

SUPPLEMENTARY TABLES

Supplementary Table 1. Myeloid cell cytokine specific signature matrix.

MCCS table of the averaged gene expression values for myeloid cell cytokine specific signatures across all three donors. Colors indicate the specific cell type and cytokine stimulation condition.

Supplementary Table 2. Description of samples and cohorts of *M. tuberculosis* infection.

Sample breakdown and number of samples from RNA-Seq and Microarray as well as sample stratification (healthy, latent or active infection status) are included. A full list of the sample IDs obtained from GEO for GSE19491, GSE28623, GSE37250, GSE39939, GSE39940, GSE40553, GSE41055, GSE56153, GSE101705, GSE107995, GSE79362, GSE89403 and GSE94438 are included.

SUPPLEMENTARY FIGURES

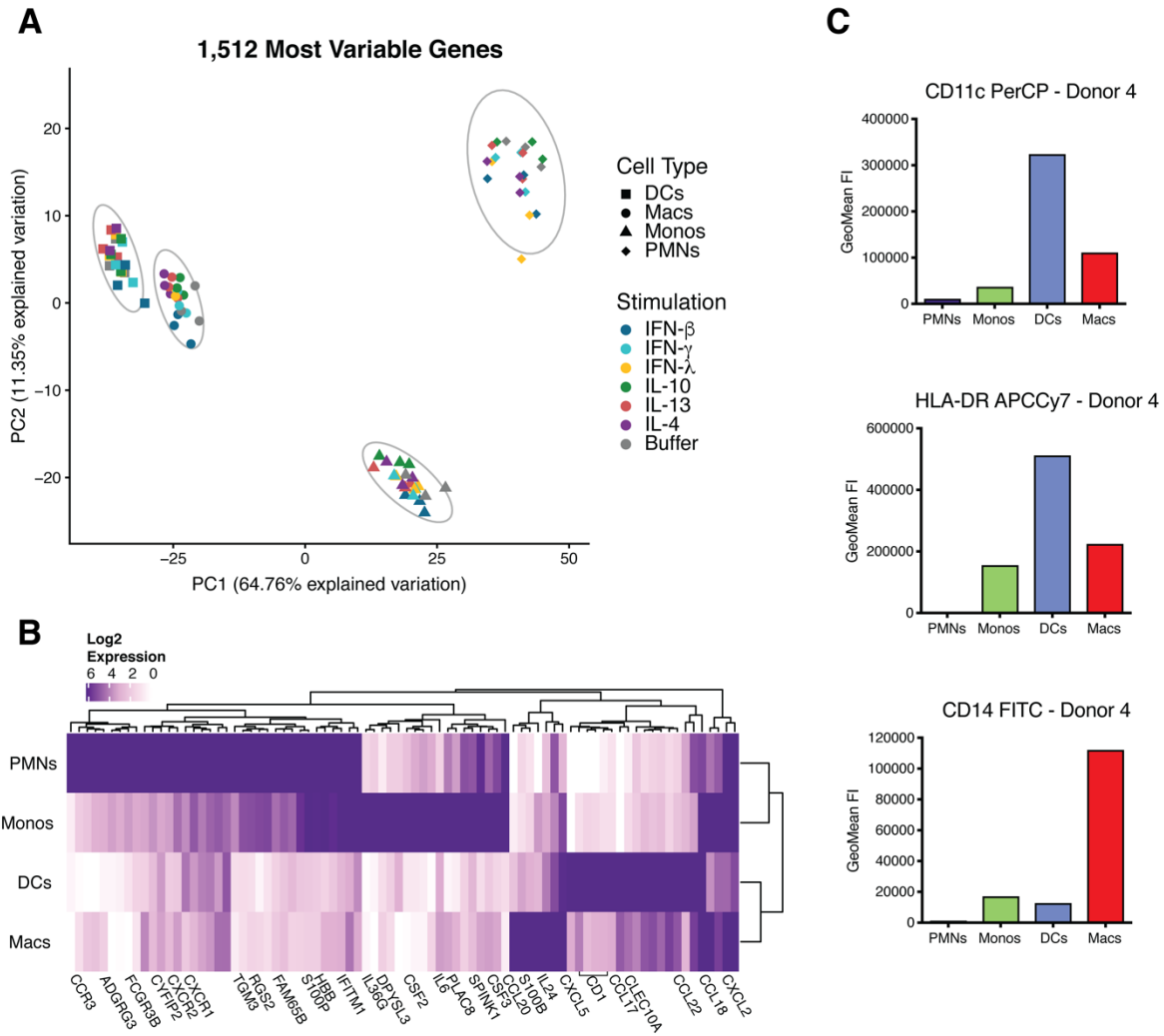


Figure S1. Transcriptional differences are cell type specific and greater than donor differences. Principle component analysis of 1,512 variably expressed genes determined by a 90% quantile filter for variance across all cell types and stimulations including buffer (A). Cell types are indicated by the corresponding symbol while stimulation is colored as indicated. Log2 expression levels of highly expressed genes in the buffer control samples from 3 donors in macrophage (Macs), dendritic cell (DCs), monocyte (Monos) and neutrophil (PMNs) subsets (B). Representative flow cytometry bar graphs of CD11c, HLA-DR and CD14 fluorescent intensity in macrophage, dendritic cell, monocyte and neutrophil subsets from Donor 4 (C).

