparative analysis of chromatin accessibility peaks across different cell types. The visualization spans 50,000 base pairs both upstream and downstream of the gene locus. Source data are provided as a Source Data file.

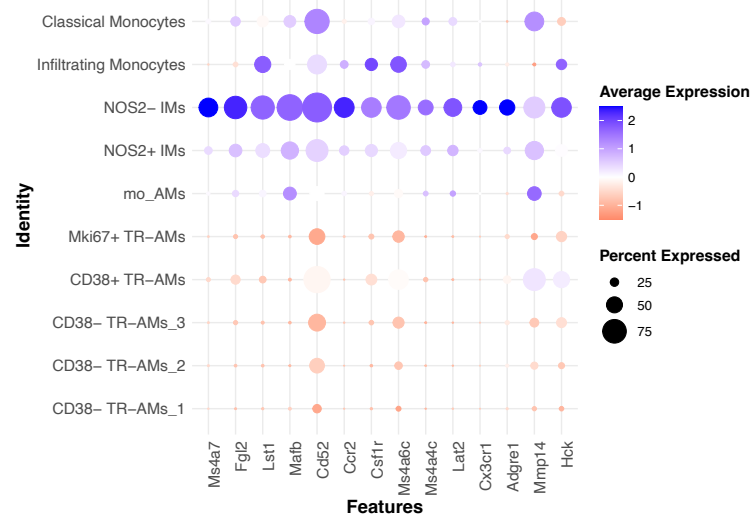
A

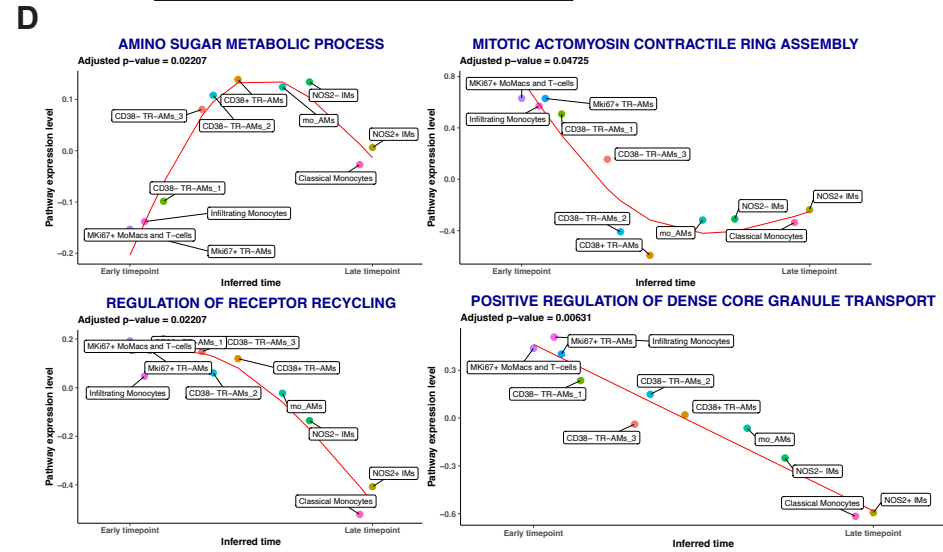
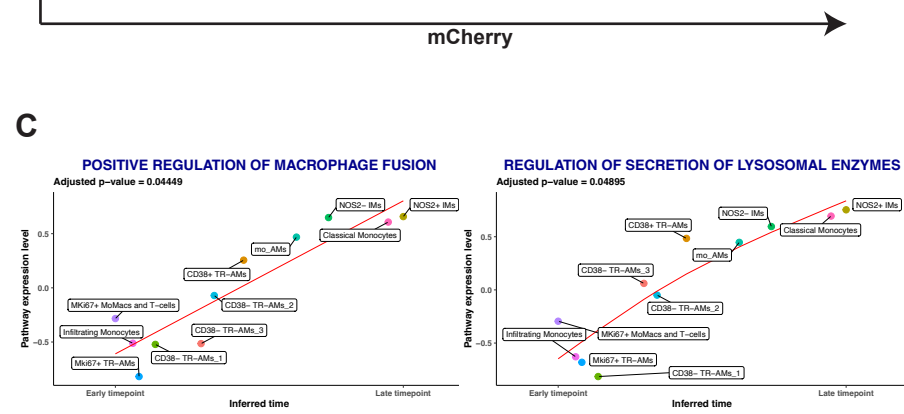
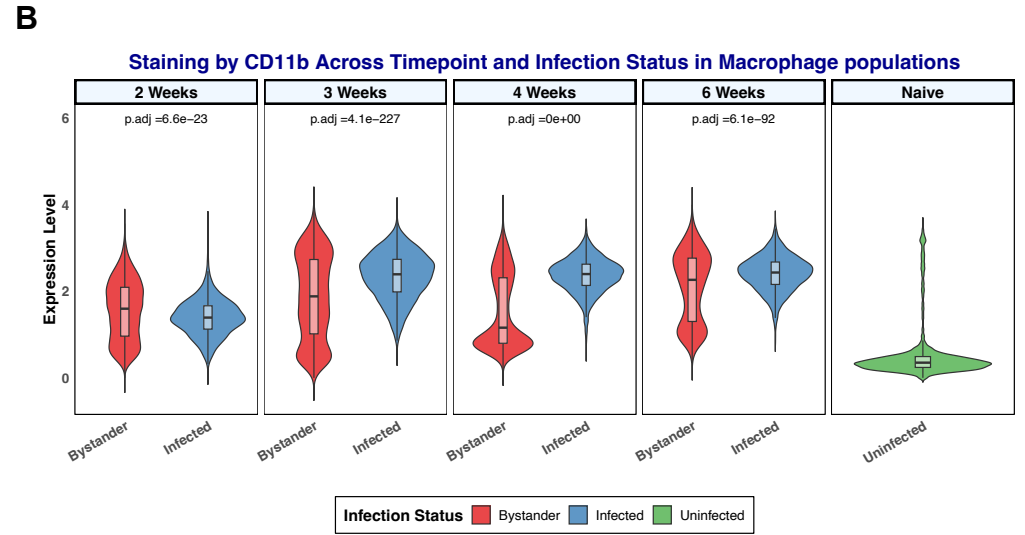
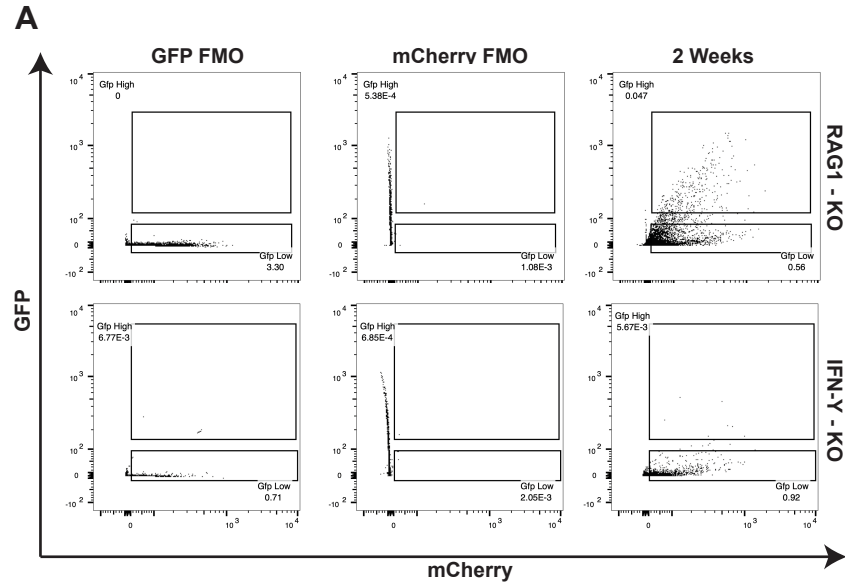
Reference-based Cluster Analysis



