TITLE

A Treatment-Naïve Cellular Atlas of Pediatric Crohn's Disease Predicts Disease Severity and Therapeutic Response

AUTHORS

Hengqi Betty Zheng[1,*], Benjamin A. Doran[2,3,4,5,6,7,*], Kyle Kimler[2,3,4,5,6,7,*], Alison Yu[2,*], Victor Tkachev[2], Veronika Niederlova[3,5,6,8], Kayla Cribbin[2], Ryan Fleming[2], Brandi Bratrude[2], Kayla Betz[2], Lorenzo Cagnin[2], Connor McGuckin[2], Paula Keskula[2], Alexandre Albanese[2], Maria Sacta[3], Joshua de Sousa Casal[3,6,9], Faith Taliaferro[3,6], Madeline Ford[1], Lusine Ambartsumyan[1], David L. Suskind[1], Dale Lee[1], Gail Deutsch[10], Xuemei Deng[10], Lauren V. Collen[3], Vanessa Mitsialis[3,11], Scott B. Snapper[3,11,12], Ghassan Wahbeh[1], Alex K. Shalek[4,5,6,7,9,13,14,#], Jose Ordovas-Montanes[3,6,9,14,#], Leslie S. Kean[2,15,#]

AFFILIATIONS

1. Division of Gastroenterology and Hepatology, Seattle Children's Hospital and University of Washington, Seattle, WA 98105, USA
2. Division of Pediatric Hematology/Oncology, Boston Children's Hospital, Boston, MA 02115, USA
3. Division of Gastroenterology, Hepatology, and Nutrition, Boston Children's Hospital, Boston, MA 02115, USA
4. Institute for Medical Engineering & Science, Massachusetts Institute of Technology, Cambridge, MA 02142, USA
5. Ragon Institute of MGH, MIT, and Harvard, Cambridge, MA 02139, USA
6. Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
7. Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
8. Laboratory of Adaptive Immunity, Institute of Molecular Genetics of the Czech Academy of Sciences, Prague, Czech Republic
9. Program in Immunology, Harvard Medical School, Boston, MA 02115, USA
10. Department of Pathology, Seattle Children's Hospital and University of Washington, Seattle, WA 98145, USA
11. Brigham and Women's Hospital Department of Medicine, Division of Gastroenterology, Hepatology, and Endoscopy, Boston, MA 02115, USA
12. Harvard Medical School, Boston, MA 02115, USA
13. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
14. Harvard Stem Cell Institute, Cambridge, MA 02138, USA
15. Dana Farber Cancer Institute, Boston, MA 02215, USA

*these authors contributed equally to the work
#these senior authors contributed equally to the work

Correspondence to Leslie S. Kean leslie_kean@dfci.harvard.edu, Jose Ordovas-Montanes jose.ordovas-montanes@childrens.harvard.edu, Alex K. Shalek shalek@mit.edu

Abstract

Crohn's disease is an inflammatory bowel disease (IBD) which most often presents with patchy lesions in the terminal ileum and colon and requires complex clinical care. Recent advances in the targeting of cytokines and leukocyte migration have greatly advanced treatment options, but most patients still relapse and inevitably progress. Although single-cell approaches are transforming our ability to understand the barrier tissue biology of inflammatory disease, comprehensive single-cell RNA-sequencing (scRNA-seq) atlases of IBD to date have largely sampled pre-treated patients with established disease. This has limited our understanding of which cell types, subsets, and states at diagnosis are predictive of disease severity and response to treatment. Here, through a combined clinical, flow cytometric, and scRNA-seq study, we profile diagnostic human biopsies from the terminal ileum of treatment-naïve pediatric patients with Crohn's disease (pediCD; n=14) and from non-inflamed pediatric controls with functional gastrointestinal disorders (FGID; n=13). To fully resolve and annotate epithelial, stromal, and immune cell states among the 201,883 single-cell transcriptomes, we develop and deploy a principled and unbiased tiered clustering approach, ARBOL, yielding 138 FGID and 305 pediCD end cell clusters. Notably, through both flow cytometry and scRNA-seq, we observe that at the level of broad cell types, treatment-naïve pediCD is not readily distinguishable from FGID in cellular composition. However, by integrating high-resolution scRNA-seq analysis, we identify significant differences in cell states that arise during pediCD relative to FGID. Furthermore, by closely linking our scRNA-seq analysis with clinical meta-data, we resolve a vector of lymphoid, myeloid, and epithelial cell states in treatment-naïve samples which can distinguish patients with less severe disease (those not on anti-TNF therapies (NOA)), from those with more severe disease at presentation who require anti-TNF therapies. Moreover, this vector was also able to distinguish those patients that achieve a full response (FR) to anti-TNF blockade from those more treatment-resistant patients who only achieve a partial response (PR). Our study jointly leverages a treatment-naïve cohort, high-resolution principled scRNA-seq data analysis, and clinical outcomes to understand which baseline cell states may predict inflammatory disease trajectory.

Introduction

Inflammatory bowel diseases (IBDs) arise when homeostatic mechanisms regulating gastrointestinal (GI) tract tissue integrity, nutrient absorption, and protective immunity are replaced by pathogenic inflammation (Baumgart and Sandborn, 2012; Chang, 2020; Corridoni et al., 2020a; Friedrich et al., 2019; Graham and Xavier, 2020; Selin et al., 2021). The initiating triggers are not fully known, but host genetics and the microbiome are being increasingly appreciated to play important, and in some cases causal roles in the IBDs (Chang, 2020; Cohen et al., 2019; Franzosa et al., 2019; Jain et al., 2021; Limon et al., 2019). Though symptoms are widespread and overlapping, the endoscopic inflammation in the IBDs is often anatomically restricted: ulcerative colitis (UC) manifests primarily as an superficial inflammatory response restricted to the colon, while Crohn's disease (CD) presents predominantly in the terminal ileum and the proximal colon, though lesions may develop anywhere along the gastrointestinal tract (Baumgart and Sandborn, 2012; Chang, 2020; Kobayashi et al., 2020; Roda et al., 2020). Among the IBDs, pediatric-onset Crohn's disease (pediCD) is particularly common (25% of all IBD cases, 60-70 % of pediatric IBD) and is a debilitating form due to its early presentation, impact on the terminal ileum and proximal colon, and the lack of disease-specific therapies developed with children in mind (Hyams et al., 1991; Ruemmele et al., 2014; Sýkora et al., 2018; Turner et al., 2012; Ye et al., 2020). In contrast to pediCD, a group of pediatric disorders termed functional gastrointestinal disorders (FGIDs) include GI symptoms but lack laboratory markers, endoscopic findings, and histologic evidence associated with inflammation (Black et al., 2020; Hyams et al., 2016; McOmber and Shulman, 2008; Santucci et al., 2020). FGID thus represents a critical non-inflamed control cohort with which to contextualize the inflammation observed in pediCD.

The current standard of care for pediCD (as with adult CD) is tailored to the patient's disease location, clinical behavior and severity, though use of prednisone, immunomodulators, as well as biologics including anti-TNF-α monoclonal antibodies, are common (Hyams et al., 1991; Ruemmele et al., 2014; Turner et al., 2012). While targeting TNF is common across many autoimmune and inflammatory conditions, it is not successful in all patients, and many go on to develop anti-TNF-refractory disease. It is of tremendous importance for the field to precisely understand and characterize for which patients anti-TNF therapy is not necessary, in which patients it may succeed in controlling disease, and which patients will be refractory to treatment. Several ideas based on individual immunogenicity and pharmacokinetics have been proposed to explain TNF-refractory disease, including gender (M>F), low albumin levels, high BMI, and high

baseline C-Reactive Protein (CRP) (Atreya et al., 2020; Digby-Bell et al., 2020). However, no single identifiable clinical or biochemical biomarker reliably predicts disease response versus resistance to anti-TNF antibodies suggesting a more complex etiology (Stevens et al., 2018).

The primary cellular lineages sampled from intestinal biopsies of CD patients (from the terminal ileum or colon) represent both the epithelium and lamina propria, and include epithelial cells, stromal cells, hematopoietic cells and neuronal processes whose cell bodies are present outside of these regions (Buisine et al., 2001; Leeb et al., 2003; Leonard et al., 1995; Lilja et al., 2000; Müller et al., 1998; Souza et al., 1999; Stappenbeck and McGovern, 2017; Takayama et al., 2010). Alterations to all cellular lineages have been implicated in CD (Furey et al., 2019; Martin et al., 2019). Histologically, CD is characterized by a granulomatous inflammation, and noted alterations in almost every leukocyte cell type studied including an increase of cytotoxic lymphocytes, potential but equivocal alterations in γδ T cells, increases in mast cells and their production of TNF, activation and shifts in antibody isotypes towards IgM and IgG from B cells and plasma cells, and cytokine production by macrophages (Catalan-Serra et al., 2017; Lilja et al., 2000; Meijer et al., 1979; Mitsialis et al., 2020; Müller et al., 1998; Sieber et al., 1984; Takayama et al., 2010). In the stromal compartment, there is evidence for enhanced vascularization and increased expression of ICAM-1 and MAdCAM-1 by vascular beds, substantial remodeling of collecting lymphatics, and altered migratory potential of fibroblasts (Leeb et al., 2003; Souza et al., 1999). Epithelial barrier dysfunction has also been noted, including alterations to mucus production, microvilli, and Paneth cell dysfunction (Buisine et al., 2001; Stappenbeck and McGovern, 2017). Collectively previous studies have identified that all cell types may be meaningfully altered during CD, and highlight the important need to comprehensively understand the concerted cellular changes that define CD. Importantly, which changes are predictive of more severe disease or treatment resistance in pediCD remain largely unknown.

Massively parallel single-cell RNA-sequencing (scRNA-seq) is enhancing our ability to comprehensively map and resolve the cell types, subsets, and states present during health and disease. This has been particularly evident in the elucidation of novel human cell subsets and states within epithelial, stromal, immune, and neuronal cell lineages. Recent work has generated cellular atlases of previously treated (pre-treated) adult CD and UC, though a comprehensive single-cell atlas for untreated pediatric disease with follow-up of patient outcomes has yet to be reported for IBD (Corridoni et al., 2020a, 2020b; Drokhlyansky et al., 2019; Elmentaite et al., 2020;

66 Huang et al., 2019; Kinchen et al., 2018; Martin et al., 2019; Parikh et al., 2019; Smillie et al.,
67 2019). The potential impact of scRNA-seq on our understanding of IBD is evidenced by studies
68 of adult UC, which have identified potential functional roles for poorly understood colonic cell
69 subsets, such as the BEST4+ enterocyte, and identified pathological alterations in UC pinch
70 biopsies compared to healthy controls, including an expansion of microfold-like cells,
71 *IL13RA2+IL11+* inflammatory fibroblasts, *CD4+CD8+IL17A+* T cells and *CD8+GZMK+* T cells
72 (Smillie et al., 2019). A single-cell study of pre-treated adult CD patients comparing non-inflamed
73 and inflamed tissue from surgically-resected bowel found IgG+ plasma cells, inflammatory
74 mononuclear phagocytes, activated T cells and stromal cells comprising the "GIMATS"
75 pathogenic cellular module (Martin et al., 2019). This module was used to derive a gene signature
76 associated with resistance to anti-TNF therapy in a distinct cohort profiled by bulk RNA-seq. Very
77 recent work has profiled how fetal transcription factors are reactivated in Crohn's disease
78 epithelium (Elmentaite et al., 2020). Together, these and other studies have demonstrated the
79 power of scRNA-seq to nominate individual and collective cell states that are associated with
80 disease, and have also underscored the unmet need to apply these techniques to untreated
81 disease and associate them with disease severity in order to more specifically identify
82 pathognomonic and prognostic cell states.
83
84 Most comprehensive scRNA-seq atlases of inflammatory disease conditions consist of patients
85 being treated with a variety of agents, and for which the biopsies included often reflect a partial
86 treatment-refractory state to combinations of antibiotics, corticosteroids, immunomodulators, and
87 biologics including anti-TNF monoclonal antibodies. A treatment-naïve single-cell atlas in an
88 inflammatory disease condition linking observed baseline cell clusters with disease trajectory and
89 treatment outcomes has yet to be reported. In order to address this unmet need in pediCD, we
90 created the prospective PREDICT study (Clinicaltrials.gov #NCT03369353) to help identify,
91 profile, and understand pediatric IBD and FGID controls. Here, we present detailed diagnostic
92 data from the first cohort of 27 patients enrolled on PREDICT, including 14 pediCD and 13 FGID
93 patients, together with flow cytometric and scRNA-seq studies of the cellular composition of the
94 terminal ileum (**Figure 1**). Furthermore, through thorough, prospective annotation of clinical
95 metadata and detailed longitudinal follow-up, we stratify our pediCD cohort by clinically-guided
96 therapeutic decisions separating patients treated with anti-TNF mAbs versus those with biopsy-
97 proven pediCD, but for whom clinical symptoms were sufficiently mild that the treating physician
98 did not prescribe anti-TNF agents (this cohort is termed "Not On Anti-TNF" or "NOA"). Importantly,
99 we were also able to separate the cohort of patients treated with anti-TNF agents into a sub-

100 cohort of those who achieved a full response (FR) to this therapy, versus those who achieved
101 only a partial response (PR). Critically, because PREDICT enrolled patients prior to their
102 diagnostic endoscopy, we were able to relate these clinical outcomes to the patients' cell states
103 at diagnosis. We contextualize our findings in pediCD relative to a cohort of 13 FGID patients,
104 which provides an age-matched comparator cohort with clinical GI symptoms, but non-
105 inflammatory disease proven by endoscopy and histologic examination.
106
107 Several analytical approaches have been developed to enable the generation and interrogation
108 of clusters during the curation of single-cell transcriptomic atlases (Hie et al., 2020). One such
109 method, sub-clustering of broad clusters, has proven to be a powerful tool for isolating highly
110 specific axes of variation that are obscured by analyses whose principal axes of variation are
111 broad cell types (La Manno et al., 2021; Tasic et al., 2018; Zeisel et al., 2018). However, while
112 sub-clustering analysis is a powerful tool allowing access to the hierarchy of cell states, this
113 method is manually intensive and there is little consensus, control, or standard in clustering
114 parameters or annotation methods. To address this issue, we developed a principled, modular,
115 automated sub-clustering routine made possible by application of parameter scanning methods
116 (Rousseeuw, 1987; Shekhar et al., 2016). We developed this tool, ARBOL, of which iterative
117 tiered clustering (ITC) is a key component, in R, integrating with Seurat functions, to make it
118 accessible and easily incorporated into common workflows and have curated a GitHub repository
119 with illustrative vignettes. Here, we use ARBOL to standardize fine-grained cell state discovery
120 by the creation and cultivation of a tree of cell states, followed by the generation of automated cell
121 names to aid in the annotation of end clusters by unique and descriptive genes.
122
123 Together we present two cellular atlases for pediatric GI disease, consisting of 94,451 cells for
124 FGID and 107,432 for pediCD. We provide key gene-list resources for further studies, identify
125 correspondence between disease states, and nominate a vector of lymphoid, myeloid and
126 epithelial cell states which predicts disease severity and treatment outcomes. This cellular vector
127 correlates strongly with both the clinical presentation of pediCD severity, and to the distinction
128 between anti-TNF full or partial response. The significant changes in cell composition associated
129 with disease severity were increases of proliferating T cells, cytotoxic NK cells, specific
130 monocytes/macrophages, and plasmacytoid dendritic cells (pDCs) accompanied by decreases of
131 metabolically-specialized epithelial cell subsets. We further validate this vector in two bulk RNA-
132 seq treatment-naïve IBD cohorts.
133

134 # Results

135 **Study cohort outcomes**

136 The PREDICT study prospectively enrolled treatment-naïve, previously undiagnosed pediatric
137 patients with GI complaints necessitating diagnostic endoscopy. The current analysis focuses on
138 patients enrolled in the first year of the study, during which time 14 patients with pediCD and 13
139 patients with FGID were enrolled and had adequate ileal samples for single cell analysis (**Figure**
140 **1; Supplemental Figure 1**). Following their initial diagnosis, patients with pediCD were followed
141 clinically for up to 3 years. Patients with FGID were followed up as needed in subspecialty/GI
142 clinic. The median time from diagnosis for the pediCD and FGID cohorts as of December 1, 2020
143 (time of database lock) was 32.5 and 31 months, respectively. Of the pediCD and FGID patients
144 analyzed, the median age at diagnosis was 12.5 years and 16 years respectively ($p = 0.095$), with
145 no significant differences in gender (**Figure 1b; Table 1**)**.** Patient weight, height, and BMI z-scores
146 were not significantly different between pediCD and FGID (**Figure 1b; Table 1**); however, in
147 addition to the diagnostic differences on histologic analysis, several key clinical laboratory values,
148 including C-reactive protein (CRP), Erythrocyte Sedimentation Rate (ESR), hemoglobin
149 concentration, albumin concentration, and platelet count were significantly different between
150 pediCD and FGID (**Figure 1c,d; Table 1**).

151

152 **Treatment with anti-TNF agents and response to therapy**

153 Patients with pediCD were initially divided into two cohorts. Those with milder disease
154 characteristics ($n = 4$) as determined by their treating physician, were not put on anti-TNF
155 therapies, and are noted as NOA. For patients with more severe disease ($n = 10$), anti-TNF
156 therapy (with either infliximab or adalimumab, **Table 2**) was initiated within 90 days of diagnostic
157 endoscopy. All pediCD patients were followed prospectively and categorized as FR ($n = 5$) or PR
158 ($n = 5$) to anti-TNF therapy based on the following criteria: FR was defined as clinical symptom
159 control and biochemical response (measuring CRP, ESR, albumin, and complete blood counts
160 (CBC)), and with a weighted Pediatric Crohn's Disease Activity Index (PCDAI) score of <12.5 on
161 maintenance anti-TNF therapy with no dose adjustments required (Cappello and Morreale, 2016;
162 Hyams et al., 1991; Sandborn, 2014; Turner et al., 2012, 2017). PR to anti-TNF therapy was
163 defined as a lack of full clinical symptom control as determined by the treating physician or lack
164 of full biochemical response, with documented escalation of anti-TNF therapy or addition of other
165 agents (**Figure 1e**; NB: patients in our cohort were dose escalated because of clinical symptoms).
166 Medication timelines and clinical laboratory data through 2 years of follow-up for all pediCD

167 patients is shown in **Supplemental Figure 1.** The designation of FR or PR was made at 2 years
168 of follow-up for all pediCD patients.
169
170 **Flow cytometry of the terminal ileum reveals minimal changes in leukocyte subsets in**
171 **FGID vs. pediCD, and no significant differences across the pediCD spectrum**
172 We collected terminal ileum biopsies from 14 pediCD patients and from 13 uninflamed FGID
173 patients, and prepared single-cell suspensions for flow cytometry and scRNA-seq. Biopsies from
174 pediCD were from actively-inflamed areas adjacent to ulcerations. Biopsies from FGID were from
175 non-inflamed terminal ileum. The epithelium was first separated from the lamina propria before
176 enzymatic dissociation, and flow cytometric analysis was performed on the viable single-cell
177 fraction, which recovered predominantly hematopoietic cells with some remnant epithelial cells
178 (<20% of all cells), likely representing those in deeper crypt regions (**Figure 2; Supplemental**
179 **Figure 2**). We utilized two flow cytometry panels, allowing us to resolve the principal lymphoid
180 (CD4 or CD8 T cells, NK cells, B cells, innate lymphoid cells, γδ T cells, CD8αα+ IELs, pDCs) and
181 myeloid (monocytes, granulocytes, HLA-DR+ mononuclear phagocyte) cell subsets
182 (**Supplemental Figure 2, Supplemental Table 1**). From these panels, which generated 32 gates
183 identifying cell lineages, types and subsets, only HLA-DR+ macrophages/DCs and pDCs were
184 significantly increased in pediCD relative to FGID (**Figure 2d**). We also analyzed within pediCD,
185 comparing the baseline samples of 4 NOA, 5 FR and 5 PR patients, and noted no significant
186 differences between NOA and patients on anti-TNF, or between FRs and PRs to anti-TNF (**Figure**
187 **2e**). Together, this suggests that despite the substantial endoscopic, histologic, and clinical
188 parameters that distinguish FGID and pediCD (**Figure 1**), the broad single-cell type composition
189 of the terminal ileum appears minimally altered in pediCD save for an increased frequency of
190 pDCs and HLA-DR+ mononuclear phagocytes.
191
192 **Traditional joint scRNA-seq clustering of FGID and pediCD patients**
193 In addition to flow cytometry, we performed droplet-based scRNA-seq on cell suspensions from
194 the 14 pediCD/13 FGID patient cohort using the 10X Genomics V2 3' platform (**Figure 1**). The
195 analyzed cell suspensions were derived from lamina propria preparations, which our flow
196 cytometry data suggested would be composed primarily of CD45+ leukocytes, alongside a small
197 fraction of epithelial cells and stromal/vascular cells. Deconstructing these tissues into their
198 component cells provided us with the ability to identify some of the corresponding cell types (e.g.
199 T or B cell) and subsets (CD8αα+ IEL or CD4+ T cell) to those we identified by flow cytometry
200 data but importantly also enabled us to: 1. characterize these major cell types and subsets without

201 needing to pre-select markers, and 2. gain substantially enhanced resolution into the cell states
202 (i.e. gene expression programs) within these types and subsets.

203 Following library preparation and sequencing, we derived a unified cells-by-genes expression
204 matrix from the 27 samples, containing digital gene expression values for all cells passing quality
205 thresholds (n=254,911 cells; **Supplemental Figure 3; Supplemental Tables 2 and 3; Methods**).
206 We then performed dimensionality reduction and graph-based clustering, noting that despite no
207 computational integration methods being used, FGID and pediCD were highly similar to each
208 other when visualized on a uniform manifold approximation and projection (UMAP) plot
209 (**Supplemental Figure 4a-c**). We recovered the following cell types from both patient groups:
210 epithelial cells, T cells, B cells, plasma cells, glial cells, endothelial cells, myeloid cells, mast cells,
211 fibroblasts, and a proliferating cell cluster. We noted that the fractional composition amongst all
212 cells of T cells, B cells, and myeloid cells was not significantly different between FGID and pediCD,
213 similar to the flow cytometric data, and this was also the case for endothelial, fibroblasts, glial,
214 mast and plasma cells, which were not measured through flow cytometry (**Supplemental Figure**
215 **4d**). This provided validation and extension of our flow cytometry data documenting that the broad
216 cell type composition of FGID and pediCD is not significantly altered, despite highly distinct
217 endoscopic and histologic diseases. Based on this joint clustering and annotation of top-level cell
218 types, we then performed differential expression testing identifying significant up- and down-
219 regulated genes across all cell types (**Supplemental Figure 4e; Supplemental Table 4**). Within
220 myeloid cells we identified some of the most significantly upregulated genes in pediCD versus
221 FGID to be *CXCL9* and *CXCL10*, canonical IFNγ-stimulated genes, and *S100A8* and *S100A9*
222 which form the biomarker fecal calprotectin (**Supplemental Figure 4f**)(Leach et al., 2007; Ziegler
223 et al., 2021). Within epithelial cells we identified that *APOA1* and *APOA4* were amongst the most
224 significantly downregulated genes, and correspondingly *REG1B*, *SPINK4* and *REG4A* were
225 amongst the most significantly upregulated indicating tradeoffs between lipid metabolism and host
226 defense (**Supplemental Figure 4f**) (Haberman et al., 2014). In T cells, we noted that the cytotoxic
227 genes *GNLY* and *GZMA* were amongst the most significantly upregulated, with almost no genes
228 downregulated (**Supplemental Figure 4f**).

229 We then systematically re-clustered each broad cell type, identifying increasing cellular
230 heterogeneity. Given that we detected changes in the frequency of HLA-DR+
231 macrophages/dendritic cells and pDCs between pediCD and FGID by flow cytometry, we initially
232 focused on the myeloid cell type sub-clustering, containing dendritic cells, macrophages,
233 monocytes, and pDCs (**Supplemental Figure 4g)**. Working within this analysis paradigm

234 revealed that a traditional clustering approach had difficulty identifying the boundaries of clusters,
235 and whether a cluster composed primarily of pediCD rather than FGID cells represented a unique
236 cell subset, or a cell state overlaid onto a core cell subset gene expression program
237 (**Supplemental Figure 4g, Methods**). These distinctions would influence whether a comparison
238 would primarily focus on differential expression testing or differential composition testing between
239 the two clinical cohorts. It also raised the possibility that this joint clustering approach, informed
240 by the inclusion of both FGID and pediCD cell types, subsets and states could muddle some of
241 the unique biology of FGID and pediCD. For instance, this analytical approach could lead to hybrid
242 clusters, informed by cells from both FGID and pediCD and correspondingly critical gene-
243 reference lists for each cluster, that may not accurately represent disease-specific cell types,
244 subsets, or states.

245 In order to approach this challenge from a more principled direction, we made five key changes
246 to our analytical workflow, which we jointly refer to as ARBOL (https://github.com/jo-m-
247 lab/ARBOL): 1. We proceeded to analyze FGID and pediCD samples *separately* to define
248 corresponding cell type, subset, and state clusters and markers, 2.We implemented an automated
249 ITC approach to optimize the silhouette score at each tier of iterative sub-clustering and stop
250 when a specific granularity was reached (**Supplemental Figure 5a,b; Methods**), 3. We
251 systematically generated descriptive names for cell types and subsets together with differentiating
252 marker genes, 4. We accounted for the number and diversity of patients which compose each
253 cluster using Simpson's Index of Diversity, and 5. We generated and optimized a Random Forest
254 classifier to identify correspondence between the resultant FGID and pediCD atlases
255 (**Supplemental Figure 5c; Methods**) (Simpson, 1949). Using this approach, each tier of analysis
256 is typically under-clustered relative to traditional empirical analyses, but the automation proceeds
257 through several more tiers (typically 6 to 7) until stop conditions (e.g. cell numbers and
258 differentially expressed genes; **Methods**) are met. We then inspected all outputs (182 FGID and
259 425 pediCD clusters) and provided descriptive cell cluster names independently for FGID and
260 pediCD (**Supplemental Figures 5b and 6**). We also focused at this stage on flagging putative
261 doublet clusters or clusters where the majority of differentially expressed genes which triggered
262 further clustering consist of known technical confounders in scRNA-seq data (e.g. mitochondrial,
263 ribosomal, and spillover genes from cells with high secretory capacity) yielding a final number of
264 118 FGID and 305 pediCD clusters (**Supplemental Figure 5b**). We note this clustering method
265 represents a data-driven approach, though it may not always reflect a cellular program or
266 transcriptional module of known biological significance.

267 We then hierarchically clustered all end cell state clusters to generate the final dendrograms for
268 FGID and pediCD, and performed 1 vs. rest within-Tier 1 clusters (i.e. broad cell types) differential
269 expression to provide systematic names for cells based on their cell type classification and two
270 genes (**Figures 3 and 4; Methods**). As several cell types contained readily identifiable and
271 meaningful cell subsets, we utilized curation of literature-based markers to provide further
272 guidance within each cell type (Blériot et al., 2020; Cherrier et al., 2018; Dutertre et al., 2019;
273 Guilliams et al., 2018; Robinette and Colonna, 2016). For example, within Tier 1 T cells, we could
274 identify T cells, NK cells and ILCs; within Tier 1 myeloid cells, monocytes, cDC1, cDC2,
275 macrophages and pDCs; within Tier 1 B cells, germinal center, germinal center dark zone and
276 light zone cells; within Tier 1 endothelial cells, arterioles, capillaries, lymphatics, mural cells and
277 venules; and so forth for other cell types. To illustrate this process for one cluster, upon automated
278 hierarchical tiered clustering of T cells, we identified a cluster that was Tier 0: pediCD, Tier 1: T
279 cells, Tier 2: cytotoxic, Tier 3: IEL_FCER1G_NKG7_TYROBP_CD160_AREG. Upon inspection
280 of CD3 genes (*CD247, CD3D,* etc.), TCR genes (*TRAC, TRBC1*, etc.), and NK cell genes
281 (*NCAM1, NCR1*), it became readily apparent these cells were NK cells (**Supplemental Figure**
282 **6**).

283 We provide complete gene lists for cell types (1 vs. rest across all cells), subsets (1 v. rest across
284 all cells), and states (1 vs. rest within-Tier 1 cell-type) in **Supplemental Tables 5-10**. To select
285 marker genes for naming in a data driven manner we used 1 vs. rest within-cell-type differential
286 expression (**Supplemental Tables 7 and 10**; Wilcoxon, Bonferroni adjusted $p<0.05$). To account
287 for genes that might be highly expressed in just a few cells we ranked the marker genes by a
288 score combining their significance, the fold change in expression, and fold change of percent
289 gene positive cells in the subset versus the percent of gene positive cells outside the subset. The
290 collected metrics were multiplied together to provide a single score by which the genes were
291 ranked: **(-log(sig+1) * avg_logFC * (pct.in / pct.out))**. For most subsets we selected the top 2 of
292 these marker genes. For T/NK/ILC cells and myeloid cells we occasionally chose a slightly lower
293 ranking gene from the top 10 if it was well supported and recognized by the literature. Using this
294 ranking system, we identified *CCL3* and *CD160* as two genes significantly enriched in one NK
295 cluster (adj. p-value = 0, expression within-cluster >40% cells positive and in other Tier 1 T cells
296 <6%). This resulted in a final name for this cluster of CD.NK.CCL3.CD160. We repeated this
297 process for all FGID (183 end clusters) and pediCD (426 end clusters) within Tier 1 B cell,
298 endothelial, epithelial, fibroblast, plasma cell, myeloid cell, mast cell, and T cell identified clusters,
299 and provide systematically generated names for all (**Supplemental Tables 11 and 12**).

300 Using this analytical workflow, we present comprehensive cellular atlases of FGID (**Figure 3**) and
301 pediCD (**Figure 4**), nominate cell states associated with disease severity and treatment outcomes
302 in pediCD (**Figure 5**), and identify correspondence between the two (**Figure 6**). We focus on
303 pediCD, and those cell states which distinguish between disease severity (NOA vs. PRs/FRs)
304 and baseline gene expression differences in anti-TNF treatment response (FRs vs. PRs).

305

306 **Comprehensive atlas of non-inflammatory FGID**

307 From 13 FGID patients, we recovered 12 Tier 1 clusters which we display on a t-stochastic
308 neighbor embedding (t-SNE) plot colored by cluster identity containing 99,488 cells (**Figure 3a;**
309 **Supplemental Figure 5b**). These Tier 1 clusters represent the main cell types found in the lamina
310 propria and remnant epithelium of an ileal biopsy. Inspecting each individual patient's contribution
311 to the t-SNE, we noted that all patients contributed to all Tier 1 clusters, though note that p044
312 was overrepresented with more terminally differentiated epithelial cells, likely from incomplete
313 EDTA separation, and thus omit the p044 unique cell clusters from further analyses of
314 composition (**Figure 3b; Supplemental Figure 4d; Supplemental Table 13**). We then
315 proceeded to generate preliminary descriptive names based on inspection of each cluster within
316 each tier, calculated a hierarchically-clustered dendrogram, and produced systematic names for
317 each end cell state within each cell Tier 1 cell type (**Figures 3c,d; Supplemental Figure 5;**
318 **Supplemental Table 11; Methods**). We present top marker genes for each main Tier 1
319 cluster/cell type, and note that we also provide complete gene lists calculated through Wilcoxon
320 with Bonferroni adjusted $p<0.05$ for Tier 1 clusters/cell types, subsets, and end cell states (**Figure**
321 **3e, Supplemental Tables 5-7**). As patient identity did not factor into ITC stop conditions, we then
322 calculated Simpson's Index of Diversity for each of the clusters (**Figure 3d; Supplemental Figure**
323 **5**; Simpson's Index >0.1). Low diversity clusters may still reflect important biology for individual
324 patients, but we comment more extensively on clusters with high patient diversity.

325 Within B cells, we identified a strong division between non-cycling and cycling B cells, with those
326 found in the cycling compartment readily identifiable by germinal center markers and further dark
327 zone (*AICDA*) and light zone (*CD83*) genes resulting in FG.B/DZ.AICDA.IGKC and
328 FG.B/LZ.CD74.CD83 clusters (**Figure 3d**) (Victora et al., 2010).

329 Within myeloid cells, we identified, and confirmed using extensive inspection of literature curated
330 markers, cell subsets corresponding to monocytes (*CD14, FCGR3A, FCN1, S100A8, S100A9*,
331 etc.), macrophages (*CSF1R, MERTK, MAF, C1QA*, etc.), cDC1 (*CLEC9A, XCR1, BATF3*), cDC2

332 (*FCER1A, CLEC10A, CD1C, IRF4* etc.), and pDCs (*IL3RA, LILRA4, IRF7*) (**Figure 3d;**
333 **Supplemental Table 6**) (Blériot et al., 2020; Dutertre et al., 2019; Guilliams et al., 2018). We
334 highlight selected cell states including a migratory dendritic cell state (FG.DC.CCR7.FSCN1),
335 extensive cDC2 heterogeneity relative to cDC1 heterogeneity, and a main distinction between
336 macrophage clusters expressing *C1Q**, *MMP**, *APOE, CD68* and *PTGDS*,
337 (FG.Mac.C1QB.SEPP1, FG.Mac.APOE.PTGDS) and a series of clusters expressing various
338 chemokines including *CCL3, CXCL3,*and *CXCL8* (FG.Mac.CCL3.HES1, FG.Mac.CXCL3.CXCL8,
339 FG.Mac.CXCL8.IL1B).

340 Within T cells, we followed a similar approach as utilized for Myeloid cells and identified principal
341 cell subsets of T cells (joint expression of *CD247, CD3D, CD3E, CD3G* with *TRAC, TRBC1,*
342 *TRBC2*, or *TRGC1, TRGC2* and *TRDC*), and a combined cluster of cytotoxic cells
343 (FG.T/NK/ILC.GNLY.TYROBP) likely including T cells, NK cells (lower expression of TCR-
344 complex genes with *NCAM1*, *NCR1* and *TYROBP*), and some ILCs (*KIT, NCR2, RORC* and low
345 expression of CD3-complex genes) (**Figure 3d, Supplemental Table 6**) (Cherrier et al., 2018;
346 Robinette and Colonna, 2016). We note that the numerical majority of CD4 T cells
347 (FG.T/NK/ILC.MAF.RPS26) and CD8 T cells (FG.T/NK/ILC.CCR7.SELL) expressed *SELL* and
348 *CCR7* thus identifying them as naïve T cells. However, regardless of clusters expressing *CD4*
349 (FG.T.GZMK.GZMA) or *CD8A/CD8B* (FG.T.GZMK.IFNG, FG.T.GZMK.CRTAM, etc.), most
350 activated T cells were characterized by expression of granzymes (Sallusto et al., 1999).

351 Within epithelial cells, most cells expressed high levels of *OLFM4*, identifying them as crypt-
352 localized cells (Moor et al., 2018). We readily identified subsets of stem cells (*LGR5*), proliferating
353 cells (*TOP2A*), goblet cells (*SPINK4*, *ZG16*, various *MUCs*), enteroendocrine cells (*SCG3, ISL1*),
354 Paneth cells (*ITLN2, PRSS2, LYZ*), tuft cells (*GNG13, SH2D6, TRPM5*) and enterocytes (*APOC3,*
355 *APOA1, FABP6,* etc.) (**Figure 3d, Supplemental Table 6**) (Barker et al., 2007; van der Flier and
356 Clevers, 2009).

357 Within endothelial cells, we readily identified vascular and lymphatic endothelial cells (*LYVE1,*
358 *PROX1*), with the vascular cells able to be further identified as capillaries (*CA4*) or venular
359 endothelial cells (*ACKR1, MADCAM1*) (Brulois et al., 2020). We also identified a subset of cells
360 (FG.Endth/Peri.FRZB.NOTCH3) expressing high levels of *FRZB* and *NOTCH3*, which, rather than
361 being arterioles, likely represent arteriole-associated pericytes or smooth muscle cells given the
362 absence of *EFNB2, SOX17, BMX,* and *HEY1,* and the presence of *ACTA2* and *MYL9*, as cluster-
363 defining genes (**Figure 3d, Supplemental Table 6**) (Travaglini et al., 2020; Whitsett et al., 2019).

