

## **Myeloid cell influx into the colonic epithelium is associated with disease severity and non-response to anti-Tumor Necrosis Factor Therapy in patients with Ulcerative Colitis.**

Divya Jha<sup>1,2</sup>, Zainab Al-Taie<sup>3,4\*</sup>, Azra Krek<sup>5,\*</sup>, Shadi Toghi Eshghi<sup>6,7</sup>, Aurelie Fantou<sup>8</sup>, Thomas Laurent<sup>8</sup>, Michael Tankelevich<sup>1,2</sup>, Xuan Cao<sup>5</sup>, Hadar Meringer<sup>1,2</sup>, Alexandra E Livanos<sup>1,2</sup>, Minami Tokuyama<sup>1,2</sup>, Francesca Cossarini<sup>1,2</sup>, Arnaud Bourreille<sup>8</sup>, Regis Josien<sup>8</sup>, Ruixue Hou<sup>3,4</sup>, Pablo Canales-Herrerias<sup>1,2</sup>, Ryan C. Ungaro<sup>1</sup>, Maia Kayal<sup>1</sup>, James Marion<sup>1</sup>, Alexandros D Polydorides<sup>9</sup>, Huaibin M. Ko<sup>10</sup>, Darwin D'souza<sup>11</sup>, Raphael Merand<sup>11</sup>, Seunghee Kim-Schulze<sup>11</sup>, Jason A. Hackney<sup>6,7</sup>, Allen Nguyen<sup>6,7</sup>, Jacqueline M. McBride<sup>6,7</sup>, Guo-Cheng Yuan<sup>5</sup>, Jean Frederic Colombel<sup>1</sup>, Jerome C. Martin<sup>8,#</sup>, Carmen Argmann<sup>3,4,#</sup>, Mayte Suárez-Fariñas<sup>3,4,#</sup>, Francesca Petralia<sup>5,#</sup>, and Saurabh Mehandru<sup>1,2,#</sup>

\* Equal contribution

### **Affiliations**

<sup>1</sup>Henry D. Janowitz Division of Gastroenterology, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

<sup>2</sup>Precision Institute of Immunology, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

<sup>3</sup>Department of Population Health Science and Policy, Icahn School of Medicine at Mount Sinai, New York, NY, USA

<sup>4</sup>Icahn Institute for Data Science and Genomic Technology, New York City, NY, USA

<sup>5</sup>Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, New York, NY, USA.

<sup>6</sup>Biomarker Discovery, OMNI, Genentech Inc. South San Francisco, CA, USA

<sup>7</sup>OMNI Biomarker Development, Genentech Inc. South San Francisco, CA, USA

<sup>8</sup>Université de Nantes, Inserm, CHU Nantes, Centre de Recherche en Transplantation et Immunologie, UMR 1064, ITUN, F-44000 Nantes, France

<sup>9</sup>Department of Pathology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

<sup>10</sup>Department of Pathology and Cell Biology, Columbia University Medical Center-New York Presbyterian Hospital, New York, New York.

<sup>11</sup>Human Immune Monitoring Core, Precision Institute of Immunology, Icahn School of Medicine at Mount Sinai, New York, NY, USA.

### **Address correspondence to**

[saurabh.mehandru@mssm.edu](mailto:saurabh.mehandru@mssm.edu)

[francesca.petralia@mssm.edu](mailto:francesca.petralia@mssm.edu)

[maytesuarezfarinas@me.com](mailto:maytesuarezfarinas@me.com)

[carmen.argmann@mssm.edu](mailto:carmen.argmann@mssm.edu)

[jerome.martin@univ-nantes.fr](mailto:jerome.martin@univ-nantes.fr)

Lead Corresponding Author: Saurabh Mehandru, MD

## Abstract

Ulcerative colitis (UC) is an idiopathic chronic inflammatory disease of the colon with sharply rising global prevalence. Dysfunctional epithelial compartment (EC) dynamics are implicated in UC pathogenesis although EC-specific studies are sparse. Applying orthogonal high-dimensional EC profiling to a Primary Cohort (PC; n=222), we detail major epithelial and immune cell perturbations in active UC. Prominently, reduced frequencies of mature *BEST4*<sup>+</sup>*OTOP2*<sup>+</sup> absorptive and *BEST2*<sup>+</sup>*WFDC2*<sup>+</sup> secretory epithelial enterocytes were associated with the replacement of homeostatic, resident *TRDC*<sup>+</sup>*KLRD1*<sup>+</sup>*HOPX*<sup>+</sup>  $\gamma\delta$ <sup>+</sup> T cells with *RORA*<sup>+</sup>*CCL20*<sup>+</sup>*S100A4*<sup>+</sup> T<sub>H17</sub> cells and the influx of inflammatory myeloid cells. The EC transcriptome (exemplified by *S100A8*, *HIF1A*, *TREM1*, *CXCR1*) correlated with clinical, endoscopic, and histological severity of UC in an independent validation cohort (n=649). Furthermore, therapeutic relevance of the observed cellular and transcriptomic changes was investigated in 3 additional published UC cohorts (n=23, 48 and 204 respectively) to reveal that non-response to anti-Tumor Necrosis Factor (anti-TNF) therapy was associated with EC related myeloid cell perturbations. Altogether, these data provide high resolution mapping of the EC to facilitate therapeutic decision-making and personalization of therapy in patients with UC.

**Key words:** Ulcerative colitis, intraepithelial compartment, RNA sequencing, spatial transcriptomics, anti-TNF treatment response

## Main

Ulcerative colitis (UC) involves epithelial barrier dysfunction in combination with immune dysregulation that leads to chronic inflammation of the large intestines<sup>1</sup>. In addition to reinforcing the physical barrier, the intestinal epithelium regulates nutrient uptake and facilitates tissue homeostasis<sup>2,3</sup>. Murine studies have delineated bi-directional epithelial-intraepithelial T cell interactions<sup>4</sup>, as being critical for response to infection<sup>5</sup>, nutrient sensing<sup>6</sup>, and regulation of gut inflammation<sup>7</sup>. Additionally, epithelium-proximate myeloid cells, especially macrophages regulate epithelial stem cells<sup>8-10</sup>, epithelial viability<sup>11</sup> and host-microbiota interactions<sup>12</sup> to reinforce the epithelial-barrier function. Despite their critical role in tissue homeostasis, human intestinal epithelium-associated myeloid cells are poorly defined.

The therapeutic landscape of UC has significantly expanded with the advent of therapies targeting tumor necrosis factor (TNF)<sup>13</sup>, integrin  $\alpha 4\beta 7$ <sup>14</sup>, IL-12/23 pathways<sup>15</sup>, JAK-STAT pathways<sup>16,17</sup> and sphingosine 1 phosphate (S1P) receptors<sup>18</sup>, with additional therapeutic options poised to become available in the near future<sup>19</sup>. However, approximately 30% of patients are non-responsive to initial therapy and an additional 50% become secondary non-responders<sup>20</sup>. To break this ‘therapeutic ceiling’, a ‘precision medicine-like’, drug-mechanism and biomarker-based approach is suggested. However, this strategy is hampered by the lack of mechanistic data to guide therapeutic decisions<sup>21</sup>. High dimensional approaches such as single cell RNA sequencing (scRNA-seq) have enabled profiling of intestinal immune cells at unprecedented resolution<sup>22-28</sup>. However, prior studies have focused on whole intestinal tissues or on the intestinal lamina propria (LP) and to-date, dedicated epithelial compartment-focused studies that relate epithelial-immune perturbations to therapeutic outcomes in patients with UC are lacking.

Here, using bulk RNA sequencing, scRNA-seq, Spatial Transcriptomics (ST), and complemented by relevant protein-level assessments, we have focused on epithelium-specific immune and non-immune alterations in patients with UC. Further, we have defined epithelium-specific pathotypes that associate with non-response to TNF inhibitors (TNFi). Taken together, our study comprehensively evaluates the role of EC in active UC and provide a rational basis for therapeutic choices in patients who are non-responsive to TNF-inhibition.

## Results

### **Intraepithelial immune compartment is extensively remodeled in active UC.**

Bulk RNA sequencing was performed on ethylenediaminetetraacetic acid (EDTA)-dissociated epithelial fraction (Epi<sub>bulk</sub>) or EDTA-dissociated, percoll-enriched intraepithelial immune cells (IIC, IIC<sub>bulk</sub>) using biopsies derived from histologically uninfamed and inflamed colon of subjects with

UC or from Normal Volunteers (NV; Fig. 1a, Extended Data Fig. 1, Methods, Supplementary Table 1). Inflamed UC samples (red) clustered distinctly from uninfamed UC (green) and NV (blue) samples for both  $Epi_{bulk}$  and  $IIC_{bulk}$  (Fig. 1b). We identified differentially expressed (DEGs) between inflamed, non-inflamed and NV within EPI and IIC (DEGs; false discovery rate (FDR) <0.05; fold change (FCH) >2.0) (Fig. 1b, Supplementary Table 2). In contrast, no significant DEGs were detected in either  $Epi_{bulk}$  or  $IIC_{bulk}$  when comparing uninfamed UC and NV samples. When comparing the  $Epi_{bulk}$  and  $IIC_{bulk}$  inflammation signatures (hereby refer to as  $IIC_{inf}$  and  $Epi_{inf}$  signatures respectively), significant overlap was noted, with 989 DEGs (562 shared, up- and 427 down-regulated genes, Fisher's exact test  $p < 2.2 \times 10^{-16}$  for both; Fig. 1d, Supplementary Table 2a,b). Next, to characterize the cell types associated with the transcriptome of  $Epi_{inf}$  ( $Epi_{bulk}$ -derived DEGs, inflamed vs. uninfamed) and  $IIC_{inf}$  ( $IIC_{bulk}$ -derived DEGs, inflamed vs. uninfamed), we performed an over-representation analysis (ORA) using a curated scRNA-seq derived atlas for cell types in the human colonic tissues<sup>23</sup>. Up-regulated  $Epi_{inf}$  and  $IIC_{inf}$  genes were significantly ( $adjP < 0.01$ ) enriched for monocytes, macrophage, dendritic cells, T cells and plasma cells suggesting an infiltration of immune cells in the inflamed epithelia while down-regulated  $Epi_{inf}$  and  $IIC_{inf}$  DEGs were enriched for epithelial cells (Fig.1e). Consistent with the cell type enrichments, ORA of KEGG<sup>29</sup> pathways revealed that the up-regulated  $Epi_{inf}$  and  $IIC_{inf}$  DEGs were involved in several inflammation-associated pathways notable for IL-17 signaling,  $T_{H17}$  cell differentiation, cytokine-cytokine receptor interactions, antigen presentation and immunoglobulin production, while the down-regulated  $Epi_{inf}$  and  $IIC_{inf}$  genes were found to be predominantly associated with pathways such as retinol and butanoate metabolic pathways indicative of epithelial cell dysfunction in active UC (Fig. 1f, Supplementary Table 2c). A subset of the  $Epi_{inf}$  and  $IIC_{inf}$  DEGs, found in the intersection of Fig.1d and associated with these pathways, is depicted in Fig. 1g. Notable for the upregulation of immune cell related transcripts are *OSM*, *IL1B*, *CXCL8*, *CCL24*, *IL17A*, *CD4*, *SAA1*, *FOXP3*, *CD14*, *MRC1*, *MNDA*, *LYZ*, *TREM1*, *CSF3R*, *IGHG* and *JCHAIN*, while epithelial cell-expressed genes including *PRAP1*, *AQP7*, *GUCA2A*, *DEFB1*, *RARB*, *SYNPO* and innate immune system-related nociceptor gene *TRPV1* were significantly down-regulated<sup>30-34</sup> (Fig. 1g, Supplementary Table 2d). Altogether, these data pointed to the influx of non-homeostatic inflammatory cells in the EC of inflamed UC, associated with the replacement of homeostatic epithelial programs with powerful pro-inflammatory programs.

Having seen a significant overlap of  $Epi_{inf}$  and  $IIC_{inf}$  gene signatures, we studied the association of  $IIC_{inf}$  signature in an independent Validation Cohort (VC-1) consisting of 406 patients with UC and 243 NV with available bulk RNA-seq data on whole biopsies (Extended Data Fig. 1, Methods). Using Gene Set Variation Analysis (GSVA), we noted significantly higher

expression of the up-regulated, primary cohort (PC)-derived IIC<sub>inf</sub> genes in inflamed intestinal biopsies from VC-1 compared to NV and uninfamed UC biopsies (Fig. 1h, Extended Data Fig. 2a). Furthermore, the expression of up-regulated IIC<sub>inf</sub> genes strongly associated with higher clinical and endoscopic disease severity and was positively correlated with histologic severity in inflamed biopsies from VC-1 ( $r=0.601$ ,  $p=1.05 \times 10^{-16}$  Fig. 1h, Extended Data Fig. 2b-d). The top genes that correlated with histologic severity (Nancy score<sup>35</sup>) included proinflammatory genes *SAA1*, *S100A8*, *S100A9*; myeloid cell-associated genes *TREM1*, *LILRA5*; neutrophil chemotactic genes *FPR1*, *FPR2*, *CXCL1*, *CXCR1*, *CXCR2*, *FCGR3B/CD16*; hypoxia related gene *HIF1A*; autophagy gene *DRAM1* and unfolded protein response gene *XBP1* (Fig. 1j)<sup>36-39</sup>. The findings from our EC-bulk transcriptomic data revealed a complex, myeloid cell-dominated milieu in UC-associated inflammation that significantly correlated with parameters of clinical, endoscopic, and histologic severity.

### **High-resolution mapping of immune cells reveals distinct cellular composition in inflamed and uninfamed colonic epithelial compartment of UC patients.**

To further define UC-associated cellular and transcriptomic changes within the IIC (percol-enriched EC), scRNA-seq was performed on 9 NV and 9 UC patients with active inflammation, stratified as mild-moderate ( $n=5$ ) or severe ( $n=4$ ) inflammation. In 5 NV and 7 UC patients, scRNA-seq was also performed on LP-derived mononuclear cells (methods) to capture compartment specific molecular and cellular perturbations between EC and LP from the same patient. We sequenced a total of 108,671 cells, with a mean of 7,698 unique molecular identifiers (UMIs) of 1437 unique genes per cell, on average (Supplementary Table 3a). Unsupervised clustering based on transcriptomic profile of scRNA-seq data resulted in 32 distinct cell clusters spanning across NV and UC (Fig. 2a), as well as IIC and LP (Fig. 2b). Grouping the clusters by expression profile similarities indicated the presence of 10 major cell types (Fig. 2c). The transcriptomic profiles confirmed the distribution at the single-cell level of UMIs corresponding to well-established immune lineage marker genes of myeloid cells (*CD14*, *CD68*, *MNDA*), plasma cells (*JCHAIN*, *XBP1*, *MZB1*), B cells (*CD79A*, *CD79B*, *MS4A1*, *BANK1*), T cells (*TRAC*, *TRBC1*, *CD3D*), innate lymphoid cells (*IL7R*, *KLRB1*, *IL1R1*, *KIT*), mast cells (*CPA3*, *CD63*, *TPSAB1*), as well as markers of epithelial cells (*EPCAM*, *KRT8*, *CD9*, *FABP1*), endothelial cells (*VWF*, *PECAM1*, *CD36*, *CD151*), pericytes (*ACTA2*, *COL3A1*, *ADAMDEC1*) and fibroblasts (*COL1A2*, *ACTR2*, *C1S*). We observed higher proportions of myeloid cells, plasma cells, mast cells, T cells and stromal cells (fibroblasts, endothelial cells and pericytes) along with lower proportion of epithelial cells and ILCs in UC samples compared to NV samples. Compartment-wise overview

revealed enrichment of epithelial cells, myeloid cells, B cells and T cells in IIC samples compared to LP enriched populations of stromal cells, plasma cells and mast cells in LP samples (Fig. 2d, Supplementary Table. 3b-d). On enumerating the cell type frequencies in IIC and LP compartments stratified by disease severity (Fig. 2e, f), we observed a significant increase in the proportion of myeloid cells and plasma cells with a concomitant decrease in the epithelial cells and ILCs when comparing patients (Wilcoxon signed-rank test  $p < 0.05$ , Fig. 2e) in IIC. Expectedly, within the stromal cells (endothelial cells, fibroblasts and pericytes) comprised a larger fraction in LP compared to IIC and LP-fibroblasts were significantly increased in severe UC compared to normal tissue (Wilcoxon signed-rank test  $p < 0.05$ , Fig. 2f, Supplementary Fig. 1a).

### **Active UC is associated with depletion of homeostatic $\gamma\delta^+$ T cells and influx of proinflammatory $T_{H17}$ cells into the EC.**

We detailed T cell-associated changes by selecting T cells *in silico* and re-clustered them using the nearest neighborhood unsupervised clustering algorithm (methods). Ten T cell sub-clusters with variable cell numbers (T0-T9, 749-6091 cells) were detected (Fig. 3a). The scRNA-seq dataset showed an enrichment of T cell specific genes in T cell subclusters such as T0, T4, T6, T7 and T8 (*CD4*, *CD3D*, *TRAC*, *TRBC1*), T1 and T3 (*CD8A*, *CD8B*, *CD3D*, *TRAC*, *TRBC2*); T2 and T9 (*CD8A*, *CD8B*, *TRDC*, *TRGC2*) and T5 (*TRAC*, *TRBC1*) (Fig. 3b). IELs (IIC related T cells hitherto referred to as intraepithelial lymphocytes/IELs) T1 ( $p=0.0056$ ) and T2 ( $p=0.0056$ ) and LP associated T0 ( $p=0.032$ ) were significantly decreased whereas T4 and T8 were significantly increased in severe UC samples in both IIC (T4,  $p=0.0028$ ; T8,  $p=0.0028$ ) and LP (T4,  $p=0.016$ ; T8,  $p=0.016$ ;  $p$ -value from Wilcoxon signed-rank test  $< 0.05$ , Fig. 3b,c, Extended Data Fig. 3a).

Based on canonical T cell markers, the sub-clusters were regrouped into cell types and annotated as  $T_{H17}$  (T4), cytotoxic  $CD8^+$  (T3), naïve (T9), Regulatory ( $T_{REG}$ , T6, T8),  $\gamma\delta^+$  (T1, T2) and memory (T0, T5, T7) T cells (Fig. 3d, Supplementary Table 4). Both IIC- and LP- associated  $T_{H17}$  cells were significantly increased in severe UC compared to NV (IIC,  $p=0.0028$ ; LP,  $p=0.032$ ) whereas  $T_{REG}$  cells were significantly increased only in IIC ( $p=0.0028$ ) and not LP ( $p=0.29$ ). In contrast,  $\gamma\delta^+$  IELs and LP specific memory T cells were significantly reduced in severe UC compared to NV (IIC,  $p=0.05$ ; LP,  $p=0.016$ ; Fig. 3d, Extended Data Fig. 3b, Supplementary Fig. 1b, c). Multiparameter flow cytometry (Supplementary Fig. 2a) confirmed a significant decrease of the frequency of EC-associated  $\gamma\delta^+$  IELs ( $p < 0.0001$ ; Fig. 3e, f). Among  $\alpha\beta^+$  IELs, a significant increase in  $CD4^+\alpha\beta^+$  IELs ( $p < 0.0001$ ) and  $CD4:CD8$  ratio ( $p < 0.0001$ ) was noted in the EC of patients with active UC (Fig. 3g, h, Extended Data Fig. 3c). Immunofluorescence (IF, Fig. 3 i,j) and immunohistochemical staining (IHC, Extended Data Fig. 3d,e) further confirmed an increase

of CD4<sup>+</sup> IELs ( $p < 0.0001$ ) in active UC patient colonic tissues. Gene Set Enrichment Analysis (GSEA) on gene level fold changes identified several UC-enriched pathways notable for T cell activation (*HLA-DRB1*, *CD3D*), cytolytic activity (*GZMB*, *PRF1*), MHC class-I antigen presentation (*TAP1*) and pro-inflammatory cytokines (*STAT1*, *IL2RG*, *JAK3*, *JAK1*, *STAT3*, *PRF1*, *IL2RG*, *JAK1*) (Extended Data Fig. 3f) suggesting an elevated secretion and activation status of IELs in the inflamed UC compared to NV samples. Altogether, a substantial remodeling of IEL was noted with a reduction in homeostatic  $\gamma\delta^+$  IELs and a concomitant increase in T<sub>H17</sub> and T<sub>REG</sub> subsets in EC.

To understand the pathogenic significance of EC-associated T<sub>H17</sub> cells, we cultured normal colonic biopsy tissue with T<sub>H17</sub> cell-associated cytokines<sup>32</sup>, using either IL-17 alone, IL-22 alone or IL-17 in combination with IL-22 (Extended Data Fig. 4a). Through subsequent RNA sequencing, we identified 119 DEGs (FCH>1.3, FDR<0.05, compared to unstimulated) upon IL-17+IL-22 stimulation (Extended Data Fig. 4b, hitherto referred to IL17+IL22 signature, Supplementary Table 5). Minimal DEGs were observed with either cytokine in monoculture. The IL17+IL22 signature genes mapped (using a published dataset<sup>23</sup>) mostly to epithelial cells (52%) with minor contributions from other cell types, including immune (1.2%) and stromal cells (1.2%) (Extended Data Fig. 4c). Up-regulated IL17+IL22 signature could be categorized into antimicrobial peptides (AMPs; *DEFB4A*, *LCN2*)<sup>40,41</sup>, chemokines (*CXCL1*, *CXCL3*), intestinal transporters (*SLC30A*, *SLC25A37*)<sup>42</sup> and immune regulators (*NFKBIZ*, *TNIP3*)<sup>43,44</sup> (Extended Data Fig. 4d). We next assessed the relevance of the ex-vivo derived IL17+IL22 signature to IBD-associated inflammation. In colonic biopsies from VC-1 patients, a significant positive correlation between the gene expression levels of IL17 and IL22 with GSVA scores of the IL17+IL22 signature were observed for inflamed biopsies (Extended Data Fig. 4e). GSVA-scores for the IL-17+IL-22 signature were significantly increased in inflamed colonic biopsies ( $p = 3.81 \times 10^{-78}$ ) compared to uninflamed biopsies ( $p = 2.45 \times 10^{-6}$ ) from VC-1 UC patients (Extended Data Fig. 4f), an association driven mainly by myeloid cell recruiting chemokine transcripts (*CXCL1*, *CXCL2*, *CXCL3*) and T<sub>H17</sub>- and T<sub>REG</sub>-recruiting chemokine transcripts (*CCL20*, *CCL28*) (Extended Data Fig. 4g). Notably, IL17+IL22 signature scored higher than IL17 ( $d = 1.21$ ,  $p < 2.2 \times 10^{-16}$ , paired t test) and IL22 ( $d = 2.6$ ,  $p < 2.2 \times 10^{-16}$ ) monocultured signatures in inflamed biopsies, suggesting a synergistic effect of those two cytokines in UC pathogenesis. Furthermore, we observed consistent transcriptional changes between the ex vivo derived IL-17+IL-22 signature and inflammation-associated EC genes shared between the Epi<sub>inf</sub> and IIC<sub>inf</sub> signatures (Fig. 1g, Extended Data Fig. 4h) with 19 genes overlapping with the IL-17+IL-22 up-(OR=18.3,  $p = 6.3 \times 10^{-}$

<sup>16</sup>, FCH>2, Fisher's exact test) and 6 down-regulated genes (OR=20.5, p=0.008, FCH>2, Fisher's exact test). Altogether, these data point to a T<sub>H17</sub> cell-driven inflammatory reprogramming of the EC and upregulation of myeloid/immune cell-recruiting chemokines.