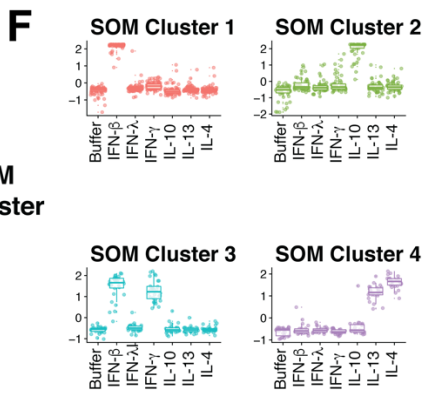
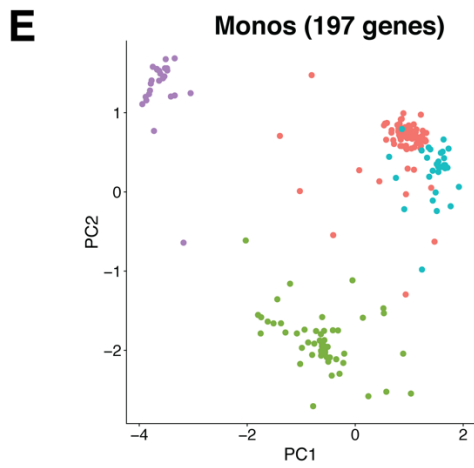
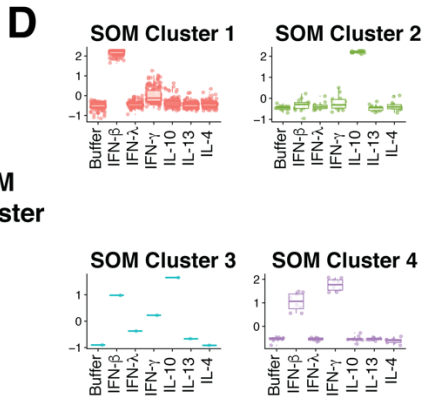
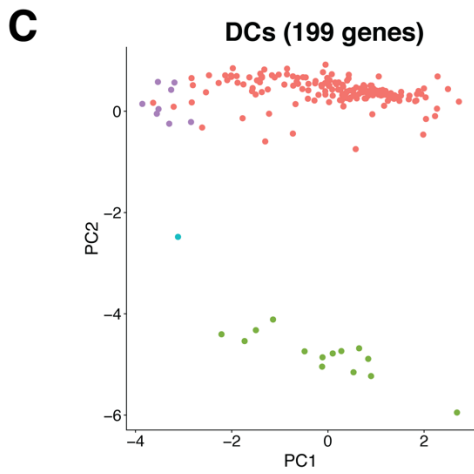
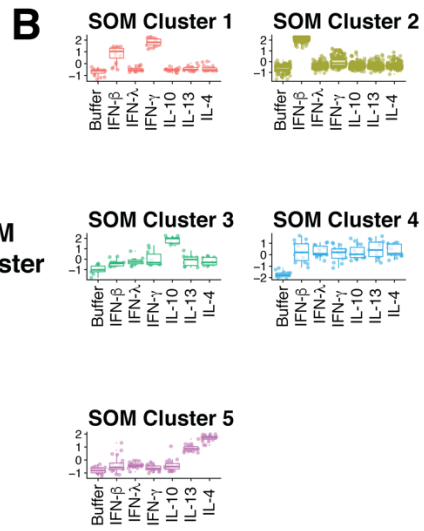
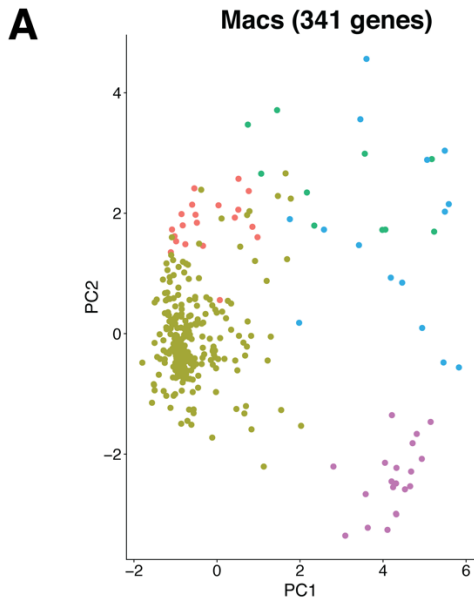


Figure S2. Specific cytokine stimulations share expression patterns within cell types. Self-organizing map analysis for cell type and stimulation clustering in Macrophages (A, B), dendritic cells (C, D), and Monocytes (E, F). Principle component analysis of each differentially expressed gene is shown to determine underlying clusters (A, C, E). Scaled gene expression levels for stimulations in each cluster for each cell type (B, D, F).

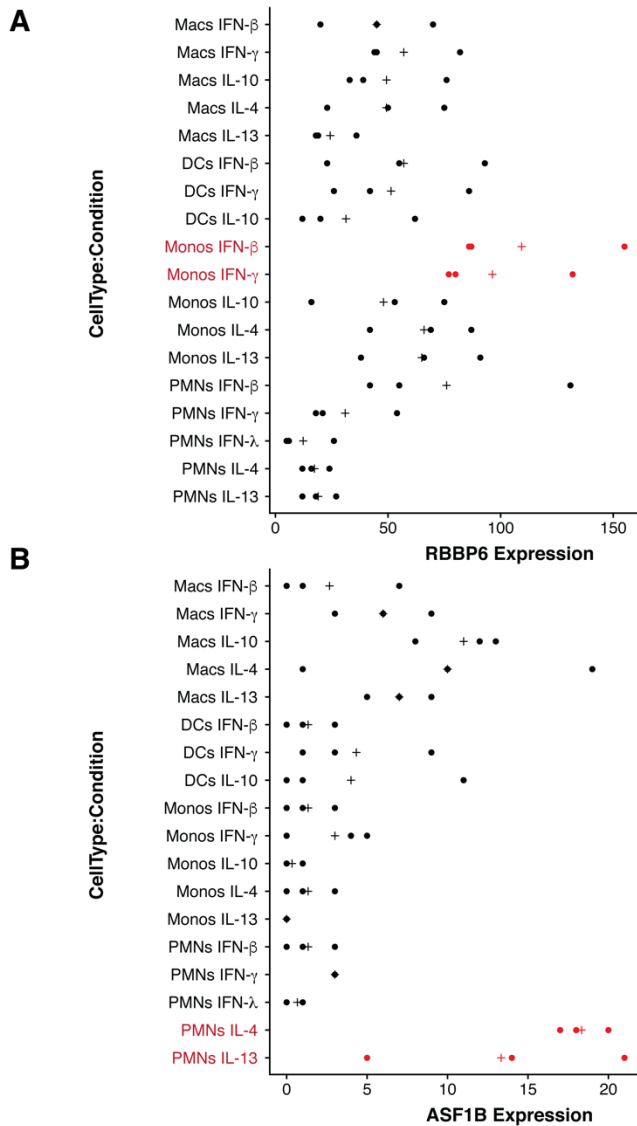


Figure S3. Outlier analysis confirms high-confidence genes within each cell type and stimulation condition. RBBP6 (A) and ASF1B (B) are shown as representative gene signatures in IFN- β and IFN- γ stimulated monocytes and IL4 and IL13 stimulated neutrophils respectively. Each dot represents one of three donors and the ‘+’ symbol indicates the mean expression level.