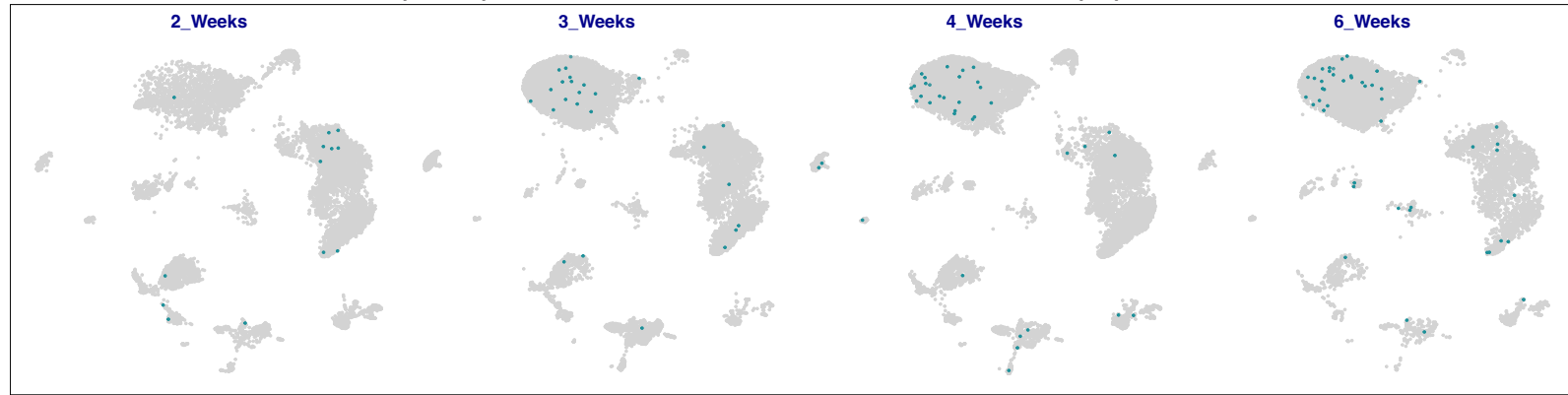
B

Monocyte to macrophage differentiation





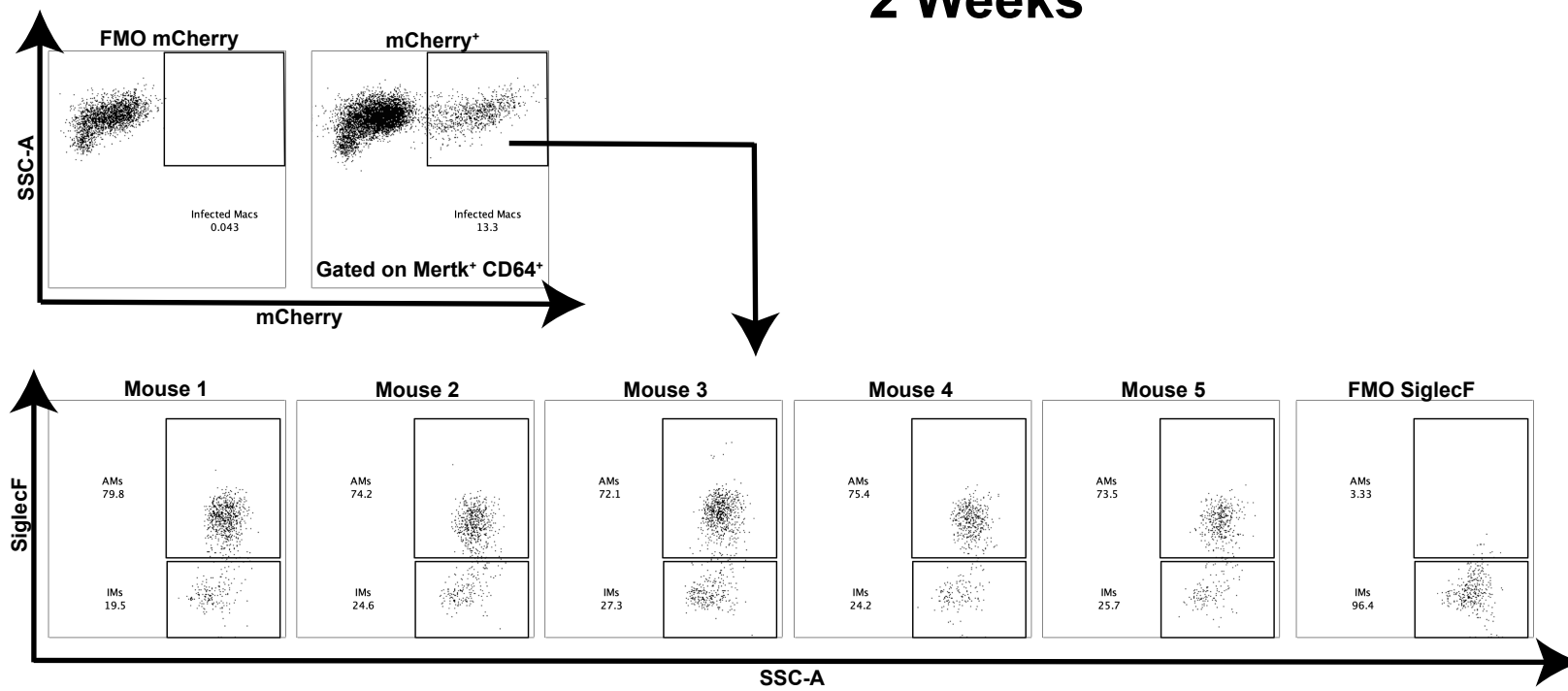
E Temporal profile of recovered Mtb Reads in host cell populations



