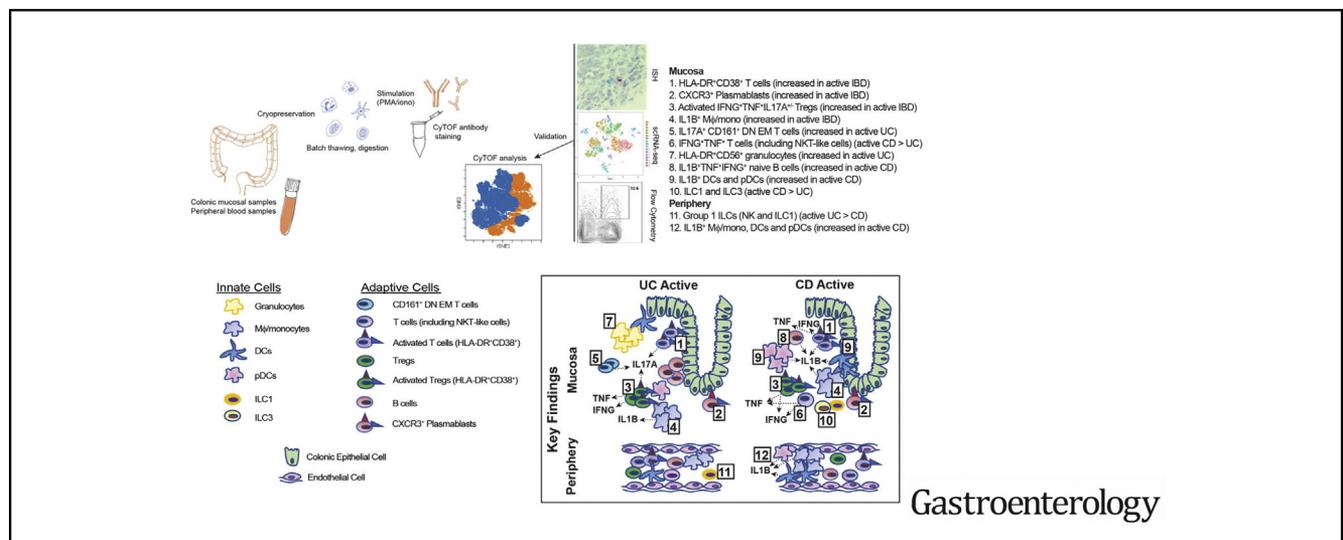


Single-Cell Analyses of Colon and Blood Reveal Distinct Immune Cell Signatures of Ulcerative Colitis and Crohn's Disease



Vanessa Mitsialis,^{1,2} Sarah Wall,² Peng Liu,³ Jose Ordovas-Montanes,^{2,4,5} Tamar Parnet,² Marko Vukovic,^{4,5,6,7,8} Dennis Spencer,² Michael Field,² Collin McCourt,⁴ Jessica Toothaker,⁵ Athos Bousvaros,² Boston Children's Hospital Inflammatory Bowel Disease Center, Brigham and Women's Hospital Crohn's and Colitis Center, Alex K. Shalek,^{4,5,6,7,8} Leslie Kean,⁹ Bruce Horwitz,² Jeffrey Goldsmith,¹⁰ George Tseng,³ Scott B. Snapper,^{1,2,§} and Liza Konnikova^{11,12,13,14,15,§}

¹Division of Gastroenterology, Brigham and Women's Hospital, Boston, Massachusetts; ²Division of Gastroenterology, Hepatology and Nutrition, Boston Children's Hospital, Boston, Massachusetts; ³Department of Biostatistics, University of Pittsburgh, Pittsburgh, Pennsylvania; ⁴Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts; ⁵Harvard Stem Cell Institute, Cambridge, Massachusetts; ⁶Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, Massachusetts; ⁷Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts; ⁸Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts; ⁹Division of Hematology Oncology, Boston Children's Hospital, Boston, Massachusetts; ¹⁰Department of Pathology, Boston Children's Hospital, Boston, Massachusetts; ¹¹Division of Newborn Medicine, Boston Children's Hospital, Boston, Massachusetts; ¹²Department of Pediatrics, University of Pittsburgh Medical Center Children's Hospital, Pittsburgh, Pennsylvania; ¹³Department of Immunology, University of Pittsburgh, Pittsburgh, Pennsylvania; ¹⁴Department of Developmental Biology, University of Pittsburgh, Pittsburgh, Pennsylvania; and ¹⁵Department of Pediatrics, Yale University School of Medicine, New Haven, Connecticut



See Covering the Cover synopsis on page 408.

BACKGROUND & AIMS: Studies are needed to determine the mechanisms of mucosal dysregulation in patients with inflammatory bowel diseases (IBDs) and differences in inflammatory responses of patients with ulcerative colitis (UC) vs Crohn's disease (CD). We used mass cytometry (CyTOF) to characterize and compare immune cell populations in the mucosa and blood from patients with IBD and without IBD (controls) at single-cell resolution. **METHODS:** We performed CyTOF analysis of colonic mucosa samples (n = 87) and peripheral blood mononuclear cells (n = 85) from patients with active or

inactive UC or CD and controls. We also performed single-cell RNA sequencing, flow cytometry, and RNA in situ hybridization analyses to validate key findings. We used random forest modeling to identify differences in signatures across subject groups. **RESULTS:** Compared with controls, colonic mucosa samples from patients with IBD had increased abundances of HLA-DR+CD38+ T cells, including T-regulatory cells that produce inflammatory cytokines; CXCR3+ plasmablasts; and IL1B+ macrophages and monocytes. Colonic mucosa samples from patients with UC were characterized by expansion of IL17A+ CD161+ effector memory T cells and IL17A+ T-regulatory cells; expansion of HLA-DR+CD56+ granulocytes; and reductions in type 3 innate lymphoid cells. Mucosal samples

from patients with active CD were characterized by IL1B+HLA-DR+CD38+ T cells, IL1B+TNF+IFNG⁺ naïve B cells, IL1B+ dendritic cells (DCs), and IL1B+ plasmacytoid DCs. Peripheral blood mononuclear cells from patients with active CD differed from those of active UC in that the peripheral blood mononuclear cells from patients with CD had increased IL1B+ T-regulatory cells, IL1B+ DCs and IL1B+ plasmacytoid DCs, IL1B+ monocytes, and fewer group 1 innate lymphoid cells. Random forest modeling differentiated active UC from active CD in colonic mucosa and blood samples; top discriminating features included many of the cellular populations identified above. **CONCLUSIONS:** We used single-cell technologies to identify immune cell populations specific to mucosa and blood samples from patients with active or inactive CD and UC and controls. This information might be used to develop therapies that target specific cell populations in patients with different types of IBD.

Keywords: Treg; MAIT Cells; ILC3; scRNA-Seq.

Inflammatory bowel disease (IBD) leads to chronic intestinal inflammation associated with significant morbidity. In the United States alone more than 3 million people have been diagnosed with IBD.¹ IBD pathogenesis is multifactorial, including contributions from genetic predisposition, breakdown of mucosal immune homeostasis, microbial dysbiosis, and environmental factors.² Subtypes of IBD include ulcerative colitis (UC), which contiguously affects the colon, and Crohn's disease (CD), which can present anywhere in the gastrointestinal tract. Macroscopic patterns of inflammation can often differentiate UC from CD, although the diagnosis can remain unclear. This becomes salient during critical management decisions, such as colectomy, which can be curative for patients with UC but not CD.

Although a number of therapies are successful in inducing and maintaining remission, including biologic agents targeting cytokine pathways and leukocyte trafficking,³ many patients are refractory to, or lose response to, available treatments. Up to one-third of patients will have disease refractory to anti-tumor necrosis factor (TNF) antibodies, and a significant proportion will ultimately require intestinal resections or colectomy.^{4,5} There is therefore an urgent need to expand the arsenal of therapies for IBD. Improved understanding of inflammatory landscapes in circulation and tissue may shed light on new therapeutic targets that can be tailored to disease type (eg, CD vs UC), subtype (eg, fistulizing vs stenotic disease), and even individual patients.

A number of studies have investigated IBD-associated patterns of inflammation in human tissue and in vitro. Before the development of multidimensional single-cell analytics using mass cytometry (CyTOF) or single-cell RNA sequencing (scRNA-seq), most analyses had been limited to the study of discrete cellular populations and signaling pathways. IBD tissue hallmarks, such as pro-inflammatory cytokine TNF generation in vitro,⁶ have been harnessed in the development of cytokine-specific biologic therapies (eg,

WHAT YOU NEED TO KNOW

BACKGROUND AND CONTEXT

Mucosal and peripheral immunologic signatures of IBD, including those that differentiate UC from CD, are not well-understood or robustly characterized.

NEW FINDINGS

Application of CyTOF to blood and colonic tissue of IBD patients and controls revealed disease-specific immune signatures at single-cell resolution. Novel findings include, but are not limited to, expansion of HLA-DR+CD38+ T cells, CXCR3+ plasmablasts, and IL1B+ Mf/mono in IBD tissue; expansion of IL17A++ CD161+ double-negative T cells in UC tissue, and expansion of IL1B+ DCs/pDCs/Mf/mono in both CD tissue and periphery.

LIMITATIONS

This study is observational with one timepoint per subject and includes patients on a variety of medical treatments with heterogeneous clinical phenotypes. Conclusions cannot be made regarding whether findings contribute to disease or represent biomarkers. Further mechanistic studies are needed to build upon these findings.

IMPACT

Our results demonstrate the strength of single-cell technology in profiling mucosal and peripheral immune signatures of disease. Some findings could be explored for targeted therapeutics and in the future possibly harnessed for personalized approaches to IBD therapy.

anti-TNF), but deeper understanding of cellular specificity in human tissue is lacking. Application of single-cell analytics to IBD is broadening these avenues of investigation, enabling the capture of a much more comprehensive cellular landscape.

To this end, we have applied CyTOF for multidimensional immunophenotyping of colonic mucosa (n = 87) and peripheral blood (n = 85) of UC, CD, and non-IBD subjects at single-cell resolution. We supplement our findings with scRNA-seq analysis of a subset of the same patients and validation cohorts using flow cytometry and RNA in situ hybridization. Further, we utilize unbiased clustering to identify cellular populations as well as predictive modeling to accurately differentiate UC and CD in the mucosa and periphery.

§ Authors share co-senior authorship.

Abbreviations used in this paper: a, active; AUC, area under the curve; CD, Crohn's disease; CyTOF, mass cytometry; DC, dendritic cell; EM, effector memory; i, inactive; IBD, inflammatory bowel disease; ILC, innate lymphoid cell; MAIT, mucosal-associated invariant T cell; pDC, plasmacytoid dendritic cell; RF, random forest; scRNA-seq, single-cell RNA sequencing; TCR, T-cell receptor; TNF, tumor necrosis factor; Treg, regulatory T cell; t-SNE, t-stochastic neighbor embedding; UC, ulcerative colitis.

 Most current article

© 2020 by the AGA Institute
0016-5085/\$36.00

<https://doi.org/10.1053/j.gastro.2020.04.074>

Materials and Methods

Human Sample Collection

Peripheral blood and/or colonic tissue from pediatric (aged older than 6 years) and adult subjects were collected after obtaining written informed consent. Samples were collected periprocedure (colectomy/colonoscopy) or during a clinical laboratory draw at Boston Children's Hospital or Brigham and Women's Hospital under Institutional Review Board protocol IRB-P00000529.

Determination of Inflammation Status and Disease Phenotype

The inflammatory status of IBD tissue was based on blinded histopathologic scoring by an experienced pathologist using the Nancy Index⁷ (Supplementary Figure 1B). An index of 0–1 was considered inactive disease and 2–4 was considered active. Nancy Index assignments and endoscopic assessments of inflammation were correlated (Supplementary Figure 1C). See Supplementary Figure 1C for further methodology for peripheral samples. Chart review was performed to obtain age, sex, treatment status, C-reactive protein values, and disease phenotype per the Montreal Classification System (Supplementary Table 1).

Peripheral Blood Mononuclear Cell and Tissue Isolation/Cryopreservation

To isolate peripheral blood mononuclear cells, whole blood collected in K2 EDTA tubes (BD Vacutainer; Becton Dickinson, Franklin Lakes, NJ) was processed using density-gradient centrifugation with Lympholyte-H Cell Separation Media (Cedarlane, Burlington, ON, Canada). Biopsies and resected surgical specimens were collected in T-cell media (500 mL RPMI 1640 medium; Thermo Fisher, Waltham, MA), 50 mL fetal bovine serum plus 5 mL of penicillin/streptomycin (Thermo Fisher), non-essential amino acids, sodium pyruvate, GlutaMAX (Thermo Fisher), and HEPES (Gibco, Gaithersburg, MD). Peripheral blood mononuclear cell suspensions (1 million cells per 100 μ L) and intestinal samples were cryopreserved in freezing media (10% dimethyl sulfoxide; Sigma, St Louis, MO) and 90% fetal bovine serum (Gibco).⁸ Samples were transferred to liquid nitrogen for long-term storage.

Mucosal Cell Isolation From Cryopreserved Samples

For CyTOF, cryopreserved intestinal samples were thawed and digested overnight in digestion media (10 μ g/mL collagenase I (Sigma) + 1 μ g/mL DNaseI (10,000 IU/mL) (Sigma) in 20 mL T-cell media/sample) at 37°C at 200 rpm. For flow cytometry, samples were digested in 300 U/mL collagenase VIII (Sigma) + 50 μ g/mL DNaseI (Sigma) in 20 mL T-cell media/sample \times 45 minutes. Digested solutions were strained and pelleted to isolate cells.

Mass Cytometry Staining and Acquisition

Single-cell suspensions were stimulated with phorbol myristate acetate (1:2000 of 100 μ g/mL stock; Sigma), ionomycin (1:1000 of 1 mg/mL stock; Sigma), and GolgiStop (1:1500; BD Biosciences, San Jose, CA) \times 4 hours. Cells were stained with Rh103 intercalator (500 μ M; Fluidigm, South San Francisco, CA) to label nonviable cells. Samples were stained with CyTOF antibody panels (Supplementary Figure 1A, Supplementary

Table 2). After staining, samples were labeled with Ir191/193 intercalator (1:1000 of 125 μ M stock; Fluidigm), then acquired on a Helios2 Mass Cytometer (Fluidigm). EQ 4 Elemental beads (Fluidigm) were used for normalization. Specifics of total cell numbers and viability of mucosal cells acquired in Supplementary Table 3. Further details regarding CyTOF staining protocol can be found in the Supplementary Methods.

Mass Cytometry Data Analysis

Normalized Flow Cytometry Standard (FCS) files were uploaded into Cytobank⁹ to define CD45⁺ viable single-cell events, and subsequently gated as demonstrated in the Results (Figures 1B, 2D, 3A, 4A, and 5A). Gated subsets were exported, uploaded into FlowJo software (FlowJo, Ashland, OR), and re-exported to ensure channel uniformity. Through *cytof-kit*,¹⁰ clustering using FlowSOM¹¹ and dimensionality reduction using t-stochastic neighbor embedding (t-SNE)¹² was performed using only markers common to all samples (Supplementary Table 4). Hierarchical clustering dendrograms were created with Euclidean agglomeration in R. Heatmaps were created using heatmap.2 in R *gplots*.¹³

All statistical computations were performed in R, including Kruskal-Wallis (for multiple group comparisons) and Wilcoxon *t* tests (for 2-group comparisons). Box-and-whisker plots were generated using *ggplot2*.¹⁴

Flow Cytometry

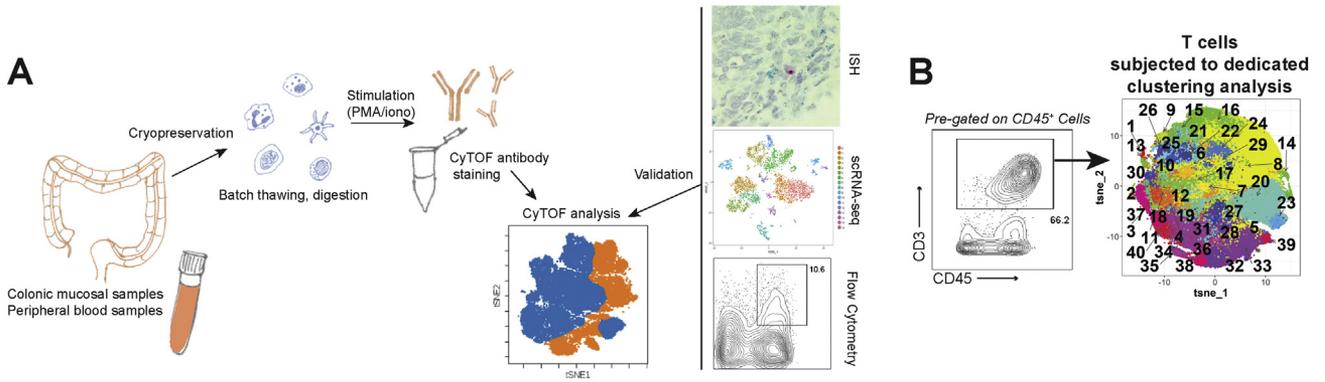
Colonic mucosal samples from a separate cohort (demographic data in Supplementary Table 5) were digested and stimulated as above. Cells were stained with fluorophore-conjugated antibodies in flow cytometry staining buffer (phosphate-buffered saline plus 2% fetal bovine serum) (antibodies listed in Supplementary Table 2). Samples were acquired on a BD LSRFortessa. Further compensation and analysis were performed in FlowJo. Additional details regarding flow cytometry staining protocol can be found in the Supplementary Methods.