364 We highlight that the FG.Endth/Ven.ACKR1.MADCAM1 cluster is characterized by expression of
365 markers for postcapillary venules specialized in leukocyte recruitment (Thiriot et al., 2017).

366 Within fibroblasts, we identified principal subsets characterized by their structural roles (*COL3A1,*
367 *ADAMDEC1, FBLN1, LUM,* etc.), myofibroblasts (*MYH11, ACTA2, ACTG2,* etc.), and
368 organization of lymphoid cells (*CCL19, CCL21* etc.) (**Figure 3d, Supplemental Table 6**)
369 (Buechler et al., 2021; Davidson et al., 2021). Within the lymphoid-organizing fibroblasts, we draw
370 attention to the FG.Fibro.C3.FDCSP, FG.Fibro.CCL19.C3, and FG.Fibro,CCL21.CCL19 subsets,
371 which appear to have some characteristics of follicular dendritic cells and variable expression of
372 *CCL19/CCL21* (T-cell or migratory dendritic cell chemoattractants) and *CXCL13* (B-cell
373 chemoattractant) (Das et al., 2017; Heesters et al., 2013). We also identified a separate Tier 1
374 cluster of glial cells characterized by *CRYAB* and *CLU*. Intriguingly within the glial cell Tier 1
375 cluster, we then recovered a cell subset expressing *FDCSP, CXCL13,* and *CR2,* a key
376 complement receptor which allows for complement-bound antigens to be recycled and presented
377 by follicular dendritic cells (Das et al., 2017; Heesters et al., 2013). This highlights the power of
378 iterative tiered clustering to recover discrete cell states that may, through the process of traditional
379 clustering, not be fully resolved. Furthermore, the presence of these discrete cell clusters within
380 larger parent cell clusters will alter the gene expression signatures of the higher-level cell types.
381 For example, the FG.Glial/fDC.FDCSP.CXCL13 in the hierarchical cluster tree then assorts within
382 the lymphoid-organizing stromal cells.

383 The mast cells recovered did not further sub-cluster in an automated fashion, and were largely
384 marked by *TPSB2* and *TPSAB1* (>97%), with minimal *CMA1* (<20%) expressing cells, suggesting
385 they are largely classical MC-T cells in FGID intestine (Dwyer et al., 2021).

386 We identified four Tier 1 clusters for plasma cells, which are characterized by their strong
387 expression of IGH* immunoglobulin heavy-chain genes together with either an IGK* (kappa light
388 chain) or IGL* (lambda light chain) genes (Cyster and Allen, 2019; James et al., 2020). This
389 resolved IgA IgK plasma cells, IgA IgL plasma cells, IgM plasma cells, and IgG plasma cells.
390 Iterative tiered clustering identified further heterogeneity within all clusters of IgA and IgG plasma
391 cells, though given the 3'-bias of this dataset, we note that a principled investigation of these
392 clusters would ideally use 5' sequencing with targeted VDJ amplification.

393 Together, our treatment-naïve cell atlas from 13 FGID patients captures 138 cell clusters from a
394 non-inflammatory state of pediatric ileum which we annotated and named in a principled fashion.

395

396 **Comprehensive atlas of pediCD**

397 From the 124,054 cells profiled from 14 pediCD patients, we recovered 12 Tier 1 clusters which
398 here we display on a t-SNE plot colored by cluster identity (**Figure 4a**). Distinct from FGID, Paneth
399 cells clustered separately at Tier 1, while glial cells were now found within the fibroblast Tier 1
400 cluster. Inspecting each individual patient's contribution to the t-SNE, we noted that all patients
401 contributed to all Tier 1 clusters (**Figure 4b; Supplemental Figure 4c, Supplemental Table 13**).
402 We then proceeded to generate preliminary descriptive names, independently from the FGID
403 atlas, based on inspection of each cluster within each tier, calculate a hierarchically-clustered
404 dendrogram, and provide systematic names for each end cell state within each cell type and
405 subset (**Figures 4c,d; Supplemental Figure 6; Supplemental Table 12; Methods**). We present
406 top marker genes for each main Tier 1 cluster/cell type, and note the complete gene lists
407 calculated through Wilcoxon with Bonferroni adjusted $p<0.05$ available for Tier 1 clusters/cell
408 types, subsets, and end cell states (**Supplemental Tables 8-10**). We then calculated Simpson's
409 Index of Diversity to denote the patient diversity present within each end cluster, identifying that
410 most clusters in pediCD are conserved across multiple patients, and 16/305 clusters are single-
411 patient clusters (**Figure 4d; Supplemental Figure 5**; Simpson's Index >0.1). We note that in
412 pediCD, relative to FGID, a higher fraction of cell clusters exhibited lower patient diversity.

413 Within B cells, we also identified a strong division between non-cycling and cycling B cells, with
414 those found in the cycling compartment readily identifiable by germinal center markers and further
415 dark zone (*AICDA*) and light zone (*CD83*) genes, as in FGID (Victora et al., 2010). Within cells
416 expressing germinal centers markers, a highly-proliferative branch including clusters such as
417 CD.B/LZ.CCL22.NPW, CD.B/GC.MKI67.RRM2, and CD.B/DZ.HIST1H1B.MKI67 emerged
418 (**Figure 4d**). The CD.B/LZ.CCL22.NPW was characterized by high levels of *MYC*, which has been
419 shown to allow for further rounds of germinal center affinity maturation (Dominguez-Sola et al.,
420 2012). More numerous B cell clusters included ones characterized by expression of *GPR183*,
421 such as CD.B.CD69.GPR183 (also expressing *IGHG1*) and CD.B.RPS29.RPS21. GPR183 has
422 been shown to regulate the positioning of B cells in lymphoid tissues (Pereira et al., 2009).

423 Within myeloid cells, we identified, and confirmed using the same extensive inspection of literature
424 curated markers as in FGID, cell subsets corresponding to monocytes (*CD14, FCGR3A, FCN1,*
425 *S100A8, S100A9*, etc.), macrophages (*CSF1R, MERTK, MAF, C1QA*, etc.), cDC1 (*CLEC9A,*
426 *XCR1, BATF3*), cDC2 (*FCER1A, CLEC10A, CD1C, IRF4* etc.), and pDCs (*IL3RA, LILRA4, IRF7*)
427 (**Figure 4d; Supplemental Figure 6, Supplemental Table 9**) (Blériot et al., 2020; Dutertre et al.,
428 2019; Guilliams et al., 2018). We highlight selected cell states including a migratory dendritic cell

429 state (CD.DC.CCR7.FSCN1), extensive cDC2 heterogeneity relative to cDC1 heterogeneity, and
430 a main distinction between macrophages expressing *C1Q**, *MMP**, *APOE, CD68* and *PTGDS*,
431 (CD.Mac.APOE.PTGDS) and a series of clusters expressing various chemokines including
432 *CXCL2, CXCL3,* and *CXCL8* (CD.Mac.SEPP1.CXCL3, CD.Mono.CXCL3.FCN1,
433 CD.Mono.CXCL10.TNF). Several of the end cell clusters initially clustering with macrophages
434 also expressed monocyte markers (*S100A8, S100A9*), and expressed detectable, but lower levels
435 of *MERTK* or *AXL* relative to *bona fide* macrophages, potentially indicative of the early stages of
436 the trajectory of monocyte-to-macrophage differentiation (Blériot et al., 2020; Dutertre et al., 2019;
437 Guilliams et al., 2018). We also noted a substantial expansion of clusters characterized by
438 expression of *CXCL9, CXCL10,* and *STAT1,* canonical interferon-stimulated genes, observed in
439 clusters such as CD.Mono/Mac.CXCL10.FCN1 (Ziegler et al., 2020, 2021). Moreover, we
440 identified a cluster of inflammatory monocytes, CD.Mono.S100A8.S100A9, characterized by both
441 *CD14* and *FCGR3A* expression.

442 Within T cells, we followed a similar approach as utilized for FGID T cells and identified cell
443 subsets of T cells (joint expression of *CD247, CD3D, CD3E, CD3G* with *TRAC, TRBC1, TRBC2*,
444 or *TRGC1, TRGC2* and *TRDC*), but in pediCD also identified several discrete clusters of NK cells
445 (lower expression of TCR-complex genes with *FCGR3A* or *NCAM1*, *NCR1* and *TYROBP*), and
446 ILCs (*KIT, NCR2, RORC* and low expression of CD3-complex genes) (**Figure 4d, Supplemental**
447 **Figure 6, Supplemental Table 9**) (Cherrier et al., 2018; Robinette and Colonna, 2016). We note
448 that T cells and NK cells with a shared expression of *GNLY*, *GZMB* and other cytotoxic effector
449 genes cluster almost indistinguishably from each other through iterative tiered clustering and
450 visualization of the hierarchical tree, but that careful inspection of literature-curated markers
451 helped resolve NK cells (CD.NK.CCL3.CD160; CD.NK.GNLY.GZMB) from *CD8A*/*CD8B* T cells
452 (CD.T.GNLY.GZMH; CD.T.GNLY.CTSW) (**Figure 4d, Supplemental Figure 6, Supplemental**
453 **Table 9**) (Cherrier et al., 2018; Robinette and Colonna, 2016). One of the specific challenges in
454 distinguishing between T cells and NK cells in scRNA-seq data is that NK cells can express
455 several CD3-complex genes, particularly *CD247*, as well as detectable aligned reads for *TRDC*
456 or *TRBC1* and *TRBC2*, and thus lower-resolution clustering approaches or datasets with lower
457 cell numbers may miss these important distinctions (Björklund et al., 2016; Renoux et al., 2015).
458 NK cell clusters also expressed the highest levels of *TYROBP*, which encodes DAP12 and
459 mediates signaling downstream from many NK receptors (French et al., 2006; Lanier, 2001;
460 Lanier et al., 1998). ILC clusters such as CD.ILC.LST1.AREG or CD.ILC.IL22.KIT were
461 characterized by an apparent ILC3 phenotype, with expression of *KIT*, *RORC* and *IL22*, though
462 they also expressed detectable transcripts of *GATA3* in the same clusters (Cherrier et al., 2018;

463 Robinette and Colonna, 2016). We detected several clusters expressing *CD4* and lacking
464 *CD8A/CD8B*, including regulatory T cells (CD.T.TNFRSF18.FOXP3), and *MAF*- and *CCR6*-
465 expressing helper T cells (CD.T.MAF.CTLA4). Perhaps most strikingly, we resolved multiple
466 subsets of proliferating lymphocytes, including regulatory T cells (CD.T.MKI67.FOXP3), *IFNG*-
467 expressing T cells (CD.T.MKI67.IFNG), and NK cells (CD.NK.MKI67.GZMA).

468 Within epithelial cells we identified substantial heterogeneity in CD. Most cells expressed high
469 levels of *OLFM4*, identifying them as crypt-localized cells (Moor et al., 2018). We readily identified
470 subsets of stem cells (*LGR5*), proliferating cells (*TOP2A*), goblet cells (*SPINK4*, *ZG16*, various
471 *MUCs*), enteroendocrine cells (*SCG3, ISL1*), Paneth cells (*ITLN2, PRSS2, LYZ*), tuft cells
472 (*GNG13, SH2D6, TRPM5*) and enterocytes (*APOC3, APOA1, FABP6,* etc.) (**Figure 4d,**
473 **Supplemental Table 9**) (Barker et al., 2007; van der Flier and Clevers, 2009). Amongst several
474 clusters characterized by *CCL25* and *OLFM4* expression, we identified a subset marked by *LGR5*
475 expression, characteristic of intestinal stem cells (CD.EpithStem.LINC00176.RPS4YA1) (Barker
476 et al., 2007). We identified several subsets expressing *CD24*, indicative of crypt localization, with
477 expression of *REG1B* (CD.Secretory.GSTA1.REG1B; CD.Secretory.REG1B.REG1A) (Moor et
478 al., 2018). We also identified early enterocyte cluster CD.EC.ANPEP.DUOX2, characterized by
479 *FABP4* and *ALDOB* and expressing *DUOX2 and MUC1*. We resolved several clusters of
480 enteroendocrine cells, including CD.Enteroendocrine.TFPI2.TPH1 and
481 CD.Enteroendocrine.NEUROG3.MLN. We also found two clusters we labeled as M cells based
482 on expression of *SPIB* (CD.Mcell.CCL23.SPIB; CD.MCell.CSRP2.SPIB) (Beumer et al., 2020;
483 Mabbott et al., 2013). Paneth cells did not further sub-cluster despite forming an independent Tier
484 1 cluster (CD.Epith.Paneth). Most strikingly, we identified a diversity of goblet cells recovered
485 across multiple patients including CD.Goblet.HES6.COLCA2 expressing *REG4* and *LGALS9*, and
486 CD.Goblet.TFF1.TPSG1 expressing *TFF1* and *ITLN1*, amongst others. We also identified a
487 cluster of Tuft cells: CD.EC.GNAT3.TRPM5.

488 Within endothelial cells, we also readily identified vascular and lymphatic endothelial cells
489 (*LYVE1, PROX1*), with the vascular cells able to be further identified as capillaries (*CA4*) or
490 venular endothelial cells (*ACKR1, MADCAM1*) (**Figure 4d, Supplemental Table 9**). We also
491 identified a subset of cells (CD.Endth/Mural.HIGD1B.NDUFA4L2) expressing high levels of *FRZB*
492 and *NOTCH3*, which, rather than being arterioles, likely represent arteriole-associated pericytes
493 or smooth muscle cells given the absence of *EFNB2, SOX17, BMX,* and *HEY1,* and the presence
494 of *ACTA2* and *MYL9*, as cluster-defining genes. In pediCD, we also identified a cluster of arteriolar
495 endothelial cells, CD.Endth/Art.SEMA3G.SSUH2, identified by expression of *HEY1, EFNB2,* and

496 *SOX17*. We also highlight that the endothelial venules characterized by expression of markers for
497 postcapillary venules specialized in leukocyte recruitment, such as
498 CD.Endth/Ven.ADGRG6.ACKR1 and CD.Endth/Ven.POSTN.ACKR1, exhibited greater diversity
499 than in FGID with multiple end cell clusters identified (Thiriot et al., 2017).

500 Within fibroblasts, we identified principal subsets characterized by their structural roles (*COL3A1,*
501 *ADAMDEC1, FBLN1, LUM,* etc.), myofibroblasts (*MYH11, ACTA2, ACTG2,* etc.), and
502 organization of lymphoid cells (*CCL19, CCL21* etc.) (**Figure 4d**) (Buechler et al., 2021; Davidson
503 et al., 2021). The principal hierarchy in fibroblasts in pediCD was between *FRZB*-, *EDRNB*- and
504 *F3*-expressing subsets such as CD.Fibro.LY6H.PAPPA2 and CD.Fibro.AGT.F3, which were also
505 enriched for *CTGF* and *MMP1* expression, and *ADAMDEC1*-expressing fibroblasts, which were
506 enriched for several chemokines such as *CXCL12,* and in some specific clusters *CXCL6, CXCL1,*
507 *CCL11,* and other chemokines. Amongst three fibroblast subsets marked by *C3* expression, we
508 identified follicular dendritic cells (CD.Fibro/fDC.FCSP.CXCL13), along with fibroblasts
509 expressing *CCL21*, *CCL19*, and the interferon-stimulated chemokines *CXCL9* and *CXCL10*
510 (CD.Fibro.CCL21.CCL19; CD.Fibro.TNFSF11.CD24) (Das et al., 2017; Heesters et al., 2013).
511 Distinct from the FGID atlas, within the pediCD atlas, glial cells clustered within fibroblasts, but
512 were also marked by *S100B, PLP1* and *SPP1* expression.

513 The mast cells recovered in pediCD did further sub-cluster in an automated fashion, were largely
514 marked by *TPSB2* (>90%), with minimal *CMA1* (<16%) expressing cells, suggesting they are
515 largely classical MC-T cells in pediCD intestine (**Figure 4d**) (Dwyer et al., 2021). Intriguingly,
516 some subsets (CD.Mstcl.AREG.ADCYAP1) were enriched for *IL13*-expression. We also detected
517 a small cluster of proliferating mast cells from several patients (CD.Mstcl.CDK1.KIAA0101).

518 We also identified four Tier 1 clusters for plasma cells, which are characterized by their strong
519 expression of IGH* immunoglobulin heavy-chain genes together with either a IGK* (kappa light
520 chain) or IGL* (lambda light chain) genes. This resolved IgA IgK plasma cells, IgA IgL plasma
521 cells, IgM plasma cells, and IgG plasma cells. Iterative tiered clustering identified further
522 heterogeneity within all clusters of IgA plasma cells, though given the 3'-bias of this dataset, we
523 note that a principled investigation of these clusters would ideally use 5' sequencing with targeted
524 VDJ amplification.

525 Together, our treatment-naïve cell atlas from 14 pediCD patients captures 305 cell clusters from
526 an inflammatory state of the pediatric ileum suggesting an increase in the number and diversity
527 of cell states present in the intestine during overt inflammatory disease.

528

529 **Clinical variables and cellular variance that associates with pediCD severity**

530 As this pediCD atlas was curated from treatment-naïve diagnostic samples, we were able to
531 interrogate the data to test if overall shifts in cellular composition, specific cell states, and/or gene
532 expression signatures underlie clinically-appreciated disease severity and treatment decisions
533 (NOA vs. FR/PR), and those that are further associated with response to anti-TNF therapies
534 (either FRs or PRs). Here, we leveraged the detailed clinical trajectories collected from all patients
535 as the ultimate functional test: resolving how cellular composition and cell states predict disease
536 and treatment outcomes.

537 In order to capture the overall principal axes of variation explaining changes in cellular
538 composition, we calculated the fractional composition of all 305 end cell clusters in pediCD within
539 its parent cell type ("per cell type"), or within all cells ("per total cells"), and performed a principal
540 component analysis (PCA) over both of these sample x cell cluster frequency tables
541 (**Supplemental Table 14**) (Mathew et al., 2020). We then used the PC1 (13.4% variation "per cell
542 type" and 13.5% variation "per total cells") and PC2 (12.7% variation "per cell type" and 11.8%
543 variation "per total cells") as numerical variables which we correlated with clinical metadata
544 including categorical variables (patient ID, ethnicity, gender, etc.), ordinal variables (Terminal
545 Ileum (TI)-macroscopic endoscopic evidence, TI-microscopic histopathology, Anti-TNF treatment
546 within 90 days of diagnosis, and treatment decision/response coded as anti-TNF_NOA_FR_PR,
547 etc.) and numerical variables (Height, BMI, CRP, ESR, PLT, PCDAI, wPCDAI, etc.) (**Figure 5a**, r
548 by Spearman-rank). Amongst the clinical variables, we noted strong correlation between Initial
549 wPCDAI and CRP (r=0.83), and moderate correlation between Initial wPCDAI and anti-TNF within
550 30 days (r=0.65) and anti-TNF_NOA_FR_PR (r=0.49). For PC1-"per total cells", we identified
551 strong correlations with anti-TNF treatment within 90 days (r=-0.76), and moderate correlation
552 anti-TNF_NOA_FR_PR (r=-0.58; **Methods**). For PC1-"per cell type", we identified strong
553 correlation with anti-TNF within 90 days (r=-0.72), and moderate correlation with anti-
554 TNF_NOA_FR_PR status (r=-0.63). PC1-"per cell type" was also strongly correlated with BMI and
555 PC1-"per total cells" (r>-0.7). PC1-"per cell type" was weakly correlated with patient ID and gender
556 (r<0.3).

557 In order to understand if any cell types were predominantly driving associations with clinical
558 disease severity at initial presentation, we then further deconvoluted the overall PCA on 305 end
559 clusters and performed PCA over each cell type's fractional composition of end clusters
560 individually (B cells: 33 clusters, Endothelial: 18 clusters, Epithelial: 68 clusters, Fibroblast 45

561 clusters, Myeloid: 54 clusters, T/NK/ILCs: 57 clusters), and correlated the first two PC's (all PC1's
562 and PC2's each accounted for >13% variance) with all of the clinical variables (**Supplemental**
563 **Table 14**). The PCs derived from T/NK/ILC cells, myeloid cells, and epithelial cells were all
564 moderately correlated with anti-TNF_NOA_FR_PR status (r>0.49) individually and had higher
565 values than the other cell types; therefore, we asked if a PCA-based metric considering all three
566 cell types would synergistically capture both disease severity and treatment response. When we
567 calculated the PCA accounting for frequencies within each cell type of T/NK/ILC cells, myeloid
568 cells, and epithelial cells, we found strong correlation for PC2 with both anti-TNF within 90 days
569 (r=-0.83) and anti-TNF-NOA_FR_PR status (r=-0.87) (**Figure 5a**). This represented the two
570 strongest correlations of any variable we tested with anti-TNF treatment and response status,
571 outperforming wPCDAI.

572

573 **Discrete cell cluster changes across the pediCD clinical severity and response spectrum**

574 We next focused on further deconstructing the disease severity vector: identifying which cell
575 clusters accounted for the most significant changes in abundance based on the relative frequency
576 of an end cell cluster within its parent cell type. We focus on this form of analysis for scRNA-seq,
577 similar to what is typically reported for flow cytometry, and further discuss approaches to
578 enumerate total cell numbers which would be critical to identify changes in overall cellularity in
579 the different pediCD treatment and response categories (**Discussion**) (Gomariz et al., 2018). We
580 first performed a Fisher's exact test between NOA vs. FR; NOA vs. PR; or FR vs. PR, and then
581 performed a Mann-Whitney U test to highlight specific clusters and discuss results from clusters
582 with high Simpson's index of diversity (i.e. recovered from multiple patients) as shown for
583 T/NK/ILCs and Myeloid Cell Types (**Fig. 5b; Methods**).

584 When comparing FR/PRs to NOAs, two subsets with significantly increased frequency in FR/PR
585 patients amongst T cells, NK cells, and ILCs were identified. These were CD.NK.MKI67.GZMA
586 and CD.T.MKI67.IL22 (**Figure 5b,c; Supplemental Figure 7a; Supplemental Table 15**). Beyond
587 the strong proliferation signature, CD.NK.MKI67.GZMA were enriched for genes such as *GNLY,*
588 *CCL3, KLRD1, IL2RB* and *EOMES*, and CD.T.MKI67.IL22 were enriched for *IFNG, CCL20, IL22,*
589 *IL26, CD40LG* and *ITGAE*. This indicates that with increasing pediCD clinical severity, there is
590 increasing local proliferation of cytotoxic NK cells, and proliferation of tissue-resident T cells with
591 the capacity to express anti-microbial and tissue-reparative cytokines, and molecules to interface
592 with antigen-presenting cells and B cells. Alongside this increase, there was a significant
593 decrease amongst fibroblasts of CD.Fibro.CCL19.IRF7, and amongst epithelial cells of

594 CD.EC.SLC28A2.GSTA2 clusters in the FR/PR patients compared to NOA (**Supplemental**
595 **Figure 7a**). The CD.Fibro.CCL19.IRF7 were enriched for *CCL19, CCL11, CXCL1, CCL2,* and
596 very specifically for *OAS1* and *IRF7*. The CD.EC.SLC28A2.GSTA2 cluster was characterized by
597 its two namesake markers, involved in purine transport and glutathione metabolism (Moor et al.,
598 2018).

599 We also detected significant decreases in FRs relative to NOAs in certain cell types, particularly
600 within Epithelial cells including CD.EpithStem.LINC00176.RPS4Y1, CD.MCell.CSRP2.SPIB,
601 CD.EC.FABP6.PLCG2, and CD.EC.FABP1.ADIRF (**Supplemental Figure 7b; Supplemental**
602 **Table 15**). We note that the relative decrease in M cells is in stark contrast to the "ectopic" M-like
603 cells that were detected in adult ulcerative colitis (Smillie et al., 2019).

604 We next focused on those cell subsets that were significantly changed only between PRs and
605 NOAs (**Figure 5c; Supplemental Figure 7c; Supplemental Table 12**). Here we note several
606 more distinct clusters within the lymphocyte cell type, including increases in the PR patients
607 compared to NOA patients of CD.T.MKI67.IFNG, CD.T.MKI67.FOXP3, CD.T.GNLY.CSF2, and
608 CD.NK.GNLY.FCER1G. The two MKI67 clusters again highlighted an increase in proliferative
609 cells, specifically cells enriched for *IFNG, GNLY, HOPX, ITGAE* and *IL26* (CD.T.MKI67.IFNG),
610 and *IL2RA, BATF, CTLA4, TNFRSF1B, CXCR3,* and *FOXP3* (CD.T.MKI67.FOXP3), the latter of
611 which may be indicative of proliferating regulatory T cells. The two GNLY clusters emphasized
612 cytotoxicity, specifically cell clusters were both enriched for *GNLY, GZMB, GZMA, PRF1* and
613 more specifically for *IFNG, CXCR6,* and *CSF2* (CD.T.GNLY.CSF2), or *AREG, TYROBP,* and
614 *KLRF1* (CD.NK.GNLY.FCER1G). Amongst myeloid cells, there was an increase in
615 CD.Mac.CXCL3.APOC1, CD.Mono/Mac.CXCL10.FCN1, and CD.Mono.FCN1.S100A4 in PR
616 versus NOA. The CD.Mac.CXCL3.APOC1 cluster was enriched for a variety of chemokines
617 including *CCL3, CCL4, CXCL3, CXCL2, CXCL1, CCL20,* and *CCL8*. It was also enriched for *TNF*
618 and *IL1B*. The CD.Mono/Mac.CXCL10.FCN1 cluster was enriched for *CXCL9, CXCL10, CXCL11,*
619 *GBP1, GBP2, GBP4, GBP5,* suggestive of activation by IFN, and more specifically Type II IFNγ,
620 based on the GBP gene cluster (Ziegler et al., 2020). CD.Mono.FCN1.S100A4 was characterized
621 by *S100A4, S100A6,* and *FCN1* expression. These two immune clusters were paralleled by
622 increases in certain clusters within endothelial cells in PR versus NOA patients
623 (CD.Endth/Ven.LAMP3.LIPG) and epithelial cells (CD.Goblet.TFF1.TPSG1).

624 Several clusters of cells were decreased in PR versus NOA, including CD.T.LAG3.BATF,
625 CD.T.IFI44L.PTGER4, and CD.T.IFI6.IRF7 amongst lymphocytes (**Supplemental Figure 7c**)
626 (Roncarolo et al., 2018). Amongst myeloid cells, CD.cDC2.CLEC10A.FCGR2B were decreased,

627 and amongst fibroblasts CD.Fibro.IFI6.IFI44L were decreased. In epithelial cells,
628 CD.Tuft.GNAT3.TRPM5 cells were decreased. Alongside the decrease in Tuft cells amongst
629 epithelial cells, two more clusters closely related to the aforementioned CD.EC.GSTA2.SLC28A3
630 cluster, also marked by *GSTA2* expression, were significantly decreased (CD.EC.GSTA2.CES3,
631 and CD.EC.GSTA2.TMPRSS15).

632 We assessed the compositional differences between FRs and PRs and only identified one cell
633 cluster which was significantly increased in PRs: CD.B/DZ.HIST1H1B.MKI67, which are
634 proliferating dark zone B cells. CD.T.EGR1.TNF T cells were significantly decreased in PR versus
635 FR (**Supplemental Figure 7d; Supplemental Table 15**). These data suggest that at the earlier
636 stages of pediCD, there are a series of changes in multiple cell types that encapsulate the
637 distinctions between NOA and FR or PR patients.

638

639 **Collective cell vectors delineating pediCD clinical severity and response spectrum**

640 Together, the significant changes in cell composition between the clinically-defined patient groups
641 were particularly notable within proliferating T cells, cytotoxic NK cells, monocytes/macrophages,
642 and epithelial cells that could be combined to calculate a numerical variable for "PC2-
643 T/NK/ILC/Myeloid/Epithelial" which correlated strongly with both the clinical presentation leading
644 to a decision to treat or not with anti-TNF therapies and to the distinction between anti-
645 TNF_NOA_FR_PR status (**Figure 5a**). Some of the top negative loadings for PC2, enriched in
646 PRs and FRs compared to NOAs, included both helper and cytotoxic T cell clusters
647 (CD.T.MAF.CTLA4; CD.T.CCL20.RORA; CD.T.GNLY.CSF2), NK cell clusters
648 (CD.NK.GNLY.FCER1G; CD.NK.GNLY.IFNG; CD.NK.GNLY.GZMB); proliferating T cells and NK
649 cells (CD.T.MKI67.FOXP3; CD.T.MKI67.IFNG; CD.T.MKI67.IL22; CD.NK.MKI67.GZMA), and
650 monocytes, macrophages, DCs and pDCs (CD.cDC2.CD1C.AREG; CD.Mac.C1QB.CD14;
651 CD.Mono.CXCL3.FCN1; CD.pDC.IRF7.IL3RA; CD.Mono/Mac.CXCL10.FCN1) (**Figure 5d;**
652 **Supplemental Table 16**). The top positive loadings for PC2 encompassing the NOA-enriched
653 clusters included several epithelial cell subsets such as Tuft cells (CD.EC.GNAT3.TRPM5) and
654 those with specialized metabolic features including retinol-binding, bile binding and export, fatty-
655 acid and cholesterol metabolism, fructose and glucose metabolism, starch metabolism
656 glutathione metabolism, sulfation, and the terminal degradation of peptides
657 (CD.EC.RBP2.CYP3A4; CD.EC.FABP6.PLCG2; CD.EC.FABP1.ADIRF;
658 CD.EC.GSTA2.TMPRSS15) (**Figure 5d; Supplemental Table 16**) (Lampen et al., 2000;
659 Mårtensson et al., 1990; Martínez-Augustin and de Medina, 2008; Sullivan et al., 2021; Wen and

660 Rawls, 2020). Furthermore, clusters also enriched in NOA PC2 such as CD.EC.ADH1C.RPS4Y1
661 and CD.EC.ADH1C.GSTA1, clustered in a separate branch together and expressed several
662 enzymes responsible for steroid hormone and dopamine biosynthesis (**Figure 4d, 5d**) (Cima et
663 al., 2004; Magro et al., 2002). Importantly for the regenerative potential of the epithelium,
664 CD.EpithStem.LINC00176.RPS4Y1 were also defining of the PC2-positive NOA direction. This
665 suggests that multiple collective changes in the composition and/or state of T/NK/ILC cells,
666 myeloid cells, and epithelial cells at diagnosis may help stratify pediCD patients not only by
667 clinically appreciated disease severity but also may influence anti-TNF responsiveness.

668 To determine whether the disease severity gene signature that we discovered in the PREDICT
669 study can be found in other cohorts, we selected the top 92 markers of the 25 cell states
670 associated with disease severity and treatment outcomes (**Supplemental Table 17**) and
671 performed a gene-set enrichment analysis (GSEA) (**Supplemental Figure 7e**). In the two
672 independent treatment-naïve cohorts that we analyzed (the pediatric RISK cohort, n = 69, and the
673 adult E−MTAB−7604 cohort, n = 43) our gene signature was significantly enriched in illeal, but
674 not colonic, mucosal biopsies from patients who did not respond to anti-TNF therapy compared
675 to those who responded (Kugathasan et al., 2017; Verstockt et al., 2019). Thus, these genes
676 (*TNFAIP6, GZMB, S100A8, CSF2, CLEC4E, S100A9, IL1RN, FCGR1A, CLIC3, CD14, PLA2G7,*
677 *FAM26F, IL3RA, NKG7, IL32, CCL3, OLR1, LILRA4, APOC1, MYBL2*) informed by the PC2
678 cellular vector, and showing best ranks in both cohorts, could potentially serve as predictive
679 markers of anti-TNF therapy outcome in newly diagnosed patients.
680

681 **Random forest classifier applied to cellular taxonomies allows for identification of**
682 **correspondence between FGID and pediCD**

683 As we had generated independent cellular atlases for FGID and pediCD to mitigate "discovery"
684 of hybrid cell clusters that may not represent bona fide biological cell states, we next sought to
685 match and identify correspondence between pediCD and FGID cell subsets. As our study
686 progressed, several analytical methods to integrate scRNA-seq emerged which utilize distinct
687 principles to either predict cell type names given reference gene lists or directly integrate two
688 datasets that were collected from distinct perturbations, tissues, or even species (Hao et al., 2021;
689 Hie et al., 2019; Korsunsky et al., 2019; Pliner et al., 2019). However, many of these methods are
690 benchmarked on broad cell type or subset integration, and thus their applicability for fine cell
691 states, as in the end clusters we identify here, remains unknown. Thus, we employed a random
692 forest (RF) classifier-based approach, which has recently also been applied successfully in work
693 to identify correspondence in fine sub-clusters in the mammalian retina (Peng et al., 2019;

694 Shekhar et al., 2016). Specifically, we employed paired RF models (one trained on FGID the other
695 trained on pediCD) to obtain cross dataset predictions per cell. We trained these models in scikit-
696 learn with 5-fold cross validation and params: min_samples_leaf=1, oob_score=True,
697 criterion="gini", max_depth=200, n_estimators=700, max_features="sqrt" (Pedregosa et al.,
698 2011). The training set (but not the test set) was sampled with replacement such that all classes
699 contained as many samples as the maximum proportioned class. This up-sampling procedure
700 provided the largest gain to our test accuracy, sensitivity, and specificity scores, increasing
701 accuracy ~10-15% across each cell type. With our final model, we attained cross-dataset
702 predictions (pediCD to FGID & FGID to pediCD) for each cell, giving a probability score of a cell
703 belonging to a subset in the other disease condition (**Supplemental Figure 5c; Methods**). We
704 applied these models to each cell type individually, and here focus our discussion on Myeloid
705 cells and T/NK/ILC cells as two cell types prominently associated with pediCD disease severity
706 (**Figure 6, Supplemental Figure 8**). As newer methods are developed, more refined integration
707 is likely to be possible.

708 Comparing across myeloid cells between pediCD and FGID, we could identify strong
709 correspondence of specific cell subsets such as cDC1s or pDCs (**Figure 6a**). We also identified
710 strong correspondence between several cDC2 clusters. We identified a gradient of monocyte and
711 macrophage correspondence of 31 clusters in pediCD to 2 FGID clusters, likely reflective of
712 inflammatory monocyte to macrophage differentiation in pediCD (Blériot et al., 2020; Dutertre et
713 al., 2019; Guilliams et al., 2018). Some clusters characterized by *STAT1* activation did not
714 demonstrate significant correspondence to any FGID cluster. We also generally noted
715 substantially increased cluster diversity in pediCD end clusters relative to their correspondence
716 in FGID. This emerged from more patient-specific clusters found in pediCD, and an overall
717 decrease in Simpson's index of diversity considering the patient composition of each end clusters
718 (**Figure 6b**).

719 For T/NK/ILC cells, we identified more discrete patterns relative to Myeloid cells based on
720 comparison of the RF result. Within the two FGID cytotoxic T cell clusters, we identified
721 correspondence by 18 pediCD clusters, representing ILC3s, and cytotoxic NK cells and T cells
722 (**Supplemental Figure 8**). The cluster of naïve T cells in FGID had correspondence with the
723 majority of pediCD non-cytotoxic T cell clusters, illustrating a substantial activation and
724 specialization to several discrete T cell states that were specific for pediCD.

