

Genomic Insights into Pediatric Intestinal Inflammatory and Eosinophilic Disorders using Single-cell RNA-sequencing

1 Marissa R. Keever-Keigher¹, Lisa Harvey¹, Veronica Williams¹, Carrie A. Vyhlidal³, Atif A.
2 Ahmed⁴, Jeffery J. Johnston¹, Daniel A. Louiselle¹, Elin Grundberg^{1,2}, Tomi Pastinen^{1,2}, Craig A.
3 Friesen^{1,2}, Rachel Chevalier^{1,2}, Craig Smail^{1, 2, †, *}, Valentina Shakhnovich^{1,2,5, †, #}

4 ¹ Children's Mercy Kansas City, Kansas City, MO, USA

5 ² University of Missouri-Kansas City School of Medicine, Kansas City, MO, USA

6 ³ KCAS Bioanalytical & Biomarker Services, Shawnee, KS, USA

7 ⁴ Seattle Children's Hospitals-University of Washington, Seattle, WA, USA

8 ⁵ Ironwood Pharmaceuticals, Boston, MA, USA

9 [†] Joint senior authors

10 [#] Current affiliation with Ironwood Pharmaceuticals has no scientific or financial bearing on the work
11 presented herein.

12 * **Correspondence:**

13 Craig Smail

14 csmail@cmh.edu

15

16 Abstract

17 Chronic inflammation of the gastrointestinal tissues underlies gastrointestinal inflammatory
18 disorders, leading to tissue damage and a constellation of painful and debilitating symptoms. These
19 disorders include inflammatory bowel diseases (Crohn's disease and ulcerative colitis), and
20 eosinophilic disorders (eosinophilic esophagitis and eosinophilic duodenitis). Gastrointestinal
21 inflammatory disorders can often present with overlapping symptoms necessitating the use of
22 invasive procedures to give an accurate diagnosis. This study used peripheral blood mononuclear
23 cells from individuals with Crohn's disease, ulcerative colitis, eosinophilic esophagitis, and
24 eosinophilic duodenitis to better understand the alterations to the transcriptome of individuals with
25 these diseases and identify potential markers of active inflammation within the peripheral blood of
26 patients that may be useful in diagnosis. Single-cell RNA-sequencing was performed on peripheral
27 blood mononuclear cells isolated from the blood samples of pediatric patients diagnosed with
28 gastrointestinal disorders, including Crohn's disease, ulcerative colitis, eosinophilic esophagitis,
29 eosinophilic duodenitis, and controls with histologically healthy gastrointestinal tracts. We identified
30 730 (FDR < 0.05) differentially expressed genes between individuals with gastrointestinal disorders
31 and controls across eight immune cell types. There were common patterns among GI disorders, such
32 as the widespread upregulation of *MTRNR2L8* across cell types, and many differentially expressed
33 genes showed distinct patterns of dysregulation among the different gastrointestinal diseases
34 compared to controls, including upregulation of *XIST* across cell types among individuals with
35 ulcerative colitis and upregulation of Th2-associated genes in eosinophilic disorders. These findings
36 indicate both overlapping and distinct alterations to the transcriptome of individuals with
37 gastrointestinal disorders compared to controls, which provide insight as to which genes may be
38 useful as markers for disease in the peripheral blood of patients.

39 Introduction

40 In genetically predisposed individuals, chronic overactivation of the inflammatory response damages
41 tissues along the gastrointestinal (GI) tract frequently resulting in painful and debilitating symptoms
42 (1, 2). GI inflammatory disorders include inflammatory bowel disease (IBD) and eosinophilic
43 gastrointestinal diseases such as eosinophilic esophagitis (EoE) and eosinophilic duodenitis (EoD).
44 IBD is characterized by chronic relapsing neutrophilic inflammation of the intestine and can be
45 divided into two main subtypes based on the site and characteristics of inflammation, with Crohn's
46 disease (CD) occurring within any portion of the gut and ulcerative colitis (UC) being confined to the
47 colon. IBD affects patients of all ages, with approximately a quarter of patients are diagnosed before
48 adulthood and incidence of pediatric IBD is increasing (3, 4). EoE and EoD are Th-2 mediated
49 inflammatory disorders which can cause dysphagia, vomiting, abdominal pain, and stricturing.
50 Histologically EoE and EoD are characterized by mucosal eosinophilia (5) and may exist
51 independently or as a comorbid condition with either form of IBD. However eosinophilic
52 infiltration of the mucosa may also precede histologic evidence of IBD (crypt distortion, cryptitis
53 with crypt abscesses, mucus depletion from goblet cells, granulomas), sometimes by years (6, 7),
54 further complicating the ability to differentiate concomitant eosinophilic disease from early
55 harbingers of IBD.

56 Differentiating between UC and CD or between early IBD and EoE/EoD can help direct therapy. As
57 example, CD patients benefit from early biologic therapy (8), and colectomy is only curative in UC
58 (9). Patients with two concomitant diseases (i.e. EoE and CD) may affect the same area but exhibit
59 different histology and symptoms and require different treatments (10). Additionally, early
60 knowledge of whether mucosal eosinophilia is burgeoning IBD can allow early and appropriate

61 intervention. The pediatric population with IBD and EoE/EoD are particularly vulnerable to growth
62 failure (11, 12) and interruptions in social-emotional development (13) and will require decades of
63 healthcare for their condition. Since disease diagnosis and follow up evaluation currently require
64 invasive endoscopic testing, non-invasive diagnostics and targeted therapies are particularly valuable
65 to this subset of patients.

66 Characterizing the role of specific immune cell populations in GI diseases has aided in recognizing
67 aberrant processes that underlie these conditions, initiating an important shift in the treatment
68 paradigm away from systemic, non-targeted immunosuppression (fraught with many unwanted side
69 effects) to targeted modulation at the site of disease activity (14). Continued identification of novel
70 therapeutic targets and molecular signatures of disease is pivotal for advancing and optimizing
71 treatment options for chronic immune-mediated inflammatory disorders. In the IBD-affected GI tract,
72 dendritic cells (DCs)—antigen presenting cells belonging to the innate immune system—exhibit up-
73 regulation of microbial recognition receptors and increased cytokine production (15) that appears to
74 induce inflammation through activation of T cells (16). T cells play a crucial role in immune
75 homeostasis (17, 18), and dysregulation of cytokine signaling in CD4⁺ T cells of the GI tract has been
76 shown to lead to pathogenic inflammation (18, 19). T cells also play a key role in eosinophilic
77 disorders of the GI tract, as overexpression of interleukin 5 (IL-5) in CD2⁺ T cells is sufficient to
78 produce eosinophilia in the esophagus and small intestine of transgenic mice (20).

79 In addition to contributing to inflammation and tissue damage at lesion sites in the GI tract, evidence
80 of altered gene expression and signaling among immune cells in peripheral blood may be reflective
81 of luminal inflammation (21-23). Information gathered from peripheral blood has the potential to
82 identify minimally-invasive, diagnosis-specific and/or disease location-specific genetic markers for
83 GI diseases. Discovery of such biomarkers could potentially decrease the need for repeat endoscopy,
84 which is invasive, associated with risks, and costly. Furthermore, identification of altered gene
85 expression within these disorders at the cellular level could yield a more complete understanding of
86 impacted pathways within specific cell types, and aid in characterizing genetic signatures for future
87 use in disease sub-typing and drug response applications. However, due to the complex and
88 multifactorial nature of GI diseases and differences in immune cell response across GI disease sub-
89 types, reliable indicators of active inflammation have been difficult to characterize within the
90 peripheral blood of patients to date.