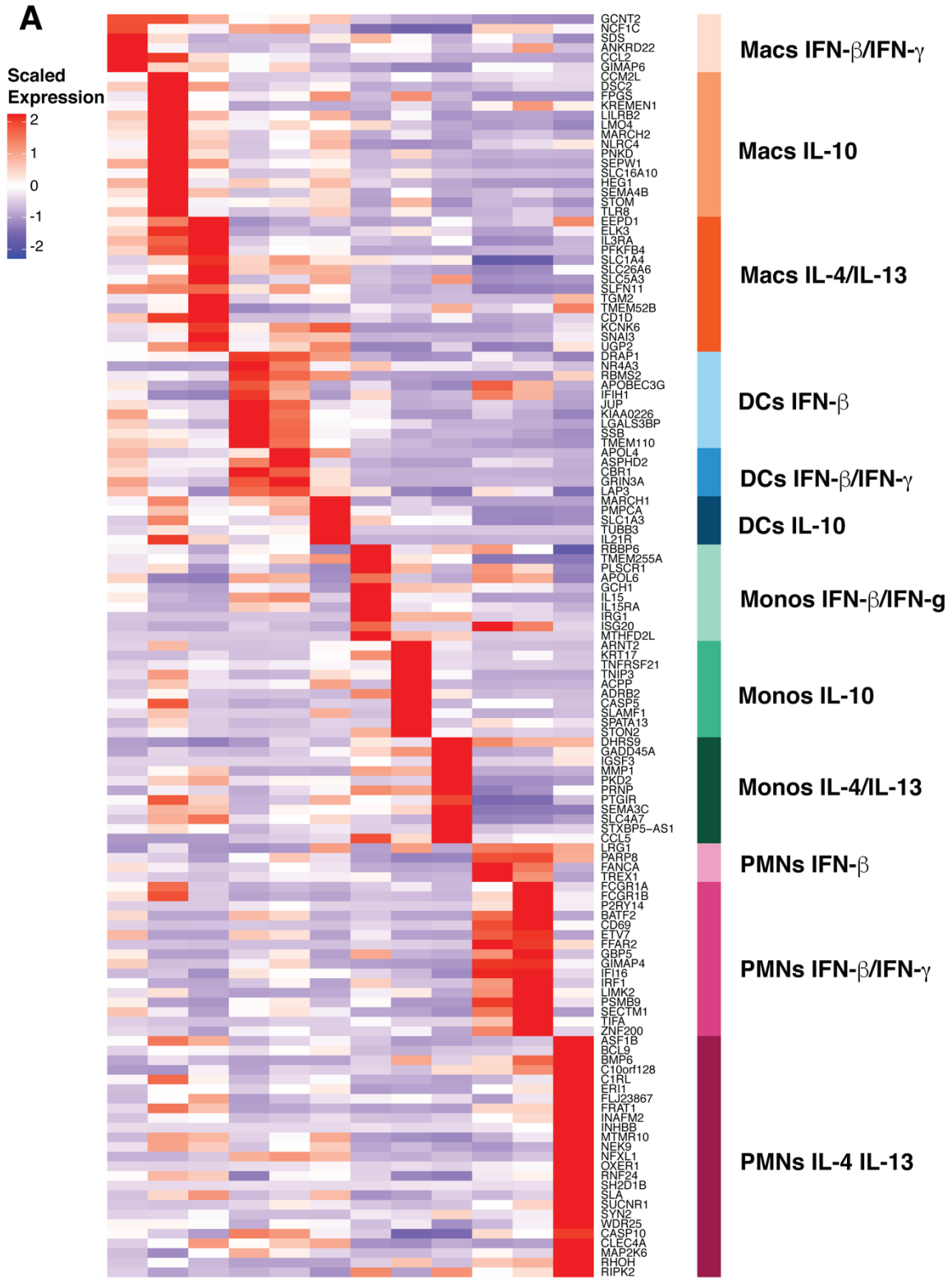
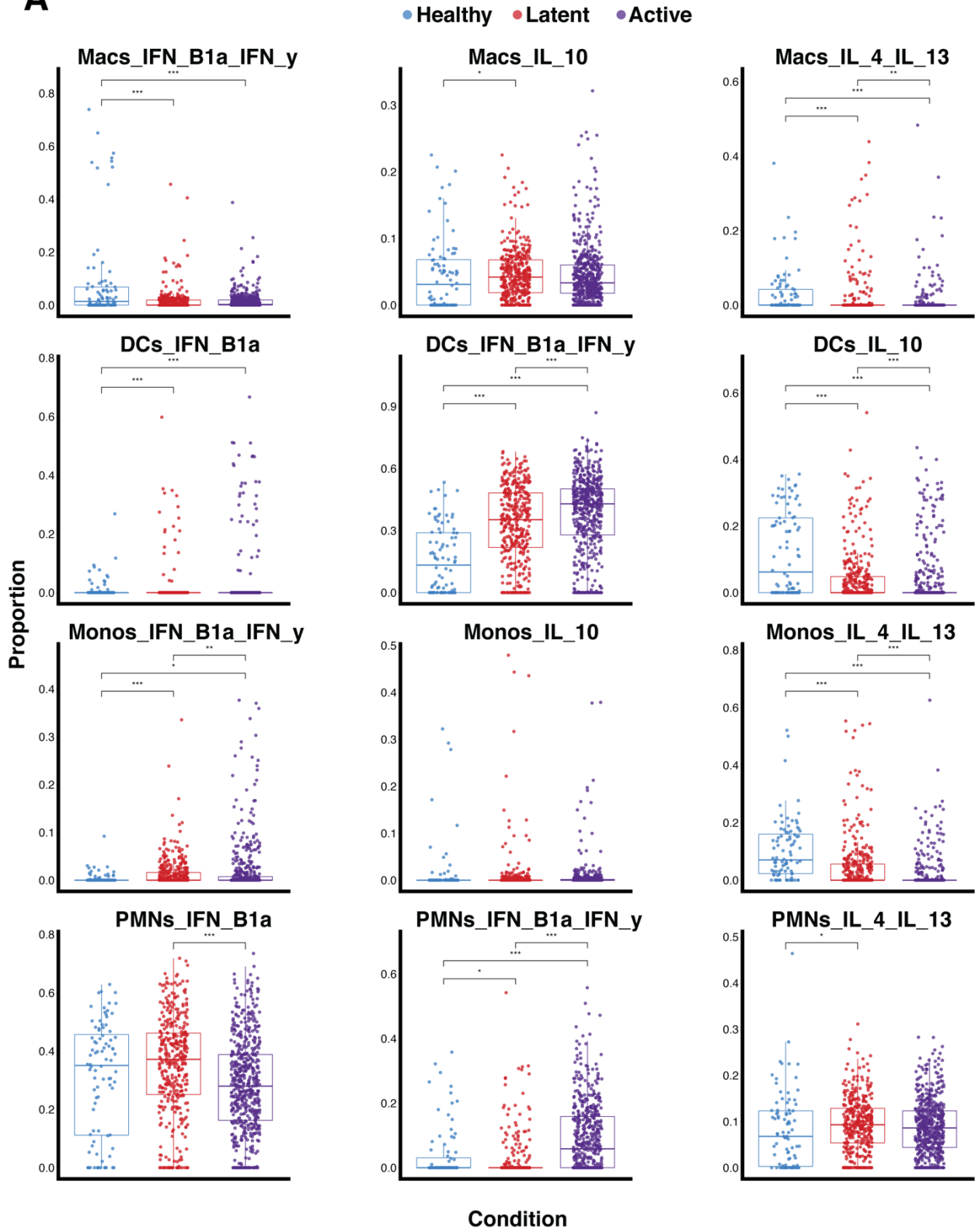