725 Importantly, when we jointly clustered macrophages from FGID and pediCD together, we
726 identified that several of the original end clusters identified through ARBOL in pediCD were

727 divided across the UMAP: being split into distinct clusters of cells (**Figure 6c-f**). This reinforces
728 the need to quantitatively approach the choice of clustering parameters and number of iterations
729 used. We also employed the STACAS package to integrate T cells between FGID, confirming the
730 higher percentage of proliferating T cells in pediCD patients compared to FGID (**Supplemental**
731 **Figure 8b**) (Andreatta and Carmona, 2021). However, as compared to ITC, the integration
732 approach on T cells resulted in lower heterogeneity, thus masking important differences revealed
733 by our newly developed ARBOL approach. Our detailed analysis of the correspondence between
734 the FGID and pediCD atlases highlights the challenge of multiple cell type, subset, and state
735 vectors which are simultaneously accounted for by clustering over a set of highly variable genes
736 jointly derived from multiple cells and disease conditions. In an atlas composed of multiple clinical
737 entities, highly homogenous cell clusters may appear dispersed across a space based on the
738 other cells that they are being compared against and the parameters used for clustering. This
739 underscores the power in the disease-specific clustering heuristic employed with this dataset,
740 which revealed principled end clusters in pediCD which could then be related back to a non-
741 inflammatory reference atlas while still maintaining fine granularity.

742

743 **The phenotypic space of macrophages and T cells is significantly different across FGID**
744 **and NOA/FR/PR pediCD**

745 Based on their over-representation within clusters showing more significant differences within
746 pediCD, we next focused on performing an analysis over a shared gene expression space of
747 FGID and pediCD of the monocytes/macrophages (**Figure 7**) and T/NK/ILCs (**Figure 8**). We
748 utilized a list of genes that were cell-type defining genes in either FGID or pediCD, but removed
749 genes that were differentially expressed between FGID and pediCD, to allow for cell type/subset
750 to drive placement on the UMAP (**Methods**) (Ordovas-Montanes et al., 2018). This allowed us to
751 place our fine-grained clusters within a joint gene-expression space related to underlying cell
752 types in FGID and pediCD, and we also contextualize our findings with an orthogonal integration
753 approach applied to the T/NK/ILCs (**Supplemental Figure 8**) (Andreatta and Carmona, 2021).

754 Within FGID monocytes/macrophages, we identified that the majority of clusters occupied the
755 periphery of the UMAP space, including chemokine-expressing clusters (FG.Mac.CCL3.HES1;
756 FG.Mac.CXCL8.IL1B) and metabolic clusters (FG.Mac.APOE.PTGDS) (**Figure 7a,c**). This was
757 in stark contrast to the pediCD monocytes/macrophages, where we identified that now many of
758 the clusters occupied the central region of the UMAP (**Figure 7a,c**). We highlight several of the
759 clusters that are significantly different in frequency between the pediCD groups which were found

760 in this central region (**Figure 7b**). Notably, NOA, FR and PR pediCD clusters had significantly
761 different distributions within this space (**Figure 7c-e**). There was a progressive increase in the
762 Hellinger Distance (**Figure legend 7e**: a measure of the distance between two distributions) of
763 the distributions from FGID to NOA, NOA to FR, and NOA to PR. PRs had the most significantly
764 different distribution relative to FGID. This was in large part driven by cells from both FRs and
765 PRs that inhabit the central region, together with a loss of density in the chemokine-expressing
766 macrophages from FGID through to PRs. Of note, the frequency of TNF+ macrophages and its
767 expression level of TNF was significantly increased in FRs relative to all other groups (**Figure 7f;**
768 permutation test shuffling anti-TNF response variable, FR had significantly more TNF+ cells than
769 expected by chance with p approximating 0). Despite the fine-grained tiered clustering approach
770 used, the majority of clusters had high Simpson's Diversity indices representing cell states found
771 in several patients (**Figure 7g**).

772 Within T/NK/ILCs, we identified that FGID cells were more uniformly mixed with the pediCD cells
773 relative to monocytes/macrophages (**Figure 8a**). FGID cells occupied naïve and quiescent states,
774 showed some signatures of activation, and also specialization towards helper and cytotoxic states
775 (**Figure 8a,c**) (Sallusto et al., 1999). The most notable changes in the Hellinger Distance
776 distribution occurred between FGID and FR, rather than between FGID and PR as may have
777 been expected (**Figure 8d,e**). Similar to the monocytes/macrophages, the main area which
778 gained density with increased disease severity was the central region: characterized by
779 proliferation of several clusters increased in frequency within FRs and PRs to anti-TNF including
780 T cells such as CD.T.MKI67.FOXP3 and CD.T.MKI67.IL22 (**Figure 8b-d**). Proliferation-
781 associated gene signatures were also seen in extreme corners driven by CD.NK.MKI67.GZMA,
782 and significantly increased from FGID through to PR (**Figure 8b,f**). Intriguingly the only T cell
783 clusters in pediCD with gini coefficient <0.1 are from pediCD patients (**Figure 8g**). Taken together
784 with several cell clusters associated with pediCD proliferation overlapping with existing areas
785 found in FGID, this indicates activation and more extreme diversification from existing T cell states
786 driving the T cell clustering that defined pediCD. This is distinct from the recruitment and failed
787 differentiation towards homeostatic cell states between FGID and pediCD that we discovered in
788 monocytes/macrophage clusters. Both joint protections confirm and extend our RF predictions.

789 In order to provide tissue-scale context and understand the impact on other cell types for these
790 anti-TNF response associated lymphocyte states, we assessed the relationship of these
791 proliferating T and NK cell clusters with epithelial and myeloid cells (**Supplemental Figure 8c**).
792 We found that CD.T.MKI67.FOXP3 were strongly positively associated with

793 CD.Goblet.RETNLB.ITLN1 and CD.EC.NUPR1.LCN2 secretory cell states. Conversely, we found
794 that CD.NK.MKI67.GZMA were significantly negatively correlated with CD.EC.ADH1C.EDN1,
795 CD.Mcell.CSRP2.SPIB, CD.EC.ADH1C.RPS4Y1, CD.EC.GSTA2.TMPRSS15, and
796 CD.EpithStem.LINC00176.RPS4Y1. This indicates the potential for cytotoxic activity of the
797 proliferating NK cells towards more homeostatic cell states of epithelial cells, and critically of
798 intestinal stem cells, with increased disease severity.

799

800 ## Discussion

801 This study addresses a critical unmet need in the fields of IBD and systems immunology: the
802 creation of an atlas of newly-diagnosed and untreated diseased tissue, coupled with detailed
803 clinical follow-up to link diagnostic cell types and states with disease trajectory. This is especially
804 true for GI disease and others like it which afflict tissues that are not easily accessible without
805 operative or endoscopic intervention, and where tissue-specific immune pathology dictates
806 disease severity and trajectory. Likewise, cross-sectional studies, as have been the norm for most
807 previous scRNA-seq studies of IBD, are not able to overlay disease trajectory and treatment
808 response onto the topography of a complex multi-cellular atlas, thus limiting the mechanistic and
809 predictive inferences that can be drawn from the generated atlas (Corridoni et al., 2020a, 2020b;
810 Drokhlyansky et al., 2019; Elmentaite et al., 2020; Huang et al., 2019; Kinchen et al., 2018; Martin
811 et al., 2019; Parikh et al., 2019; Smillie et al., 2019). Furthermore, mouse models of CD, and of
812 IBD more broadly, may not be the most appropriate models for understanding treatment
813 resistance in pediCD (Neurath, 2019). To surmount these limitations, we created a prospective
814 clinical study, and enrolled patients requiring a diagnostic biopsy for possible IBD, prior to
815 diagnosis. This allowed us to capture a tremendously valuable control group: those patients with
816 FGID, who experience GI symptoms without evidence of GI inflammation or autoimmunity. These
817 uninflamed controls served as a critical comparator to contextualize the evidence of immune
818 pathology that we observed in patients with pediCD. With these detailed clinical phenotypes as
819 our foundation, we developed an automated ITC algorithm for scRNA-seq data, ARBOL, which
820 defines a vector of T cells, myeloid cells and epithelial cells that cleanly stratifies both Crohn's
821 disease severity and response to treatment.

822 The availability of comprehensive clinical, flow cytometric and scRNA-seq data from patients with
823 pediCD and from uninflamed FGID controls created an unprecedented opportunity for
824 comparative atlas creation. We took the opportunity to develop a methodical, unbiased, approach

825 to cell state discovery, ARBOL, released alongside our manuscript (https://github.com/jo-m-
826 lab/ARBOL). ARBOL iteratively explores axes of variation in scRNA-seq data by clustering and
827 subclustering until variation between cells becomes noise. The philosophy of ARBOL is that every
828 axis of variation could be biologically meaningful so each should be explored, and that axes of
829 variation are relative to the comparative outgroup, meaning that similar cell states may arise at
830 distinct tiers. Once every possibility is explored, curation and a statistical interrogation of
831 resolution are used to collapse clusters into the elemental transcriptomes of the dataset. ARBOL
832 inherently builds a tree of subclustering events. As data is separated by major axes of variation
833 in each subset, later rounds capture less pronounced variables. This comes with some caveats:
834 variation shared by all cell types (for example, cell cycle stage) can make up one of the major
835 axes of variation in the first round of clustering. Cell types can split up at the beginning, so the
836 same splitting of B and T cells, for example, may happen further down in separate branches. The
837 resulting tree of clustering events (**Supplemental Figure 5a**) is therefore neither indicative of true
838 distances between end clusters nor a tree of unique groupings. We address this problem by
839 calculation of a binary tree of manually and computationally curated end clusters. Using a
840 standardized method of end cluster naming, which we describe in ARBOL's tutorial (https://jo-m-
841 lab.github.io/ARBOL/ARBOLtutorial.html), we found the resulting binary tree assorted end
842 clusters into appreciated cell types and subsets (**Figures 3 and 4**), and also reveals further
843 previously unappreciated granularity that will serve as the foundation for future work into the
844 cellular composition of the gastrointestinal tract.

845 One of the primary remaining challenges going forward will be to identify which clusters are truly
846 patient-unique, or are simply patient-unique at the cohort size to which we are currently limited to.
847 We calculate a diversity metric for each end cluster to highlight those which are largely conserved
848 between patients, and provide complete cluster-defining gene lists for both FGID and pediCD at
849 three levels of clustering. We also provide links to our data visualization portal to enable cross-
850 atlas comparisons: https://singlecell.broadinstitute.org/single_cell/study/SCP1422/predict-2021-
851 paper-fgid and https://singlecell.broadinstitute.org/single_cell/study/SCP1423/predict-2021-
852 paper-cd.

853 One of the chief advantages of enrolling pediCD patients at diagnosis, and prior to any therapeutic
854 intervention, was that we were able to relate their diagnostic immune landscape with disease
855 trajectory. In the pediCD group, we identified 3 clinical subgroups. The first distinction was made
856 by treating physicians, and classified patients with milder versus more severe clinical disease
857 characteristics at diagnosis. The milder patients were not placed on anti-TNF agents (NOA), while

858 the more severe patients were treated with monoclonal antibodies that neutralize TNF including
859 infliximab and adalimumab. The second distinction between patient groups could not be made at
860 diagnosis, but rather, was based on clinical and biochemical response to anti-TNF agents. Thus,
861 of those patients treated with anti-TNF therapeutics, some were FRs, and some were only PRs,
862 with PRs requiring anti-TNF dose modifications and the addition of other agents, and with
863 ongoing, uncontrolled disease signs and symptoms. While differences in ant-TNF
864 pharmacokinetics have been partially implicated in the need to dose-escalate anti-TNF agents in
865 some pediCD patients, our study identifies foundational differences in the immune state at
866 diagnosis in PR patients compared to the NOA and FR subgroups (D'Haens and Deventer, 2021;
867 Ordás et al., 2012; Yarur et al., 2016). While standard flow cytometry was not able to distinguish
868 the immune phenotype of NOA versus treated patients, scRNA-seq identified significant
869 differences. The contextualization of our scRNA-seq derived predictive cellular vector with two
870 other treatment-naïve bulk RNA-seq studies of Crohn's disease, underscores the broader
871 applicability of our findings (Kugathasan et al., 2017; Verstockt et al., 2019).

872 We noted significant cell state changes at diagnosis underlying clinically-appreciated disease
873 severity that impacted the clinical decision to treat or not to treat with anti-TNF agents. These
874 occurred within multiple clusters of T, NK, fibroblast, epithelial, monocyte, macrophage, and
875 dendritic cells. For anti-TNF response, very few clusters exhibited significantly differential
876 composition between FR and PR individuals. This suggests that multiple collective changes in
877 several cell types may conspire to lead to differences in treatment outcomes. Indeed, when we
878 jointly considered a cellular principal component vector comprising epithelial cells, T/NK/ILCs,
879 and myeloid cells, we identified several clusters that together could delineate the full spectrum of
880 NOA, FR, and PR. This cellular vector indicated that multiple T cell subsets, NK cells, monocytes,
881 macrophages, and epithelial cells were altered in disease. Intriguingly, by finely clustering each
882 cell type, we found that proliferating T and NK cells do not represent a uniform population, but
883 rather reflect functional specialization capturing *FOXP3, IFNG, IL22,* and *GZMA* as cluster-
884 defining genes. Enriched in NOA individuals were epithelial cells involved in chemosensation
885 (Tuft.GNAT3.TRPM5) and absorption of metabolites (EC.GSTA3.TMPRSS15), as well as stem
886 cells (Banerjee et al., 2020; von Moltke et al., 2016; Sido et al., 1998). That pediCD severity is
887 not uniquely predicted by a singular cell subset or gene is reflective of the complex genetics and
888 environmental factors that have been implicated, along with the rich literature that has found
889 significant changes by histology, flow cytometry, or mass cytometry in CD relative to control
890 tissue(Buisine et al., 2001; Leeb et al., 2003; Leonard et al., 1995; Lilja et al., 2000; Mitsialis et
891 al., 2020; Müller et al., 1998; Souza et al., 1999; Stappenbeck and McGovern, 2017; Takayama

et al., 2010). *However, with the PREDICT study, we have discovered precisely which changes in CD cellular composition come together to form a predictive vector for both disease severity and treatment response*. Intriguingly, the quantification and visualization of this response vector predicted a later escalation of one of our patients (p022; who appeared as an outlier FR in **Figure 5d**) from FR to PR, which occurred after our database lock in December of 2020.

When considering the relationships between T cells and NK cells along with epithelial cells, we captured that proliferating cytotoxic NK cell subsets like CD.NK.MKI67.GZMA were significantly negatively correlated with critical metabolic and progenitor epithelial cell subsets in pediCD. Conversely, proliferating regulatory CD.T.MKI67.FOXP3 were positively associated with secretory epithelial cells in pediCD, but did not appear related to the decrease in metabolic or progenitor cells. How T cell-derived cytokines impact intestinal regeneration and differentiation has recently been the focus of several studies, but the relationship of these fine-grained T cell subsets with specific epithelial cell states observed in the human intestine remained unknown (Biton et al., 2018; Lindemans et al., 2015). Our work suggests that in the context of the ileum impacted by CD that there is further complexity to understand, particularly as it pertains to cytotoxic NK cells and T cells and their impact on epithelial cell homeostasis and regeneration.

The mapping of these disease severity-associated cell networks identifies a host of new potential therapeutic targets for pediCD, for many of which there are clinical-stage therapeutics that could be investigated. These include CD40L-blocking antibodies, IL-22 agonists, and targeted anti-proliferation agents (Betts et al., 2017; Lindemans et al., 2015; Miura et al., 2021; Ramanujam et al., 2020; Sootome et al., 2020).. A case can also be built for targeting inflammatory cytokines such as IL-1, and for interrogating agents aimed at mucosal healing including new anti-GM-CSF antibodies, given that several prominent cell subsets marked by CSF2 were enriched in the PR patients (Ai et al., 2021; Aschenbrenner et al., 2021; Castro-Dopico et al., 2020; Mehta et al., 2020; Mitsialis et al., 2020; Muro and Mrowiec, 2015). This atlas therefore provides a rigorous evidence-based rationale for proposing new therapeutic interventions, as well as a mechanism for interrogating the impact of new agents on the longitudinal immune landscape of pediCD patients.

Recent work on COVID-19 has also highlighted the challenges faced by systems approaches to capture baseline cell states that predict disease trajectory (Kaczorowski et al., 2017; Lucas et al., 2020; Mathew et al., 2020; Schulte-Schrepping et al., 2020; Su et al., 2020). In a disease of known infectious etiology with SARS-CoV-2, monocytes, macrophages, granulocytes, T cells, B cells, antibodies, and interferon state have all independently been associated with disease outcomes.

925 Few studies have considered how multiple collective changes at baseline may influence outcome,
926 yet are likely more reflective of the disease. With the complex and protracted presentation of a
927 multifactorial disease like Crohn's disease, we posit that multiple concerted effects are required
928 to dictate both the severity (NOA vs FR/PR) and the treatment-response (FR vs PR). Future work
929 will aim to consider which cell subsets are recovered during mucosal healing, and how closely
930 the treated state reflects each individual patient's baseline presentation.

931

932 **Limitations of the Study**

933 Our study has several limitations that are important to consider. Here we are limited by the cohort
934 size to the strict inclusion/exclusion criteria used, follow-up period, and the ability to capture
935 untreated patients early during the diagnosis process. As a result, we are underpowered to
936 directly correlate in our cohort genetic and environmental influences such as the enteric
937 microbiome, infections, and dietary factors (Dovrolis et al., 2020; Rajca et al., 2014; Yilmaz et al.,
938 2019). We also are limited in our ability to capture spatial information and context from our cohort.
939 It will also be critical to understand how disease progresses and the reason for persistent NOA,
940 FR, or PR patient states in some individuals which will require retention in the trial and repeat
941 possibility of for-cause biopsy procurement. Furthermore, our cohort is representative of the
942 demographics of FGID and IBD in Seattle, WA. It will be essential to understand how both FGID
943 and IBD may present differently across the world, and multiple cohorts will be required to
944 understand location-unique and generalizable findings. We also recognized the number of
945 pediCD patients within this cohort limits our ability to subdivide CD phenotype into location and
946 behavior of disease (via the Paris/Montreal classification systems) (Levine et al., 2011). Based
947 on this study, we will seek to expand our efforts to validate and extend our findings.

948
949
950 **ACKNOWLEDGMENTS**

951 We thank the patients and their families for helping and contributing to our study. We thank Kathy
952 McConville and Sarah Mbonde for identifying and recruiting patients for PREDICT. We thank
953 Carly G.K. Ziegler for discussions and advice related to iterative tiered clustering, and all members
954 of the Kean, Ordovas-Montanes and Shalek labs for thoughtful discussions. J.O.-M is a New York
955 Stem Cell Foundation – Robertson Investigator. J.O.-M was supported by the Richard and Susan
956 Smith Family Foundation, the AGA Research Foundation's AGA-Takeda Pharmaceuticals
957 Research Scholar Award in IBD – AGA2020-13-01, the HDDC Pilot and Feasibility P30
958 DK034854, the Food Allergy Science Initiative, the Leona M. and Harry B. Helmsley Charitable

Trust, The Pew Charitable Trusts Biomedical Scholars and The New York Stem Cell Foundation. L.S.K is supported by NIH P01 1P01HL158504, R01 5R01HL095791, U19 U19AI051731, and by the Helmsley Charitable Trust. V.N. was supported by the International mobility of research, technical and administrative staff of research organizations (CZ.02.2.69/0.0/0.0/18_053/0016981). A.K.S. was supported, in part, by the Searle Scholars Program, the Beckman Young Investigator Program, a Sloan Fellowship in Chemistry, and the NIH (5U24AI118672, 2R01HL095791). V.M. reports research support from Novartis. V.T. Is supported by ASTCT New Investigator Award and CIBMTR/Be the Match Foundation Amy Strelzer Manasevit Research Program Award. S.B.S. is supported by NIH grants P30 DK034854 and RC2 DK122532, and the Helmsley Charitable Trust.

AUTHOR CONTRIBUTIONS

Conceptualization: G.W., A.Y., A.K.S., L.S.K., H.B.Z., J.O.-M

Methodology: B.A.D., K.K., A.Y., L.A., G.W., D.L.S., D.L., V.N., A.K.S., L.S.K., V.T., H.B.Z., J.O.-M

Formal analysis: B.A.D., K.K., A.Y., V.N., X.D., V.T., H.B.Z., M.F., M.S., J.S.C., J.O.-M

Investigation: B.A.D., K.K., L.A., K.C., R.F., P.K., A.A., L.C., C.M., A.Y., G.W., D.L.S., D.L., M.F., V.N., L.S.K., X.D., G.D., B.B., K.B., V.T., L.V.C., V.M., H.B.Z., S.B.S., J.O.-M

Resources: B.A.D., K.K., A.K.S. L.S.K., H.B.Z., J.O.-M

Data Curation: B.A.D., K.K., A.Y., P.K., M.F., V.N., V.T., H.B.Z., J.O.-M

Writing – Original Draft: B.A.D., K.K., V.N., A.K.S., L.S.K., H.B.Z., M.S., J.S.C., J.O.-M

Writing – Review & Editing: B.A.D., K.K., L.A., G.W., D.L.S., A.Y., A.A., P.K., K.C., R.F., B.B., K.B., L.C., C.M., D.L., M.F., V.N., G.D., A.K.S., L.S.K., V.M., V.T., L.V.C., H.B.Z., M.S., J.S.C., F.T., S.B.S., J.O.-M

Visualization: B.A.D., K.K., V.N., A.K.S. H.B.Z., M.S., J.O.-M

Supervision: A.K.S., L.S.K., J.O.-M

Project Administration: A.K.S., L.S.K., B.B., H.B.Z., F.T., J.O.-M

Funding Acquisition: A.K.S., L.S.K., J.O.-M

DECLARATION OF INTERESTS

J.O.-M. reports compensation for consulting services with Cellarity and Hovione. A.K.S. reports compensation for consulting and/or SAB membership from Merck, Honeycomb Biotechnologies, Cellarity, Repertoire Immune Medicines, Hovione, Third Rock Ventures, Ochre Bio, FL82, and Dahlia Biosciences unrelated to this work. A.K.S. has received research support from Merck, Novartis, Leo Pharma, Janssen, the Bill and Melinda Gates Foundation, the Moore Foundation,

994 the Pew-Stewart Trust, Foundation MIT, the Chan Zuckerberg Initiative, Novo Nordisk and the
995 FDA unrelated to this work. Dr. Kean is on the scientific advisory board for HiFiBio and Mammoth
996 Biosciences. She reports research funding from Kymab Limited, Magenta Therapeutics, BlueBird
997 Bio, and Regeneron Pharmaceuticals. She reports consulting fees from Equillium, FortySeven
998 Inc, Novartis Inc, EMD Serono, Gilead Sciences, Vertex Pharmaceuticals, and Takeda
999 Pharmaceuticals. Dr. Kean reports grants and personal fees from Bristol Myers Squibb that are
1000 managed under an agreement with Harvard Medical School. L.A. is a consultant for Takeda
1001 Pharmaceuticals. G.W. reports Research funding from Abbvie, Jansen, Takeda, Allakos; SAB for
1002 Abbvie, Bristol Myers Squibb; DSMB for Abbvie. D.L.S., is the co-founder, CM, president of
1003 NiMBAL Health. S.B.S. declares the following interests: Scientific advisory board participation for
1004 Pfizer, Lilly, IFM therapeutics, Merck, Pandion, and Takeda Inc., and grant support from Pfizer,
1005 Novartis, Amgen, Takeda Consulting for Hoffman La Roche and Amgen. A.K.S., L.S.K., J.O.-M.,
1006 H.B.Z., and B.A.D., are co-inventors on a provisional patent application relating to methods of
1007 stratifying and treating IBD.

1008

1009

1010

1011

1012

1013

1014

1015

1016

1017

1018

1019

1020

1021

1022 ## Supplemental Tables

1023

1024 **Supplemental Table 1: Flow cytometry panels**

1025 **Supplemental Table 2: scRNA-seq sample filtering thresholds**

1026 **Supplemental Table 3: scRNA-seq quality control metrics for all single-cells**

1027 **Supplemental Table 4: Traditional joint clustering for broad cell types and differential**
1028 **expression testing by Wilcoxon to determine CD or FGID-enriched genes.**

1029 **Supplemental Table 5: FGID cell type markers**

1030 **Supplemental Table 6: FGID cell subset markers**

1031 **Supplemental Table 7: FGID cell state (i.e. end cell cluster) markers**

1032 **Supplemental Table 8: CD cell type markers**

1033 **Supplemental Table 9: CD cell subset markers**

1034 **Supplemental Table 10: CD cell state (i.e. end cell cluster) markers**

1035 **Supplemental Table 11: FGID end cell cluster descriptive names and short curated names**
1036 **look up table**

1037 **Supplemental Table 12: CD end cell cluster descriptive names and short curated names**
1038 **look up table Supplemental Table 13: CD cell type markers**

1039 **Supplemental Table 13: Number of cells per patient per end cell cluster**

1040 **Supplemental Table 14: PCA Loadings for joint Epithelial, Myeloid, T/NK/ILC vectors**

1041 **Supplemental Table 15: Differential composition testing for NOA vs FR, NOA vs PR and PR**
1042 **vs FR categories of anti-TNF response with CD patients**

1043 **Supplemental Table 16: PC2 top positive and negative cell clusters**

1044 **Supplemental Table 17: 92 markers derived from PC2 used for bulk RNA seq extension**

1045

1046

1047

1048

1049

1050

1051

1052

1053

1054 # METHODS

1055 ## RESOURCE AVAILABILITY

1056 **Corresponding Authors**

1057 Further information and requests for resources and reagents should be directed to and will be
1058 fulfilled by Leslie S. Kean, Jose Ordovas-Montanes, and Alex K. Shalek.

1059

1060

1061 **Data and Code Availability**

1062 Single-cell RNA-seq data can be visualized via the Single Cell Portal links for each atlas of FGID
1063 https://singlecell.broadinstitute.org/single_cell/study/SCP1422/predict-2021-paper-fgid and
1064 pediCD https://singlecell.broadinstitute.org/single_cell/study/SCP1423/predict-2021-paper-cd.
1065 The cell-by-gene matrices will be available with the peer-reviewed version of this manuscript. The
1066 raw human FASTQ files will be available from the Broad controlled access repository DUOS with
1067 the peer-reviewed version of this manuscript. Please contact the authors

1068 All original code has been assembled as the ARBOL package and deposited at GitHub and is
1069 publicly available together with this manuscript at: https://github.com/jo-m-lab/ARBOL and
1070 https://jo-m-lab.github.io/ARBOL/ARBOLtutorial.html.

1071

1072 ## EXPERIMENTAL MODEL AND SUBJECT DETAILS

1073 **Study Population and Clinical Parameters**

1074 Pediatric patients less than 20 years of age with suspected inflammatory bowel disease were
1075 enrolled on the PREDICT Study (ClinicalTrials.gov# NCT03369353). Enrollment took place
1076 between November 9, 2017 to December 21, 2018 in accordance with the Fred Hutch Institutional
1077 Review Board (Protocol #9730, ethical approval given) with written informed consent and assent
1078 when applicable. Patients diagnosed with Crohn's Disease (CD) were included and patients
1079 without gut inflammation on endoscopy and histology, and who were diagnosed with Functional
1080 GI Disease (FGID), served as a comparative cohort for this study. Terminal ileum and blood
1081 samples were taken during the diagnostic endoscopy procedures prior to initiation of therapy.
1082 Patients diagnosed with other inflammatory or infectious etiologies on endoscopy and biopsy were
1083 excluded from the analysis.

1084 Clinical course and variables were monitored at the time of enrollment and for 3 years after initial
1085 endoscopy, with median follow up for CD being 32.5 months and FGID being 31 months at the
1086 time of clinical database lock (December 1, 2020). Medical management was dictated by
1087 clinicians. Clinical variables obtained included sex, race, age at diagnosis, weight z-score, height
1088 z-score, BMI z-score, clinical disease severity using the Pediatric Crohn's Disease Activity Index
1089 (PCDAI), and disease location and phenotype using the Montreal Criteria (Hyams et al., 1991;
1090 Silverberg et al., 2005). Laboratory evaluation included C-reactive protein, ESR, hemoglobin,
1091 albumin, white blood cell count, and platelet count.
1092
1093 **Response to Anti-TNF therapy**
1094 Early anti-TNF or immunomodulator therapy was defined as initiation of immunosuppression
1095 within 90 days of diagnostic endoscopy. Anti-TNF monoclonal antibody was started in 10 patients
1096 with CD. All patients were followed prospectively and categorized as full responders (FR), partial
1097 responders (PR), or not on anti-TNF (NOA). Full response to anti-TNF is defined as clinical
1098 symptom control and biochemical response with wPCDAI score of <12.5 on maintenance anti-
1099 TNF therapy and partial response defined as lack of clinical symptom control and biochemical
1100 response with documented escalation of anti-TNF therapy.
1101
1102 **Clinical Statistical Analysis**
1103 Clinical variables are expressed as median (lower and upper confidence interval; range) and
1104 compared using the Mann-Whitney U test. Categorical variables were described as frequencies
1105 and percentages and compared using the chi-square test. Clinical laboratory values are
1106 represented by mean and standard error of the mean (range) and compared with the Mann-
1107 Whitney U test. Significance is indicated by a P value of <0.05. Clinical statistical analysis was
1108 performed using GraphPad Prism version 8.3.0.
1109
1110 **METHOD DETAILS**
1111 **Tissue Dissociation into Single-Cell Suspensions**
1112 *Human Ileum*
1113 Single-cell suspensions were collected from intestinal biopsies using a modified version of a
1114 previously published protocol (Persson et al., 2013) as described in (Smillie et al., 2019). One
1115 biopsy from the terminal ileum was received directly in hand and processed with an average time
1116 from patient to loading on the 10X Chromium platform of 2.5 total hours, and never exceeding 3.5
1117 hours. While intact, biopsy bites were handled using a P1000 pipette applying gentle suction, and

1118 all centrifugation steps done in a temperature controlled 4°C centrifuge. Biopsy bites were first
1119 rinsed in 30 mL of ice-cold PBS (ThermoFisher 10010-049) and allowed to settle. Each individual
1120 bite was then transferred to 10 mL epithelial cell solution (HBSS Ca/Mg-Free [ThermoFisher
1121 14175-103], 10 mM EDTA [ThermoFisher AM9261], 100 U/ml penicillin [ThermoFisher 15140-
1122 122], 100 µg/mL streptomycin [ThermoFisher 15140-122], 10 mM HEPES [ThermoFisher 15630-
1123 080], and 2% FCS [ThermoFisher 10082-147]) freshly supplemented with 200 µL of 0.5M EDTA.
1124 Separation of the epithelial layer from the underlying lamina propria was performed for 15 minutes
1125 at 37°C with rotation at 120RPM. The tube was then removed and placed on ice immediately for
1126 10 minutes before shaking vigorously 15 times. Visual macroscopic inspection of the tube at this
1127 point yielded visible epithelial sheets, and microscopic examination confirmed the presence of
1128 single-layer sheets and crypt-like structures.
1129
1130 The remnant tissue bite was carefully removed and placed into a large volume of ice-cold PBS to
1131 rinse before transferring to 5mL of enzymatic digestion mix (Base: RPMI1640, 100 U/ml penicillin
1132 [ThermoFisher 15140-122], 100 µg/mL streptomycin [ThermoFisher 15140-122], 10 mM HEPES
1133 [ThermoFisher 15630-080], 2% FCS [ThermoFisher 10082-147], & 50 µg/mL gentamicin
1134 [ThermoFisher 15750-060]), freshly supplement immediately before with 100 µg/mL of Liberase
1135 TM [Roche 5401127001] and 100 µg/mL of DNase I [Roche 10104159001]), at 37°C with 120
1136 rpm rotation for 30 minutes. During this 30-minute lamina propria (LP) digestion, the epithelial
1137 (EPI) fraction was spun down at 400g for 7 minutes and resuspended in 1 mL of epithelial cell
1138 solution before transferring to a 1.5mL Eppendorf tube in order to minimize time spent centrifuging
1139 and provide a more concentrated cell pellet. Cells were spun down at 800g for 2 minutes and
1140 resuspended in TrypLE express enzyme [ThermoFisher 12604013] for 5 minutes in a 37°C bath
1141 followed by gentle trituration with a P1000 pipette. Cells were spun down at 800g for 2 minutes
1142 and resuspended in ACK lysis buffer [ThermoFisher A1049201] for 3 minutes on ice to remove
1143 red blood cells, even if no RBC contamination was visibly observed in order to maintain
1144 consistency across samples. Cells were spun down at 800g for 2 minutes and resuspended in 1
1145 mL of epithelial cell solution and placed on ice for 3 minutes before triturating with a P1000 pipette
1146 and filtering into a new Eppendorf tube through a 40 µM cell strainer [Falcon/VWR 21008-949].
1147 Cells were spun down at 800g for 2 minutes and then resuspended in 200 µL of epithelial cell
1148 solution and placed on ice while final steps of LP dissociation occurred. After 30 minutes, the LP
1149 enzymatic dissociation was quenched by addition of 1ml of 100% FCS [ThermoFisher 10082-147]
1150 and 80 µL of 0.5M EDTA and placing on ice for five minutes. Samples were typically fully
1151 dissociated at this step and after gentle trituration with a P1000 pipette filtered through a 40µM

1152 cell strainer into a new 50 mL conical tube and rinsed with PBS to 30 mL total volume. This tube
1153 was spun down at 400g for 10 minutes and resuspended in 1 mL of ACK and placed on ice for 3
1154 minutes. LP cells were spun down at 800g for 2 minutes and resuspended in 1 mL of epithelial
1155 cell solution and spun down at 800g for 2 minutes and resuspended in 200 µL of epithelial cell
1156 solution and placed on ice. Following centrifugation, the cells from both EPI and LP fractions were
1157 counted and prepared as a single-cell suspension for scRNA-seq. Since the full EPI isolation was
1158 not performed on all patients limiting sample sizes, here we focus our analysis on LP fractions.
1159

1160 **Flow Cytometry**

1161 Multicolor flow cytometry was performed on tissue samples to examine the immune composition
1162 for enrolled patients. Flowjo software was used to phenotypically define cell populations that will
1163 be analyzed and compared in patients using two-way ANOVAs (or non-parametric equivalent).
1164 Antibodies used include: CD3 APC, SP34-2 (BD Biosciences); CD3 BUV661, UCHT1 (BD
1165 Biosciences); CD3 BV711, OKT3, (Biolegend); CD3 PE, SP34 (BD Biosciences); CD4 BV785,
1166 OKT4 (Biolegend); CD8a BUV395, RPA-T8 (BD Biosciences); CD8b FITC, REA715 (Miltenyi
1167 Biotec); CD11b APC-Cy7, ICRF44 (BD Biosciences); CD11c APC-eFlour 780, BU15 (Fisher
1168 Scientific); CD11c BUV661, B-ly6 (BD Biosciences); CD14 APC-eFluor 780, 61D3 (Fisher
1169 Scientific); CD14 BUV737, M5E2 (BD Biosciences); CD20 APC-eFluor 780, 2H7 (Fisher
1170 Scientific); CD20 PE-Cy7, L27 (BD Biosciences); CD38 APC, HIT2 (BD Biosciences); CD45
1171 PerCP/Cy5.5, HI30 (Biolegend); CD45RA BV605, HI100 (Biolegend); CD56 (NCAM) FITC,
1172 TULY56 (Fisher); CD94 APC-Vio770, REA113 (Miltenyi Biotec); CD117 (c-kit) BV421, 104D2
1173 (Biolegend); CD123 BV711, 9F5 (BD Biosciences); CD127 Biotin, HIL-7R-M21 (BD Biosciences);
1174 CD161 BV711, DX12 (BD Biosciences); CD197 (CCR7) BV421, GO43H7 (Biolegend); CD294
1175 (CRTH2) BV605, BM16 (Biolegend); CD326 (Epcam) APC, HEA-125 (Miltenyi Biotec); HLA-DR
1176 APC-H7, L243 (G46-6) (BD Biosciences); TCR PAN γδ PE-Cy7, IMMU510 (Beckman Coulter);
1177 α4-β7 integrin (Act-1), (NIH AIDS Reagent Program); Streptavidin BUV737 (Fisher); Live/dead
1178 Fix Aqua (Fisher); R-PE Antibody Labeling Kit (300 mcg) (Abcam).
1179

1180 **Methods to Generate Single-Cell RNA-seq Libraries and Sequencing**

1181 *10X v2 3'*

1182 Single cells were loaded onto 3' library chips as per the manufacturers protocol for Chromium
1183 Single Cell 3' Library (v2) (10X Genomics). The LP fraction was captured in its own channel of
1184 the 10X Chromium Single Cell Platform, in order to recover sufficient numbers of cells for
1185 downstream analyses. An input of 10,000 single cells was added to each channel with a recovery

1186 rate of 9,514 cells per sample based on median across samples. Briefly, single cells were
1187 portioned into Gel Beads in Emulsion (GEMs) in the Chromium controller with cell lysis and
1188 barcoded reverse transcription of RNA, followed by cDNA amplification, enzymatic fragmentation
1189 and 5' adaptor and sample index attachment. Libraries were sequenced on a HiSeq or NovaSeq
1190 flow cell. The read structure was paired end with length of read 1 26bp, length of read 2 91bp,
1191 and the length if index 1 (i7 primer) 8bp. Quality-filtered base calls were converted to
1192 demultiplexed FASTQ files.
1193
1194 **Alignment and Filtering**
1195 FASTQ files were aligned to GRCh38 using Cellranger v2.2 pipeline on the Cumulus/Terra cloud
1196 pipeline https://portal.firecloud.org/?return=firecloud#methods/cumulus/cellranger_workflow/10
1197 generating 27 cell-by-gene matrices (13 FGID, 14 CD), one for each patient. We used default
1198 parameters of the 10th snapshot version of the pipeline, aside from requiring that it use cellranger
1199 v2.2.0.
1200
1201 Every sample was first filtered excluding genes measured in fewer than 3 cells and cells with
1202 fewer than 200 unique genes. To control for doublets and low-quality cells we then further filtered
1203 individually, attempting to match the approximate 10,000 cells loaded onto the sample lane and
1204 balancing the thresholds to not cut out dense regions of a Ncounts by Nfeatures scatter plot. Pre-
1205 filtering, we looked for outlier samples, based on proportion of percent mitochondrial genes,
1206 number of counts, and number of features, none fell beyond the 1.5 times the IQR threshold.
1207
1208 Exact thresholds used for each sample:

Sample ID	Diagnosis	min_cells	min_genes	max_genes	max_count	N_cells_post
p011_T0D_ILE_LPS_3p	CD	3	200	3000	20000	11087
p014_T0D_ILE_LPS_3p	CD	3	200	3000	20000	7285
p016_T0D_ILE_LPS_3p	CD	3	200	3000	20000	7713
p018_T0D_ILE_LPS_3p	CD	3	300	4000	20000	15469
p019_T0D_ILE_LPS_3p	CD	3	200	5000	20000	11200
p022_T0D_ILE_LPS_3p	CD	3	200	3000	20000	7798
p024_T0D_ILE_LPS_3p	CD	3	200	3000	20000	9705
p026_T0D_ILE_LPS_3p	CD	3	300	4000	30000	10847
p027_T0D_ILE_LPS_3p	CD	3	200	4000	40000	8124
p031_T0D_ILE_LPS_3p	CD	3	200	3000	20000	12177
p032_T0D_ILE_LPS_3p	CD	3	200	3000	20000	9052
p041_T0D_ILE_LPS_3p	CD	3	200	4000	40000	7508

p042_T0D_ILE_LPS_3p	CD	3	200	4000	30000	12219
p048_T0D_ILE_LPS_3p	CD	3	200	4000	40000	9158
p009_T0D_ILE_LPS_3p	FGID	3	200	3000	20000	11426
p017_T0D_ILE_LPS_3p	FGID	3	200	3000	20000	11107
p023_T0D_ILE_LPS_3p	FGID	3	200	3000	20000	11262
p028_T0D_ILE_LPS_3p	FGID	3	200	4000	20000	10293
p030_T0D_ILE_LPS_3p	FGID	3	200	4000	30000	9590
p033_T0D_ILE_LPS_3p	FGID	3	200	3000	20000	9514
p034_T0D_ILE_LPS_3p	FGID	3	200	3000	20000	8142
p035_T0D_ILE_LPS_3p	FGID	3	200	3000	20000	5668
p037_T0D_ILE_LPS_3p	FGID	3	200	3000	20000	4210
p043_T0D_ILE_LPS_3p	FGID	3	200	5000	30000	5045
p044_T0D_ILE_LPS_3p	FGID	3	350	4000	30000	12727
p049_T0D_ILE_LPS_3p	FGID	3	200	3000	20000	8003
p050_T0D_ILE_LPS_3p	FGID	3	200	4000	30000	8582

Post filtering, we merged sample matrices using an outer join to create an FGID dataset (115,569 cells) and a pediCD dataset (139,342 cells).

