bioRxiv preprint doi: https://doi.org/10.1101/2023.05.08.539701; this version posted May 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## Supplementary Figures and Legends

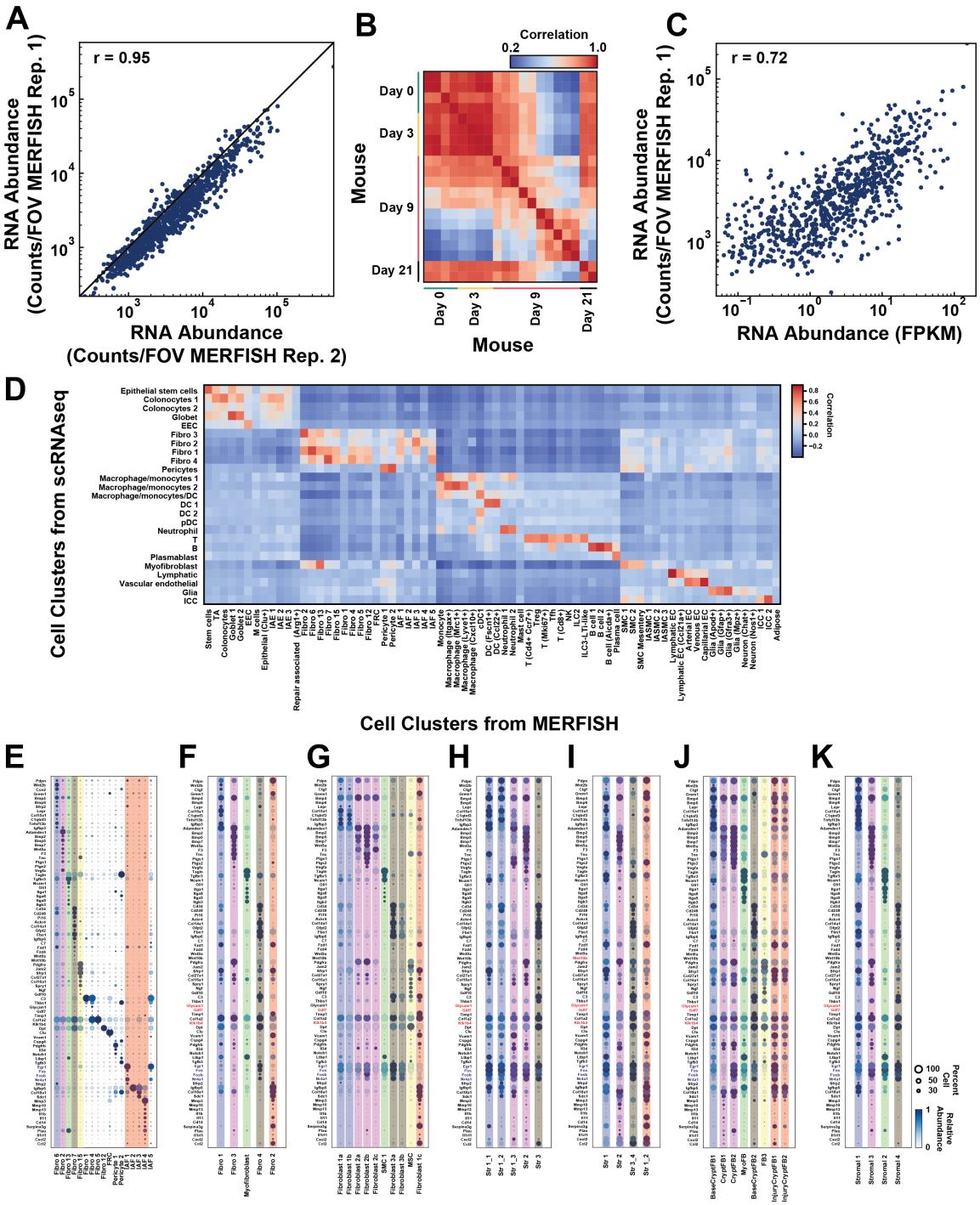


**Figure S1: Reproducibility and cross-validation of MERFISH measurements**

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.08.539701; this version posted May 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(A) Scatter plot of counts per cell for each RNA averaged across all fields-of-view (FOV) for one MERFISH measurement in a Day 0 mouse versus those for a replicate MERFISH measurement in a different Day 0 mouse. The Pearson correlation coefficient (r) between the logarithm of the average expression values is 0.95.