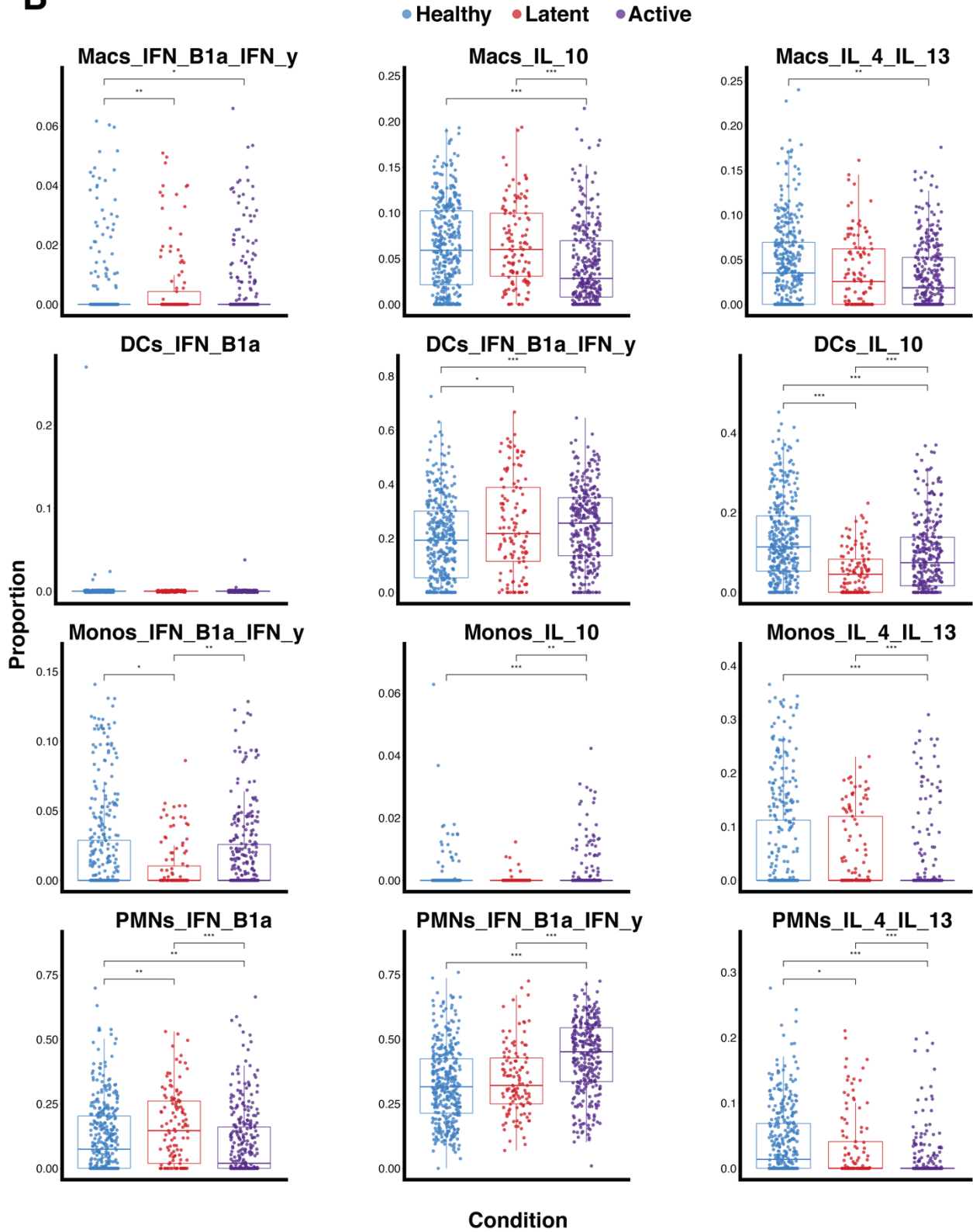