QUANTIFICATION AND STATISTICAL ANALYSIS

Preprocessing & Clustering of scRNA-seq Data

1st Approach – Classical Methods on Combined Dataset

We began initial analysis following traditional clustering and annotation techniques; however, these methods using manual and at times subjective metrics scaled poorly to the size and scope of our dataset and moreover did not give clear distinction between disease specific cell states and compositional shifts within cell states across disease.

For our first pass at analysis, we grouped the FGID and pediCD datasets together (254,911 cells) and proceeded with the standard Seurat v3.1.5 pipeline (Stuart et al., 2019). We used manual heuristics of gene marker specificity to choose cluster resolution and isolate 9 major cell types (T, B, plasma, epithelial, endothelial, fibroblast, myeloid, mast cell, and glial) and 1 aggregate cluster of T, B, myeloid, and epithelial cells with a strong proliferation signature. We then subclustered the proliferating group and manually merged the proliferating cells with their corresponding cell type based on marker gene expression, and separately re-preprocessed and clustered each cell

type annotating based on one vs. rest differential expression (Wilcoxon, fdr < 0.05) within the cell type.

We found several disadvantages to this approach. First, we found it difficult to determine for each cluster whether we should be looking for changes in compositional frequency or gene expression. Particularly within the myeloid major cell group we would find extremely disease biased sub-clusters, as much as a 9:1 ratio between pediCD to FGID. It was unclear whether there was massive compositional shift within a conserved cell state or if instead a base cell state was split into multiple clusters based on phenotypic differences in disease and we should perform a differential expression test between it and neighboring FGID biased clusters. Second, after two rounds of manual processing we were still unsure if we had reached a base level with each end cluster corresponding to a unique and biologically homogeneous cell state. Third, at that point, having partitioned over 100 distinct clusters, individually supervising each subsets processing and sub-clustering was infeasible. We needed a more systematic method to address these challenges.

2nd approach – ARBOL: Iterative Tiered Clustering (ITC) on Separated Disease Conditions

It is common to organize cell identity ontologies in a tree structure. With major groups such as immune, stromal, and epithelial at the top and branching down a level, you might set more nuanced identities like T, B, endothelial and goblet cell types as a second tier, and even more nuanced identities like CD4+ vs CD8+ T cells as a third. In ideal circumstances, this mental model conforms well to RNA-seq data where we can layer gene modules with more and more specific variation together to describe highly particular cell identities and states. And, by clustering at a high level with genes that vary across the entire dataset, then sub-clustering with genes that vary only within a particular parent cluster we are able reveal this hierarchy of cell identity. Many additional factors to cell identity contribute to the variation in gene expression within actual datasets, particularly as we found with disease condition during our first approach.

To be able to choose the appropriate future analyses and comparisons, we need a highly accurate representation of cell identity and state. The underlying issue in our first pass at clustering was that in combining the disease conditions together, the variable genes selected at each stage represented a combination of differences between cell identity & disease. This combination could have been manageable if either disease or cell identity were consistently more variable. We could isolate one factor at a specific tier in the hierarchy before sub-clustering to isolate the other. In

our case, disease and cell identity both had many overlapping scales of variation. To address this problem, we isolated cell identity by separating our dataset by disease and clustering for cell identity within each disease set (FGID 115,569 cells, pediCD 139,342 cells). This approach did then require us to perform an additional stage of analysis to find corresponding clusters between the two datasets, but allowed us to far more effectively distinguish type, scale, and specificity of disease differences.

Within each disease set we still needed a method to ensure we were reaching the bottom level of biological heterogeneity, and preferably an automated method as our first pass had shown the potential for isolating hundreds of cell states. To efficiently cluster and isolate these cell states we wrote a cloud-based pipeline to systematically optimize parameter selection and stop when biological heterogeneity is exhausted. Homogenous cell subsets were isolated by recursively normalizing, selecting variable genes, and clustering based on silhouette score. We stopped recursing into sub-clusters once we reached one of four end conditions defined as:

1. Having a group of less than 100 cells (though we did partition many clusters smaller than 100 cells after clustering groups just larger than that cutoff).
2. Isolating an optimized clustering of only one cluster.
3. Finding two clusters that have fewer than 25 genes (fdr< 0.05 & |log fold change| > 1.5 & percent expression >= 20% in at least one cluster) differentially upregulated between each cluster using a bimodal test developed in (Shekhar2016 10.1016/j.cell.2016.07.054). For this last condition, if reached, we reject the clustering and return back the cells as a single end cluster.
4. Having reached a max tier limit. Setting this value to 10, we never triggered this condition with either FGID or pediCD datasets, but included it to prevent runaway recursion.

Code for generating this tree of cell clusters is currently available here: (https://jo-m-lab.github.io/ARBOL/ARBOLtutorial.html. Within each recursion, the established steps were processed using Seurat version 3.1.5 (https://github.com/satija.lab/seurat). Normalization and variable gene selection were processed with SCTransform (https://github.com/ChristophH/sctransform) (Hafemeister and Satija, 2019). Clustering for major cell types was performed using Louvain clustering on dimensionally reduced principal components.

Parameters depend on the size of the dataset, and thus must be adjusted based on how many cells are being partitioned for each recursive step. When calculating nearest neighbor graphs, and clustering we set the K parameter to `ceiling(0.5*sqrt(N))`. We chose the number of principal components based on the top 15 percentile of calculated improvement of variance explained. For subsets less than 500 cells we used Jackstraw to calculate significant principal components. If neither method succeeded, we chose the first two principal components. This only occurred rarely at late stages where Jackstraw was unable to find significance. We set clustering resolution via a grid search optimizing for maximum average silhouette score (silhouette measures the ratio of intra-cluster distance to inter-cluster distance, where a high score means highly distinct clusters). For stages where we were clustering more than 500 cells a randomized subsample of N cells / 10 was used to calculate the average silhouette score.

Additionally, at each recursive step we output quality metrics and basic plots, such as 1:rest differential expression from the optimal partitioning at each stage and UMAP representations painted by sample metadata (sample ID, cluster number). Our pipeline, saved output as a directory structure matching the tree discovered by this recursive clustering. This tree represents the lower levels of variance of discovered at each tier. At any tier level we are able to extract the cell's partitioning. Due to the intermixing of patient and cell identity effects at multiple levels of the tree (a fraction of a single patient's cells might separate out at a high level, but then continue to separate into identifiable cell types, or vice versa), we found the most meaningful levels at the top and bottom of the tree. The clustering tree is useful for understanding the levels of variance in the dataset, but we found it contains too much noise to be easily interpretable. Thus, we later generated a hierarchical clustering of the bottom level clusters based on pairwise differential expression, which is displayed in figures (**Figure 3**: FGID atlas, **Figure 4**: pediCD atlas). See the hierarchical clustering of cell subsets section for details.

Cell Type and Subset Annotation from Tiered Clustering

After running the hierarchical tiered clustering pipeline we manually curated the generated tree of clusters. Specifically, tree generation was reinitiated for the B Cells within the FGID dataset as it had stopped at the first tier on two clusters with < 50 genes differentially expressed, however we could see in this case that there was additional biological stratification based on strong differential expression of CXCR4 (Wilcoxon; logFC=1.22860917, Bonferroni.p=1.0E-300) CD69 (Wilcoxon; logFC=1.27527652, Bonferroni.p=2.99E-151), HMGN1 (Wilcoxon; logFC=-1.1688612, Bonferroni.p=1.62E-227) and HMGA1 (Wilcoxon; logFC=-1.28838294, Bonferroni.p= 1.06E-209)

among others. This formed a clear divide between non-proliferating and proliferating B Cells, further validated by a clear separation within the UMAP based on PCA reduced variable genes within the B cells. We further examined each branching point of the tree to determine its splitting cause, noting splits based on spillover, doublet, and singular patient effects. Splits at higher tiers based on doublets often split again allowing us to recover cells that did not have the dual expression profile. Splits that only had patient splits below (measured by having only clusters of single patients) were manually marked as end clusters, thereby merging all clusters below that split. With these manual steps made, we performed pairwise differential expression to ensure each partitioned subset is distinct from its neighbors.

We annotated these final clusters with four methods attempting to balance descriptiveness, ease of understanding, and ease of name generation: The first method, is generated during the hierarchical tiered clustering by following the path from the end cluster up to the original tier. An example annotation is *T0C0.T1C3.T2C3.T3C5* marking an end cluster that split at tier 1 into cluster 0 and at tier 2 into cluster 3. These annotations do not provide any biological information to the reader, but do provide a unique ID for the end cluster. Our second method is far more descriptive, where we manually annotate the main reason for each particular split. This still follows the original ranking of variation as found by the hierarchical tiered clustering, while also providing biological interpretation, as an example: *CD.Mloid.macrophage_chemokine.S100A8_S100A9_CXCL9_CXCL10_TNF_inflamonocyte*. This method of annotation was particularly useful during analysis as we were immediately able to see how early or late two clusters had split from each other, as well as seeing a number of the subset defining genes. Unfortunately, as is apparent this method also produces extremely long names that are difficult to display and to which to refer. It is also a highly manual process, and difficult to reproduce precisely. To better present our findings and aid others in reproducing our results, our third method automates this annotation. This method is performed by taking each major cell type, which in our case matched the tier one splits, and performing 1:rest differential expression testing (Wilcoxon; adj.p<0.05, only.pos=True) within each major cell type. We then ranked the genes based on the product of `-log(Bonferroni.p)`, `avg_logFC`, and `pct.exp.1/pct.exp.rest` and took the top 5, forming a name like *CD.Mloid_CCL3_CCL4_CCL3L3_TNF_TNFAIP6*. This scheme, again was useful, but did not quite meet our demands of recognizability and brevity. Thus for T and Myeloid cells we adjusted these names to a finer degree of specificity by visualizing the expression profiles of each subset with a dotplot of canonical marker genes based off of current literature, and limiting to the top 2

genes based off our method 3 rankings and the dotplot of canonical markers, thereby producing the fourth and final annotations in the form: *CD.Mono.CXCL10.TNF.* Due to the limited nature of current characterization of stromal and epithelial cells we were unable to match the same degree of specificity as the T and Myeloid cells, however we did where possible adjust from the major cell type, to the most specific that we could be confident of. For instance, adjusting “Epith” to “Goblet” based on marker expression of TFF3 and MUC13.

Hierarchical Clustering of Subsets on Unified Gene Space and Removal of Doublets

At this point we had generated a hierarchical representation of the datasets from the top down showing the splits of highest variation at every level. By necessity that means that each level is controlled by and represents different selections of genes, which may have no relation to the genes selected in another branch. To understand the relations of cell subsets and compare across cell type we needed a unified set of genes. For each dataset (FGID and pediCD) we performed pairwise differential expression (Wilcoxon; Bonferroni.p<0.01, max.cells.per.ident=500) and selected the top 50 most significant genes from each test. Gene lists were merged as a union, finding 4445 unique genes for FGID and 1760 unique genes for pediCD that best differentiate the subsets. Subset centers were calculated from these selected genes as the median expression of cells grouped by subset. The resulting table was then hierarchically clustered using correlation distance and complete linkage. Clustering was performed in R using the pvclust package (https://github.com/cran/pvclust)

The resulting tree shows from the bottom up the relationships between cell subsets, and allows cell subsets that were potentially misclassified at a high split in hierarchical tiered clustering to find their biological neighboring subsets. As previously mentioned within our description of hierarchical tiered clustering we did not find any end cluster subsets that met our thresholds for merging. This does not mean that we did not observe shuffling from the initial tiered splits. While overall there was good agreement between the two methods, we noted subsets jumping between major cell types as defined by the first splits of our tiered clustering. We identified the majority of these jumping subsets as doublet clusters by exploring their differential gene results at multiple levels of the tiered clustering tree. We removed these doublet subsets and others based on flipping expression programs at different tiers. For instance, looking like T cells expressing *TRAC*, *IL7R* within an epithelial cluster, then at the next tier expressing *KRT18* and *PIGR*. After removing doublets, we recalculated subset distances and dimensional reductions, as presented in the main figures.

Finding Corresponding Cell Subsets Between Diseases

Separating the data on disease condition into two datasets was important as it allowed us to isolate the axis of cell identity within each disease and be confident in the homogeneity of each subset. But does present the problem of how to make comparisons across the disease condition, which is our main research question.

1st Approach – KNN classifier

Our first attempt to find corresponding clusters followed the methods of *Tasic et al. 2018*. We used the best differentiating genes sets created for our unified gene space clustering to as the mapping space for a nearest-neighbor classifier. For each cell within the disease condition, we could map it to the nearest cell subset within the other disease condition. As a trial run, we created this gene space for each major cell type of the FGID disease condition and performed 5-fold cross-validation. Unfortunately, we could only achieve accuracies of up to 55%.

Rethinking approach

From this trial run we realized that we needed a more automated system to choose genes as the most significantly differentially expressed genes did not create enough separation between cluster centers to effectively classify new cells. We chose to use an RF classifier as it allowed us to train for the optimal selection of genes, required little to no preparation of data, and provided probabilities of each cell being predicted to each class. These probabilities for each class proved particularly useful do to our second realization. Because the number of subsets differs between disease conditions, we cannot make the assumption that there is a one-to-one relationship between conditions. We also cannot make the assumption that the many-to-one relationships are unidirectional with one base subset splitting into many states only from FGID toward CD. A single classifier would not allow us to distinguish between these many types of relationships. However, we realized that by creating a classifier for both directions (FGID to pediCD & pediCD to FGID) we could take advantage of the difference in confidence between the two classifications to discover the direction and type of relationship. In a perfect world of 1-to-1 relationships, we would expect all cells of subset A in condition X to match with 100% confidence to subset B in condition Y. In that particular case the summed probability equal 2 and there would be zero difference in confidence of one classifier to its matching classifier. In our imperfect world we might instead see 90% of cells of subset A in condition X to matching with > 85% confidence to subset B in condition Y, and only 30% of subset B in condition Y matching with > 85% confidence. From this discrepancy we can infer that subset A may be a cell state in condition X that is layered on top of

a base state B in condition Y. Low confidence in both directions indicates subsets unique to a particular condition.

2nd Approach – Training Random Forest Model

After these realizations we trained random forest classifiers for each cell type in each disease condition using SciKit-Learn v0.22.2, with the intent to classify each cell to the subset in the opposed dataset the cell is most similar to (Pedregosa et al., 2011). For each cell type we optimized a classifier for accuracy using grid-based search tuning number of trees, depth, number of features, criterion, and min samples per leaf with 5-fold cross validation for each set of tuning parameters. We never observed full overfitting where the accuracy on test folds began to drop with increased size of model, but we did quickly find diminishing returns as we increased model size. For simplicity and because optimal tuning parameters were robust to overfitting, we chose to use the same largest model parameters for all models (number of trees = 500, depth = 200, number of features = sqrt, criterion = gini, min samples per leaf = 1). Our initial training rounds found accuracies in the mid 60%. A definite increase from the NN classifier, but not high enough for us to be confident in the results. The main issue we eventually determined to be the uneven class distributions (far more cells in subset A than subset B). This caused the smaller subsets to be under trained. To compensate we up-sampled with replacement each subset within the training fold to contain at least the 75th quantile number of cells. This single change improved accuracy on the unmodified test fold the most, varying from 5-15% improvement of accuracy, precision and recall across each cell type, and provided accuracies ranging from high 70 to low 90 percent per major cell type.

Applying Random Forest Model

Satisfied with the validation results within each disease condition, we were finally ready to fully run the random forest model across the disease conditions. We trained each random forest model with optimized parameters on all folds of its dataset, then proceeded to get probability predictions for each cell from the disease condition to the trained disease condition. With these class probabilities per cell we could aggregate for each disease condition by taking the mean class probabilities for each group, leaving us with 2 *n* by *m* table where *n* equals the number of subset groups and *m* equals the number of subset classes in the opposing disease condition. Using the mean probabilities for the group allowed more information from the cell level to rise to the aggregated levels than using the individual class prediction alone (computed as the class with

max confidence of cell membership). These tables also provide confidences to all classes which is important for understanding the transverse confidence in both directions.

It is especially important to understand the many-to-one relationships between disease conditions and find where a base cell state becomes layered in additional expression profiles, as these are the exact cases where we can infer the underlying signaling patterns that diversify or concentrate cell state profiles. In diverse splitting of a subset across disease we can start to understand the heterogeneity of patient response to treatment as it becomes clear which particular cell profiles are correlated with strong and poor response. To gain insight to these changes, we care about where there is strong confidence in both directions and where there is strong confidence in only one direction. The simplest method to calculate these is to separately take the sum of the pairwise prediction confidences and the difference. We call the sum of confidences the correspondence of a subset, and the difference the bias.

Visualizing Correspondence and Bias

We plot these metrics on a dot plot where each possible connection is laid out on a grid. For each dot we set the size to match the correspondence, and color the dot based on the bias, such that a perfect match would appear as a large white circle. A more unidirectional match would be tinted darker in the color matching the disease condition with more confidence. Matches with more bias tend to indicate a subset matching a base cell state but also expressing some additional gene modules. To aid the human eye on picking up the major patterns we filter to only show the top 10% highest correspondences. This parameter was chosen after looking at the distribution of correspondence scores and selecting the majority of the right tail of the distribution. It keeps the strongest matches in both ways and keeps the strongest in highly biased matches. To also aid the human eye we perform a hierarchical clustering using cosine distance and complete linkage on the prediction confidences and compute an optimal ordering based on the cosine distances using the “cba” package in R: https://cran.r-project.org/web/packages/cba/index.html. This allows us to sort subsets on the rows and columns such that subsets that get predicted similarly are next to each other. From this visualization we are able to easily discern which are the subsets FGID that split into many phenotypes within pediCD from high correspondence and bias, which subsets don’t change phenotype much at all based on high correspondence and low bias, and which are the subsets are potentially unique to a disease condition based on very low correspondence and bias.

1499

1500 **Association of Cell Subsets to anti-TNF Response**

1501 Compositional differences are an important metric for understanding the baseline differences that
1502 prognose a patent's response to treatment. We measure these differences with proportional
1503 enrichment of particular cell subsets within each patient, and finding the significantly reproducible
1504 enrichments across disease. As an extreme toy example we might find that subset A cells
1505 comprise as much 80% of cells sampled in one condition whereas they might only comprise 30%
1506 in a different condition. This type of compositional analysis is highly affected by the number and
1507 choice of subsets included, and the sampling depth per patient (how may cells are collected). The
1508 first factor is controlled by the confidence in our clustering and using computationally optimized
1509 parameters. We further control this factor by limiting analysis of compositional shifts of cell states
1510 to within major cell types. This isolates the chance of error from affecting the entire analysis and
1511 allows us to gain a more direct biological insight of the rise and fall of particular cell states in the
1512 context of similar subsets. We control the second factor of sampling depth differences by
1513 computing a normalized cell count score per patient of the form (ncells in subset / ncells in
1514 patient's major cell type) * 1e6. This score provides us with the number of cells expected per
1515 million.

1516

1517 **Mann-Whitney tests**

1518 We input our cells per million score into a two-sample Wilcoxon test in base R, which is equivalent
1519 to the Mann-Whitney rank score test. We set a significance threshold of p_value < 0.05. We made
1520 5 different pairwise comparisons (FGID vs FR, FGID vs PR, NOA vs FR, NOA vs PR, FR vs PR).
1521 Comparisons between FGID and pediCD groups were determined by finding maximum
1522 correspondence between the disease conditions for each subset.

1523

1524 **Fisher's Exact tests**

1525 A similar compositional analysis to that done with the Mann-Whitney was performed with a
1526 Fisher's Exact test. Do the difference of the tests we input for each subset the number of cells for
1527 that subset against the number of cells not of that subset within the major cell type split on rows
1528 by pairwise comparisons (NOA vs FR, NOA vs PR, FR vs PR). We computed FDR correction of
1529 p_values at major cell type and entire dataset levels and found significance subsets at both levels.
1530 But, most interestingly in comparing the two tests we found that the Mann-Whitney discovered as
1531 significant (pval < 0.05) the portion of cell subsets with largest effect sizes. Understanding our
1532 limited patient number at these within pediCD comparisons and wanting to only report results

most likely to be reproducible biology, we determined to only follow those subsets reported as significant within both Mann-Whitney and Fisher's exact tests.

Principal component analysis of cell frequencies and correlation to clinical metadata

Cell frequencies were calculated per patient for cell subsets (i.e. end clusters) within parent cell types and cell subsets (i.e. end clusters) within all cells as CPM = ((count/sum(count)) * 1e6. Principal component analysis (PCA) was performed on the resulting patient x CPM matrices using the R package *stats*::prcomp(., scale=TRUE). Variance explained per PC was calculated as std^2/sum(std^2). PCA loadings per patient and per cell subset were extracted from the prcomp() result. PCA1 and PCA2 from the total PCA x patient and from each celltype's PCA x patient matrix were correlated with clinical metadata using Spearman rank correlation as calculated by the R package *stats*::cor.test(., method='spearman'). P values were recorded from the cor.test() call, and FDR was calculated using R's *fdrtool*::fdrtool(p.values, statistic="pvalue"). For combined celltype PCA's, patient x CPM tables were concatenated before PCA.

Gene set enrichment analysis

Fold changes between patients responding or not responding to anti-TNF therapy from RISK and E-MTAB-7604 cohorts were calculated with Seurat (v4.0.3) (Haberman et al., 2014; Hao et al., 2021) and DESeq2 (v1.30.1) (Love et al., 2014) packages, respectively. GSEA analysis was performed using the fgsea R package (v1.16.0) (Korotkevich et al., 2021). Genes with similar fold changes were preranked in a random order. The code for this analysis can be found in the GitHub repository https://jo-m-lab.github.io/3p-PREDICT-Paper/4_GSEA/PREDICT_GSEA_final.html.

Pseudotime analysis of expression landscape

The micrograin structure found through hierarchical tiered clustering is vital for being able to directly compare like cells across disease conditions, and find significant changes in phenotype and composition within individual subsets. It is also vital to understand how those like subsets relate to each other within a disease condition and how the larger macrograin structure differs across conditions. This macrograin structure can be explored through the gradients of gene expression among cells of a major type. Pseudotime and RNA-velocity are both excellent tools for exploring these gradients. For both tools, the choice of genes directly determines the structure found within the dimensional reduction, and thus what genes are chosen as significantly location

specific within the resulting landscape of cells. for our purposes, as we knew we would be exploring a single cell lineage, and exploring the relationships of cell states within that space, we required for our dimensional reduction the genes common to that space. We selected genes by performing differential expression between the major cell type and all other cell types within that disease. We took the outer union of those genes. Then removed genes from the list found to be differentially expressed between disease conditions at the major cell type level. From these genes we performed PCA to 50 principal components and then computed a UMAP reduction to 2 components. This selection process allows the dimensional reduction to find smooth gradients between cells and provided a common space for cells of multiple disease conditions to exist.

From this common expression landscape we utilized Monocle3 https://cole-trapnell-lab.github.io/monocle3/ (Cao et al., 2019) to extract a best estimate linear path through the space, which we calculated a diffusion pseudotime on allowing use to numerically estimate the distribution of cells within the expression landscape. To compute the significance of changes in that distribution we used a permutation test of Hellinger distance between distributions. At each of 10,000 permutations we shuffled the group ordering within the comparison pair. We performed this test five times for comparisons between FGID vs FR, FGID vs PR, NOA vs FR, NOA vs PR and FR vs PR. Our threshold was set as Bonferroni corrected p_value < 0.05.

FGID and pediCD integration using STACAS

Integration of T cells from the FGID and pediCD datasets (n = 29640 and 38031, respectively) was performed using the STACAS package (v1.1.0) (Andreatta and Carmona, 2021) Sankey plot was created using RAWGraphs 2.0 beta (https://github.com/rawgraphs) (Mauri et al., 2017).

Differential expression testing

To calculate differential expression between FR and PR groups, for each subset with a least 50 cells in each condition we used a Wilcoxon test thresholded to 0.05 Bonferroni corrected p-value and down sampled using the “max.cells.per.ident” argument within Seurat’s `FindMarkers` function to a maximum of 10000 cells. The limits on minimum and maximum number of cells were chosen to mitigate issues with comparisons between disproportionate populations and computational efficiency. There does still exist 2 orders of magnitude between the minimum and maximum; however, the subsets most of interest and reported through the manuscript are within the same order of magnitude

There are noted spillover effects within the expression tests. We observe ubiquitous contamination of genes such as *IGHA1*, *IGHG1* and *DEFA5* across all cell types and subsets. These genes are routinely found as enriched within more severe inflammation, beyond even this dataset. This is a real effect, but less than useful for understanding driving factors within individual cell subsets. So, we focus on significant differentially expressed genes that also have a high pct.cells.expressing.in/ pct.cells.expressing.out ratio.

General Statistical Testing

Parameters such as sample size, number of replicates, number of independent experiments, measures of center, dispersion, and precision (mean +/- SEM) and statistical significances are reported in Figures and Figure Legends. A p-value less than 0.05 was considered significant. Where appropriate, a Bonferroni or FDR correction was used to account for multiple tests, as noted in the figure legends or Methods. All statistical tests corresponding to differential gene expression are described above and completed using R language for Statistical Computing.

Table 1. Demographics and clinical characteristics of patients analyzed on PREDICT.
* indicates significance detected between Crohn's Disease (CD) and Functional Gastrointestinal Disorder (FGID).

		Crohn's Disease	FGID	P-values
		n = 14	n = 13	
Demographics				
	Age at diagnosis (year)	12.5 (9-19)	16 (10-18)	0.096
	Female sex	47% (9 patients)	46% (6 patients)	>0.99
	Caucasian	79% (11 patients)	84% (11 patients)	
	African American	0% (0 patients)	8% (1 patient)	
	Native American or Alaskan Native	0% (0 patients)	0% (0 patients)	
	Asian	7% (1 patient)	0% (0 patients)	
	Mixed race	14% (2 patients)	0% (0 patients)	
	Other	0% (0 patients)	8% (1 patient)	
Growth				
	Weight Z score	-0.61 (-3.31 to 2.31)	0.12 (-1.34 to 2.54)	0.12
	Height Z score	0.09 (-2.04 to 2.32)	0.01 (-1.82 to 2.3)	0.86
	BMI Z score	-0.44 (-3.74 to 2.1)	0.2 (-2.15 to 2.4)	0.23
Disease location/activity				
	Location			
	L1	0% (0 patients)		
	L2	0% (0 patients)		
	L3	93% (13 patient)		
	P modifier	14% (2 patients)		
	Behavior			
	B1	93% (13 patient)		
	B2	0% (0 patients)		
	B3	0% (0 patients)		
	Severity in wPCDAI			
Laboratory evaluation				
	CRP (mg/dl)	3.9 ± 0.8 (0.8 to 8)	0.7 ± 0.1 (0.1 to 1)	0.0019*
	ESR (mm/hr)	35.8 ± 6.2 (7 to 69)	5.2 ± 0.9 (2 to 8)	<0.0001*
	Hgb (g/dl)	11.2 ± 0.5	13.1 ± 0.5	0.0018*

		(8.3 to 13.7)	(9.4 to 15.5)	
	PLT (k/mm3)	424.4 ± 38.8 (176 to 704)	274.2 ± 19.3 (186 to 433)	0.0213*
	WBC (k/mm3)	7.1 ± 0.4 (4.6 to 10.1)	7.1 ± 0.6 (4.2 to 11.5)	0.914
	Albumin (g/dl)	3.5 ± 0.2 (2.5 to 4.6)	4.5 ± 0.1 (4.3 to 5)	0.0023*

Table 2. Demographics and clinical characteristics of Crohn's Disease cohorts.
* indicates significance detected between

		Partial Responder	**Full Responder**	**Not on anti-TNF**
		n = 5	n = 5	n = 4
Demographics and follow-up				
	Age at diagnosis (year)	12 (10 to 14)	12 (9 to 14)	15.5 (9-19)
	Female sex	40% (2 patients)	40% (2 patients)	50% (2 patients)
	Caucasian	80% (4 patients)	80% (4 patients)	50% (2 patients)
	African American	0% (0 patient)	0% (0 patient)	0% (0 patient)
	Native American or Alaskan Native	0% (0 patient)	0% (0 patient)	0% (0 patient)
	Asian	0% (0 patient)	0% (0 patient)	0% (0 patient)
	Mixed race	20% (1 patient)	0% (0 patient)	25% (1 patient)
	Other	0% (0 patient)	20% (1 patient)	0% (0 patient)
Disease location/activity				
	Location			
	L1	0% (0 patient)	0% (0 patient)	0% (0 patient)
	L2	0% (0 patient)	0% (0 patient)	0% (0 patient)
	L3	100% (5 patients)	100% (5 patients)	75% (3 patients)
	P modifier	20% (1 patient)	0% (0 patients)	25% (1 patient)
	Behavior			
	B1	100% (5 patients)	100% (5 patients)	75% (3 patients)
	B2	0% (0 patients)	0% (0 patients)	0% (0 patients)
	B3	0% (0 patients)	0% (0 patients)	0% (0 patients)
	Severity in wPCDAI	65 (35 to 77.5)	65 (22.5 to 87.5)	26.3 (15 to 57.5)
Medications				
	Mesalamine	0% (0 patient)	20% (1 patients)	50% (2 patients)
	Azathioprine	20% (1 patient)	20% (1 patient)	25% (1 patient)
	Methotrexate	80% (4 patient)	60% (3 patient)	0% (0 patient)

	Anti - TNF	100% (5 patients)	100% (5 patients)	0% (0 patient)
	Nutritional therapy	20% (1 patient)	20% (1 patient)	25% (1 patient)
	Steroids	60% (3 patients)	60% (3 patients)	50% (2 patients)
Median time since diagnosis				
	Median	31.7 (29.6 to 37.2)	36.3 (29.6 to 37.2)	32 (24.2 to 33.7)