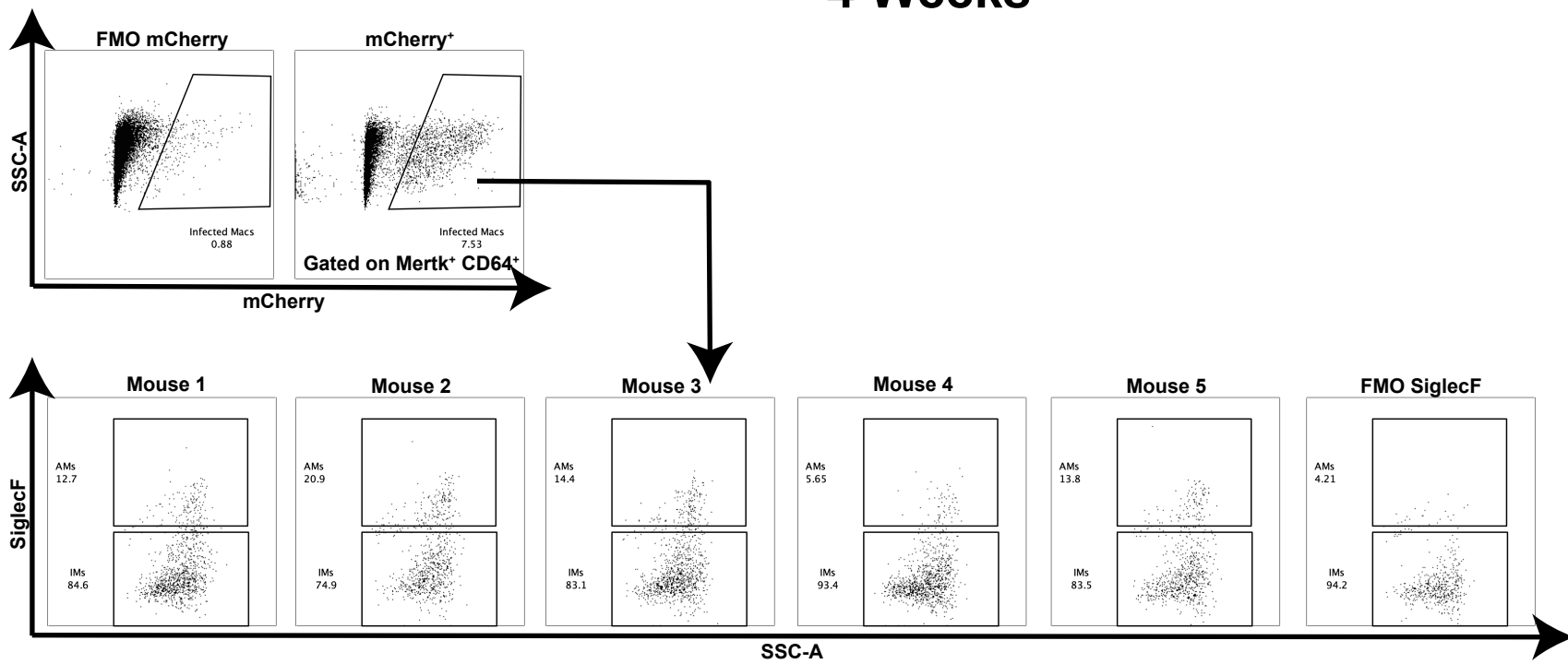
● Host cell associated with Mtb Reads

A

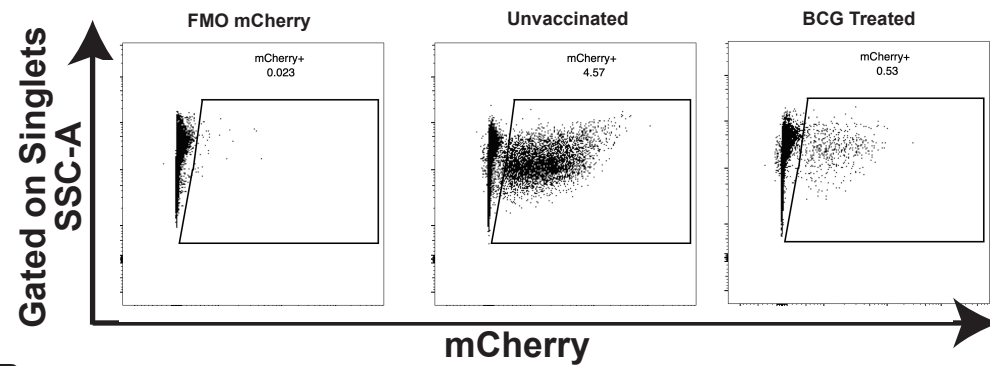
2 Weeks



4 Weeks



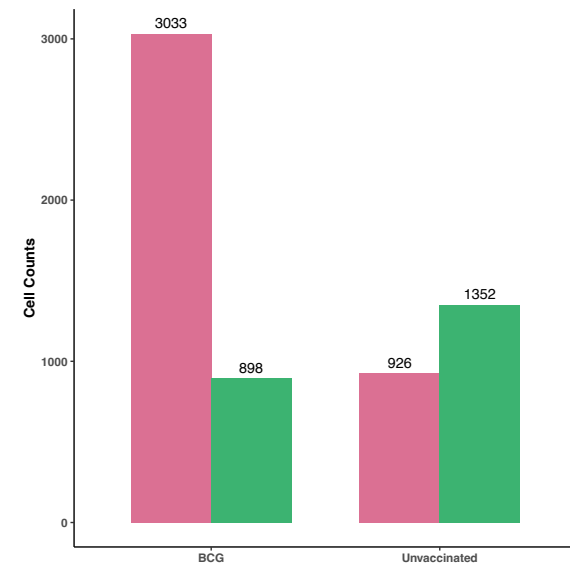
A



B

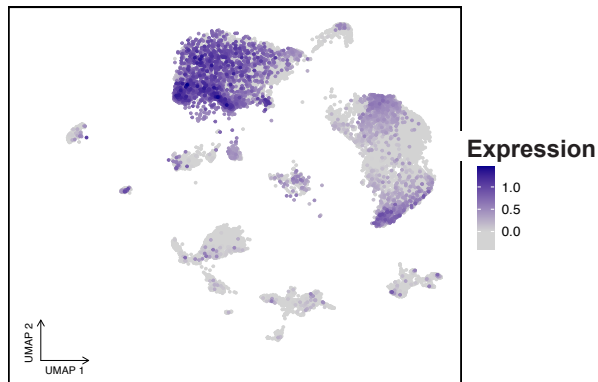
Distribution of hsp α -high and low cells by Vaccination Status