Figure S4 Myeloid cell cytokine stimulation signature matrix. 131 myeloid cell cytokine stimulation (MCCS) signature matrix with gene names.

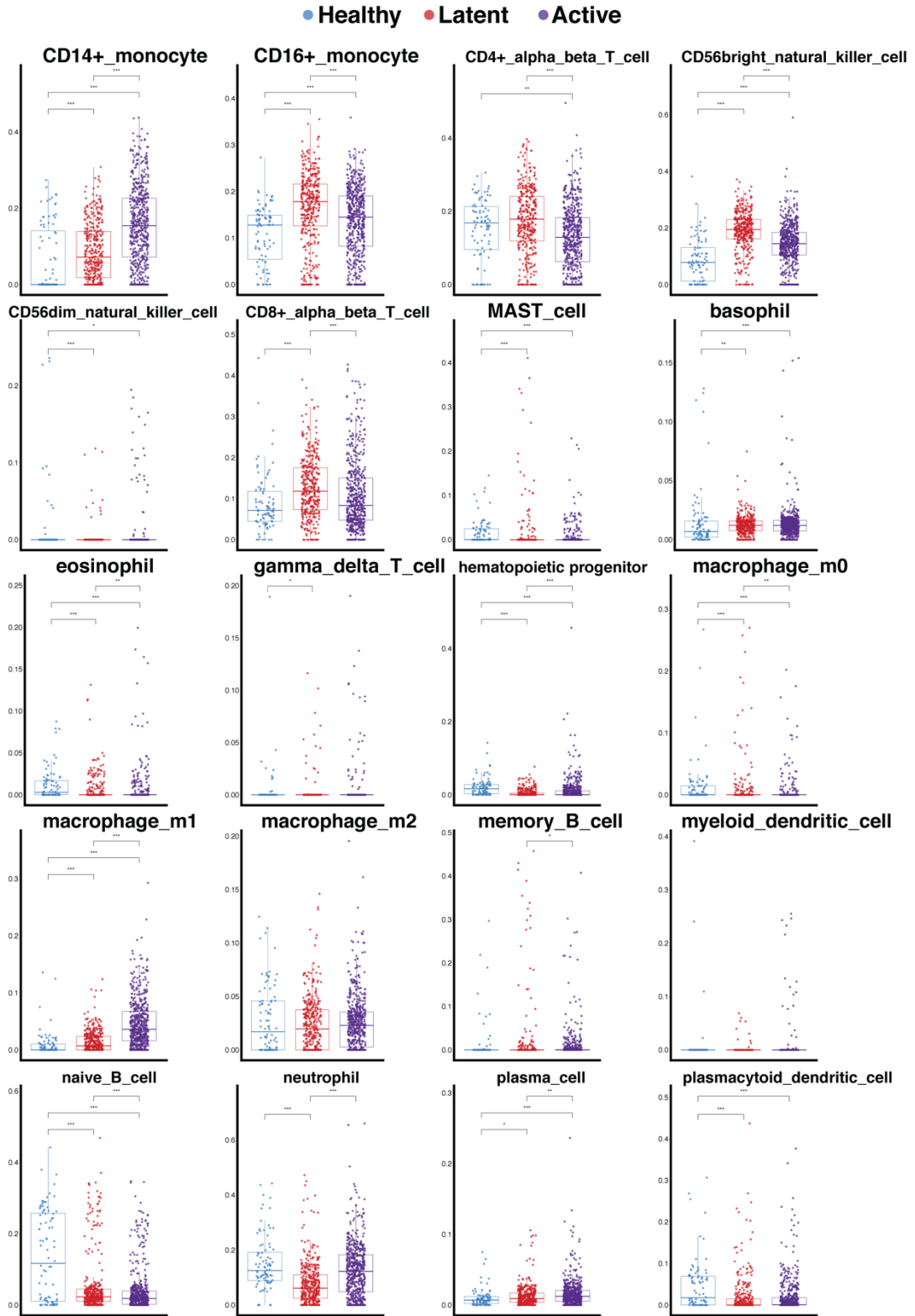
A



B

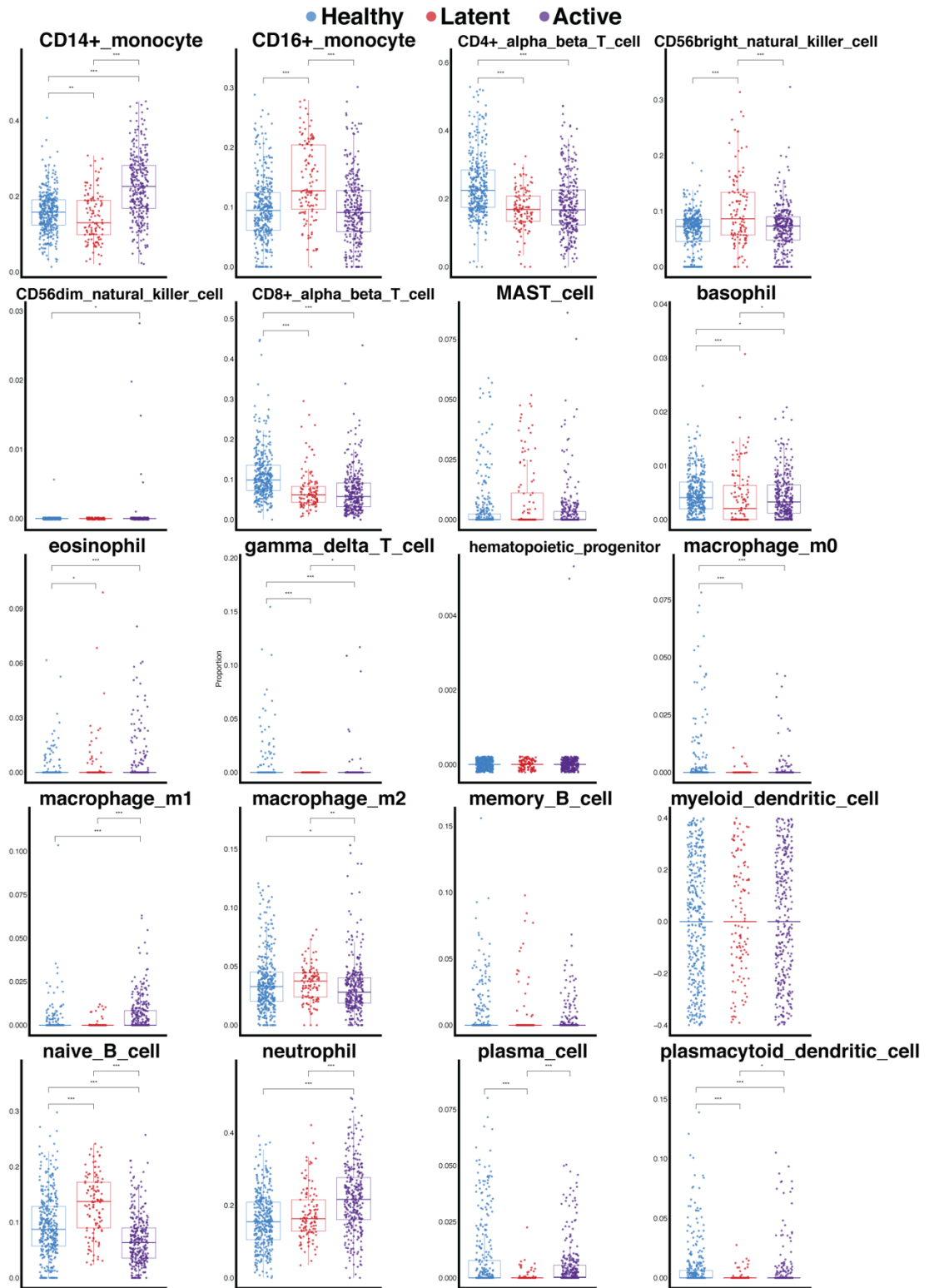


C



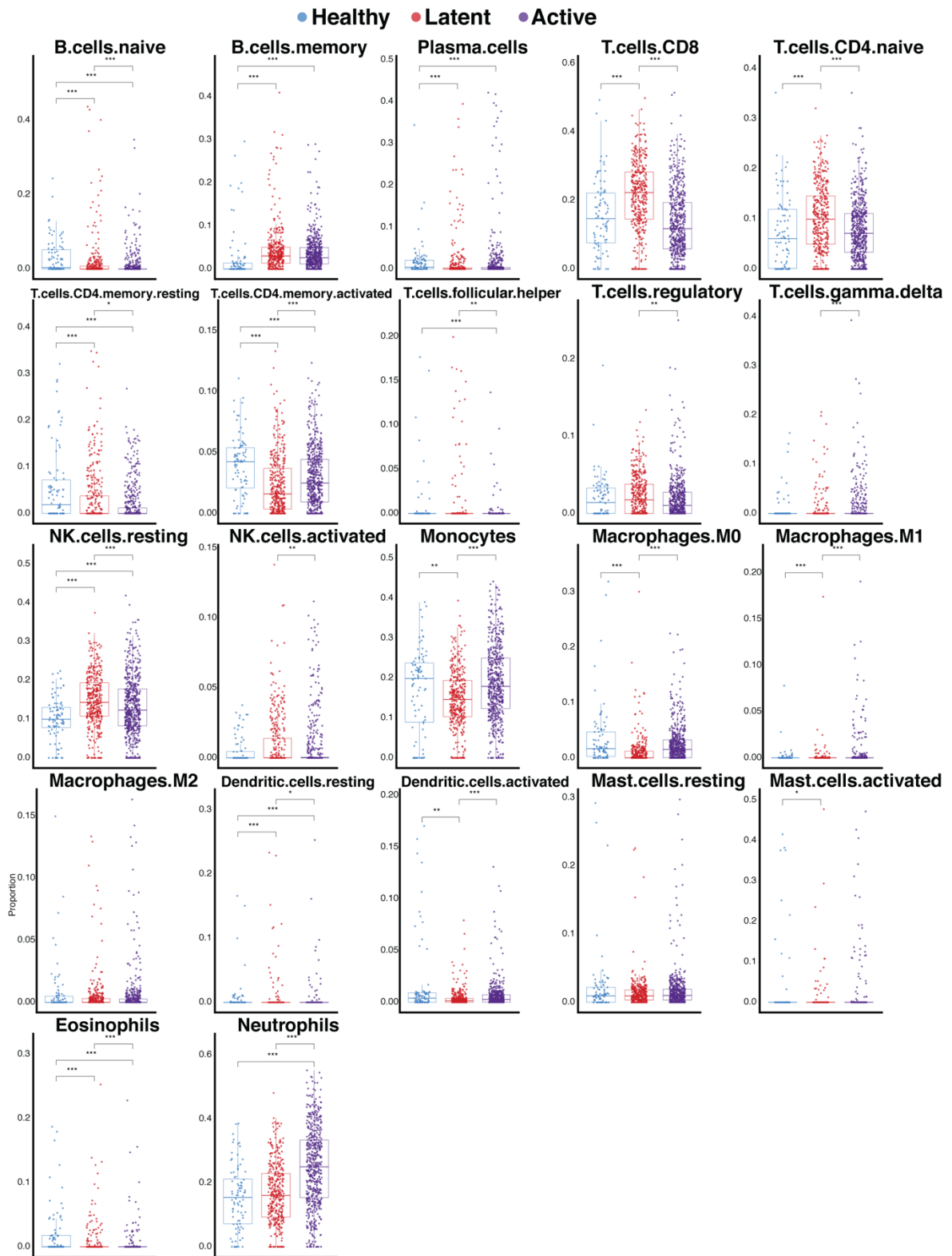
Matrix: ImmunoStates
Data Type: Microarray