### **Active UC is associated with a major influx of pro-inflammatory myeloid cells into the EC.**

Having observed the potential for T<sub>H17</sub> cell-derived cytokines to recruit myeloid cells, we further elaborated on the myeloid cells in the scRNA-seq data and re-clustered them using the nearest neighborhood unsupervised clustering algorithm (Fig. 4a, methods) to derive 10 sub-clusters with variable cell numbers (M0-M9, 142-1099 cells, Supplementary Table 6). The transcriptomic profiles of the myeloid sub-clusters were distinct, and their distribution varied within the two intestinal compartments- IIC and LP. IIC-associated M0, M1, M2 and M3 were significantly higher in severe UC compared to NV (p = 0.0021; p= 0.0065; p= 0.0069; p= 0.0066); while M5, M6, M7 and M8 were significantly lower in severe UC compared to NV (p = 0.0028; p=0.0069; p= 0.0028; p=0.0028). The same differences between severe UC and NV were noted in the LP except for sub-clusters M1, M5 and M7 (Fig. 4b,c; Extended Data Fig. 5a, Wilcoxon signed-rank test p< 0.05). Based on canonical markers, myeloid sub-clusters were coalesced and annotated into 4 cell types: neutrophils (*CSF3R*), monocytes-macrophages (MoMac) (*EREG*, *MNDA*), macrophages (*MRC1*, *C1QA/B/C*) and dendritic cells (DCs, *LGALS2*) (Supplementary Table 6). Neutrophils and MoMac were significantly increased in the severe UC compared to NV both in IIC (p=0.0066; p= 0.0028) and LP (p=0.018; p=0.016) while macrophages were significantly increased only in the IIC (p=0.034) in the severe UC compared to NV (Fig. 4d, Extended Data Fig. 5b, Supplementary Fig. 1d,e). Using multi-parameter flow cytometry (Supplementary Fig. 2b), we observed a significant increase in the frequency of neutrophils (CD45<sup>+</sup>CD3<sup>-</sup>CD19<sup>-</sup>CD16<sup>hi</sup>CD66b<sup>+</sup> cells; p<0.0001) and CD14<sup>hi</sup> MoMac (live CD45<sup>+</sup>CD3<sup>-</sup>CD19<sup>-</sup>CD14<sup>hi</sup>; p<0.0001) in the EC of inflamed UC compared to the uninfamed UC and NV (Fig. 4e,f, Supplementary Table 6). Though neutrophils showed an analogous increase in the inflamed LP, their increase was more pronounced within the EC (Extended Data Fig. 5c).

To gain insights on the biological functions of EC-specific myeloid cell DEGs in UC (Extended Data Fig. 5d), we performed pathway enrichment analysis (GSEA) to find biological pathways dysregulated in UC samples compared to NV within the myeloid compartment (methods). This pathway analysis revealed an elevated granulocyte adhesion and diapedesis (*CXCL8*, *ICAM1*, *SELL*, *ITGAM*), monocyte-associated surface molecules (*CD44*, *ICAM1*, *SELL*, *ITGAM*), endogenous TLR signaling (*S100A8*, *S100A9*, *TLR2*, *CD14*, *VCAN*), interferon alpha/beta (*IFITM2*, *IFITM3*, *IRF1*) and interferon gamma signaling (*STAT1*, *SOCS3*, *IFNGR2*,



*PTPN1*), TRAF3-dependent IRF activation (*TRIM25*, *IRF7*, *EP300*) and NF- $\kappa$ B signaling (*NFKBIA*, *TNFAIP3*, *TNFRSF1B*) in UC. In contrast, pathways up-regulated in NV included oxidative phosphorylation (*COX7C*, *COX41*, *UQCRB*), electron transport chain (*SLC25A6*, *SLC25A5*) and polyamine metabolism (*ADI1*, *SMS*, *SRM*) (Extended Data Fig. 5d). Altogether, these data demonstrate a major alteration in the composition of myeloid cells, associated with the emergence of neutrophils and MoMac within the epithelium leading to elaboration of potent pro-inflammatory programs along with a concomitant decrease of the homeostatic metabolic processes during active disease.

### **Increase in IgG<sup>+</sup> plasma cells in the EC and LP of UC.**

Plasma cells (PC) and B cells were selected *in silico* and sub-clustered into 10 PC sub-clusters and 6 B cell sub-clusters (Extended Data Fig. 6a, Supplementary Table 7, methods). P0, P1, P2, P5, P6, P7 were annotated as IgA<sup>+</sup> PCs (*IGHA1*, *IGHA2*) while P3, P4, P8 and P9 were annotated as IgG<sup>+</sup> PC (*IGHG1*, *IGHG2*, *IGHG3*, *IGHG4*) (Extended Data Fig. 6b). EC associated P6 and LP associated P0 were decreased whereas EC associated P3 and LP associated P3, P4, P8 and P9 were significantly increased in severe UC samples compared to NV (Extended Data Fig. 6b,c, Supplementary Fig. 3a). Consistent with our recent data<sup>28</sup> IgG<sup>+</sup> PC were significantly increased in both IIC (p=0.006) and LP (p=0.032) and IgA<sup>+</sup> PC were significantly decreased in LP (p=0.016) of UC vs NV (Extended Data Fig. 6e, Supplementary Fig. 1f, 3c). Multiparameter flow cytometry (Extended Data Fig. 6g, h, Supplementary Fig. 2c) further confirmed a significant increase in IgG<sup>+</sup> PC (p=0.0016) within the EC of UC vs NV (Extended Data Fig. 6i-k). Finally, the pathways enrichment analysis within the plasma cell compartment revealed the upregulation of class I antigen presentation (*HLA-C*, *HLA-A*, *HLA-B*), noncanonical NF $\kappa$ B pathway (*PSME2*, *UBC*) and interferon alpha pathways (*IFNAR2*, *IFNAR1*, *STAT1*) (Supplementary Fig. 3e) in UC compared to NV; highlighting elevated ER stress and interferon (IFN) imprinting on plasma cells in patients with UC samples.

Among B cells, we annotated germinal center (GC) B cells (B5, *AICDA*, *MME*, *CD38* and *CD27*), naïve B cells (B2 and B3; *IGHD*, *FCER2*) and memory B cells (B0 and B1; *TNFRSF13B/TACI*) (Extended Data Fig. 6a,b,d,f; Supplementary Fig. 1g, 3b,d) with comparable distribution between UC and NV. Consistent with the scRNA-seq data, we did not detect significant changes in total, naïve and switched memory B cell frequency between UC and NV (Extended Data Fig. 6l-p, Supplementary Fig. 2d) by flow cytometry, although and in consistence with our recent report<sup>28</sup>, EC-associated B cells were enriched in IFN alpha signaling pathway

(*STAT1*, *IRF9*) (Supplementary Fig. 3f). Enhanced type-1 interferon signaling in B cells and plasma cells is not only reported to be associated with autoimmune conditions but also has detrimental proinflammatory effects on non-autoimmune diseases<sup>45,46</sup>. Altogether, these data show a remodeling of EC-associated PC, analogous to LP-derived PC, with an increase in IgG<sup>+</sup> PC in UC.

### **Spatial transcriptomics reveals depletion of epithelial subtypes and enhanced cell-cell interactions within colonic tissues of patients with UC.**

To study the inflammation in a spatially coherent context, we used spatial transcriptomics (ST) to map the regional distribution of cell types in UC (n=2) and NV (n=2) (Supplementary Fig. 4a). The 4 samples contained between 11660 and 15846 genes, with the mean of 4438 unique genes per spot in the two UC samples, and 2873 in NV. We performed GSEA<sup>47</sup> to characterize the relative spatial distributions of different cell-types using cell-type specific signatures from scRNA-seq data, together with three epithelial (*LGR5*<sup>+</sup> epithelial stem, *BEST4*<sup>+</sup> *OTOP2*<sup>+</sup> absorptive and *BEST2*<sup>+</sup> *WFDC2*<sup>+</sup> secretory cells) signatures from the literature<sup>23</sup>. Enrichment scores indicated an infiltration of the immune cells such as neutrophils and T cells in active UC samples and a depletion of epithelial cells (Fig. 5a, Supplementary Fig. 4a). To estimate cell type fractions across spatial spots, we performed the deconvolution of ST data via spatialDWLS<sup>48</sup> (Fig. 5b) leveraging our scRNA-seq data as reference (Methods). Consistent with the enrichment analysis, deconvolution analysis revealed that a significant fraction of UC samples was comprised of different types of immune cells whereas NV mostly consisted of epithelial cells (Fig. 5b). Both enrichment analyses and cell-type deconvolution revealed increased proportions of immune cells including the four myeloid cell types, T cells, B cells, plasma cells, epithelial stem cells and the three stromal cell types in UC with a concomitant decrease in mature epithelial absorptive and secretory cell types (Wilcoxon signed-rank test, all  $p < 10^{-16}$ ). While epithelial stem cells showed a prominent presence in both UC and NV, these cells comprise a significantly higher portion of epithelium in UC (64% and 57% in UC, 15% and 8% in NV, Fig. 5c, Supplementary Fig. 4b-c, 5a-b, Supplementary Table 8).

To investigate cell-cell interactions, we performed co-localization analysis of different cell types in ST data for each patient (methods) and observed co-localization between myeloid cells and T cells as well as co-localization between neutrophils and epithelial absorptive cells solely in UC (Fig. 5d, Supplementary Table 8). Given the significant proximity of neutrophils and epithelial absorptive cells in UC samples (Fig. 5d), we investigated if neutrophils and epithelial cells were associated with up-regulated gene signature from IL-17+IL-22 treated biopsies. Our association

analysis based on ST data confirmed higher positive correlations with the activity of the IL-17+IL-22 signature in UC (neutrophils:  $r=0.37$ ,  $r=0.51$ ,  $p<10^{-43}$ ; epithelial absorptive cells:  $r=0.67$ ,  $r=0.78$ ,  $p<10^{-170}$ ) compared to NV (neutrophils:  $r=0.15$ ,  $r=0.12$ ; epithelial absorptive cells:  $r=0.07$ ,  $r=0.09$ ,  $p<10^{-2}$ , Fig. 5e). These results further confirm that myeloid cells could be stimulated by  $T_{H17}$  associated signals.

Quantitative IHC revealed a significant increase in  $CD14^+$  MoMac in UC vs NV, both infiltrating the colonic crypts ( $p=0.02$ ) and in a peri-crypt distribution ( $p=0.0001$ ; Extended Data Fig. 7a, b, Supplementary Fig. 6). Further, IF staining showed a significant increase in  $CD68^+$  macrophages in UC vs NV ( $p=0.01$ ; Extended Data Fig. 7c,d, Supplementary Fig. 7). Finally, a significant increase in EC-associated  $MPO^+$  neutrophils was noted in inflamed UC vs NV ( $p=0.008$ ; Extended Data Fig. 7e,f, Supplementary Fig. 8). The microscopy findings thus affirmed the epithelial cell-myeloid cell proximity observed using ST. Overall, consistent with our sequencing data, ST-derived data reveal a major loss of mature enterocytes and increased T cell-myeloid cell-epithelial cell proximity in UC, with  $T_{H17}$ -cell driven imprinting of UC tissues.

### **Persistence of myeloid cell and plasma cell signatures in the EC associates with anti-TNF treatment non-response in three distinct validation cohorts.**

We first tested the clinical relevance of DEGs associated with inflammation in the EC ( $IIC_{inf}$ ) (Fig. 1c) using three published data sets of UC patients treated with TNFi (Extended Data Fig.1, cohort details in methods). The activity of up-regulated  $IIC_{inf}$  genes was significantly reduced for both, 5 and 10 mg/kg doses of Infliximab/IFX- at week 8 ( $\Delta_{IFX-10} = -11.8$ (-25.4%),  $\Delta_{IFX-5} = -8$  (-19.9%)) and week 30 ( $\Delta_{IFX-10} = -20.4$  (-43.7%),  $\Delta_{IFX-5} = -12.1$  (-30%)) , in VC-3 (Fig. 6a,  $p<0.05$ ) while placebo patients had no significant change over 8 weeks. Similar effects were observed in VC-2 with IFX ( $\Delta_{IFX} = -13$  (-28%),  $p<10^{-5}$ ) and VC-4 with Adalimumab/ ADA ( $p<10^{-3}$ ), Extended Data Fig. 8a)

Next, we determined if expression of  $IIC_{inf}$  up-regulated genes differed between responders (R) and non-responders (NR) to TNFi. The  $IIC_{inf}$  score after 6 weeks of IFX treatment ( $p<0.05$ ) was significantly lower in patients that respond to IFX treatment (defined as mucosal healing) as compared to NR in VC-2 (R:  $\Delta_{IFX} = -21$  (-51.7%), NR:  $\Delta_{IFX} = -8$  (-17.5%)) (Fig. 6b). Similarly, in VC-3, TNFi R (defined as clinical response) demonstrated a significantly greater decrease ( $p<0.05$ ) in the  $IIC_{inf}$  score (R  $\Delta_{IFX-5} = -20$  (-54.1%),  $\Delta_{IFX-10} = -26$  (-56.9%)) as compared to NR ( $\Delta_{IFX-5} = -1$  (2.2%),  $\Delta_{IFX-10} = -2$  (-3.6%). Finally, in VC-4,  $IIC_{inf}$  up-regulated genes were significantly reduced in endoscopic R to TNFi therapy ( $\Delta_{ADA}$  R= -27.6(-45.3%),  $p=2.25*10^{-13}$ ; NR=-

5.14 (-7.8%),  $p=0.005$ ) (Fig. 6b, Extended Data Fig. 8b). Overall, these observations support that inflammation associated genes expressed in the IIC compartment are potentially targeted by TNFi and are important for disease remission.

Having observed a significant decrease in  $IIC_{inf}$  scores with TNFi therapy, we sought to determine the  $IIC_{inf}$  genes that reversed with TNFi response using the VC-3 RCT dataset. From the  $IIC_{inf}$  bulk signature, we identified 255 genes (159 up-regulated and 96 down-regulated, *IIC<sub>inf</sub> IFX response signature*, detailed in methods, Supplementary Fig. 9, Supplementary Table 9). Pathway enrichment analyses on the  $IIC_{inf}$  IFX-response signature (Supplementary Table 10) revealed proinflammatory pathways such as IL-17, TNF and NF- $\kappa$ B signaling pathways together with Th1, Th2 and Th17 differentiation pathways as IIC specific mechanisms attenuated with TNFi therapy in patients responding to the treatment. To delineate which cell types (in the epithelium compartment) express these treatment response genes we projected the  $IIC_{inf}$  IFX response signature onto our scRNA-seq data and noted higher expression of genes down-regulated with TNFi in myeloid cells including neutrophils, MoMac, macrophages and DCs, while genes up-regulated with TNFi were expressed in epithelial cells (Fig. 6c). To identify the likely TNFi responsive genes within cell types, we next intersected the  $IIC_{inf}$  IFX response signature with our scRNA-seq IIC-DEGs (UC vs NV). Genes up-regulated in active UC samples that were down-regulated in TNFi responders (VC-3 cohort) primarily included IIC-associated myeloid cells (*CXCL8*, *CSF3R*, *FPR2*, *FCGR2A*, *SOD2*, *TLR2*, *TREM1* and *STEAP4*) while the down-regulated genes in active UC samples that were up-regulated such as *ABCB1*, *ADH1C* in epithelial cells were up-regulated with treatment (Fig. 6d).

To further explore the importance of the IIC in TNFi response, we projected the cell type defining signature sets (Fig. 2b) onto the clinical trial associated datasets (VC-2, VC-3 and VC-4, Fig. 6e, Extended Data Fig. 9a-e, Supplementary Fig. 10-13). In this way, we explored the contribution of both IIC and LP-resident cells together in relation to TNFi response. Overall, we observed consistent decreased expression of genes enriched in myeloid cells (neutrophils, MoMac, macrophages, DCs), plasma cells and increased expression for genes marking epithelial cells (secretory, absorptive and stem cells) (Fig. 6e). Fig. 6f highlights the key cell types associated with treatment response in VC-3, where we found that both R and NR showed significant difference in genes associated with neutrophils, MoMac, macrophages, DCs, plasma cells at 5mg/kg dose of IFX. Furthermore, both VC-3 and VC-4 suggested that early W8-W10 reduction in myeloid cell- and plasma cell-associated genes was associated with treatment response. In VC-4, additional genes associated with mast cells, stromal cells, T cells, epithelial cells and B cells were also observed significantly changed at W10 compared to W0 (Fig. 6e) with

TNFi. Conversely, increased expression of epithelial cell-associated genes with TNFi therapy was associated with therapeutic response (Extended Data Fig. 8e, f, Extended Data Fig. 9b, d,e, Supplementary Fig. 11,12,13).

By correlating expression of the cell type defining signature genes with anti-TNF response in VC-3 (Pearson's correlation for genes that have correlation $>0.5$  and FDR $<0.05$ ), we identified leading-edge genes (LEGs) considered as 'driving' anti-TNF response for each of the cell types. Pathway enrichment analysis of these LEGs revealed TNF- $\alpha$  signaling, interferon- $\gamma$  response, and IL-6-JAK-STAT signaling as well as IL-17 signaling and Th1, Th2 and Th17 differentiation and neutrophil extracellular trap formation as pathways associated with anti-TNF therapeutic response in neutrophil-, MoMac- and macrophages. LEGs in plasma cells associated with TNFi response included unfolded protein response and protein export while LEGs in epithelial cell genes were enriched in fatty acid and xenobiotic metabolism along with cell adhesion and tight junction related pathways (Fig. 6g, Supplementary Fig. 14, Supplementary Table 10). Using both bulk sequencing and scRNA-seq approaches, we narrowed down to an overlapping set of cells and pathways to be associated with therapeutic non-response. Since these pathways remain elevated in TNFi NR, these could be potentially good targets for primary NR to therapy.

## Discussion

EC-based investigations are highly relevant to understanding the pathogenesis of UC. However, EC-focused high-dimensional studies in UC are sparse<sup>23,25</sup>. Additionally, EC-specific immune details get overlooked with whole tissue-based, and LP-based analyses of intestinal biopsies<sup>22,24</sup>. Here, using two distinct cohorts of patients, and integrating bulk RNA-seq-, scRNA-seq-, spatial transcriptomic-, flow cytometry- and microscopy-derived data, we have identified novel cellular and transcriptional signatures specific to the EC in patients with UC. We show that the molecular signature associated with colonic inflammation ( $IIC_{int}$ ) highly correlates with UC severity (clinical, endoscopic and histologic). By identifying high resolution intestinal inflammation signatures in our data and intersecting them with 3 published longitudinal patient cohorts, including subjects from a seminal clinical trial of TNFi in UC<sup>49-51</sup>, we have identified key transcriptional and cellular correlates of TNFi failure in UC. Identification of such cellular modules associated with anti-TNF non-response can be harnessed to identify early correlates of treatment non-response (based on the trajectory of response to anti-TNF in UC at early time points such as week 6 (based on VC-3) and week 10 (based on VC-4), as well as to stratify patients as per the likelihood of response. Thus, the present study details high resolution EC dynamics in UC, paves the way for

development of novel biomarkers to identify TNFi failure, and provides a rational basis of the design of sequential and combinatorial therapies in UC.

In addition to nutrient absorptive and secretory functions, to facilitate mucosal homeostasis, colonic epithelial cells tether mucosal immune cells such as IELs and mediate innate and adaptive mucosal immunity<sup>52</sup> through bidirectional communication with immune cells<sup>9,11,53-55</sup>. During the inflamed state in UC, we observed a notable loss of differentiated enterocytes, both absorptive and secretory, while the stem cell-like enterocytes were numerically preserved. Accordingly and associated with the loss of epithelial tethers, physiological EC residents, especially  $\gamma\delta^+$  IELs which are known to enhance barrier function<sup>56</sup>, dampen inflammation<sup>57</sup> and regulate  $\alpha\beta^+$  IELs<sup>58</sup> were significantly depleted from the EC in UC as was previously reported in CD<sup>26</sup>.  $\gamma\delta^+$  IEL loss from the EC was accompanied by a major influx of CD4<sup>+</sup> $\alpha\beta^+$  IELs, especially T<sub>H17</sub> and T<sub>REG</sub> cells encoding a number of pathogenic, pro inflammatory programs, associated with tissue damage and worsening of disease activity<sup>57,59</sup> that included TNFR2 signaling<sup>60</sup>, TCR activation<sup>61,62</sup>, IL-2/STAT5 signaling<sup>63</sup>, IL-7 signaling<sup>64</sup>, IL-9 signaling<sup>65</sup> and IL-12 signaling<sup>66</sup>.

In addition to T cell-associated changes, multiple orthogonal approaches demonstrated a dominant presence of myeloid cells (including neutrophils, inflammatory monocytes, inflammatory macrophages and DCs), infiltrating or proximate to the EC during active inflammation. Notably, we could capture previously unreported neutrophils in our scRNA seq dataset and ascribe this to the use of fresh biopsies, gentle EDTA-based dissociation protocols and short processing times in our scRNA-seq methodologies. In addition to infiltrating neutrophils, and inflammatory monocytes, homeostatic myeloid cells, especially cDCs and resident macrophages, key sentinels and drivers of tissue homeostasis at steady state<sup>67</sup>, switched to inflammatory phenotypes (including inflammatory macrophages) and were noted to encode powerful inflammatory programs including interferon alpha/beta signaling, interferon gamma signaling, IRF activation and NF- $\kappa$ B signaling in UC, likely contributing in a major way to disease pathogenesis<sup>67,68</sup>. Spatial transcriptomic data and microscopic approaches corroborated a significantly increased proximity between myeloid cells and epithelial cells during UC. Furthermore, IL-17/IL-22-stimulated colonic tissues up-regulated key myeloid-recruiting chemokine genes including *CXCL1*, *CXCL2* and *CXCL3*, providing a mechanistic link between the observed T cell and myeloid cell perturbations in UC. Top drivers of histologic activity in UC were also dominated by myeloid-cell associated genes, including *S100A8/S9*, *FPR1* and *FPR2*<sup>69</sup>, *FCGR3B*<sup>70</sup>, *CLEC4A*, *TREM1*<sup>39,71</sup>, *CXCR1/2*<sup>72</sup> and *CXCL1*. Other notable genes associated with increased histological activity included

epithelium-related *SAA1*<sup>73</sup>, innate immune related *LILRA5*, oxidative stress sensor /responder genes *VNN3*<sup>74</sup>, *HIF1A*<sup>75</sup>, *DRAM1*, *XBP1* and *AQP9*. These data provide potential biomarkers for monitoring of disease activity, in addition to some established biomarkers such as myeloid cell-derived calprotectin<sup>76,77</sup>.