RNA In Situ Hybridization

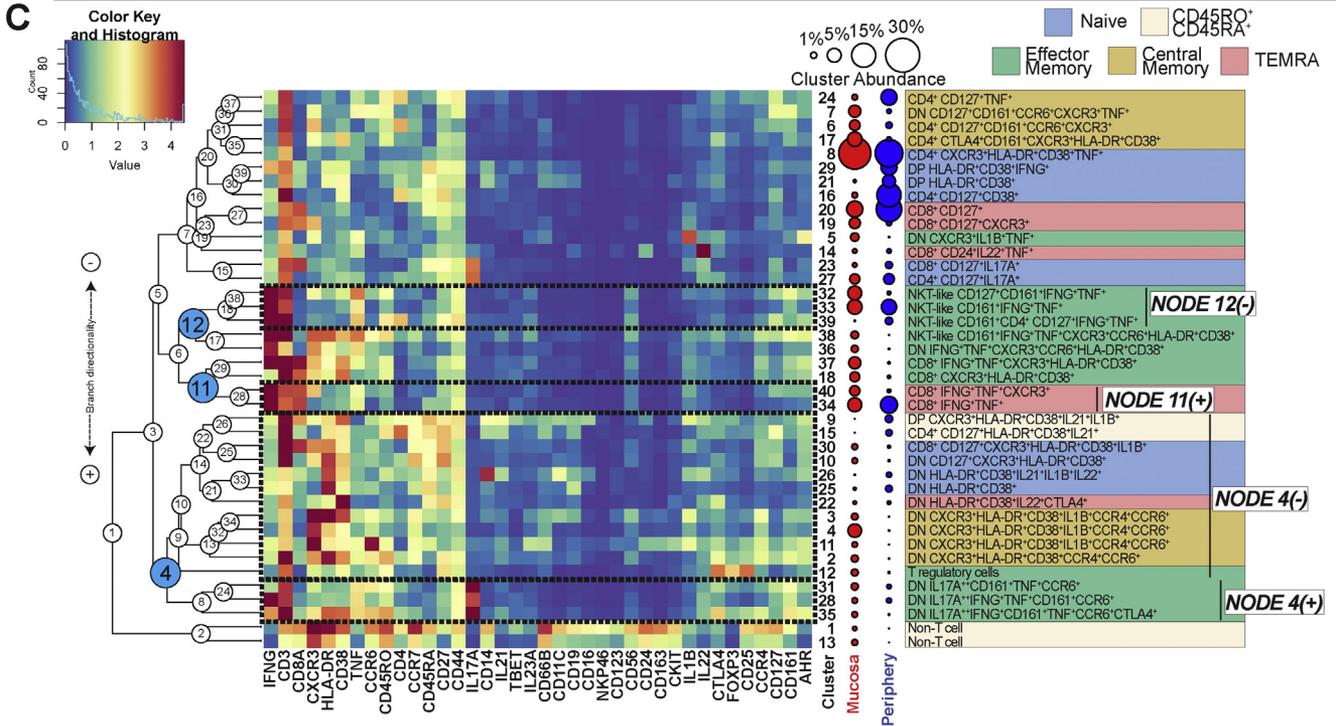
Formalin-fixed tissue samples were paraffin-embedded and sectioned by the University of Pittsburgh Biospecimen Core. Tissue sections were deparaffinized using xylene and 100% ethanol. RNAscope probes (ACD Bio, Newark, CA) were used to stain targets, including *IL1B* (C1) and *CD14* (C2). RNAscope Duplex Detection reagents (ACD Bio) were used according to manufacturer's instructions to amplify target RNA signal and bind probes to chromophores. Sections were stained with hematoxylin for detection of cell bodies. An Echo Revolve microscope (Echo, San Diego, CA) was used to image stained sections at 40 \times . Cells positive for one or both probes were counted on 40 \times images using ImageJ software, version 1.52 (National Institutes of Health, Bethesda, MD) in a blinded fashion. Available demographic data of this cohort are provided in Supplementary Table 5.

Single-Cell RNA-Sequencing: Tissue Preparation, Protocols, and Analysis

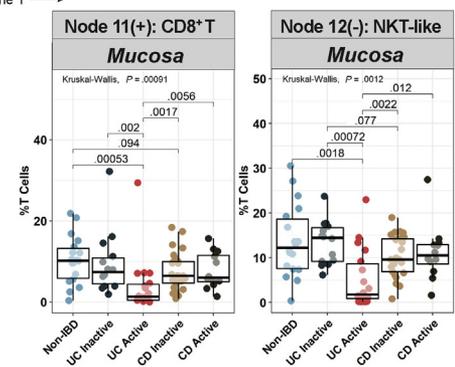
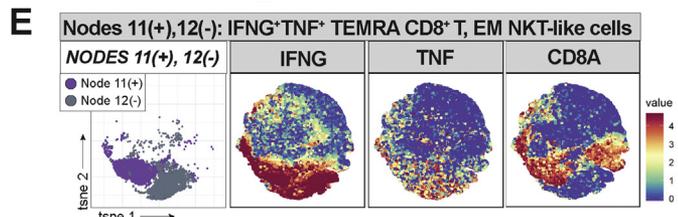
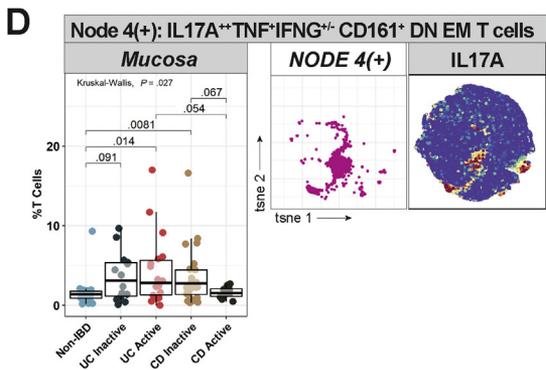
Please refer to the Supplementary Methods for full protocol. In brief, processing of single-cell biopsies was based on a



Automated CyTOF Analysis: T Cells



BASIC AND TRANSLATIONAL AT



protocol adapted from Smillie et al.¹⁵ Cryopreserved mucosal samples (2–3 biopsies/sample or equivalently sized pieces of resected surgical tissue, as described in Konnikova et al.⁸) were thawed and separation of epithelial layer from lamina propria was performed. The 2 cellular fractions were enzymatically digested and run on separate Seq-Well arrays as described in Ordovas-Montanes et al.,¹⁶ with improved protocol for library preparation based on second-strand synthesis for complementary DNA (S-3) from Hughes et al.¹⁷ Two arrays were sequenced per sequencing run with an Illumina 75 Cycle NextSeq500/550v2.5 kit on an Illumina NextSeq (Illumina, San Diego, CA). Read alignment was performed as done previously.^{16,17} Quality-filtered base calls were converted to demultiplexed FASTQ files and aligned to the Hg19 genome using Cell Ranger on the Galaxy portal maintained by the Broad Institute.

Merging of data sets, clustering, and differential gene expression analyses were performed using Seurat, version 3.1¹⁸ (<https://satijalab.org/seurat/>). Lists of differentially expressed genes defining the cellular clusters for all cells and for T cells (from Figure 2A and B) can be found in Supplementary Tables 6 and 7.

Random Forest Predictive Modeling

Random forest (RF) classifiers¹⁹ were constructed to differentiate between 2 subject groups (pairwise comparisons depicted in Figure 6A and B and Supplementary Figure 6). RF modeling was performed using the leave-one-out cross-validation strategy, removing 1 sample as a testing set and using the rest as training set. This strategy was repeated until each sample was left out at least once. The input features were the isometric log ratios²⁰ of branch abundance of each node in the hierarchical clustering dendrograms from T, B, and innate cell analyses. Prediction accuracy (with a cutoff point equally weighing false-positive and false-negative predictions) and area under the curve (AUC) were calculated to evaluate model performance. To choose the proper number of features to include in classifiers, each time in leave-one-out cross-validation, we selected top n ($n = 5, 10, 20, 30, 50, \text{ or } 100$) differential features by Wilcoxon test based on training data to construct classifiers; the n that yielded the highest prediction accuracy was selected as the number of features for the full models.

Data Availability Statement

Mass cytometry. All CyTOF files (.fcs files) can be accessed through a publicly available experiment on Cytobank

platform (<https://premium.cytobank.org>) entitled “Raw mass cytometry data from: Mitsialis et. al. Single-Cell Analyses of Colon and Blood Reveal Distinct Immune Cell Signatures of Ulcerative Colitis and Crohn’s Disease. Gastroenterology. 2020.” File names are deidentified and correlate with a metadata spreadsheet (Supplementary Table 8).

Single-cell RNA-sequencing. The cells by genes matrices generated from colonic mucosa from the 5 UC subjects included in this study can be accessed on National Institutes of Health Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>) with the accession number GSE150115. Associated clinical metadata for these subjects can be found in Supplementary Table 8.

Results

CytoF was performed on colonic mucosa from 18 non-IBD, 32 UC (14 inactive [i], and 18 active [a]) and 37 CD ($i = 25, a = 12$) subjects and peripheral blood from 24 non-IBD, 27 UC ($i = 5, a = 22$), and 34 CD ($i = 12, a = 22$) subjects (Supplementary Table 1). Of these, 24 subjects had matched blood and biopsy samples. CyTOF markers included surface antigens and cytokines (Supplementary Figure 1A, Supplementary Table 2), with the intent to broadly immunophenotype innate and adaptive immune cells. Mucosal analyses were performed on cryopreserved samples with robust viability, as we have described⁸ (Figure 1A, Supplementary Table 3).

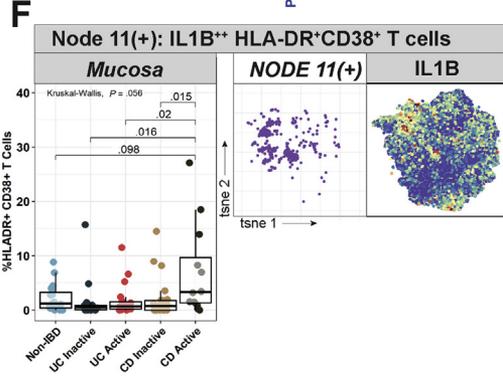
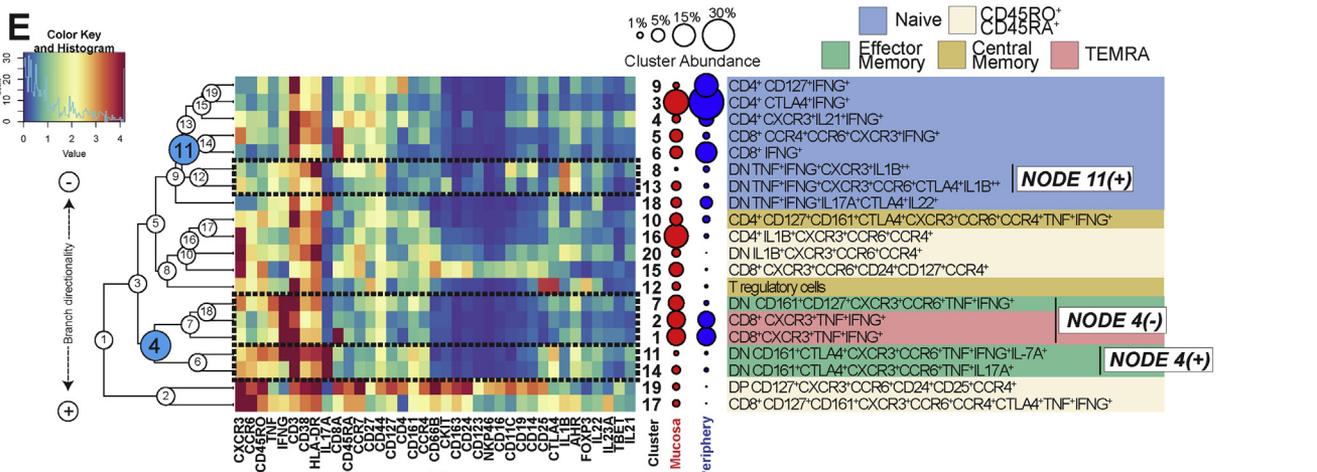
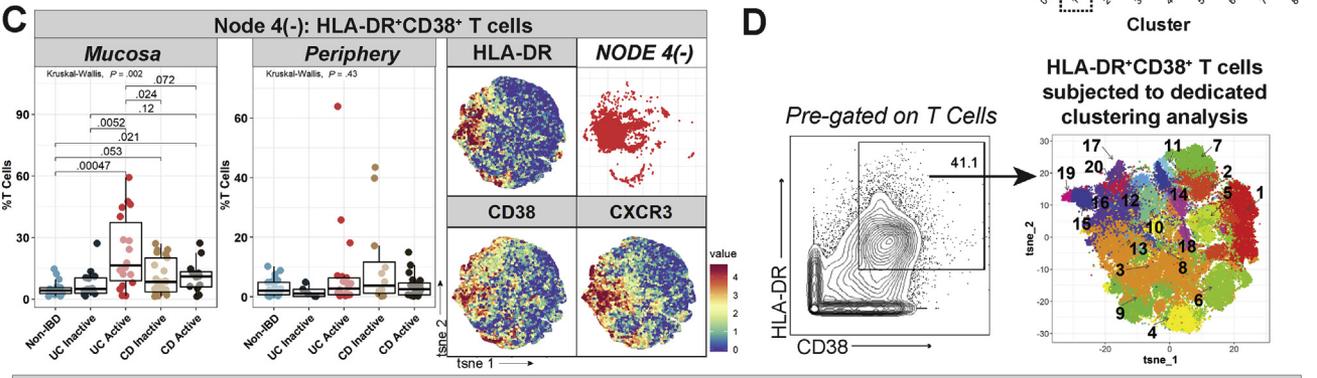
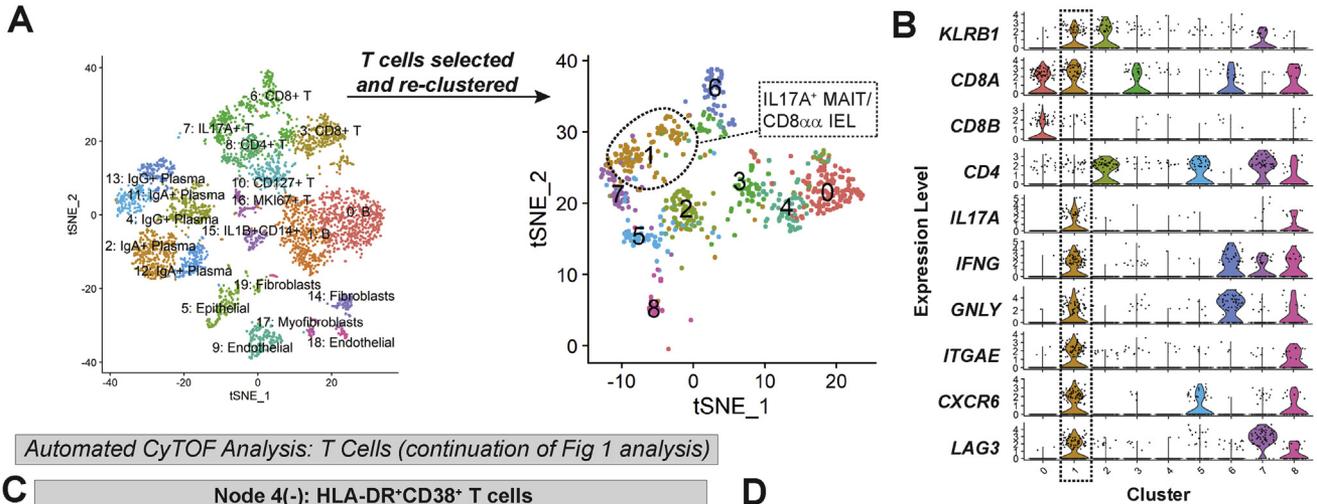
Immune cell subsets of interest were exported for dedicated automated analyses with unbiased clustering. Phenotypically similar clusters were organized into branches of hierarchical clustering dendrograms. On relevant figures, dendrogram nodes are labeled in circles and node branches are labeled as *Node X(-)* or *Node X(+)*, depending on the direction of branch splitting. The abundances of cellular populations (individual clusters and/or node branches) were compared across subject groups. Finally, predictive modeling using RF was applied to identify features that distinguish subject groups in the mucosa and periphery.

Active Ulcerative Colitis Mucosa Is Characterized by an Increased B:T Cell Ratio

Before automated analyses, we assessed the relative proportions of manually gated T cells ($CD3^+CD45^+$),

Figure 1. T-cell analysis demonstrates UC- and CD-specific cytokine signatures. (A) Schematic of sample analysis methodology. (B) Gating strategy to define T cells ($CD3^+CD45^+$) that were exported and subjected to dimensionality reduction (t-SNE) and clustering (FlowSOM) as depicted in labeled t-SNE cluster plot (right). (C) CyTOF marker heatmap (rows are clusters, columns are markers, tile colors represent mean metal intensity per color key and histogram legend scaled across all tiles). On the right are manually labeled cluster identities and a schematic of cluster abundance in mucosa vs periphery. Hierarchical clustering dendrogram on left of heatmap, with nodes numbered and selected nodes highlighted. Selected node branches outlined on heatmap and labeled on right. T-cell memory subsets were assigned and color-coded (naïve: $CD45RA^+CD45RO^-CCR7^+CD27^+$; central memory: $CD45RA^-CD45RO^+CCR7^+CD27^+$; EM: $CD45RA^-CD45RO^+CCR7^-CD27^{+/-}$; terminally differentiated EM cells re-expressing CD45RA [TEMRA]: $CD45RA^+CD45RO^{+/-}CCR7^-CD27^{+/-}$; and a subtype external to these established categories, $CD45RO^+CD45RA^+$). (D) Node 4(+) labeled on t-SNE plot, with IL17A heatmap for reference. Box-and-whisker plot of node 4(+) abundance across subject groups on the left. (E) Nodes 11(+), 12(-) t-SNE plot, abundance plots, and IFNG/TNF/CD8A heatmaps.

scRNA-seq T Cell Analysis



BASIC AND TRANSLATIONAL AT

B cells (CD19⁺CD45⁺), and innate immune cells (CD3⁻CD19⁻CD45⁺) across all subject groups (Supplementary Figure 1D). UCa mucosa had proportionally fewer T cells and increased B cells compared to all other IBD groups with an enhanced B:T cell ratio (Supplementary Figure 1D). The T and B cell findings in UC compared to non-IBD mucosa were validated by flow analysis of an independent cohort (n = 10 non-IBD, n = 9 UCa, n = 10 CDa) (Supplementary Table 5 and Supplementary Figure 1E). No overall differences in the proportion of innate cells were observed (Supplementary Figure 1D).

T cells were clustered in a dedicated automated analysis (Figure 1B and C). T cell memory status (naïve, effector memory [EM], central memory, terminally differentiated EM cells re-expressing CD45RA, and a subtype that did not clearly fall into these categories, labeled CD45RO⁺CD45RA⁺) was assigned based on differential expression of CD45RA, CD45RO, CCR7, and CD27 (Figure 1C). A range of CD4⁺, CD8⁺, CD4⁺CD8⁺ (DP), and CD4⁻CD8⁻ (DN) T cell subtypes were identified.

IL17A⁺⁺ CD161⁺ Effector Memory T Cells Are Enriched While IFNG⁺ TNF⁺ Effector Memory T-Cell Subsets Are Diminished in Active Ulcerative Colitis Mucosa

We identified disease-specific compositional differences in T cells pertaining to cytokine-producing EM subsets. Specifically, DN IL17A⁺⁺ CD161⁺ EM T cell clusters (belonging to node 4[+]) were expanded in UCa mucosa (Figure 1D). Given cytokine co-expression (not only IL17A⁺⁺ but also TNF⁺ and IFNG⁺), as well as CD161 and CCR6 co-expression, these DN EM T cells may represent mucosal-associated invariant T cells (MAIT) (or related lymphocytes with innate characteristics),²¹ as our CyTOF panel was not equipped to capture MAIT-specific semi-invariant T-cell receptors (TCRs) for more definitive classification. While these clusters were expanded, IFNG⁺ TNF⁺ EM T-cell subsets were decreased compared to all other subject groups in UCa mucosa, including node 11(+) (cytotoxic CD8⁺ terminally differentiated EM cells re-expressing CD45RA) and node 12(-), which were CD161⁺CD56⁺ EM T cells that could represent natural killer T-like cells (Figure 1E).

A decrease in overall IFNG⁺ T cells in UCa compared to both non-IBD and CDa mucosa was confirmed with flow cytometry, as was an increased ratio of IL17A⁺ to IFNG⁺

T cells in UCa mucosa compared to non-IBD (Supplementary Figure 2A).