91 In this study, we identified cell-type specific differential gene expression and enrichment of
92 functional gene ontology terms and pathways in individuals diagnosed with CD, UC, EoE, and EoD
93 using single-cell RNA-sequencing of peripheral blood mononuclear cells (PBMCs) in a pediatric
94 patient cohort. Results from this study assist in uncovering the genomic landscape of these
95 phenotypes, which often present with overlapping symptoms in patients, and aid in identifying robust
96 markers of disease types within the peripheral blood mononuclear cells of patients.

97 **Materials and Methods**

98 **Patient Information**

99 Potential study participants were identified via review of the clinical endoscopy schedule and the
100 electronic medical record (EMR) at the Children's Mercy Hospital (CMH), a tertiary regional
101 pediatric hospital in the Midwestern United States. To be considered for study inclusion, patients had
102 to be between 1 month and 21 years of age (inclusive), undergoing both upper and lower endoscopy
103 with biopsies for clinical purposes, having a reasonable clinical suspicion for a new diagnosis of
104 immune-mediated inflammatory disease or another clinical indication for undergoing endoscopy

105 (e.g., abdominal pain), and not receiving systemic immunomodulating, immunosuppressive, or
106 biological drugs. Subjects were recruited on the day of procedure, prior to endoscopy. All subjects
107 were fasting at least 8 hours for procedural purposes as part of routine medical care. Only those
108 subjects who provided informed consent (if 18 years of age), or informed assent with parental/legal
109 guardian consent (if under 18 years of age) were included. All research activities were approved by
110 the CMH Institutional Review Board. A total of 35 patients seen in the CMH operating room for
111 routine endoscopy (Kansas City, MO, USA) were included in the study. This cohort consists of 16
112 males and 19 females ranging in age from 6.17 to 19.25 years with a mean age of 13.3 years. Seven
113 patients were subsequently diagnosed with CD, nine with EoD, ten with EoE, and three with UC. Six
114 patients were identified as controls who had no relevant GI pathology on visual or histologic
115 examination of tissue. Review of individuals' medical charts indicated no bias toward a single drug
116 therapy in any sub-cohort.

117 **PBMC Isolation**

118 Up to 4 mL of whole blood was collected from patients in a sodium heparin tube and stored on ice
119 until PBMCs were isolated. Automated PBMC isolation was performed using a STEMCELL
120 Technologies RoboSep-S using the EasySep Direct Human PBMC Isolation Kit (STEMCELL
121 Technologies Cat No. 19654RF) and following the manufacturer's protocol. After PBMC isolation,
122 the resulting cell suspension was centrifuged at 300 x g for 8 min, and the supernatant was carefully
123 aspirated. The cell pellet was resuspended in 1 mL of ACK Lysing Buffer (Thermo Fisher Cat No.
124 A1049201) and incubated at room temperature for 5 min to remove any remaining RBCs. The cell
125 suspension was centrifuged at 300 x g for 8 min, and the supernatant was carefully aspirated. Cells
126 were washed twice with PBS (Thermo Fisher Cat No. 14190144) supplemented with 2% heat-
127 inactivated FBS (GE Healthcare Cat No. SH30088.03HI), and cell count and viability were assessed
128 using a Countess II automated cell counter. An aliquot of 300,000 cells was diluted in a total volume
129 of 200 μ L of PBS + 2% FBS and frozen at -80°C for downstream DNA isolation and genotyping.
130 The remaining cells were cryopreserved in aliquots of at least one million cells by centrifuging at 300
131 x g for 8 min, aspirating the supernatant, and resuspending the cell pellets in Recovery Cell Culture
132 Freezing Medium (Thermo Fisher Cat No. 12648010). The cell suspensions were transferred to
133 cryogenic storage vials and were slow-frozen overnight to a temperature of -80°C in a Corning
134 CoolCell FTS30.

135 **DNA Isolation and Genotyping**

136 Aliquots of 300,000 PBMCs frozen in PBS + 2% FBS were thawed at room temperature, and DNA
137 was isolated using the Qiagen DNeasy Blood & Tissue Kit (Qiagen Cat No. 69506) according to the
138 manufacturer's protocol. Eluate was concentrated to approximately 50 μ L using an Eppendorf
139 Vacufuge Plus, and DNA was quantified using a Qubit dsDNA BR Assay Kit following the
140 manufacturer's protocol. All DNA samples were selected for high-density genotyping using the
141 Illumina Global Screening Array (GSAMD-24v1-0) according to protocols recommended by
142 Illumina.

143 **Cell Pooling**

144 Two pools of PBMCs were made. Thawing Medium for PBMC samples consisted of IMDM (ATCC
145 Cat No. 30-2005) supplemented with 10% heat-inactivated FBS, 100 units/mL of penicillin, and 100
146 μ g/mL of streptomycin. For each sample to be thawed, 10 mL of Thawing Medium was prewarmed
147 in a 37°C bead bath. Cells were thawed in groups of up to five samples at a time. The cryovials were
148 placed in a 37°C bead bath. When thawed, the cryovials and 15-mL conical tubes containing

149 Thawing Medium were aseptically transferred to the biosafety cabinet. For each sample, 1 mL of
150 Thawing Medium was added, dropwise, to the cell suspension. The cell suspension was pipette-
151 mixed and then diluted in the remaining 9 mL of Thawing Medium. The thawed and diluted cells
152 were left at room temperature while the remaining cells to be pooled were similarly thawed. When all
153 samples were thawed, the samples were centrifuged at 300 x g for 8 min. The supernatant was
154 carefully aspirated, and the cell pellets were resuspended in 0.5 mL of room-temperature Thawing
155 Medium. All samples were placed on ice and then pooled together. The pool was passed through a
156 40- μ m nylon mesh cell strainer to remove cell aggregates. The pool was centrifuged at 300 x g for 8
157 min at 4°C, and the supernatant was carefully aspirated. The cell pellet was resuspended in 1 mL of
158 cold Thawing Medium, and cell count and viability were assessed using a Countess II automated cell
159 counter. No fewer than three aliquots per pool were cryopreserved by centrifuging at 300 x g for 8
160 min at 4°C and resuspending the cell pellets in Recovery Cell Culture Freezing Medium. The cell
161 suspensions were transferred to cryogenic storage vials and were slow-frozen overnight to a
162 temperature of -80°C in a Corning CoolCell FTS30.

163 **Single-cell Sequencing**

164 Aliquots from each pool were thawed for scRNA-seq. The 10x Genomics Chromium Single Cell 3
165 Reagent Kit v3 was used according to the manufacturer's protocol to target approximately 15,000
166 cells per scRNA-seq capture. Libraries were sequenced on an Illumina NovaSeq 6000 platform using
167 2x94 cycle paired-end sequencing.

168 **scRNA-seq Alignment and Quality Control**

169 Cell Ranger v 4.0.0 (10x Genomics) was used for read alignment to the GRCh38 (2020) reference
170 genome, gene counting, and cell calling. Demuxlet (24) was used to demultiplex single-cell data,
171 assigning reads back to the patient of origin using VCF files associated with each patient.
172 Additionally, deumuxlet was used to remove data in instances where barcodes were assigned to more
173 than a single cell.

174 Quality control of cells was performed with the Seurat v 4.9.9 (25) package in R v 4.2.1. Cells with
175 greater than 20% MT-RNA, fewer than 500 UMIs, and fewer than 0.8 \log_{10} genes per UMI were
176 removed. Annotation of the cell types of remaining cells was performed with the Azimuth v 0.4.6
177 package in R using a previously published PBMC reference (26) to annotate cells to eight broad level
178 one cell types and 30 more specific level two cell types (**Supplementary File 1 Table A**).