D



Matrix: ImmunoStates
Data Type: RNA-Seq

E



Matrix: LM22

Data Type: Microarray

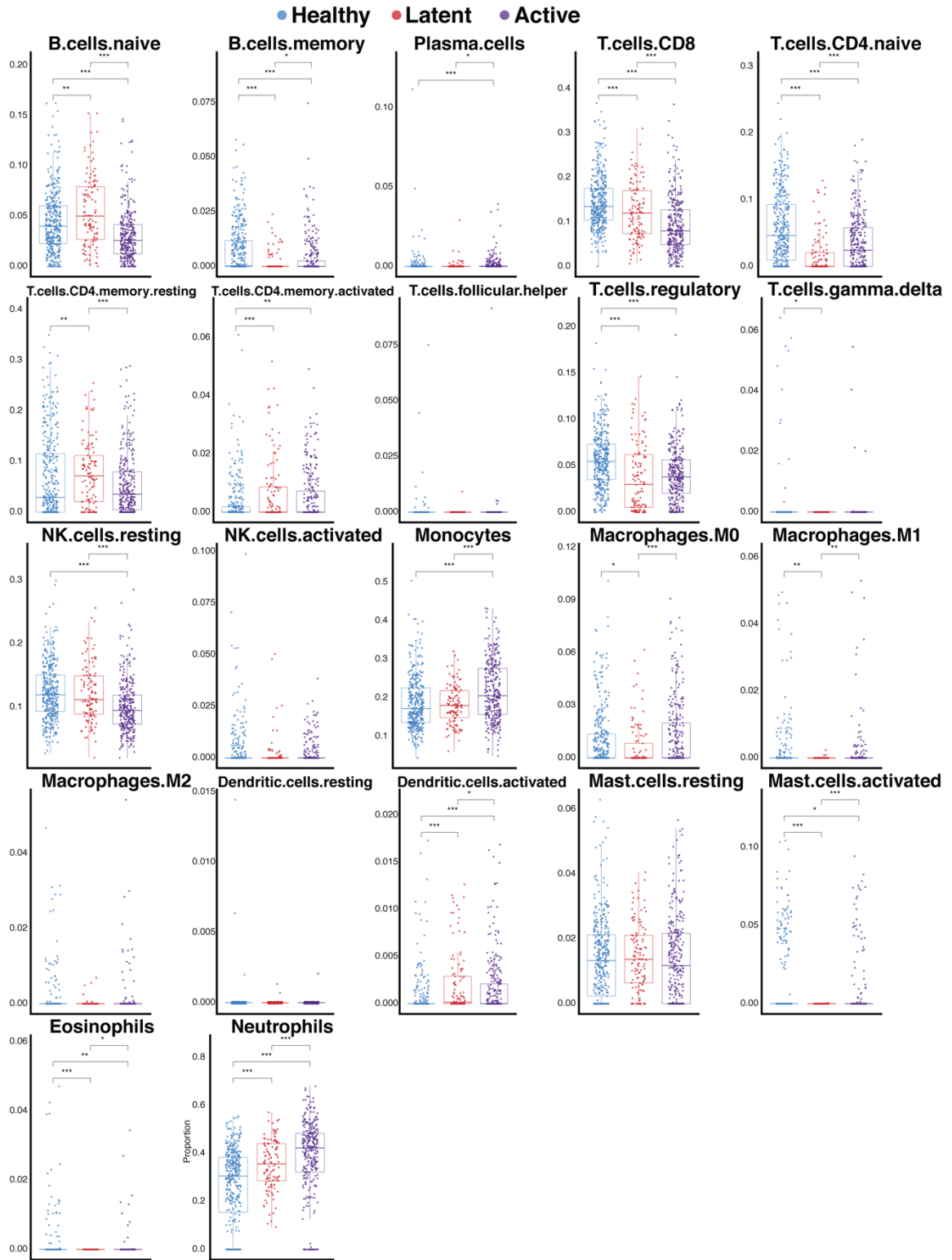
F**Matrix: LM22****Data Type: RNA-Seq**

Figure S5 Statistical deconvolution of 13 clinical cohorts of active and latent TB infections.

Box plots of TB infection status after deconvolution with our MCCA curated matrix for microarray (A) and RNA-Seq studies (B), the immunoStates matrix [10] for microarray (C) and RNA-Seq (D) studies and the LM22 matrix from CIBERSORT [1] for microarray (E) and RNA-Seq (F) studies. Significance determined by Kruskal-Wallis rank sum test with p -value $< 0.05 = *$, p -value $< 0.01 = **$ and p -value $< 0.001 = ***$. Sample sizes for each disease state and data type are as follows; healthy (microarray $n = 88$, RNA-Seq $n = 365$), latently infected (microarray $n = 376$, RNA-Seq $n = 117$) and active disease individuals (microarray $n = 547$, RNA-Seq $n = 306$).