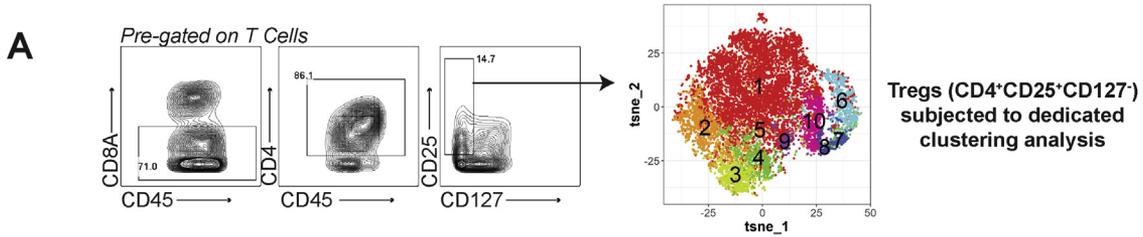
To gain greater molecular insights into these cell subsets, we performed scRNA-seq on matched cryopreserved mucosal samples from 5 UCa subjects in our CyTOF cohort. We recovered all expected epithelial, stromal, and immune cell subsets¹⁵ and focused further analyses on the T-cell clusters (Figure 2A). T cells were identified, subsetted, and reclustered independently (Figure 2A; Supplementary Methods). We identified a T-cell cluster co-expressing *IL17A*, *IFNG*, *CD8A* (but not *CD8B*),²² and *KLRB1* (CD161), potentially representing IL17A⁺ CD161⁺ cells of similar phenotype to those expanded in UCa mucosa (node 4[+]) identified with CyTOF (Figure 2A and B). As our sequencing data did not capture enrichment of specific TCR transcripts, such as *TRAV1-2* or the specific *TRAJ* transcripts that make up the semi-invariant MAIT TCR, these cells may represent MAIT cells or polyclonal CD8 α IELs. Specific sequencing of the TCR region (<https://www.nature.com/articles/s41590-019-0544-5>) will be necessary to identify whether this subset is composed of public or private TCRs. Nevertheless, our scRNA-seq analysis did highlight several MAIT-related markers^{23,24} in this cluster, which were not included in our CyTOF panels, including *ITGAE* (CD103), *CXCR6*, *LAG3*, *GPLY*, *CCR5*, and *RORA*, a transcription factor belonging to the same family as *RORGT*, associated with T-helper 17 responses (Supplementary Table 7).

IBD Mucosa Is Characterized by an Abundance of HLA-DR⁺CD38⁺ T-Cell Populations

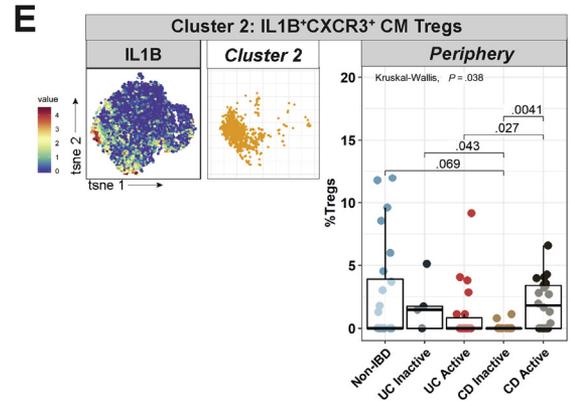
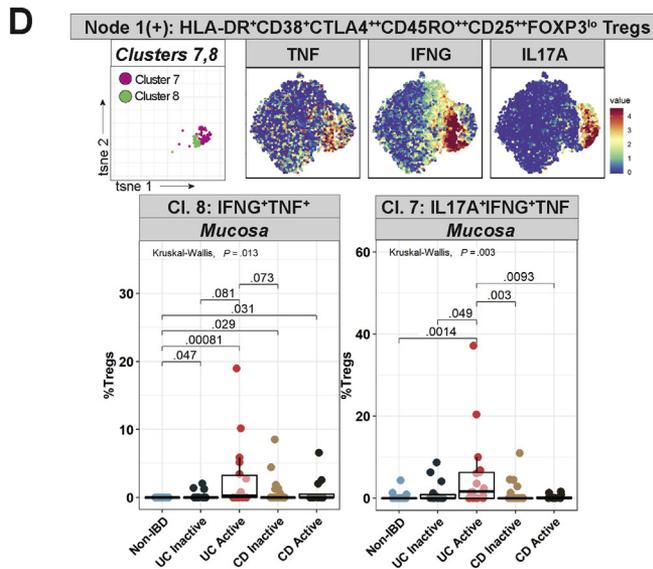
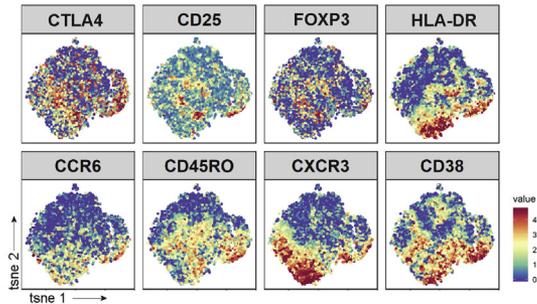
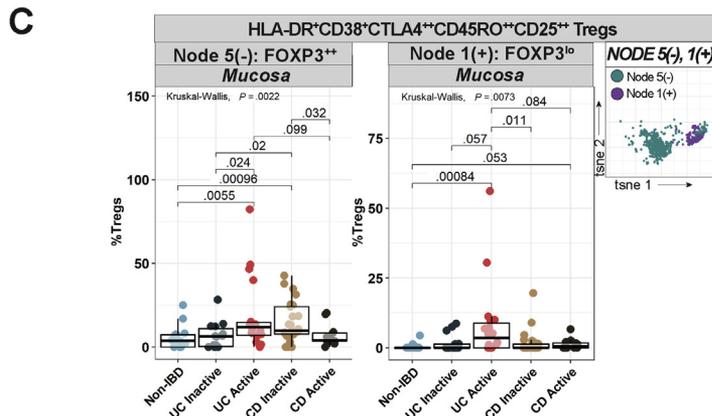
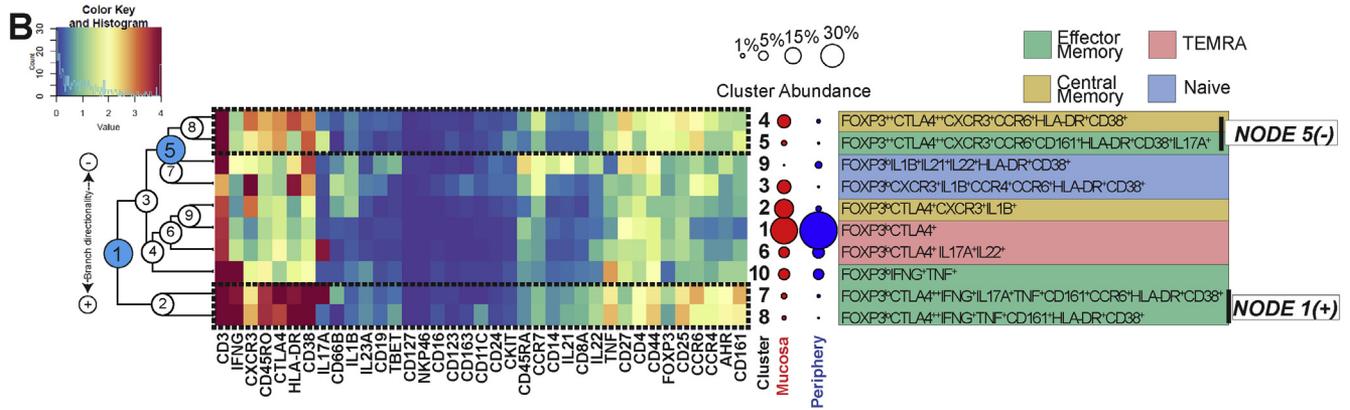
In our CyTOF analysis, T-cell clusters co-expressing HLA-DR and CD38 (node 4[-]) were expanded in IBD mucosa (Figure 2C) and this finding was validated by flow cytometry (Supplementary Figure 2B). HLA-DR⁺CD38⁺ T cells were characterized by expression of chemokine receptors, such as *CXCR3*, *CCR6*, and *CCR7*, as well as co-expression of the CD45R isoforms CD45RA and CD45RO (eg, clusters 9 and 15) (Figure 1C). In contrast to some earlier reports,^{25,26} HLA-DR⁺CD38⁺ T cells were not found to be expanded in peripheral IBD (Figure 2C).

To better phenotype HLA-DR⁺CD38⁺ T cells, we manually gated this population and performed a dedicated clustering analysis (Figure 2D and E). HLA-DR⁺CD38⁺ T cells

Figure 2. Mucosal scRNA-seq identifies *IL17A⁺* T cells in UC and CyTOF analysis demonstrates expansion of HLA-DR⁺CD38⁺ T-cell subsets in IBD. (A) scRNA-seq of UCa subjects from CyTOF cohort (n = 5 samples) represented by t-SNE plot colored by cluster (left; total cells analyzed = 3979) with cluster identity labeled according to differential gene expression (see Supplementary Table 6). Data were subsetted on T cells, which were then re-clustered (right; total cells analyzed = 718) with generation of 9 clusters labeled “0” through “8” (see Supplementary Table 7 and Supplementary Methods); cluster 1, corresponding to *IL17A⁺* MAIT cells or *IL17A⁺* CD8 α intraepithelial lymphocytes (IEL), is circled. (B) Violin plots of selected transcripts, with expression profile of cluster 1 outlined. *KLRB1*, CD161; *ITGAE*, CD103; *GPLY*, granulysin. All violin plots generated using standard Seurat, version 3.1 implementation; y-axis represents normalized and log-transformed expression data (log[scaled transcript counts + 1]). (C) Node 4(-) (from Figure 1C) abundance plot, t-SNE, and HLA-DR/CD38/CXCR3 heatmaps. (D) Gating strategy defining HLA-DR⁺CD38⁺ T cells (left) with labeled t-SNE cluster plot (right) from dedicated clustering analysis. (E) CyTOF marker heatmap, cluster dendrogram, cluster abundance, and identity for HLA-DR⁺CD38⁺ T-cell analysis with labeling strategy as per Figure 1C. (F) Node 11(+) (from Figure 2E) abundance plot, t-SNE, and IL1B heatmap.



Automated CyTOF Analysis: T Regulatory Cells (Tregs)



were found to represent a mix of T-cell subtypes (CD4, CD8, and DN T cells, regulatory T cells [Tregs]) and memory states (Figure 2E). Similar to our findings in T cells at large, HLA-DR⁺CD38⁺ T cells co-expressing IFNG⁺ TNF⁺ (node 4 [-] in Figure 2E) were diminished in UCa mucosa whereas IL17A⁺HLA-DR⁺CD38⁺ CD161⁺ DN EM T cells (node 4[+] in Figure 2E) were enhanced (Supplementary Figure 3A). IL1B⁺HLA-DR⁺CD38⁺ T cells (node 11[+] in Figure 2E) demonstrated a strong trend towards expansion in CDa mucosa (Figure 2F).

Memory Regulatory T Cells Are Increased in Abundance in Active Inflammatory Bowel Disease Mucosa and Express Pro-Inflammatory Cytokines

Tregs were manually gated (as CD4⁺CD25⁺CD127⁻ T cells) and subjected to dedicated automated analysis (Figure 3A and B). Memory Treg subsets were assigned according to the same algorithm detailed for other T cells. We identified disease-specific signatures in Tregs that mirrored findings in our preceding T-cell analyses, including HLA-DR and CD38 co-expression in IBD mucosa and IL17A expression in UCa mucosa.

Treg populations enhanced in IBD mucosa were HLA-DR⁺CD38⁺, especially in UCa (node 5[-], node 1[+] in Figure 3B and C), similar to our findings in preceding T-cell analyses. HLA-DR⁺CD38⁺ IBD-associated Tregs co-expressed various chemokine receptors, including CCR6 and CXCR3, and were uniformly CD25⁺⁺CTLA4⁺⁺CD45RO⁺⁺, suggesting an activated memory phenotype (Figure 3B and C). Some branches were characterized by FoxP3⁺⁺ expression, perhaps signifying high suppressive capacity (node 5[-]). Others, for example, clusters 7 and 8 (node 1[+]), were FOXP3^{lo} and expressed pro-inflammatory cytokines (IFNG⁺ TNF⁺IL17A^{+/}), possibly indicating non-suppressive or even non-Treg properties^{27,28} (Figure 3B and C).

Cluster 8, which co-expressed IFNG and TNF, was increased in both UCa and CDa, whereas cluster 7, which co-expressed IFNG, TNF, and IL17A, was specifically increased in UCa mucosa (Figure 3D). Only 1 Treg cluster differed in peripheral disease: cluster 2, IL1B⁺CXCR3⁺ central memory Tregs, enhanced in CDa compared to UCa and CDi (Figure 3E). Given the novelty of identifying IL1B expression in Tregs, we confirmed this finding using flow cytometry (Supplementary Figure 4A). Our data supports the possibility that Tregs in active IBD tissue may be altered, possibly with pathogenic potential, as has been described in IBD²⁹ and other inflammatory disorders.^{30,31}

Finally, as has been reported previously,^{32,33} manual gating of 2 independent cohorts with CyTOF or flow

cytometry revealed increased Tregs in IBD compared to non-IBD mucosa (Supplementary Figure 1D, Supplementary Figure 4B). In contrast to some prior reports,³² no difference was observed in peripheral abundance of Tregs in IBD (Supplementary Figure 1D).

IL1B⁺IFNG⁺ TNF⁺ Naïve B Cells Are Enriched in Active Crohn's Disease Mucosa

As noted, manual gating of CyTOF data demonstrated enrichment of B cells in UCa mucosa (Supplementary Figure 1D). To investigate differences within B-cell populations across subject groups, CD19⁺ B cells were clustered in a dedicated analysis (Figure 4A and B). Multiple B-cell subtypes, including CD27⁻ naïve, CD27⁺ memory, CD27⁻CD24⁺CD38⁺ transitional, and CD27⁺CD24⁻CD38⁺⁺ plasmablasts, were identified (Figure 4B).

IL1B⁺IFNG⁺TNF⁺ naïve B-cell clusters (node 8[+] in Figure 4B) were increased in CDa mucosa (Figure 4C). These clusters were CD44⁺⁺ (marker of activated B cells), CCR7⁺, AHR⁺, HLA-DR⁺, CD38⁺, and CD11C⁺, a marker expressed in B cells capable of antigen presentation associated with autoimmunity³⁴ (Figure 4B and C). Cluster 33 within this node branch was found to express CD14, which could reflect a non-B-cell macrophage/monocyte population or CD14-expressing B cells that have been reported previously.³⁵

CXCR3⁺ Plasmablasts Are Increased in Inflammatory Bowel Disease Mucosa

A population of CXCR3⁺ plasmablasts (cluster 18 in Figure 4B) was found to be expanded in active IBD mucosa, especially in UCa (Figure 4D). CXCR3⁺ plasmablast clusters (node 10[-] in Figure 4B) expressed FOXP3, which was further investigated through manual gating of CyTOF data (Supplementary Figure 5A), also supporting a FOXP3 signal in mucosal plasmablasts. Although FOXP3 expression in non-T-cell lineages has been controversial,³⁶ our findings may suggest a regulatory function for mucosal plasmablasts as has been suggested by others.³⁷

Active Ulcerative Colitis Mucosa Is Enriched With Unconventional Granulocytes and Active Crohn's Disease Mucosa With IL1B⁺ Dendritic Cells

CD3⁻CD19⁻CD45⁺ cells (innate immune cells) were clustered in a dedicated analysis (Figure 5A and B). Multiple subsets of innate immune cells were identified, including granulocytes (CD66B⁺), macrophages/monocytes (CD11C⁺CD14⁺), dendritic cells (DCs; HLA-DR⁺CD11C⁺CD14⁻), plasmacytoid DCs (pDCs; CD123⁺HLA-DR⁺CD11C⁺CD14⁻), and innate lymphoid cells (ILCs) groups 1 (CD56⁺TBET⁺CD161⁺IFNG⁺TNF⁺), 2

Figure 3. Treg analysis demonstrates expansion of pro-inflammatory memory Treg subsets in IBD. (A) Gating strategy defining Tregs cells (left) with labeled t-SNE cluster plot (right) from dedicated clustering analysis. (B) CyTOF marker heatmap, cluster dendrogram, cluster abundance, and identity for Treg analysis with labeling strategy as per Figure 1C. (C) Node 5(-) and node 1(+) (from Figure 3B) abundance plots, t-SNE, and selected Treg marker heatmaps for reference. (D) Clusters 7 and 8 (from node 1[+] in Figure 3B) abundance plots, t-SNE, and TNF/IFNG/IL17A marker heatmaps. (E) Cluster 2 (from Figure 3B) abundance plot, t-SNE, and IL1B marker heatmaps.