A major aspect of our report details cell signatures in UC that are associated with TNFi-nonresponse. This was facilitated by intersecting high-resolution tissue data with three published cohorts, of which two were large, randomized placebo-controlled, clinical trials with a cumulative total of 252 patients<sup>50,51</sup>. In both clinical trials, we noted unequivocal and a significant increase in the frequency of neutrophils, inflammatory MoMacs, inflammatory macrophages and PC in non-responders to TNFi. Analogously, early, week 8-10 reductions in the frequencies of epithelium-associated myeloid cells and absorptive and secretory enterocytes were associated with TNFi treatment-response. These data, first in UC, have potentially significant implications for patients who do not respond to TNFi and are reminiscent of the GIMATs-associated myeloid cell drivers of TNFi-non response in CD<sup>22,24</sup>. Findings from our study indicates the prominence of epithelium associated myeloid cells to be associated with lack of early response in UC patients on TNFi<sup>78</sup>. In sight of multitude of treatment options in UC, our findings support the development of myeloid-targeted therapies particularly in TNFi-non responders. Among the clinically approved drugs, JANUS Kinase inhibitors (JAKi) target many of the pathways that persist in TNFi-non responders. Our work proposes a rationale to use JAK inhibitors in combination with or sequentially in the face of TNFi non-response in alignment with a recent meta-analysis report<sup>79</sup>. Thus, by detailing EC-compartment-associated changes in UC and elaborating on cell- and transcriptomic-signatures associated with non-response to TNFi, here we provide rational underpinnings for therapeutic decision-making with major clinical implications.

**Acknowledgements:** We thank the patients who participated in the study. We thank Travis Dawson, Rachel Chen and Dr. Zhihong Chen in Human Immune Monitoring Centre at the Icahn School of Medicine for their single cell transcriptomic assay technical assistance and the Mount Sinai Flow Cytometry Core for support. This work was supported by the following grants: NIH/NIDDK R01 123749 (SM). CA was supported in part by The Leona M. and Harry B. Helmsley Charitable Trust and RC2 DK122532/DK/NIDDK NIH. JCM is supported by NExT “Junior Talent”, ANR JCJC (ANR-20-CE17-0009) and ARC Programmes labellisés (PGA). RCU is supported by an NIH K23 Career Development Award (1K23DK111995-01A1).

**Author contributions:** DJ, FP, MSF, CA and SM drafted the manuscript. SM designed the study, supervised experimental data collection and coordinated integration of collaboration between all participating laboratories. DJ led the work, conducted experiments, analyzed the data and coordinated collaboration between all participating laboratories. ZAT, MSF and CA performed bulk RNA seq analyses and TNFi treatment response analyses. AK and FP performed single cell RNA seq analyses. AK, FP and GCY performed spatial transcriptomic analyses. STE, JBM, JAH and AN performed analysis on VC4. JFC, RU, MK and JM contributed patient-related samples and clinical analyses. JCM, AF, TL, AB and RJ performed in vitro stimulation of biopsies with IL-17/IL-22 and the relevant analyses. All other authors contributed to experimental data and analyses, and along with SM, DJ, FP, MSF and CA, they critically reviewed and edited the final version of the manuscript.

**Competing interests:** SM reports receiving research grants from Genentech and Takeda; receiving payment for lectures from Takeda, Genentech, Morphic; and receiving consulting fees from Takeda, Morphic, Ferring and Arena Pharmaceuticals. JM reports receiving consulting fees from Janssen Pharmaceuticals. JFC reports receiving research grants from AbbVie, Janssen Pharmaceuticals, Takeda and Bristol Myers Squibb; receiving payment for lectures from AbbVie, and Takeda; receiving consulting fees from AbbVie, Amgen, AnaptysBio, Allergan, Arena Pharmaceuticals, Boehringer Ingelheim, Bristol Myers Squibb, Celgene Corporation, Celltrion, Eli Lilly, Ferring Pharmaceuticals, Galmed Research, Glaxo Smith Kline, Genentech (Roche), Janssen Pharmaceuticals, Kaleido Biosciences, Immunicon, Invea, Iterative Scopes, Merck, Landos, Microba Life Science, Novartis, Otsuka Pharmaceutical, Pfizer, Protagonist Therapeutics, Prometheus, Sanofi, Seres, Takeda, Teva, TiGenix, Vifor; and hold stock options in Intestinal Biotech Development. JM reports receiving consulting fees from Janssen Pharmaceuticals. RCU has served as an advisory board member or consultant for AbbVie, Bristol Myers Squibb, Celltrion, Janssen, Pfizer, Roivant, and Takeda; research support from AbbVie, Boehringer Ingelheim, Eli Lilly, and Pfizer. MK has served as a consultant for AbbVie, Pfizer, Bristol Myers Squibb, Fresenius, GoodRx.

## Figure Legends

**Fig. 1: Intraepithelial immune compartment is extensively remodeled in active UC.**

**a**, Experimental design outlining the methodologies and assays used. **b**, *Left*, Principal component analysis (PCA) plot of the epithelial compartment (Epi<sub>bulk</sub>) expression profiles from the



healthy (blue) (n=6), uninfamed (green) (n=4), and inflamed (red) (n=5) biopsies. *Right*; Venn diagram showing the differentially expressed genes (DEGs) from inflamed UC samples compared to NV samples (blue) and the DEGs from inflamed UC samples compared to the uninfamed UC samples (green). Up-regulated DEGs are in red while down-regulated DEGs in blue. **c**, *Left*, PCA plot of the intraepithelial immune cells (IIC<sub>bulk</sub>) RNA-seq expression profiles of the samples from the IIC from the healthy (blue) (n=6), uninfamed (green) (n=4), and inflamed (red) (n=5) biopsies. *Right*; Venn diagram showing the DEGs from inflamed UC samples compared to NV samples (blue) and the DEGs from inflamed UC samples compared to the uninfamed UC samples (green). Up-regulated DEGs are in red while down-regulated DEGs are in blue. **d**, Venn diagram showing the intersection of inflamed vs uninfamed UC DEGs from the Epi<sub>bulk</sub> and IIC<sub>bulk</sub>. **e-f**, Overexpression analysis (ORA) of DEGs between inflamed and uninfamed UC samples interrogated from the cell type signature<sup>23</sup> for selected significantly (adj-p<0.05) enriched cell types (e) and KEGG<sup>29</sup> pathways (f). Color intensity represents -log<sub>10</sub> transformed p-adjusted and the size represents the log<sub>2</sub> transformed Odds Ratio (OR) using Fisher's exact test. OR from enrichments with <5 IIC<sub>inf</sub> DEG are not presented. **g**, Heatmap depicting expression profiles of 42 DEGs selected for relevance with disease pathogenesis among the DEGs in the intersection of Fig.1d. **h**, Association of the IIC<sub>inf</sub> signature with UC disease and severity in the validation cohort-1 (VC-1). Each plot in **h** represents the estimated marginal mean expression and 95% CI for the GSVA-scores of the up-regulated genes in the IIC<sub>inf</sub> signature across conditions. *Upper left*; Activity of IIC<sub>inf</sub> signature in endoscopically inflamed colon compared to uninfamed biopsies of IBD patients and non-IBD normal volunteers (NV) from VC-1 cohort using fora function from fgsea R package. *Upper right*; The mean activity of up-regulated DEGs from IIC<sub>inf</sub> signature was compared between clinically active IBD patients compared to IBD patients with no active disease and was also associated with endoscopic scores. *Lower*; Activity of IIC<sub>inf</sub> up-regulated genes was associated with histological scores (measured by Nancy score 0-4) of endoscopically inflamed biopsies (red) and uninfamed biopsies (blue) from UC patients using Pearson's correlation test \*p ≤ 0.05; \*\*p ≤ 0.01; \*\*\*p ≤ 0.001; \*\*\*\*p ≤ 0.0001 (ic test and BH-adjusted p-value). **i**, Heatmap showing the expression of the top 20 genes (among DEGs in inflamed versus uninfamed or NV comparisons) correlated with histological disease severity (Nancy index 0-4) measured from endoscopically inflamed gut biopsies of UC patients in VC-1 cohort.

**Fig. 2: High-resolution mapping of immune cells reveals distinct cellular composition in the colonic EC of ulcerative colitis patients.**

**a-c**, UMAP of scRNA-seq data colored by **a**, disease status (i.e., UC samples (n=9) and NV samples (n=9)), **b**, intraepithelial immune cells (IIC) and lamina propria (LP) compartments, **c**, 32 cell-type annotated clusters identified by unsupervised clustering **d**, Heatmap showing the average z-score normalized log expression of canonical cell type markers across different clusters. Pie charts on the left indicate overall composition for each cell type, in terms of disease status, severity and tissue compartments. **e-f**, Boxplots of per-sample cell type fractions in IIC (e) and LP (f), stratified by disease status and severity. NV (blue), mild/moderate UC (orange) and severe UC samples (red). P-values from Wilcoxon signed-rank test are reported.

**Fig.3: Active UC is associated with depletion of homeostatic  $\gamma\delta^+$  IELs and influx of proinflammatory  $T_{H17}$  cells into the EC.**

**a**, UMAP of T cells from NV (n=9) and active UC (n=9) colonic biopsies colored by 10 sub-clusters identified via unsupervised clustering. **b**, Heatmap showing average z scores of log normalized expression of the T cell sub-clusters defining key genes. The sub-cluster composition in terms of intestinal compartments (IIC and LP), and disease severity (NV, mild/moderate UC and severe UC) is depicted in the pie charts on the left side. **c**, Boxplots showing the distribution of per-sample relative fractions of T cell sub-clusters in the IIC stratified by disease severity. Fractions are computed with respect to all T cells in each sample. P-values from Wilcoxon signed-rank test are reported. **d**, Boxplots of per-sample fractions of  $T_{H17}$ ,  $T_{REG}$ ,  $\gamma\delta^+$  IELs, cytotoxic  $CD8^+$  T, memory and naïve T cells stratified by disease severity in IIC. Fractions are computed with respect to the total number all cells in each sample. P-values from Wilcoxon signed-rank test are reported. **e-h**, Representative flow cytometry plots to identify  $\gamma\delta^+$  (e) and  $CD4^+\alpha\beta^+$  IELs (g) in NV, uninfamed UC and inflamed UC samples. Bar plots depicting the frequencies of  $\gamma\delta^+$  IELs (f), the frequencies of  $CD4^+\alpha\beta^+$  IELs (h, *left*) and  $CD4:CD8$  ratio (h, *right*) in the inflamed samples compared to the uninfamed UC and NV samples. Cells were gated on live singlet  $CD45^+CD3^+$  and on live singlet  $CD45^+CD3^+TCR\alpha\beta^+$  respectively. P-values from one-way ANOVA test are reported. \*\*\*\*p<0.0001. **i**, Immunofluorescence (IF) staining for  $CD4^+$  T cells (green), EPCAM (red) and DAPI (blue) in formalin-fixed paraffin embedded colonic sections derived from NV (*left, 10X and 20X magnification*) and patients with UC (*right, 10X and 20X magnification*). **j**, IF- based quantification of  $CD4^+$  T cells within the epithelial compartment with p-values computed using Mann Whitney test. \*\*\*\*p<0.0001.

**Fig. 4: Active UC is associated with a major influx of pro-inflammatory myeloid cells into the EC.**

**a**, UMAP of myeloid cells from NV (n=9) and active UC (n=9) colonic biopsies colored by 10 sub-clusters (M0-M9). **b**, Heatmap showing the average z-scores of log transformed expression for myeloid sub-clusters defining genes. The sub-cluster composition in terms of intestinal compartments (IIC and LP), and disease severity (NV, mild/moderate UC and severe UC) is depicted in the pie charts on the left side. **c**, Boxplots showing the distribution of per-sample relative fractions of myeloid sub-clusters (M0-M9) in IIC stratified by disease severity. Fractions are computed with respect to all myeloid cells in each sample. P-values from Wilcoxon signed-rank test are reported. **d**, Boxplots showing the distribution of per-sample fractions of myeloid cell subtypes (i.e., neutrophils, MoMac, macrophages and DC) in the IIC stratified by disease severity. Fractions are computed with respect to the total number of all cells within each sample. P-values from Wilcoxon signed-rank test are reported. **e**, Representative flow cytometry plots identifying CD14<sup>hi</sup> MoMac, CD16<sup>hi</sup> MoMac, eosinophils and neutrophils in NV, uninfamed UC and inflamed UC samples. **f**, Bar plots show an increased frequency of CD14<sup>hi</sup> MoMac (*left*) and neutrophils (*right*) in the epithelial compartment from inflamed UC samples compared to uninfamed UC and NV samples.

**Fig. 5: Spatial transcriptomics (ST) reveals depletion of epithelial subtypes and enhanced cell-cell interactions within colonic tissues of patients with UC.**

**a**, Spatial distribution of enrichment scores of neutrophils, T cells, epithelial absorptive and epithelial secretory in the four ST samples. **b**, Cell type fractions obtained via deconvolution. Spatial distribution of cell types, represented as cell type fraction pie charts at each spatial location in representative inflamed and healthy colonic biopsies. The spot diameter is 55 $\mu$ m and center-to-center distance between the spots is 100 $\mu$ m. Bar plot (*right*) showing the average fraction across all spatial spots inferred via deconvolution analysis for each sample. **c**, Per-sample distribution across spatial spots of enrichment scores and cell-type fractions inferred via deconvolution analysis for different cell types. **d**, Cell-cell co-localization analysis for NV samples (left panel) and UC samples (right). For significantly enriched co-localization pairs, the relative effect magnitude is represented by the bubble size (log<sub>2</sub> fold change) while the significance by the bubble color (-log<sub>10</sub> p-value). Co-localized cell type pairs detected in both NV samples and not in UC samples are highlighted with a square, while those detected in both UC samples only by a circle. **e**, Scatterplots showing enrichment scores from ST data of neutrophils and epithelial absorptive (y-axis) versus IL-17+ IL-22 gene signatures (x-axis) across spatial spots. For each

scatterplot, Pearson's correlation and corresponding correlation test p-values are shown. These statistics were separately derived for each sample.

**Fig. 6: Persistence of myeloid cell and plasma cell signatures in the EC associates with anti-TNF treatment non-response in three distinct validation cohorts.**

**a**, Changes in score (mean  $\pm$  SEM) for the activity (GSVA scores) of the IIC<sub>inf</sub> signatures at baseline and after 6 weeks (in VC-2, left) or 8 and 30 weeks (in VC-3, middle) of anti-TNF therapy (Infliximab (IFX) in VC-2 and at week 10 (in validation cohort 4, VC-4, right) after adalimumab (ADA) therapy. **b**, IFX-induced changes in IIC<sub>inf</sub> signatures responders (R) and non-responders (NR), where treatment response was defined by mucosal healing in VC-2, clinical response in VC-3 and by endoscopic remission in VC-4. Changes with treatment were modeled using LMEM and comparisons tested using contrasts (emmeans package in R). p values above the error bars denote significant change from baseline within the group at each time point, while p-values at the top indicate that treatment changes over time are significantly different between R vs NR, \*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$  **c**, Boxplots showing the average z-score of IIC<sub>inf</sub> up- and down-regulated response signatures (methods) based on the transcriptomic profile of each cell from scRNA-seq derived cell types. For each cell type, we report one-sided AUC reflecting the specific z-score to discriminate between that cell type and the rest of the cell types. **d**, Heatmap representing the expression profile of the IIC<sub>inf</sub> IFX-response signature that overlaps with scRNA-seq derived cell type signatures in VC-3. On the left side of the heatmap [IIC<sub>inf</sub> (UC vs NV) scRNA-seq], the status of the gene in the differential expression analysis comparing UC vs NV profiles for each cell type is indicated (up-regulated (red) or down-regulated (blue)). **e**, Differential anti-TNF effect between R and NR for the activity (GSVA scores) of cell-type scRNA-seq derived signatures in VC-3 and VC-4 cohorts. Bar plots indicate  $-\log_{10}(p\text{-values})$  for the comparison of the changes (from W0) in scores between R and NR at respective time points (W8, W10, W30). The dashed line indicates  $p\text{-value}=0.05$ , indicating a significant change in R vs NR at W30 with respect to W0. **f**, Estimated marginal mean for the GSVA scores of scRNA-seq derived cell types that were significantly different between R and NR at week 30 from the start of IFX therapy at dose 5 mg/kg. **g**, Alluvial plots showing the pathway ORA for the leading-edge genes (LEGs) of each cell type (only those associated with IFX response in f are shown) using mSigDB Hallmark<sup>80</sup> database. Color density denotes adjusted p-values while the band width denotes odds ratio of the pathway enrichment.

**Extended Data Figure Legends**

**Extended Data Fig. 1: Outline of the patients and samples included in the study.**

**Extended Data Fig. 2: Association of IIC<sub>inf</sub> GSVA activity scores in biopsies from the Validation Cohort-1/ VC-1 stratified by inflammation status and disease severity.**

**a**, Expression of IIC<sub>inf</sub> GSVA activity scores of the inflamed UC samples from the Primary Cohort (PC) in the inflamed UC, uninfamed UC samples and NV (control) samples in the VC-1. **b**, Association of the mean marginal expression for GSVA score of IIC<sub>inf</sub> genes from PC with clinical severity in VC-1. **c**, Association of mean marginal expression for GSVA score of IIC<sub>inf</sub> genes with endoscopic severity in VC-1. **d**, Association of mean marginal expression for GSVA score of IIC<sub>inf</sub> genes with histologic severity in VC-1. Up and down indicate direction associated with inflammation.

**Extended Data Fig. 3: Distribution of T cell types in the LP in patients with active UC and pathway analysis of IELs.**

**a**, Boxplots showing the distribution of per-sample relative fractions within the T cell compartment of different sub-clusters in LP stratified by disease severity. Fractions are computed with respect to all T cells in each sample. P-values from Wilcoxon signed-rank test are reported. **b**, Boxplots of per-sample fractions of T<sub>H17</sub>, T<sub>REG</sub>,  $\gamma\delta^+$  IELs, cytotoxic CD8<sup>+</sup> T, memory and naïve T cells stratified by disease severity in LP. Fractions are computed with respect to the total number of all cells in each sample. P-values from Wilcoxon signed-rank test are reported. **c**, Flow cytometry-based quantification of the frequencies of  $\alpha\beta^+$  T cells and CD8<sup>+</sup>  $\alpha\beta^+$  T cells. \*\*p<0.01, \*\*\*\*p<0.0001. **d**, Immunohistochemical (IHC) staining showing the infiltration of CD4<sup>+</sup> T cells (pink cells) within the EC. Representative sections from an inflamed UC biopsy (black arrows, *right panel*) uninfamed UC biopsy (*middle panel*) and NV biopsy (*left panel*) are shown. **e**, Bar plot comparing intraepithelial CD4<sup>+</sup> T cell count between NV, uninfamed UC and inflamed UC samples, based on IHC staining. P-values from one-way ANOVA test are reported, ns=not significant, \*p<0.05; \*\*\*p<0.001. **f**, Heat map shows pathway enrichment analysis from IELs in UC versus NV samples. The bubble plot on the left side of the figure shows the summary statistics from pathway analysis; with the size of the bubble corresponding to the enrichment score (ES) while the bubble color to the signed -log<sub>10</sub> adjusted p-value. The heatmap at the right side shows the log fold change between UC and NV for the top five leading edge genes in each pathway.

**Extended Data Fig. 4: IL-17 and IL-22 stimulation led to proinflammatory transcriptional reprogramming of healthy colonic tissue.**

**a**, Methodological approach adopted to examine the effect of IL-17 and IL-22 stimulation on colonic intestinal biopsies. **b**, Volcano plots depicting top differentially expressed genes post- IL-22, IL-17 and IL17+ IL-22 stimulation of colonic biopsies derived from RNA sequencing. **c**, Cell types corresponding to the DEGs from RNA sequencing done on colonic biopsies stimulated by IL-17 + IL-22 cytokines. **d**, Heatmap showing biological functions for gene expression levels for up-regulated genes after colonic biopsies were stimulated with medium containing IL-17 and IL-22 compared to medium only. **e**, Correlation plots for IL17 and IL22 gene expression in VC-1 with the GSVA scores of IL17+IL22 signature (up-regulated genes) from the ex-vivo experiment using Pearson's correlation. **f**, Box plots indicating GSVA scores of up-regulated IL-17+IL-22 signature post stimulation of colonic tissues in VC-1 cohort stratified by inflammation status (NV, uninfamed and inflamed samples from UC patients) used estimated marginal means using lsmeans function from emmeans R package. **g**, Heatmap showing the expression of the IL-17+ IL-22 signature per intestinal segment (large intestine) in VC-1. **h**, Venn diagram showing the intersection of genes that were up-/down-regulated by IL-17+IL-22 stimulation and intersecting DEGs common to  $Epi_{inf}$  and  $IIC_{inf}$  (as shown in Fig. 1d) using Fisher's Exact Test.

**Extended Data Fig. 5: Distribution of myeloid cell types in the LP in patients with active UC and pathway analysis of IIC enriched myeloid cell types.**

**a**, Boxplots showing the distribution of per-sample relative fractions within the myeloid cell compartment of different sub-clusters in LP stratified by disease severity. Fractions are computed with respect to all myeloid cells in each sample. P-values from Wilcoxon signed-rank test are reported. **b**, Boxplots of per-sample fractions of myeloid cell types (i.e. Neutrophils, MoMac, Macrophages and DCs) stratified by disease severity in LP. Fractions are computed with respect to the total number of all cells in each sample. P-values from Wilcoxon signed-rank test are reported. **c**, Flow cytometry comparing of EC-associated neutrophils and LP-associated neutrophils in the UC samples. **d**, Pathway enrichment analysis of epithelium-enriched myeloid cells in UC samples compared to NV samples. The bubble plot on the left side of the figure shows the summary statistics from pathway analysis; with the size of the bubble corresponding to the enrichment score (ES) while the bubble color to the signed  $-\log_{10}$  adjusted p-value. The heatmap at the right side shows the log fold change between UC and NV for the top five leading edge genes in each pathway.

**Extended Data Fig. 6: Altered transcriptomic profiles of plasma cells and B cells in active UC patients and NV.**

**a**, UMAP of plasma cells and B cells from NV (n=9) and active UC (n=9) colonic biopsies by unsupervised clustering to demonstrate 10 plasma cell sub-clusters and 6 B cell sub-clusters. **b**, Heatmap showing the average expression z-scores of plasma cells and B cells defining genes for each sub-cluster. The sub-cluster composition in terms of intestinal compartments (IIC and LP), and disease severity (NV, mild/moderate UC and severe UC) is depicted in the pie charts on the right side of the heatmap. **c-d**, Boxplots showing the distribution of plasma cell sub-clusters (c) and B cell subclusters (d) in the IIC stratified by disease severity. Fractions are computed with respect to all plasma (c) or B (d) cells in each sample. P-values from Wilcoxon signed-rank test are reported. **e-f**, Boxplots showing the distribution of per-sample fractions of plasma cell isotypes, IgA<sup>+</sup> and IgG<sup>+</sup> (e), and naïve, memory, GC and unclassified (other) B cell subtypes (f) in the IIC stratified by disease severity. Fractions are computed with respect to the total number of all cells in each sample. P-values from Wilcoxon signed-rank test are reported. **g-h**, Representative flow cytometry plots to identify plasma cell types in NV, uninfamed UC and inflamed UC samples. **i-p**, Bar plots showing frequencies of plasma cells (i), IgA<sup>+</sup> plasma cells (j), IgG<sup>+</sup> plasma cells (k), B cells (l) and B cell types (naïve (m), switched memory (n), switched memory IgA<sup>+</sup> (o) and switched memory IgG<sup>+</sup> B cells (p) in healthy and inflamed samples. P-values from Kruskal-Wallis tests are reported. \*\*p<0.01, \*p<0.05.