(B) Heat map of the pairwise Pearson correlation coefficients calculated as in (A) for all MERFISH measurements (each measurement often contains multiple slices), indicating strong reproducibility between measurements of mice from different disease stages.

(C) Scatter plot of the average counts per cell for each RNA measured with MERFISH from a representative Day 0 replicate versus the average expression determined from bulk RNA sequencing of tissue harvested from Day 0 mice. The Pearson correlation coefficient between the logarithm of the average expression values is 0.72.

(D) Heatmap of the pairwise Pearson correlation coefficients determined between the expression profiles measured for cell clusters identified in MERFISH and those previously determined via scRNAseq[39]. Expression was determined by averaging the z-score of the logarithmic expression for all cells within a given cluster.

(E) Dot plot representation of the expression of key genes for fibroblast populations determined with MERFISH. Color indicates the relative expression of each gene across the listed cell populations while the circle size represents the fraction of cells with at least RNA copy. The colored bars represent corresponding fibroblast populations observed in published data as judged by the co-expression of key marker genes.

(F-K) As in (E) but for all fibroblast populations identified via scRNAseq of the mouse colon in healthy and DSS-induced colitis from Ho et. al.[39] (F), Jasso et. al.[18] (G), Kinchen et. al. (healthy)[9] (H), Kinchen et. al. (DSS-treated)[9] (I), Xie et. al.[40] (J), and Brügger et. al.[15] (K).

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.08.539701; this version posted May 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