179 **Pseudobulk Differential Expression Analysis and Functional Analysis**

180 Gene expression data was aggregated by genotype within cell type using AggregateExpression() in
181 the Seurat R package v 4.9.9. Pseudobulk differential expression analysis of the single-cell data was
182 performed on the aggregated count data for each defined cell type with edgeR v 3.40.2 (27) in R v
183 4.2.0. Patient sex and sequencing pool were added to the statistical model to account for biological
184 variation and batch effects, and a generalized linear model was used to identify differentially
185 expressed genes (DEGs) between each GI disorder group (CD, UC, EoE, and EoD) and controls
186 across eight level one cell types and 29 level two cell types. P-values were adjusted for multiple
187 testing using the Benjamini-Hochberg false discovery rate (FDR), and genes with and FDR < 0.05
188 were considered significantly differentially expressed.

189 Functional and pathway analysis for DEGs was carried out using the ToppFun (Transcriptome,
190 ontology, phenotype, proteome, and pharmacome annotations based gene list functional enrichment

191 analysis) tool with default settings from the web-based software ToppGene Suite
192 (<http://toppgene.cchmc.org>) to identify enriched gene ontology (GO) terms and pathways from
193 databases including KEGG, Reactome, and BioCarta (28). Terms and pathways with a Benjamini-
194 Hochberg FDR < 0.05 and a minimum number of three hits in query list were considered to be
195 significantly enriched.

196 **Protein-Protein Interaction Networks**

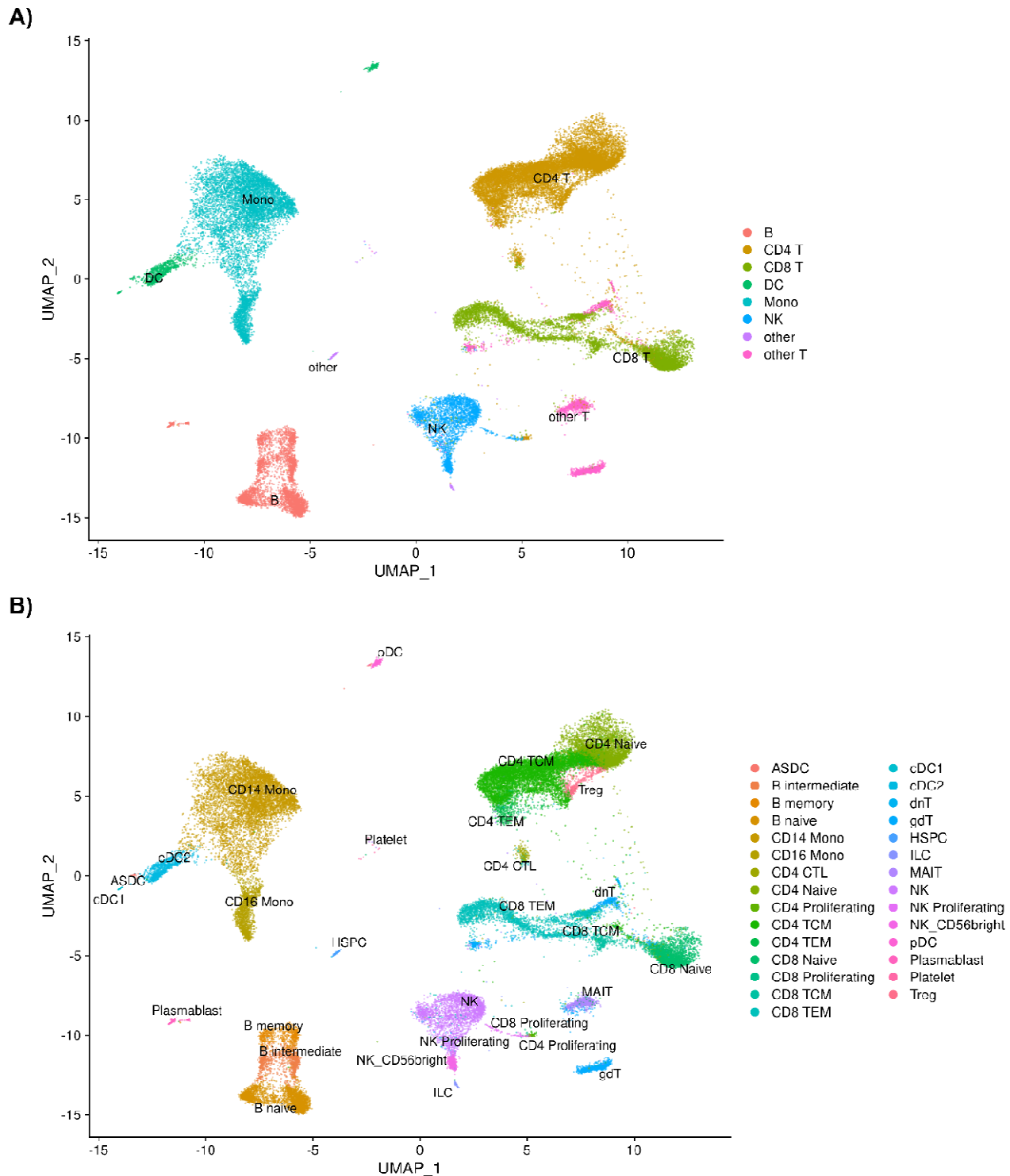
197 Networks of interactions of DEGs were generated in STRING v 12.0 (29) to visualize gene
198 relationships and trends within and between cell types and GI disorders. Networks generated with
199 STRING were imported into Cyotscape v 3.10.0 (30) where node color was used to designate the
200 direction of fold change observed in GI disorders compared to controls, with red corresponding to
201 upregulation and blue corresponding to downregulation. Furthermore, functional terms of interest
202 identified to be enriched in STRING were added to networks.

203 **Results**

204 **Cell Clustering**

205 A total of 39,622 cells from the 35 individuals in this study passed quality control measures and were
206 mapped to eight level one cell types and 29 level two cell types (**Figure 1; Supplementary File 1**
207 **Tables A, B, and C**).

208 **Figure 1. A)** UMAP of 39,622 cells passing quality control measures sorted into eight level one cell
209 types and into **B)** 29 level two cell types.