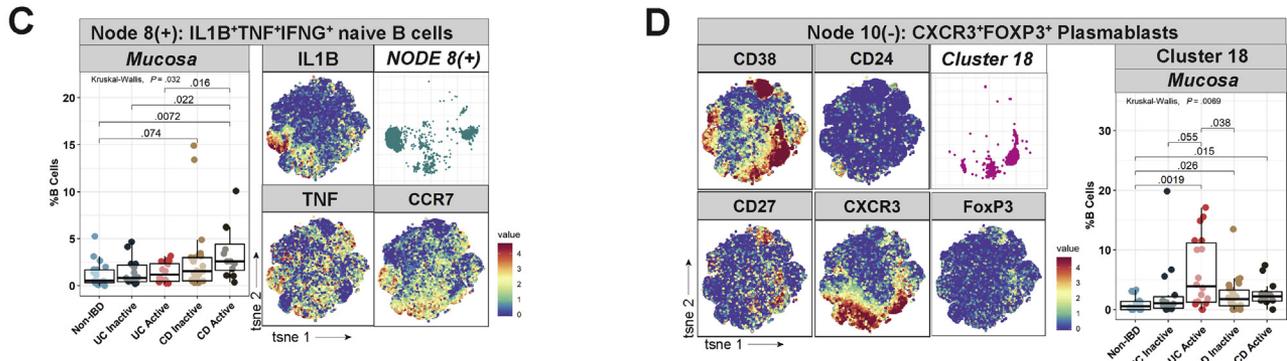
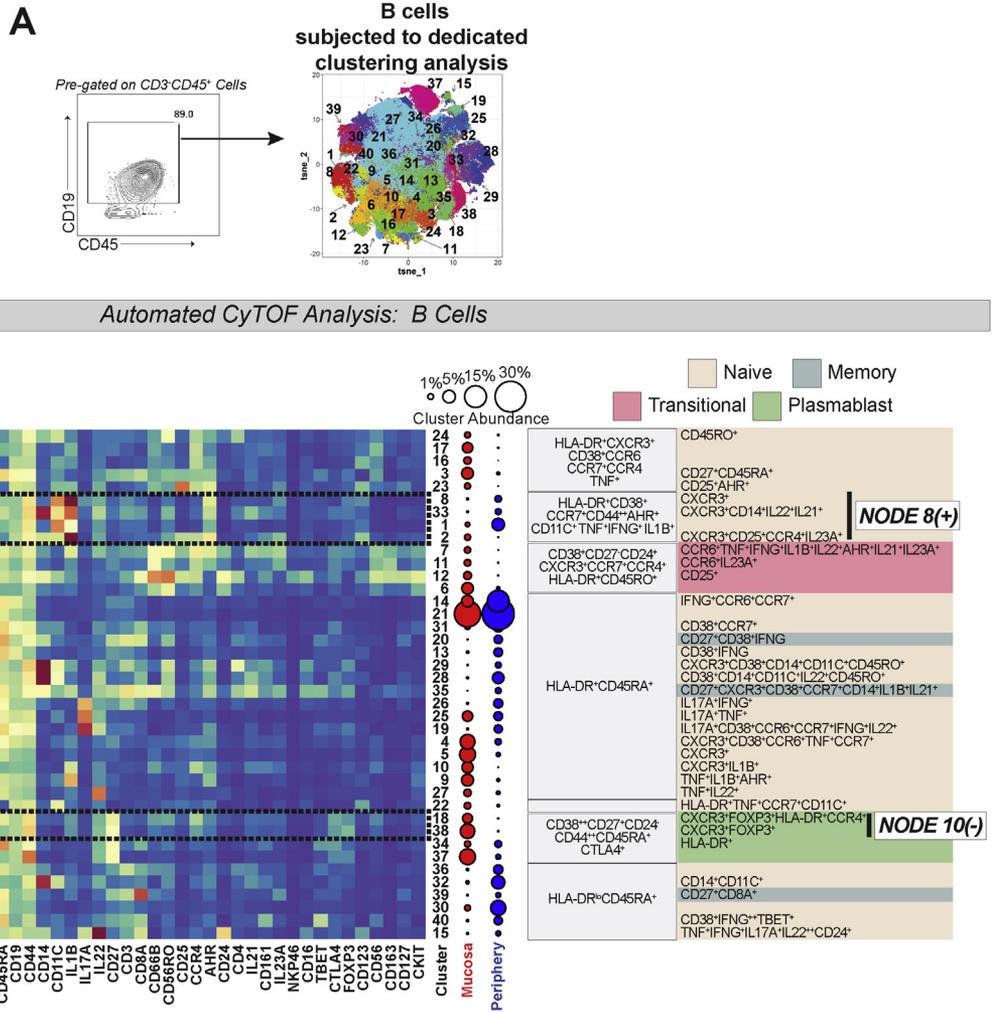
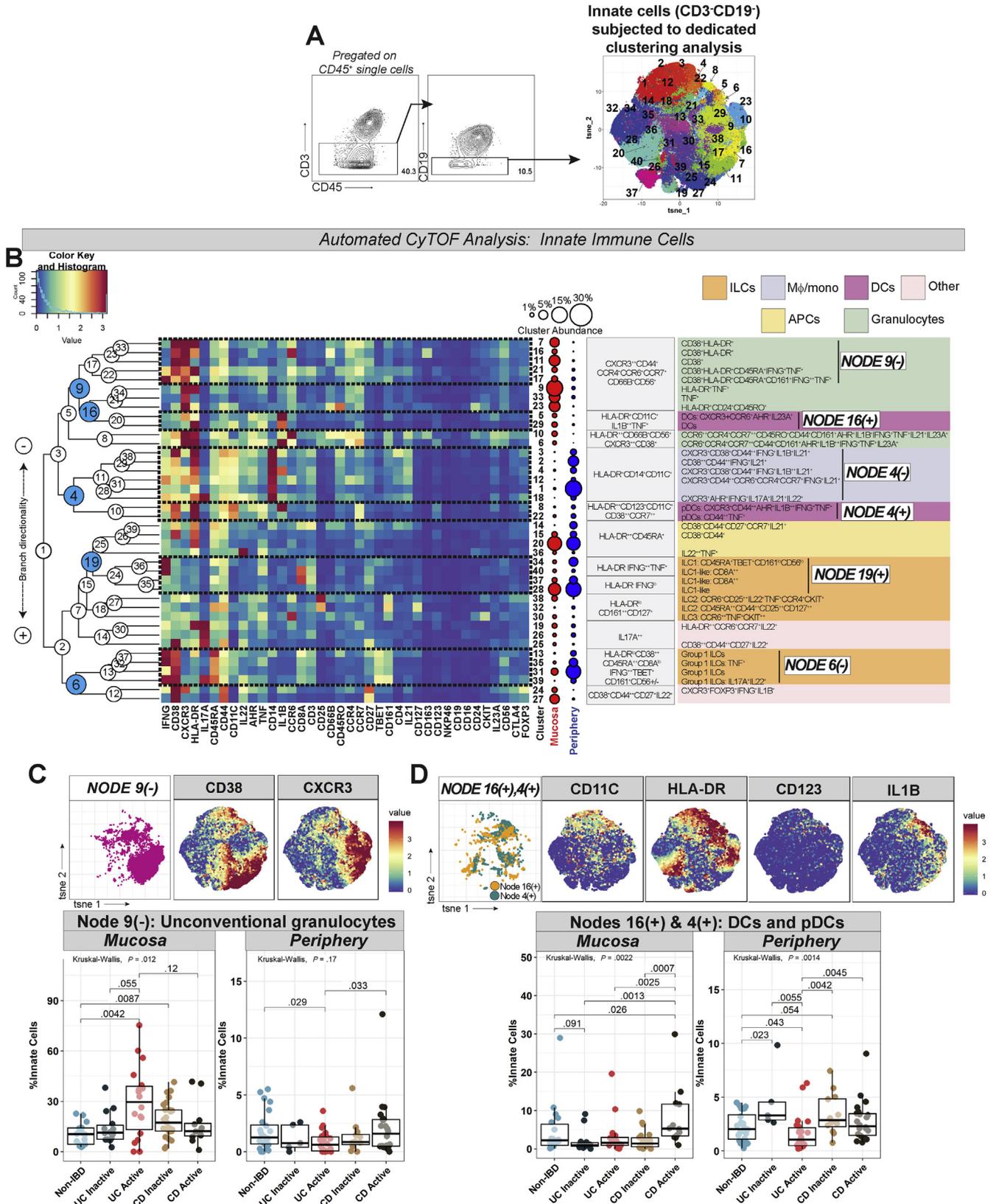


Figure 4. B-cell analysis highlights unique naïve B-cell and plasmablast populations in IBD. (A) Gating strategy defining B cells (left) with labeled t-SNE cluster plot (right) from dedicated clustering analysis. (B) CyTOF marker heatmap, cluster dendrogram, cluster , and identity (color-coded by subset) organized as per Figure 1C. (C) Node 8(+) (from Figure 4B) abundance plots, t-SNE, and IL1B/TNF/CCR7 marker heatmaps. (D) Cluster 18 (from node 10(-) in Figure 4B) abundance plots, t-SNE, and selected marker heatmaps.

(CD161⁺CD127⁺CD25⁺⁺CKIT^{+/-}) and 3 (CD161⁺CD127⁺CCR6⁺⁺CKIT⁺) (Figure 5B).

UCa mucosa was characterized by increased granulocytes (node 9[-] in Figure 5B) with a concomitant decrease (trend) in the periphery (Figure 5C). This

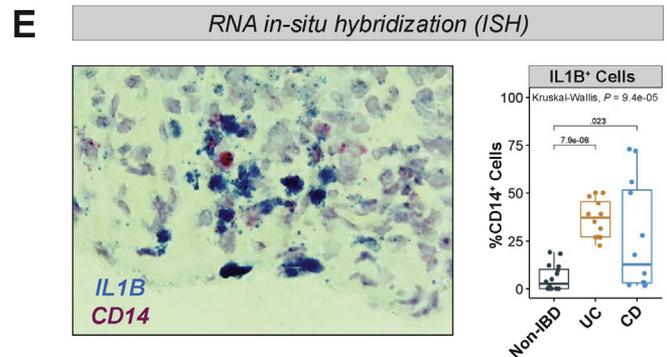
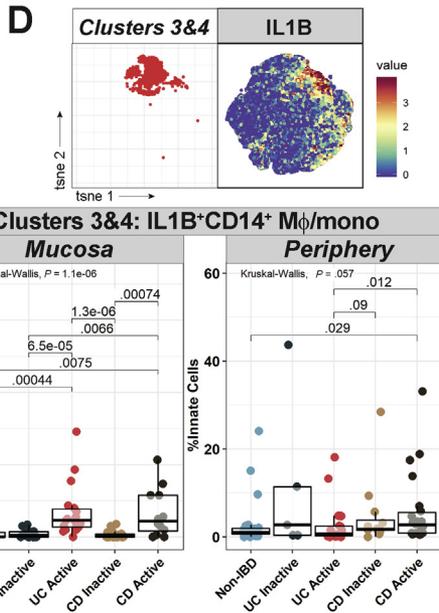
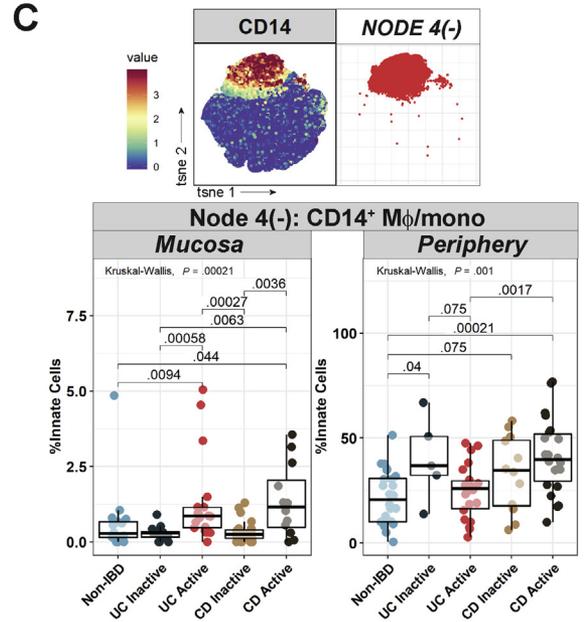
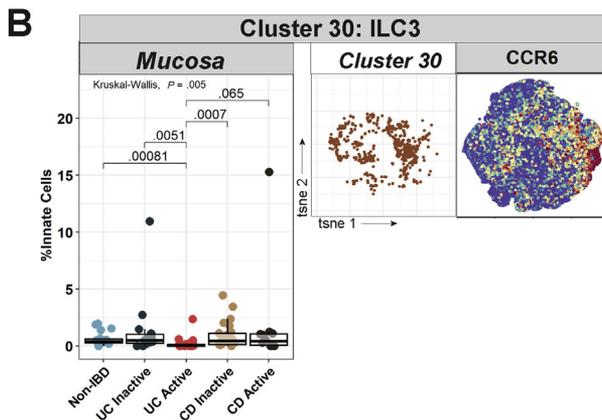
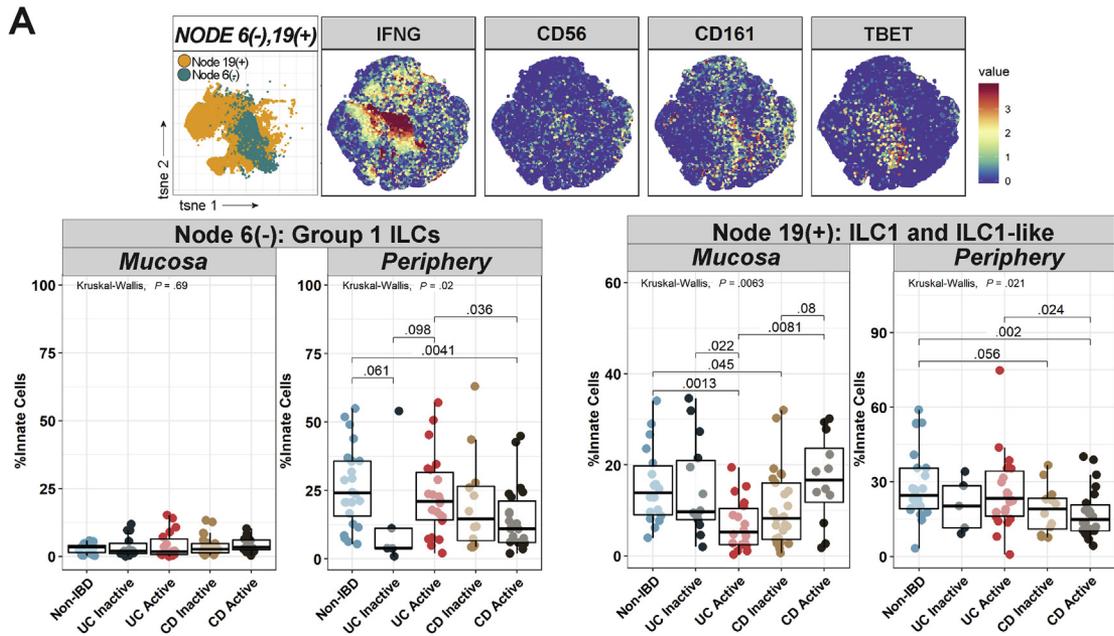
population expressed chemokine receptors (CXCR3, CCR6), as well as unconventional granulocyte markers, including HLA-DR, CD38, and CD56, which have been reported to be up-regulated on granulocytes in other human diseases^{38,39} (Figure 5B and C). CDa mucosa, however, was



BASIC AND TRANSLATIONAL AT

Figure 5. Innate immune cell analysis reveals granulocyte and DC differences in UC and CD. (A) Gating strategy defining innate immune cells (left) with labeled t-SNE cluster plot (right) from dedicated clustering analysis. (B) CyTOF marker heatmap, cluster dendrogram, cluster abundance and identity (color-coded by subset) organized as per Figure 1C. (C) Node 9(-) (from Figure 5B) abundance plots, t-SNE, and CD38/CXCR3 marker heatmaps. (D) Nodes 16(+) and 4(+) (from Figure 5B) abundance plots (combined), t-SNE, and selected marker heatmaps.

Automated CyTOF Analysis: Innate Immune Cells, continued



characterized by increased DCs and pDCs (nodes 16[+] and 4[+], respectively in [Figure 5B](#)), which were also increased in peripheral CDa>UCa ([Figure 5D](#)). The majority of these DC and pDC clusters, notably, highly expressed IL1B ([Figure 5B](#) and [D](#)).

Innate Lymphoid Cells Signatures Differentiate Crohn's Disease From Ulcerative Colitis in Mucosa and Periphery

Group 1 ILCs (likely including both ILC1 and natural killer cells) (node 6[-] in [Figure 5B](#)) were decreased in peripheral CDa, without subject group differences in the mucosa ([Figure 6A](#)). Node 19(+) (in [Figure 5B](#)), which included both ILC1 (cluster 34) as well as "ILC1-like" clusters (identified by lack of expression of lineage markers [eg, CD14, CD11c], but with expression of the ILC1-hallmark cytokine IFNG and by CD8A, which has been reported in ILCs⁴⁰) was also found to be decreased in peripheral CDa ([Figure 6A](#)). ILC1 and ILC1-like clusters conversely were increased in the mucosa in CDa>UCa ([Figure 6A](#)), while cluster 30 (in [Figure 5B](#)), representing ILC3s, was specifically reduced in UCa mucosa ([Figure 6B](#)).

IL1B⁺ Macrophages/Monocytes Are Increased in Active Inflammatory Bowel Disease Mucosa and in Peripheral Active Crohn's Disease

Total CD14⁺ macrophages/monocytes, corresponding to node 4(-) (in [Figure 5B](#)), were increased in active IBD mucosa and in peripheral CDa ([Figure 6C](#)) as were IL1B⁺ macrophages/monocytes clusters (clusters 3 and 4) ([Figure 6D](#)). Expansion of IL1B⁺ macrophages/monocytes in active IBD mucosa was confirmed in an independent cohort through in situ hybridization ([Figure 6E](#), [Supplementary Table 5](#)). scRNA-seq analysis of colonic mucosa of 5 active UC subjects also identified a cluster of myeloid cells expressing *IL1B*, *CD14*, and *FCGR3A* (CD16) (cluster 15 in [Supplementary Table 6](#)). Additional cluster-defining genes included inflammatory mediators, such as *IL8*, *S100A8/A9* (calprotectin), *CD163*, *SOD2*, and *LYZ* ([Supplementary Table 6](#)).

Random Forest Modeling Discriminates Active Crohn's Disease vs Active Ulcerative Colitis in Mucosa and Periphery

We applied machine learning with RF to determine whether differences between subject groups based on automated T, B, and innate clustering could classify mucosal and peripheral disease. In brief, the relative abundance of the daughter branches of every node was computed in all samples, and RF analysis was used to model how these

compositional differences vary across subject groups. RF models accurately differentiated CDa and UCa both in the mucosa and periphery ([Figure 7A](#) and [B](#)). Model receiver operating characteristic curves associated with AUC *P* values demonstrate that a 6-feature model discriminates CDa vs UCa mucosa and periphery (*P* = .017 and *P* = .0095, respectively) ([Figure 7A](#) and [B](#)).

Interestingly, half of the top 6 nodes differentiating CDa from UCa mucosa were T-cell nodes ([Figure 7A](#)), and those in the periphery were predominantly innate cell nodes ([Figure 7B](#)). The top 6 nodes differentiating CDa from UCa mucosa were T-cell nodes 35 and 33, which both branch into clusters of HLA-DR⁺CD38⁺ T cells; T-cell node 3, which has a daughter branch giving rise to HLA-DR⁺CD38⁺ T cells and IL17A⁺ CD161⁺ DN EM T cells; B-cell node 23, which branches into clusters of CXCR3⁺ plasmablasts; and innate cell nodes 15 and 18, which branch into ILC1/ILC2/ILC3 populations. The ILC3 branch of innate cell node 18 (the number 1 feature discriminating UCa from CDa mucosa) was enriched in CDa>UCa compared to the ILC2 branch ([Figure 7A](#)). Among the top 6 nodes differentiating peripheral CDa from UCa were multiple innate cell nodes branching into IL1B^{+/−} macrophages/monocytes and DCs, complementing our findings demonstrating enrichment of IL1B⁺ macrophages/monocytes/DCs in peripheral CDa>UCa (in [Figures 5D](#) and [6D](#)).

RF was additionally used to differentiate other subject group pairings ([Supplementary Figure 6A](#) and [B](#)). Based on AUC *P* values, in the mucosa, RF modeling accurately classified non-IBD:UCa, non-IBD:CDa (strong trend with *P* = .079), and UCa:UCi, but not CDa:CDi, perhaps suggesting that differences between CDa and CDi are more subtle ([Supplementary Figure 6A](#)). In the periphery, apart from UCa:CDa, only the comparison between non-IBD and CDa was significant, suggesting that perhaps CD is associated with more systemic inflammation than UC ([Supplementary Figure 6B](#)).