#### **Extended Data Fig. 7: Neutrophils, monocytes and macrophages penetrate the colonic epithelium in patients with active UC.**

**a**, IHC staining showing the infiltration of CD14<sup>+</sup> myeloid cells (in brown) within the EC of active UC patients (*right*) at the inflamed sites in comparison to the uninfamed sites (*middle*) and NV samples (*left*). Red arrows indicate the CD14<sup>+</sup> cells entering the crypts and green arrows indicate the CD14<sup>+</sup> cells surrounding the crypts. **b**, Bar plots comparing CD14<sup>+</sup> cells surrounding the intestinal crypts (*left*) and CD14<sup>+</sup> cells entering the crypts (*right*) between inflamed UC-, uninfamed UC- and NV-derived colonic tissues. P-values from one-way ANOVA test are reported. \*\*\*p<0.001, \*\*p<0.01, \*p<0.05, ns-p>0.05. **c**, IF staining depicting the expression of CD68 (green), EPCAM (red), DAPI (blue) in NV (*left, 10X and 20X magnification*) and UC (*right, 10X and 20X magnification*) samples. **d**, Bar plots comparing epithelium associated macrophages between inflamed UC and NV-derived tissues. P-value from Mann-Whitney test is reported. \*\*\*p<0.001. **e**, IF staining depicting the expression of MPO (green), EPCAM (red) and DAPI (blue) in NV (*left, 10X and 20X magnification*) and UC (*right, 10X and 20X magnification*) samples. **f**, Bar plot comparing epithelium-associated neutrophils between inflamed UC and NV-derived tissues. P-value from Mann-Whitney test is reported. \*\*p<0.01.

**Extended Data Fig. 8: Down-regulated IIC<sub>inf</sub> signature associates with non-response to anti-TNF therapy.**

**a**, Estimated marginal means (mean  $\pm$  SEM) for the activity (GSVA scores) of the down-regulated IIC<sub>inf</sub> signatures at baseline and after 6 weeks (in VC-2) or 8 and 30 weeks (in VC-3) and 10 weeks (in VC-4) of anti-TNF therapy (Infliximab (IFX) in VC-2 and IFX 5 mg/kg and 10 mg/kg in VC-3); ADA in VC-4 and placebo. **b**, IFX-induced changes in down-regulated IIC<sub>inf</sub> signature in responders (R) and non-responders (NR), where treatment response was defined by mucosal healing in VC-2, clinical response in VC-3 and endoscopic response in VC-4. P-values above the error bars denote significant change from baseline within the group at each time point, while p-values at the top indicate that treatment changes over time are significantly different between treatment groups (a-b) or R vs NR (c-d), \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ . **c-d**, Estimated marginal means (mean  $\pm$  SEM) for the activity (GSVA scores) of the down-regulated Epi<sub>inf</sub> subtype signatures at baseline and after 6 weeks (**c**, in VC-2) or 8 and 30 weeks (**d**, in VC-3) of anti-TNF therapy (Infliximab (IFX) in VC-2 and IFX 5 mg/kg and 10 mg/kg in VC-3) and placebo. **e-f**, Effect of anti-TNF therapy on the activity (GSVA scores) of the epithelial cell subtype signatures in VC-2 (**e**) and VC-3 (**f**) in R and NR.

**Extended Data Fig. 9: Effect of anti-TNF therapy on cell types in VC-2, VC-3 and VC-4 cohorts in patients on anti-TNF.**

**a,c**, Effect of anti-TNF therapy on the activity of cell-type signatures derived from scRNA-seq between active treatment and placebo in VC-2 (**a**) and VC-3 (**c**). **b,d,e**, Effect of anti-TNF therapy on the activity of cell-type signatures derived from scRNA-seq between responders (R) and non-responders (NR) in VC-2 (**b**), VC-3 (**d**) and VC-4 (**e**). Gene activity is assessed by GSVA scores from scRNA-seq derived cell type signatures. Connecting line shows  $p < 0.05$  between R and NR at respective timepoints.

**Data availability statement**

All requests for raw and analyzed data and materials will be promptly reviewed by corresponding author and the study team. We will provide source data files for all the figures. Raw sequencing reads of scRNA-seq samples as well as UMI tables are available on the Gene Expression Omnibus (GEO) under accession number **GSE231817** that may be accessible with the following token **cbefiwcgvvsxrwf**.



## Code Availability Statement

The R code developed for clustering and analyses in this study will be made available.

## References-

1. Ungaro, R., Mehandru, S., Allen, P.B., Peyrin-Biroulet, L. & Colombel, J.F. Ulcerative colitis. *Lancet* **389**, 1756-1770 (2017).
2. Liang, S., *et al.* Nutrient Sensing by the Intestinal Epithelium Orchestrates Mucosal Antimicrobial Defense via Translational Control of Hes1. *Cell Host Microbe* **25**, 706-718 e707 (2019).
3. Kopf, A. & Sixt, M. Gut Homeostasis: Active Migration of Intestinal Epithelial Cells in Tissue Renewal. *Curr Biol* **29**, R1091-R1093 (2019).
4. Di Marco Barros, R., *et al.* Epithelia Use Butyrophilin-like Molecules to Shape Organ-Specific gammadelta T Cell Compartments. *Cell* **167**, 203-218 e217 (2016).
5. Hoytema van Konijnenburg, D.P., *et al.* Intestinal Epithelial and Intraepithelial T Cell Crosstalk Mediates a Dynamic Response to Infection. *Cell* **171**, 783-794 e713 (2017).
6. Sullivan, Z.A., *et al.* gammadelta T cells regulate the intestinal response to nutrient sensing. *Science* **371**(2021).
7. Matsuzawa-Ishimoto, Y., *et al.* The gammadelta IEL effector API5 masks genetic susceptibility to Paneth cell death. *Nature* (2022).
8. Kang, B., *et al.* Commensal microbiota drive the functional diversification of colon macrophages. *Mucosal Immunol* **13**, 216-229 (2020).
9. Pull, S.L., Doherty, J.M., Mills, J.C., Gordon, J.I. & Stappenbeck, T.S. Activated macrophages are an adaptive element of the colonic epithelial progenitor niche necessary for regenerative responses to injury. *Proc Natl Acad Sci U S A* **102**, 99-104 (2005).
10. Sehgal, A., *et al.* The role of CSF1R-dependent macrophages in control of the intestinal stem-cell niche. *Nat Commun* **9**, 1272 (2018).
11. Bai, R., *et al.* Myeloid cells protect intestinal epithelial barrier integrity through the angiogenin/plexin-B2 axis. *EMBO J* **39**, e103325 (2020).
12. Chikina, A.S., *et al.* Macrophages Maintain Epithelium Integrity by Limiting Fungal Product Absorption. *Cell* **183**, 411-428 e416 (2020).
13. Rutgeerts, P., *et al.* Infliximab for induction and maintenance therapy for ulcerative colitis. *N Engl J Med* **353**, 2462-2476 (2005).
14. Feagan, B.G., *et al.* Vedolizumab as induction and maintenance therapy for ulcerative colitis. *N Engl J Med* **369**, 699-710 (2013).
15. Sands, B.E., *et al.* Ustekinumab as Induction and Maintenance Therapy for Ulcerative Colitis. *N Engl J Med* **381**, 1201-1214 (2019).
16. Sandborn, W.J., Su, C. & Panes, J. Tofacitinib as Induction and Maintenance Therapy for Ulcerative Colitis. *N Engl J Med* **377**, 496-497 (2017).
17. Sandborn, W.J., *et al.* Tofacitinib as Induction and Maintenance Therapy for Ulcerative Colitis. *N Engl J Med* **376**, 1723-1736 (2017).
18. Sandborn, W.J., *et al.* Ozanimod as Induction and Maintenance Therapy for Ulcerative Colitis. *N Engl J Med* **385**, 1280-1291 (2021).
19. Cohen, N.A. & Rubin, D.T. New targets in inflammatory bowel disease therapy: 2021. *Curr Opin Gastroenterol* **37**, 357-363 (2021).
20. Chang, J.T. Pathophysiology of Inflammatory Bowel Diseases. *N Engl J Med* **383**, 2652-2664 (2020).
21. Stalgis, C., Deepak, P., Mehandru, S. & Colombel, J.F. Rational Combination Therapy to Overcome the Plateau of Drug Efficacy in Inflammatory Bowel Disease. *Gastroenterology* **161**, 394-399 (2021).

22. Friedrich, M., *et al.* IL-1-driven stromal-neutrophil interactions define a subset of patients with inflammatory bowel disease that does not respond to therapies. *Nat Med* **27**, 1970-1981 (2021).
23. Smillie, C.S., *et al.* Intra- and Inter-cellular Rewiring of the Human Colon during Ulcerative Colitis. *Cell* **178**, 714-730 e722 (2019).
24. Martin, J.C., *et al.* Single-Cell Analysis of Crohn's Disease Lesions Identifies a Pathogenic Cellular Module Associated with Resistance to Anti-TNF Therapy. *Cell* **178**, 1493-1508 e1420 (2019).
25. Parikh, K., *et al.* Colonic epithelial cell diversity in health and inflammatory bowel disease. *Nature* **567**, 49-55 (2019).
26. Jaeger, N., *et al.* Single-cell analyses of Crohn's disease tissues reveal intestinal intraepithelial T cells heterogeneity and altered subset distributions. *Nat Commun* **12**, 1921 (2021).
27. Kinchen, J., *et al.* Structural Remodeling of the Human Colonic Mesenchyme in Inflammatory Bowel Disease. *Cell* **175**, 372-386 e317 (2018).
28. Uzzan, M., *et al.* Ulcerative colitis is characterized by a plasmablast-skewed humoral response associated with disease activity. *Nat Med* **28**, 766-779 (2022).
29. Kanehisa, M. The KEGG database. *Novartis Found Symp* **247**, 91-101; discussion 101-103, 119-128, 244-152 (2002).
30. Zhang, W., *et al.* Gut-innervating nociceptors regulate the intestinal microbiota to promote tissue protection. *Cell* **185**, 4170-4189 e4120 (2022).
31. Yang, D., *et al.* Nociceptor neurons direct goblet cells via a CGRP-RAMP1 axis to drive mucus production and gut barrier protection. *Cell* **185**, 4190-4205 e4125 (2022).
32. Brenna, O., *et al.* Cellular localization of guanylin and uroguanylin mRNAs in human and rat duodenal and colonic mucosa. *Cell Tissue Res* **365**, 331-341 (2016).
33. Wang, R.X., Lee, J.S., Campbell, E.L. & Colgan, S.P. Microbiota-derived butyrate dynamically regulates intestinal homeostasis through regulation of actin-associated protein synaptopodin. *Proc Natl Acad Sci U S A* **117**, 11648-11657 (2020).
34. Gallo, R.L. & Hooper, L.V. Epithelial antimicrobial defence of the skin and intestine. *Nat Rev Immunol* **12**, 503-516 (2012).
35. Marchal-Bressenot, A., *et al.* Development and validation of the Nancy histological index for UC. *Gut* **66**, 43-49 (2017).
36. Kawano, T., *et al.* ER proteostasis regulators cell-non-autonomously control sleep. *Cell Rep* **42**, 112267 (2023).
37. Zhang, Y., *et al.* FUT2 Facilitates Autophagy and Suppresses Apoptosis via p53 and JNK Signaling in Lung Adenocarcinoma Cells. *Cells* **11**(2022).
38. Dorward, D.A., *et al.* The role of formylated peptides and formyl peptide receptor 1 in governing neutrophil function during acute inflammation. *Am J Pathol* **185**, 1172-1184 (2015).
39. Caer, C., *et al.* TREM-1+ Macrophages Define a Pathogenic Cell Subset in the Intestine of Crohn's Disease Patients. *J Crohns Colitis* **15**, 1346-1361 (2021).
40. Giudici, F., *et al.* Multiplex gene expression profile in inflamed mucosa of patients with Crohn's disease ileal localization: A pilot study. *World J Clin Cases* **7**, 2463-2476 (2019).
41. Hau, C.S., *et al.* Lipocalin-2 exacerbates psoriasiform skin inflammation by augmenting T-helper 17 response. *J Dermatol* **43**, 785-794 (2016).
42. Shi, X., *et al.* Combinatorial GxGxE CRISPR screen identifies SLC25A39 in mitochondrial glutathione transport linking iron homeostasis to OXPHOS. *Nat Commun* **13**, 2483 (2022).
43. Kakiuchi, N., *et al.* Frequent mutations that converge on the NFKBIZ pathway in ulcerative colitis. *Nature* **577**, 260-265 (2020).
44. Cording, S., *et al.* Oncogenetic landscape of lymphomagenesis in coeliac disease. *Gut* **71**, 497-508 (2022).

45. King, K.R., *et al.* IRF3 and type I interferons fuel a fatal response to myocardial infarction. *Nat Med* **23**, 1481-1487 (2017).
46. Bengtsson, A.A. & Ronnblom, L. Role of interferons in SLE. *Best Pract Res Clin Rheumatol* **31**, 415-428 (2017).
47. 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).
48. Dong, R. & Yuan, G.C. SpatialDWLS: accurate deconvolution of spatial transcriptomic data. *Genome Biol* **22**, 145 (2021).
49. Arijis, I., *et al.* Effect of vedolizumab (anti-alpha4beta7-integrin) therapy on histological healing and mucosal gene expression in patients with UC. *Gut* **67**, 43-52 (2018).
50. Toedter, G., *et al.* Gene expression profiling and response signatures associated with differential responses to infliximab treatment in ulcerative colitis. *Am J Gastroenterol* **106**, 1272-1280 (2011).
51. Rubin, D.T., *et al.* Etrolizumab versus adalimumab or placebo as induction therapy for moderately to severely active ulcerative colitis (HIBISCUS): two phase 3 randomised, controlled trials. *Lancet Gastroenterol Hepatol* **7**, 17-27 (2022).
52. Peterson, L.W. & Artis, D. Intestinal epithelial cells: regulators of barrier function and immune homeostasis. *Nat Rev Immunol* **14**, 141-153 (2014).
53. Groh, V., Steinle, A., Bauer, S. & Spies, T. Recognition of stress-induced MHC molecules by intestinal epithelial gammadelta T cells. *Science* **279**, 1737-1740 (1998).
54. Bergstrom, K.S., *et al.* Goblet Cell Derived RELM-beta Recruits CD4+ T Cells during Infectious Colitis to Promote Protective Intestinal Epithelial Cell Proliferation. *PLoS Pathog* **11**, e1005108 (2015).
55. Mayassi, T., *et al.* Chronic Inflammation Permanently Reshapes Tissue-Resident Immunity in Celiac Disease. *Cell* **176**, 967-981 e919 (2019).
56. Chen, Y., Chou, K., Fuchs, E., Havran, W.L. & Boismenu, R. Protection of the intestinal mucosa by intraepithelial gamma delta T cells. *Proc Natl Acad Sci U S A* **99**, 14338-14343 (2002).
57. Kuhl, A.A., *et al.* Aggravation of intestinal inflammation by depletion/deficiency of gammadelta T cells in different types of IBD animal models. *J Leukoc Biol* **81**, 168-175 (2007).
58. Bhagat, G., *et al.* Small intestinal CD8+TCRgammadelta+NKG2A+ intraepithelial lymphocytes have attributes of regulatory cells in patients with celiac disease. *J Clin Invest* **118**, 281-293 (2008).
59. Hu, M.D., *et al.* gammadelta Intraepithelial Lymphocytes Facilitate Pathological Epithelial Cell Shedding Via CD103-Mediated Granzyme Release. *Gastroenterology* **162**, 877-889 e877 (2022).
60. Alam, M.S., *et al.* TNF plays a crucial role in inflammation by signaling via T cell TNFR2. *Proc Natl Acad Sci U S A* **118**(2021).
61. Korn, T., Bettelli, E., Oukka, M. & Kuchroo, V.K. IL-17 and Th17 Cells. *Annu Rev Immunol* **27**, 485-517 (2009).
62. Shale, M., Schiering, C. & Powrie, F. CD4(+) T-cell subsets in intestinal inflammation. *Immunol Rev* **252**, 164-182 (2013).
63. Joosse, M.E., *et al.* Duplication of the IL2RA locus causes excessive IL-2 signaling and may predispose to very early onset colitis. *Mucosal Immunol* **14**, 1172-1182 (2021).
64. Shalapour, S., *et al.* Interleukin-7 links T lymphocyte and intestinal epithelial cell homeostasis. *PLoS One* **7**, e31939 (2012).
65. Gerlach, K., *et al.* TH9 cells that express the transcription factor PU.1 drive T cell-mediated colitis via IL-9 receptor signaling in intestinal epithelial cells. *Nat Immunol* **15**, 676-686 (2014).

66. Uchiyama, K., *et al.* Increased mucosal IL-12 expression is associated with relapse of ulcerative colitis. *BMC Gastroenterol* **21**, 122 (2021).
67. Bain, C.C. & Mowat, A.M. Macrophages in intestinal homeostasis and inflammation. *Immunol Rev* **260**, 102-117 (2014).
68. Viola, M.F. & Boeckxstaens, G. Intestinal resident macrophages: Multitaskers of the gut. *Neurogastroenterol Motil* **32**, e13843 (2020).
69. Chen, K., *et al.* Formylpeptide receptor-2 contributes to colonic epithelial homeostasis, inflammation, and tumorigenesis. *J Clin Invest* **123**, 1694-1704 (2013).
70. Asano, K., *et al.* Impact of allele copy number of polymorphisms in FCGR3A and FCGR3B genes on susceptibility to ulcerative colitis. *Inflamm Bowel Dis* **19**, 2061-2068 (2013).
71. Schenk, M., Bouchon, A., Seibold, F. & Mueller, C. TREM-1--expressing intestinal macrophages crucially amplify chronic inflammation in experimental colitis and inflammatory bowel diseases. *J Clin Invest* **117**, 3097-3106 (2007).
72. Foroutan, P., *et al.* Expression of CXC Chemokine Receptors in Acute Ulcerative Colitis: Initial Study from an Animal Model. *Adv Biomed Res* **8**, 56 (2019).
73. Davis, T.A., *et al.* Serum Amyloid A Promotes Inflammation-Associated Damage and Tumorigenesis in a Mouse Model of Colitis-Associated Cancer. *Cell Mol Gastroenterol Hepatol* **12**, 1329-1341 (2021).
74. Low, E.N.D., Mokhtar, N.M., Wong, Z. & Raja Ali, R.A. Colonic Mucosal Transcriptomic Changes in Patients with Long-Duration Ulcerative Colitis Revealed Colitis-Associated Cancer Pathways. *J Crohns Colitis* **13**, 755-763 (2019).
75. Zhou, C., *et al.* SCFAs induce autophagy in intestinal epithelial cells and relieve colitis by stabilizing HIF-1alpha. *J Mol Med (Berl)* **98**, 1189-1202 (2020).
76. Yamaguchi, S., *et al.* Fecal calprotectin is a clinically relevant biomarker of mucosal healing in patients with quiescent ulcerative colitis. *J Gastroenterol Hepatol* **31**, 93-98 (2016).
77. Wakai, M., *et al.* Serum amyloid A is a better predictive biomarker of mucosal healing than C-reactive protein in ulcerative colitis in clinical remission. *BMC Gastroenterol* **20**, 85 (2020).
78. Narula, N., *et al.* Early Change in Epithelial Neutrophilic Infiltrate Predicts Long-Term Response to Biologics in Ulcerative Colitis. *Clin Gastroenterol Hepatol* **20**, 1095-1104 e1099 (2022).
79. Burr, N.E., Gracie, D.J., Black, C.J. & Ford, A.C. Efficacy of biological therapies and small molecules in moderate to severe ulcerative colitis: systematic review and network meta-analysis. *Gut* (2021).

## Materials and Methods

### Primary cohort

Subjects with UC and NV (Primary cohort) were recruited, from the Inflammatory Bowel Disease (IBD) Center, the Gastroenterology Department, and the Endoscopy Unit at Mount Sinai Hospital. The protocol was approved by the Institutional Review Board #17-01304 and informed consent was obtained from all participants. Patients who were enrolled in the study were asked to provide research biopsies during the standard-of-care colonoscopy. All samples were fully deidentified before analyses. The public datasets derived from the patients will be made available after the

acceptance of the paper. Clinical characteristics of patients with research biopsies in the Primary Cohort are displayed in Supplementary Table 1.

### **Validation cohorts**

*Validation Cohort-1 (VC-1)*<sup>1</sup>: GSE193677 series include RNAseq-based gene expression profiles from gut biopsies taken at endoscopy visit for 401 MSCCR UC patients and 243 non-IBD controls. Multiple biopsies from rectum and colonic regions were available for each patient, and they were defined as inflamed or non-inflamed based on endoscopy. Besides clinical and endoscopic disease severity measures, histological scores were also available for a subset of UC patients (n=178). Gene expression matrices from biopsy were generated from the count matrices using the voom transformation as previously described<sup>1,2</sup>. Expression profiles adjusted for technical variables (batch, rRNA rate, exonic rate and RIN score) using all the MSCCR cohort and after subsetting to our population of interest, used to generate signature scores.

*Validation Cohort-2 (VC-2)*<sup>3</sup>: GSE73661 series includes colonic gene expression (Affymetrix, HGU-1.0 ST) profiles from colonic biopsies from 23 UC patients before and 4-6 weeks after first infliximab (IFX) treatment. Response to therapy was defined as endoscopic mucosal healing (Mayo endoscopic score 0-1).

*Validation Cohort-3 (VC-3)*<sup>4</sup>: GSE23597 series includes mRNA expression of 113 mucosal colonic biopsies taken from 48 moderate-to-severe UC patients (6-12 Mayo score) at baseline and 8 and 30 weeks after treatment with 5 mg/kg (n=15) or 10 mg/kg (n=17) of infliximab or placebo (n=13). Patients were enrolled in the ACT1 randomized clinical trial (RCT, NCT00036439). Patients were considered responsive to therapy after 30 weeks if the total Mayo score decreased by at least 3 points from the baseline with an accompanying decrease in the sub-score for rectal bleeding of at least 1 point or an absolute sub-score for rectal bleeding of 0 or 1.

*Validation Cohort-4 (VC-4)*<sup>5</sup>: This cohort includes bulk RNAseq gene expression profiles of colonic biopsies collected from patients enrolled in the Hibiscus I trial (ClinicalTrials.gov: NCT02163759), a multicenter phase 3, randomized double-blind placebo-controlled and active-controlled study of etrolizumab, adalimumab and placebo in adult patients with moderately to severely acute ulcerative colitis (6-12 Mayo score, with endoscopic subscore  $\geq 2$ , and rectal bleeding and stool frequency subscores  $\geq 1$ ). This dataset includes samples collected from patients in the placebo (n = 70) and adalimumab (n = 134) arms before treatment and at week 10 after treatment. MCS remission is defined as a Mayo clinic score of 2 or lower, a stool frequency and hybrid endoscopy subscore of 1 or lower and a rectal bleeding subscore of 0. Endoscopic improvement is defined as an endoscopic subscore of 1 or lower.