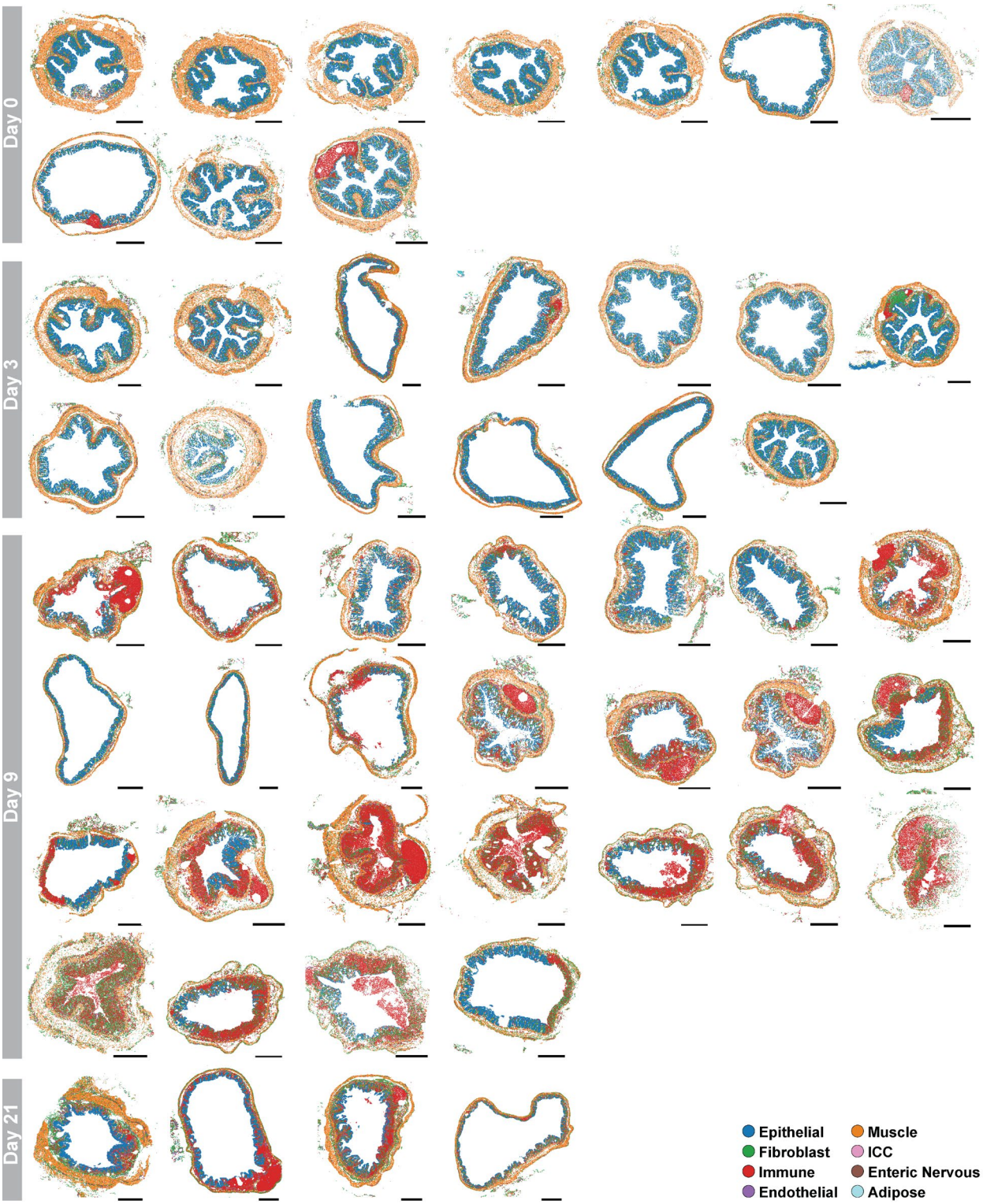


**Figure S2. The spatial distribution of cell classes in all measured slices**

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.08.539701; this version posted May 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Each dot represents the center of a cell, and the color indicates the cell class assignment. Slices are grouped in rows by the disease stage from which they were collected. Scale bars: 500 μm.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.08.539701; this version posted May 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

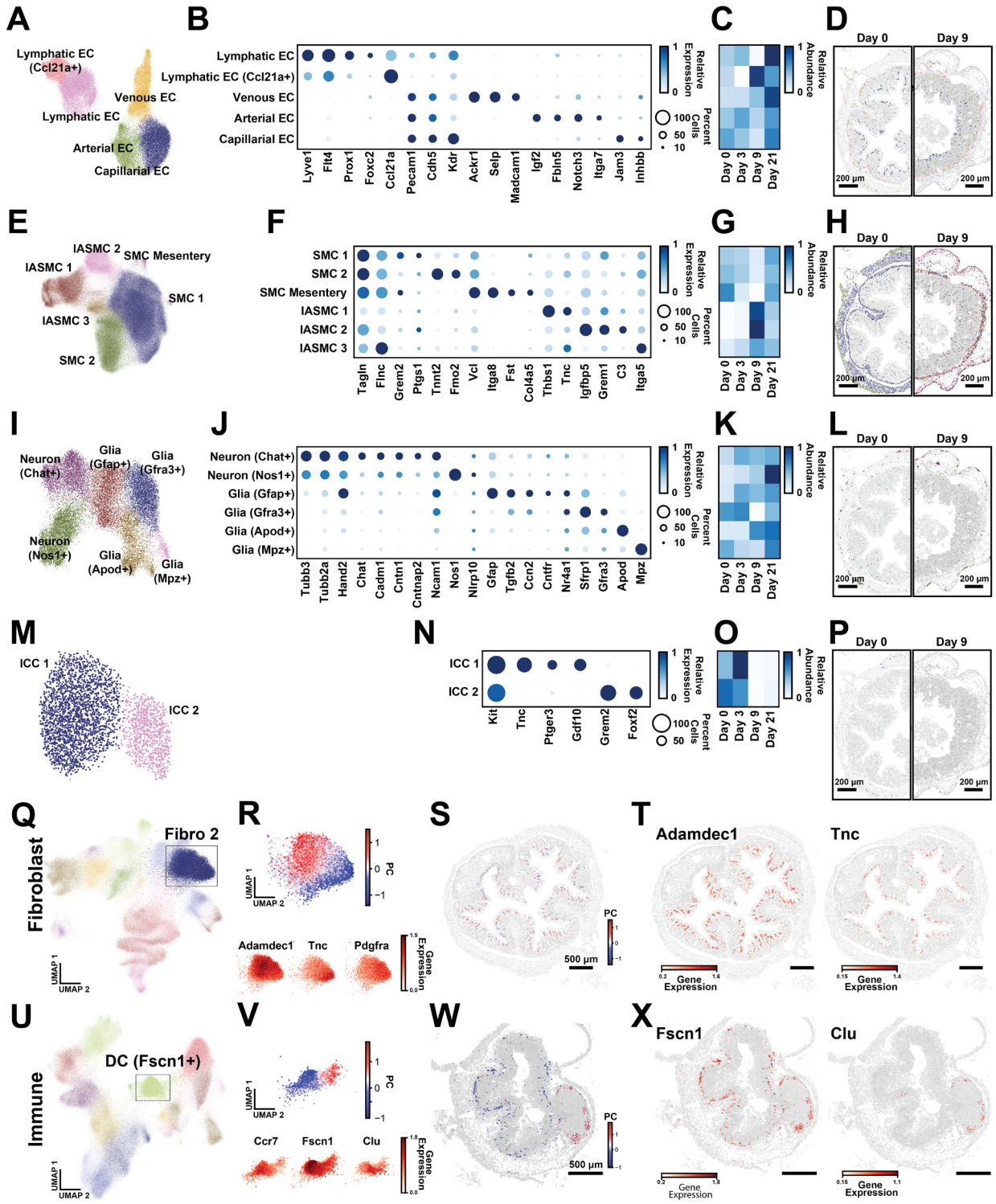


**Figure S3. Gene expression and spatial distributions for sub-populations within major cell classes identified with MERFISH.**