Gfp-high Gfp-low



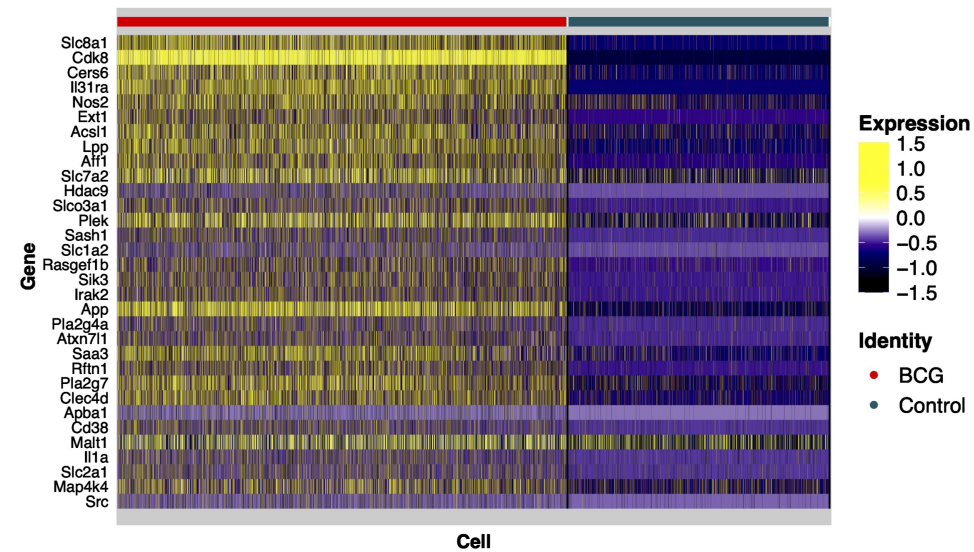
C

Module 2 – 463 Genes



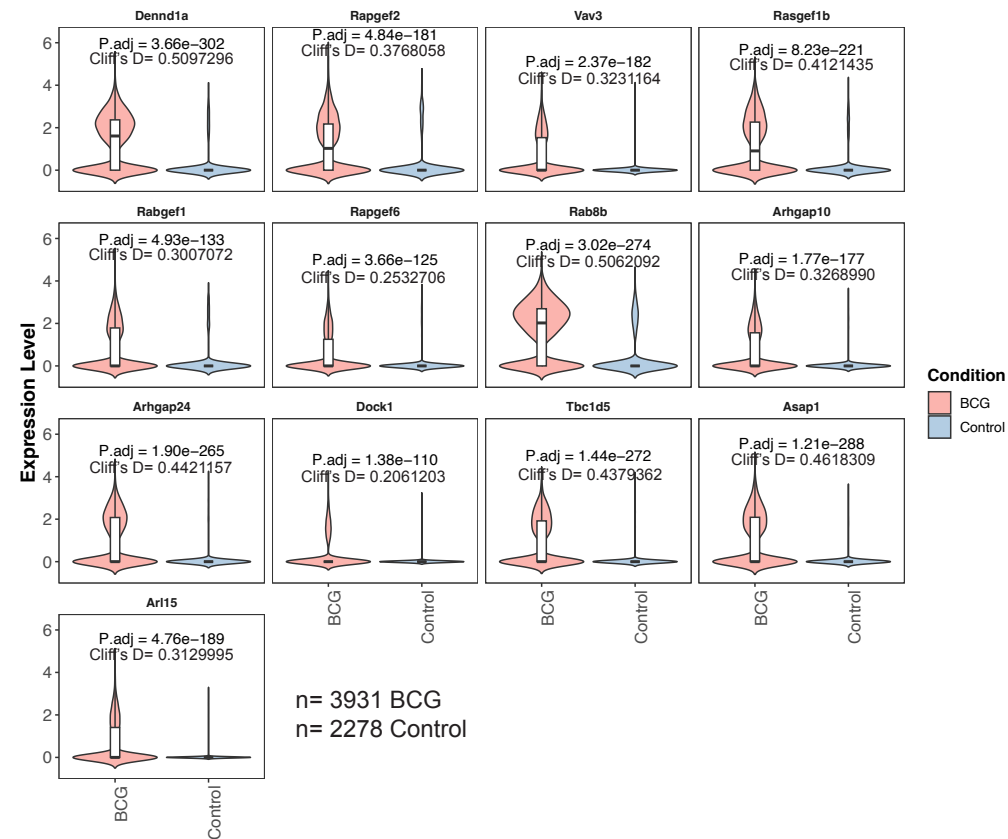
D