### **Preparation of single cell suspension from pinch biopsies**

Eight to 12 biopsies from inflamed and/or noninflamed parts of the colon of each patient were collected in RPMI on ice. Samples were processed within 2 h. Biopsies were first transferred in the dissociation medium consisting of 10 ml HBSS (free of calcium and magnesium) containing EDTA (0.5M, pH 8, Invitrogen) and HEPES (1M, Lonza) and incubated for 40 min at 37°C. The biopsies were then vortexed, washed with HBSS. After 40 min, the supernatant containing cells was removed, filtered through a cell strainer, spun down and resuspended. This cell suspension was referred to as the 'epithelial cell' fraction. To obtain LP cells, the remaining tissue was transferred into digestion media (10 ml of RPMI containing 0.005 mg of collagenase IV (Sigma-Aldrich) and DNase I (Sigma-Aldrich). After an incubation of 40 minutes at 37°C in a rotating incubator (180 rpm), single cell suspension was made by mechanical dissociation, followed by subsequent filtering through a 100- $\mu$ m and a 40- $\mu$ m cell strainer and washed with RPMI twice. The digestion supernatant was then centrifuged, and the pellet was resuspended. This cell suspension was referred to as 'LP cell fraction'. Both, the epithelial cell and LP fractions were resuspended into 500ul of RPMI (Gibco) containing 10% FBS. The single cell suspensions were then used for different assays as detailed.

### **RNA extraction for IIC<sub>bulk</sub> and Epi<sub>bulk</sub> RNA sequencing**

RLT plus buffer was added to the cell pellet obtained post EDTA dissociation (Epi<sub>bulk</sub>) and post-percoll (IIC<sub>bulk</sub>) and vortexed to obtain the lysate. The lysate was transferred to a gDNA eliminator spin column and centrifuged. Flow through was collected while the column was discarded. 1 volume of 70% ethanol was added to the flow-through and mixed well by pipetting. The sample was transferred to a RNeasy MinElute spin column and centrifuged. The flow-through was discarded and 700  $\mu$ l Buffer RW1 was added to the RNeasy MinElute spin column, centrifuged. 500  $\mu$ l Buffer RPE was added to the RNeasy MinElute spin column and centrifuged. Flow through was discarded and 500  $\mu$ l of 80% ethanol added to the RNeasy MinElute spin column and centrifuged. The collection tube was discarded, and a dry spin was run for 5min to dry the membrane. RNeasy MinElute spin column was placed in a new 1.5 ml collection tube. RNase-free water was added, and the collection tube was centrifuged for 1 min at full speed to elute the RNA.

*Pre-processing:* Reads from FASTQ format files were mapped against Homo sapiens reference genome hg38 using STAR aligner. FastQC and RseqQC were used to check quality of raw read

and aligned reads. The mapping results were analyzed by featureCounts of the Subread package to quantify transcript reads. Features with fewer than 4 normalized counts across all samples were filtered out. The counts were normalized by using *voom*<sup>6,7</sup>, which estimates the mean variance relationship and generates a precision weight for individual normalized observations. Voom expression profiles were adjusted for RIN score using a linear model<sup>8-10</sup> prior to statistical analysis.

### **3' DGE RNA-sequencing**

3'DGE RNA-sequencing was performed according to Cacchiarelli's protocol. Briefly, libraries were prepared from 10 ng of total RNA obtained from *ex-vivo* stimulated biopsies. The mRNA poly(A) tail was tagged with universal adapters, well-specific barcodes and unique molecular identifiers (UMIs) during template-switching reverse transcriptase. Barcoded cDNA from multiple samples were pooled, amplified and tagmented using a transposon-fragmentation approach, which enriches for 3' ends of cDNA. Libraries of 350–800 base pairs (bp) were sequenced with an Illumina HiSeq® 2500 using a TruSeq Rapid SBS Kit. Read pairs used for the analysis fulfilled the following QCs: all 16 bases of the first read must have quality scores of at least 10 and the first 6 bases must exactly align to a designed well-specific barcode. The second reads were aligned to RefSeq human mRNA sequences (GRCh37) using Burrows-Wheeler Aligner version 0.7.15.

### **DGE RNAseq analysis for IICb<sub>ulk</sub> and Epi<sub>bulk</sub> RNA sequencing**

DGE profiles were generated by counting for each sample the number of UMIs associated with each RefSeq genes. The raw count matrices were normalized with a variance-stabilizing transformation (vst) using the R package *DESeq2*. Differential gene expression analysis across culture conditions was performed with the *DESeq2* R package (version 1.30) and visualized using volcano plots. Genes were considered differentially expressed (DE) if adjusted p-values < 0.05. Each DE gene in the comparison of IL-17+IL-22 vs. medium conditions were centred and scaled across patients and the IL-17/IL-22 signature score was defined as the sum of their z-scores. R squared and p-values of the signature score to scaled gene correlations were computed with linear regression.

### **Sample preparation for single-cell RNA sequencing**

After single cell suspensions were prepared from the epithelial cell fraction (as described), the cell pellets were resuspended in 40% Percoll (BD Pharmigen) complemented with RPMI, 2%

FCS, and separated by density gradient centrifugation in a discontinuous Percoll gradient (70%/40%) at 2000g for 20 min at 20°C. Cells were isolated from the 70-40% interphase were referred to as intraepithelial immune cells (IIC) in the study.

### **10x Processing**

Single cell suspensions were counted using a Nexcelom Cellometer Auto2000 and each loaded on one lane of the 10x Genomics NextGem 3'v3.1 assay as per the manufacturer's protocol with a targeted cell recovery of 8,000 cells. Gene expression libraries were prepared as per 10x Genomics demonstrated protocol<sup>11</sup>. Libraries were quantified via Agilent 2100 hsDNA Bioanalyzer and KAPA library quantification kit (Roche Cat# 0796014001) and sequenced at a targeted depth of 25,000 reads per cell. Libraries were pooled in equimolar concentration and sequenced on the Illumina NovaSeq 100 cycle kit with run parameters set to 28x8x0x60 (R1xi7xi5xR2).

### **Single-cell RNA data processing and analysis**

*Filtering, Normalization, Batch correction:* Cells were filtered based on the total number of UMI counts and mitochondria genes fractions. Only cells with total UMI counts greater than 1,000 and mitochondria gene fraction less than 10% were considered for the analysis. The analysis of single-cell RNA data including data normalization and batch effect correction was performed via the Seurat package<sup>12,13</sup>. First, each cell was normalized independently using function *NormalizeData* which divides feature counts for each cell by the total counts for that cell. Data was then natural log transformed. Next, we found variable features using function *FindVariableFeatures*<sup>12,13</sup>. This function first fits a line to the relationship of log(variance) and log(mean) using local polynomial regression; then, standardizes the feature values using the observed mean and expected variance (given by the fitted line). Feature variance was then calculated based on the standardized values. Finally, we selected features which were repeatedly variable across datasets for integration via function *SelectIntegrationFeatures*<sup>12,13</sup>. We then identified anchors using the *FindIntegrationAnchors* function, which takes a list of Seurat objects as input, and use these anchors to integrate different datasets together via the *IntegrateData* function<sup>12,13</sup>. Dimensionality reduction to identify anchors was performed using reciprocal principal component analysis. For visualization, the dimensionality of each dataset was further reduced using Uniform Manifold Approximation and Projection (UMAP) implemented with Seurat functions *RunUMAP*. The PCs used to calculate the embedding were as the same as those used for clustering.



Clustering: Clustering was performed via function *Find.Cluster* available in the Seurat package, with resolution parameter  $r=0.5$ . This function identifies clusters of cells using a shared nearest neighbor algorithm. Once that unsupervised clustering was performed, we identified cluster-specific markers using function *FindMarkers* available in the Seurat package<sup>12,13</sup>. Markers were ranked based on area under the receiver operating characteristic curve (AUC). Based on the cluster-specific markers and leveraging existing signatures from the literature<sup>14</sup>, clusters were annotated into 10 cell-types<sup>12</sup>.

Cell-type signatures:

We considered the lineage cell-type categories to obtain gene-signature from the single-cell RNA data. This was performed via function *FindMarkers* available in the Seurat package. Only markers with an AUC greater than 0.7 were considered as cell-type specific markers. Mitochondrial and ribosomal protein genes were removed in the final list of genes since they were found expressed in multiple cell-types<sup>12</sup>.

Re-clustering of lymphoid and myeloid compartments:

T-cells, B-cells, Plasma and Myeloid cells were re-clustered as smaller individual subsets using the same nearest neighborhood clustering algorithm<sup>12</sup>. The resolution parameters used in re-clustering were 0.5 for T cells, Plasma and Myeloid cells and 0.3 for B cells. After identifying markers for each of the subclusters we could assign most of them to specific subtypes of each cell type: Neutrophils, Macrophages, Monocyte-Macrophages and DCs for Myeloid; Memory, Naïve, Cytotoxic, T<sub>REG</sub>, Gamma-Delta ( $\gamma\delta$ ) and T<sub>H17</sub> for T cells; Memory, Naïve, and Germinal Center (GC) for B cells; and IgA and IgG isotypes for Plasma cells.

Pathway enrichment analysis:

For each cell-type compartment (i.e., T, Myeloid and Plasma cells), the log fold-change was computed between the average gene expression in cells belonging to UC and that of cells belonging to NV samples. Enrichment scores and adjusted p values (FDR) were derived via gene set enrichment analysis<sup>15</sup> for UC vs NV and NV vs UC log fold-change separately. Specifically, genes were ordered in decreasing order of fold-change and a Kolmogorov-Smirnov statistics was utilized to find pathways significantly enriched at the top of the list. A permutation-based strategy was utilized to derive p-values and FDR for different pathways<sup>15</sup>. Molecular pathway gene sets used in this analysis were from KEGG<sup>16</sup>, Reactome<sup>17</sup> and BioPlanet<sup>18</sup> databases.

## **Spatial transcriptomic sample collection, data processing and analysis**

*Sample preparation and preprocessing:* Biopsy specimen from inflamed and/or noninflamed parts of the colon of each patient were collected into OCT medium in cryomolds and snap frozen over dry ice during the colonoscopy procedure. Additional OCT medium was added to cover the cryomolds. Once frozen, samples were transported on dry ice and stored at -80 until further use. For the Visium Gene Expression library preparation, each frozen section was mounted on Tissue Optimization Slide and permeabilized for increasing times to determine optimal permeabilization incubation following the manufacturer's instructions (CG000238 Rev E, 10X Genomics). The sections were fixed in pre-chilled methanol for 30 min, H&E stained and imaged using Aperio AT2 scanner (Leica Biosystems) 20X magnification. The sections were then permeabilized for 10 min and spatially tagged cDNA libraries were constructed using the 10x Genomics Visium Spatial Gene Expression 3' Library Construction V1 Kit (CG000239 Rev F, 10X Genomics). 10  $\mu$ l of amplified cDNA from each sample was used for library construction and has undergone fragmentation, adapter ligation, index PCR, and purification. Each final library was sequenced on an Illumina NextSeq using 300 cycle high output kits with sequencing depth of ~25,000 reads per spot. Sequencing data was mapped to GRCH38-2020-A reference transcriptome using the Space Ranger software (Space ranger v1.3) to derive a feature spot-barcode expression matrix.

*Mapping, filtering, normalization:* The four Visium data sets (two UC and two NV) were initially processed by 10x Space Ranger. Higher level analysis and visualizations were performed by *Giotto* R package<sup>19</sup>. In each sample the spots with fewer than 500 unique genes were removed. Genes present in fewer than 30 spots were removed. Expression levels were normalized using default *Giotto* procedure.

*Cell Type Enrichment:* Enrichment scores were computed on normalized expression levels in each spatial spot independently using Kolmogorov-Smirnov statistics<sup>15</sup>. Enrichment scores were derived based on signatures from scRNA-seq data. Cell-type specific markers from scRNA-seq data with an AUC greater than 0.70 were considered to derive enrichment scores (Supplemental Table 8). In addition to these cell-types derived based on our scRNA-seq data, we considered different types of epithelial cells including stem cells, absorptive and secretory. Specifically, we used the gene sets- absorptive\_all, secretory\_all and stem cells from the epithelial subset<sup>14</sup> (Supplementary Table 8).

*Spatial deconvolution:*

In addition to enrichment analysis, we performed cell type deconvolution via *spatialDWLS*<sup>20</sup>. For the deconvolution, we utilized the same gene signatures described in the enrichment analysis. In addition to gene signatures, *spatialDWLS* requires as input the gene expression profile of different cell types. Briefly, we utilized the gene expression profiles of different cell-types from our scRNA-seq data. For different types of epithelial cells, we re-clustered the epithelial subset of our scRNA-seq data the same way as described for lymphoid and myeloid compartments, with resolution parameter 0.7. We obtained 15 epithelial sub-clusters and mapped some of them to epithelial absorptive, epithelial stem cells and epithelial secretory cells based on elevated levels of corresponding gene set signatures from the literature<sup>14</sup> (absorptive\_all, secretory\_all and stem cells from the epithelial subset, Supplementary Fig 5b). Epithelial sub-clusters 0, 5 and 10 were annotated as Absorptive, sub-clusters 4,6 and 9 as Secretory and sub-clusters 3 and 12 as Stem cells. The gene expression profiles corresponding to these sub-clusters were utilized to perform the deconvolution of ST (Supplementary Table 8).

#### Cell-cell colocalization:

The identification of cell types frequently co-localized together was performed using *cellProximityEnrichmentSpots* function in *Giotto* package<sup>19</sup>. As input, we used the deconvoluted fractions of different cell types in each spot. Spatial network was used to quantify cell type fractions occurrence in neighboring spots. We performed 1000 random permutations of spot locations to estimate the background co-localization distribution. The significance of the co-localization (p-value for each pair of cell types) was computed as the fraction of random permutations co-localized more frequently than the observed value.

#### **IIC and Epi bulk RNA-seq data analysis**

Differential Expression Analysis: Principal component analysis (PCA) was used for exploratory analysis. Expression profiles were modeled using linear mixed-effects models in the limma framework. Models included group variables (Inflamed, uninflamed and normal volunteers) in IIC and Epi as fixed effects, and a random intercept for each subject. Hypotheses of interests were tested using contrasts and p-values from the moderated t-test were adjusted for multiple hypotheses using the Benjamini-Hochberg approach, which controls the FDR. Genes were considered differentially expressed, if false discovery rate (FDR) <0.05; fold change (FCH) >2.0. . IIC<sub>inf</sub> and Epi<sub>inf</sub> signatures consist of DEG between inflamed and uninflamed samples of IIC and Epi compartments, respectively.

Pathways and cell type Enrichment analysis: Gene sets were tested for over-representation analysis using clusterProfiler 3.16.1, p-values from the Fisher's exact test were adjusted using Benjamini-Hochberg (BH) multiple test correction. The collection of genesets included KEGG<sup>16</sup> pathways, cell-types signatures from our scRNA-seq experiment and a published dataset<sup>14</sup>.

Defining  $IIC_{inf}$  IFX-response signature: In VC-3, expression profiles were modeled using linear mixed-effect models with the interaction between time and treatment group (5mg /10mg dose of IFX and placebo). Treatment DEG were identified as those with significant changes in expression compared to baseline (FDR<0.05 and FCH>1.5) at any time point in each treatment group. The difference between responder and non-responder to anti-TNF was assessed for treatment DEGs, using a linear model with interaction between the time, treatment dose, and response. The IFX - response signature thus included genes where the IFX-induced changes in the direction of disease resolution were significantly different between responders (R) vs non-responders (NR) at either week 8 or week 30 (FDR<1.5).  $IIC_{inf}$  IFX-response signature was thus defined as  $IIC_{inf}$  genes that were IFX-responsive by the criteria above and included 159 genes up-regulated (96 down-regulated) in inflamed gut and down-regulated (up-regulated) with treatment.

Scoring methods: We estimated the overall expression activity of various gene signatures in the VC-1-4 using gene set variation analysis (GSVA). For each gene signature ( $IIC_{inf}$ , scRNA-seq based cell type markers and subclusters), GSVA, a non-parametric and non-supervised method, estimated a sample-wise enrichment score for each gene set on the validation cohorts. Such enrichment scores, representing overall gene set activity, were then used for hypothesis testing with respect to phenotype information as described in the sections below.

Association of  $IIC_{inf}$  activity with UC severity in VC-1: Linear mixed-effect model was used with region, disease status and tissue type (inflamed vs non-inflamed) with a random intercept of patient. Association between gut signature scores with clinical and endoscopic severity for MSCCR UC patients were analyzed with a mixed-effect model including clinical activity or disease severity, age, sex and region as a fixed-effects. Clinical disease activity measures were based on Clinician-based Simple Clinical Colitis Activity Index (SCCAI), clinically inactive disease defined as an SCCAI < 5 and active disease as SCCAI  $\geq$  5. The Mayo endoscopic scores were used to categorize UC endoscopic severity as: normal/inactive disease (0); mild disease (1); moderate disease (2); or severe disease (3). Pearson correlation between histology scores (Nancy Index) with signature scores was estimated. The expressions of genes in the signatures were scaled as z-scores in MSCCR data before examining the correlation with Nancy index in inflamed UC biopsies.

Association with anti-TNF/ADA response in VC-2, VC-3, VC-4: The effect of IFX on the activity (GSVA-scores) of each gene signature ( $IIC_{Inf}$ , cell-types and subcluster) was compared between active treatment and placebo (using VC-2 and VC-3) and between week 30 responders and non-responders (VC-3). In the treatment effect analysis, changes in GSVA scores across time in different treatment groups (IFX 5mg/kg, IFX 10 mg/kg and placebo) were estimated by the linear mixed-effect model (LMEM) with visit, treatment and its interaction as fixed effects, and random intercept for each subject. Differences in IFX-induced activity between anti-TNF responders and non-responders was evaluated via LMEM with a three-way interaction of visit, treatment dose, and response status as fixed effects, and a random intercept for each subject. Similarly, in the treatment effect analysis for VC-4, the difference between Adalimumab-induced activity between responders and non-responders for each gene set was evaluated through a LMEM with visit, response, and their interaction term as fixed terms and a random intercept for each subject.

### **Statistical analysis:**

Statistical analysis was conducted in R version 4.1.3.

### **Flow cytometry**

Multi-parametric flow cytometry was performed on intestinal cells from patients with UC and controls, sampling the non-inflamed and inflamed sites. The gating strategy for colonic T cells, myeloid cells, plasma cells and B cells are shown in Supplementary Fig. 2 a-d respectively. Single cell suspensions were incubated with PBS containing the flow antibody cocktail for 30 min at 4°C in dark. Cells were washed twice with PBS and then fixed using 2% formaldehyde for 30 min at 4°C. Finally, cells were harvested in PBS before acquisition. Dead cells and doublets were excluded from all analyses.

### **Immunofluorescence staining**

5 µm thick sections were cut from formalin-fixed, paraffin-embedded tissue blocks. These sections were dewaxed twice in 100% xylene for 5 min. These were then rehydrated in graded alcohol starting from 100% ethanol, 90% ethanol, 80% ethanol and then 70% ethanol with 20 dips in each of the alcohol. The sections were then washed in 1X PBS twice for 5 min each. Heat-induced epitope retrieval was performed by incubating slides in a pressure cooker for 15 minutes on in target retrieval solution. Once slides cooled to room temperature, they were washed twice in PBS. Nonspecific binding was blocked with 10% goat serum for 1 hour at room temperature. Sections were incubated in primary antibodies diluted in blocking solution overnight at 4°C. Slides were

washed in 0.1% Tween-20 plus PBS thrice and then incubated in secondary antibody and DAPI (4',6-diamidino-2-phenylindole, 1  $\mu\text{g}/\text{mL}$ ) for 1 hour at room temperature. Sections were washed twice in 0.1% Tween 20 plus PBS and once in PBS and then mounted with Fluoromount-G. Controls included omitting the primary antibody (no primary control) or substituting primary antibodies with nonreactive antibodies of the same isotype (isotype control). Tissue was visualized and imaged using a Nikon Eclipse Ni microscope and digital SLR camera (Nikon, DS-Qi2).

### **Ex vivo stimulation of colonic biopsies**

Colonic biopsies (n=8) for *ex vivo* stimulations were collected in unaffected areas obtained after surgical resection of colorectal tumors (at least 10 cm distant to the tumor). Written informed consents were obtained under institutional review board protocol # DC-2008-402.

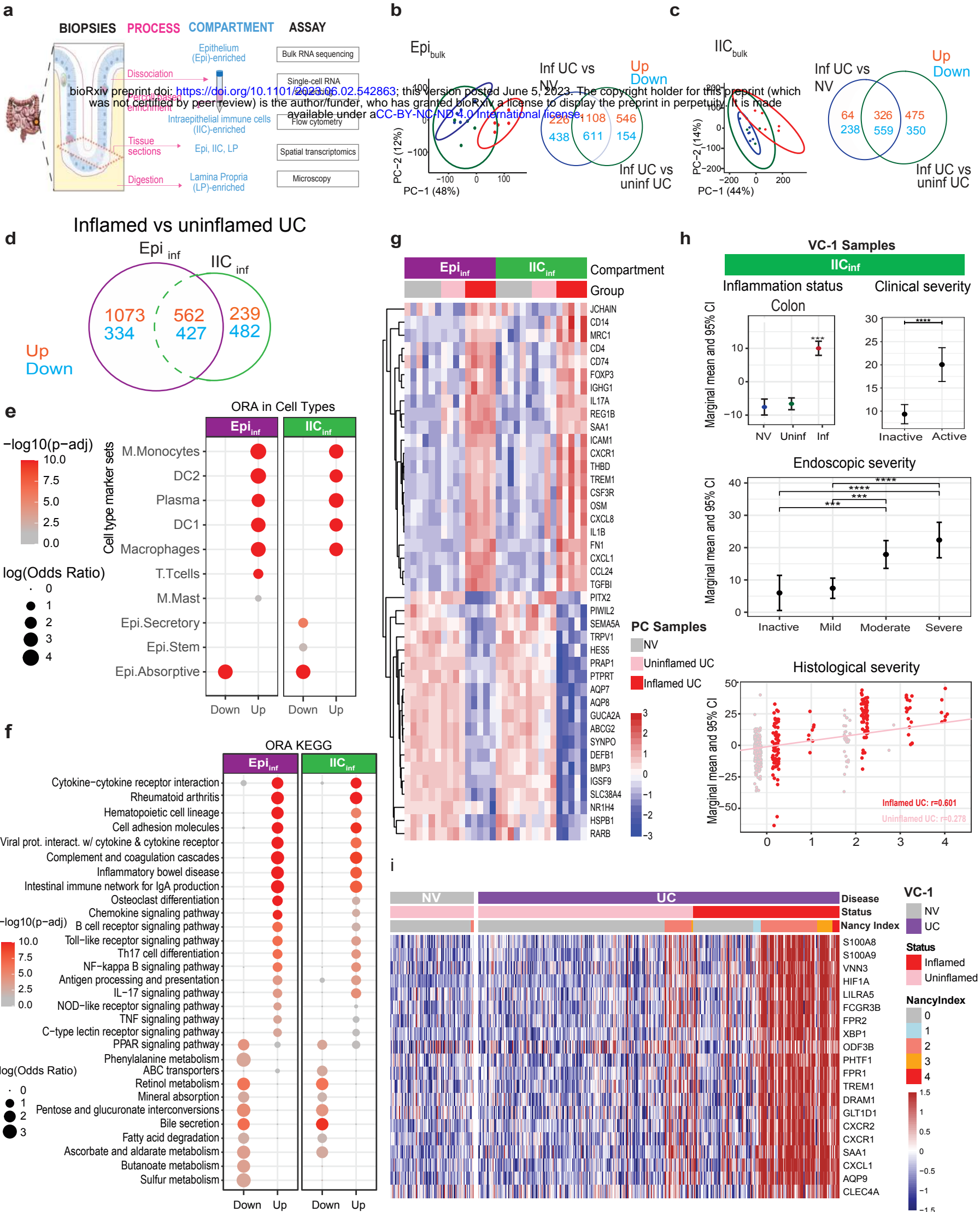
Biopsies were transported in ice-cold RPMI and processed within a maximum of 2h. Biopsies were controlled for weight and placed into 4-well Petri dishes filled with in 500 $\mu\text{L}$  serum-free medium (RPMI 1640, Gibco<sup>TM</sup>) supplemented with BSA (0.01%), 200 $\mu\text{g}/\text{mL}$  Penicillin/Streptomycin (Gibco<sup>TM</sup>; #15140-122) and 0.25 $\mu\text{g}/\text{mL}$  Fungizone (Gibco<sup>TM</sup>; #15290-026). Biopsies were cultured during 6 h at 37°C in a 95% O<sub>2</sub> / 5% CO<sub>2</sub> atmosphere on a low-speed rocking platform in the presence of cytokines (IL-22; 100ng/mL [Miltenyi, # 130-096-295], IL-17; 100ng/mL [Miltenyi, #130-093-959], IL-22BP; 200ng/mL [Miltenyi, #8498-BP-025]). Supernatants were collected and stored at -80°C until use.