(A) UMAP representation of the endothelial cells (EC), colored by identified clusters.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.08.539701; this version posted May 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(B) Dot plot representation of key marker genes for the populations of EC. The color indicates the relative expression of each gene across all listed populations, and the size of the marker indicates the fraction of cells that express at least one copy of that RNA.

(C) The fractional abundance of each of the identified EC populations across disease stage, normalized to the maximum abundance observed at all disease stages.

(D) The spatial distribution of EC in representative slices taken from Day 0 and Day 9. Markers represent the center of cells. Gray represents all cells, and all EC are colored corresponding to their cluster identities as shown in the UMAP in (A). Scale bars: 200 μm.

(E-H) As in (A-D) but for smooth muscle cells (SMC).

(I-L) As in (A-D) but for cells of the enteric nervous system.

(M-P) As in (A-D) but for interstitial cells of Cajal (ICC).

(Q) UMAP representation of fibroblasts cells highlighting the cells assigned to Fibro 2.

(R) Zoom in on the boxed region of the UMAP in (Q) with cells colored by the amplitude of a principal component (PC) associated with a PC analysis of gene expression within cells assigned to Fibro 2 (top) and colored by the expression level of three genes that contribute to the loading associated with this PC (bottom).

(S) Spatial distribution of Fibro 2 cells within a representative slice. All cells are colored gray, and all cells assigned to Fibro 2 are colored by the amplitude of the PC as seen in (B). Scale bar: 500 μm.

(T) Spatial distribution of Fibro 2 cells colored by the expression of two genes that define the loading associated with the highlighted PC in (R). Scale bar as in (S).

(U-X) As in (R-T) but for DC (Fscn1+).

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.08.539701; this version posted May 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

