

Figure S1, related to Figure 1. QC Metrics and cell type distribution by disease and inflammation status. (A) Violin plots of layer preparation effects on (top to bottom): Number of detected genes in cells, mitochondrial percentages, and total number of UMIs, from all three cell compartments (left to right), and between colon and TI locations. P-values: **<0.01, *<0.05. P-values are from the *lmerTest* package in R with formula $y \sim \text{Layer} + (1 | \text{Donor}) + (1 | \text{Channel})$). The number of donors contributing to each violin plot is shown below. Violin plots were weighted by the inverse of the number of cells per donor to reduce the weight of donors that contributed many cells. (B) Fraction of the cells from each sample belonging to each cell type, broken down by cell compartment, in TI (left) and CO (right). For epithelial cell types, only samples from the epithelial layer or non-separated samples were used, while for stromal and immune cell types, lamina propria and non-separated samples were used. Top bars: E: epithelium, L: lamina propria. Non-separated samples are represented with both E and L. (C) Principal coordinates analysis of Bray-Curtis dissimilarities from cell type composition using data only from colon. Left to right: Study origin (CD: from this study UC Control: from Smillie, et al and Jasso, et al^{16,20}; Disease status (Healthy, Non-inflamed or Inflamed); Layer (E: epithelium, L: lamina propria, N: not separated). (D) Principal coordinates analysis of Bray-Curtis dissimilarities from sample-level cell type composition profiles as **Fig. 1D**, split by layer and colored by disease status. (E-F) Barplots show significant differences in cell type frequency for non-inflamed (blue) and inflamed (red) samples relative to healthy (green) samples in immune, stromal and epithelial compartments in TI and CO. (*adjusted $p < 0.05$, **adjusted $p < 0.01$, red asterisks means overrepresented in inflamed or non-inflamed samples vs. healthy, blue asterisks means underrepresented).

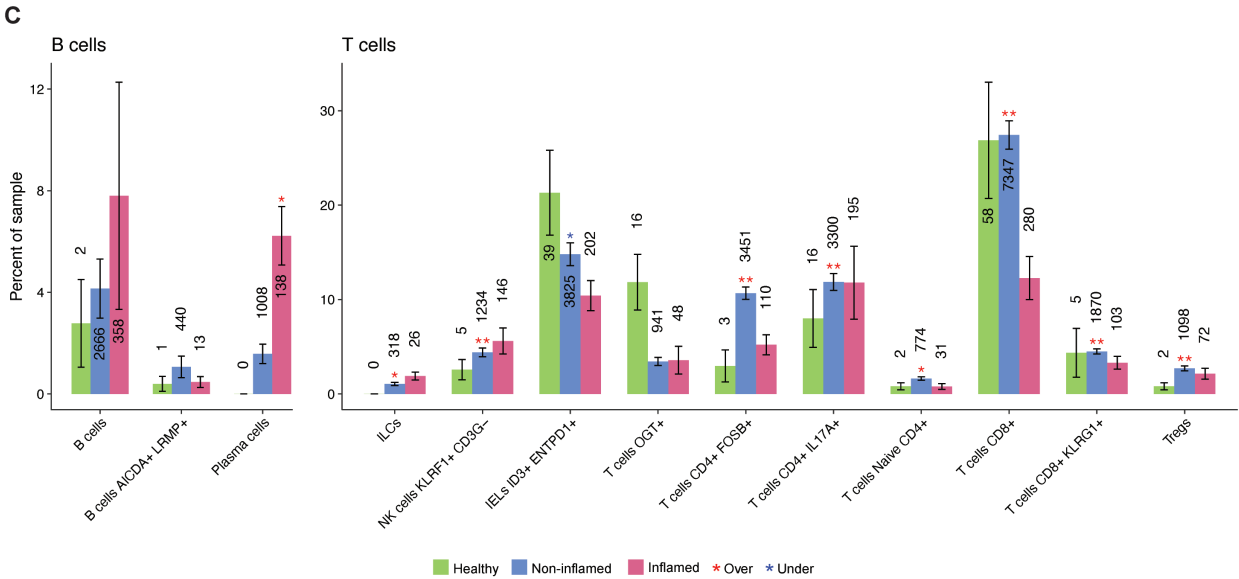
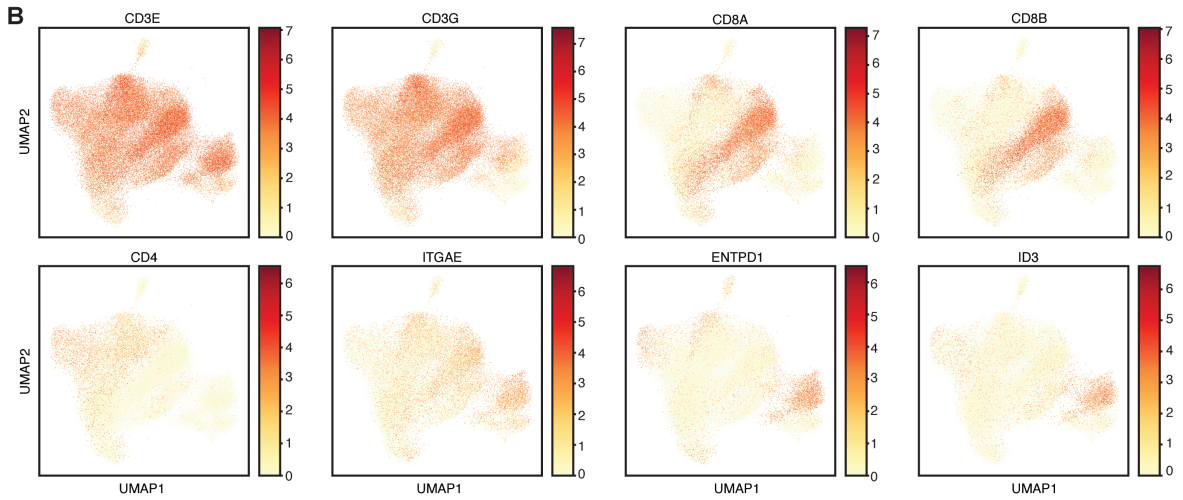
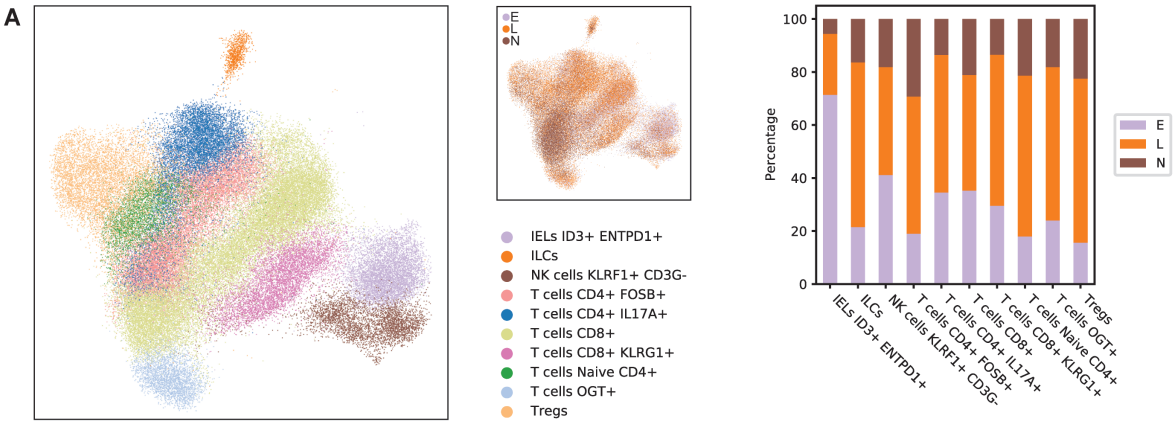


Figure S2, related to Figure 1. Intraepithelial lymphocytes (IELs) in TI. (A) UMAP visualization of T cells in TI colored by cell type (left) and layer preparation (middle). Right: Fraction of cells from each layer for each T cell type. **(B)** Expression profiles of selected marker genes for the IELs ID3+ ENTPD1+ population in TI. **(C)** Significant differences in cell type frequency for non-inflamed (blue) and inflamed (red) samples relative to healthy (green) samples in the immune compartment in TI, among epithelial layer samples. (*adjusted $p < 0.05$, **adjusted $p < 0.01$, red* means overrepresented in inflamed or non-inflamed samples vs. healthy, blue* means underrepresented). The total number of cells contributing to each bar is also shown.