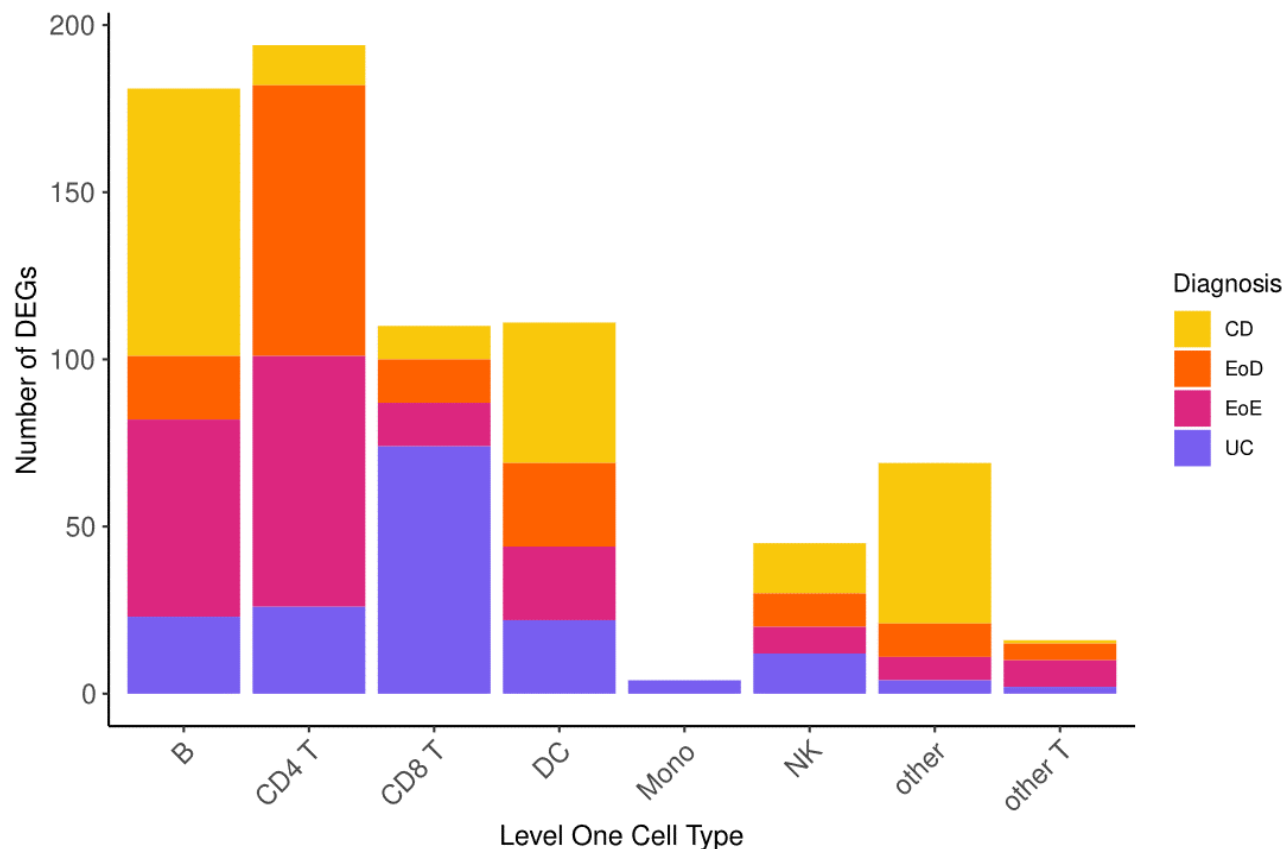
210

211 Effect of Gastrointestinal Disorders on the Transcriptome of Immune Cells

212 Comparison of gene expression between individuals diagnosed with GI disorders (CD, UC, EoD, and
 213 EoE) and controls yielded 730 (FDR < 0.05) DEGs (**Figure 2**) across eight level one cell types and
 214 807 (FDR < 0.05) across 25 level two cell types. A list of significant DEGs found across GI disorders
 215 within each cell type can be found in **Supplementary File 1 Tables D and E** and volcano plots

216 depicting the results of differential expression analysis for between GI disorders and controls within
217 all level one cell types can be found in **Supplementary File 2 Figures A-H**.

218 **Figure 2.** Number of differentially expressed genes (DEGs) across level one cell types and GI
219 disorder subtypes.

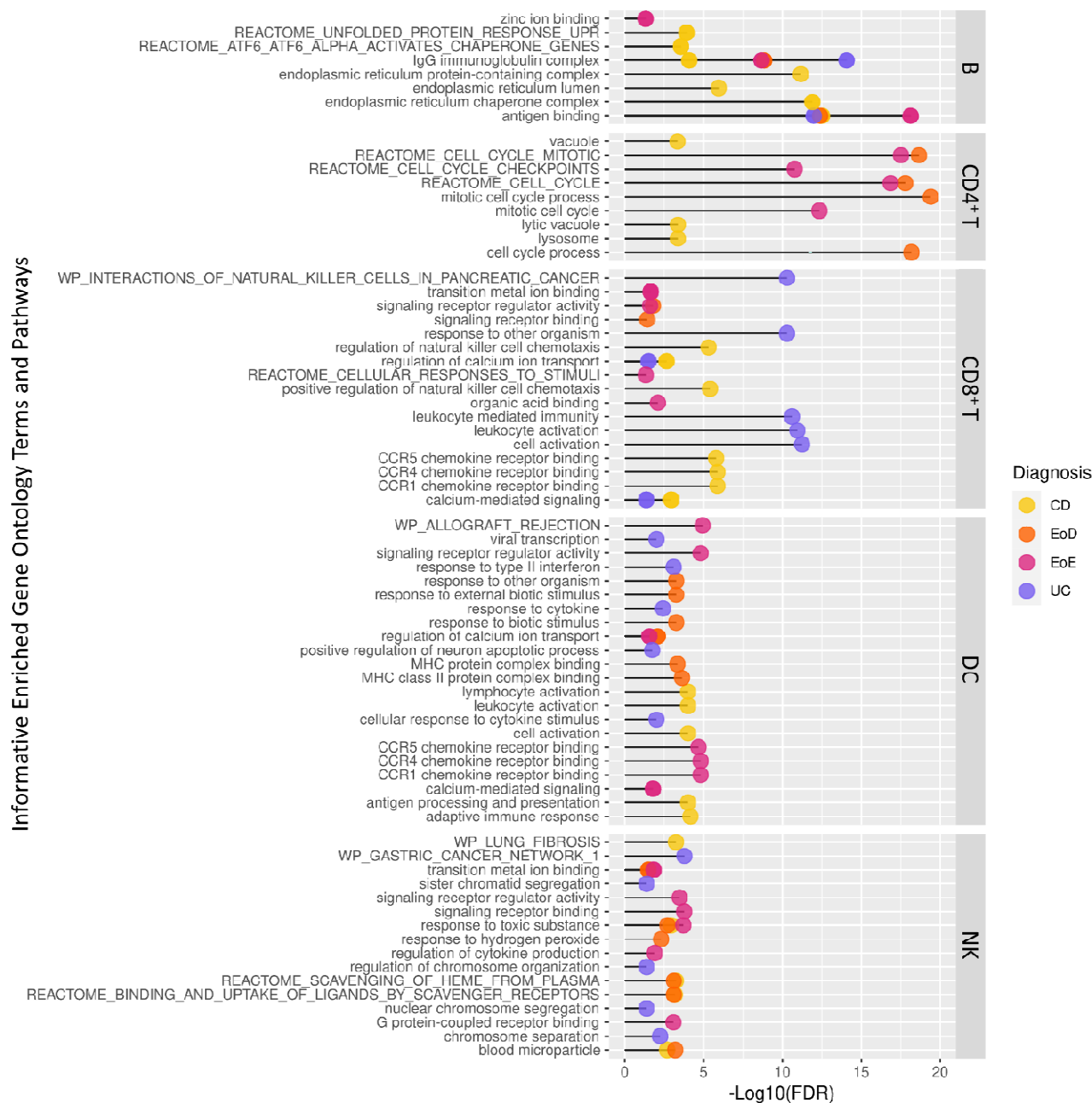


220

221 Relatively few genes showed common patterns of dysregulation across all GI disorders within cell
222 types: four genes in B cells and DCs, three genes in CD4⁺ and CD8⁺ T cells, and one gene in both
223 NK cells and other T cells (**Supplementary File 1 Table D**). Among these genes with shared
224 patterns of dysregulation were upregulated *MTRNR2L8* in six of the eight level one cell types (B,
225 CD4⁺ T, CD8⁺ T, DC, NK, and other T), downregulated *CPA5* in both CD4⁺ and CD8⁺ T cells, and
226 downregulated *CST3* in CD4⁺ T cells (**Supplementary File 1 Table D**).