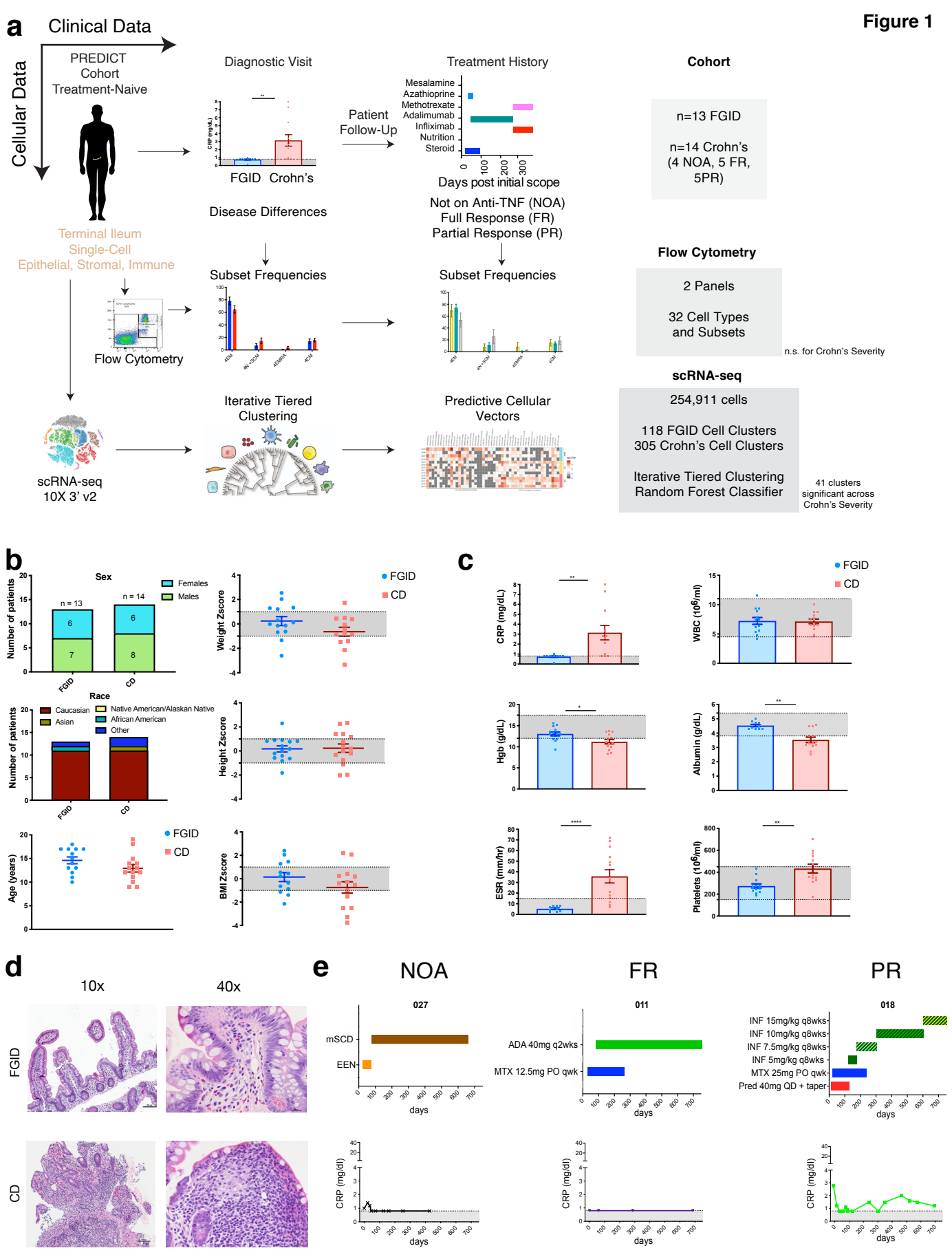


Figure 1
a
Clinical Data
Cellular Data
PREDICT
Cohort
Treatment-Naive
Terminal Ileum
Single-Cell
Epithelial, Stromal, Immune
Flow Cytometry
scRNA-seq
10X 3' v2
Diagnostic Visit
FGID
Crohn's
Disease Differences
Subset Frequencies
Iterative Tiered
Clustering
Patient
Follow-Up
Treatment History
Mesalamine
Azathioprine
Methotrexate
Adalimumab
Infliximab
Nutrition
Steroid
Days post initial scope
Not on Anti-TNF (NOA)
Full Response (FR)
Partial Response (PR)
Subset Frequencies
Predictive Cellular
Vectors
Cohort
n=13 FGID
n=14 Crohn's
(4 NOA, 5 FR,
5PR)
Flow Cytometry
2 Panels
32 Cell Types
and Subsets
n.s. for Crohn's Severity
scRNA-seq
254,911 cells
118 FGID Cell Clusters
305 Crohn's Cell Clusters
Iterative Tiered Clustering
Random Forest Classifier
41 clusters
significant across
Crohn's Severity
b
Sex
Females
Males
n = 13
n = 14
Number of patients
FGID
CD
Race
Caucasian
Asian
Native American/Alaskan Native
African American
Other
Age (years)
Weight Zscore
Height Zscore
BMI Zscore
c
CRP (mg/dL)
WBC (10^6/ml)
Hgb (g/dL)
Albumin (g/dL)
ESR (mm/hr)
Platelets (10^6/ml)
d
10x
40x
FGID
CD
e
NOA
FR
PR
027
011
018
mSCD
EEN
ADA 40mg q2wks
MTX 12.5mg PO qwk
INF 15mg/kg q8wks
INF 10mg/kg q8wks
INF 7.5mg/kg q8wks
INF 5mg/kg q8wks
MTX 25mg PO qwk
Pred 40mg QD + taper
days
CRP (mg/dl)

Main Figure Titles and Legends

Figure 1. PREDICT Study Design with Patient Diagnostic Criteria and Histopathology

a. Study overview depicting clinical and cellular measurements from 13 functional gastrointestinal disorder (FGID) patients and 14 pediatric Crohn's disease (pediCD) patients. Terminal ileum biopsies were isolated at a treatment-naïve diagnostic visit, and pediCD patients were followed up to determine their anti-TNF response and categorized as not on anti-TNF (NOA), Full Response (FR), or Partial Response (PR) (see **Methods**). Two panels of flow cytometry allowed for relative frequency quantification of 32 cell types and subsets, and 10X 3' v2 single-cell RNA-sequencing (scRNA-seq) captured 254,911 total cells including 118 FGID and 305 pediCD end clusters through an iterative tiered clustering (ITC) approach, ARBOL (see **Methods**).
b. Demographic data, weight, height, and BMI for cohort (see **Table 1** and **Supplemental Figure 1.**
c. Clinical inflammatory laboratory values for cohort (see **Table 1** and **Supplemental Figure 1**).
d. Representative histopathology of FGID (top) and pediCD (bottom) at 10x (scale bar = 100um) and 40x (scale bar = 20um) magnification.
e. Representative treatment history and clinical inflammatory parameters used for determination of NOA, FR and PR status (see **Methods**, **Table 1**, and **Supplemental Figure 1;** ADA: adalimumab, INF: infliximab; MTX: methotrexate; Pred: prednisone; mSCD: modified specific carbohydrate diet; EEN: exclusive enteral nutrition).

Figure 2

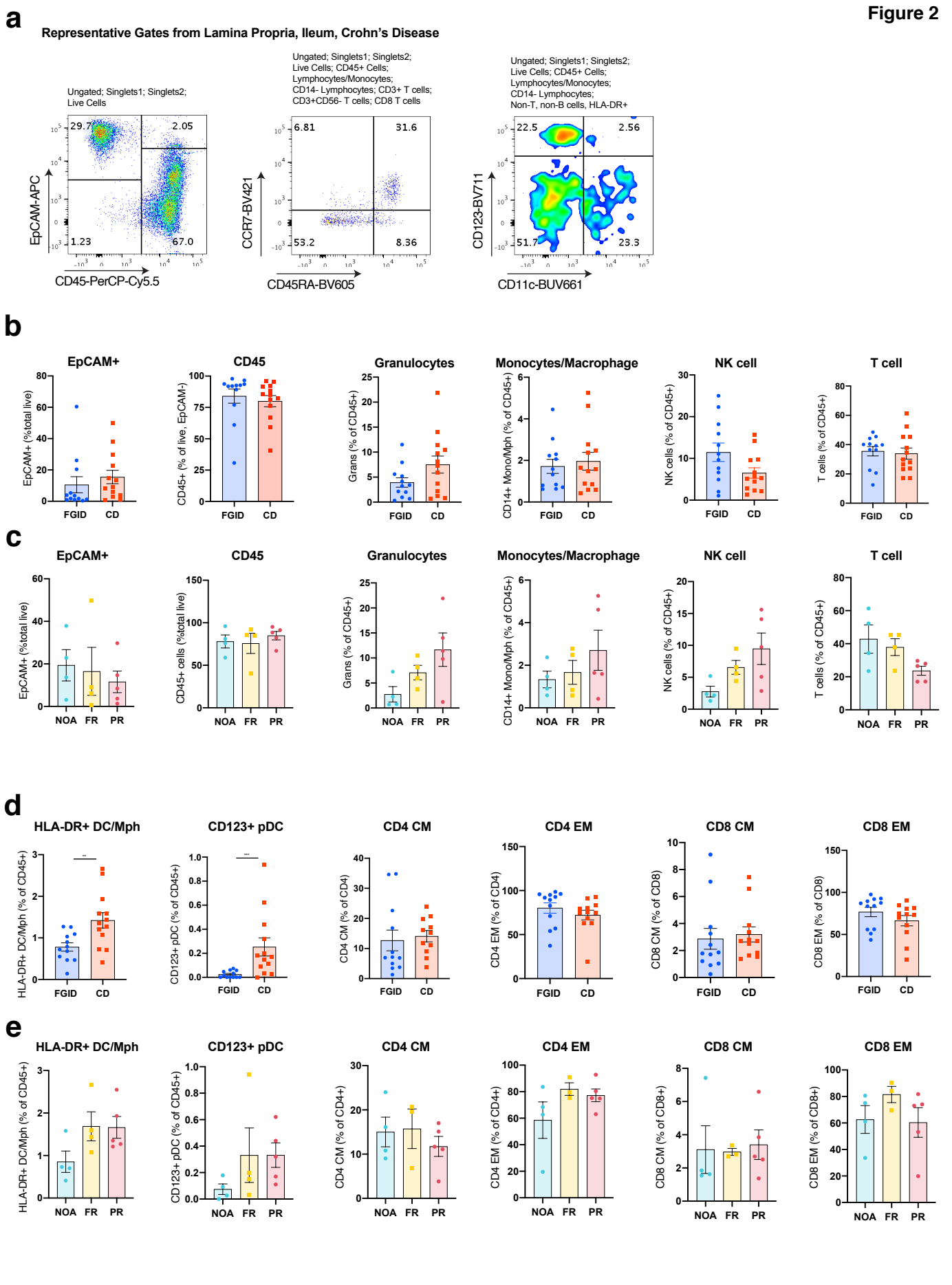

a
Representative Gates from Lamina Propria, Ileum, Crohn's Disease
Ungated; Singlets1; Singlets2; Live Cells
29.7
2.05
1.23
67.0
EpCAM-APC
CD45-PerCP-Cy5.5
Ungated; Singlets1; Singlets2; Live Cells; CD45+ Cells; Lymphocytes/Monocytes; CD14- Lymphocytes; CD3+ T cells; CD3+CD56- T cells; CD8 T cells
6.81
31.6
53.2
8.36
CCR7-BV421
CD45RA-BV605
Ungated; Singlets1; Singlets2; Live Cells; CD45+ Cells; Lymphocytes/Monocytes; CD14- Lymphocytes; Non-T, non-B cells, HLA-DR+
22.5
2.56
51.7
23.3
CD123-BV711
CD11c-BUV661
b
EpCAM+
EpCAM+ (%total live)
CD45
CD45+ (% of live, EpCAM-)
Granulocytes
Grans (% of CD45+)
Monocytes/Macrophage
CD14+ Mono/Mph (% of CD45+)
NK cell
NK cells (% of CD45+)
T cell
T cells (% of CD45+)
FGID
CD
c
EpCAM+
EpCAM+ (%total live)
CD45
CD45+ cells (%total live)
Granulocytes
Grans (% of CD45+)
Monocytes/Macrophage
CD14+ Mono/Mph (% of CD45+)
NK cell
NK cells (% of CD45+)
T cell
T cells (% of CD45+)
NOA
FR
PR
d
HLA-DR+ DC/Mph
HLA-DR+ DC/Mph (% of CD45+)
CD123+ pDC
CD123+ pDC (% of CD45+)
CD4 CM
CD4 CM (% of CD4)
CD4 EM
CD4 EM (% of CD4)
CD8 CM
CD8 CM (% of CD8)
CD8 EM
CD8 EM (% of CD8)
FGID
CD
e
HLA-DR+ DC/Mph
HLA-DR+ DC/Mph (% of CD45+)
CD123+ pDC
CD123+ pDC (% of CD45+)
CD4 CM
CD4 CM (% of CD4+)
CD4 EM
CD4 EM (% of CD4+)
CD8 CM
CD8 CM (% of CD8+)
CD8 EM
CD8 EM (% of CD8+)
NOA
FR
PR

Figure 2. Flow Cytometry of Ileal Biopsies Does Not Reveal Significant Changes in Cell Composition in FGID vs. pediCD or across the pediCD Treatment Response Spectrum

a. Representative flow cytometry end gates for selected cell subsets (left: epithelial and hematopoietic; middle: naïve and effector T cells; right: pDCs and antigen-presenting cells) from single-cell dissociated samples from one terminal ileum biopsy for pediCD patients (see **Supplemental Figure 2** for full gating strategy).
b. Fractional composition of selected cell subsets of CD45+ cells from 13 FGID and 14 pediCD patients (error bars are s.e.m).
c. Fractional composition of selected cell subsets of CD45+ cells from 4 NOA, 5 FR and 5 PR patients.
d. Fractional composition of dendritic, pDC, central memory (CM) and effector memory (EM) CD4+ and CD8+ cells from 13 FGID vs 14 pediCD patients. Dendritic cells and pDC plotted as percentage of CD45+ cells. CM/EM CD4+ and CD8+ cells plotted as percentage of total CD4+ and CD8+ cells, respectively. $p < 0.05$ by Mann-Whitney for pediCD versus FGID and 1-way ANOVA for pediCD cohorts).
e. Fractional composition of dendritic cells, pDCs, central memory (CM) and effector memory (EM) CD4+ and CD8+ cells from 4 NOA, 5FR, and 5 PR patients. Graphs plotted as in **d**.

Figure 3

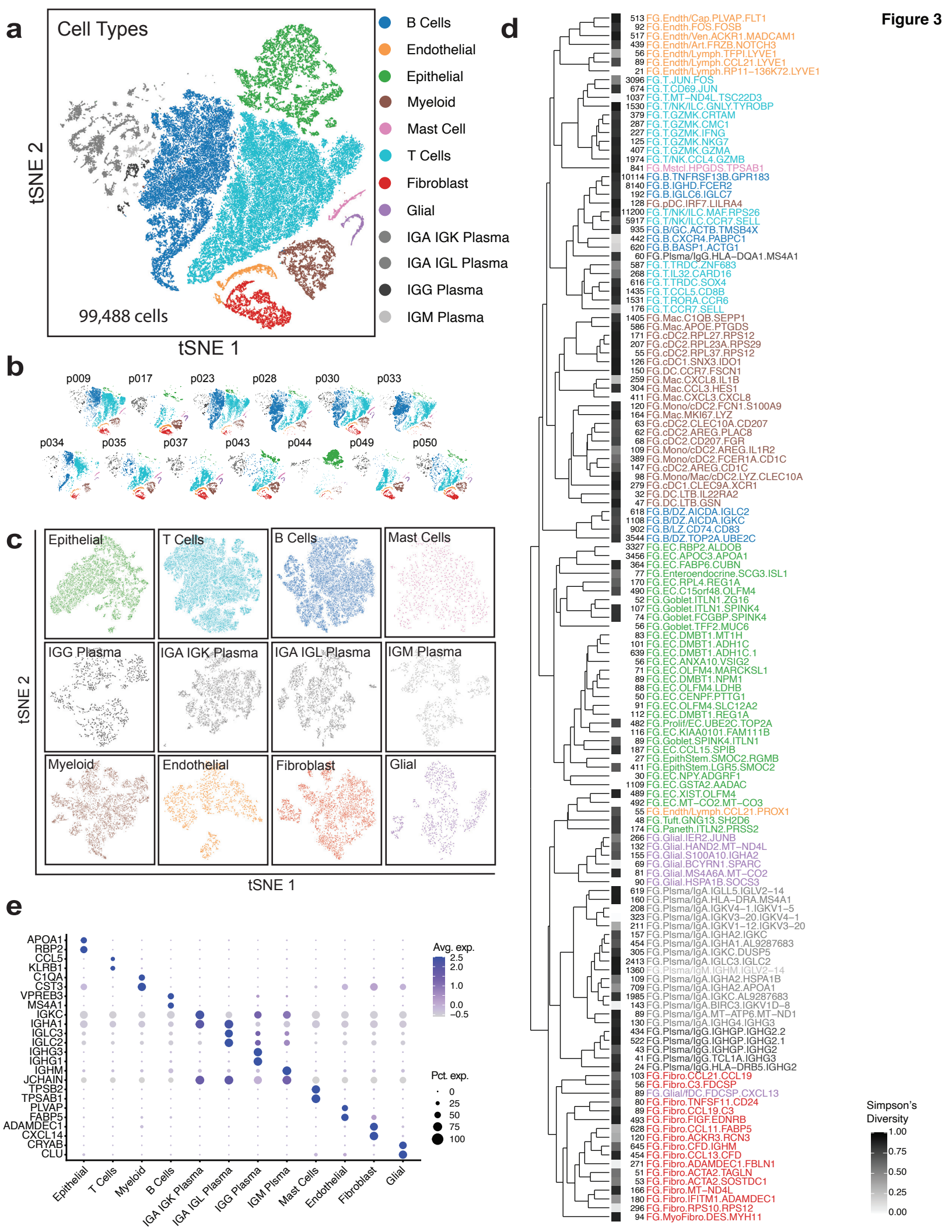

a
Cell Types
99,488 cells
tSNE 2
tSNE 1
B Cells
Endothelial
Epithelial
Myeloid
Mast Cell
T Cells
Fibroblast
Glial
IGA IGK Plasma
IGA IGL Plasma
IGG Plasma
IGM Plasma
b
p009
p017
p023
p028
p030
p033
p034
p035
p037
p043
p044
p049
p050
c
Epithelial
T Cells
B Cells
Mast Cells
IGG Plasma
IGA IGK Plasma
IGA IGL Plasma
IGM Plasma
Myeloid
Endothelial
Fibroblast
Glial
tSNE 2
tSNE 1
e
APOA1
RBP2
CCL5
KLRB1
C1QA
CST3
VPREB3
MS4A1
IGKC
IGHA1
IGLC3
IGLC2
IGHG3
IGHG1
IGHM
JCHAIN
TPSB2
TPSAB1
PLVAP
FABP5
ADAMDEC1
CXCL14
CRYAB
CLU
Epithelial
T Cells
Myeloid
B Cells
IGA IGK Plasma
IGA IGL Plasma
IGG Plasma
IGM Plsma
Mast Cells
Endothelial
Fibroblast
Glial
Avg. exp.
2.5
2.0
1.0
0.0
−0.5
Pct. exp.
0
25
50
75
100
d
513 FG.Endth/Cap.PLVAP.FLT1
92 FG.Endth.FOS.FOSB
517 FG.Endth/Ven.ACKR1.MADCAM1
439 FG.Endth/Art.FRZB.NOTCH3
56 FG.Endth/Lymph.TFPI.LYVE1
89 FG.Endth/Lymph.CCL21.LYVE1
21 FG.Endth/Lymph.RP11−136K72.LYVE1
3096 FG.T.JUN.FOS
674 FG.T.CD69.JUN
1037 FG.T.MT−ND4L.TSC22D3
1530 FG.T/NK/ILC.GNLY.TYROBP
379 FG.T.GZMK.CRTAM
287 FG.T.GZMK.CMC1
227 FG.T.GZMK.IFNG
125 FG.T.GZMK.NKG7
407 FG.T.GZMK.GZMA
1974 FG.T/NK.CCL4.GZMB
841 FG.Mstcl.HPGDS.TPSAB1
10114 FG.B.TNFRSF13B.GPR183
8140 FG.B.IGHD.FCER2
192 FG.B.IGLC6.IGLC7
128 FG.pDC.IRF7.LILRA4
11200 FG.T/NK/ILC.MAF.RPS26
5917 FG.T/NK/ILC.CCR7.SELL
935 FG.B/GC.ACTB.TMSB4X
442 FG.B.CXCR4.PABPC1
620 FG.B.BASP1.ACTG1
60 FG.Plsma/IgG.HLA−DQA1.MS4A1
587 FG.T.TRDC.ZNF683
268 FG.T.IL32.CARD16
616 FG.T.TRDC.SOX4
1435 FG.T.CCL5.CD8B
1531 FG.T.RORA.CCR6
176 FG.T.CCR7.SELL
1405 FG.Mac.C1QB.SEPP1
586 FG.Mac.APOE.PTGDS
171 FG.cDC2.RPL27.RPS12
207 FG.cDC2.RPL23A.RPS29
55 FG.cDC2.RPL37.RPS12
126 FG.cDC1.SNX3.IDO1
150 FG.DC.CCR7.FSCN1
259 FG.Mac.CXCL8.IL1B
304 FG.Mac.CCL3.HES1
411 FG.Mac.CXCL3.CXCL8
120 FG.Mono/cDC2.FCN1.S100A9
164 FG.Mac.MKI67.LYZ
63 FG.cDC2.CLEC10A.CD207
62 FG.cDC2.AREG.PLAC8
68 FG.cDC2.CD207.FGR
109 FG.Mono/cDC2.AREG.IL1R2
389 FG.Mono/cDC2.FCER1A.CD1C
147 FG.cDC2.AREG.CD1C
98 FG.Mono/Mac/cDC2.LYZ.CLEC10A
279 FG.cDC1.CLEC9A.XCR1
32 FG.DC.LTB.IL22RA2
47 FG.DC.LTB.GSN
618 FG.B/DZ.AICDA.IGLC2
1108 FG.B/DZ.AICDA.IGKC
902 FG.B/LZ.CD74.CD83
3544 FG.B/DZ.TOP2A.UBE2C
3327 FG.EC.RBP2.ALDOB
3456 FG.EC.APOC3.APOA1
364 FG.EC.FABP6.CUBN
77 FG.Enteroendocrine.SCG3.ISL1
170 FG.EC.RPL4.REG1A
490 FG.EC.C15orf48.OLFM4
52 FG.Goblet.ITLN1.ZG16
107 FG.Goblet.ITLN1.SPINK4
74 FG.Goblet.FCGBP.SPINK4
56 FG.Goblet.TFF2.MUC6
83 FG.EC.DMBT1.MT1H
101 FG.EC.DMBT1.ADH1C
639 FG.EC.DMBT1.ADH1C.1
56 FG.EC.ANXA10.VSIG2
71 FG.EC.OLFM4.MARCKSL1
89 FG.EC.DMBT1.NPM1
88 FG.EC.OLFM4.LDHB
50 FG.EC.CENPF.PTTG1
91 FG.EC.OLFM4.SLC12A2
112 FG.EC.DMBT1.REG1A
482 FG.Prolif/EC.UBE2C.TOP2A
116 FG.EC.KIAA0101.FAM111B
89 FG.Goblet.SPINK4.ITLN1
187 FG.EC.CCL15.SPIB
27 FG.EpithStem.SMOC2.RGMB
411 FG.EpithStem.LGR5.SMOC2
30 FG.EC.NPY.ADGRF1
1109 FG.EC.GSTA2.AADAC
489 FG.EC.XIST.OLFM4
492 FG.EC.MT−CO2.MT−CO3
55 FG.Endth/Lymph.CCL21.PROX1
48 FG.Tuft.GNG13.SH2D6
174 FG.Paneth.ITLN2.PRSS2
266 FG.Glial.IER2.JUNB
132 FG.Glial.HAND2.MT−ND4L
155 FG.Glial.S100A10.IGHA2
69 FG.Glial.BCYRN1.SPARC
81 FG.Glial.MS4A6A.MT−CO2
90 FG.Glial.HSPA1B.SOCS3
619 FG.Plsma/IgA.IGLL5.IGLV2−14
160 FG.Plsma/IgA.HLA−DRA.MS4A1
208 FG.Plsma/IgA.IGKV4−1.IGKV1−5
323 FG.Plsma/IgA.IGKV3−20.IGKV4−1
211 FG.Plsma/IgA.IGKV1−12.IGKV3−20
157 FG.Plsma/IgA.IGHA2.IGKC
454 FG.Plsma/IgA.IGHA1.AL9287683
305 FG.Plsma/IgA.IGKC.DUSP5
2413 FG.Plsma/IgA.IGLC3.IGLC2
1360 FG.Plsma/IgM.IGHM.IGLV2−14
109 FG.Plsma/IgA.IGHA2.HSPA1B
709 FG.Plsma/IgA.IGHA2.APOA1
1985 FG.Plsma/IgA.IGKC.AL9287683
143 FG.Plsma/IgA.BIRC3.IGKV1D−8
89 FG.Plsma/IgA.MT−ATP6.MT−ND1
130 FG.Plsma/IgA.IGHG4.IGHG3
434 FG.Plsma/IgG.IGHGP.IGHG2.2
522 FG.Plsma/IgG.IGHGP.IGHG2.1
43 FG.Plsma/IgG.IGHGP.IGHG2
41 FG.Plsma/IgG.TCL1A.IGHG3
24 FG.Plsma/IgG.HLA−DRB5.IGHG2
103 FG.Fibro.CCL21.CCL19
56 FG.Fibro.C3.FDCSP
89 FG.Glial/fDC.FDCSP.CXCL13
80 FG.Fibro.TNFSF11.CD24
89 FG.Fibro.CCL19.C3
493 FG.Fibro.FIGF.EDNRB
628 FG.Fibro.CCL11.FABP5
120 FG.Fibro.ACKR3.RCN3
645 FG.Fibro.CFD.IGHM
454 FG.Fibro.CCL13.CFD
271 FG.Fibro.ADAMDEC1.FBLN1
51 FG.Fibro.ACTA2.TAGLN
53 FG.Fibro.ACTA2.SOSTDC1
166 FG.Fibro.MT−ND4L
180 FG.Fibro.IFITM1.ADAMDEC1
296 FG.Fibro.RPS10.RPS12
94 FG.MyoFibro.DES.MYH11
Simpson's Diversity
1.00
0.75
0.50
0.25
0.00

Figure 3. A Comprehensive Cell Atlas of Terminal Ileum in Non-Inflammatory FGID

a. tSNE of 99,488 single-cells isolated from terminal ileal biopsies of 13 FGID patients. Colors represent major cell type groups determined via Louvain clustering with resolution set by optimized silhouette score.
b. tSNE as in **a** with individual patients plotted. For specific proportions please see **Supplemental Figure 4**.
c. tSNE of each major cell type which was used as input into iterative tiered clustering (ITC).
d. Hierarchical clustering of complete FGID data set with input clusters determined based on results of ITC and performed on the median expression of 4,428 pairwise differentially expressed genes, using complete linkage and distance calculated with Pearson correlation, between each end cell cluster. Simpson's Index of Diversity represented as 1-Simpson's where 1 (black) indicates equivalent richness of all patients in that cluster, and 0 (white) indicates a completely patient-specific subset. Numbers represent the number of cells in that cluster. Names of subsets are determined by Disease.CellType.GeneA.GeneB as in **Methods**.
e. Dot plot of 2 defining genes for each cell type. Dot size represents fraction of cells expressing the gene, and color intensity represents binned count-based expression level (log(scaled UMI+1)) amongst expressing cells. All cluster defining genes are provided in **Supplemental Tables 5 to 7**.

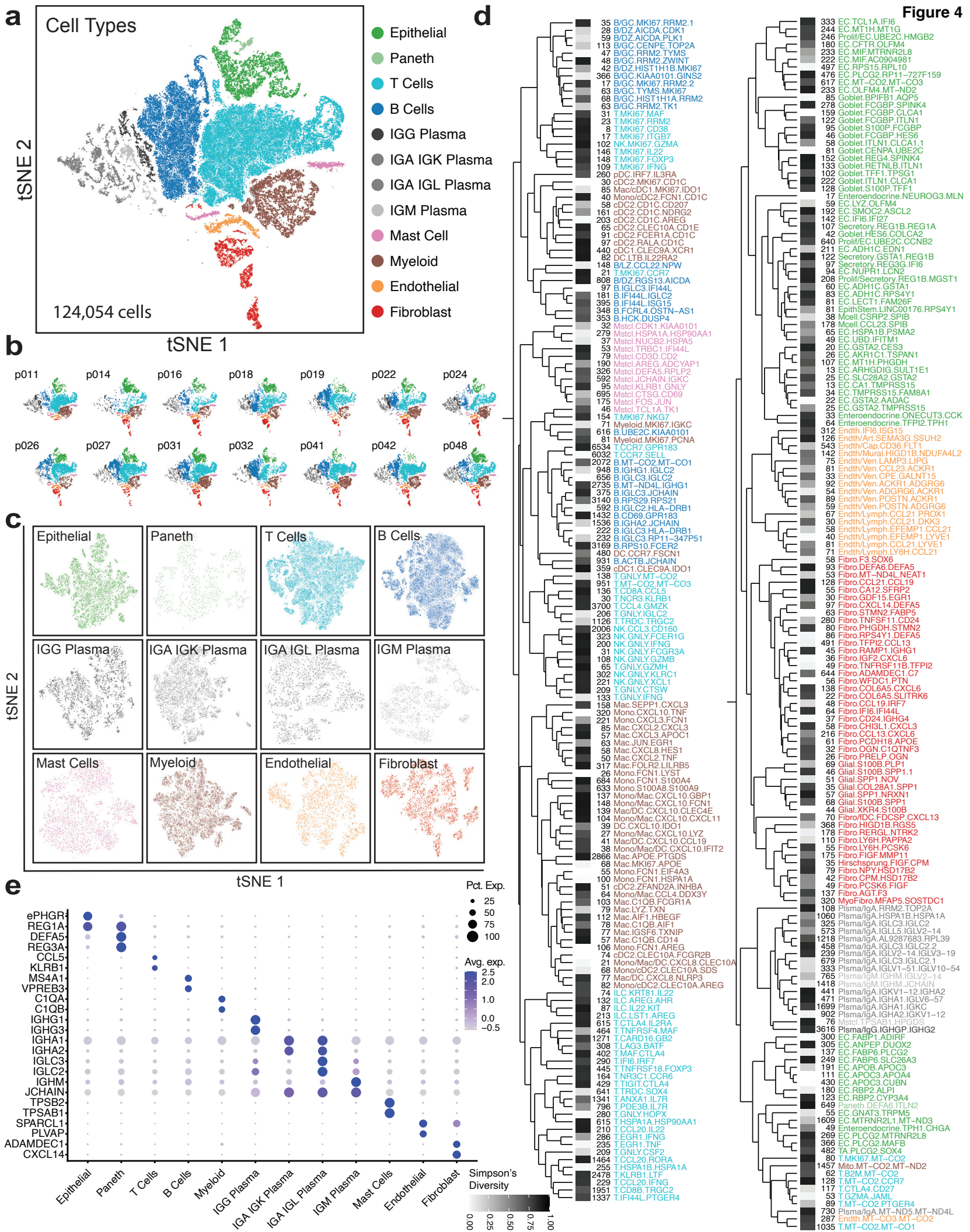


Figure 4
a
Cell Types
124,054 cells
tSNE 2
tSNE 1
Epithelial
Paneth
T Cells
B Cells
IGG Plasma
IGA IGK Plasma
IGA IGL Plasma
IGM Plasma
Mast Cell
Myeloid
Endothelial
Fibroblast
b
p011
p014
p016
p018
p019
p022
p024
p026
p027
p031
p032
p041
p042
p048
c
Epithelial
Paneth
T Cells
B Cells
IGG Plasma
IGA IGK Plasma
IGA IGL Plasma
IGM Plasma
Mast Cells
Myeloid
Endothelial
Fibroblast
tSNE 2
tSNE 1
d
e
Pct. Exp.
25
50
75
100
Avg. exp.
2.5
2.0
1.0
0.0
−0.5
ePHGR
REG1A
DEFA5
REG3A
CCL5
KLRB1
MS4A1
VPREB3
C1QA
C1QB
IGHG1
IGHG3
IGHA1
IGHA2
IGLC3
IGLC2
IGHM
JCHAIN
TPSB2
TPSAB1
SPARCL1
PLVAP
ADAMDEC1
CXCL14
Epithelial
Paneth
T Cells
B Cells
Myeloid
IGG Plasma
IGA IGK Plasma
IGA IGL Plasma
IGM Plasma
Mast Cells
Endothelial
Fibroblast
Simpson's Diversity
0.00
0.25
0.50
0.75
1.00

Figure 4. A Comprehensive Cell Atlas of Terminal Ileum in pediCD

a. tSNE of 124,054 single-cells isolated from terminal ileal biopsies of 14 pediCD patients. Colors represent major cell type groups determined via Louvain clustering with resolution set by optimized silhouette score.
b. tSNE as in **a** with individual patients plotted. For specific proportions please see **Supplemental Figure 4**.
c. tSNE of each major cell type which was used as input into iterative tiered clustering (ITC).
d. Hierarchical clustering of complete pediCD data set with input clusters determined based on results of ITC, and performed on the median expression of 4,428 pairwise differentially expressed genes, using complete linkage and distance calculated with Pearson correlation, between each end cell cluster. Simpson's Index of Diversity represented as 1-Simpson's where 1 (black) indicates equivalent richness of all patients in that cluster, and 0 (white) indicates a completely patient-specific subset. Numbers represent the number of cells in that cluster. Names of subsets are determined by Disease.CellType.GeneA.GeneB as in **Methods**.
e. Dot plot of 2 defining genes for each cell type. Dot size represents fraction of cells expressing the gene, and color intensity represents binned count-based expression level (log(scaled UMI+1)) amongst expressing cells. All cluster defining genes are provided in **Supplemental Tables 8 to 10**.

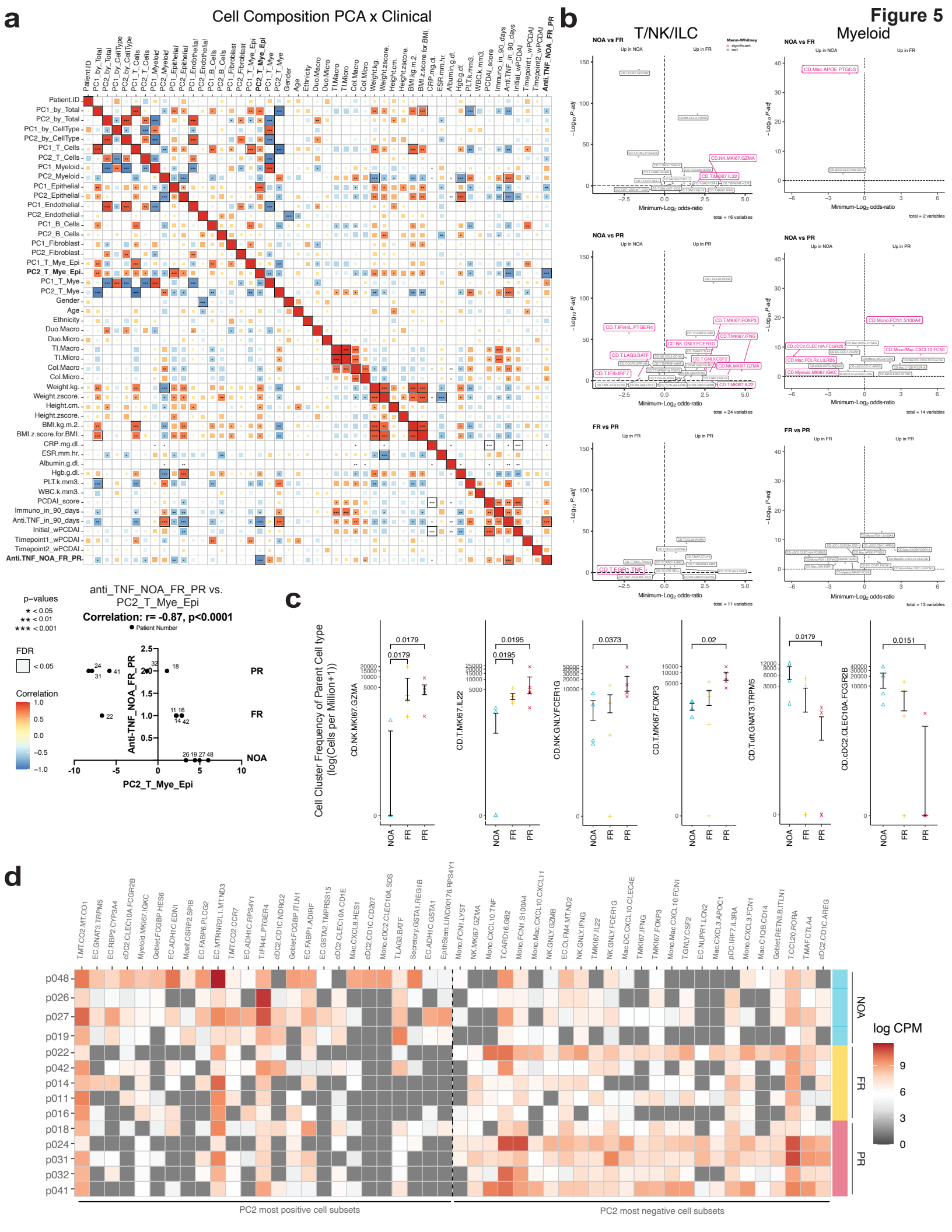


Figure 5
a
Cell Composition PCA x Clinical
b
T/NK/ILC
Myeloid
NOA vs FR
NOA vs PR
FR vs PR
Minimum-Log2 odds-ratio
−Log10 P-adj
c
anti_TNF_NOA_FR_PR vs. PC2_T_Mye_Epi
Correlation: r= -0.87, p<0.0001
Patient Number
p−values
FDR
Correlation
Anti-TNF_NOA_FR_PR
PC2_T_Mye_Epi
PR
FR
NOA
Cell Cluster Frequency of Parent Cell type (log(Cells per Million+1))
CD.NK.MKI67.GZMA
CD.T.MKI67.IL22
CD.NK.GNLY.FCER1G
CD.T.MKI67.FOXP3
CD.Tuft.GNAT3.TRPM5
CD.cDC2.CLEC10A.FCGR2B
d
PC2 most positive cell subsets
PC2 most negative cell subsets
log CPM

Figure 5. A Collective Cell Vector in pediCD Reveals Predictive Axes of Disease Trajectory and Treatment Response

a. Spearman rank correlation heatmap of principal components calculated from the frequencies of each end cluster per main cell type together with clinical metadata. Correlation is represented by both the intensity and size of the box and those which are FDR < 0.05 have a bounding box (inset highlights the specific correlation between PC2 of the T, Myeloid, Epithelial cell frequency analysis with anti-TNF response).
b. Volcano plots for T/NK/ILC and myeloid cell clusters between NOA, FR and PR, where named clusters are significant by Fisher's exact test and those in pink are significant by Mann-Whitney U test.
c. Cell cluster frequencies of the parent cell type found to be significant by Mann-Whitney U test between selected clusters (see **Supplemental Figure 7** for all graphs; **Supplemental Table 15**).
d. Heatmap showing cell frequencies per patient of most positive and negative cell subsets of PC2 from PCA performed on T/NK/ILC, myeloid and epithelial cell subsets (**Supplemental Table 16**). Cell subsets are sorted by PC2 score, and patients were sorted by anti-TNF response. Heatmap is not normalized and displaying the log counts-per million of each cell subset normalized per cell type. *Patient p022's response category changed from FR to PR after database lock in December of 2020. No other patient's categorization has changed.