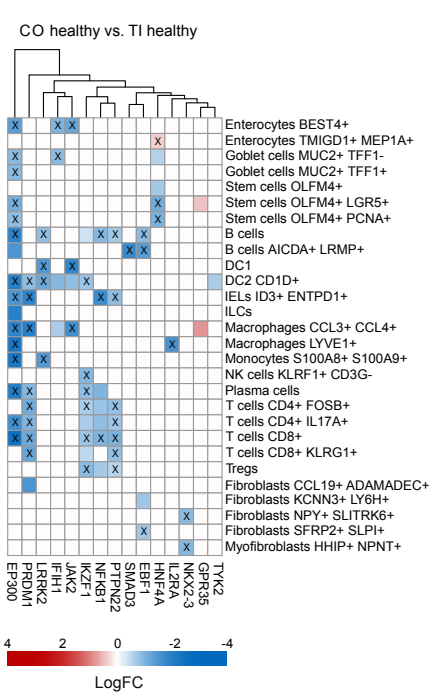
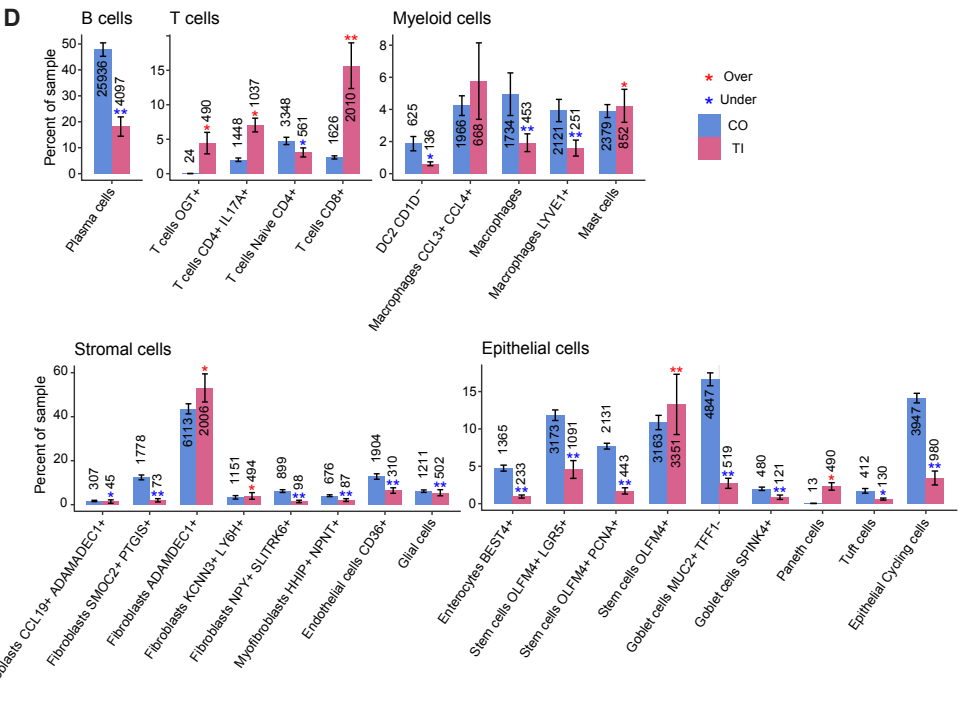
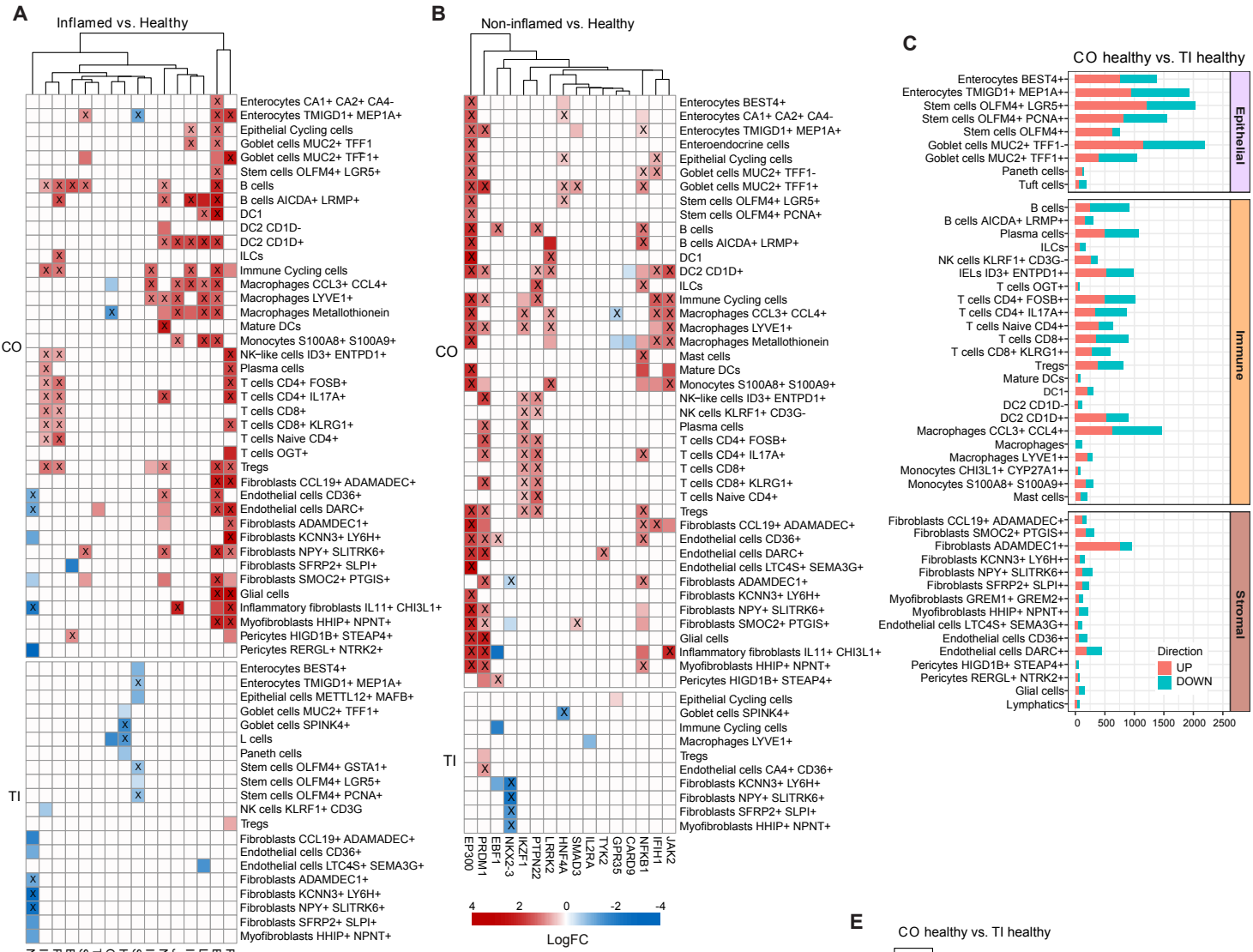


Figure S3, related to Figure 2. Expression of IBD-related genes in disease in TI and CO. (A) DE coefficients of core IBD risk genes from inflamed vs. healthy samples. Genes that have $FDR < 0.1$ in at least one cell type were shown. X in the heatmap indicates $FDR < 0.05$. **(B)** Same as **(A)**, but in non-inflamed vs. healthy samples. **(C)** Number of differentially-expressed genes between healthy colon and healthy TI samples, broken down by cell type (discrete component of a MAST model; $FDR < 0.05$). **(D)** Significant differences in cell type frequency for colon (blue) healthy samples relative to TI (red) healthy samples in immune, stromal and epithelial compartments in TI, among epithelial layer samples. (*adjusted $p < 0.05$, **adjusted $p < 0.01$, red* means overrepresented in inflamed or non-inflamed samples vs. healthy, blue* means underrepresented). **(E)** DE coefficients of core IBD risk genes from healthy colon and healthy TI samples. Genes that have $FDR < 0.1$ in at least one cell type were shown. X in the heatmap indicates $FDR < 0.05$.