227 Additionally, relatively few genes showed similar patterns of differential expression specific to
228 eosinophilic disorders (EoD and EoE) or IBD (CD and UC) compared to controls. Notable among
229 genes that shared expression patterns among eosinophilic disorders compared to controls was the
230 upregulation of *MTRNR2L1* in four level one cell types (B, CD4⁺ T, CD8⁺ T, and DC), upregulation
231 of several cell cycle associated genes (*MKI67*, *MELK*, *CLSPN*, *CCNA2*, *TOP2A*, *MCM5*, *CDCA8*,
232 *MYBL2*, *PCLAF*, *PCNA*, *DLGAP5*, *MCM7*, *UBE2C*, *ZWINT*, *SKA1*, *CCNB1*, *CDK1*, *KIF11*,
233 *CDKN3*, and *ASPM*) in CD4⁺ T cells (**Figure 3; Supplementary File 1 Table D**), and
234 downregulation of *YBX3* in CD8⁺ T cells (**Supplementary File 1 Table D**). In IBD subtypes,
235 *OR11G2* was upregulated in CD4⁺ T and CD8⁺ T cells relative to controls (**Supplementary File 1**
236 **Table D**).

237 **Figure 3.** Informative significantly enriched Gene Ontology terms and pathways among
 238 differentially expressed genes across level one cell types and GI disorder subtypes identified through
 239 ToppGene.



240

241 The majority of DEGs did not present with dysregulation in common between GI disorder subtypes;
 242 thus, many genes had patterns of dysregulation specific to each GI disorder. Notably, in CD, we
 243 observed upregulation of *CC8T* in CD4⁺ T cells, downregulation of *BTN3A2* in CD4⁺ TCM cells, and
 244 upregulation of *ETS2* in DCs. In EoD, we observed upregulation of *FOXM1* and *CCR9* in CD4⁺ T
 245 cells and downregulation of *GADD45B* and *GADD45G* in CD4⁺ T cells. In EoE, we observed
 246 upregulation of *IFNG* in CD4⁺ T cells; upregulation of *IER2* in B and CD8⁺ naïve T cells;
 247 upregulation of *EGR1* in three level one cell types (B, CD4⁺ T, and CD8⁺ T cells) and several level
 248 two cell types, including naïve CD8⁺ T cells and CD4⁺ TCM cells, and upregulation of both *EGR3*

249 and *EGR4* in B cells. Within UC, we observed downregulation of several genes in CD8⁺ T cells
250 associated with cytotoxicity, such as *CCL5*, *FGFBP2*, *GZMA*, *GZMB*, *GZMH*, *IFNG*, *NKG7*, and
251 *PRF1*; upregulation of *NOG*, *REG4*, and *AIF1* in CD8⁺ T cells; and upregulation *XIST* in five level
252 one cell types (B, CD4⁺ T, DC, monocytes, and NK) (**Supplementary File 1 Tables D and E**).

253 A subset of DEGs along with citations of literature supporting their role in GI disorders or in the
254 pathogenesis of inflammation can be found in **Supplementary File 1 Table F**.

255 **Enrichment of Functional Gene Ontology Terms among DEGs Associated with** 256 **Gastrointestinal Disorders**

257 Functional analysis of DEGs to identify enriched gene ontology terms and pathways yielded 1037
258 terms in B cells, 892 terms in CD4⁺ T cells, 760 terms in CD8⁺ T cells, 421 terms in DCs, and 62
259 terms in NK. To better understand the context of gene dysregulation observed within cell types, and
260 uncover which cellular processes may be affected, functional enrichment analysis of gene ontology
261 terms and pathways was performed. Informative enriched terms for each set of DEGs among GI
262 disorders (CD, UC, EoD, and EoE) within each cell type is shown in **Figure 4**. A full list of
263 significantly enriched GO terms and pathways can be found in **Supplementary File 1 Tables G-K**.

264 While there were few enriched GO terms and pathways detected in common among genes
265 dysregulated in CD, UC, EoD, and EoE, there was evidence of altered expression of
266 immunoglobulins in B cells among all GI disorders compared to controls, including antigen binding
267 (GO:0003823) and IgG immunoglobulin complex (GO:0071735) (**Figure 4; Supplementary File 1**
268 **Table G**). Furthermore, calcium homeostasis-associated terms were enriched among DEGs for each
269 GI disorder, terms included regulation of calcium ion transport (GO:0051924) for DEGs found in CD
270 and UC in CD8⁺ T cells and among DEGs in both EoE and EoD in DCs and regulation of calcium-
271 mediated signaling (GO:0050848) for DEGs detected in CD and UC in CD8⁺ T cells and among
272 DEGs found in EoE within B cells and DCs (**Figure 3; Supplementary File 1 Tables G, I, and J**)

273 Similarities within eosinophilic disorders (EoD and EoE) were also present. GO terms and pathways
274 associated with cell cycle activity in CD4⁺ T cells were enriched among genes differentially
275 expressed in EoD and EoE compared to controls, including Reactome cell cycle mitotic (M5336),
276 Reactome cell cycle (M543) and mitotic cell cycle process (GO:1903047). Other commonly enriched
277 terms among DEGs found in EoD and EoE included transition metal ion binding (GO:0046914)
278 within NK cells and PID AP1 pathway (M167) within CD4⁺ T cells (**Figure 3; Supplementary File**
279 **1 Tables H and K**)

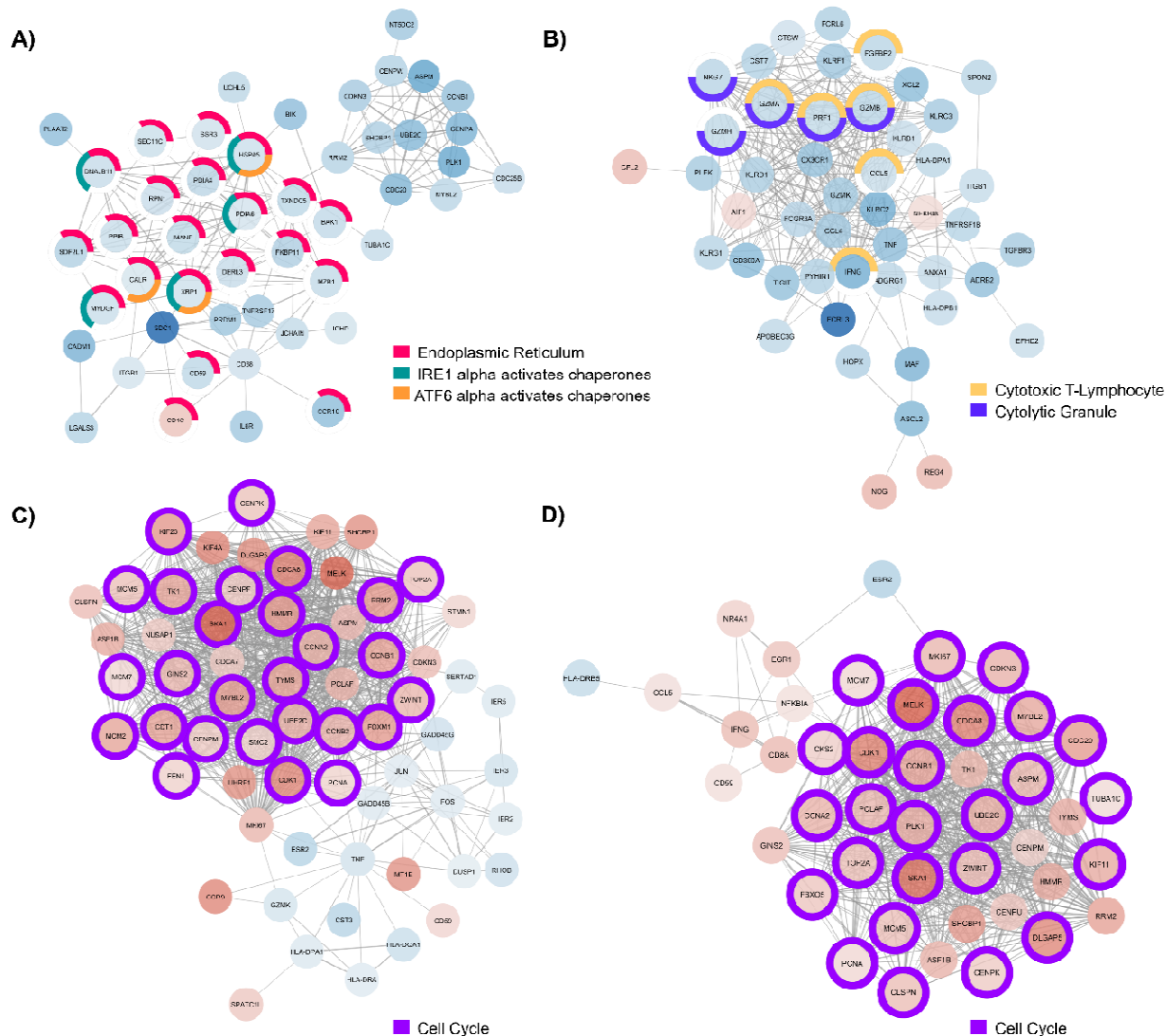
280 Specifically, among DEGs in CD, there was an enrichment of GO terms and pathways associated
281 with endoplasmic reticulum function and protein folding in B cells, such as endoplasmic reticulum
282 protein-containing complex (GO:0140534), endoplasmic reticulum chaperone complex
283 (GO:0034663), and Reactome pathways ATF6 alpha activates chaperone genes (M801). Notably,
284 among DEGs in EoD, there was enrichment of terms associated with the p38 MAPK pathway in
285 CD4⁺ T cells, including p38MAPK cascade (GO:0038066) and regulation of p38 MAPK cascade
286 (GO:1900744). Analysis of DEGs within EoE yielded enrichment of zinc ion binding (GO:0008270)
287 within B cells (**Figure 3; Supplementary File 1 Tables G and H**).