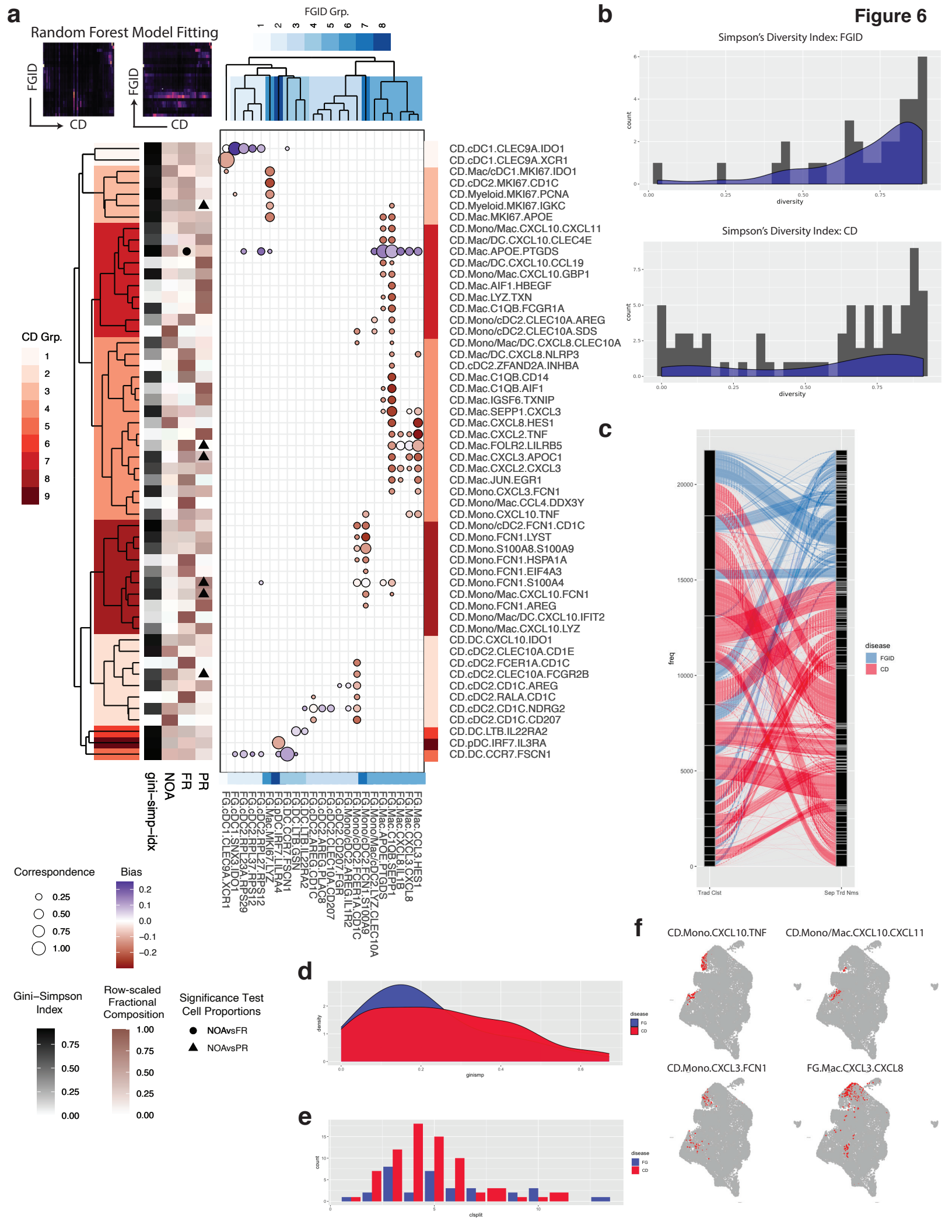


Figure 6
a
Random Forest Model Fitting
FGID
CD
FGID Grp.
CD Grp.
gini–simp–idx
NOA
FR
PR
CD.cDC1.CLEC9A.IDO1
CD.cDC1.CLEC9A.XCR1
CD.Mac/cDC1.MKI67.IDO1
CD.cDC2.MKI67.CD1C
CD.Myeloid.MKI67.PCNA
CD.Myeloid.MKI67.IGKC
CD.Mac.MKI67.APOE
CD.Mono/Mac.CXCL10.CXCL11
CD.Mac/DC.CXCL10.CLEC4E
CD.Mac.APOE.PTGDS
CD.Mac/DC.CXCL10.CCL19
CD.Mono/Mac.CXCL10.GBP1
CD.Mac.AIF1.HBEGF
CD.Mac.LYZ.TXN
CD.Mac.C1QB.FCGR1A
CD.Mono/cDC2.CLEC10A.AREG
CD.Mono/cDC2.CLEC10A.SDS
CD.Mono/Mac/DC.CXCL8.CLEC10A
CD.Mac/DC.CXCL8.NLRP3
CD.cDC2.ZFAND2A.INHBA
CD.Mac.C1QB.CD14
CD.Mac.C1QB.AIF1
CD.Mac.IGSF6.TXNIP
CD.Mac.SEPP1.CXCL3
CD.Mac.CXCL8.HES1
CD.Mac.CXCL2.TNF
CD.Mac.FOLR2.LILRB5
CD.Mac.CXCL3.APOC1
CD.Mac.CXCL2.CXCL3
CD.Mac.JUN.EGR1
CD.Mono.CXCL3.FCN1
CD.Mono/Mac.CCL4.DDX3Y
CD.Mono.CXCL10.TNF
CD.Mono/cDC2.FCN1.CD1C
CD.Mono.FCN1.LYST
CD.Mono.S100A8.S100A9
CD.Mono.FCN1.HSPA1A
CD.Mono.FCN1.EIF4A3
CD.Mono.FCN1.S100A4
CD.Mono/Mac.CXCL10.FCN1
CD.Mono.FCN1.AREG
CD.Mono/Mac/DC.CXCL10.IFIT2
CD.Mono/Mac.CXCL10.LYZ
CD.DC.CXCL10.IDO1
CD.cDC2.CLEC10A.CD1E
CD.cDC2.FCER1A.CD1C
CD.cDC2.CLEC10A.FCGR2B
CD.cDC2.CD1C.AREG
CD.cDC2.RALA.CD1C
CD.cDC2.CD1C.NDRG2
CD.cDC2.CD1C.CD207
CD.DC.LTB.IL22RA2
CD.pDC.IRF7.IL3RA
CD.DC.CCR7.FSCN1
Correspondence
Bias
Gini–Simpson Index
Row-scaled Fractional Composition
Significance Test Cell Proportions
NOAvsFR
NOAvsPR
b
Simpson's Diversity Index: FGID
Simpson's Diversity Index: CD
count
diversity
c
freq
disease
FGID
CD
Trad Clst
Sep Trd Nms
d
density
ginismp
e
count
clsplit
f
CD.Mono.CXCL10.TNF
CD.Mono/Mac.CXCL10.CXCL11
CD.Mono.CXCL3.FCN1
FG.Mac.CXCL3.CXCL8

Figure 6. Random Forest (RF) Classifier Applied to Myeloid Cellular Taxonomies Identifies Correspondence between FGID and pediCD

a. Correspondence between cell subsets from FGID-to-pediCD and pediCD-to-FGID.

   Top left heatmaps: RF probabilities for each cell averaged over subset to gain probability of each FGID matching onto each pediCD subset (left), and pediCD onto FGID (right).

   Bubble plot (center): size = sum(probability matrices) for confidence of predictions, marker color = diff(probability matrices) to show which direction the RF model is more confident on, e.g. more likely for FGID subset to belong to pediCD subset or pediCD subset to belong to FGID subset. Markers are filtered to show the top 10th percentile of correspondence.

   Dendrograms: separated-tiered clustering on prediction probabilities of FGID (blue) and pediCD (red) using complete linkage with correlation distance metric, clusters are cut at height 0.7 (range 0-1).

   Heatmap: 1-Gini-Simpson index based on patient diversity, mono-patient clusters (white), full representation (black). Right 3 columns show row-normalized of frequency of NOA, FR, PR representation in each CD cell subset. Significant differences (Mann-Whitney, alpha=0.05) are marked, triangle NOA vs. PR and circle NOA vs. FR.

b. Distribution of Gini-Simpson's index of patient diversity in FGID (top) and pediCD (bottom) for myeloid cell clusters.

c. Sankey plot comparing joined traditional single-level clustering (left) to disease-separated iterative tiered clustering (right). Each line follows each cell as it moves between in the two cluster sets (back bar split based on cluster identity).

d. Gini-Simpson index on representation of traditional clusters in each of the separated tiered clusters (i.e., from how many of the higher-level clusters does the deep clustering pull). Calculated separately for FGID (blue) and pediCD (red).

e. Similar to **d** but showing the total counts of how many traditional clusters are represented in a single tiered cluster per disease.

f. UMAP of combined Myeloid cells: red shows example end clusters from ITC that are split across the traditional-clustering joint-disease UMAP.

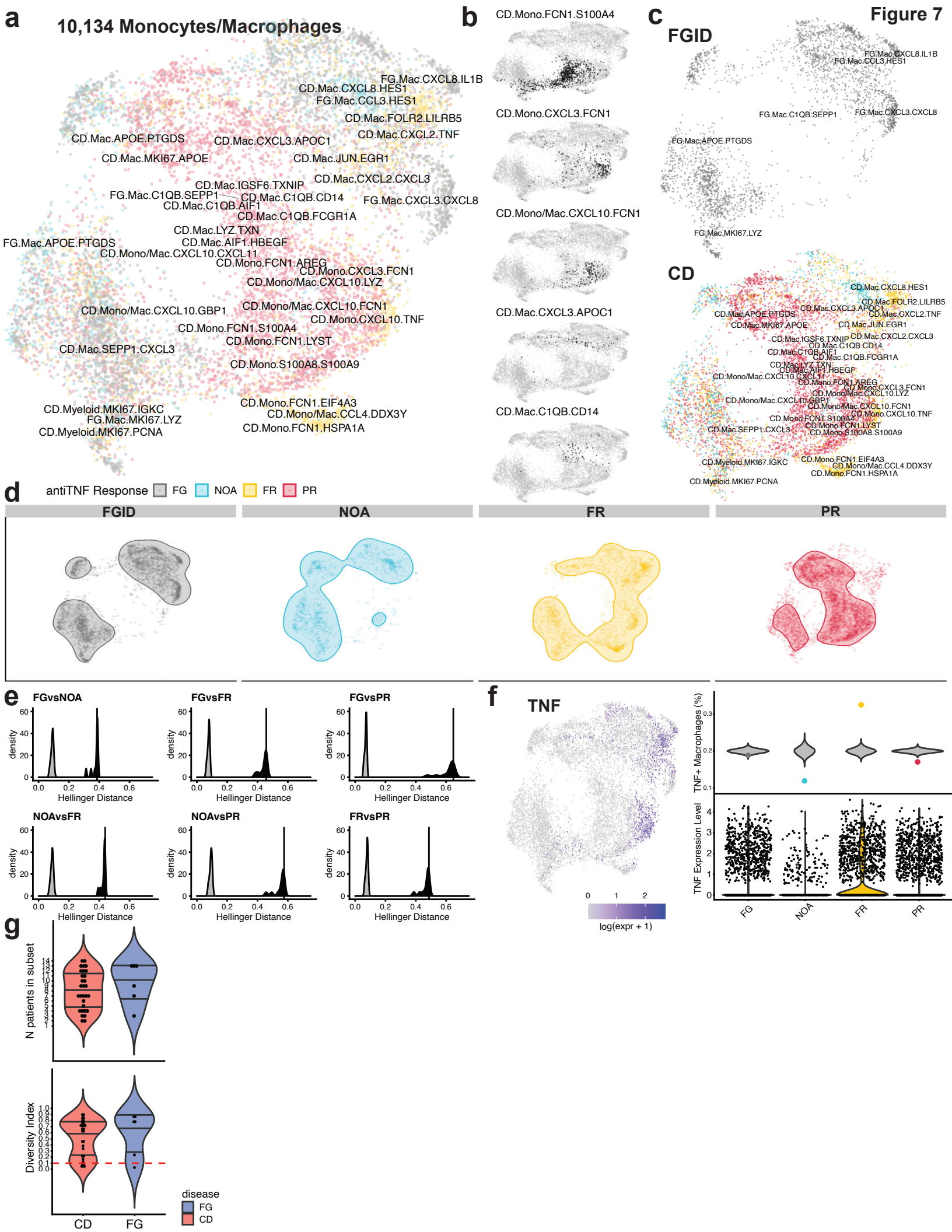


Figure 7
a
10,134 Monocytes/Macrophages
FG.Mac.CXCL8.IL1B
CD.Mac.CXCL8.HES1
FG.Mac.CCL3.HES1
CD.Mac.FOLR2.LILRB5
CD.Mac.CXCL2.TNF
CD.Mac.APOE.PTGDS
CD.Mac.CXCL3.APOC1
CD.Mac.MKI67.APOE
CD.Mac.JUN.EGR1
CD.Mac.CXCL2.CXCL3
CD.Mac.IGSF6.TXNIP
FG.Mac.C1QB.SEPP1
CD.Mac.C1QB.CD14
CD.Mac.C1QB.AIF1
FG.Mac.CXCL3.CXCL8
CD.Mac.C1QB.FCGR1A
CD.Mac.LYZ.TXN
FG.Mac.APOE.PTGDS
CD.Mac.AIF1.HBEGF
CD.Mono/Mac.CXCL10.CXCL11
CD.Mono.FCN1.AREG
CD.Mono.CXCL3.FCN1
CD.Mono/Mac.CXCL10.LYZ
CD.Mono/Mac.CXCL10.GBP1
CD.Mono/Mac.CXCL10.FCN1
CD.Mono.CXCL10.TNF
CD.Mono.FCN1.S100A4
CD.Mono.FCN1.LYST
CD.Mac.SEPP1.CXCL3
CD.Mono.S100A8.S100A9
CD.Mono.FCN1.EIF4A3
CD.Myeloid.MKI67.IGKC
CD.Mono/Mac.CCL4.DDX3Y
FG.Mac.MKI67.LYZ
CD.Mono.FCN1.HSPA1A
CD.Myeloid.MKI67.PCNA
b
CD.Mono.FCN1.S100A4
CD.Mono.CXCL3.FCN1
CD.Mono/Mac.CXCL10.FCN1
CD.Mac.CXCL3.APOC1
CD.Mac.C1QB.CD14
c
FGID
FG.Mac.CXCL8.IL1B
FG.Mac.CCL3.HES1
FG.Mac.C1QB.SEPP1
FG.Mac.CXCL3.CXCL8
FG.Mac.APOE.PTGDS
FG.Mac.MKI67.LYZ
CD
d
antiTNF Response
FG
NOA
FR
PR
FGID
NOA
FR
PR
e
FGvsNOA
FGvsFR
FGvsPR
NOAvsFR
NOAvsPR
FRvsPR
density
Hellinger Distance
f
TNF
log(expr + 1)
TNF+ Macrophages (%)
TNF Expression Level
FG
NOA
FR
PR
g
N patients in subset
Diversity Index
CD
FG
disease
FG
CD

Figure 7: Distinct Distributions of Macrophages Across the pediCD Treatment Response Spectrum Relative to FGID

a. UMAP representation of macrophages (27 patients; 10,134 cells) from FG and pediCD datasets, run across 50 principal components based on 539 genes significantly upregulated (Wilcoxon; p.adj<0.05) in macrophages versus all other cell types and not significantly differentially expressed between FG and pediCD sets. UMAP parameters [min-dist=0.1, N-neighbors=ceiling(sqrt(Ncells)/2)]. Labels are set at median of IQR for each subset. Clinical metadata showing future response to anti-TNF treatment: FGID (grey), NOA (blue), FR (yellow), PR (red).

b. Same UMAP as in **a** colored to isolate single subsets. Subsets chosen based on significant Mann-Whitney tests (**Figure 5; Supplemental Figure 7**), (black) cells from subset, (grey) rest of macrophages.

c. Same UMAP as in **a** separated into FGID and pediCD.

d. Same UMAP as in **a** split into each treatment response group. Shaded area captures 80% most densely populated regions of plot area calculated using 2d KDE estimate from MASS R package.

e. Permutation test results using Hellinger distance to measure if 2 conditions are sampled from the same distribution (0 = complete overlap, 1 = no overlap). Hellinger distance is computed with sqrt(1 - sum(sqrt(kde1*kde2))) with a KDE estimation for each condition group calculated across 1000 points uniformly distributed across plot area, with bandwidth selected using ks::Hpi() function. Black distribution shows test statistic varying min-dist parameter with 11 evenly spaced values between 0.01 and 1. Vertical line shows test statistic using UMAP parameters [min-dist=0.1, Neighbors=ceiling(sqrt(Ncells)), Npcs=50, nDim=2]. Grey distribution shows results of 11,000 permutations to treatment response group varied across same min-dist umap parameters between 0.01 and 1. All tests are significant beyond a 0.001 threshold.

f. Clockwise from left: UMAP of macrophages with color intensity displaying amount of TNF expression based on ((log(scaledUMI+1)). Plot (top) showing fraction of macrophages expressing TNF with colored dots showing fraction of TNF+ cells within each treatment response group and grey violins showing results of 10,000 permutations of treatment response labels. Violin plot (bottom) of ((log(scaledUMI+1)) TNF expression split on treatment response group.

g. Diversity of macrophage clusters in FGID and pediCD: (top) each dot represents a cell subset, y-axis shows how many patients are included within the subset. (bottom) each dot represents a subset, with y position showing (1-Gini-Simpson's Diversity Index), Subsets below red dashed line set at 0.1 diversity were excluded.

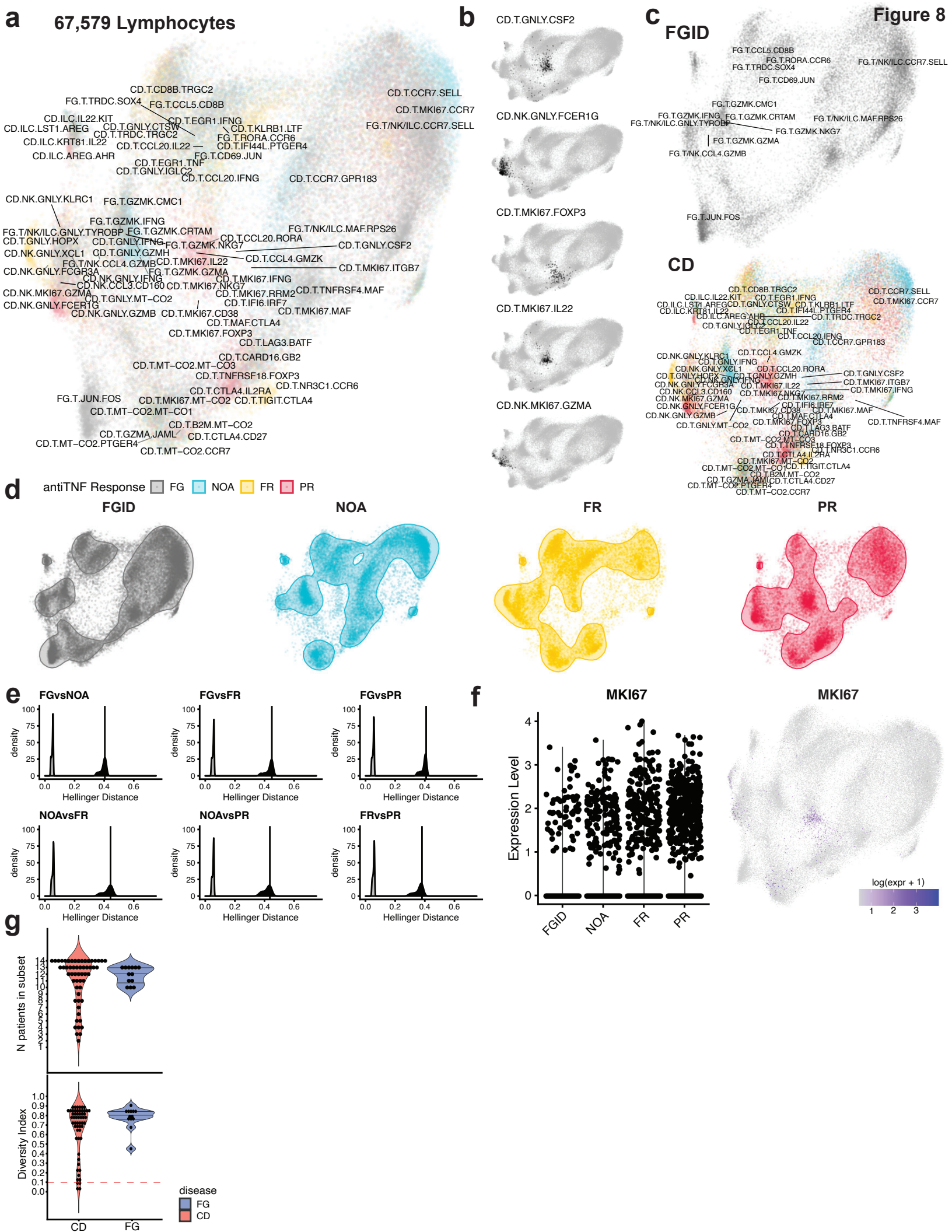


Figure 8
a
67,579 Lymphocytes
b
CD.T.GNLY.CSF2
CD.NK.GNLY.FCER1G
CD.T.MKI67.FOXP3
CD.T.MKI67.IL22
CD.NK.MKI67.GZMA
c
FGID
CD
d
antiTNF Response
FG
NOA
FR
PR
e
FGvsNOA
FGvsFR
FGvsPR
NOAvsFR
NOAvsPR
FRvsPR
density
Hellinger Distance
f
MKI67
Expression Level
FGID
NOA
FR
PR
log(expr + 1)
g
N patients in subset
Diversity Index
disease
FG
CD

Figure 8. Distinct Distributions of Lymphocytes Across the pediCD Treatment Response Spectrum Relative to FGID

a. UMAP representation of T/NK/ILCs (27 patients; 67,579 cells) from FG and CD datasets, run across 50 principal components based on 345 genes significantly upregulated (Wilcoxon; p.adj<0.05) in lymphocytes versus all other cell types and not significantly differentially expressed between FG and pediCD sets. UMAP parameters [min-dist=0.1, N-neighbors=ceiling(sqrt(Ncells)/2)] Labels are set at median of IQR for each subset. Clinical metadata showing future response to anti-TNF treatment: FGID (grey), NOA (blue), FR (yellow), PR (red).

b. Same UMAP as in **a** colored to isolate single subsets. Subsets chosen based on significant Mann-Whitney tests (**Figure 5**), (black) cells from subset, (grey) rest of lymphocytes.

c. Same UMAP as in **a** separated into FG and pediCD.

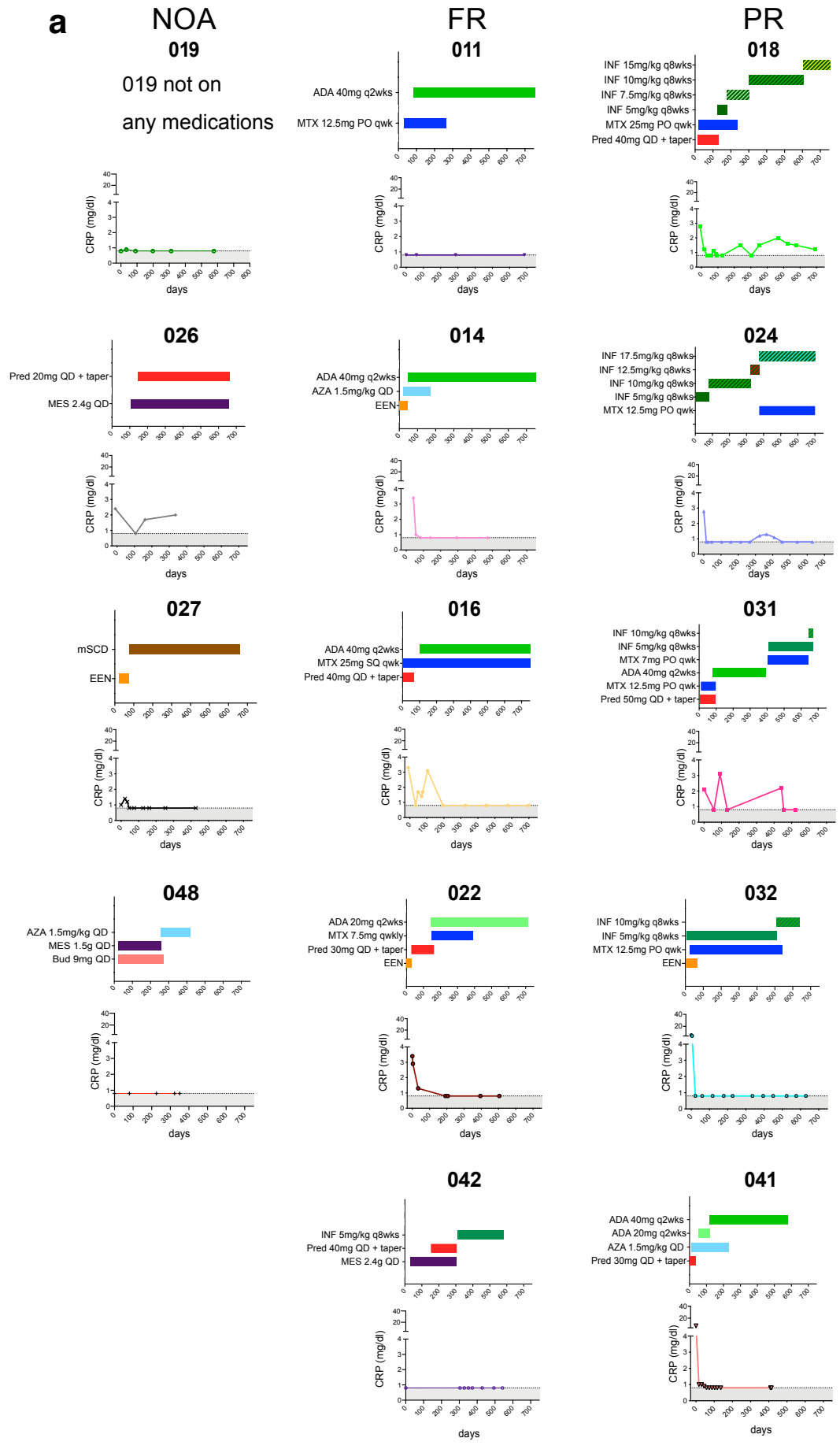
d. Same UMAP as in **a** split into each treatment response group. Shaded area captures 80% most densely populated regions of plot area calculated using 2d KDE estimate from MASS R package.

e. Permutation test results using Hellinger distance to measure if 2 conditions are sampled from the same distribution (0 = complete overlap, 1 = no overlap). Hellinger distance is computed with sqrt(1 - sum(sqrt(kde1*kde2))) with a KDE estimation for each condition group calculated across 1000 points uniformly distributed across plot area, with bandwidth selected using ks::Hpi() function. Black distribution shows test statistic varying min-dist parameter with 11 evenly spaced values between 0.01 and 1. Vertical line shows test statistic using UMAP parameters [min-dist=0.1, Neighbors=ceiling(sqrt(Ncells)), Npcs=50, nDim=2]. Grey distribution shows results of 11,000 permutations to treatment response group varied across same min-dist umap parameters between 0.01 and 1. All tests are significant beyond a 0.001 threshold.

f. Violin plot (left) of ((log(scaledUMI+1)) *MKI67* expression split on treatment response group. UMAP (right) of lymphocytes with color intensity displaying *MKI67* expression based on ((log(scaledUMI+1)) (right).

g. Diversity of lymphocyte clusters in FGID and CD: (top) each dot represents a cell subset, y-axis shows how many patients are included within the subset. (bottom) each dot represents a subset, with y position showing (1-Gini-Simpson's Diversity Index), Subsets below red dashed line set at 0.1 diversity were excluded.

Supplemental Figure 1



Supplemental Figure Titles and Legends

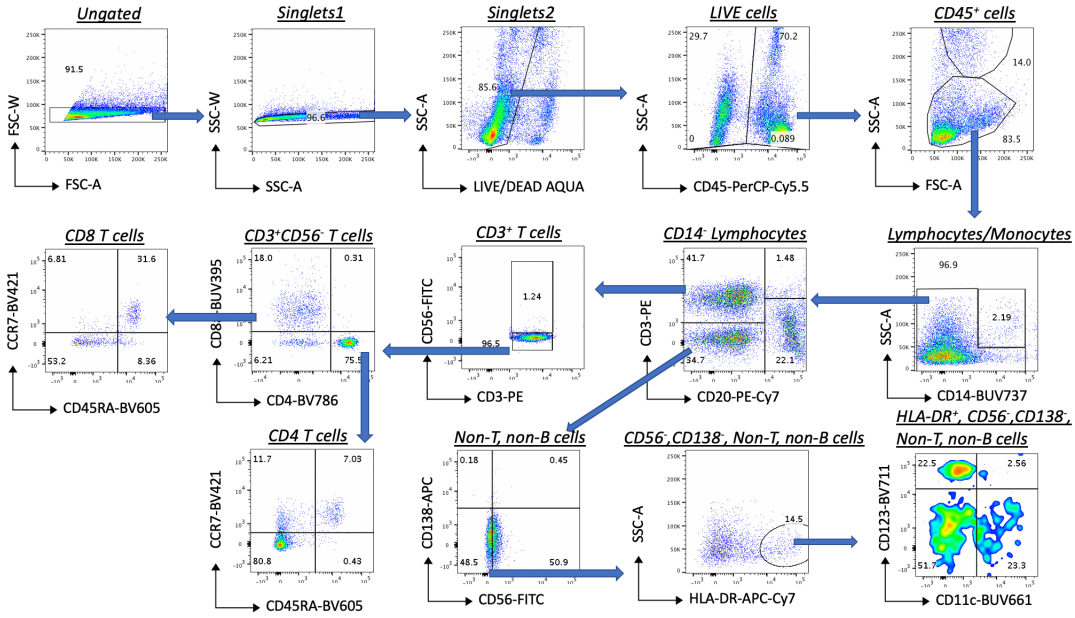
Supplemental Figure 1. Clinical trajectory and treatments for all pediCD patients

a. Representative treatment history and clinical inflammatory parameters used for determination of NOA, FR and PR status for all pediCD patients (see **Methods**, **Table 1**, and **Figure 1;** ADA: adalimumab, INF: infliximab; MES: mesalamine MTX: methotrexate; Pred: prednisone; mSCD: modified specific carbohydrate diet; EEN: exclusive enteral nutrition)

Supplemental Figure 2

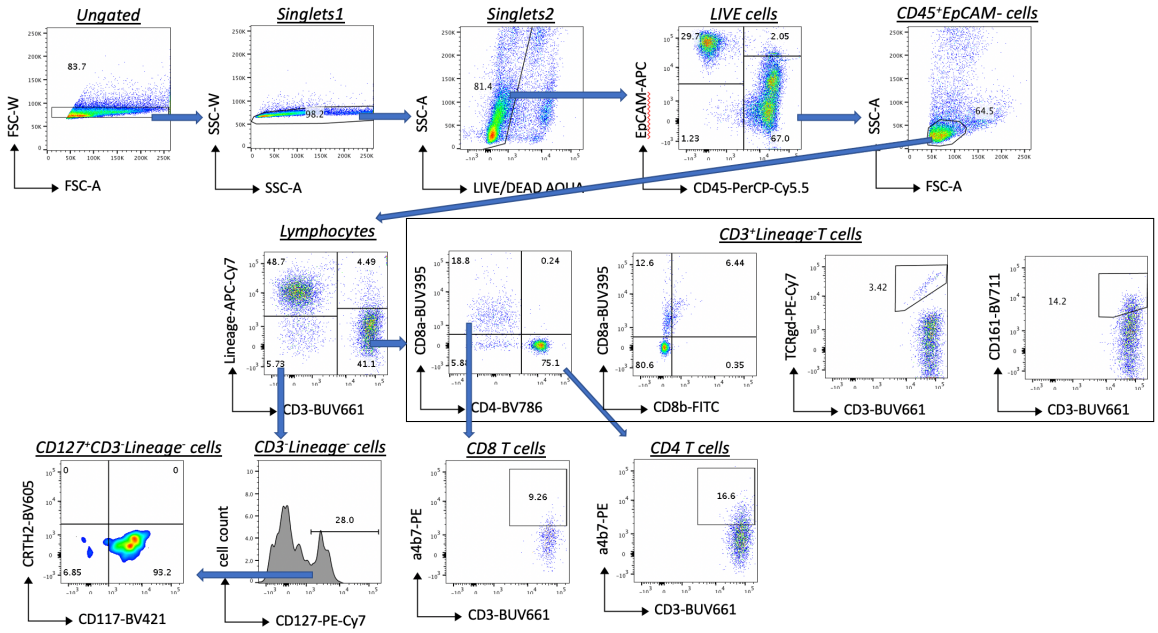
a

Panel 1: Representative Gates from Lamina Propria, Ileum, Crohn's Disease



b

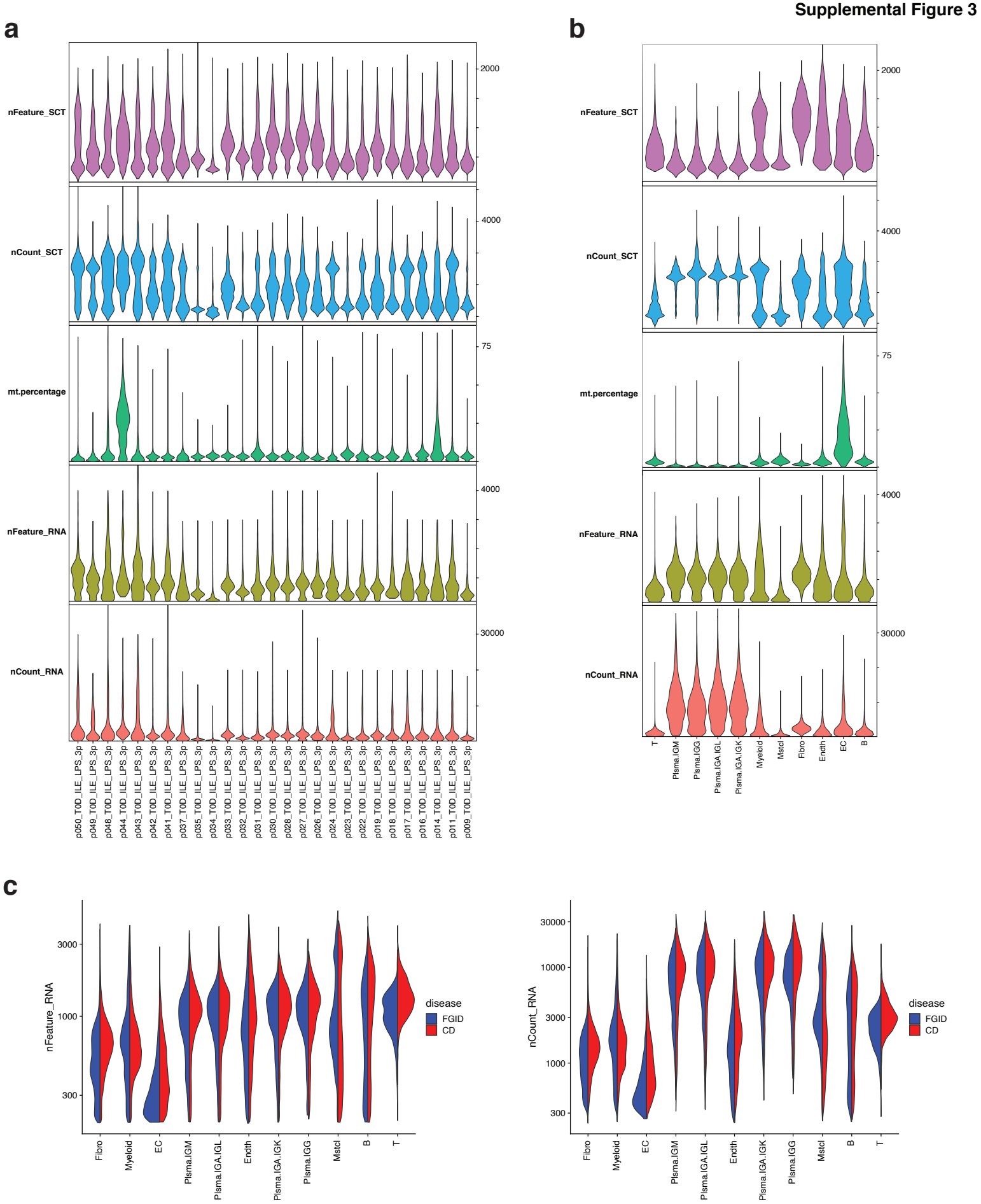
Panel 2: Representative Gates from Lamina Propria, Ileum, Crohn's Disease



Supplemental Figure 2. Representative gating strategies for flow cytometry

a. Representative gating strategy for Panel 1 focused on T cells and myeloid cells, for antibodies please see **Supplemental Table 1**.
b. Representative gating strategy for Panel 2 focused on non-classical T cells and innate lymphoid cells (NB: Lineage = CD14, CD20, CD11c, CD11b, CD56), for antibodies please see **Supplemental Table 1**.