Figure S4, related to Figure 3. Expression of five IBD-associated loci, non-inflamed vs healthy differential expression, and the Ketogenesis and *PPAR* signaling pathways in CD. (A) Scaled mean expression (**Methods**) of five new IBD-associated loci in both TI and colon. **(B)** Heatmap of DE coefficients for five new IBD-associated loci in CO. “X” in the heatmap indicates $FDR < 0.05$ and “|” indicates $FDR < 0.1$. Results in TI showed less significance; full results in **Table S3**. **(C)** Number of differentially-expressed genes between non-inflamed CD and healthy samples in TI and colon **(D)**, broken down by cell type. **(E)** Heatmap of key genes involved in Ketogenesis and *PPAR* signaling pathway in Inflamed-Healthy in TI and CO. X in the heatmap indicates $FDR < 0.05$.

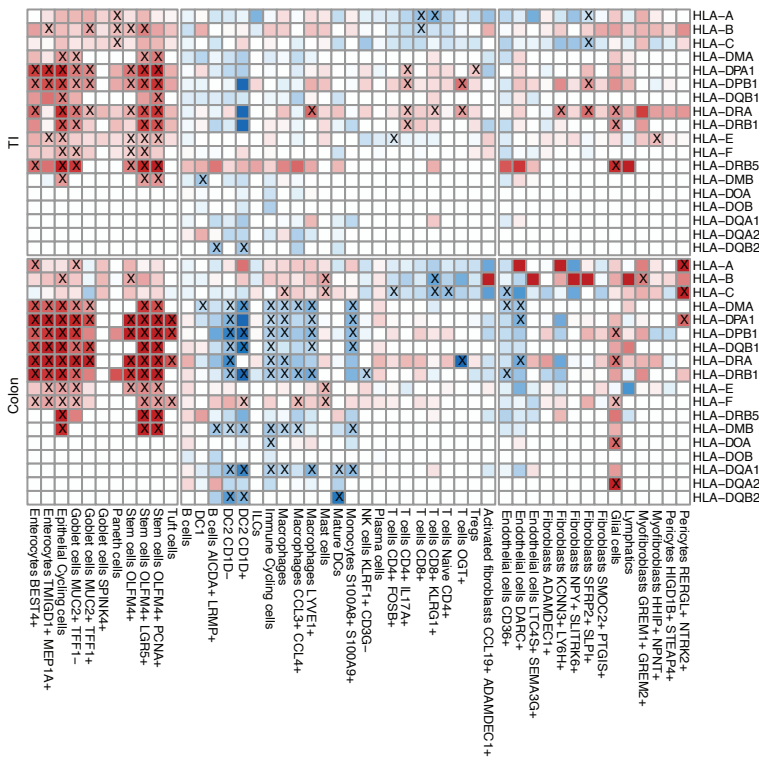
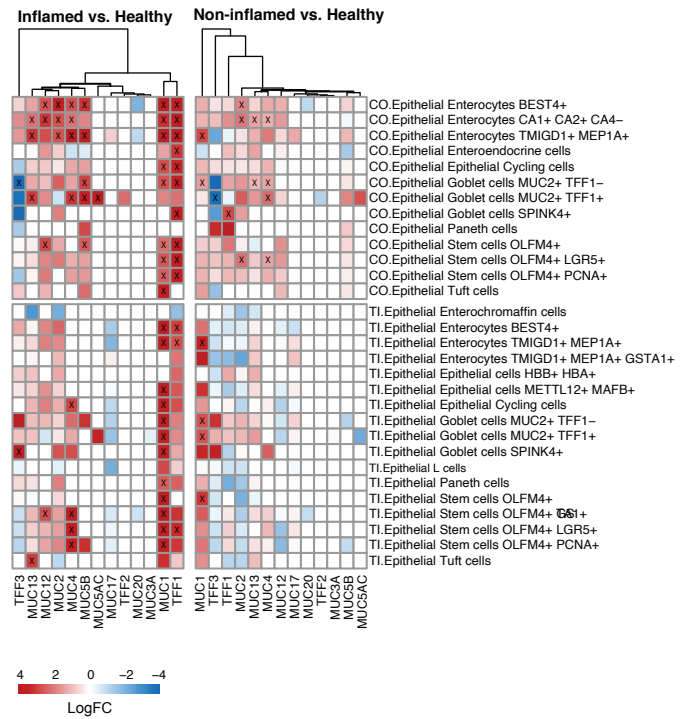
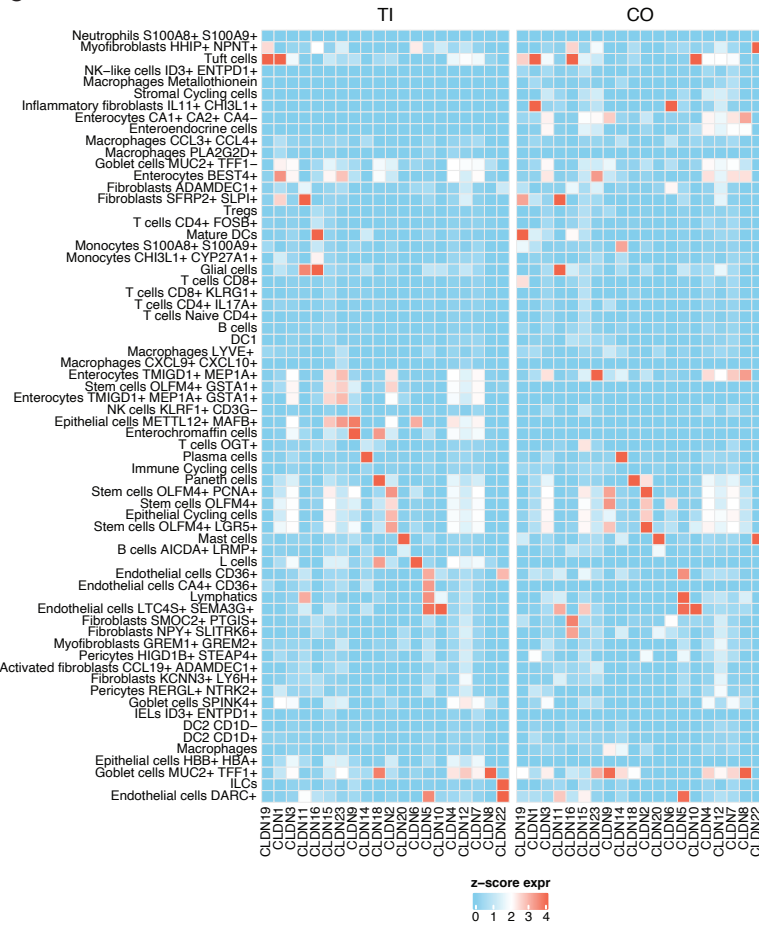
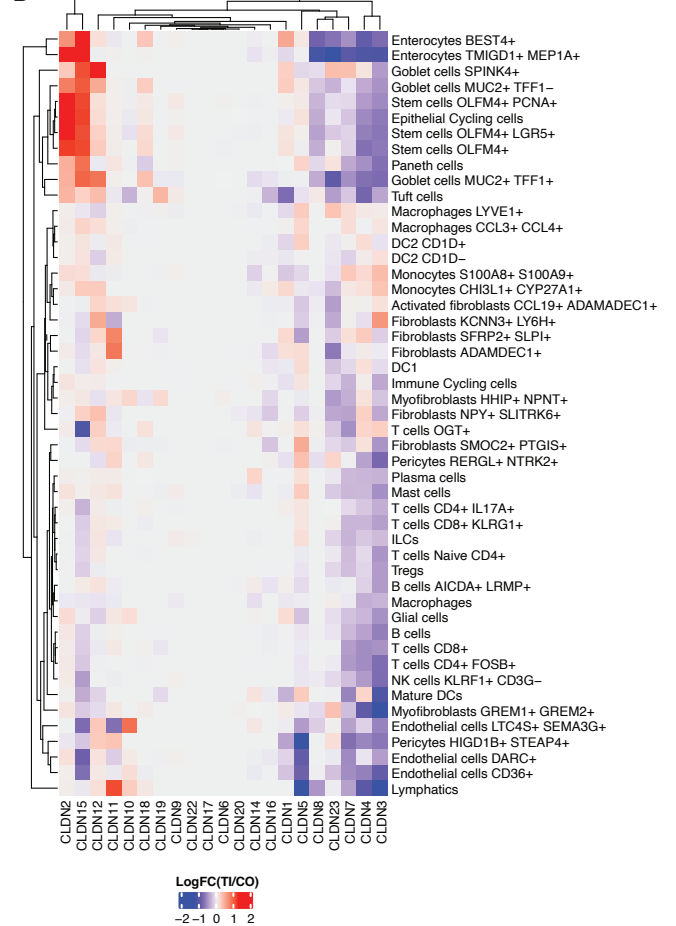
A**B****C****D**

Figure S5, related to Figure 2-3. Differential expression in HLA, mucin and TFF genes in TI and CO. (A) Heatmaps of DE coefficients between inflamed vs. healthy samples in HLA- genes. X in the heatmap indicates $FDR < 0.05$. **(B)** Heatmaps of DE coefficients among 14 mucin and TFF genes. Left: inflamed vs. healthy; Right: noninflamed vs. healthy. X in the heatmap indicates $FDR < 0.05$. **(C)** Scaled mean expression (**Methods**) of claudin genes in both TI and colon. **(D)** logFC of TI/CO in mean expression among claudins.

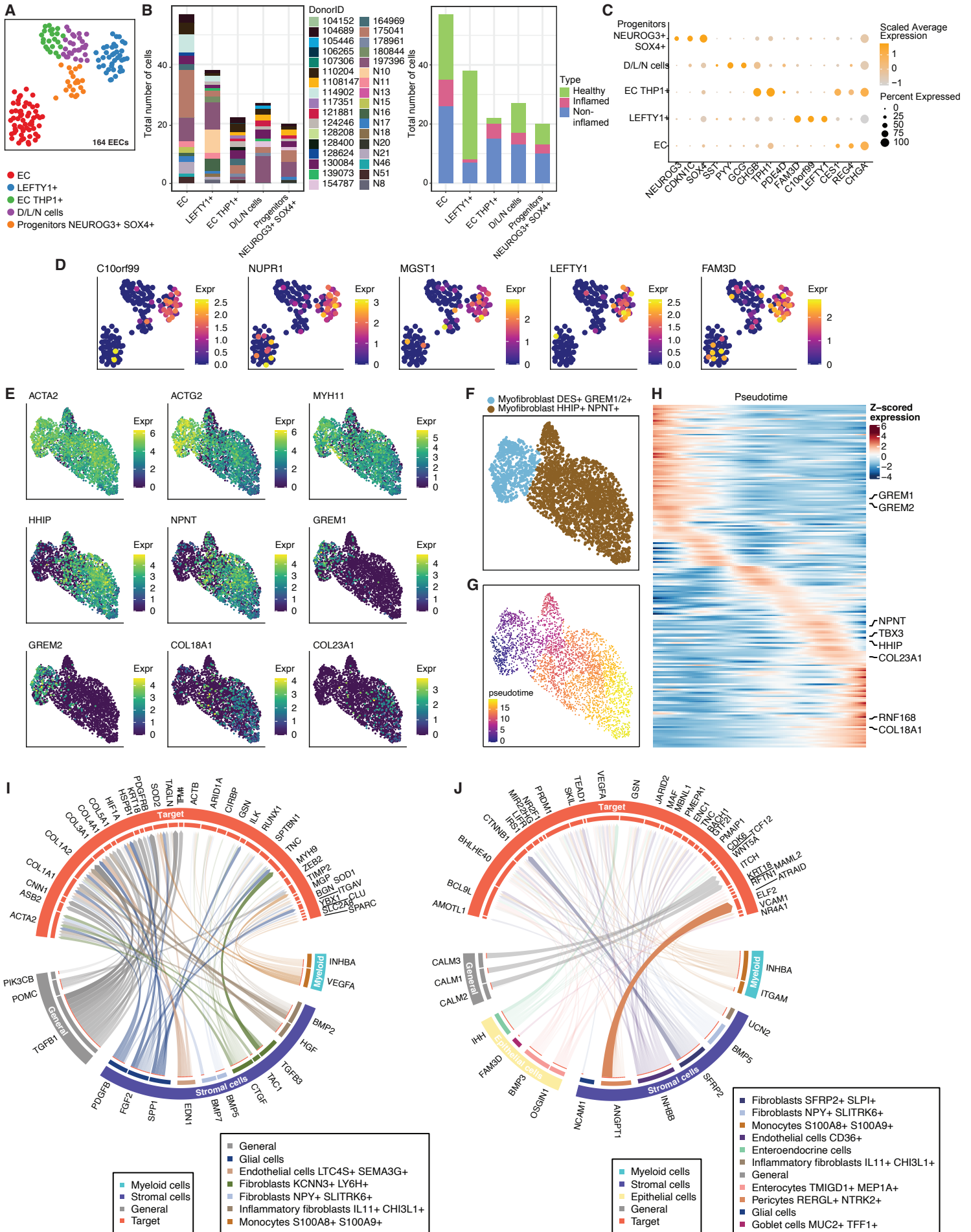