288 **Protein-Protein Interaction Networks of DEGs**

289 We generated networks for protein interaction among significant DEGs to identify relationships from
290 curated databases and mined from high-throughput studies and primary literature. A selection of the

291 major components of PPI networks generated in STRING and enhanced with Cytoscape illustrating
 292 the relationships of DEGs detected within cell types are shown in **Figure 4**. Additionally, genes
 293 associated with enriched gene ontology terms and pathways of interest have been highlighted.

294 **Figure 4.** Major component of protein-protein interaction (PPI) network generated with **A)**
 295 differentially expressed genes (DEGs) detected in B cells of individuals with CD, **B)** DEGs detected
 296 in CD8⁺ T cells from individuals with UC, **C)** DEGs detected in CD4⁺ T cells from individuals with
 297 EoD, and **D)** DEGs detected in CD4⁺ T cells from individuals with EoE. Node color indicates the
 298 direction of fold change of expression of each gene, with red indicating upregulation of gene
 299 gene expression in patients with GI disorders compared to control and blue indicating downregulation of
 300 gene expression in patients with GI disorders compared to control. Node outline color indicates
 301 enriched functional annotations detected in STRING.



302

303 In B cells of individuals with CD, most of the genes associated with endoplasmic reticulum terms are
 304 downregulated. Additionally, many of those genes are associated with chaperone functions of the
 305 endoplasmic reticulum (**Figure 4A**). We also observed widespread downregulation of DEGs within
 306 CD8⁺ T cells among individuals with UC, including markers of cytotoxicity (*CCL5*, *FGFBP2*,

307 *GZMA*, *GZMB*, *GZMH*, *IFNG*, *NKG7*, and *PRF1*) along with the upregulation *NOG*, *REG4*, and
308 *AIF1* (**Figure 4B**). We further illustrate through PPIs, the similarity of gene dysregulation and
309 dominance of DEGs associated with the cell cycle for both EoE and EoD within CD4⁺ T cells
310 (**Figure 4C and D**), highlighting the similarity of these disorders. However, within these networks
311 we can also see distinct profiles of genes dysregulated between EoE and EoD, including the
312 downregulation of genes associated with p38 MAPK (*GADD45B*, *GADD45G*, and *DUSP1*) in EoD
313 (**Figure 4C**).

314 Discussion

315 Aberrant immune signaling due to genetic and environmental factors contributes to the development
316 of GI disorders (31). In this study we focused on the characterization of transcriptomic patterns in
317 PBMCs of pediatric patients with active CD, UC, EoE, and EoD to identify genes and pathways
318 associated with active inflammation and the pathogenesis of each disease. Insights into these
319 transcriptomic phenotypes provide potential indicators of active inflammation and identify genetic
320 markers for improved diagnosis, as well as possible therapeutic targets for treatment.

321 Genes showing similar patterns of differential expression among all GI disorders (CD, UC, EoE, and
322 EoD) include the upregulation of *MTRNR2L8* within six level one cell types (B, CD4⁺ T, CD8⁺ T,
323 DC, NK, and other T), may be helpful in identifying active inflammation in GI disorders.
324 *MTRNR2L8* is believed to be a marker of cellular stress (32), and upregulation of *MTRNR2L8* has
325 been observed among PBMCs of patients with primary mitochondrial disease (33) and among
326 immune cell types in individuals with aspirin-exacerbated respiratory disease (34).

327 The common upregulation of *MTRNR2L1* for EoD and EoE was identified in four level one cell types
328 (B, CD4⁺ T, CD8⁺ T, and DC). Similar to *MTRNR2L8*, *MTRNR2L1* also appears to be upregulated in
329 response to cellular stress (32) and has been found to be upregulated in myeloid cells of patients with
330 autoimmune disorders (35). There was also enrichment of functional terms associated with cell cycle
331 activity in CD4⁺ T cells among DEGs in EoE and EoD compared to controls, prominent among
332 which were Reactome cell cycle mitotic (M5336), Reactome cell cycle (M543) and mitotic cell cycle
333 process (GO:1903047), which may be indicative of dysregulation associated with Th2-mediated
334 inflammation; genes associated with cell cycle progression and proliferation, specifically *MKI76*,
335 *RRM2*, and *TOP2A*, which were found among upregulated cell cycle genes in EoD and EoE, have
336 also been observed to be upregulated in the epithelium of asthmatics with high levels of Th2 (36).

337 Functional analysis of DEGs in CD detected in B cells yielded enrichment of terms associated with
338 endoplasmic reticulum function and protein folding. Endoplasmic reticulum stress and the
339 dysregulation of the unfolded protein response have been implicated in the development of IBD,
340 which are essential for the maintenance of homeostasis in both the intestine and immune cells (37,
341 38). Furthermore, the specific enrichment of Reactome pathway ATF6 alpha activates chaperone
342 genes in CD is supported by evidence of that diminished *ATF6* activity can contribute to intestinal
343 barrier dysfunction in mouse models of IBD (39).

344 In individual with EoD, upregulation of *CCR9* and downregulation of *GADD45B* and *GADD45G*
345 was found in CD4⁺ T cells of patients compared to controls. *CCR9* has demonstrated to play a key
346 role in Th2-mediated inflammation, with *Ccr9*^{-/-} mice treated with ovalbumin to induce allergic
347 inflammation exhibiting an impaired immune response and diminished recruitment of eosinophils to
348 inflamed tissue (40). Deficiency in *GADD45* expression has been associated with autoimmune
349 disease (41), and the increased expression of *GADD45G* through administration of IL-27, has been

350 shown to attenuate Th2-mediated allergic possibly through the activation of the p38 MAPK pathway
351 (42), a pathway that was found to be enriched among DEGs in CD4⁺ T cells in EoD.