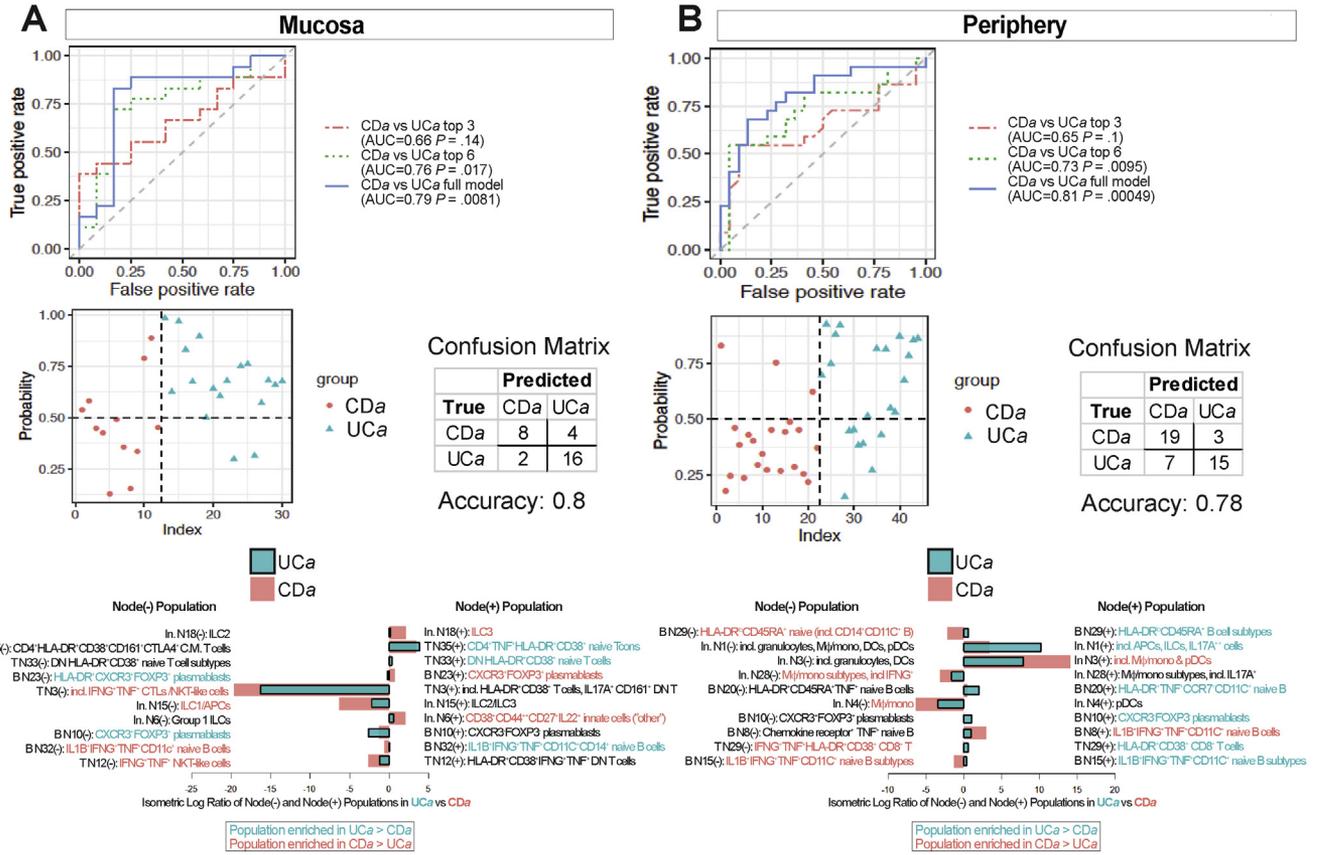
Discussion

UC and CD result from immune dysregulation in the context of genetics and environment that is not adequately understood. Translational investigation using human samples is critical in identifying hallmarks of disease to improve understanding of pathogenesis, discern new therapeutic targets, and expand our diagnostic repertoire.

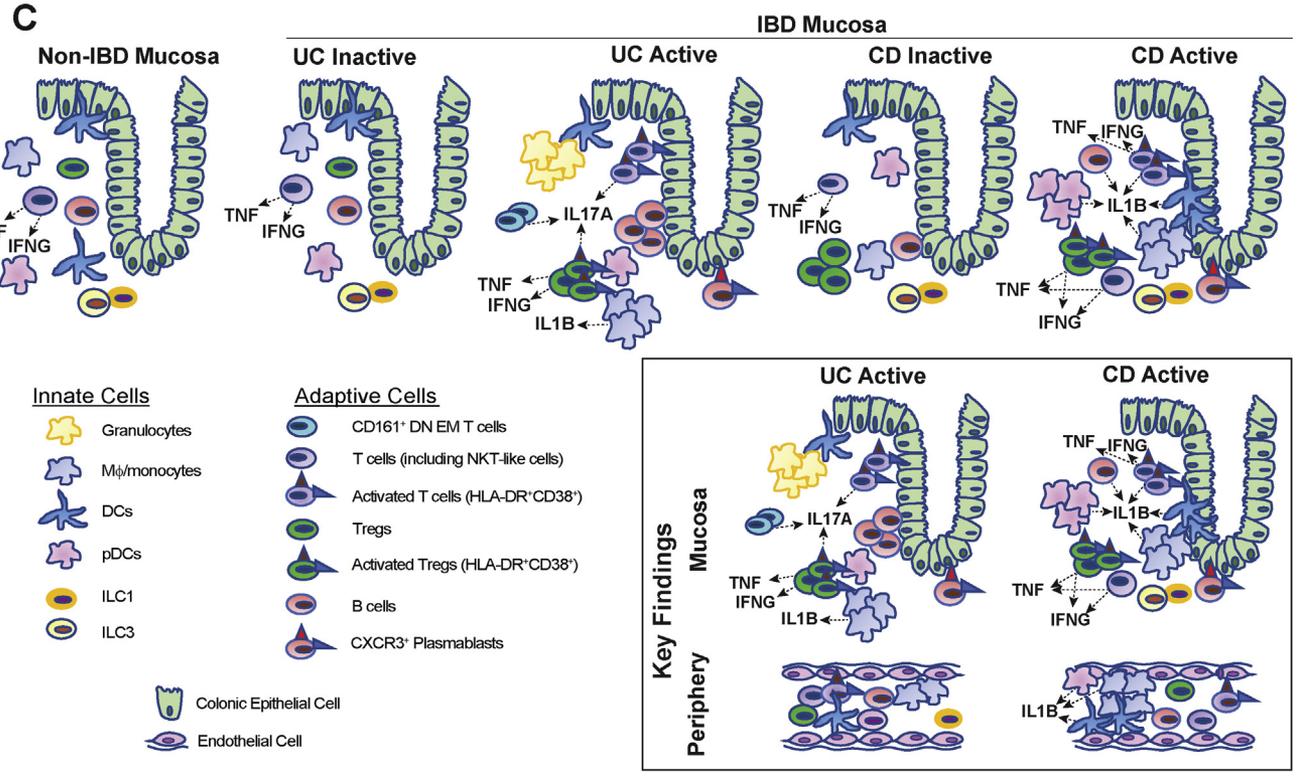
With these relevant issues in mind, we employed single-cell analysis with CyTOF on IBD and non-IBD colonic mucosa and blood to identify disease-specific immune signatures (summarized in [Figure 7C](#)). Many findings have yet to be reported and could potentially be harnessed for disease- or even patient-specific treatments. Furthermore, we used

Figure 6. CyTOF demonstrates differential innate lymphoid cell and IL1B-related macrophage/monocyte signatures in UC vs CD. (A) Nodes 6(-) and 19(+) (from [Figure 5B](#)) abundance plots, t-SNE, and selected marker heatmaps. (B) Cluster 30 (from [Figure 5B](#)) abundance plot, t-SNE, and CCR6 marker heatmap. (C) Node 4(-) (from [Figure 5B](#)) abundance plot, t-SNE, and CD14 marker heatmap. (D) Clusters 3 and 4 (from [Figure 5B](#)) abundance plot (combined), t-SNE, and IL1B marker heatmap. (E) Box-and-whisker plot of IL1B+CD14+ cells counted in in situ hybridization images (4 images per subject) with representative image.

Random Forest: Active CD versus Active UC



Disease-specific Immune Signatures



machine learning to create models based on our data that can accurately classify disease type/activity in both tissue and periphery, potentially heralding new diagnostic avenues in this field.

As expected, immune cell populations in active IBD mucosa were significantly altered. We identified abundant HLA-DR⁺CD38⁺ T cells in *UCa* and *CDa* mucosa, previously reported only in the circulation of immune-mediated disease states, including human immunodeficiency virus⁴¹ and IBD.^{26,42} CD38 has been implicated in colitis in mice⁴³ and CD38⁺ effector T cells in the severity of pediatric IBD.²⁵ Our results suggest that CD38 could be targeted as an IBD therapeutic. Interestingly, we did not reliably observe *HLA-DRA/CD38* transcripts in T cells in scRNA-seq data (or in mining available published scRNA-seq data from human UC colonic mucosa¹⁵), suggesting T cells may acquire HLA-DR/CD38 proteins through membrane fusion upon cellular interaction.⁴⁴

CXCR3⁺ plasmablasts were enriched in active IBD tissue. Plasmablasts, precursors to antibody-secreting plasma cells, migrate to sites of inflammation via the chemokine receptor CXCR3.⁴⁵ Circulating CXCR3⁺ plasmablasts have been described in immune-mediated diseases, including UC,⁴⁶ as have increased plasma cells in UC colonic mucosa.⁴⁷ Our finding of IBD-associated mucosal CXCR3⁺ plasmablasts complements and builds upon this knowledge. Classically associated with Tregs and suppression, FOXP3 expression in non-T cells is controversial; nevertheless, we did detect FOXP3 expression in the CXCR3⁺ plasmablast population, suggesting these cells may possibly play a regulatory role in inflamed mucosa.

Our results identified several disease-specific mucosal signatures related to differential cytokine expression. In T cells, EM subsets were stratified in *UCa* vs *CDa* mucosa along a cytokine axis, with IL17A⁺ T cells expanded in *UCa* and IFNG⁺ TNF⁺ T cells expanded in *CDa*. The role of T-helper 17 cells in IBD has been controversial, with some studies suggesting pathogenic implication in CD,^{48,49} while others suggesting a UC-specific signature.⁵⁰ MAIT cells, known to express pro-inflammatory cytokines, including IFNG and IL17A, have not been rigorously studied in IBD although enrichment of MAIT cells in IBD mucosa has been reported.⁵¹ In this study, we identify a UC-specific IL17A⁺ CD161⁺ EM T-cell subset, which could represent MAIT cells or T cells with a similar innate phenotype,⁸ and based on scRNA-seq analysis, may be characterized by the T-helper 17-related transcription factor *RORA*. While IL17A blockade was not efficacious in treating CD,⁵² its use in UC might be considered.

Our analysis of Tregs similarly identified expansion of IL17A⁺ memory subsets specific to *UCa*, belonging to umbrella populations of highly activated IBD-associated memory Tregs. Whereas one population of IBD-enriched, activated memory Tregs was FOXP3⁺⁺, the other was FOXP3^{lo} with intense pro-inflammatory cytokine expression (IFNG⁺TNF⁺IL17A^{+/-}), suggesting that Treg function in IBD may be altered, as has been suggested by one prior study in CD.²⁹ In addition to these mucosal Treg signatures, we identified the unusual finding of an IL1B⁺ Treg population expanded in peripheral *CDa*.

IL1B signatures specific to CD also involved HLA-DR⁺CD38⁺ T cells, naïve B cells, and DCs. IL1B⁺ macrophages/monocytes were expanded in both *CDa* and *UCa* mucosa, which has been reported previously employing *in vitro* studies⁵³; however, here we also identify expansion of IL1B⁺ monocytes specific to peripheral *CDa*. We hypothesize that in a subgroup of CD patients, targeting IL1B may be a promising strategy.

Peripheral immune population differences distinguishing *UCa* from *CDa* were primarily found in innate immune cells. DCs, pDCs, and CD14⁺ macrophages/monocytes were all expanded, while group 1 ILCs were decreased in peripheral *CDa*. In peripheral *UCa*, we found a decrease (trend) in granulocytes expressing unconventional granulocyte markers, including HLA-DR and CD56 with a concomitant increase in these cells in *UCa* mucosa, potentially reflective of tissue homing. This is consistent with previous data suggesting that intestinal neutrophil infiltration correlates with severity of UC.^{54,55}

Predictive modeling with RF built upon these findings. Many of the top RF nodes branched into clusters that we had identified through our initial analysis as being hallmarks of disease, including HLA-DR⁺CD38⁺ T cells, FOXP3⁺CXCR3^{+/-} plasmablasts, and IL1B^{+/-} macrophages/monocytes and DCs. Other nodes identified through RF revealed important patterns related to specific clusters, such as the enrichment of ILC3>ILC2 in *CDa* mucosa. ILC1 (which we found to be increased in *CDa*>*UCa* mucosa) can transition to an ILC3 phenotype in an IL1-dependent manner,⁵⁶ complementing our data demonstrating IL1-specific signatures in *CDa*. Importantly, RF models could distinguish nearly all mucosal subject group comparisons, but in the periphery could only distinguish *UCa*:*CDa* and *CDa*:non-IBD. This suggests that although robust immunologic signatures are more likely to be mucosal, there are some that can be peripherally detected (especially in *CDa*) potentially of great biological and diagnostic relevance.

Figure 7. Signatures of immune dysregulation in IBD. RF-generated models of *UCa*:*CDa* discrimination in mucosa (A) and periphery (B). Receiver operator characteristic (ROC) curves (*top panels*) with models using top 3, top 6, and full model of input features shown. AUC for each model shown with associated *P* value. Probability plots with associated confusion matrices (*middle panels*) demonstrating model accuracy of the full model. Top 10 features (in descending order of importance) representing isometric log ratio of depicted nodes, with relative abundance of node branches in *bar graphs* and are colored based on preferential abundance in *UCa* vs *CDa* (*bottom panels*). In., innate. (C) Schematic of key findings in IBD mucosa and periphery: IL17A expression by several immune populations in *UCa*>*CDa* mucosa (CD161⁺ DN EM T cells, HLA-DR⁺CD38⁺ T cells and Tregs), IL1B expression in *CDa*>*UCa* mucosa (in DCs, pDCs, HLA-DR⁺CD38⁺ T cells), expansion of granulocytes in *UCa*>*CDa* and ILC1 and ILC3 in *CDa*>*UCa* mucosa; and in the periphery, expansion of IL1B⁺ macrophages/monocytes, DCs, and pDCs in *CDa*>*UCa* and expansion of Group 1 ILCs in *UCa*>*CDa*.

There are several limitations to these studies. Our IBD cohort was heterogenous and given the variety of clinical phenotypes, we were unable to extract immune signatures specific to treatment response or other characteristics. Only a subset of subjects had matched tissue and blood, limiting our interpretation of cellular interchange between mucosal and peripheral compartments. Additional investigation with functional studies is needed to understand whether our findings contribute to pathogenesis and/or severity of disease.

In summary, our results highlight that CyTOF analysis of mucosal and peripheral immune cells can be used to define unique signatures differentiating disease type and activity. A number of these signatures, such as IL17A in UC, IL1B in CD, and CD38 in IBD, already have approved biologic inhibitors that can allow for targeted therapeutics. Further translational work in this field can build upon these findings and potentially identify patient-specific signatures in an evolving era of personalized therapeutic approaches.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at <https://doi.org/10.1053/j.gastro.2020.04.074>.

References

- Dahlhamer JM, Zammitti EP, Ward BW, et al. Prevalence of Inflammatory bowel disease among adults aged \geq 18 years – United States, 2015. *MMWR Morb Mortal Wkly Rep* 2016;65:1166–1169.
- de Souza HS, Fiocchi C. Immunopathogenesis of IBD: current state of the art. *Nat Rev Gastroenterol Hepatol* 2016;13:13–27.
- Neurath MF. Current and emerging therapeutic targets for IBD. *Nat Rev Gastroenterol Hepatol* 2017;14:269–278.
- Cohen BL, Sachar DB. Update on anti-tumor necrosis factor agents and other new drugs for inflammatory bowel disease. *BMJ* 2017;357:j2505.
- Ford AC, Sandborn WJ, Khan KJ, et al. Efficacy of biological therapies in inflammatory bowel disease: systematic review and meta-analysis. *Am J Gastroenterol* 2011;106:644–659; quiz 660.
- Reinecker HC, Steffen M, Witthoef T, et al. Enhanced secretion of tumour necrosis factor-alpha, IL-6, and IL-1 beta by isolated lamina propria mononuclear cells from patients with ulcerative colitis and Crohn's disease. *Clin Exp Immunol* 1993;94:174–181.
- Marchal-Bressenot A, Scherl A, Salleron J, et al. A practical guide to assess the Nancy histological index for UC. *Gut* 2016;65:1919–1920.
- Konnikova L, Boschetti G, Rahman A, et al. High-dimensional immune phenotyping and transcriptional analyses reveal robust recovery of viable human immune and epithelial cells from frozen gastrointestinal tissue. *Mucosal Immunol* 2018;11:1684–1693.
- Kotecha N, Krutzik PO, Irish JM. Web-based analysis and publication of flow cytometry experiments. *Curr Protoc Cytom* 2010. Chapter 10:Unit10 7.
- Chen H, Lau MC, Wong MT, et al. Cytofkit: A Bioconductor Package for an Integrated Mass Cytometry Data Analysis Pipeline. *PLoS Comput Biol* 2016;12:e1005112.
- Van Gassen S, Callebaut B, Van Helden MJ, et al. FlowSOM: using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry A* 2015; 87:636–645.
- van der Maaten L. Visualizing Data using t-SNE. *J Mach Learn Res* 2008;9:2579–2605.
- Warnes GR, Bolker B, Bonebakker L, et al. gplots: Various R Programming Tools for Plotting Data. <https://cran.r-project.org/web/packages/gplots/index.html>. Published 2020. Accessed June 10, 2020.
- Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag, 2016.
- Smillie CS, Biton M, Ordovas-Montanes J, et al. Intra- and Inter-cellular rewiring of the human colon during ulcerative colitis. *Cell* 2019;178:714–730.e22.
- Ordovas-Montanes J, Dwyer DF, Nyquist SK, et al. Allergic inflammatory memory in human respiratory epithelial progenitor cells. *Nature* 2018;560:649–654.
- Hughes TK, Wadsworth MH, Gierahn TM, et al. Highly efficient, massively-parallel single-cell RNA-Seq reveals cellular states and molecular features of human skin pathology. *bioRxiv* 2019:689273.
- Satija R, Farrell JA, Gennert D, et al. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol* 2015;33:495–502.
- Svetnik V, Liaw A, Tong C, et al. Random forest: a classification and regression tool for compound classification and QSAR modeling. *J Chem Inf Comput Sci* 2003;43:1947–1958.
- Egozcue JJ, Pawlowsky-Glahn V, Mateu-Figueras G, et al. Isometric logratio transformations for compositional data analysis. *Math Geol* 2003;35:279–300.
- Fergusson JR, Smith KE, Fleming VM, et al. CD161 defines a transcriptional and functional phenotype across distinct human T cell lineages. *Cell Rep* 2014;9:1075–1088.
- Walker LJ, Kang YH, Smith MO, et al. Human MAIT and CD8alpha cells develop from a pool of type-17 precommitted CD8+ T cells. *Blood* 2012;119:422–433.
- Koay HF, Su S, Amann-Zalcenstein D, et al. A divergent transcriptional landscape underpins the development and functional branching of MAIT cells. *Sci Immunol* 2019;4(41):eaay6039.
- Shaler CR, Choi J, Rudak PT, et al. MAIT cells launch a rapid, robust and distinct hyperinflammatory response to bacterial superantigens and quickly acquire an anergic phenotype that impedes their cognate antimicrobial function: Defining a novel mechanism of superantigen-induced immunopathology and immunosuppression. *PLoS Biol* 2017;15:e2001930.
- Joose ME, Menckeborg CL, de Ruiter LF, et al. Frequencies of circulating regulatory TIGIT(+)/CD38(+) effector T cells correlate with the course of inflammatory bowel disease. *Mucosal Immunol* 2019;12:154–163.