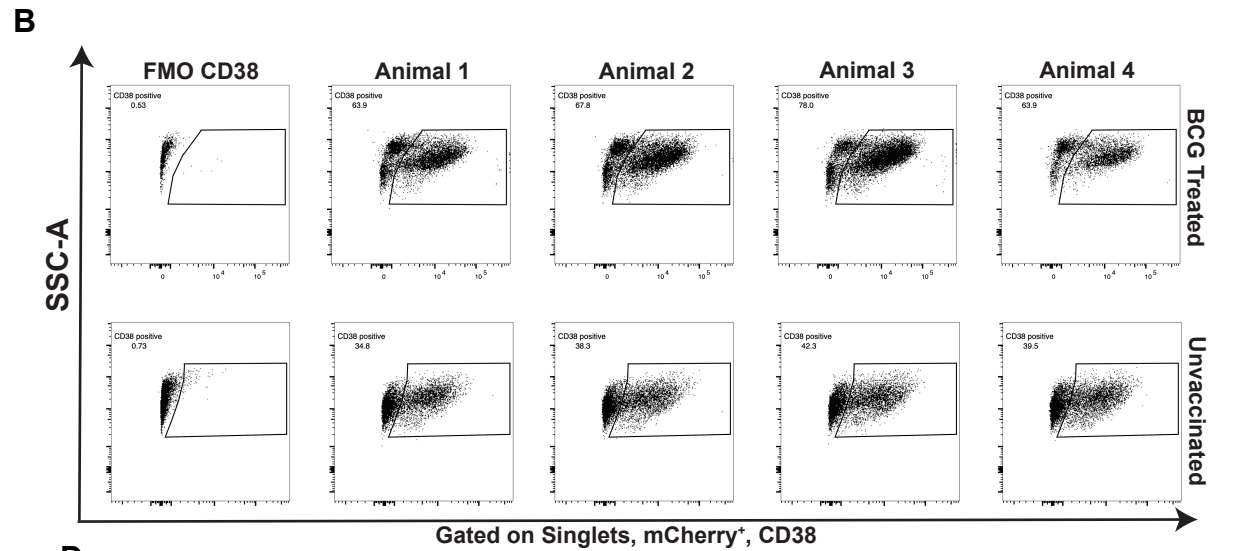
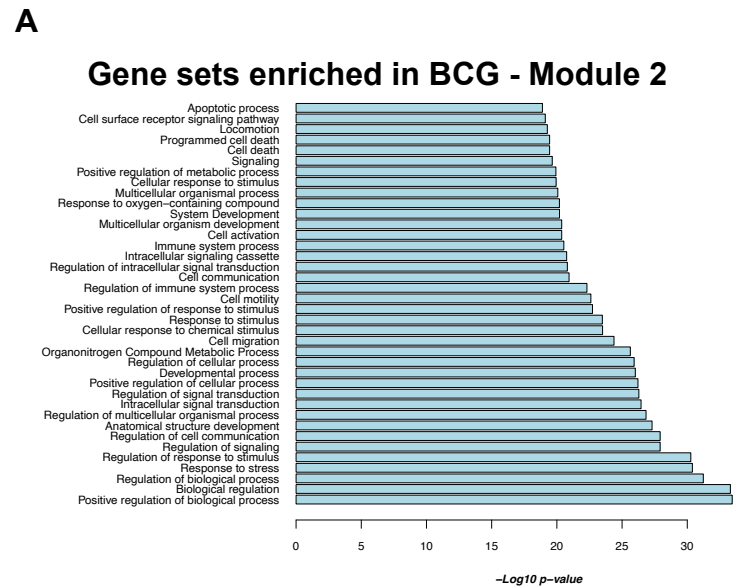
Pro-inflammatory genes between BCG and Unvaccinated