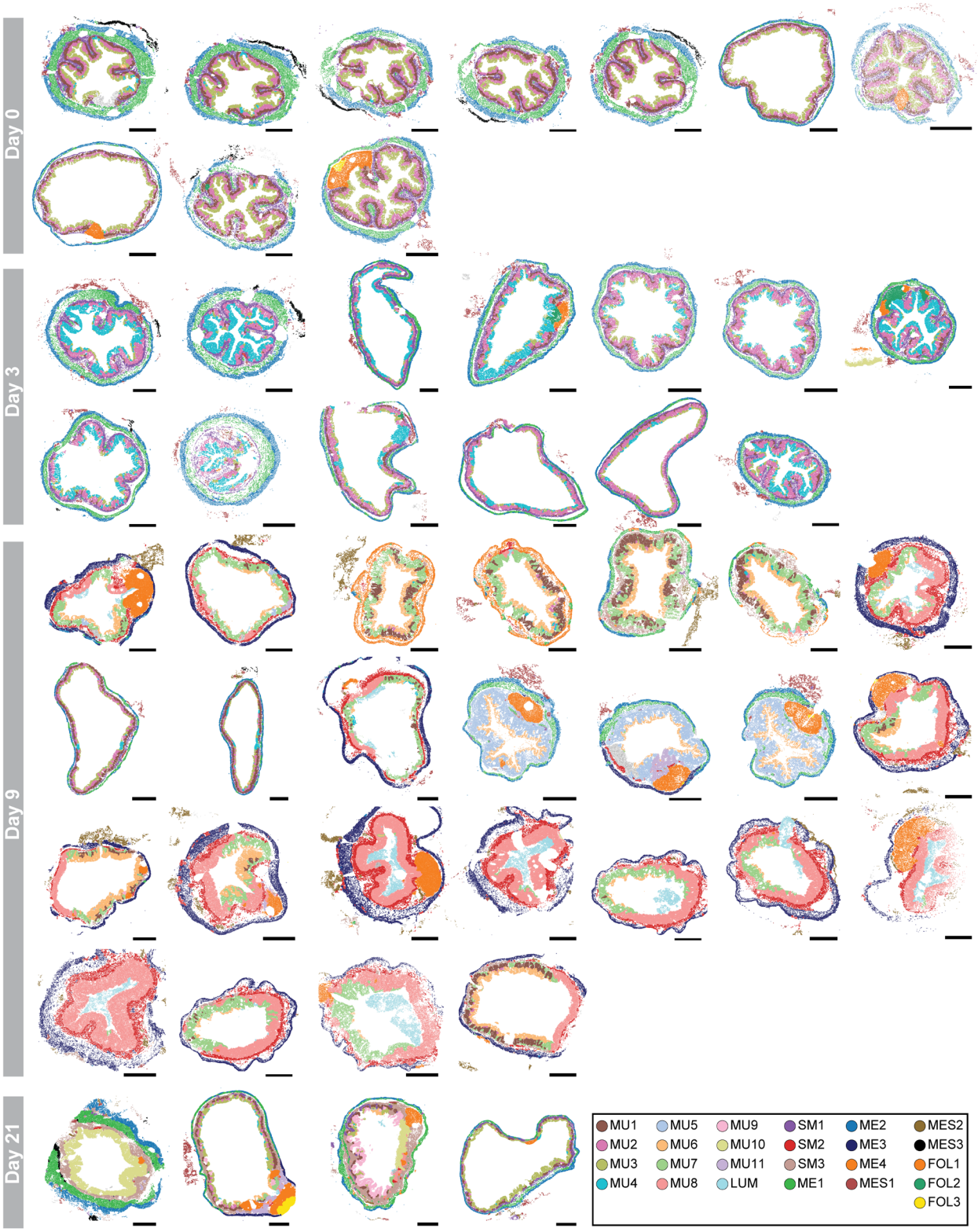


**Figure S4. The spatial distribution of cellular neighborhoods in all slices**

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.08.539701; this version posted May 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Each dot represents the center of individual cells and color indicates the neighborhood assignment. The rare gray cells represent the 0.5% of cells not assigned to one of our neighborhoods. Slices are grouped in rows by the disease stage from which they were collected. Scale bars: 500 μm.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.08.539701; this version posted May 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