26. Funderburg NT, Stubblefield Park SR, Sung HC, et al. Circulating CD4(+) and CD8(+) T cells are activated in inflammatory bowel disease and are associated with plasma markers of inflammation. *Immunology* 2013; 140:87–97.
27. Baecher-Allan C, Wolf E, Hafler DA. MHC class II expression identifies functionally distinct human regulatory T cells. *J Immunol* 2006;176:4622–4631.
28. Miyara M, Yoshioka Y, Kitoh A, et al. Functional delineation and differentiation dynamics of human CD4+ T cells expressing the FoxP3 transcription factor. *Immunity* 2009;30:899–911.
29. Hovhannisyan Z, Treatman J, Littman DR, et al. Characterization of interleukin-17-producing regulatory T cells in inflamed intestinal mucosa from patients with inflammatory bowel diseases. *Gastroenterology* 2011; 140:957–965.
30. Koenecke C, Lee CW, Thamm K, et al. IFN-gamma production by allogeneic Foxp3+ regulatory T cells is essential for preventing experimental graft-versus-host disease. *J Immunol* 2012;189:2890–2896.
31. Esposito M, Ruffini F, Bergami A, et al. IL-17- and IFN-gamma-secreting Foxp3+ T cells infiltrate the target tissue in experimental autoimmunity. *J Immunol* 2010; 185:7467–7473.
32. Maul J, Loddenkemper C, Mundt P, et al. Peripheral and intestinal regulatory CD4+ CD25(high) T cells in inflammatory bowel disease. *Gastroenterology* 2005;128:1868–1878.
33. Holmen N, Lundgren A, Lundin S, et al. Functional CD4+CD25high regulatory T cells are enriched in the colonic mucosa of patients with active ulcerative colitis and increase with disease activity. *Inflamm Bowel Dis* 2006;12:447–456.
34. Ma S, Wang C, Mao X, et al. B cell dysfunction associated with aging and autoimmune diseases. *Front Immunol* 2019;10:318.
35. Ziegler-Heitbrock HW, Pechumer H, Petersmann I, et al. CD14 is expressed and functional in human B cells. *Eur J Immunol* 1994;24:1937–1940.
36. Vadasz Z, Toubi E. FoxP3 expression in macrophages, cancer, and B cells—is it real? *Clin Rev Allergy Immunol* 2017;52:364–372.
37. Mao H, Pan F, Wu Z, et al. CD19(lo)CD27(hi) plasmablasts suppress harmful Th17 inflammation through interleukin 10 pathway in colorectal cancer. *DNA Cell Biol* 2017;36:870–877.
38. Lok LSC, Dennison TW, Mahubani KM, et al. Phenotypically distinct neutrophils patrol uninfected human and mouse lymph nodes. *Proc Natl Acad Sci U S A* 2019; 116:19083–19089.
39. Lin A, Lore K. Granulocytes: new members of the antigen-presenting cell family. *Front Immunol* 2017; 8:1781.
40. Roan F, Stoklasek TA, Whalen E, et al. CD4+ group 1 innate lymphoid cells (ILC) form a functionally distinct ilc subset that is increased in systemic sclerosis. *J Immunol* 2016;196:2051–2062.
41. Kestens L, Vanham G, Gigase P, et al. Expression of activation antigens, HLA-DR and CD38, on CD8 lymphocytes during HIV-1 infection. *AIDS* 1992;6:793–797.
42. Rubin SJS, Bai L, Haileselassie Y, et al. Mass cytometry reveals systemic and local immune signatures that distinguish inflammatory bowel diseases. *Nat Commun* 2019;10:2686.
43. Schneider M, Schumacher V, Lischke T, et al. CD38 is expressed on inflammatory cells of the intestine and promotes intestinal inflammation. *PLoS One* 2015;10: e0126007.
44. Game DS, Rogers NJ, Lechler RI. Acquisition of HLA-DR and costimulatory molecules by T cells from allogeneic antigen presenting cells. *Am J Transplant* 2005;5:1614–1625.
45. Radbruch A, Muehlinghaus G, Luger EO, et al. Competence and competition: the challenge of becoming a long-lived plasma cell. *Nat Rev Immunol* 2006;6:741–750.
46. Hosomi S, Oshitani N, Kamata N, et al. Increased numbers of immature plasma cells in peripheral blood specifically overexpress chemokine receptor CXCR3 and CXCR4 in patients with ulcerative colitis. *Clin Exp Immunol* 2011;163:215–224.
47. Mitsuishi T. Correlation between histological findings and endoscopic findings in patients with ulcerative colitis: Basal plasmacytosis is an important finding suggesting active inflammation. *JGH Open* 2019; 3:100–104.
48. Brand S. Crohn's disease: Th1, Th17 or both? The change of a paradigm: new immunological and genetic insights implicate Th17 cells in the pathogenesis of Crohn's disease. *Gut* 2009;58:1152–1167.
49. Seiderer J, Elben I, Diegelmann J, et al. Role of the novel Th17 cytokine IL-17F in inflammatory bowel disease (IBD): upregulated colonic IL-17F expression in active Crohn's disease and analysis of the IL17F p.His161Arg polymorphism in IBD. *Inflamm Bowel Dis* 2008;14:437–445.
50. Rosen MJ, Karns R, Vallance JE, et al. Mucosal expression of type 2 and type 17 immune response genes distinguishes ulcerative colitis from colon-only Crohn's disease in treatment-naïve pediatric patients. *Gastroenterology* 2017;152:1345–1357 e7.
51. Tominaga K, Yamagiwa S, Setsu T, et al. Possible involvement of mucosal-associated invariant T cells in the progression of inflammatory bowel diseases. *Biomed Res* 2017;38:111–121.
52. Hueber W, Sands BE, Lewitzky S, et al. Secukinumab, a human anti-IL-17A monoclonal antibody, for moderate to severe Crohn's disease: unexpected results of a randomised, double-blind placebo-controlled trial. *Gut* 2012; 61:1693–1700.
53. Mahida YR, Wu K, Jewell DP. Enhanced production of interleukin 1-beta by mononuclear cells isolated from mucosa with active ulcerative colitis of Crohn's disease. *Gut* 1989;30:835–838.
54. Raab Y, Gerdin B, Ahlstedt S, et al. Neutrophil mucosal involvement is accompanied by enhanced local production of interleukin-8 in ulcerative colitis. *Gut* 1993; 34:1203–1206.
55. Bressenot A, Salleron J, Bastien C, et al. Comparing histological activity indexes in UC. *Gut* 2015;64:1412–1418.

56. Bernink JH, Krabbendam L, Germar K, et al. Interleukin-12 and -23 control plasticity of CD127(+) group 1 and group 3 innate lymphoid cells in the intestinal lamina propria. *Immunity* 2015;43:146–160.

Received February 26, 2019. Accepted April 24, 2020.

Correspondence

Address correspondence to: Scott B. Snapper, MD, PhD, Division of Gastroenterology, Hepatology and Nutrition, Boston Children's Hospital, 300 Longwood Avenue, Boston, Massachusetts 02115. e-mail: scott.snapper@childrens.harvard.edu; fax: 617-730-0498 and Liza Konnikova, MD, PhD, Section of Neonatal-Perinatal Medicine, Department of Pediatrics, Yale School of Medicine, 375 Congress Ave, LSOG 405B, New Haven, Connecticut 06519. e-mail: liza.konnikova@yale.edu.

Acknowledgments

Boston Children's Hospital Inflammatory Bowel Disease Center: Sonia Ballal, MD, Silvana Bonilla, MD, MS, Rima Fawaz, MD, Laurie N. Fishman, MD, Alejandro Flores, MD, Victor Fox, MD, Amit S. Grover, MB, BCH, BAO, Leslie Higuchi, MD, Susanna Huh, MD, Stacy Kahn, MD, Christine Lee, MD, Munir Mobassaleh, MD, Jodie Ouahed, MD, Randi G. Pleskow, MD, Brian Regan, DO, Paul A. Rufo, MD, MMSc, Sabina Sabharwal, MD, Jared Silverstein, MD, Menno Verhave, MD, Anne Wolf, MD, Lori Zimmerman, MD, Naamah Zitomersky, MD. Brigham and Women's Hospital Crohn's and Colitis Center: Jessica R. Allegretti, MD, MPH, Punyanganie De Silva, MD, MPH, Sonia Friedman, MD, Matthew Hamilton, MD, Joshua Korzenik, MD, Frederick Makrauer, MD, Beth-Ann Norton, MS, RN, ANP-BC, Rachel W. Winter, MD, MPH.

CRedit Authorship Contributions

Vanessa Mitsialis, MD (Conceptualization: Equal; Data curation: Equal; Formal analysis: Lead; Project administration: Equal; Writing – original draft: Lead; Writing – review & editing: Lead). Sarah Wall, BA (Data curation: Supporting; Methodology: Supporting). Peng Liu, BA (Formal analysis: Equal). Jose Ordovas-Montanes, PhD (Formal analysis: Equal). Tamar Parmet, BA (Data curation: Supporting; Methodology: Supporting). Marko Vukovic, BA (Data curation: Supporting; Formal analysis: Supporting). Dennis Spencer, MD (Data curation: Supporting; Formal analysis: Supporting). Michael Field, BA (Data curation: Supporting; Methodology: Supporting). Collin McCourt, BA

(Data curation: Supporting; Formal analysis: Supporting). Jessica Toothaker, BA (Formal analysis: Supporting). Athos Bousvaros, MD (Investigation: Supporting; Resources: Equal). Alex K Shalek, PhD (Methodology: Equal; Supervision: Equal). Leslie Kean, MD, PhD (Methodology: Equal; Supervision: Equal). Bruce Horwitz, MD, PhD (Conceptualization: Supporting; Formal analysis: Supporting; Supervision: Supporting). Jeffrey Goldsmith, MD (Data curation: Supporting; Formal analysis: Supporting; Resources: Supporting). George Tseng, PhD (Methodology: Supporting; Supervision: Equal). Scott Snapper, MD, PhD (Conceptualization: Lead; Funding acquisition: Lead; Supervision: Lead; Writing – original draft: Equal; Writing – review & editing: Equal). Liza Konnikova, MD, PhD (Data curation: Supporting; Formal analysis: Supporting; Funding acquisition: Equal; Investigation: Equal; Supervision: Lead; Writing – original draft: Equal; Writing – review & editing: Equal).

Conflicts of interest

These authors disclose the following: Scott B. Snapper is supported by grants or in-kind contributions from Pfizer, Novartis, Janssen, Merck, and Regeneron. He is on the scientific advisory boards of Pfizer, Janssen, IFM Therapeutics, Lycera, Inc., Celgene, Lilly, Pandion Therapeutics, and Applied Molecular Transport. He has consulted for Amgen and Hoffman La-Roche. Liza Konnikova is supported by funding from NGMB Biopharmaceuticals. Alex K. Shalek has received compensation for consulting and/or SAB membership from Merck, Honeycomb Biotechnologies, Cellarity, Cogen Therapeutics and Dahlia Biosciences, as well as grants from Novartis, Janssen, and Merck. The remaining authors disclose no conflicts.

Funding

Vanessa Mitsialis is a recipient of a Research Fellowship Award from the Crohn's and Colitis Foundation (RFA 622723). Scott B. Snapper is supported by National Institutes of Health grants R01 DK115217, R56 AI125766; P30DK034854, the Helmsley Charitable Trust, the Wolpov Family Chair in IBD Treatment and Research, and the Translational Investigator Service at Boston Children's Hospital. Liza Konnikova is a recipient of a Career Development Award Grant from the Crohn's and Colitis Foundation (CDA 422348). Alex K. Shalek is supported by the Beckman Young Investigator Program, a Sloan Fellowship in Chemistry, and NIH 5U24AI118672. Jose Ordovas-Montanes was supported by the Richard and Susan Smith Family Foundation, the HHMI Damon Runyon Cancer Research Foundation Fellowship (DRG-2274-16), the AGA Research Foundation's AGA-Takeda Pharmaceuticals Research Scholar Award in IBD – AGA2020-13-01, the HDDC Pilot and Feasibility P30 DK034854, and the Food Allergy Science Initiative.

Supplementary Methods

Mass Cytometry: Staining Protocol

After stimulation with phorbol myristate acetate/ionomycin/GolgiSTOP, cell suspensions were spun down at 1400 rpm \times 5 minutes and resuspended in 500 μ L of CyTOF staining buffer (CSB; 500 mL low-barium phosphate-buffered saline (PBS) plus 2.5 g bovine serum albumin [Sigma #A3059] plus 100 mg sodium azide [Sigma #71289]) in 1.5-mL Eppendorf tubes. Cells were spun down at 2000 rpm \times 5 minutes and resuspended in 500 μ L of solution containing rhodium viability dye (Cell-ID Intercalator-Rh—500 μ M; Fluidigm #201103A) (1:2000 dilution in CSB) for 10 minutes at room temperature. Cells were spun down at 2000 rpm \times 5 minutes, supernatant carefully discarded, and 5 μ L of FcX-block (Human TruStain FcX; BioLegend, San Diego, CA; #422302) was added followed by gentle disruption of pellet, followed by incubation for 10 minutes at room temperature.

At this point, a cocktail of surface marker antibodies was added to each sample (first, a master mix was created—each sample is allotted 0.5 μ L of each surface marker antibody) followed by gentle disruption of the pellets. Surface stains were incubated at room temperature for 30 minutes. At the end of incubation, 500 μ L of CSB was added directly to the samples, which were then spun down at 2000 rpm \times 5 minutes and then washed again by resuspension in 500 μ L CSB, followed by spinning down at 2000 rpm \times 5 minutes.

Samples were then resuspended in 500 μ L FoxP3 fix/perm (Thermo Fisher; #00-5523) for 45 minutes at room temperature. At the end of incubation, 1 mL of FoxP3 permeabilization buffer (Thermo Fisher; #00-5523) was added to each sample, followed by spinning down at 2600 rpm \times 5 minutes. Samples were then resuspended in 1 mL FoxP3 perm buffer and spun down again at 2600 rpm with careful removal of supernatant.

At this point, a cocktail of intracellular marker antibodies was added directly to the pellet (first, a master mix was created—each sample is allotted 0.5 μ L of each intracellular marker antibody) followed by gentle disruption of the pellets. Intracellular stains were incubated at room temperature \times 45 min. After incubation, 500 μ L of CSB was added to the cells for washing, then the samples were spun down at 2600 rpm, followed by another wash of 500 μ L CSB and spinning down at 2600 rpm. Pellets were then resuspended in 500 μ L of 1.6% paraformaldehyde for 10 minutes at room temperature. The paraformaldehyde was then spun out at 2600 rpm \times 5 minutes and cells were resuspended in 500 μ L CSB and kept at 4°C overnight or up to 3 days.

On the day of acquisition, the samples were spun down at 2600rpm \times 5 minutes and resuspended in 500 μ L solution inclusive of an iridium intercalator (Cell-ID Intercalator-Ir—125 μ M; Fluidigm #201192A; 1:1000 in CSB) and incubated at room temperature \times 20 minutes. Cells were then spun down at 2600 rpm, resuspended in 500 μ L CSB twice, for a total of 2 washes. Then, cells were washed twice with 500 μ L water (Maxpar Water from Fluidigm

#201069 or MilliQ water). Cells were ultimately resuspended in 250–500 μ L water, depending on cellular concentration (goal of approximately 500,000 cells per mL) and then EQ beads (EQ Four Element Calibration Beads; Fluidigm #201078) were added for a 1:10 dilution, for normalization. Cell suspensions were then transferred through a filter (cap of Falcon 5 mL Round Bottom Polystyrene Test Tube, with Cell Strainer Snap Cap; Corning, Corning, NY; #352235) into 5-mL polypropylene tubes (Falcon 5 mL Round Bottom Polypropylene Test Tube) and brought to the Longwood Medical Area CyTOF Core for acquisition on Helios2 cytometer.

Flow Cytometry: Staining Protocol

For the main validation cohort. Stimulated samples were spun down at 1400 rpm \times 5 minutes, resuspended in T-cell media, and distributed into a 96-well plate. Extra tissue was processed for use as single stain controls, fluorescence-minus-one (FMO) controls (only for intracellular stains), and unstained sample for calibration of voltages during acquisition. The plate was spun down at 1400 rpm \times 5 minutes and all samples and all FMOs (plus the appropriate single-stain control) were resuspended in 50 μ L of PBS with 1:1000 of viability stain (eBioscience Fixable Viability Dye eFluor 506; Thermo Fisher Scientific; #65-0866-14); other samples were resuspended in 50 μ L PBS, and the plate was incubated \times 10 minutes in the dark at room temperature. Two hundred microliters flow cytometry staining buffer was then added to all wells and the plate was spun down at 1400 rpm \times 5 minutes and supernatant discarded. Fifty microliters of surface staining cocktail was then added to each sample well and FMO wells (antibodies were diluted ranging from 1:100 to 1:25 in flow cytometry staining buffer in a master mix, first). Single-stain controls were treated with 50 μ L flow cytometry staining buffer plus 1 μ L of a surface-stain antibody specific to each channel (used for downstream calibration of cytometer and for compensation). Surface stains were incubated at 4°C in dark \times 30 minutes.

Each well was then treated with 200 μ L flow cytometry staining buffer and the plate was spun down at 1400 rpm \times 5 minutes, and the supernatant discarded. Cells were resuspended in 100 μ L of FoxP3 fix/perm (Thermo Fisher; #00-5523) \times 35 minutes at room temperature in the dark. Two hundred microliters of FoxP3 permeabilization buffer (Thermo Fisher; #00-5523) was added to each sample, the plate spun down at 2600 rpm \times 5 minutes and supernatant discarded. Cells were resuspended again in FoxP3 permeabilization buffer and kept overnight at 4°C in the dark.

The next day, the cells were spun down at 2600 rpm \times 5 minutes, supernatant discarded, and were resuspended in 50 μ L of an intracellular antibody cocktail (antibodies first diluted ranging from 1:100 to 1:25 in FoxP3 permeabilization buffer) and incubated at 4°C in the dark \times 35 minutes. FMOs were incubated with an intracellular cocktail that lacked one of the targets, for example, antibodies against IFNG, IL17A, or FOXP3). Two hundred microliters FoxP3 permeabilization buffer was added to the cells, the plate

was spun down at 2600 rpm, and the cells were resuspended in 300 μ L FoxP3 permeabilization buffer and acquired on BD LSR Fortessa using FACSDiva software.