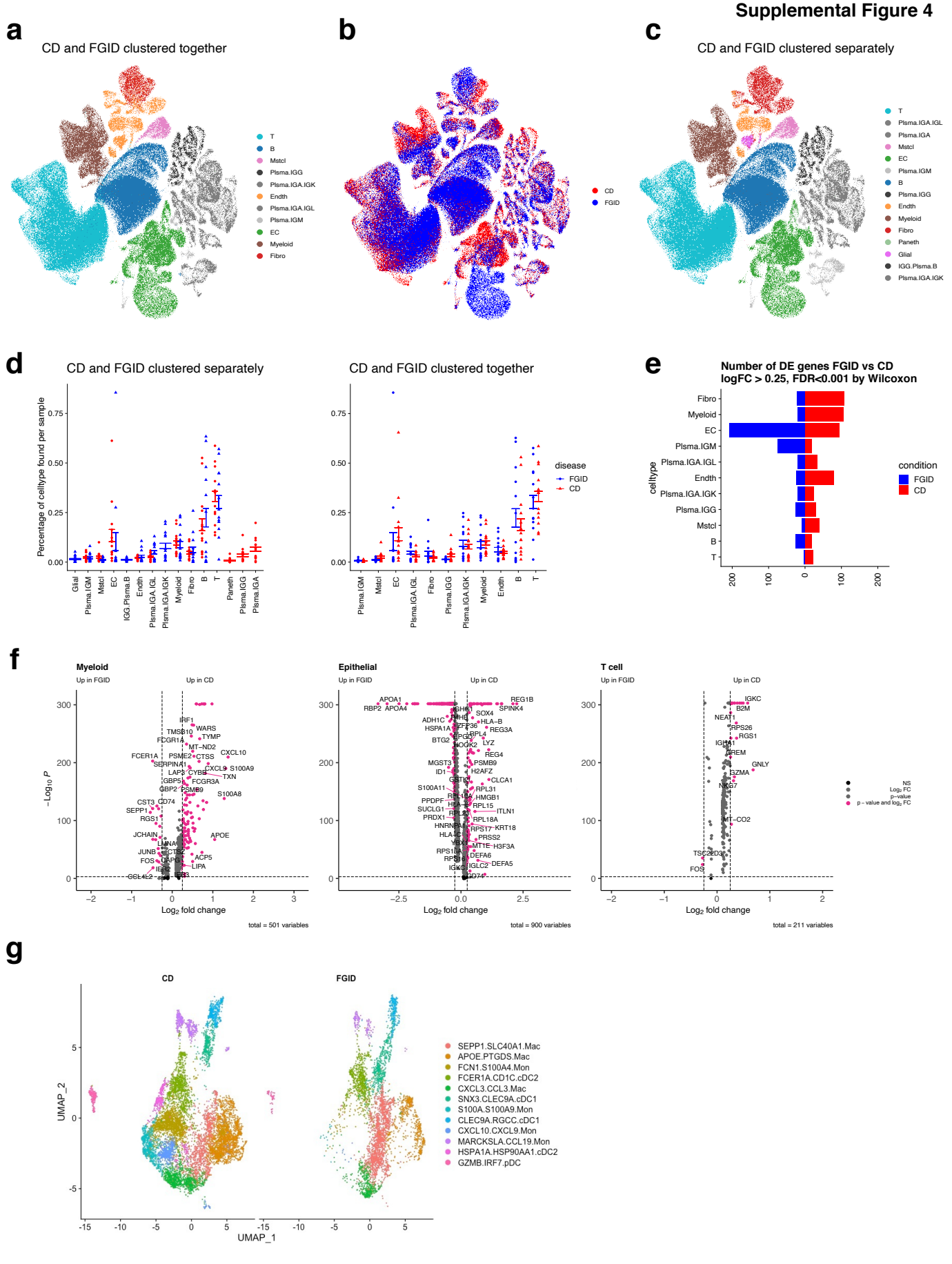
Supplemental Figure 3



Supplemental Figure 3. Comparison of quality control measures reveals similar sequencing depths and gene capture between FGID and pediCD

a. Quality control measures for scRNA-seq of ileal biopsies of 27 patients (13 FGID, 14 pediCD) included in the study. Top two graphs denote total genes (nFeature) and UMIs (nCount) after normalization with SCTransform. Lower graphs denote total genes (nFeature), UMIs (nCount) and mitochondrial read percentage (mt.percentage) of pre-processed 10X 3' v2 single-cell RNA-sequenced samples. Cutoffs for individual samples are represented in **Supplemental Table 2**.
b. Quality control measures as in **a** split by cell type.
c. Comparison of total genes captured (nFeature, left) and total UMIs (nCount, right) between FGID (blue) and pediCD (red) split by cell type.

Supplemental Figure 4

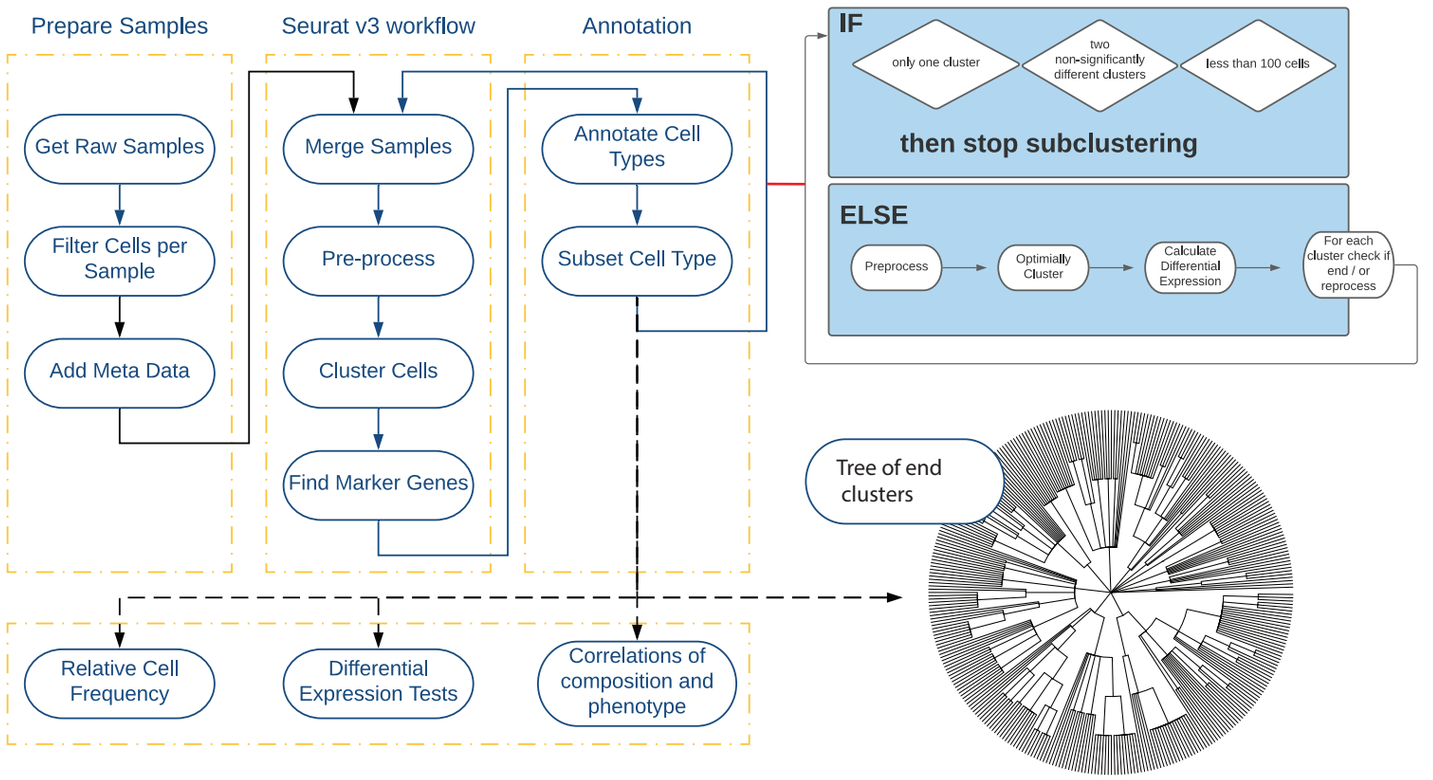


Supplemental Figure 4. Traditional clustering with SCTransform normalization reveals similarities across cell types in FGID and pediCD

a. UMAPs representing one round of clustering of 197,281 single-cells across FGID and pediCD samples. Traditional clustering performed on disease states together. Colors represent major cell types determined by one round of clustering with Seurat RunUMAP parameters (PCs = 1:50, n.neighbors = 50, min.dist = 1). Cell types were assigned based on significantly upregulated marker genes (Wilcoxon; p.adj<0.05) obtained from comparison of specific cell type versus all other cell types.
b. UMAPs as in **a** colored to highlight FGID (blue) and pediCD (red) cells.
c. UMAP as in **a** colored by Tier 1 ITC clusters performed separately for FGID and pediCD.
d. Comparison of cell cluster frequencies between FGID (blue) and pediCD (red). Patient contributions denoted by circles (FGID) and triangles (pediCD).
e. Differentially expressed genes across cell type in FGID vs pediCD determined to be significant by Wilcoxon test (logFC>0.25, FDR<0.001).
f. Volcano plots for Myeloid, Epithelial, T-cell clusters denoting differentially expressed genes in FGID vs. pediCD. Those in pink are significant by Wilcoxon test; **Supplemental Table 4** for full gene lists.
g. UMAPs of jointly clustered pediCD and FGID Myeloid cells.

Supplemental Figure 5

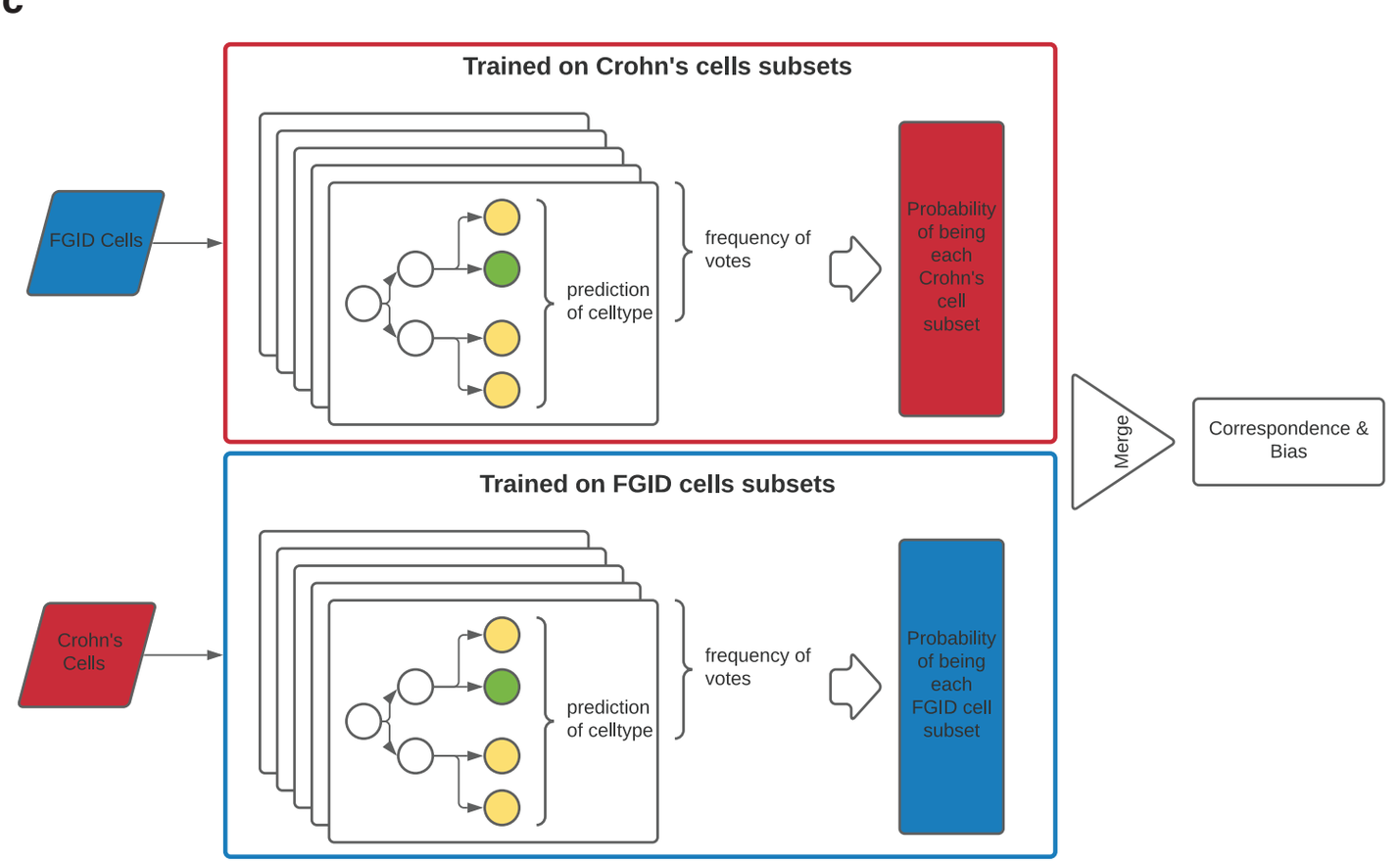
a



b

	CD	FGID
	cells	
Cells passing quality thresholds	139,342	115,569
Included in atlas tSNE	124,054	99,488
Included in ITC tree - with p044 but no doublets	107,432	94,451
Final numbers used for analysis - no p044	107,432	89,849
	clusters	
Raw ITC	514	837
Included in atlas tSNE	425	182
Included in ITC tree - with p044 but no doublets	305	138
Final number used for analysis - no p044	305	118

c

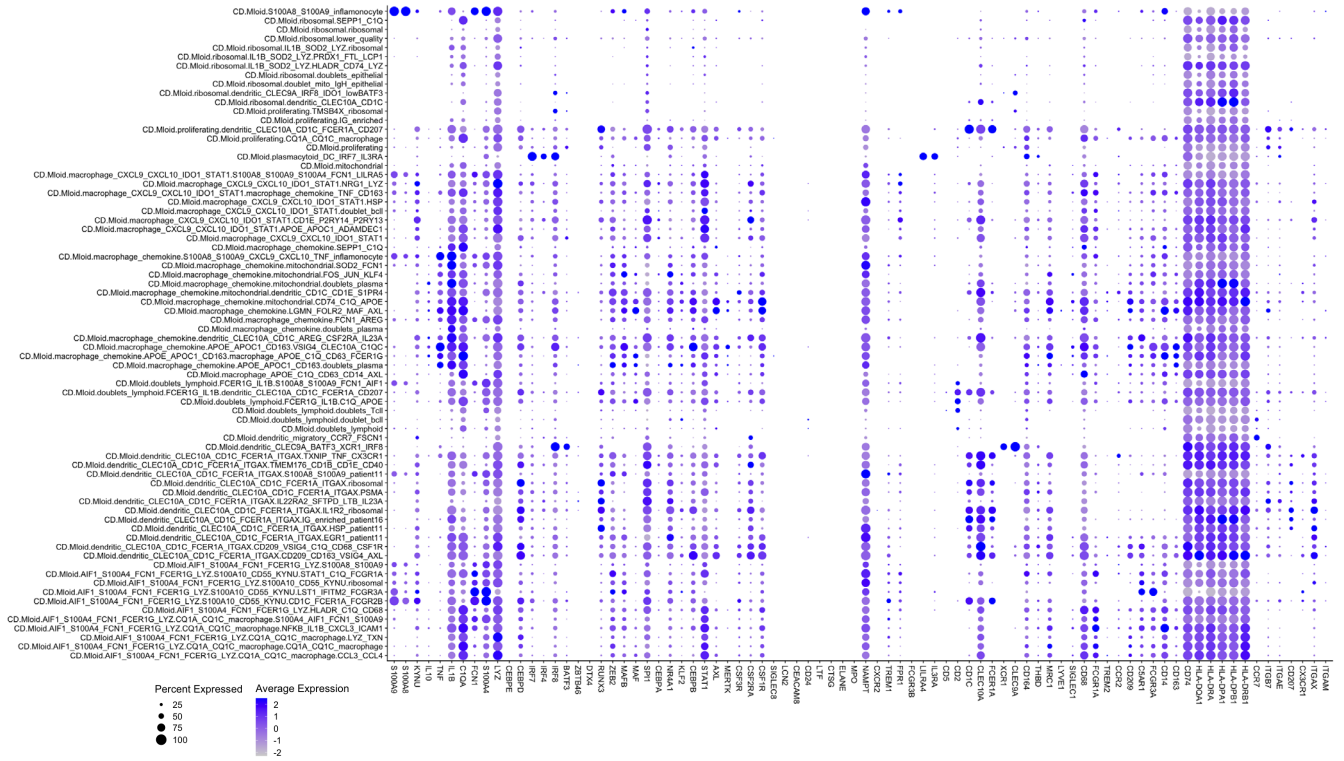


Supplemental Figure 5. Schematic for iterative tiered clustering and random forest classifier approach

a. Flowchart depicting iterative tiered clustering (ITC) used for generating FGID and pediCD cellular atlases. After sequencing, cells underwent quality control and a cell by gene expression matrix was derived from the 27 ileal samples. Dimensionality reduction and graph-based clustering were performed using the standard Seurat workflow to annotate cell types. Resulting clusters were then iteratively processed through the same pipeline unless end conditions were met. Each cluster was checked for three end conditions which included: only one cluster remaining, two clusters remaining with no more than 5 up and down regulated genes as determined by Wilcoxon test (logFC > 1.5, FDR < 0.001), and/or less than 100 cells in the cluster. Iterative clustering stopped if any of the three conditions are met. Unlike traditional Seurat clustering, in ITC principal component and clustering resolution parameters are chosen automatically. Stop conditions are built in as parameters to the ITC pipeline, allowing customization to the dataset.
b. Cell and cluster numbers after various processing steps tabulated
c. Random forest classifier approach for integrating FGID and pediCD datasets. FGID and pediCD datasets were used as training datasets to create random forest predictors used in downstream sub-clustering of cell types and subsets. The opposing dataset was then tested by each algorithm independently to determine correspondence and bias as depicted in **Figure 6 and Supplemental Figure 8**.

Supplemental Figure 6

a End Clusters for Crohn's Disease Myeloid Cells with Descriptive Names and Biologically Curated Marker Genes



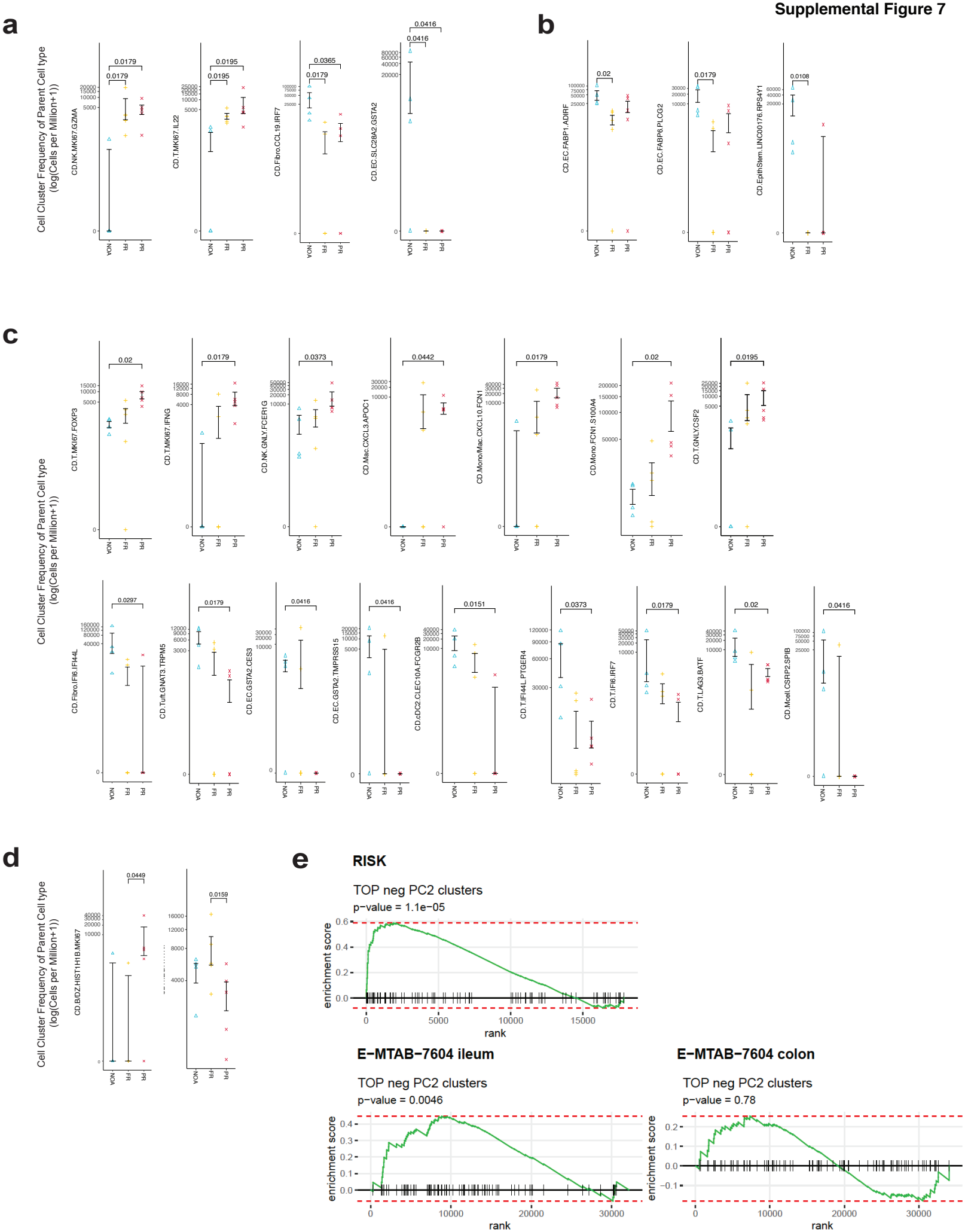
b End Clusters for Crohn's Disease Lymphocytes (T/NK/ILC) with Descriptive Names and Biologically Curated Marker Genes

Supplemental Figure 6. Representative marker genes for myeloid and T cells.

a. Dot plot of curated genes related to myeloid biology. Dot size represents fraction of cells expressing the gene, and color intensity represents binned count-based expression level (log(scaled UMI+1)) amongst expressing cells. All cluster defining genes are provided in **Supplemental Table 7 and Supplemental Table 10**. Dot size is only plotted if more than 5% of cells are expressing the transcript. Names are descriptive names generated from inspection of ITC output which were then converted to standardized naming scheme as in **Methods**.

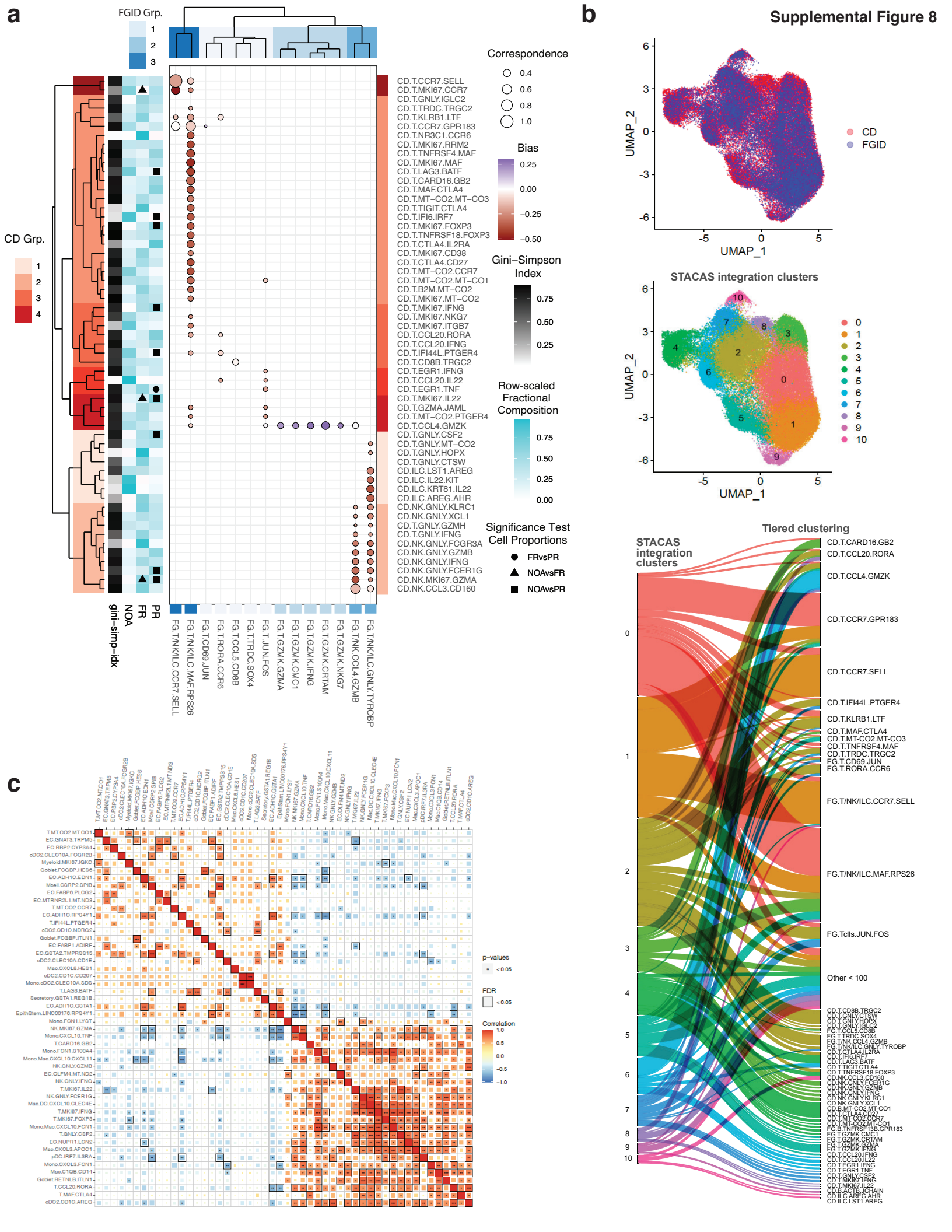
b. Dot plot of marker genes related to T/NK/ILC lymphoid biology as in **a**.

Supplemental Figure 7



Supplemental Figure 7. Cell types associated with pediCD severity after PCA analysis

a. Cell cluster frequencies of the parent cell type found to be significant by Mann-Whitney U test between selected clusters (see **Supplemental Figure 7** for all graphs).
b. Cell cluster frequencies of the parent cell type between NOA and FR (as above).
c. Cell cluster frequencies of the parent cell type between NOA and PR (as above).
d. Cell cluster frequencies of the parent cell type between FR and PR (as above).
e. GSEA analysis showing the ranks of 92 PREDICT markers (markers of top 25 cell states associated with disease severity and treatment outcomes) in bulk RNA sequencing of illeal or colonic mucosa of two other treatment-naïve cohorts (pediatric RISK cohort, n = 69, adult E−MTAB−7604 cohort, n = 43) comparing pediCD patients who did or did not respond to anti-TNF therapy. P-value is estimated based on an adaptive multi-level split Monte-Carlo scheme.



Supplemental Figure 8
a
FGID Grp.
CD Grp.
Correspondence
Bias
Gini–Simpson Index
Row-scaled Fractional Composition
Significance Test Cell Proportions
FRvsPR
NOAvsFR
NOAvsPR
gini–simp–idx
NOA
FR
PR
b
UMAP_1
UMAP_2
CD
FGID
STACAS integration clusters
Tiered clustering
STACAS integration clusters
c
p–values
FDR
Correlation

Supplemental Figure 8. Random Forest classification applied to T cell subsets and integration using STACAS

a. Correspondence between cell subsets from FGID-to-pediCD and pediCD-to-FGID.

   Top left heatmaps: RF probabilities for each cell averaged over subset to gain probability of each FGID matching onto each pediCD subset (left), and pediCD onto FGID (right).

   Bubble plot (center): size = sum(probability matrices) for confidence of predictions, marker color = diff(probability matrices) to show which direction the RF model is more confident on, e.g. more likely for FGID subset to belong to pediCD subset or pediCD subset to belong to FGID subset. Markers are filtered to show the top 10th percentile of correspondence.

   Dendrograms: separated-tiered clustering on prediction probabilities of FGID (blue) and pediCD (red) using complete linkage with correlation distance metric, clusters are cut at height 0.7 (range 0-1).

   Heatmap: 1-Gini-Simpson index based on patient diversity, mono-patient clusters (white), full representation (black). Right 3 columns show row-normalized of frequency of NOA, FR, PR representation in each pediCD cell subset. Significant differences (Mann-Whitney, alpha=0.05) are marked, triangle NOA vs. PR and circle NOA vs. FR.

b. T cells from the main FGID (n = 29,640 cells) and pediCD (n = 38,031) datasets were integrated using identification of mutual nearest neighbors in a reduced space (reciprocal PCA method) with the STACAS package. UMAP plots show distribution of cells coming from FGID (blue) and pediCD (red) datasets and 11 clusters obtained using Louvain algorithm. Sankey plot shows the contribution of ARBOL clusters to each Louvain cluster in the integrated dataset.

c. Spearman rank correlation heatmap of the counts-per-million for each of the top 25 clusters defining PC2 positive (NOA-associated) and PC2 negative (PR-associated) vectors. Correlation is represented by both the intensity and size of the box and those which are FDR < 0.05 have a bounding box

REFERENCES

Ai, L., Ren, Y., Zhu, M., Lu, S., Qian, Y., Chen, Z., and Xu, A. (2021). Synbindin restrains proinflammatory macrophage activation against microbiota and mucosal inflammation during colitis. Gut gutjnl-2020-321094.

Andreatta, M., and Carmona, S.J. (2021). STACAS: Sub-Type Anchor Correction for Alignment in Seurat to integrate single-cell RNA-seq data. Bioinformatics *37*, 882–884.

Aschenbrenner, D., Quaranta, M., Banerjee, S., Ilott, N., Jansen, J., Steere, B., Chen, Y.-H., Ho, S., Cox, K., Arancibia-Cárcamo, C.V., et al. (2021). Deconvolution of monocyte responses in inflammatory bowel disease reveals an IL-1 cytokine network that regulates IL-23 in genetic and acquired IL-10 resistance. Gut *70*, 1023–1036.

Atreya, R., Neurath, M.F., and Siegmund, B. (2020). Personalizing Treatment in IBD: Hype or Reality in 2020? Can We Predict Response to Anti-TNF? Frontiers in Medicine *7*, 517.

Banerjee, A., Herring, C.A., Chen, B., Kim, H., Simmons, A.J., Southard-Smith, A.N., Allaman, M.M., White, J.R., Macedonia, M.C., Mckinley, E.T., et al. (2020). Succinate Produced by Intestinal Microbes Promotes Specification of Tuft Cells to Suppress Ileal Inflammation. Gastroenterology *159*, 2101-2115.e5.

Barker, N., van Es, J.H., Kuipers, J., Kujala, P., van den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P.J., et al. (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature *449*, 1003–1007.

Baumgart, D.C., and Sandborn, W.J. (2012). Crohn's disease. The Lancet *380*, 1590–1605.

Betts, B.C., Veerapathran, A., Pidala, J., Yang, H., Horna, P., Walton, K., Cubitt, C.L., Gunawan, S., Lawrence, H.R., Lawrence, N.J., et al. (2017). Targeting Aurora kinase A and JAK2 prevents GVHD while maintaining Treg and antitumor CTL function. Sci Transl Med *9*, eaai8269.