352 Among identified DEGs in EoE relative to controls was *EGR1* in three level one cell types (B, CD4⁺
353 T, and CD8⁺ T cells) and several level two cell types, including naïve CD8⁺ T cells and CD4⁺ TCM
354 cells. Among CD4⁺ T cells, *EGR1* is preferentially expressed by Th2 and plays a role in the
355 production of IL-4 cytokines (43), which mediate the allergic inflammatory response (44).
356 Additionally, increased intracellular zinc levels have been shown to upregulate *EGR1* (45).
357 Relevantly, the term zinc ion binding was enriched among DEGs in B cells in EoE. Zinc exposure
358 has been demonstrated to elicit cellular damage (46), induce eosinophilia in mice, and evoke Th2
359 cytokine production (47). Furthermore, dysregulation of zinc signaling resulting from depletion of
360 zinc within mucosal tissues and the release of zinc from airway discharge has been associated with
361 eosinophilia (48). Together, these data support evidence that zinc homeostasis is critical in regulating
362 inflammatory responses (49) and suggests that zinc homeostasis may be involved the development of
363 eosinophilic disorders.

364 Several DEGs were found to be dysregulated between UC and controls within CD8⁺ T cells, such as
365 the downregulation of several genes involved in T cell cytotoxicity (*CCL5*, *FGFBP2*, *GZMA*, *GZMB*,
366 *GZMH*, *IFNG*, *NKG7*, and *PRF1*). Previously, it has been shown that while elevated levels of *GZMB*
367 were present in mucosal biopsies of treatment-naïve individuals with CD, there was not upregulation
368 of *GZMB* in treatment-naïve individuals with UC compared to controls, suggesting enhanced CD8⁺ T
369 cytotoxicity in CD but not UC (50). Reduced T-cell cytotoxicity has also been detected in sub-
370 populations of individuals with UC, including those who develop persistent low-grade dysplasia (51).
371 Additionally, widespread upregulation of *XIST* was detected within five level one cell types (B, CD4⁺
372 T, DC, monocytes, and NK) in UC compared to controls. In mouse models of IBD, *Xist* expression
373 has been demonstrated to be upregulated after inducing colitis, and silencing of *Xist* in these models
374 has shown to reduce colitis-associated symptoms (52). Moreover, previous transcriptome analysis of
375 intestinal mucosa biopsies from individuals with UC have identified *XIST* as a key mediator of
376 inflammation within this disease, as well as a possible therapeutic target (53).

377 The results of this study hold future implications for clinicians. DEGs that identify active
378 inflammation may be of utility in differentiating patients with active disease and disease in remission
379 without the need for more invasive endoscopy. Additionally, these DEGs show promise of indicating
380 degree of inflammation allowing clinicians to stratify severity of active disease and adjust treatment
381 plans accordingly. Zinc homeostasis dysregulation in eosinophilic disorders provides a potential
382 treatment target in a condition where presently there are few pharmacologic treatment options (54).
383 Patients with EoE are known to have zinc deficiencies associated with elimination diets (55), but
384 further investigation as to how these diets affect overall zinc regulation is needed.

385 This study demonstrates preliminary evidence for the utility of single-cell RNA-sequencing of patient
386 blood cells to characterize the genomic landscape of pediatric IBD subtypes and eosinophilic
387 disorders. These data indicate both overlapping and distinct DEGs, enriched Gene Ontology terms,
388 and enriched pathways associated the pathogenesis of inflammation among individuals with CD, UC,
389 EoD, and EoE, offering further insight into which genes and pathways may serve as useful markers
390 of disease in the peripheral blood mononuclear cells of patients.

391

392

393 **References**

- 394 1. Abraham C, Cho JH. MECHANISMS OF DISEASE Inflammatory Bowel Disease. *New Engl*
395 *J Med.* 2009;361(21):2066-78.
- 396 2. Silva FAR, Rodrigues BL, Ayrizono MDS, Leal RF. The Immunological Basis of
397 Inflammatory Bowel Disease. *Gastroent Res Pract.* 2016;2016.
- 398 3. Molodecky NA, Soon IS, Rabi DM, Ghali WA, Ferris M, Chernoff G, et al. Increasing
399 incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review.
400 *Gastroenterology.* 2012;142(1):46-54.
- 401 4. Kaplan GG. The global burden of IBD: from 2015 to 2025. *Nat Rev Gastroenterol Hepatol.*
402 2015;12(12):720-7.
- 403 5. Furuta GT, Katzka DA. Eosinophilic Esophagitis. *New Engl J Med.* 2015;373(17):1640-8.
- 404 6. Bass JA, Friesen CA, Deacy AD, Neilan NA, Bracken JM, Shakhnovich V, et al.
405 Investigation of potential early histologic markers of pediatric inflammatory bowel disease. *BMC*
406 *Gastroenterol.* 2015;15:1-7.
- 407 7. Jacobs I, Ceulemans M, Wauters L, Breynaert C, Vermeire S, Verstockt B, et al. Role of
408 Eosinophils in Intestinal Inflammation and Fibrosis in Inflammatory Bowel Disease: An Overlooked
409 Villain? *Front Immunol.* 2021;12:754413.
- 410 8. Berg DR, Colombel JF, Ungaro R. The Role of Early Biologic Therapy in Inflammatory
411 Bowel Disease. *Inflamm Bowel Dis.* 2019;25(12):1896-905.
- 412 9. Kelay A, Tullie L, Stanton M. Surgery and paediatric inflammatory bowel disease. *Transl*
413 *Pediatr.* 2019;8(5):436-48.
- 414 10. Urquhart SA, Quinn KP, Ravi K, Loftus EV, Jr. The Clinical Characteristics and Treatment
415 Outcomes of Concomitant Eosinophilic Esophagitis and Inflammatory Bowel Disease. *Crohns*
416 *Colitis* 360. 2021;3(2):otab018.
- 417 11. Sawczenko A, Ballinger AB, Savage MO, Sanderson IR. Clinical features affecting final
418 adult height in patients with pediatric-onset Crohn's disease. *Pediatrics.* 2006;118(1):124-9.
- 419 12. Heuschkel R, Salvestrini C, Beattie RM, Hildebrand H, Walters T, Griffiths A. Guidelines for
420 the management of growth failure in childhood inflammatory bowel disease. *Inflamm Bowel Dis.*
421 2008;14(6):839-49.
- 422 13. Greenley RN, Hommel KA, Nebel J, Raboin T, Li SH, Simpson P, et al. A meta-analytic
423 review of the psychosocial adjustment of youth with inflammatory bowel disease. *J Pediatr Psychol.*
424 2010;35(8):857-69.
- 425 14. Cohen NA, Rubin DT. New targets in inflammatory bowel disease therapy: 2021. *Curr Opin*
426 *Gastroenterol.* 2021;37(4):357-63.
- 427 15. Hart AL, Al-Hassi HO, Rigby RJ, Bell SJ, Emmanuel AV, Knight SC, et al. Characteristics of
428 intestinal dendritic cells in inflammatory bowel diseases. *Gastroenterology.* 2005;129(1):50-65.
- 429 16. Coombes JL, Powrie F. Dendritic cells in intestinal immune regulation. *Nat Rev Immunol.*
430 2008;8(6):435-46.