To Assess IL1B Staining. Digested tissue was spun down and resuspended in T-cell media and transferred to a 96-well plate. No stimulation was applied to these samples. Samples were spun down at 1400 rpm \times 5 minutes and resuspended in 50 μ L of viability stain, as above, \times 10 minutes at room temperature in the dark. The cells were washed with PBS. Cells were then treated with Fc block (Human TruStain FcX; BioLegend #422302) \times 10 minutes at room temperature in the dark. Cells were washed with PBS, and then resuspended in FoxP3 fix/perm (100 μ L) \times 35 minutes at room temperature. Cells were washed with FoxP3 permeabilization buffer and resuspended in 50 μ L FoxP3 permeabilization buffer with 1:10 dilution of mouse anti-human IL1B primary antibody (clone CRM56, Invitrogen, Carlsbad, CA; catalog no. 14-7018-81, RRID: AB_468400) and incubated \times 45 min at 4°C. Cells were then washed with FoxP3 permeabilization buffer and resuspended in 50 μ L of FoxP3 permeabilization buffer with 1:25 dilution of goat anti-mouse IgG1 conjugated to phycoerythrin (Clone RMG1-1; BioLegend; catalog no. 406607, RRID: AB_10551439) \times 45 minutes at 4°C in the dark. Cells were then washed with flow cytometry staining buffer \times 2. Cells were then resuspended in 50 μ L of a master mix of both surface and intracellular antibodies diluted in FoxP3 permeabilization buffer (ranging from 1:100 to 1:25 dilution) and incubated \times 30 minutes at 4°C in the dark. Cells were then washed with FCS buffer and resuspended in 300 μ L flow cytometry staining buffer for subsequent acquiring on BD LSR Fortessa. Appropriate controls, including single stains, unstained, and specific controls for IL1B staining (eg, omission of primary anti-IL1B antibody \pm additional omission of anti-mouse IgG1 secondary antibody) were performed accordingly.

Single-Cell RNA-Sequencing: Tissue Preparation

Cryopreserved mucosal samples (2–3 biopsies/sample, or equivalently-sized pieces of resected surgical tissue, as described in Konnikova et al⁸) were thawed in a 37°C rock bath. Biopsy bites were rinsed in 30 mL of ice-cold PBS (Thermo Fisher; catalog no. 10010-049) and allowed to settle. Each individual bite was then transferred to 10 mL epithelial cell solution (HBSS Ca/Mg-Free, 10 mM EDTA, 100 U/mL penicillin, 100 mg/mL streptomycin, 10 mM HEPES, and 2% FCS [all Thermo Fisher]) freshly supplemented with 100 μ L 0.5M EDTA. Separation of the epithelial layer from the underlying lamina propria was performed for 15 minutes at 37°C at 140 rpm in a dry shaking incubator. Samples were placed on ice for 10 minutes, then shaken vigorously 15 times. Visual macroscopic inspection of the tube at this point yielded visible epithelial sheets, and microscopic examination confirmed the presence of single-layer sheets and crypt-like structures. The remnant lamina propria (LP) was carefully transferred into a large volume (30 mL) of ice-cold PBS to rinse, and then into 10 mL enzymatic digestion mix (Base: RPMI1640, 100 U/mL

penicillin, 100 mg/mL streptomycin, 10 mM HEPES, 2% FCS, and 50 mg/mL gentamicin [all Thermo Fisher]), supplemented with 100 mg/mL Liberase TM [Roche] and 100 mg/mL DNaseI [Roche]), at 37°C at 120 rpm for 30 minutes. During this 30-minute LP digestion, the epithelial fraction was spun down at 400g for 7 minutes and resuspended in 500 μ L of epithelial cell solution. Cells were resuspended in ACK lysis buffer (Thermo Fisher; catalog no. A1049201) with gentle resuspension, incubated on ice for 4 minutes, and spun down at 300g for 4 minutes. This step removes red blood cells and was used even if no red blood cell contamination was visibly observed in order to maintain consistency across samples. Cells were resuspended in TrypLE express enzyme (Thermo Fisher; catalog no. 12604013) for 4 minutes in a 37°C bath, followed by gentle trituration with a P1000 pipette and filtering into a new conical tube through a 40- μ m cell strainer (Falcon/VWR 21008-949). Filter was washed with 20 mL of ice-cold PBS. Cells were spun down at 300g for 4 minutes and resuspended in 1 mL of epithelial cell solution and placed on ice, while final steps of LP dissociation occurred. The LP enzymatic dissociation was quenched with 1 mL 100% FCS (Thermo Fisher) and 80 μ L 0.5M EDTA and placed on ice \times 5 minutes, followed by filtering through 40- μ m cell strainer and rinsing with PBS to 30 mL total volume.

This tube was spun down at 400g for 10 minutes and resuspended in 1 mL of ACK and placed on ice for 3 minutes. Cells were spun down at 400g for 4 minutes and resuspended in 1 mL of epithelial cell solution and spun down at 400g for 2 minutes and resuspended in 400 μ L of epithelial cell solution and placed on ice.

Single-Cell RNA-Sequencing: Seq-Well, Sequencing, Alignment

Epithelial and lamina propria samples were then run on separate Seq-Well arrays as described in Ordovas-Montanes et al,¹⁶ with the improved protocol for library preparation based on second-strand synthesis for complementary DNA (S-3) from Hughes et al.¹⁷ Two arrays were sequenced per sequencing run with an Illumina 75 Cycle NextSeq500/550v2.5 kit on an Illumina NextSeq at a final concentration of 2.2 pM. The read structure was paired end with read 1 starting from a custom read 1 primer containing 20 bases with a 12-bp cell barcode and 8-bp unique molecular identified and read 2 containing 50 bases of transcript information. Read alignment was performed as done previously.^{16,17} Quality-filtered base calls were converted to demultiplexed FASTQ files and aligned to the Hg19 genome using Cell Ranger on the Galaxy portal maintained by the Broad Institute.

Single-Cell RNA-Sequencing: Data Analysis

Merging of datasets, clustering, and differential gene expression analyses were performed using Seurat, version 3.1¹⁸ (<https://satijalab.org/seurat/>). In brief, the unique molecular identified count matrices from the LP sequencing data of the 5 included subjects were read into Seurat objects. Each Seurat object was subsetted so that only those

cells with at least 200 and no more than 2500 features (genes), and no more than 15% mitochondrial-specific genes, were included. Seurat objects were then merged. Count data were scaled using the default scale factor of 10,000 and then log-normalized (natural log +1) as per Seurat, version 3.1, implementation. Then, the top 2000 differentially expressed genes were identified using the *vst* method and the data were scaled using a linear transformation to enable dimensionality reduction techniques, such as principal component analysis and t-SNE. Principal component analysis was applied and the top 50 principal components were used to construct a Shared Nearest Neighbor Graph (KNN), which was then used for clustering with the default Louvain algorithm per Seurat, version 3.1, with a resolution of 1.2 (total cells analyzed = 3979) and generation of 20 clusters. The Wilcoxon rank sum test was used to identify differentially expressed genes in each cluster compared to all others (this list can be found in [Supplementary Table 6](#)). To subset on T cells, a T-cell feature list (comprised of all genes belonging to *TRAV*, *TRB*,

TRG groups, in addition to other T-cell-related genes *TRDC*, *TRDJ1*, *CD8G*, *CD8D*, *CD3E*, *CD247*, *CD4*, *CD8A*, *CD8B*) was applied to the data using the *AddModuleScore* feature. First, the complete Seurat object data set was subsetted on all clusters that appeared to be T cells based on their T cell feature/module score (clusters 3, 6, 7, 8, 10, 16 in [Supplementary Table 6](#)). Then, this new data set was subsetted so that only those cells with a T-cell feature/module score >0 were included. The top 2000 differentially expressed genes in this new data set were computed, data then scaled, principal component analysis run, KNN neighbors were identified using the first 20 principal components and clustered with a resolution of 1.2. Two of the generated clusters were clearly B cells and plasma cells based on differential gene expression and the data set was subsetted once again to exclude those 2 clusters. On this new Seurat object (total cells = 718), the same pipeline was run, with generation of 9 clusters that appeared to all be T cells based on differential gene expression (this list can be found in [Supplementary Table 7](#)).

A CyTOF Panel A and B (Shared Targets)

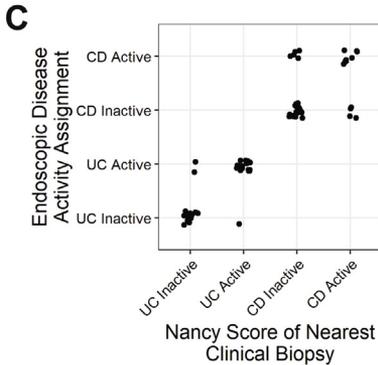
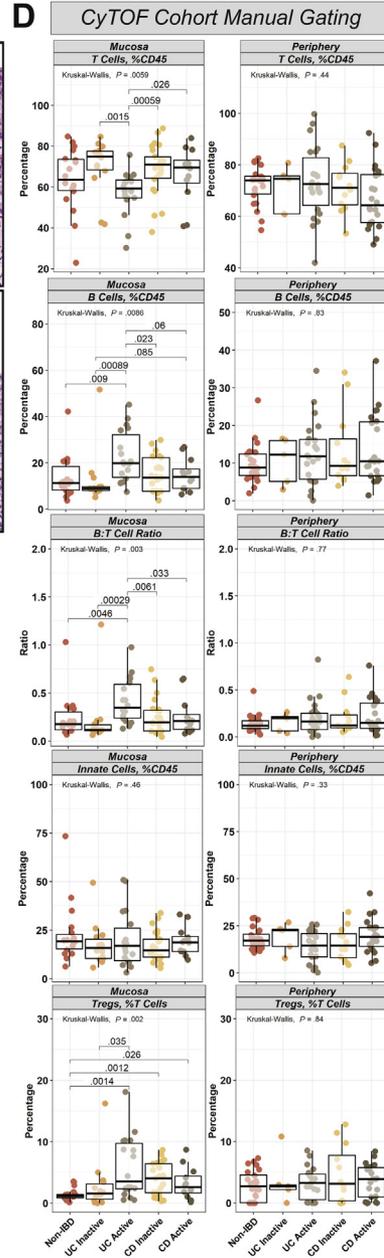
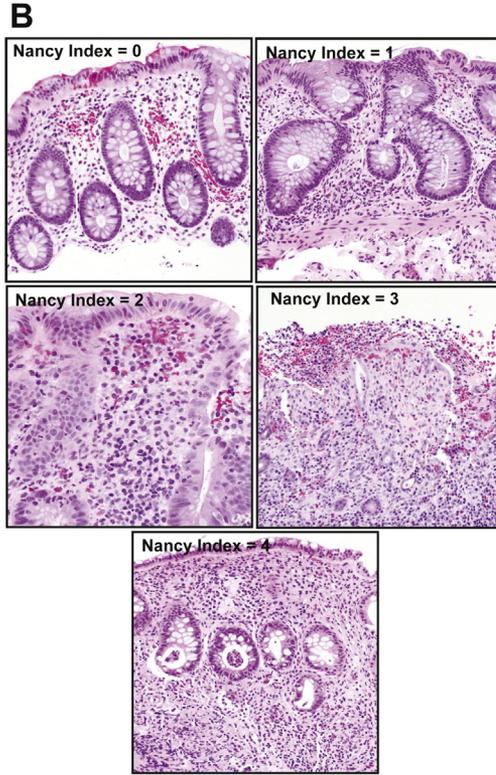
Marker	Tag	Clone
CD45	Y89	HI30
CD19	142Nd	HIB19
HLA-DR	143Nd	L243
TNF	144Nd	MAB11
CD8A	146Nd	RPA-T8
CD45RO	147Sm	UCHL1
CD14	148Nd	RMO52
CD25	149Sm	2A3
IL22	150Nd	22URTI
CD123	151Eu	6H6
CD45RA	153Eu	HI100
CD38	154Sm	HIT2
CD27	155Gd	L128
CD3	158Gd	UCHT1
CCR7	159Tb	G043H7
IFNG	160Gd	4S.B3
IL1B	162Dy	CRM56
CXCR3	163Dy	G025H7
CD161	164Dy	HP-3610
CCR6	168Er	G034E3
IL17A	169Tm	BL168
CD127	171Yb	eBioRDR5
IL21	172Yb	3A3-N2
CD4	174Yb	SK3
CD56	176Yb	HCD56

CyTOF Panel A (Unique Targets)

Marker	Tag	Clone
CD44	115In	IM7
CK1T	141Pr	104D2
CD16	145Nd	3G8
CTLA4	152Sm	L3D10
CCR4	156Gd	L291H4
AHR	161Dy	FF3399
FOXP3	165Ho	PCH101
CD24	166Er	ML5
NKP46	173Yb	9E2
TBET	175Lu	4B10

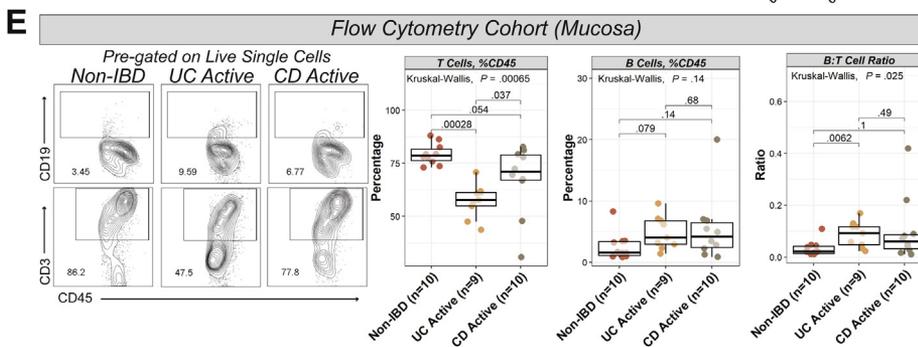
CyTOF Panel B (Unique Targets)

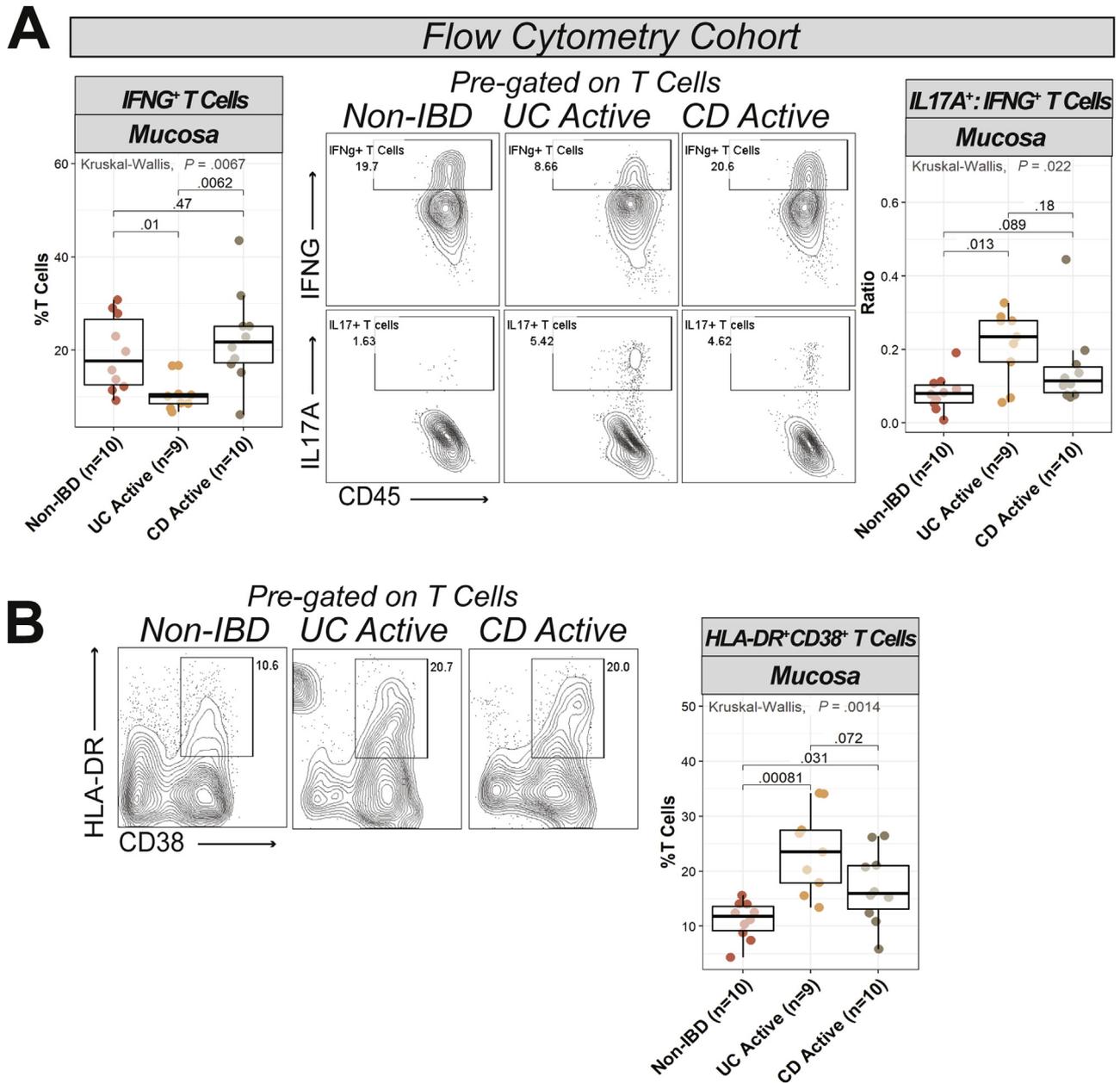
Marker	Tag	Clone
CD11C	115In	Bu15
CD66B	141Pr	G10F5
IL23A	161Dy	23cdcp
CD163	165Ho	GHI/61



Nancy Index	Mucosa*		Periphery†		Assignment
	UC	CD	UC	CD	
0	11	20	3	8	Inactive
1	3	5	2	8	Inactive
2	7	8	10	7	Active
3	6	2	3	2	Active
4	5	2	6	7	Active

*Nancy index derived from most colonic area (to CyTOF biopsies) sampled for clinical pathology
 †Nancy index derived from most inflamed intestinal area (in colon or terminal ileum) sampled for clinical pathology on endoscopy or surgery within 1 month of blood draw. 3 UC and 2 CD patients did not have intestinal pathology within 1 month of blood draw and are not featured in this table (their disease activity was assigned based on PUCAI or HBI scores, respectively. PUCAI >10 and HBI >3 deemed active).

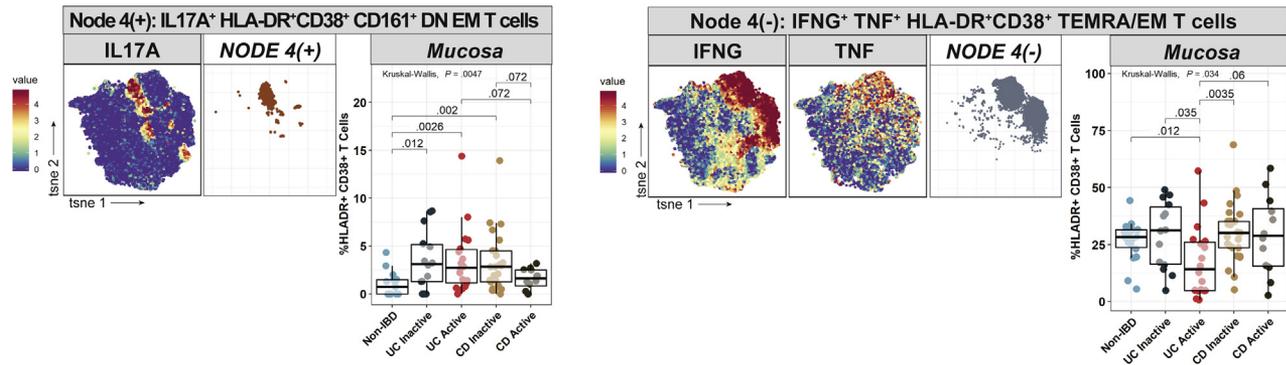




Supplementary Figure 2. T-cell findings in flow validation cohort and UC-specific HLA-DR⁺CD38⁺ cluster. (A) Flow cytometry data from an independent cohort demonstrating proportion of IFNG⁺ T cells (out of total T cells) and the ratio of IL17A to IFNG⁺ T cells across subject groups with representative biaxial plots shown. (B) Flow cytometry data demonstrating increase in HLA-DR⁺CD38⁺ T cells (out of total T cells) in active IBD mucosa of an independent cohort with representative biaxial gating plots.