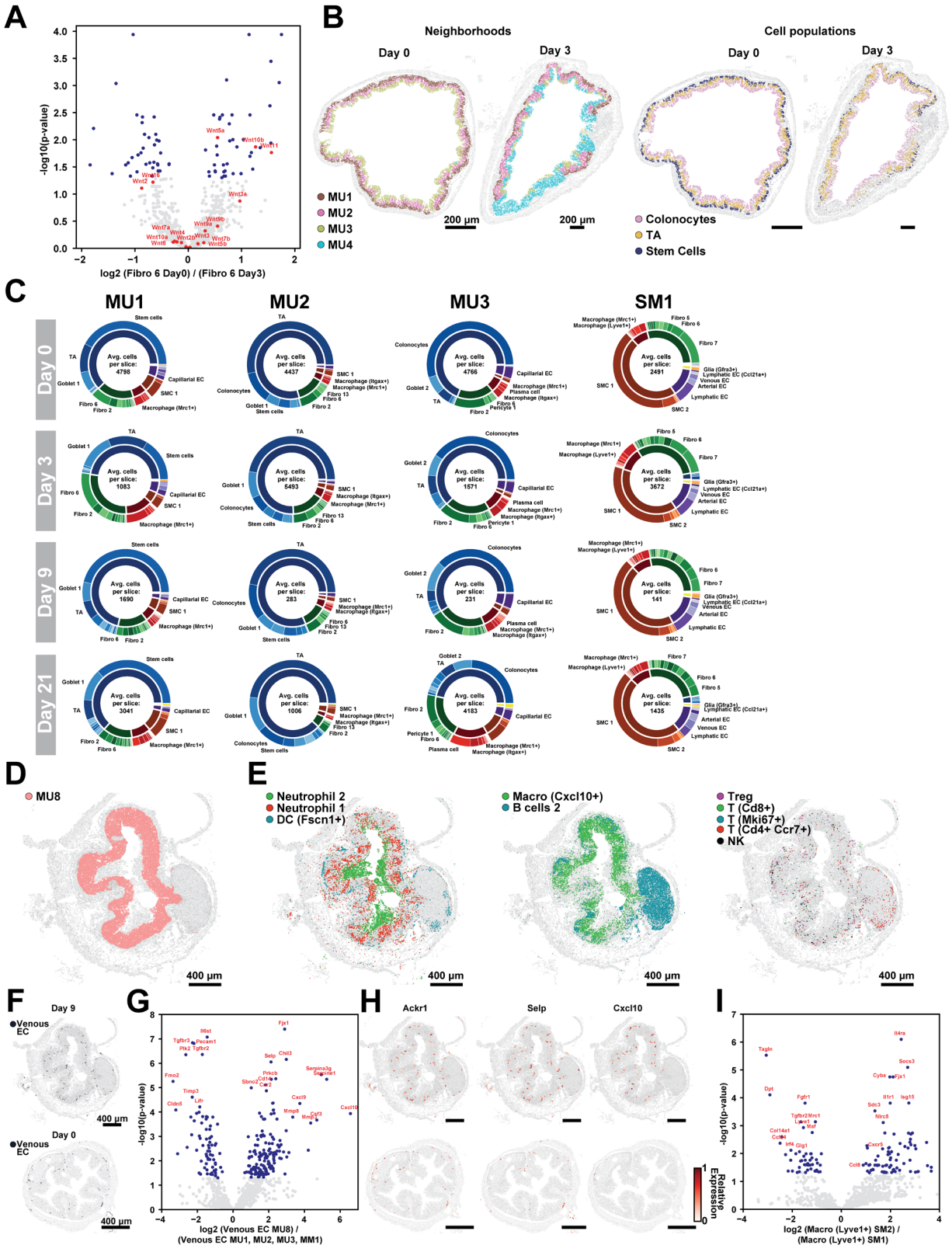


**Figure S5. Additional properties of cellular neighborhoods**

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.08.539701; this version posted May 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(A) Volcano plot of the differential expression of genes in Fibro 6 cells seen at Day 3 versus that seen at Day 0. Genes highlighted in blue pass a false discovery threshold of 5% while named genes marked in red represent all Wnt-family members.

(B) The spatial distribution of cells in a representative Day 0 and Day 3 slice with cells colored by the cellular neighborhood to which they were assigned (left) or the cell population to which they were assigned (right). Scale bars: 200 μm.

(C) Pie charts representing the abundance of cell types in MU1, MU2, MU3, and SM1 for only the cells associated with Day 0, Day 3, Day 9, or Day 21. Pie charts are displayed as described in **Figure 3**.

(D, E) The spatial distribution of cells in a representative Day 9 slice colored by cells assigned to MU8 (left) or major clusters found in MU8 (right). Scale bars: 400 μm.

(F) The spatial distribution of cells in a representative Day 9 (top) or Day 0 (bottom) slice with Venous EC cells in blue. Scale bars: 400 μm.

(G) Volcano plot of the expression of all genes within Venous EC in MU8 or that in MU1, MU2, MU3, or MM1 as in (A).

(H) The spatial distribution of all cells in the representative slices shown in (F). Only Venous EC cells are colored, and they are colored based on the relative expression of the listed genes.

(I) Volcano plot of the differential expression of genes in Lyve1+ macrophages see in SM2 versus those seen in SM1.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.08.539701; this version posted May 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