*DGE RNAseq analysis for ex-vivo stimulation:* DGE profiles were generated by counting for each sample the number of UMIs associated with each RefSeq genes. The raw count matrices were normalized with a variance-stabilizing transformation (vst) using the R package *DESeq2*. Differential gene expression analysis across culture conditions was performed with the *DESeq2* R package (version 1.30) and visualized using volcano plots. Genes were considered differentially expressed (DE) if adjusted p-values < 0.05. Each DE gene in the comparison of IL-17+IL-22 vs. medium conditions were centered and scaled across patients and the IL-17/IL-22 signature score was defined as the sum of their z-scores. R squared and p-values of the signature score to scaled gene correlations were computed with linear regression.

### **References-**

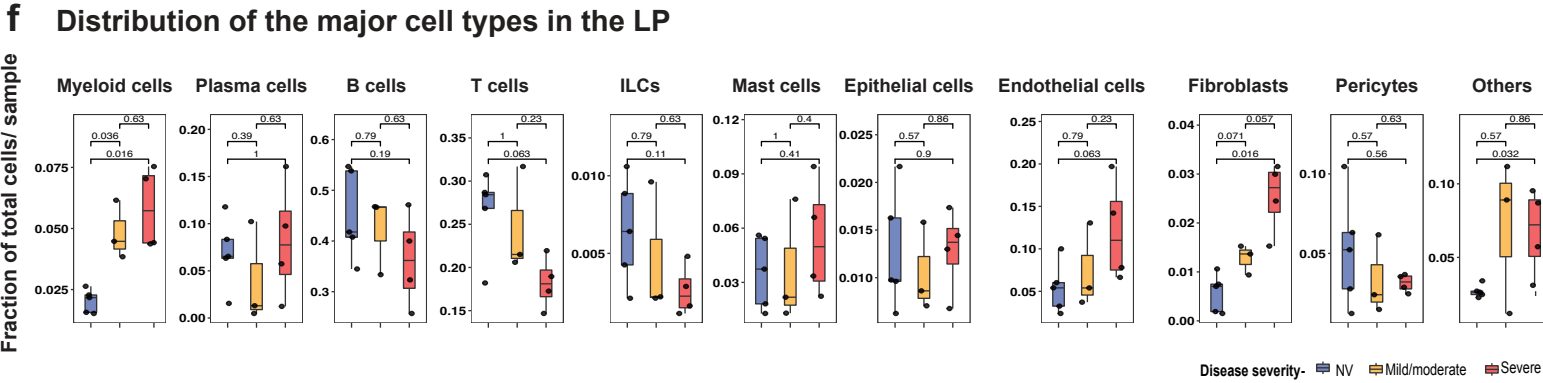
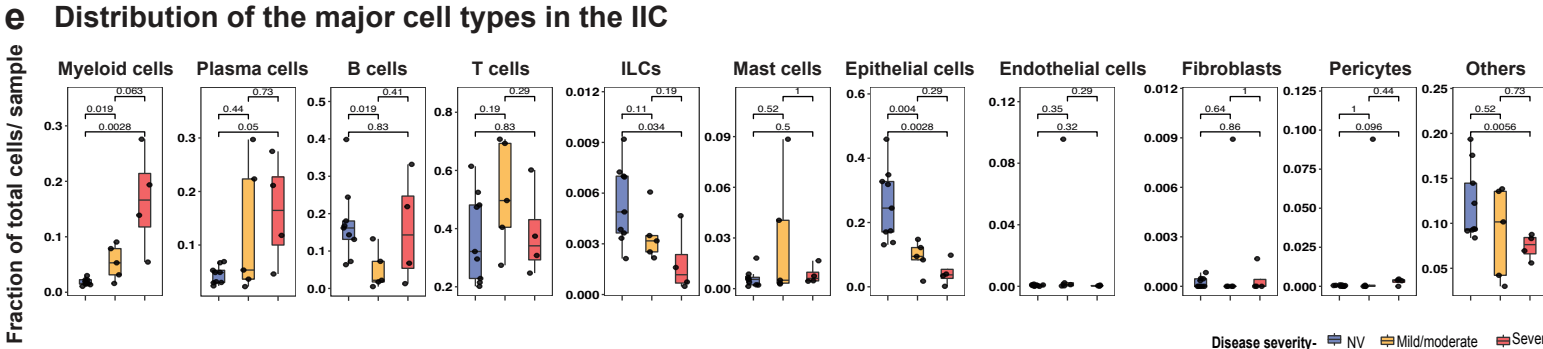
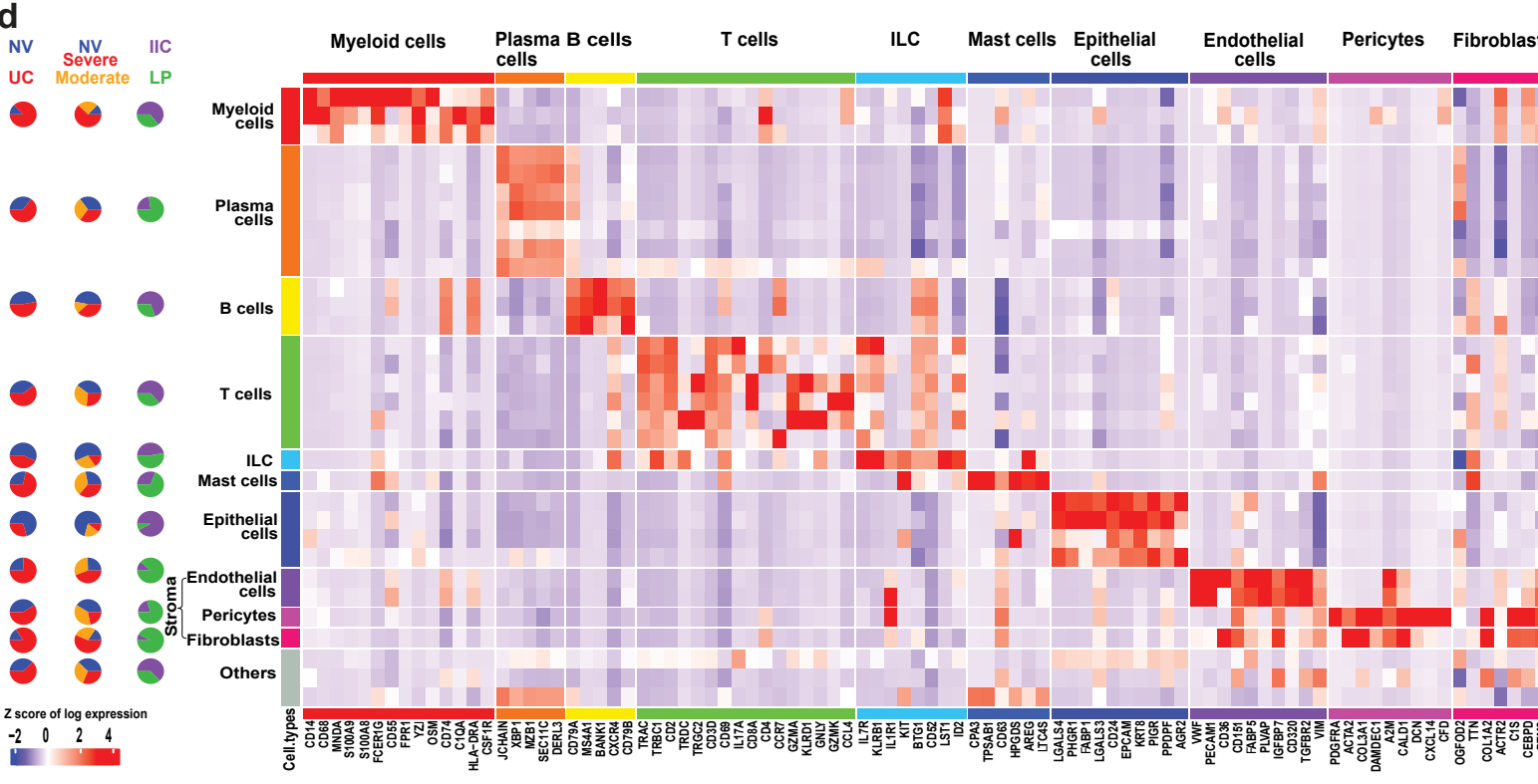
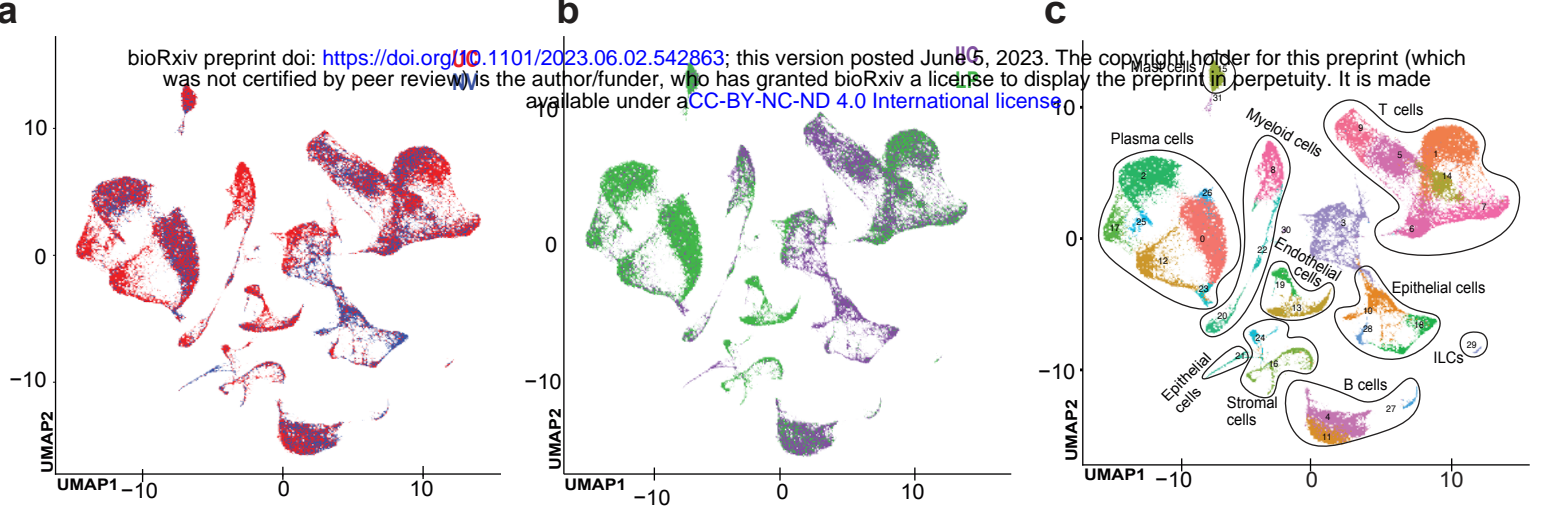
1. Suarez-Farinas, M., *et al.* Intestinal Inflammation Modulates the Expression of ACE2 and TMPRSS2 and Potentially Overlaps With the Pathogenesis of SARS-CoV-2-related Disease. *Gastroenterology* **160**, 287-301 e220 (2021).
2. Argmann, C., *et al.* Molecular Characterization of Limited Ulcerative Colitis Reveals Novel Biology and Predictors of Disease Extension. *Gastroenterology* **161**, 1953-1968 e1915 (2021).
3. Arijs, I., *et al.* Effect of vedolizumab (anti-alpha4beta7-integrin) therapy on histological healing and mucosal gene expression in patients with UC. *Gut* **67**, 43-52 (2018).
4. Toedter, G., *et al.* Gene expression profiling and response signatures associated with differential responses to infliximab treatment in ulcerative colitis. *Am J Gastroenterol* **106**, 1272-1280 (2011).
5. Rubin, D.T., *et al.* Etrolizumab versus adalimumab or placebo as induction therapy for moderately to severely active ulcerative colitis (HIBISCUS): two phase 3 randomised, controlled trials. *Lancet Gastroenterol Hepatol* **7**, 17-27 (2022).
6. Law, C.W., Chen, Y., Shi, W. & Smyth, G.K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**, R29 (2014).
7. Law, C.W., *et al.* RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Res* **5**(2016).
8. Phipson, B., Lee, S., Majewski, I.J., Alexander, W.S. & Smyth, G.K. Robust Hyperparameter Estimation Protects against Hypervariable Genes and Improves Power to Detect Differential Expression. *Ann Appl Stat* **10**, 946-963 (2016).
9. Ritchie, M.E., *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
10. Smyth, G.K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* **3**, Article3 (2004).
11. 10x, G. Chromium Next GEM Single Cell 3' Reagent Kits v3.1 User Guide. (2019).
12. Stuart, T., *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902 e1821 (2019).
13. Hao, Y., *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587 e3529 (2021).
14. Smillie, C.S., *et al.* Intra- and Inter-cellular Rewiring of the Human Colon during Ulcerative Colitis. *Cell* **178**, 714-730 e722 (2019).
15. 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).
16. Kanehisa, M. The KEGG database. *Novartis Found Symp* **247**, 91-101; discussion 101-103, 119-128, 244-152 (2002).
17. Croft, D., *et al.* Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res* **39**, D691-697 (2011).
18. Huang, R., *et al.* The NCATS BioPlanet - An Integrated Platform for Exploring the Universe of Cellular Signaling Pathways for Toxicology, Systems Biology, and Chemical Genomics. *Front Pharmacol* **10**, 445 (2019).
19. Dries, R., *et al.* Giotto: a toolbox for integrative analysis and visualization of spatial expression data. *Genome Biol* **22**, 78 (2021).
20. Dong, R. & Yuan, G.C. SpatialDWLS: accurate deconvolution of spatial transcriptomic data. *Genome Biol* **22**, 145 (2021).

**Fig 1. The intraepithelial immune compartment of patients with ulcerative colitis is remodeled with disease.**

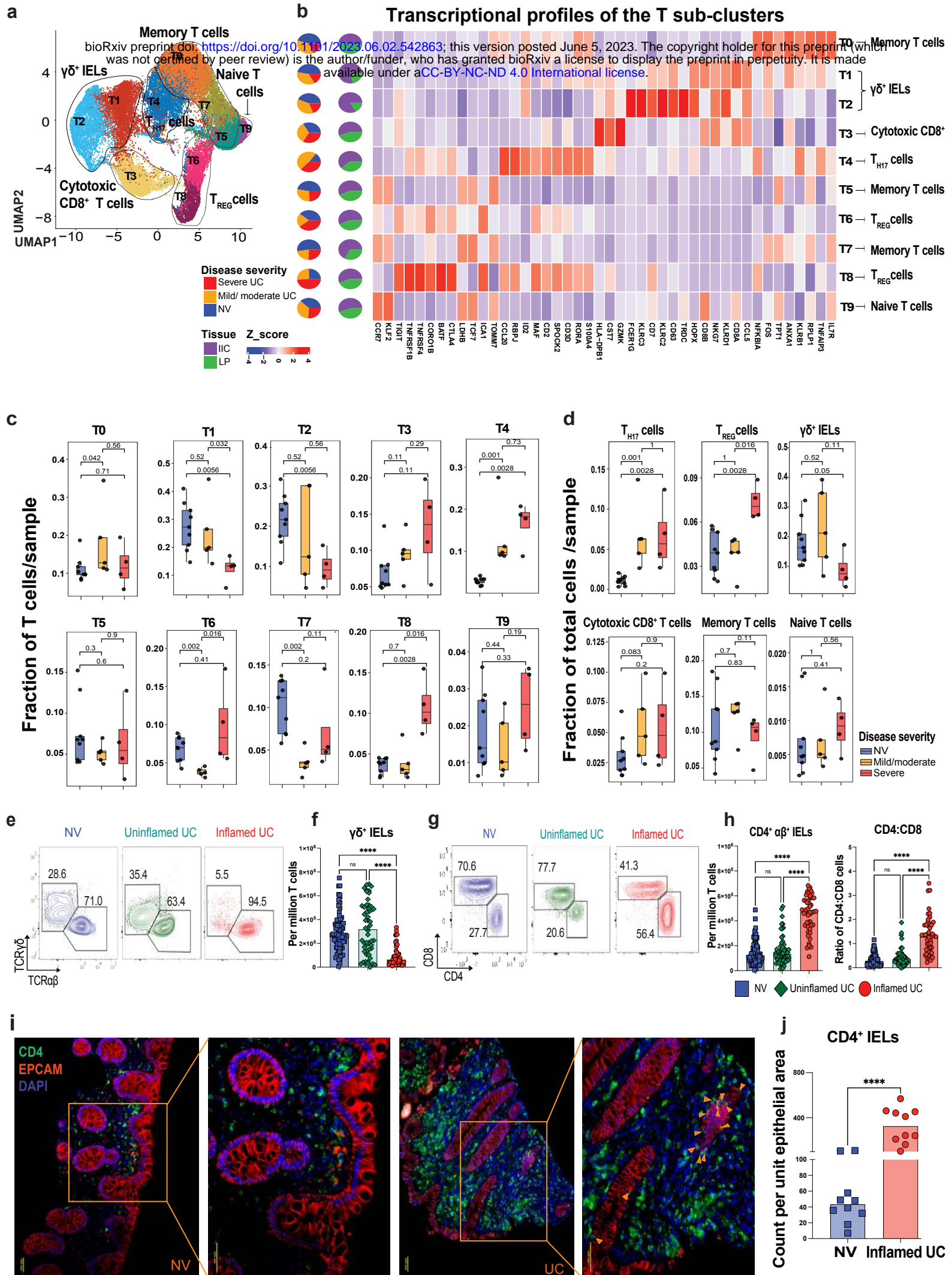




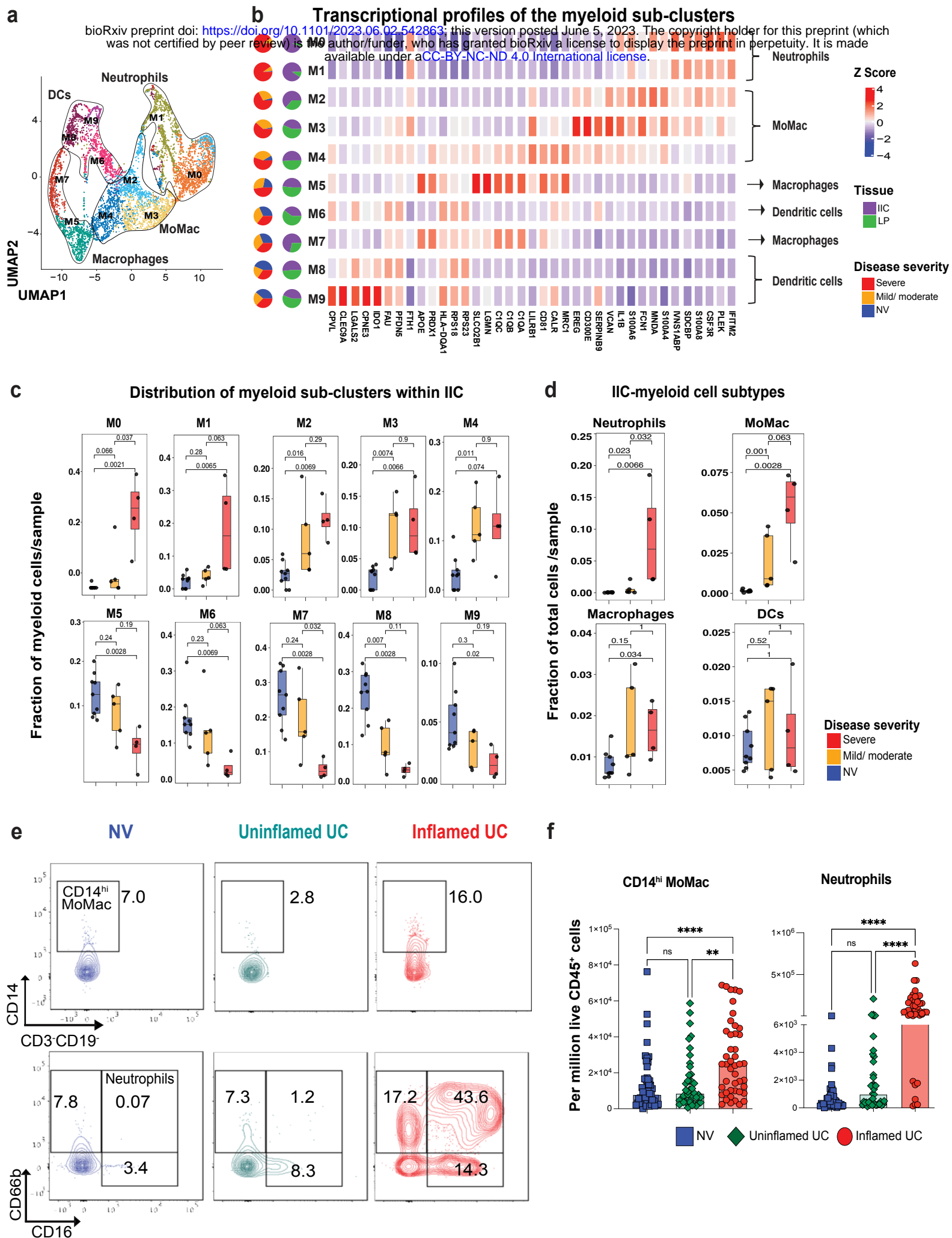
**Fig.2: High-resolution mapping of immune cells reveals distinct cellular composition in the colonic epithelial compartment of ulcerative colitis patients.**