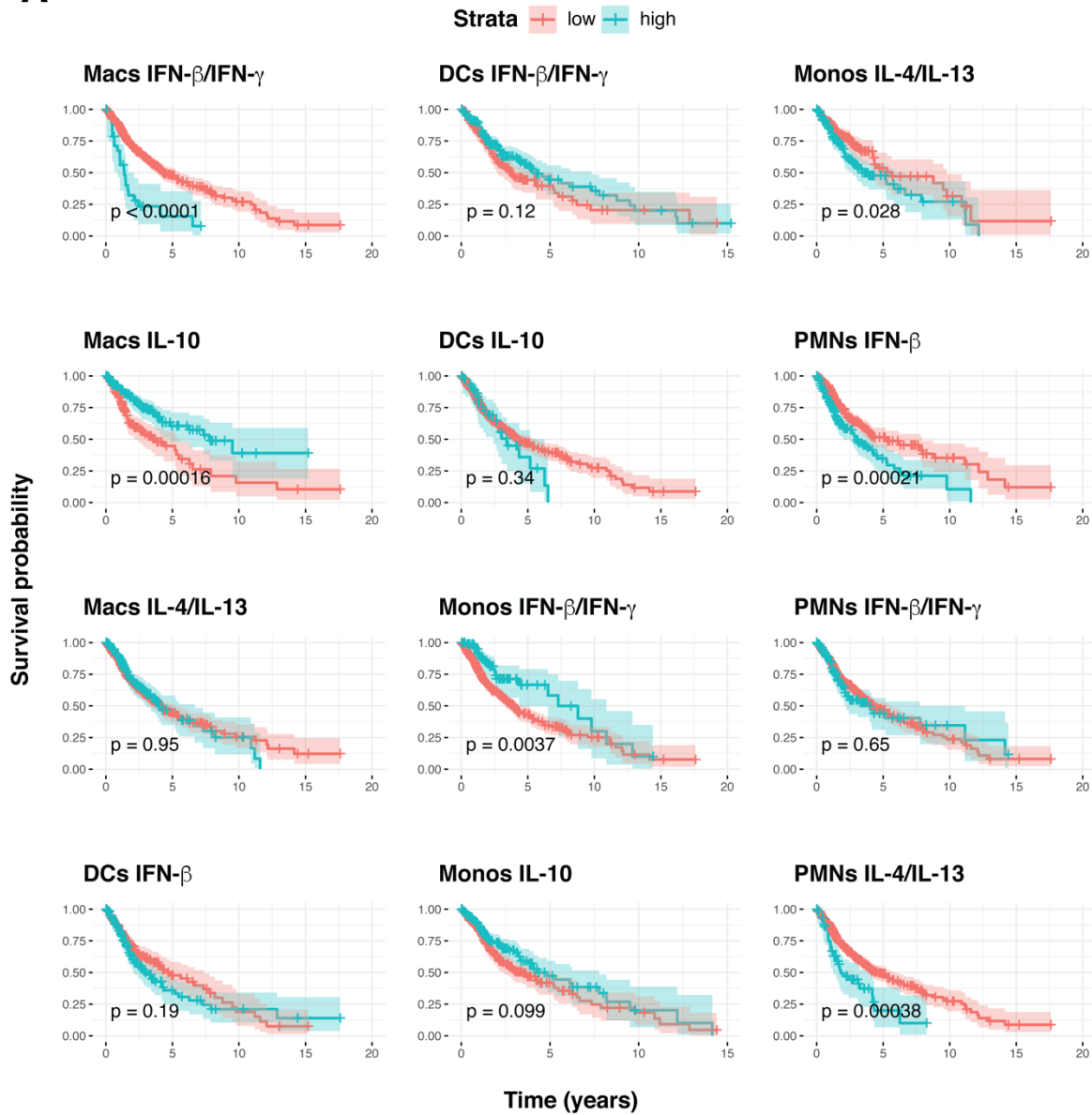
A

Figure S6. MCCS signatures are linked to survival in glioma patients. Survival analysis of 671 primary glioma tumors after cell type specific deconvolution through CIBERSORT using our MCCS signature matrix. Samples are stratified by low and high quartiles for resulting deconvolution proportion estimate. Significance determined by log rank test.

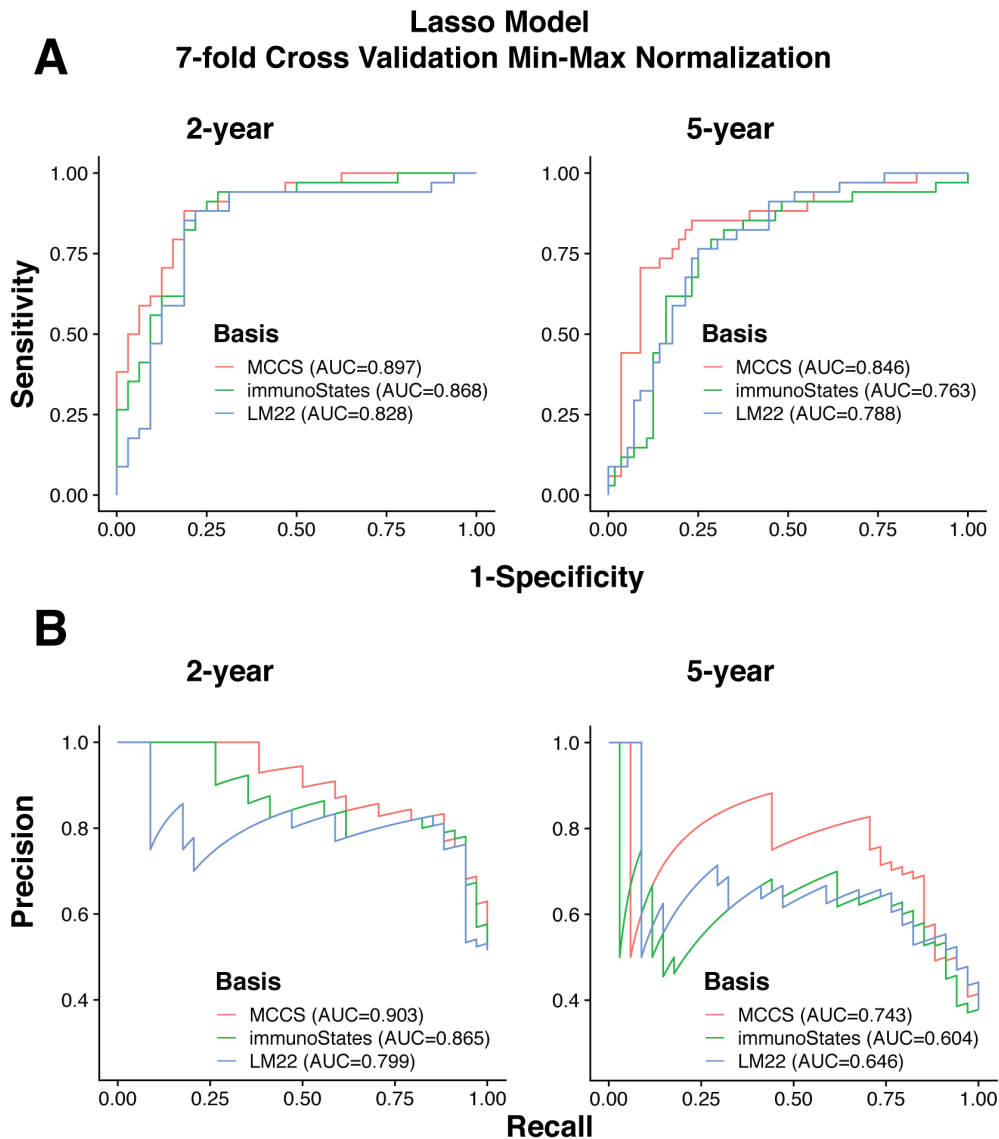


Figure S7. MCCS gene signature is highly successful at predicting long-term survival in glioma patients. Area under ROC curves (A) and Precision Recall curves (B) for 2 and 5-year LASSO models with 7-fold cross validation and min-max normalization. Models trained on 131 MCCS gene signature (red), 341 immunoStates signature (green) or 547 LM22 gene signature set (blue).