E

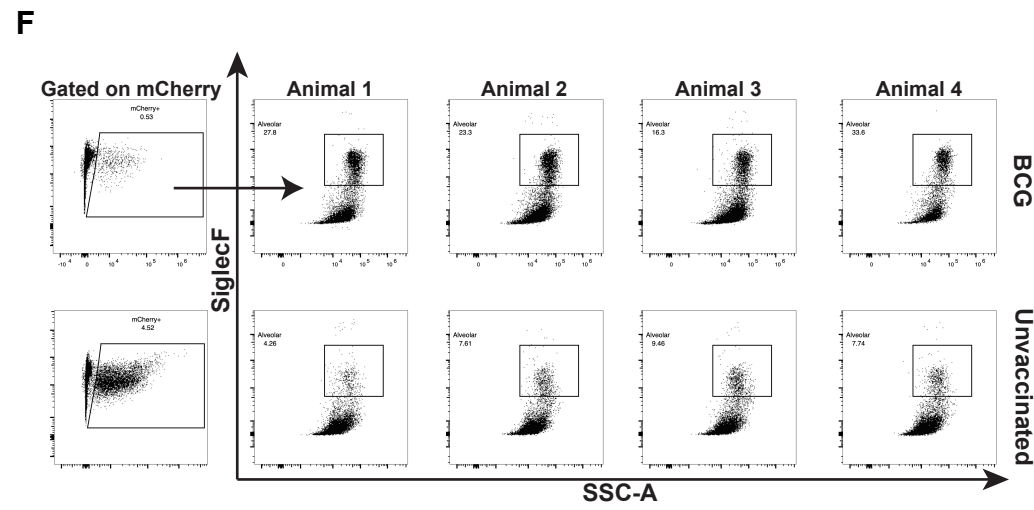
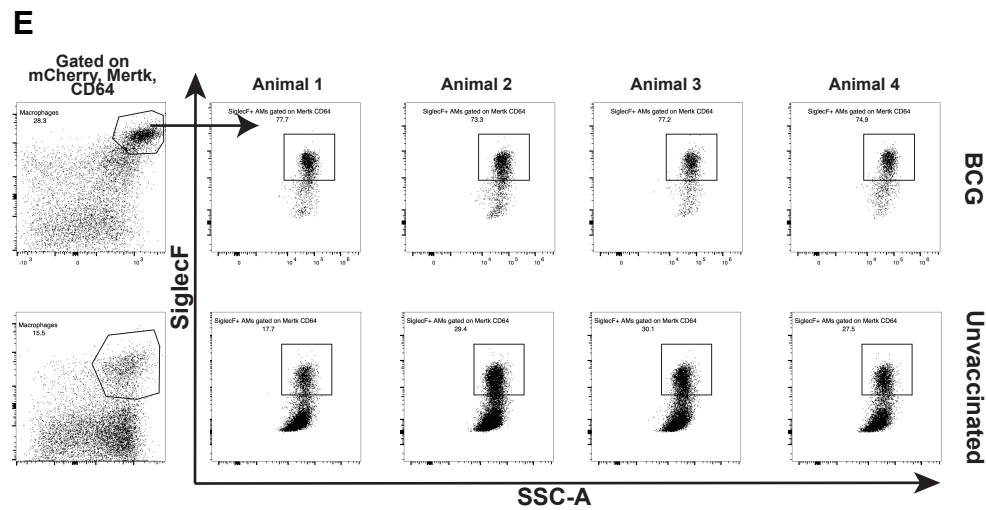
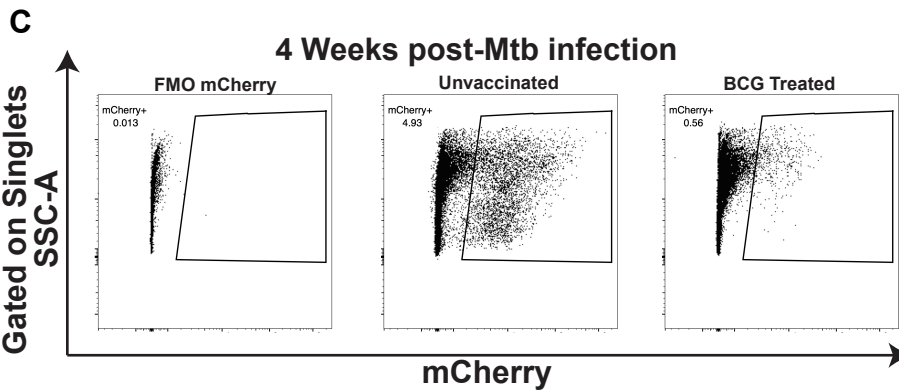
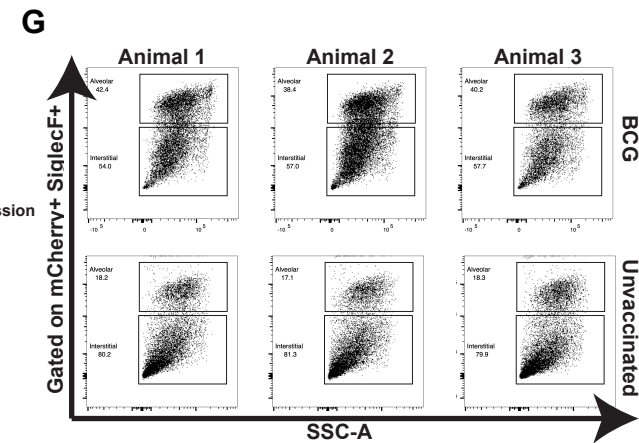
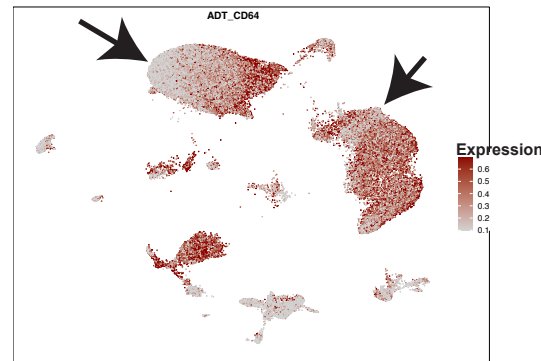
Small GTPase signal transduction