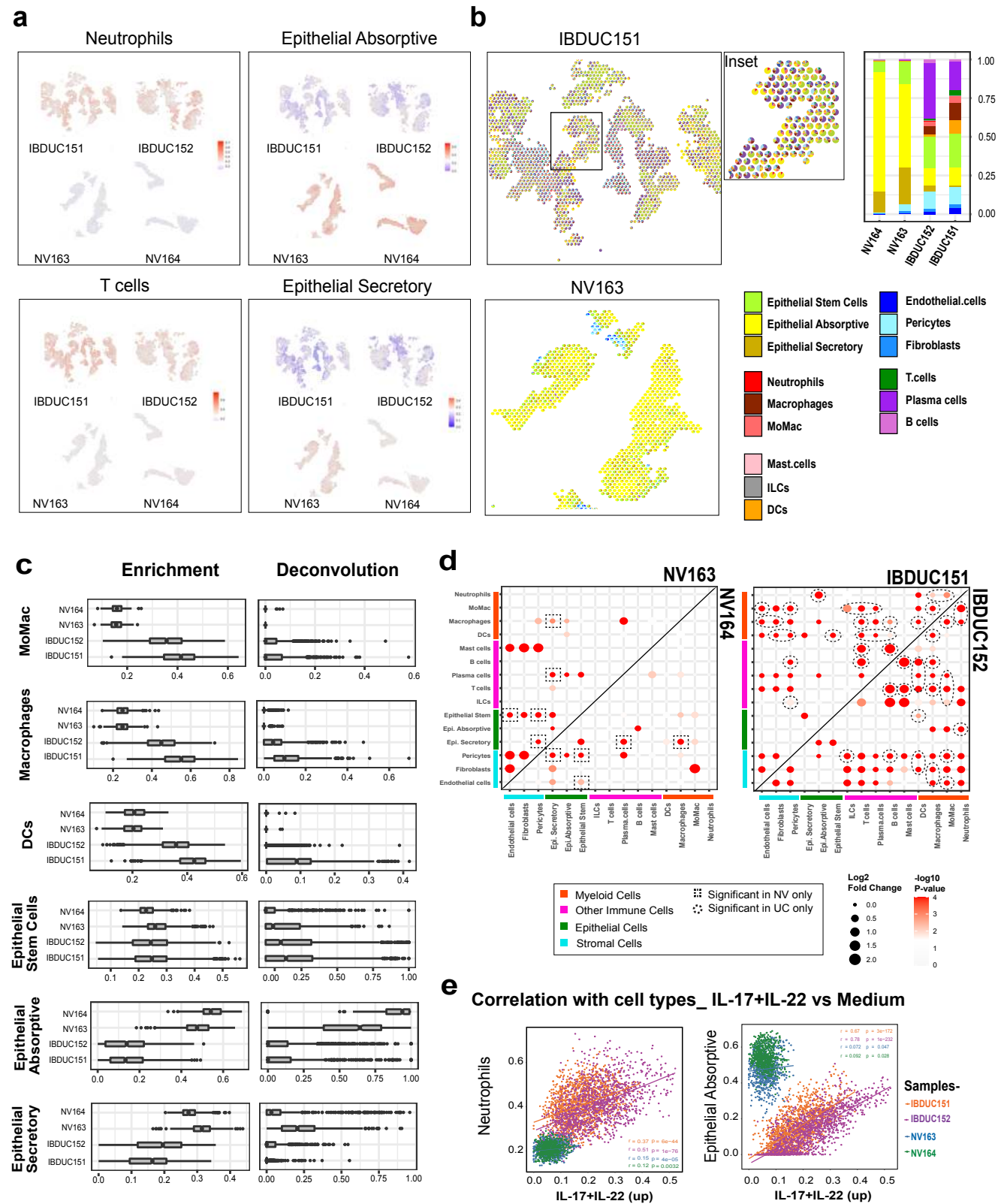
**Fig. 3: Active UC is associated with depletion of homeostatic  $\gamma\delta^+$  IELs and influx of proinflammatory  $T_{H17}$  cells into the EC.**



**Fig 4. Active UC is associated with a major influx of pro-inflammatory myeloid cells into the EC.**

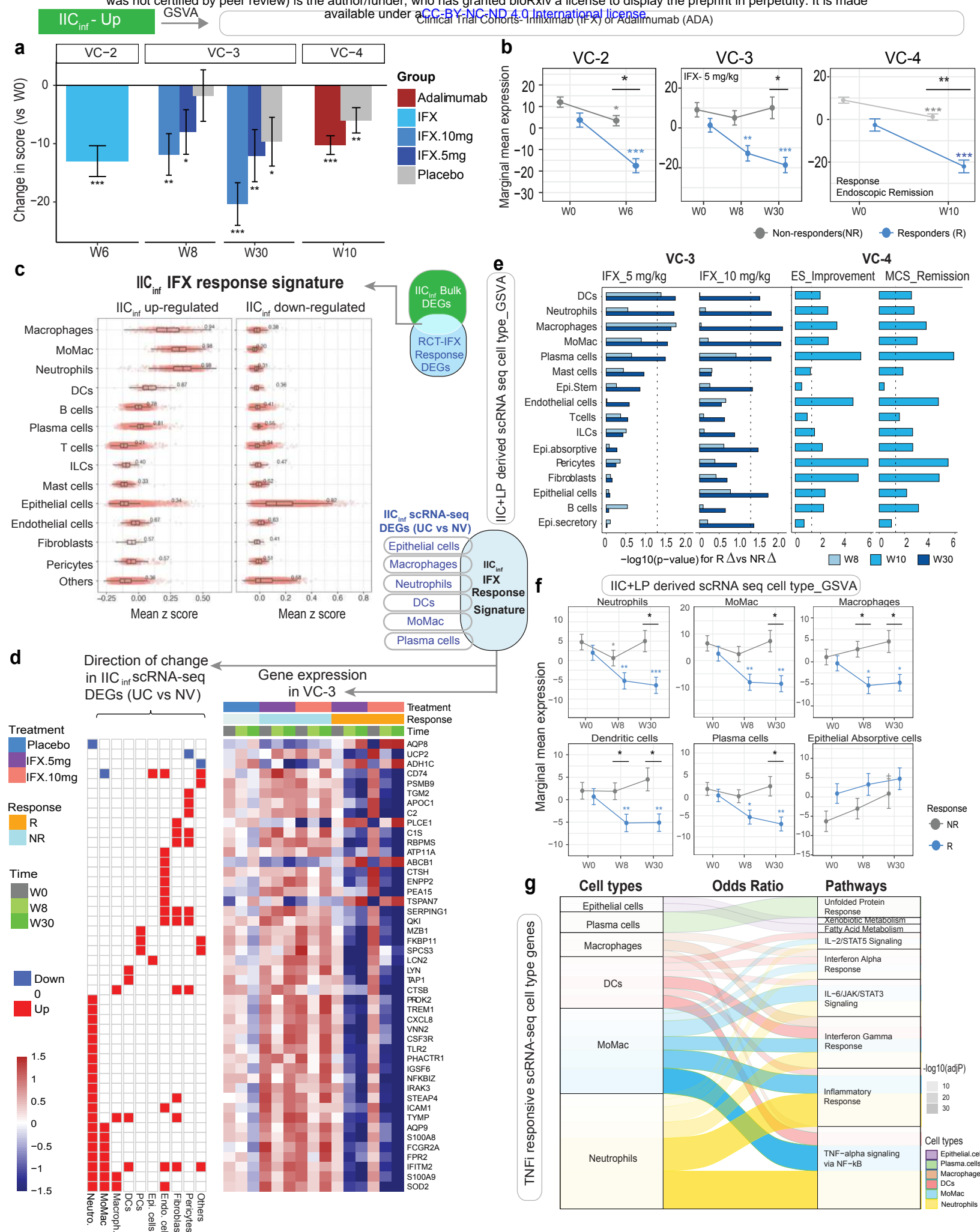


**Fig 5. Spatial transcriptomics reveals depletion of epithelial subtypes and enhanced cell-cell interactions within colonic tissues of patients with UC.**



# Fig. 6: Persistence of myeloid and plasma cell signatures in the EC associates with anti-TNF treatment non-response in three distinct validation cohorts!

bioRxiv preprint doi: <https://doi.org/10.1101/2023.06.02.126601>; this version posted June 5, 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.



## Extended Data Fig. 1: Outline of the patients and samples included in the study.

### Primary cohort\* Mount Sinai Hospital

Total n = 222

NV: n= 121

UC: n= 101

Single-cell RNA sequencing

NV: n= 9

UC: n= 9

Epi\_bulk RNA sequencing

NV: n= 6

UC: n= 5

IIC\_bulk RNA sequencing

NV: n= 6

UC: n= 5

Flow cytometry

NV: n= 110

UC: n= 94

Microscopy

NV: n= 18

UC: n= 22

### Validation Cohort - 1 (Observational cohort) Mount Sinai Crohn's and Colitis registry (MSCCR) GSE 193677, PMID-32980345

Total n = 649

NV: n= 243

UC: n= 406

### Validation Cohort - 2, GSE73661, PMID-27802155

UC: n= 23

Treatment response assessment- Mucosal healing

### Validation Cohort - 3, GSE23597, PMID- 21448149

UC: n= 48

Treatment response assessment- Clinical response

### Validation Cohort - 4, PMID: 34798036

UC: n= 204

Treatment response assessment- ES improvement  
MCS remission

UC - Ulcerative colitis

NV - Normal volunteers

Epi - Epithelial compartment (*Entire epithelial fraction*)

IIC - Intraepithelial Immune Cells (*Percoll based immune cells enriched epithelial fraction*)

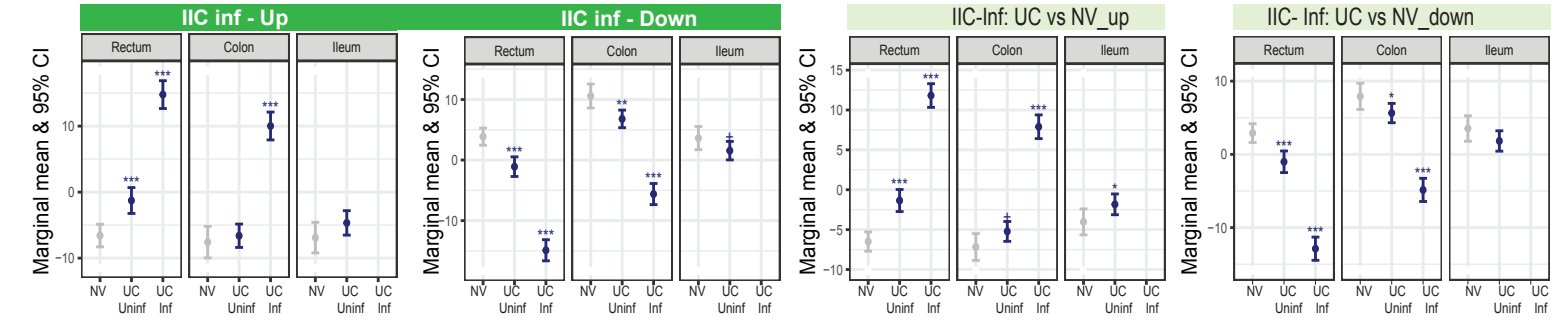
ES- Endoscopic remission Score

MCS- Mayo Clinical Score

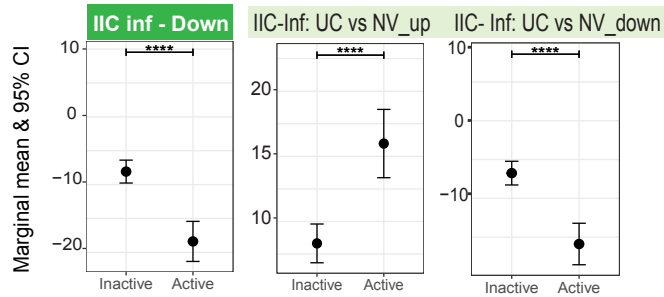
\*Clinical details and assay allocation for each patient are included in Supplementary Table1.

# Extended Data Fig.2: Association of IIC<sub>inf</sub> GSVA activity scores in biopsies from the validation cohort-1/ VC-1 stratified by inflammation status and disease severity.

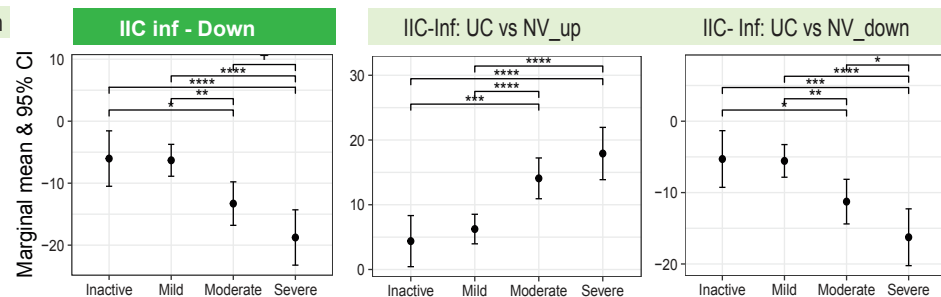
**a** GSVA scores of IIC<sub>inf</sub> genes defined in the Primary Cohort (PC) in the inflamed UC, uninfamed UC samples and healthy (NV) biopsies in the VC-1 cohort.



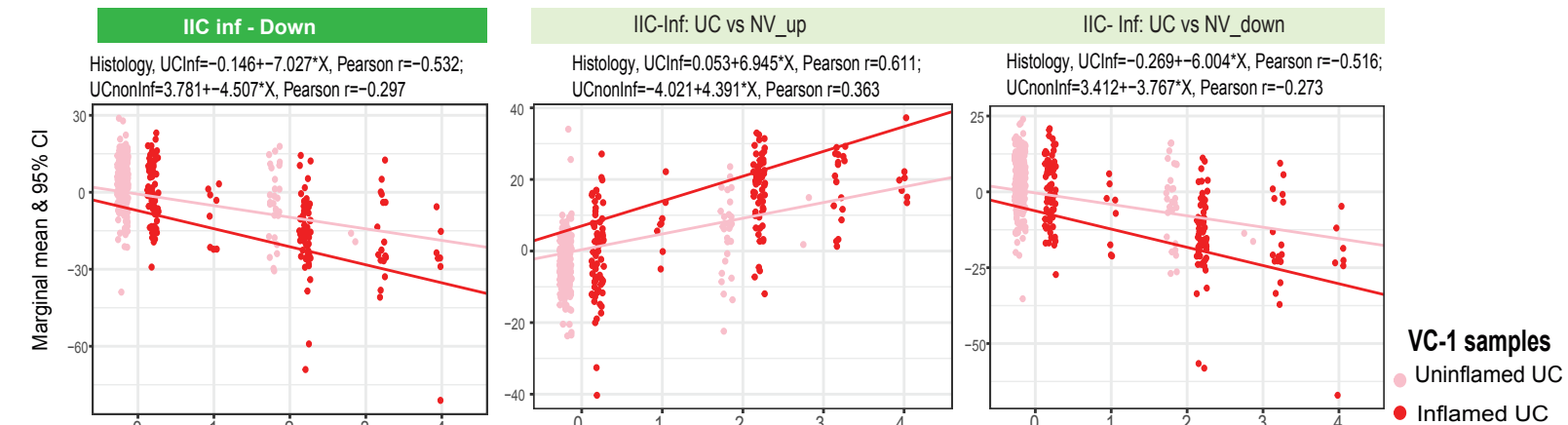
**b** IIC<sub>inf</sub> genes vs clinical severity in VC-1.



**c** IIC<sub>inf</sub> genes vs endoscopic severity in VC-1.

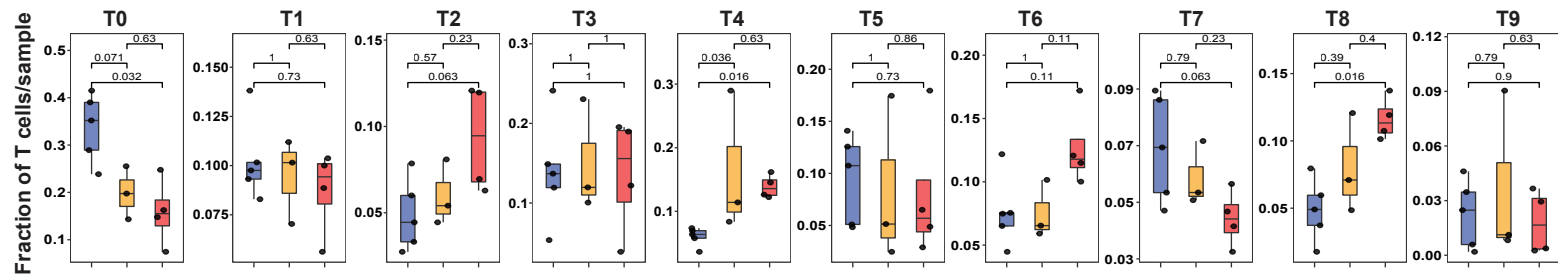


**d** Association of GSVA scores of IIC<sub>inf</sub> genes with histologic severity in VC-1.

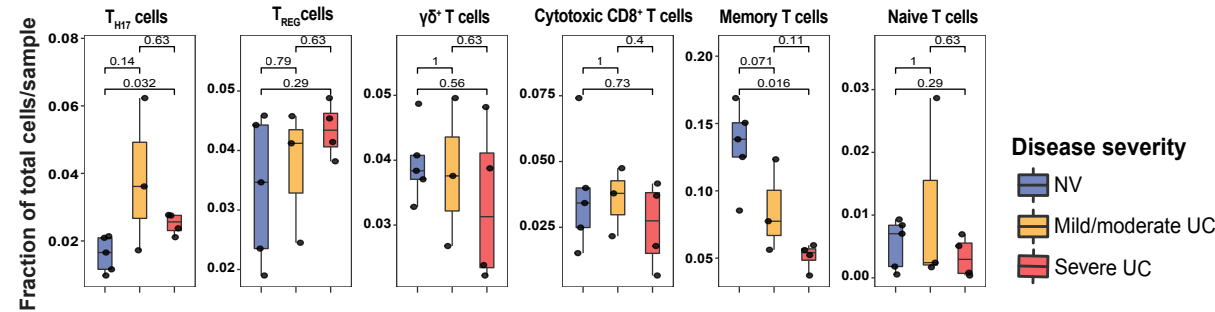


# Extended Data Fig.3: Distribution of T cell types in the LP in patients with active UC and pathway analysis of IELs.

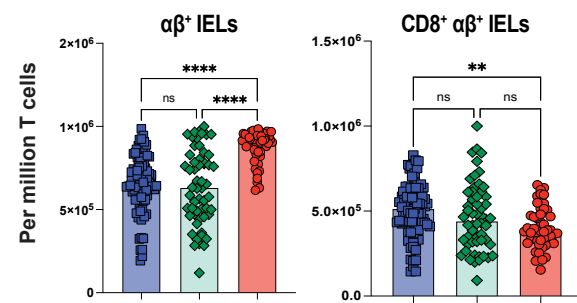
**a** Distribution of LP enriched T cell sub-clusters in the UC samples segregated by disease status and severity



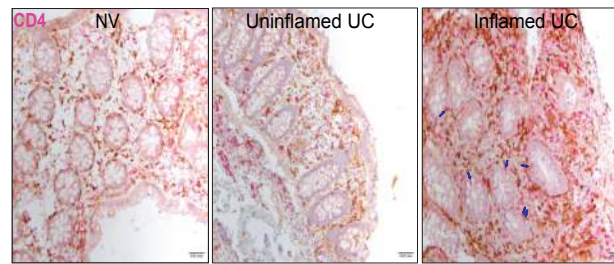
**b** Distribution of LP enriched T cell types in the UC samples segregated by disease status and severity



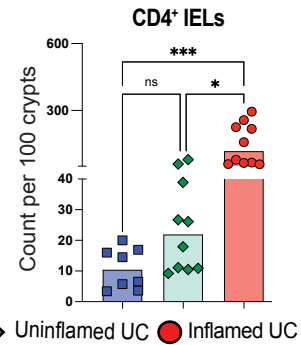
**c** Flow cytometry analyses



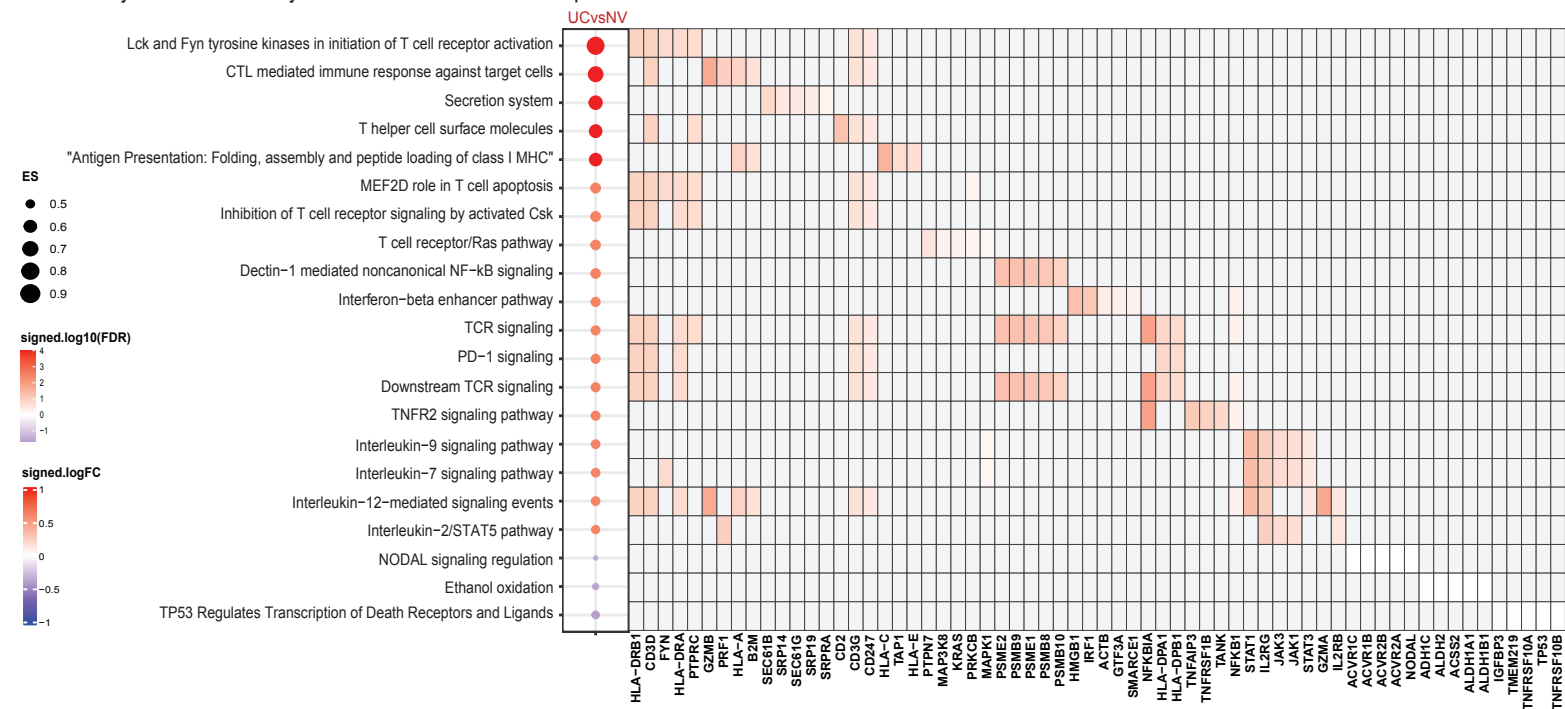
**d** Quantitative immunohistochemistry of CD4<sup>+</sup> IELs