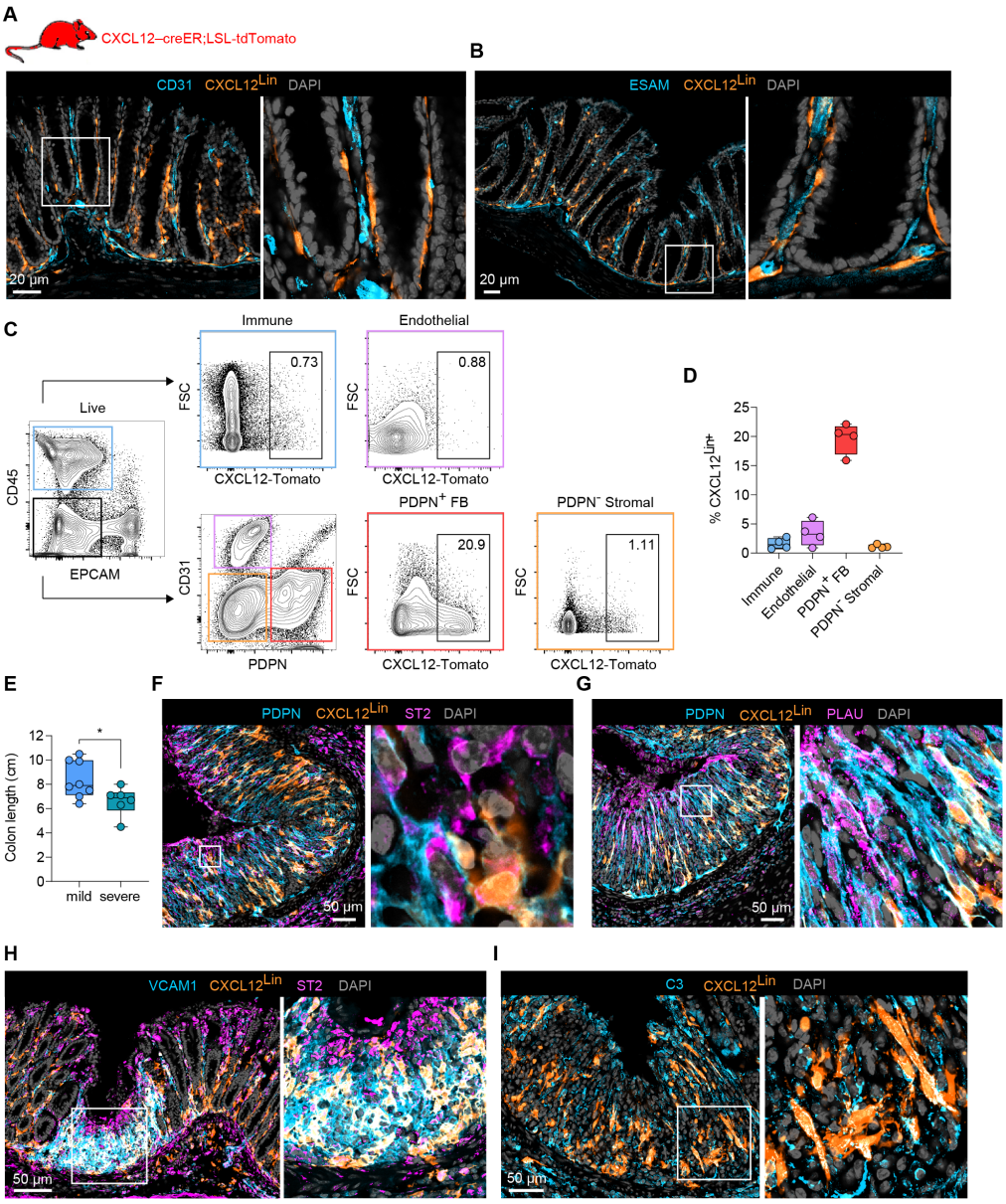


**Figure S6. Lineage tracing reveals distinct origins of inflammation-associated fibroblasts.**

(A, B) Representative immunofluorescence images of distal colon from healthy CXCL12$^{Lin}$ mice (CXCL12-creER; LSL-tdTomato; *n*=5) for CD31 (A) or ESAM (B). Scale bars: 20 µm.

(C, D) Flow cytometry analysis of colonic CXCL12-Tomato expression in immune (CD45+), endothelial (CD31+), PDPN+ fibroblasts (PDPN+ +CD45-CD31-

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.08.539701; this version posted May 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

EpCAM-) and PDPN- stromal (PDPN-CD45-CD31-EpCAM-) cells at steady-state (*n*=6).

(E) Colon length measured for the $CXCL12^{Lin}$ mice sorted into groups that had mild or severe degrees of weight loss as measured at Day 9 (*n*=14). The center line represents mean, the width of the bars represents standard deviation, and the whiskers extend to include 95% of the data. Individual markers represent distinct mice. *p<0.05.

(F-I) Representative immunofluorescence images of inflammatory regions in distal colon from mice, harvested on Day 9 (n=6). The degree of overlap between the $CXCL12^{Lin}$ marker and other markers of IAF populations, ST2 (F), PLAU (G), VCAM1 (H) and C3 (I), are highlighted. Scale bars: 50 μm.

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.08.539701; this version posted May 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