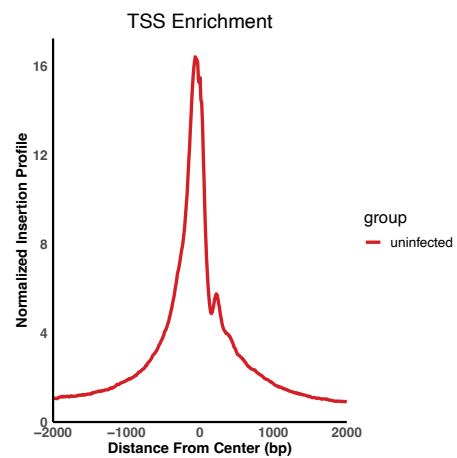
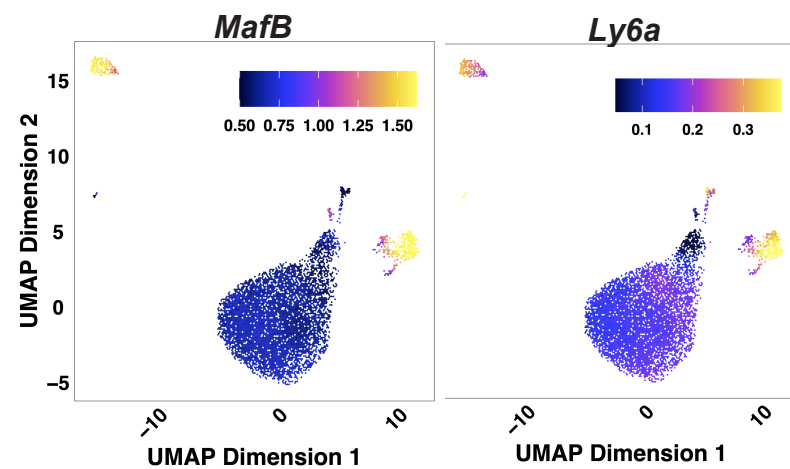
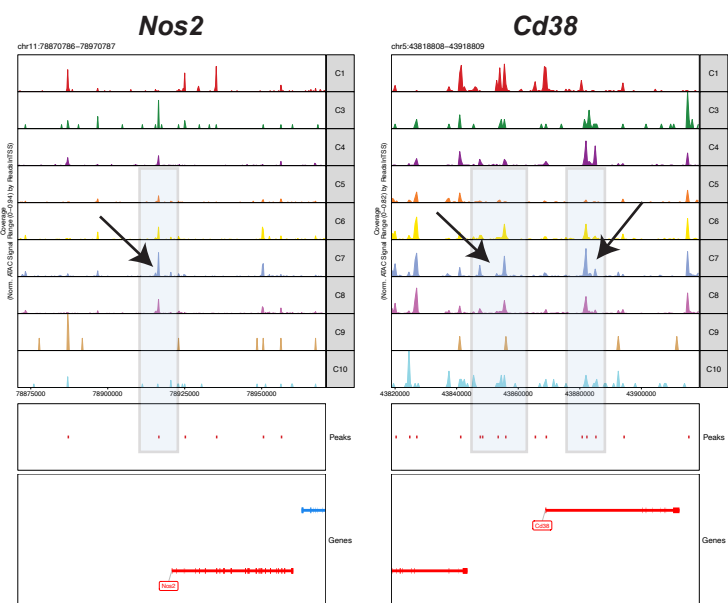




D

Lack of CD64 staining on monocyte-derived pro-inflammatory macrophages



A**B****C****C**