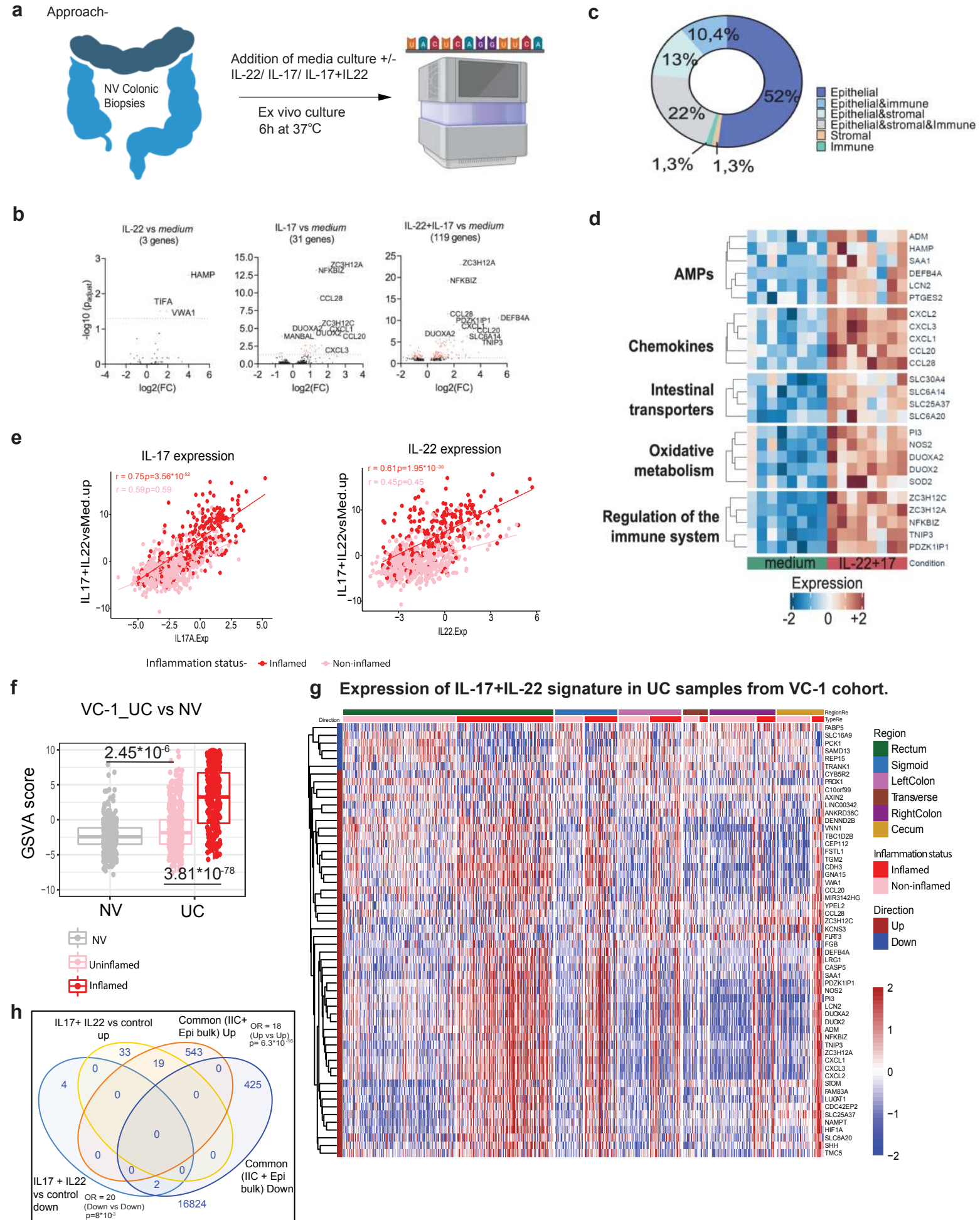
**e**



**f** Pathway enrichment analysis of the IELs in the UC samples.

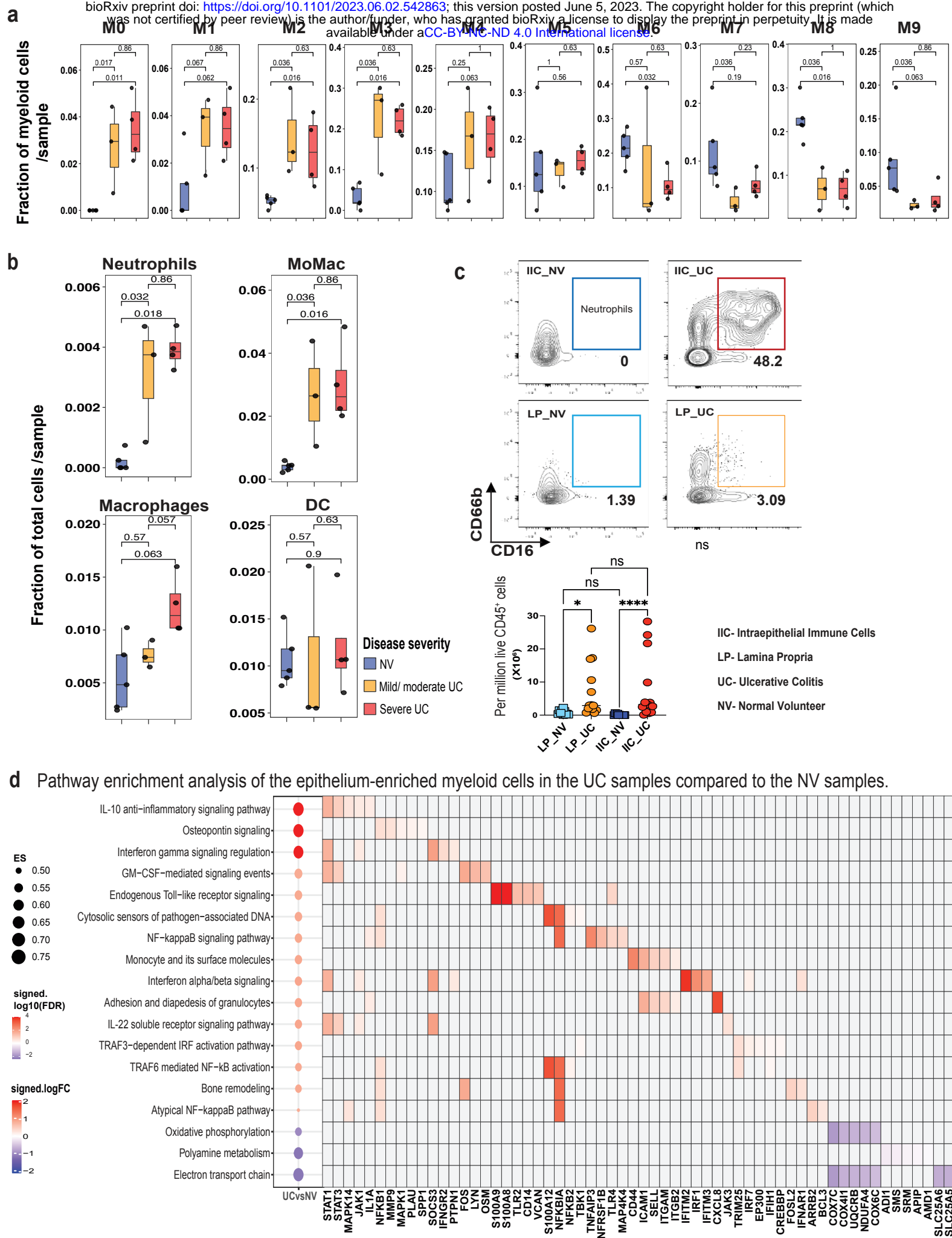




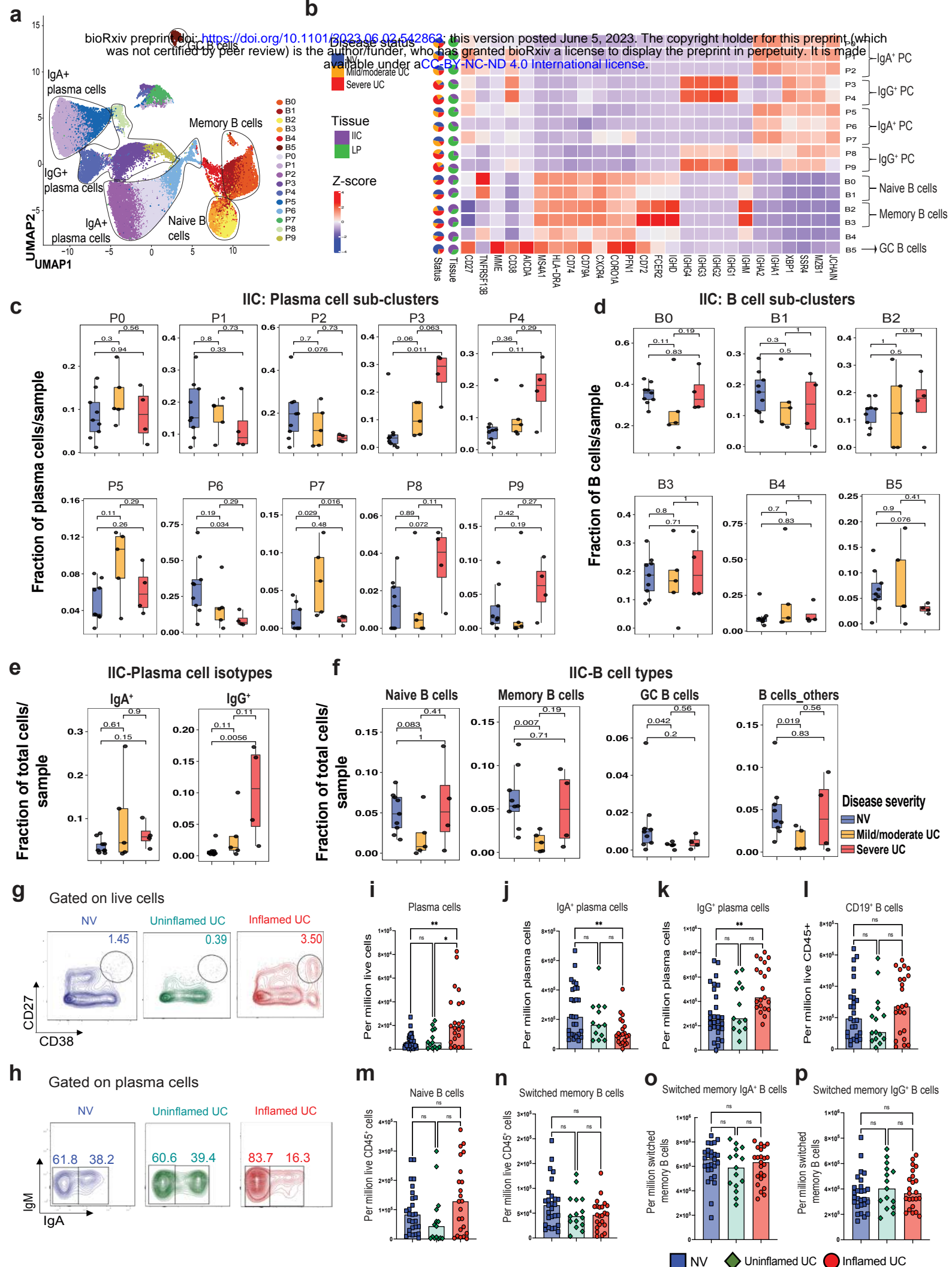


**Extended Data Fig. 5: Distribution of myeloid cell types in the LP in patients with active UC and pathway analysis of IIC enriched myeloid cell types.**

bioRxiv preprint doi: <https://doi.org/10.1101/2023.06.02.542863>; this version posted June 5, 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.

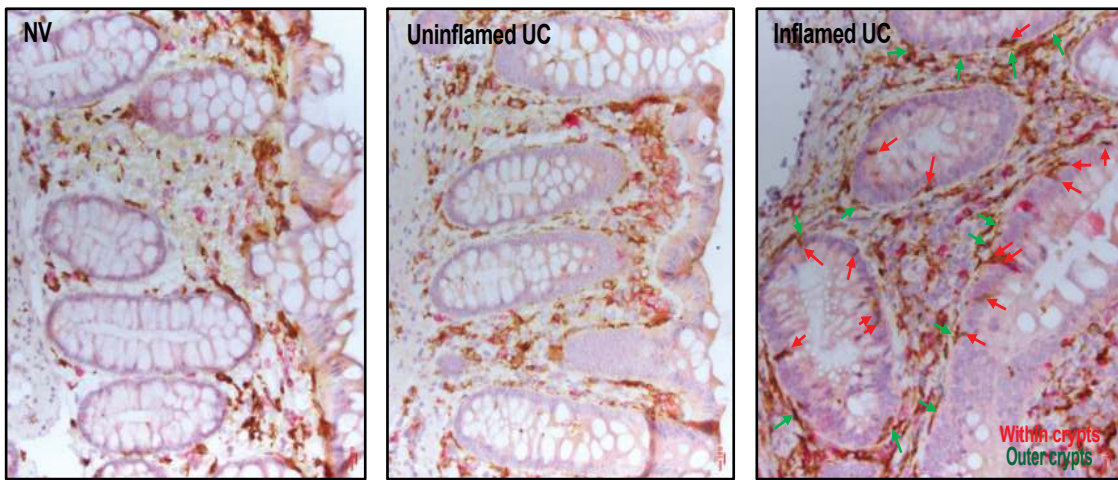


# Extended Data Fig.6: Altered transcriptomic profiles of plasma cells and B cells in active UC patients and NV.

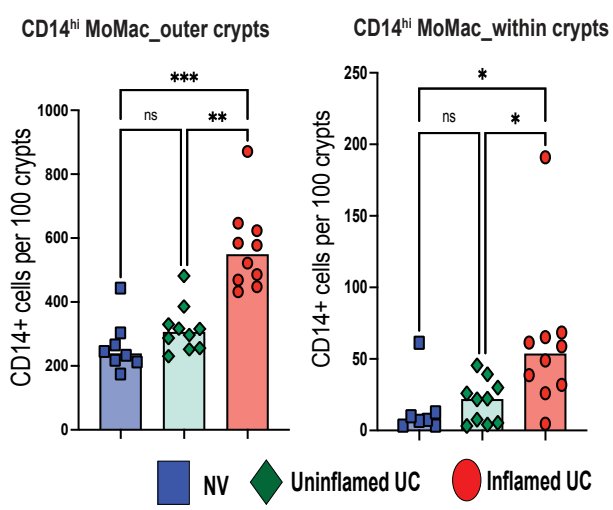


**Extended Data Fig. 7: Neutrophils, monocytes and macrophages penetrate the colonic epithelium in patients with active UC.**

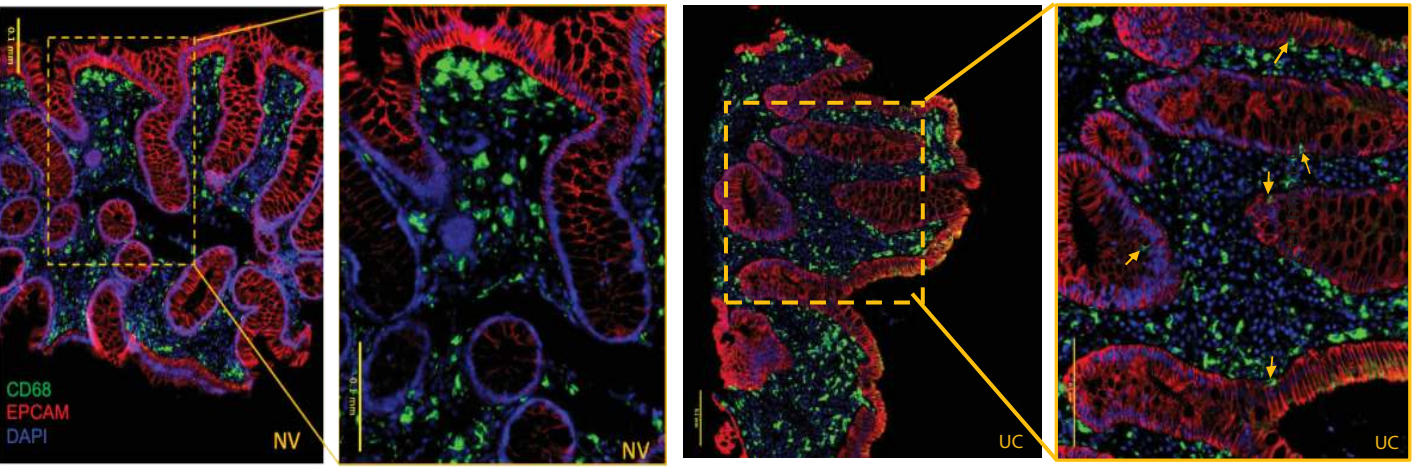
**a**



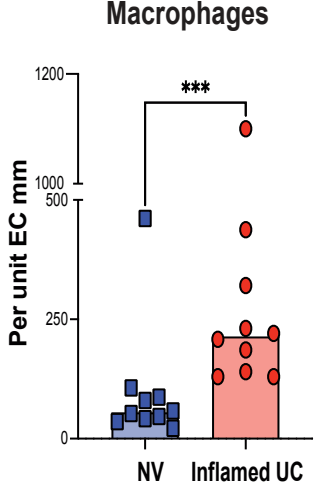
**b**



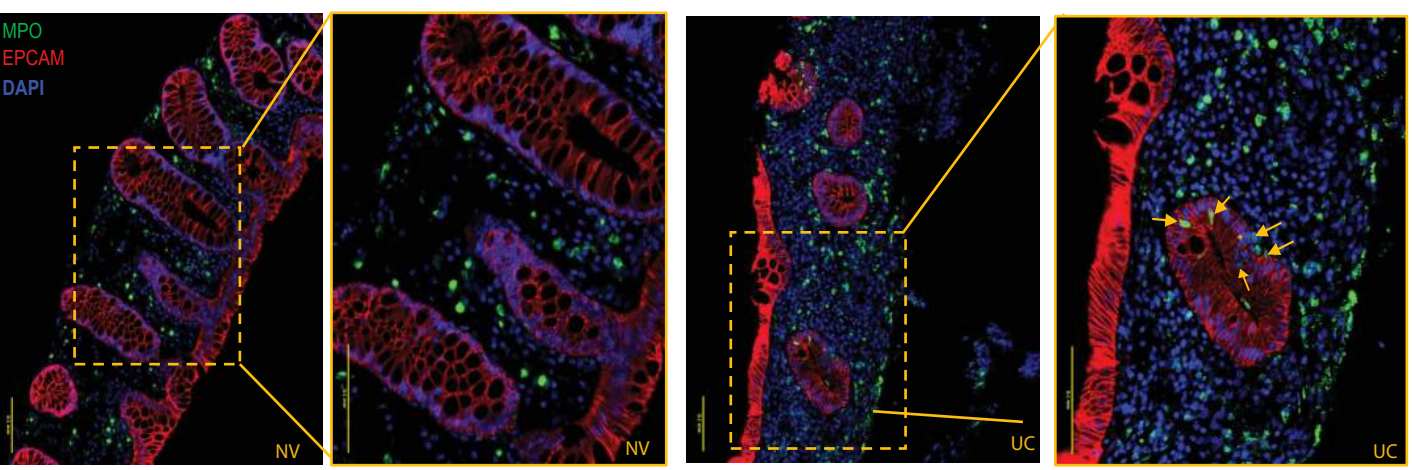
**c**



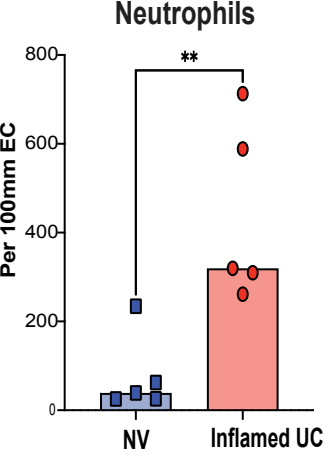
**d**

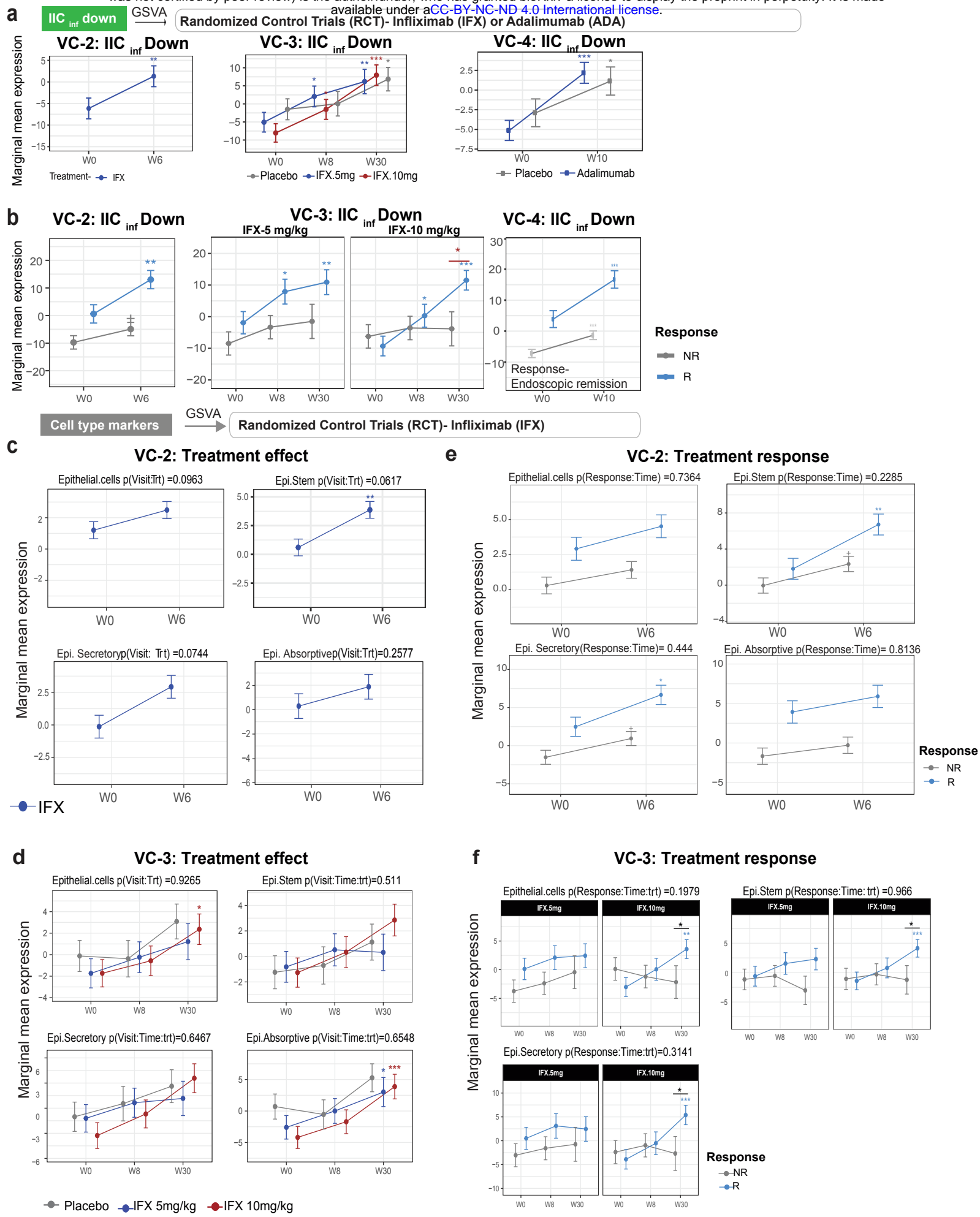


**e**



**f**





**Extended Data Fig. 9: Effect of anti-TNF therapy on cell types in VC-2, VC-3 and VC-4 cohorts in patients stratified by treatment response.**