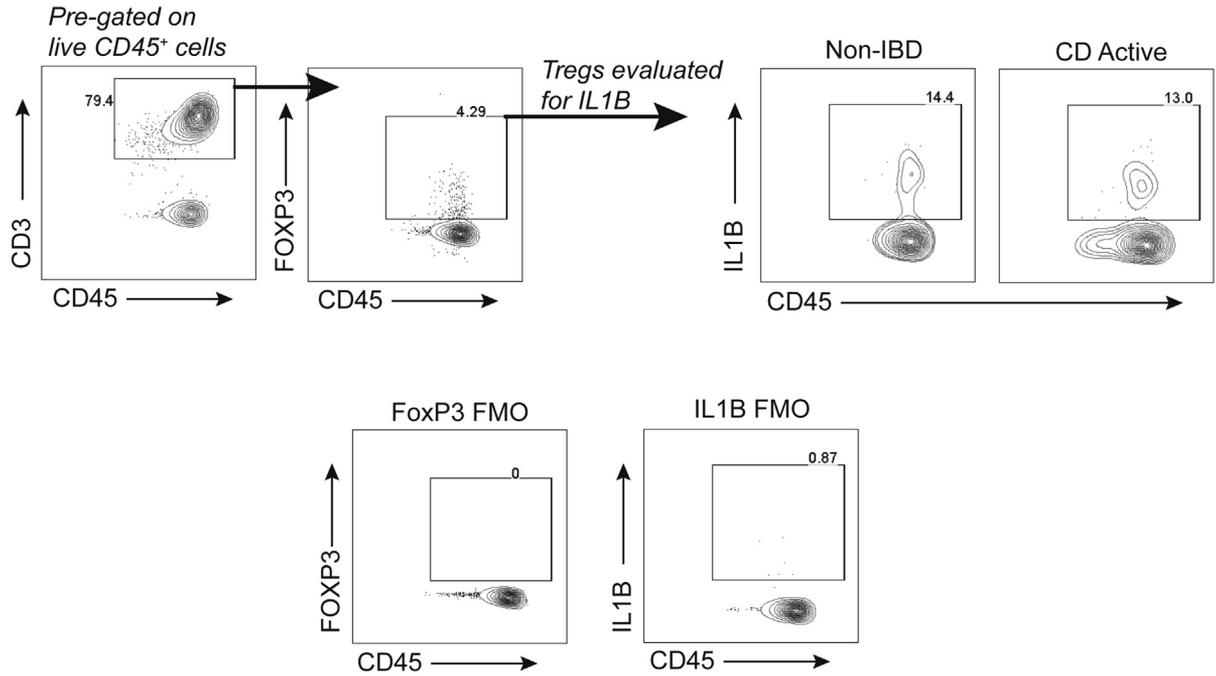
Supplementary Figure 1. CyTOF panels, Nancy Indices, and analysis of major immune subsets in IBD. (A) CyTOF antibody panels (A) and (B), including target marker, heavy metal tag, and clone. Samples were stained with one of either panel. Panels were designed to include the same 25 shared targets/channels to enable dimensionality reduction and clustering across all samples. Unique antibody targets of each panel also listed. (B) Representative images from Nancy Index 0–4 in colonic mucosa. (C) Comparison of endoscopic activity assessment by the endoscopist at time of specimen collection (for samples obtained during diagnostic colonoscopies) with the corresponding Nancy Index assigned by a blinded pathologist. Alongside, graphical representation of group assignments in mucosa and peripheral blood. (D) Box-and-whisker plots demonstrating proportion of manually gated immune cell subsets (total T cells, total B cells, total innate cells, T regulatory cells [Tregs]) out of all CD45⁺ cells (for Tregs, out of total T cells) in the mucosa and periphery as well as the B:T cell ratio across subject groups. (E) Percentage of total T cells, total B cells (out of total CD45⁺ cells), and B:T ratio from flow cytometry data of an independent cohort, with representative biaxial gating plots.

A

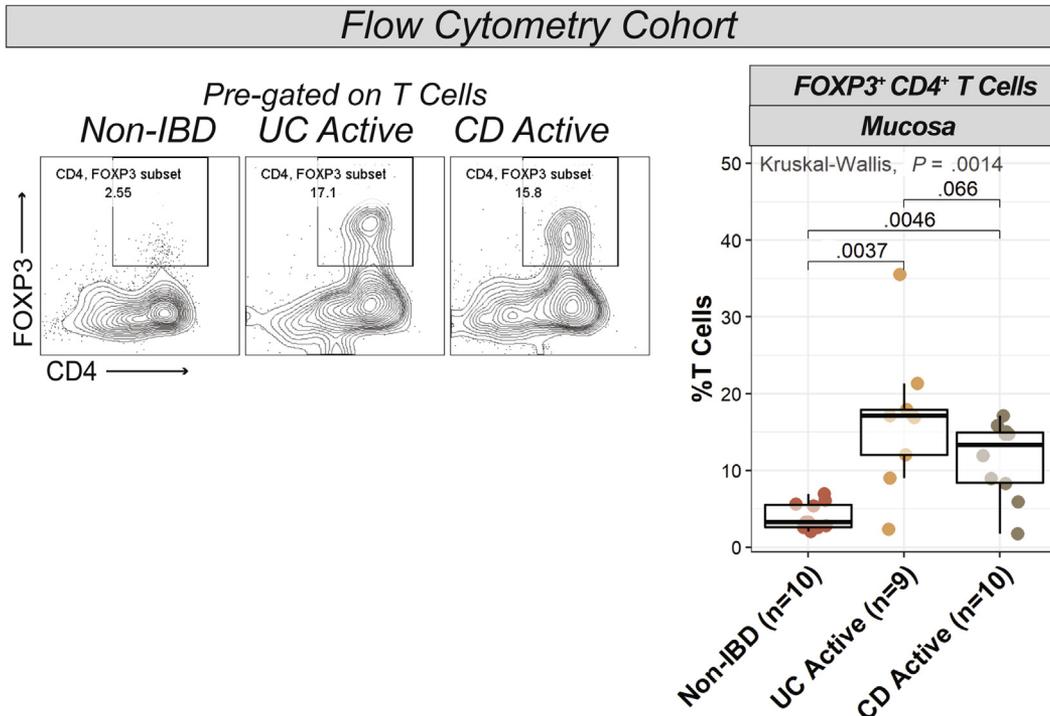


Supplementary Figure 3. Disease-specific cytokine signatures in HLA-DR⁺CD38⁺ T cells. (A) Node 4(+) and node 4(-) (derived from HLA-DR⁺CD38⁺ T cell clustering analysis in Figure 2E) t-SNE plots, abundance plots, and selected cytokine marker heatmaps for reference.

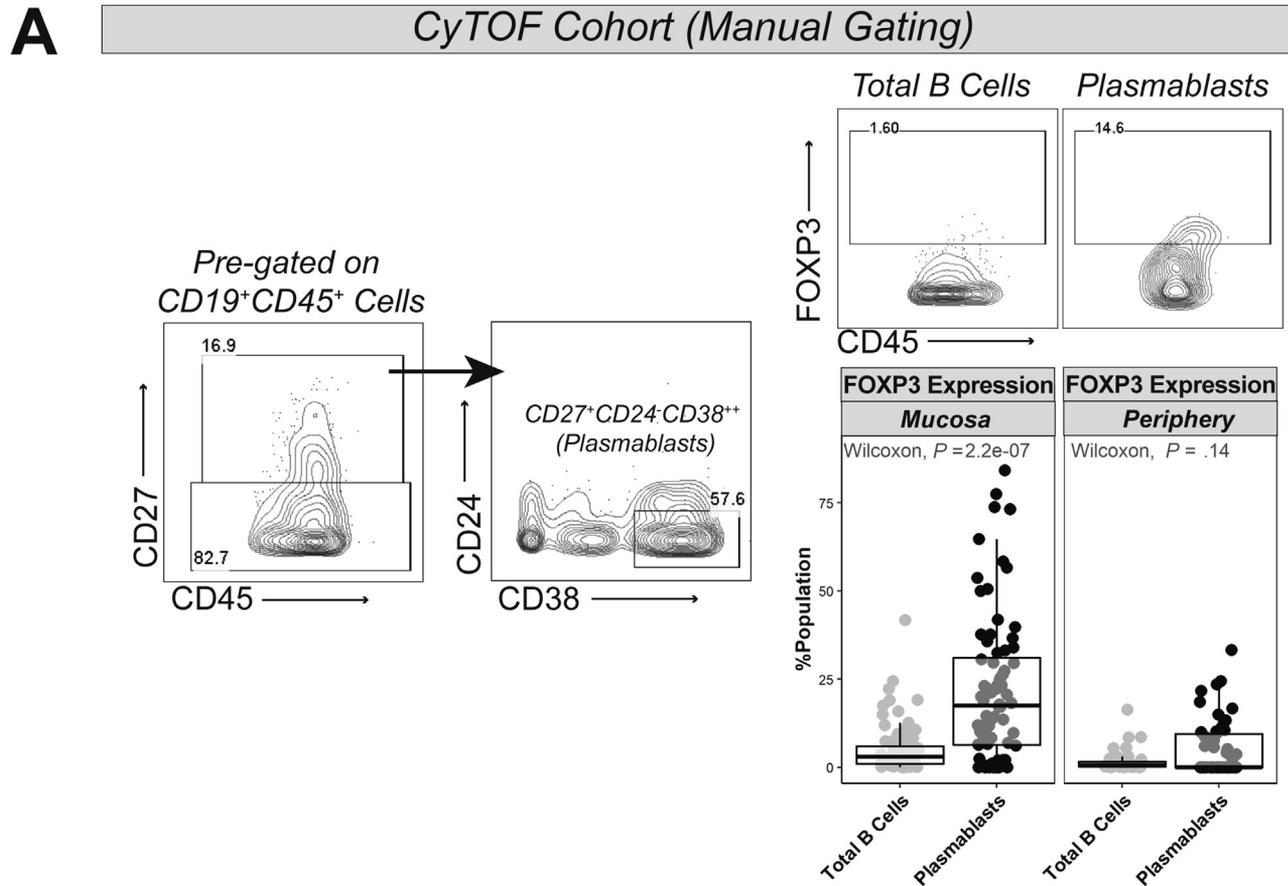
A



B



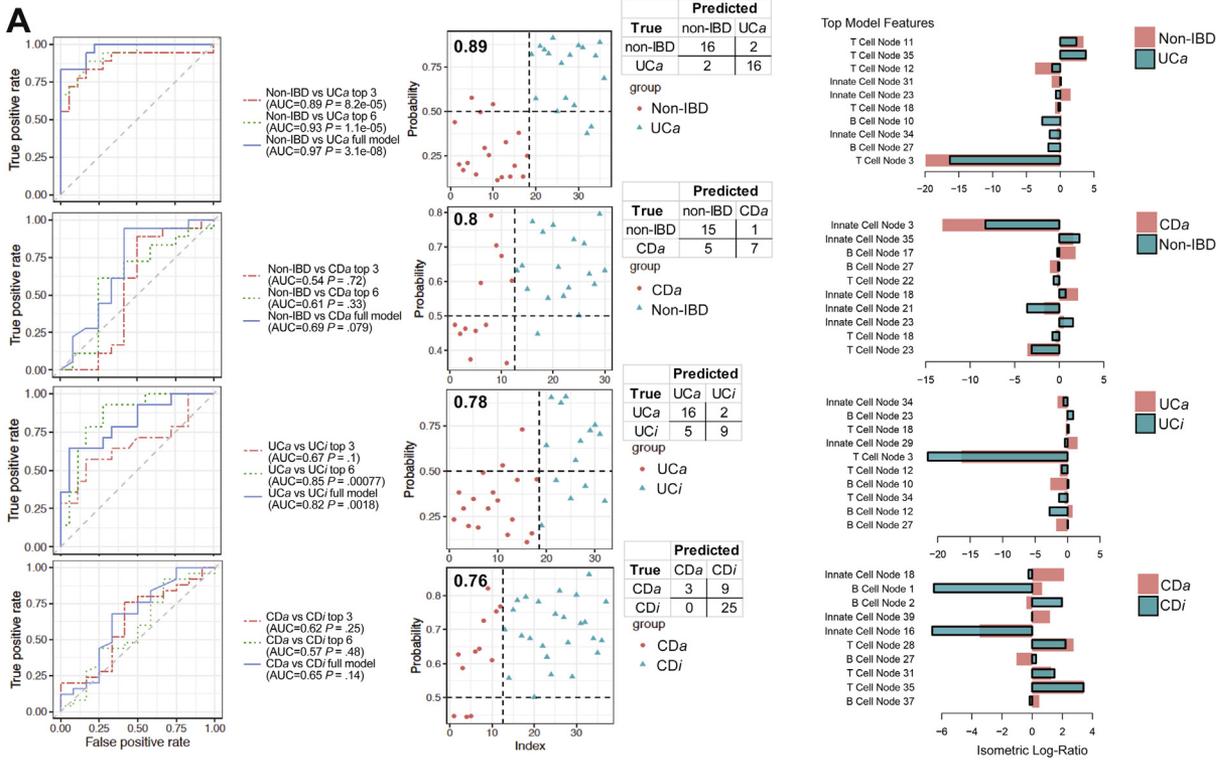
Supplementary Figure 4. Treg expansion in IBD mucosa and identification of IL1B⁺ Treg population with flow cytometry. (A) An IL1B⁺ Treg population was identified using flow cytometry in healthy control intestinal tissue as well as in IBD tissue (active CD patient included here). Gating strategy as well as fluorescence minus one (FMO) controls for FOXP3 and IL1B are included. (B) Tregs (gated as CD45⁺CD3⁺CD4⁺FOXP3⁺ cells) using flow cytometry in an independent cohort are shown to be expanded in active IBD mucosa; representative biaxial gating plots on left with *box-and-whisker abundance plot* on right.



Supplementary Figure 5. FOXP3 expression as a hallmark of mucosal plasmablasts. (A) Biaxial plots of CyTOF data demonstrating gating strategy to define plasmablasts (CD45⁺CD19⁺CD27⁺CD24⁻CD38⁺⁺); gating strategy to define FOXP3⁺ populations in both total B cells and plasmablasts; accompanied by plot comparison of proportion of FOXP3⁺ plasmablasts (out of total plasmablasts) vs proportion of FOXP3⁺ B cells (out of total B cells) in mucosa and in periphery.

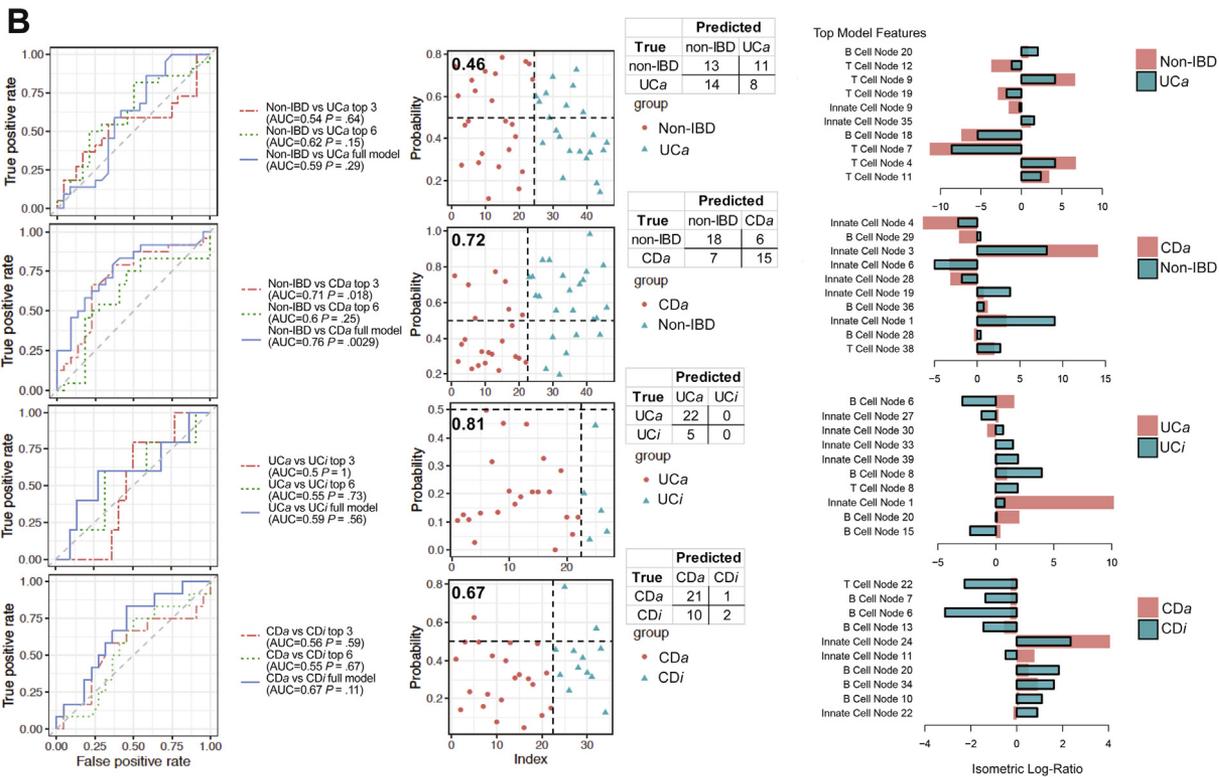
RF: Mucosa

A



RF: Periphery

B



Supplementary Figure 6. RF predictive modeling of pairwise subject group comparisons. RF classifier was used to generate models differentiating non-IBD:UCa, non-IBD:CDa, UCa:UCi, and CDa:CDi in both the mucosa (A) and in the periphery (B). *Leftmost panels* depict receiver operator characteristic curves (ROC) of models using the top 3, top 6, or full model of input features. Each curve is accompanied with AUC and the associated *P* value. *Middle panels* show probability plots with associated confusion matrices of prediction accuracy of the full model (prediction accuracy listed in **bold** in the *upper left corner* of the probability plots). *Rightmost plots* depict the isometric log-ratio of the top 10 features (nodes derived from the automated clustering analysis of total T, total B, and total innate immune cells (Figures 1C, 4B, and 5B) in descending order of importance. The *barplots* are *color-coded* by subject group comparison. The more “negative” the isometric log ratio of a particular node, the more expanded node X(-)>node X(+) branch abundance. The more “positive” the isometric log ratio of a particular node, the more expanded node X(+)>node X(-) branch abundance.