Beumer, J., Puschhof, J., Bauzá-Martinez, J., Martínez-Silgado, A., Elmentaite, R., James, K.R., Ross, A., Hendriks, D., Artegiani, B., Busslinger, G.A., et al. (2020). High-Resolution mRNA and Secretome Atlas of Human Enteroendocrine Cells. Cell *181*, 1291-1306.e19.

Biton, M., Haber, A.L., Rogel, N., Burgin, G., Beyaz, S., Schnell, A., Ashenberg, O., Su, C.-W., Smillie, C., Shekhar, K., et al. (2018). T Helper Cell Cytokines Modulate Intestinal Stem Cell Renewal and Differentiation. Cell *175*, 1307-1320.e22.

Björklund, Å.K., Forkel, M., Picelli, S., Konya, V., Theorell, J., Friberg, D., Sandberg, R., and Mjösberg, J. (2016). The heterogeneity of human CD127+ innate lymphoid cells revealed by single-cell RNA sequencing. Nat Immunol *17*, 451–460.

Black, C.J., Drossman, D.A., Talley, N.J., Ruddy, J., and Ford, A.C. (2020). Functional gastrointestinal disorders: advances in understanding and management. The Lancet *396*, 1664–1674.

Blériot, C., Chakarov, S., and Ginhoux, F. (2020). Determinants of Resident Tissue Macrophage Identity and Function. Immunity *52*, 957–970.

Brulois, K., Rajaraman, A., Szade, A., Nordling, S., Bogoslowski, A., Dermadi, D., Rahman, M., Kiefel, H., O'Hara, E., Koning, J.J., et al. (2020). A molecular map of murine lymph node blood vascular endothelium at single cell resolution. Nat Commun *11*, 3798.

Buechler, M.B., Pradhan, R.N., Krishnamurty, A.T., Cox, C., Calviello, A.K., Wang, A.W., Yang, Y.A., Tam, L., Caothien, R., Roose-Girma, M., et al. (2021). Cross-tissue organization of the fibroblast lineage. Nature 1–5.

Buisine, M.P., Desreumaux, P., Leteurtre, E., Copin, M.C., Colombel, J.F., Porchet, N., and Aubert, J.P. (2001). Mucin gene expression in intestinal epithelial cells in Crohn's disease. Gut *49*, 544–551.

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. Nature *566*, 496–502.

Cappello, M., and Morreale, G.C. (2016). The Role of Laboratory Tests in Crohn's Disease. Clin Med Insights Gastroenterol *9*, 51–62.

Castro-Dopico, T., Fleming, A., Dennison, T.W., Ferdinand, J.R., Harcourt, K., Stewart, B.J., Cader, Z., Tuong, Z.K., Jing, C., Lok, L.S.C., et al. (2020). GM-CSF Calibrates Macrophage Defense and Wound Healing Programs during Intestinal Infection and Inflammation. Cell Reports *32*, 107857.

Catalan-Serra, I., Sandvik, A.K., Bruland, T., and Andreu-Ballester, J.C. (2017). Gammadelta T Cells in Crohn's Disease: A New Player in the Disease Pathogenesis? Journal of Crohn's and Colitis *11*, 1135–1145.

Chang, J.T. (2020). Pathophysiology of Inflammatory Bowel Diseases. New England Journal of Medicine *383*, 2652–2664.

Cherrier, D.E., Serafini, N., and Di Santo, J.P. (2018). Innate Lymphoid Cell Development: A T Cell Perspective. Immunity *48*, 1091–1103.

Cima, I., Corazza, N., Dick, B., Fuhrer, A., Herren, S., Jakob, S., Ayuni, E., Mueller, C., and Brunner, T. (2004). Intestinal Epithelial Cells Synthesize Glucocorticoids and Regulate T Cell Activation. J Exp Med *200*, 1635–1646.

Cohen, L.J., Cho, J.H., Gevers, D., and Chu, H. (2019). Genetic Factors and the Intestinal Microbiome Guide Development of Microbe-Based Therapies for Inflammatory Bowel Diseases. Gastroenterology *156*, 2174–2189.

Corridoni, D., Chapman, T., Antanaviciute, A., Satsangi, J., and Simmons, A. (2020a). Inflammatory Bowel Disease Through the Lens of Single-cell RNA-seq Technologies. Inflammatory Bowel Diseases *26*, 1658–1668.

Corridoni, D., Antanaviciute, A., Gupta, T., Fawkner-Corbett, D., Aulicino, A., Jagielowicz, M., Parikh, K., Repapi, E., Taylor, S., Ishikawa, D., et al. (2020b). Single-cell atlas of colonic CD8+ T cells in ulcerative colitis. Nat Med *26*, 1480–1490.

Cyster, J.G., and Allen, C.D.C. (2019). B Cell Responses: Cell Interaction Dynamics and Decisions. Cell *177*, 524–540.

Das, A., Heesters, B.A., Bialas, A., O'Flynn, J., Rifkin, I.R., Ochando, J., Mittereder, N., Carlesso, G., Herbst, R., and Carroll, M.C. (2017). Follicular Dendritic Cell Activation by TLR Ligands Promotes Autoreactive B Cell Responses. Immunity *46*, 106–119.

Davidson, S., Coles, M., Thomas, T., Kollias, G., Ludewig, B., Turley, S., Brenner, M., and Buckley, C.D. (2021). Fibroblasts as immune regulators in infection, inflammation and cancer. Nature Reviews Immunology 1–14.

D'Haens, G.R., and Deventer, S. van (2021). 25 years of anti-TNF treatment for inflammatory bowel disease: lessons from the past and a look to the future. Gut *70*, 1396–1405.

Digby-Bell, J.L., Atreya, R., Monteleone, G., and Powell, N. (2020). Interrogating host immunity to predict treatment response in inflammatory bowel disease. Nat Rev Gastroenterol Hepatol *17*, 9–20.

Dominguez-Sola, D., Victora, G.D., Ying, C.Y., Phan, R.T., Saito, M., Nussenzweig, M.C., and Dalla-Favera, R. (2012). The proto-oncogene MYC is required for selection in the germinal center and cyclic reentry. Nat Immunol *13*, 1083–1091.

Dovrolis, N., Michalopoulos, G., Theodoropoulos, G.E., Arvanitidis, K., Kolios, G., Sechi, L.A., Eliopoulos, A.G., and Gazouli, M. (2020). The Interplay between Mucosal Microbiota Composition and Host Gene-Expression is Linked with Infliximab Response in Inflammatory Bowel Diseases. Microorganisms *8*, 438.

Drokhlyansky, E., Smillie, C.S., Wittenberghe, N.V., Ericsson, M., Griffin, G.K., Dionne, D., Cuoco, M.S., Goder-Reiser, M.N., Sharova, T., Aguirre, A.J., et al. (2019). The enteric nervous system of the human and mouse colon at a single-cell resolution. BioRxiv 746743.

Dutertre, C.-A., Becht, E., Irac, S.E., Khalilnezhad, A., Narang, V., Khalilnezhad, S., Ng, P.Y., van den Hoogen, L.L., Leong, J.Y., Lee, B., et al. (2019). Single-Cell Analysis of Human Mononuclear Phagocytes Reveals Subset-Defining Markers and Identifies Circulating Inflammatory Dendritic Cells. Immunity *51*, 573-589.e8.

Dwyer, D.F., Ordovas-Montanes, J., Allon, S.J., Buchheit, K.M., Vukovic, M., Derakhshan, T., Feng, C., Lai, J., Hughes, T.K., Nyquist, S.K., et al. (2021). Human airway mast cells proliferate and acquire distinct inflammation-driven phenotypes during type 2 inflammation. Sci Immunol *6*, eabb7221.

Elmentaite, R., Ross, A.D.B., Roberts, K., James, K.R., Ortmann, D., Gomes, T., Nayak, K., Tuck, L., Pritchard, S., Bayraktar, O.A., et al. (2020). Single-Cell Sequencing of Developing Human Gut Reveals Transcriptional Links to Childhood Crohn's Disease. Developmental Cell *55*, 771-783.e5.

van der Flier, L.G., and Clevers, H. (2009). Stem Cells, Self-Renewal, and Differentiation in the Intestinal Epithelium. Annu. Rev. Physiol. *71*, 241–260.

Franzosa, E.A., Sirota-Madi, A., Avila-Pacheco, J., Fornelos, N., Haiser, H.J., Reinker, S., Vatanen, T., Hall, A.B., Mallick, H., McIver, L.J., et al. (2019). Gut microbiome structure and metabolic activity in inflammatory bowel disease. Nat Microbiol *4*, 293–305.

French, A.R., Sjölin, H., Kim, S., Koka, R., Yang, L., Young, D.A., Cerboni, C., Tomasello, E., Ma, A., Vivier, E., et al. (2006). DAP12 Signaling Directly Augments Proproliferative Cytokine Stimulation of NK Cells during Viral Infections. The Journal of Immunology *177*, 4981–4990.

Friedrich, M., Pohin, M., and Powrie, F. (2019). Cytokine Networks in the Pathophysiology of Inflammatory Bowel Disease. Immunity *50*, 992–1006.

Furey, T.S., Sethupathy, P., and Sheikh, S.Z. (2019). Redefining the IBDs using genome-scale molecular phenotyping. Nat Rev Gastroenterol Hepatol *16*, 296–311.

Gomariz, A., Helbling, P.M., Isringhausen, S., Suessbier, U., Becker, A., Boss, A., Nagasawa, T., Paul, G., Goksel, O., Székely, G., et al. (2018). Quantitative spatial analysis of haematopoiesis-regulating stromal cells in the bone marrow microenvironment by 3D microscopy. Nat Commun *9*, 2532.

Graham, D.B., and Xavier, R.J. (2020). Pathway paradigms revealed from the genetics of inflammatory bowel disease. Nature *578*, 527–539.

Guilliams, M., Mildner, A., and Yona, S. (2018). Developmental and Functional Heterogeneity of Monocytes. Immunity *49*, 595–613.

Haberman, Y., Tickle, T.L., Dexheimer, P.J., Kim, M.-O., Tang, D., Karns, R., Baldassano, R.N., Noe, J.D., Rosh, J., Markowitz, J., et al. (2014). Pediatric Crohn disease patients exhibit specific ileal transcriptome and microbiome signature. J Clin Invest *124*, 3617–3633.

Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biology *20*, 296.

Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell.

Heesters, B.A., Chatterjee, P., Kim, Y.-A., Gonzalez, S.F., Kuligowski, M.P., Kirchhausen, T., and Carroll, M.C. (2013). Endocytosis and Recycling of Immune Complexes by Follicular Dendritic Cells Enhances B Cell Antigen Binding and Activation. Immunity *38*, 1164–1175.

Hie, B., Bryson, B., and Berger, B. (2019). Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. Nat Biotechnol *37*, 685–691.

Hie, B., Peters, J., Nyquist, S.K., Shalek, A.K., Berger, B., and Bryson, B.D. (2020). Computational Methods for Single-Cell RNA Sequencing. Annu. Rev. Biomed. Data Sci. *3*, 339–364.

Huang, B., Chen, Z., Geng, L., Wang, J., Liang, H., Cao, Y., Chen, H., Huang, W., Su, M., Wang, H., et al. (2019). Mucosal Profiling of Pediatric-Onset Colitis and IBD Reveals Common Pathogenics and Therapeutic Pathways. Cell *179*, 1160-1176.e24.

Hyams, J.S., Ferry, G.D., Mandel, F.S., Gryboski, J.D., Kibort, P.M., Kirschner, B.S., Griffiths, A.M., Katz, A.J., Grand, R.J., Boyle, J.T., et al. (1991). Development and Validation of a Pediatric Crohn's Disease Activity Index. Journal of Pediatric Gastroenterology and Nutrition *12*, 439.

Hyams, J.S., Di Lorenzo, C., Saps, M., Shulman, R.J., Staiano, A., and van Tilburg, M. (2016). Childhood Functional Gastrointestinal Disorders: Child/Adolescent. Gastroenterology *150*, 1456-1468.e2.

Jain, U., Heul, A.M.V., Xiong, S., Gregory, M.H., Demers, E.G., Kern, J.T., Lai, C.-W., Muegge, B.D., Barisas, D.A.G., Leal-Ekman, J.S., et al. (2021). Debaryomyces is enriched in Crohn's disease intestinal tissue and impairs healing in mice. Science *371*, 1154–1159.

James, K.R., Gomes, T., Elmentaite, R., Kumar, N., Gulliver, E.L., King, H.W., Stares, M.D., Bareham, B.R., Ferdinand, J.R., Petrova, V.N., et al. (2020). Distinct microbial and immune niches of the human colon. Nat Immunol *21*, 343–353.

Kaczorowski, K.J., Shekhar, K., Nkulikiyimfura, D., Dekker, C.L., Maecker, H., Davis, M.M., Chakraborty, A.K., and Brodin, P. (2017). Continuous immunotypes describe human immune variation and predict diverse responses. PNAS *114*, E6097–E6106.

Kinchen, J., Chen, H.H., Parikh, K., Antanaviciute, A., Jagielowicz, M., Fawkner-Corbett, D., Ashley, N., Cubitt, L., Mellado-Gomez, E., Attar, M., et al. (2018). Structural Remodeling of the Human Colonic Mesenchyme in Inflammatory Bowel Disease. Cell *175*, 372-386.e17.

Kobayashi, T., Siegmund, B., Le Berre, C., Wei, S.C., Ferrante, M., Shen, B., Bernstein, C.N., Danese, S., Peyrin-Biroulet, L., and Hibi, T. (2020). Ulcerative colitis. Nat Rev Dis Primers *6*, 1–20.

Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N., and Sergushichev, A. (2021). Fast gene set enrichment analysis.

Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P., and Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods *16*, 1289–1296.

Kugathasan, S., Denson, L.A., Walters, T.D., Kim, M.-O., Marigorta, U.M., Schirmer, M., Mondal, K., Liu, C., Griffiths, A., Noe, J.D., et al. (2017). Prediction of complicated disease course for children newly diagnosed with Crohn's disease: a multicentre inception cohort study. Lancet *389*, 1710–1718.

La Manno, G., Siletti, K., Furlan, A., Gyllborg, D., Vinsland, E., Mossi Albiach, A., Mattsson Langseth, C., Khven, I., Lederer, A.R., Dratva, L.M., et al. (2021). Molecular architecture of the developing mouse brain. Nature *596*, 92–96.

Lampen, A., Meyer, S., Arnhold, T., and Nau, H. (2000). Metabolism of vitamin A and its active metabolite all-trans-retinoic acid in small intestinal enterocytes. J Pharmacol Exp Ther *295*, 979–985.

Lanier, L.L. (2001). On guard—activating NK cell receptors. Nat Immunol *2*, 23–27.

Lanier, L.L., Corliss, B., Wu, J., and Phillips, J.H. (1998). Association of DAP12 with Activating CD94/NKG2C NK Cell Receptors. Immunity *8*, 693–701.

Leach, S.T., Yang, Z., Messina, I., Song, C., Geczy, C.L., Cunningham, A.M., and Day, A.S. (2007). Serum and mucosal S100 proteins, calprotectin (S100A8/S100A9) and S100A12, are elevated at diagnosis in children with inflammatory bowel disease. Scandinavian Journal of Gastroenterology *42*, 1321–1331.

Leeb, S.N., Vogl, D., Gunckel, M., Kiessling, S., Falk, W., Göke, M., Schölmerich, J., Gelbmann, C.M., and Rogler, G. (2003). Reduced migration of fibroblasts in inflammatory bowel disease: role of inflammatory mediators and focal adhesion kinase. Gastroenterology *125*, 1341–1354.

Leonard, N., Hourihane, D.O., and Whelan, A. (1995). Neuroproliferation in the mucosa is a feature of coeliac disease and Crohn's disease. Gut *37*, 763–765.

Levine, A., Griffiths, A., Markowitz, J., Wilson, D.C., Turner, D., Russell, R.K., Fell, J., Ruemmele, F.M., Walters, T., Sherlock, M., et al. (2011). Pediatric modification of the Montreal classification for inflammatory bowel disease: the Paris classification. Inflamm Bowel Dis *17*, 1314–1321.

Lilja, I., Gustafson-Svärd, C., Franzén, L., and Sjödahl, R. (2000). Tumor Necrosis Factor-Alpha in Ileal Mast Cells in Patients with Crohn's Disease. DIG *61*, 68–76.

Limon, J.J., Tang, J., Li, D., Wolf, A.J., Michelsen, K.S., Funari, V., Gargus, M., Nguyen, C., Sharma, P., Maymi, V.I., et al. (2019). Malassezia Is Associated with Crohn's Disease and Exacerbates Colitis in Mouse Models. Cell Host & Microbe *25*, 377-388.e6.

Lindemans, C.A., Calafiore, M., Mertelsmann, A.M., O'Connor, M.H., Dudakov, J.A., Jenq, R.R., Velardi, E., Young, L.F., Smith, O.M., Lawrence, G., et al. (2015). Interleukin-22 promotes intestinal-stem-cell-mediated epithelial regeneration. Nature *528*, 560–564.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology *15*, 550.

Lucas, C., Wong, P., Klein, J., Castro, T.B.R., Silva, J., Sundaram, M., Ellingson, M.K., Mao, T., Oh, J.E., Israelow, B., et al. (2020). Longitudinal analyses reveal immunological misfiring in severe COVID-19. Nature *584*, 463–469.

Mabbott, N.A., Donaldson, D.S., Ohno, H., Williams, I.R., and Mahajan, A. (2013). Microfold (M) cells: important immunosurveillance posts in the intestinal epithelium. Mucosal Immunol *6*, 666–677.

Magro, F., Vieira-Coelho, M.A., Fraga, S., Serräo, M.P., Veloso, F.T., Ribeiro, T., and Soares-da-Silva, P. (2002). Impaired Synthesis or Cellular Storage of Norepinephrine, Dopamine, and 5-Hydroxytryptamine in Human Inflammatory Bowel Disease. Dig Dis Sci *47*, 216–224.

Mårtensson, J., Jain, A., and Meister, A. (1990). Glutathione is required for intestinal function. Proc Natl Acad Sci U S A *87*, 1715–1719.

Martin, J.C., Chang, C., Boschetti, G., Ungaro, R., Giri, M., Grout, J.A., Gettler, K., Chuang, L., Nayar, S., Greenstein, A.J., et al. (2019). Single-Cell Analysis of Crohn's Disease Lesions

Identifies a Pathogenic Cellular Module Associated with Resistance to Anti-TNF Therapy. Cell *178*, 1493-1508.e20.

Martínez-Augustin, O., and de Medina, F.S. (2008). Intestinal bile acid physiology and pathophysiology. World J Gastroenterol *14*, 5630–5640.

Mathew, D., Giles, J.R., Baxter, A.E., Oldridge, D.A., Greenplate, A.R., Wu, J.E., Alanio, C., Kuri-Cervantes, L., Pampena, M.B., D'Andrea, K., et al. (2020). Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications. Science *369*.

Mauri, M., Elli, T., Caviglia, G., Uboldi, G., and Azzi, M. (2017). RAWGraphs: A Visualisation Platform to Create Open Outputs. In Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter, (New York, NY, USA: Association for Computing Machinery), pp. 1–5.

McOmber, M.A., and Shulman, R.J. (2008). Pediatric Functional Gastrointestinal Disorders. Nutrition in Clinical Practice *23*, 268–274.

Mehta, P., Porter, J.C., Manson, J.J., Isaacs, J.D., Openshaw, P.J.M., McInnes, I.B., Summers, C., and Chambers, R.C. (2020). Therapeutic blockade of granulocyte macrophage colony-stimulating factor in COVID-19-associated hyperinflammation: challenges and opportunities. The Lancet Respiratory Medicine *8*, 822–830.

Meijer, C.J.L.M., Bosman, F.T., and Lindeman, J. (1979). Evidence for Predominant Involvement of the B-Cell System in the Inflammatory Process in Crohn's Disease. Scandinavian Journal of Gastroenterology *14*, 21–32.

Mitsialis, V., Wall, S., Liu, P., Ordovas-Montanes, J., Parmet, T., Vukovic, M., Spencer, D., Field, M., McCourt, C., Toothaker, J., et al. (2020). Single-Cell Analyses of Colon and Blood Reveal Distinct Immune Cell Signatures of Ulcerative Colitis and Crohn's Disease. Gastroenterology *159*, 591-608.e10.

Miura, A., Sootome, H., Fujita, N., Suzuki, T., Fukushima, H., Mizuarai, S., Masuko, N., Ito, K., Hashimoto, A., Uto, Y., et al. (2021). TAS-119, a novel selective Aurora A and TRK inhibitor, exhibits antitumor efficacy in preclinical models with deregulated activation of the Myc, β-Catenin, and TRK pathways. Invest New Drugs *39*, 724–735.

von Moltke, J., Ji, M., Liang, H.-E., and Locksley, R.M. (2016). Tuft-cell-derived IL-25 regulates an intestinal ILC2–epithelial response circuit. Nature *529*, 221–225.

Moor, A.E., Harnik, Y., Ben-Moshe, S., Massasa, E.E., Rozenberg, M., Eilam, R., Bahar Halpern, K., and Itzkovitz, S. (2018). Spatial Reconstruction of Single Enterocytes Uncovers Broad Zonation along the Intestinal Villus Axis. Cell *175*, 1156-1167.e15.

Müller, S., Lory, J., Corazza, N., Griffiths, G.M., Z'graggen, K., Mazzucchelli, L., Kappeler, A., and Mueller, C. (1998). Activated CD4+ and CD8+ cytotoxic cells are present in increased numbers in the intestinal mucosa from patients with active inflammatory bowel disease. Am J Pathol *152*, 261–268.

Muro, M., and Mrowiec, A. (2015). Interleukin (IL)-1 Gene Cluster in Inflammatory Bowel Disease: Is IL-1RA Implicated in the Disease Onset and Outcome? Dig Dis Sci *60*, 1126–1128.

Neurath, M.F. (2019). Targeting immune cell circuits and trafficking in inflammatory bowel disease. Nat Immunol *20*, 970–979.

Ordás, I., Mould, D.R., Feagan, B.G., and Sandborn, W.J. (2012). Anti-TNF monoclonal antibodies in inflammatory bowel disease: pharmacokinetics-based dosing paradigms. Clin Pharmacol Ther *91*, 635–646.

Ordovas-Montanes, J., Dwyer, D.F., Nyquist, S.K., Buchheit, K.M., Vukovic, M., Deb, C., Wadsworth, M.H., Hughes, T.K., Kazer, S.W., Yoshimoto, E., et al. (2018). Allergic inflammatory memory in human respiratory epithelial progenitor cells. Nature *560*, 649–654.

Parikh, K., Antanaviciute, A., Fawkner-Corbett, D., Jagielowicz, M., Aulicino, A., Lagerholm, C., Davis, S., Kinchen, J., Chen, H.H., Alham, N.K., et al. (2019). Colonic epithelial cell diversity in health and inflammatory bowel disease. Nature *567*, 49–55.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research *12*, 2825–2830.

Peng, Y.-R., Shekhar, K., Yan, W., Herrmann, D., Sappington, A., Bryman, G.S., van Zyl, T., Do, M.Tri.H., Regev, A., and Sanes, J.R. (2019). Molecular Classification and Comparative Taxonomics of Foveal and Peripheral Cells in Primate Retina. Cell *176*, 1222-1237.e22.

Pereira, J.P., Kelly, L.M., Xu, Y., and Cyster, J.G. (2009). EBI2 mediates B cell segregation between the outer and centre follicle. Nature *460*, 1122–1126.

Persson, E.K., Uronen-Hansson, H., Semmrich, M., Rivollier, A., Hägerbrand, K., Marsal, J., Gudjonsson, S., Håkansson, U., Reizis, B., Kotarsky, K., et al. (2013). IRF4 Transcription-Factor-Dependent CD103+CD11b+ Dendritic Cells Drive Mucosal T Helper 17 Cell Differentiation. Immunity *38*, 958–969.

Pliner, H.A., Shendure, J., and Trapnell, C. (2019). Supervised classification enables rapid annotation of cell atlases. Nat Methods *16*, 983–986.

Rajca, S., Grondin, V., Louis, E., Vernier-Massouille, G., Grimaud, J.-C., Bouhnik, Y., Laharie, D., Dupas, J.-L., Pillant, H., Picon, L., et al. (2014). Alterations in the Intestinal Microbiome (Dysbiosis) as a Predictor of Relapse After Infliximab Withdrawal in Crohn's Disease. Inflammatory Bowel Diseases *20*, 978–986.

Ramanujam, M., Steffgen, J., Visvanathan, S., Mohan, C., Fine, J.S., and Putterman, C. (2020). Phoenix from the flames: Rediscovering the role of the CD40–CD40L pathway in systemic lupus erythematosus and lupus nephritis. Autoimmunity Reviews *19*, 102668.

Renoux, V.M., Zriwil, A., Peitzsch, C., Michaëlsson, J., Friberg, D., Soneji, S., and Sitnicka, E. (2015). Identification of a Human Natural Killer Cell Lineage-Restricted Progenitor in Fetal and Adult Tissues. Immunity *43*, 394–407.

Robinette, M.L., and Colonna, M. (2016). Immune modules shared by innate lymphoid cells and T cells. J Allergy Clin Immunol *138*, 1243–1251.

Roda, G., Chien Ng, S., Kotze, P.G., Argollo, M., Panaccione, R., Spinelli, A., Kaser, A., Peyrin-Biroulet, L., and Danese, S. (2020). Crohn's disease. Nat Rev Dis Primers *6*, 1–19.

Roncarolo, M.G., Gregori, S., Bacchetta, R., Battaglia, M., and Gagliani, N. (2018). The Biology of T Regulatory Type 1 Cells and Their Therapeutic Application in Immune-Mediated Diseases. Immunity *49*, 1004–1019.

Rousseeuw, P.J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics *20*, 53–65.

Ruemmele, F.M., Veres, G., Kolho, K.L., Griffiths, A., Levine, A., Escher, J.C., Amil Dias, J., Barabino, A., Braegger, C.P., Bronsky, J., et al. (2014). Consensus guidelines of ECCO/ESPGHAN on the medical management of pediatric Crohn's disease. J Crohns Colitis *8*, 1179–1207.

Sallusto, F., Lenig, D., Förster, R., Lipp, M., and Lanzavecchia, A. (1999). Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. Nature *401*, 708–712.

Sandborn, W.J. (2014). Crohn's Disease Evaluation and Treatment: Clinical Decision Tool. Gastroenterology *147*, 702–705.

Santucci, N.R., Saps, M., and van Tilburg, M.A. (2020). New advances in the treatment of paediatric functional abdominal pain disorders. The Lancet Gastroenterology & Hepatology *5*, 316–328.

Schulte-Schrepping, J., Reusch, N., Paclik, D., Baßler, K., Schlickeiser, S., Zhang, B., Krämer, B., Krammer, T., Brumhard, S., Bonaguro, L., et al. (2020). Severe COVID-19 Is Marked by a Dysregulated Myeloid Cell Compartment. Cell *182*, 1419-1440.e23.

Selin, K.A., Hedin, C.R.H., and Villablanca, E.J. (2021). Immunological networks defining the heterogeneity of inflammatory bowel diseases. Journal of Crohn's and Colitis.

Shekhar, K., Lapan, S.W., Whitney, I.E., Tran, N.M., Macosko, E.Z., Kowalczyk, M., Adiconis, X., Levin, J.Z., Nemesh, J., Goldman, M., et al. (2016). COMPREHENSIVE CLASSIFICATION OF RETINAL BIPOLAR NEURONS BY SINGLE-CELL TRANSCRIPTOMICS. Cell *166*, 1308-1323.e30.

Sido, B., Hack, V., Hochlehnert, A., Lipps, H., Herfarth, C., and Dröge, W. (1998). Impairment of intestinal glutathione synthesis in patients with inflammatory bowel disease. Gut *42*, 485–492.

Sieber, G., Herrmann, F., Zeitz, M., Teichmann, H., and Rühl, H. (1984). Abnormalities of B-cell activation and immunoregulation in patients with Crohn's disease. Gut *25*, 1255–1261.

Silverberg, M.S., Satsangi, J., Ahmad, T., Arnott, I.D.R., Bernstein, C.N., Brant, S.R., Caprilli, R., Colombel, J.-F., Gasche, C., Geboes, K., et al. (2005). Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. Can J Gastroenterol *19 Suppl A*, 5A-36A.

Simpson, E.H. (1949). Measurement of Diversity. Nature *163*, 688–688.

Smillie, C.S., Biton, M., Ordovas-Montanes, J., Sullivan, K.M., Burgin, G., Graham, D.B., Herbst, R.H., Rogel, N., Slyper, M., Waldman, J., et al. (2019). Intra- and Inter-cellular Rewiring of the Human Colon during Ulcerative Colitis. Cell *178*, 714-730.e22.

Sootome, H., Miura, A., Masuko, N., Suzuki, T., Uto, Y., and Hirai, H. (2020). Aurora A Inhibitor TAS-119 Enhances Antitumor Efficacy of Taxanes In Vitro and In Vivo: Preclinical Studies as Guidance for Clinical Development and Trial Design. Mol Cancer Ther *19*, 1981–1991.

Souza, H.S., Elia, C.C.S., Spencer, J., and MacDonald, T.T. (1999). Expression of lymphocyte-endothelial receptor-ligand pairs, α4β7/MAdCAM-1 and OX40/OX40 ligand in the colon and jejunum of patients with inflammatory bowel disease. Gut *45*, 856–863.

Stappenbeck, T.S., and McGovern, D.P.B. (2017). Paneth Cell Alterations in the Development and Phenotype of Crohn's Disease. Gastroenterology *152*, 322–326.

Stevens, T.W., Matheeuwsen, M., Lönnkvist, M.H., Parker, C.E., Wildenberg, M.E., Gecse, K.B., and D'Haens, G.R. (2018). Systematic review: predictive biomarkers of therapeutic response in inflammatory bowel disease-personalised medicine in its infancy. Aliment Pharmacol Ther *48*, 1213–1231.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888-1902.e21.

Su, Y., Chen, D., Yuan, D., Lausted, C., Choi, J., Dai, C.L., Voillet, V., Duvvuri, V.R., Scherler, K., Troisch, P., et al. (2020). Multi-Omics Resolves a Sharp Disease-State Shift between Mild and Moderate COVID-19. Cell *183*, 1479-1495.e20.

Sullivan, Z.A., Khoury-Hanold, W., Lim, J., Smillie, C., Biton, M., Reis, B.S., Zwick, R.K., Pope, S.D., Israni-Winger, K., Parsa, R., et al. (2021). γδ T cells regulate the intestinal response to nutrient sensing. Science *371*.

Sýkora, J., Pomahačová, R., Kreslová, M., Cvalínová, D., Štych, P., and Schwarz, J. (2018). Current global trends in the incidence of pediatric-onset inflammatory bowel disease. World J Gastroenterol *24*, 2741–2763.

Takayama, T., Kamada, N., Chinen, H., Okamoto, S., Kitazume, M.T., Chang, J., Matuzaki, Y., Suzuki, S., Sugita, A., Koganei, K., et al. (2010). Imbalance of NKp44+NKp46− and NKp44−NKp46+ Natural Killer Cells in the Intestinal Mucosa of Patients With Crohn's Disease. Gastroenterology *139*, 882-892.e3.

Tasic, B., Yao, Z., Graybuck, L.T., Smith, K.A., Nguyen, T.N., Bertagnolli, D., Goldy, J., Garren, E., Economo, M.N., Viswanathan, S., et al. (2018). Shared and distinct transcriptomic cell types across neocortical areas. Nature *563*, 72–78.

Thiriot, A., Perdomo, C., Cheng, G., Novitzky-Basso, I., McArdle, S., Kishimoto, J.K., Barreiro, O., Mazo, I., Triboulet, R., Ley, K., et al. (2017). Differential DARC/ACKR1 expression distinguishes venular from non-venular endothelial cells in murine tissues. BMC Biology *15*, 45.

Travaglini, K.J., Nabhan, A.N., Penland, L., Sinha, R., Gillich, A., Sit, R.V., Chang, S., Conley, S.D., Mori, Y., Seita, J., et al. (2020). A molecular cell atlas of the human lung from single-cell RNA sequencing. Nature *587*, 619–625.

Turner, D., Griffiths, A.M., Walters, T.D., Seah, T., Markowitz, J., Pfefferkorn, M., Keljo, D., Waxman, J., Otley, A., LeLeiko, N.S., et al. (2012). Mathematical weighting of the pediatric Crohn's disease activity index (PCDAI) and comparison with its other short versions. Inflamm Bowel Dis *18*, 55–62.

Turner, D., Levine, A., Walters, T.D., Focht, G., Otley, A., López, V.N., Koletzko, S., Baldassano, R., Mack, D., Hyams, J., et al. (2017). Which PCDAI Version Best Reflects Intestinal Inflammation in Pediatric Crohn Disease? Journal of Pediatric Gastroenterology and Nutrition *64*, 254–260.

Verstockt, B., Verstockt, S., Dehairs, J., Ballet, V., Blevi, H., Wollants, W.-J., Breynaert, C., Van Assche, G., Vermeire, S., and Ferrante, M. (2019). Low TREM1 expression in whole blood predicts anti-TNF response in inflammatory bowel disease. EBioMedicine *40*, 733–742.

Victora, G.D., Schwickert, T.A., Fooksman, D.R., Kamphorst, A.O., Meyer-Hermann, M., Dustin, M.L., and Nussenzweig, M.C. (2010). Germinal Center Dynamics Revealed by Multiphoton Microscopy with a Photoactivatable Fluorescent Reporter. Cell *143*, 592–605.

Wen, J., and Rawls, J.F. (2020). Feeling the Burn: Intestinal Epithelial Cells Modify Their Lipid Metabolism in Response to Bacterial Fermentation Products. Cell Host & Microbe *27*, 314–316.

Whitsett, J.A., Kalin, T.V., Xu, Y., and Kalinichenko, V.V. (2019). Building and Regenerating the Lung Cell by Cell. Physiol Rev *99*, 513–554.

Yarur, A.J., Jain, A., Sussman, D.A., Barkin, J.S., Quintero, M.A., Princen, F., Kirkland, R., Deshpande, A.R., Singh, S., and Abreu, M.T. (2016). The association of tissue anti-TNF drug levels with serological and endoscopic disease activity in inflammatory bowel disease: the ATLAS study. Gut *65*, 249–255.

Ye, Y., Manne, S., Treem, W.R., and Bennett, D. (2020). Prevalence of Inflammatory Bowel Disease in Pediatric and Adult Populations: Recent Estimates From Large National Databases in the United States, 2007-2016. Inflamm Bowel Dis *26*, 619–625.

Yilmaz, B., Juillerat, P., Øyås, O., Ramon, C., Bravo, F.D., Franc, Y., Fournier, N., Michetti, P., Mueller, C., Geuking, M., et al. (2019). Microbial network disturbances in relapsing refractory Crohn's disease. Nat Med *25*, 323–336.

Zeisel, A., Hochgerner, H., Lönnerberg, P., Johnsson, A., Memic, F., van der Zwan, J., Häring, M., Braun, E., Borm, L.E., La Manno, G., et al. (2018). Molecular Architecture of the Mouse Nervous System. Cell *174*, 999-1014.e22.

Ziegler, C.G.K., Allon, S.J., Nyquist, S.K., Mbano, I.M., Miao, V.N., Tzouanas, C.N., Cao, Y., Yousif, A.S., Bals, J., Hauser, B.M., et al. (2020). SARS-CoV-2 Receptor ACE2 Is an Interferon-Stimulated Gene in Human Airway Epithelial Cells and Is Detected in Specific Cell Subsets across Tissues. Cell *181*, 1016-1035.e19.

Ziegler, C.G.K., Miao, V.N., Owings, A.H., Navia, A.W., Tang, Y., Bromley, J.D., Lotfy, P., Sloan, M., Laird, H., Williams, H.B., et al. (2021). Impaired local intrinsic immunity to SARS-CoV-2 infection in severe COVID-19. Cell.