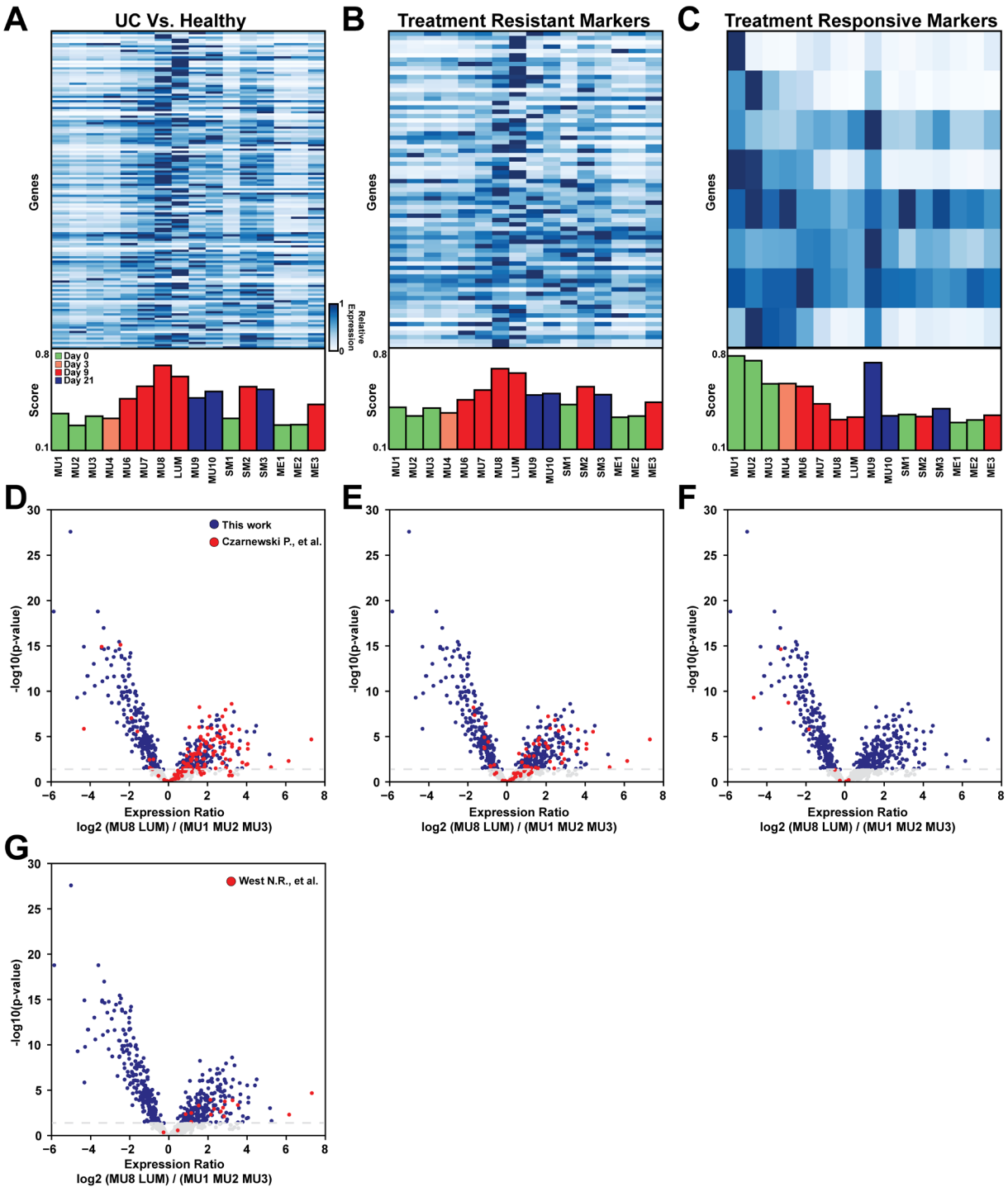


**Figure S7. Human biomarker panel expression in mouse neighborhoods supports the presence of similar neighborhoods in human ulcerative colitis**

(A-C) The average expression of the mouse homologs of the human biomarkers reported by Czarnewski et. al.[84] to distinguish UC from healthy samples (A), anti-

bioRxiv preprint doi: https://doi.org/10.1101/2023.05.08.539701; this version posted May 9, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TNF resistant UC samples from anti-TNF sensitive samples (B), anti-TNF sensitive samples from anti-TNF resistant samples (C), within the listed mouse neighborhoods (top) or the average biomarker score (as in **Figure 7**) for these neighborhoods (bottom).

(D-G) Volcano plot of the differential expression of mouse genes observed in the MU8 and LUM neighborhoods relative to that observed in the healthy neighborhoods, MU1, MU2, and MU3. Blue marks all genes that pass a false-discovery-rate threshold of 5% (dashed gray line), and red marks the mouse homologs of the human biomarkers reported by Czarnewski et. al.[84] to distinguish UC from healthy samples (D), anti-TNF resistant UC samples from anti-TNF sensitive samples (E), anti-TNF sensitive samples from anti-TNF resistant samples (F), or from the anti-TNF-resistant UC biomarkers reported by West et al.[72].