Figure S6, related to Figure 4-5. Enteroendocrine Cells (EEC) and Myofibroblasts in the colon. (A) UMAP of 164 EECs in colon, colored by subset. (B) Donor and disease composition among EEC subsets. (C) Expression of markers for major EEC subsets. (D) Expression profiles of select top genes expressed in the LEFTY1+ cluster. (E) UMAP of the myofibroblast sub-groups in colon (HHIP+ NPNT+ and GREM1+ GREM2+), overlaid with expression of myofibroblast markers and sub-group markers. (F) Myofibroblast sub-groups clustered into two distinct clusters. (G) Pseudotime trajectory UMAP in TI myofibroblasts. (H) Expression profiles of 200 selected genes with significant expression changes with respect to pseudotime (**Methods**). Expression was first smoothed with a cubic spline and standardized to highlight expression differences (**Methods**). (I) Circos plot summarizes the predicted upstream ligands (bottom half genes) based on the DEGs in Myofibroblasts GREM1+ GREM2+ (Target, top half genes) by NicheNet. Predicted ligands were assigned to the cell type with maximum expression if Gini>0.5, otherwise the ligand is “General”. Cell types were further grouped into cell categories (overall and detailed cell type legends were shown next to the circo plot). (J) Heatmap (left, orange) showed the predicted ligand activity driving DEGs in myofibroblasts GREM1+ GREM2+ in inflamed samples, ranked by Pearson correlation coefficient (**Methods**). The potential of each ligand for inducing differential expression of myofibroblast target genes was shown by regulatory potential (right, purple heatmap).