- 431 17. Chaudhry A, Rudensky AY. Control of inflammation by integration of environmental cues by
432 regulatory T cells. *J Clin Invest*. 2013;123(3):939-44.
- 433 18. Weaver CT, Elson CO, Fouser LA, Kolls JK. The Th17 pathway and inflammatory diseases
434 of the intestines, lungs, and skin. *Annu Rev Pathol*. 2013;8:477-512.
- 435 19. Ohnmacht C, Marquee R, Presley L, Sawa S, Lochner M, Eberl G. Intestinal microbiota,
436 evolution of the immune system and the bad reputation of pro-inflammatory immunity. *Cellular*
437 *Microbiology*. 2011;13(5):653-9.
- 438 20. Mishra A, Hogan SP, Lee JJ, Foster PS, Rothenberg ME. Fundamental signals that regulate
439 eosinophil homing to the gastrointestinal tract. *Journal of Clinical Investigation*. 1999;103(12):1719-
440 27.
- 441 21. Burczynski ME, Peterson RL, Twine NC, Zuberek KA, Brodeur BJ, Casciotti L, et al.
442 Molecular classification of Crohn's disease and ulcerative colitis patients using transcriptional
443 profiles in peripheral blood mononuclear cells. *J Mol Diagn*. 2006;8(1):51-61.
- 444 22. Martin JC, Chang C, Boschetti G, Ungaro R, Giri M, Grout JA, et al. Single-Cell Analysis of
445 Crohn's Disease Lesions Identifies a Pathogenic Cellular Module Associated with Resistance to Anti-
446 TNF Therapy. *Cell*. 2019;178(6):1493-508.
- 447 23. Chen P, Zhou GS, Lin JX, Li L, Zeng ZR, Chen MH, et al. Serum Biomarkers for
448 Inflammatory Bowel Disease. *Front Med-Lausanne*. 2020;7.
- 449 24. Kang HM, Subramaniam M, Targ S, Nguyen M, Maliskova L, McCarthy E, et al.
450 Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat Biotechnol*.
451 2018;36(1):89-94.
- 452 25. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary
453 learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol*.
454 2024;42(2):293-304.
- 455 26. Hao YH, Hao S, Andersen-Nissen E, Mauck WM, Zheng SW, Butler A, et al. Integrated
456 analysis of multimodal single-cell data. *Cell*. 2021;184(13):3573-87.
- 457 27. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential
458 expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139-40.
- 459 28. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis
460 and candidate gene prioritization. *Nucleic Acids Res*. 2009;37:W305-W11.
- 461 29. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING
462 database in 2023: protein-protein association networks and functional enrichment analyses for any
463 sequenced genome of interest. *Nucleic Acids Res*. 2023;51(D1):D638-D46.
- 464 30. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A
465 software environment for integrated models of biomolecular interaction networks. *Genome Res*.
466 2003;13(11):2498-504.
- 467 31. Turpin W, Goethel A, Bedrani L, Croitoru K. Determinants of IBD Heritability: Genes, Bugs,
468 and More. *Inflamm Bowel Dis*. 2018;24(6):1133-48.
- 469 32. Yen K, Lee C, Mehta H, Cohen P. The emerging role of the mitochondrial-derived peptide
470 humanin in stress resistance. *J Mol Endocrinol*. 2013;50(1):R11-R9.

- 471 33. Gordon-Lipkin EM, Banerjee P, Franco JLM, Tarasenko T, Kruk S, Thompson E, et al.
472 Primary oxidative phosphorylation defects lead to perturbations in the human B cell repertoire. *Front*
473 *Immunol.* 2023;14:1142634.
- 474 34. Bangert C, Villazala-Merino S, Fahrenberger M, Krausgruber T, Bauer WM, Stanek V, et al.
475 Comprehensive Analysis of Nasal Polyps Reveals a More Pronounced Type 2 Transcriptomic Profile
476 of Epithelial Cells and Mast Cells in Aspirin-Exacerbated Respiratory Disease. *Frontiers in*
477 *Immunology.* 2022;13.
- 478 35. Taft J, Markson M, Legarda D, Patel R, Chan M, Malle L, et al. Human TBK1 deficiency
479 leads to autoinflammation driven by TNF-induced cell death. *Cell.* 2021;184(17):4447-63.
- 480 36. Hachim MY, Elemam NM, Ramakrishnan RK, Salameh L, Olivenstein R, Hachim IY, et al.
481 Derangement of cell cycle markers in peripheral blood mononuclear cells of asthmatic patients as a
482 reliable biomarker for asthma control (vol 11, 11873, 2021). *Sci Rep-Uk.* 2021;11(1).
- 483 37. Cao SS. Endoplasmic Reticulum Stress and Unfolded Protein Response in Inflammatory
484 Bowel Disease. *Inflamm Bowel Dis.* 2015;21(3):636-44.
- 485 38. Li C, Grider JR, Murthy KS, Bohl J, Rivet E, Wiegand N, et al. Endoplasmic Reticulum
486 Stress in Subepithelial Myofibroblasts Increases the TGF- β 1 Activity That Regulates Fibrosis in
487 Crohn's Disease. *Inflamm Bowel Dis.* 2020;26(6):809-19.
- 488 39. Huang SS, Xie Z, Han J, Wang HL, Yang G, Li MY, et al. Protocadherin 20 maintains
489 intestinal barrier function to protect against Crohn's disease by targeting ATF6. *Genome Biol.*
490 2023;24(1).
- 491 40. López-Pacheco C, Soldevila G, Du Pont G, Hernández-Pando R, García-Zepeda EA. CCR9 Is
492 a Key Regulator of Early Phases of Allergic Airway Inflammation. *Mediat Inflamm.* 2016;2016.
- 493 41. Schmitz I. Gadd45 Proteins in Immunity. *Gadd45 Stress Sensor Genes.* 2013;793:51-68.
- 494 42. Su XQ, Pan J, Bai FX, Yuan HL, Dong N, Li DD, et al. IL-27 attenuates airway inflammation
495 in a mouse asthma model via the STAT1 and GADD45 γ /p38 MAPK pathways. *J Transl Med.*
496 2016;14.
- 497 43. Lohoff M, Giaisi M, Köhler R, Casper B, Krammer PH, Li-Weber M. Early Growth
498 Response Protein-1 (Egr-1) Is Preferentially Expressed in T Helper Type 2 (Th2) Cells and Is
499 Involved in Acute Transcription of the Th2 Cytokine Interleukin-4. *J Biol Chem.* 2010;285(3):1643-
500 52.
- 501 44. Ricci M, Matucci A, Rossi O. IL-4 as a key factor influencing the development of allergen-
502 specific Th2-like cells in atopic individuals. *J Invest Allerg Clin.* 1997;7(3):144-50.
- 503 45. Barbato JC, Catanescu O, Murray K, DiBello PM, Jacobsen DW. Targeting of
504 metallothionein by L-homocysteine - A novel mechanism for disruption of zinc and redox
505 homeostasis. *Arterioscl Throm Vas.* 2007;27(1):49-54.
- 506 46. Fukui H, Horie M, Endoh S, Kato H, Fujita K, Nishio K, et al. Association of zinc ion release
507 and oxidative stress induced by intratracheal instillation of ZnO nanoparticles to rat lung. *Chem Biol*
508 *Interact.* 2012;198(1-3):29-37.
- 509 47. Huang KL, Lee YH, Chen HI, Liao HS, Chiang BL, Cheng TJ. Zinc oxide nanoparticles
510 induce eosinophilic airway inflammation in mice. *J Hazard Mater.* 2015;297:304-12.

- 511 48. Suzuki M, Ramezanzpour M, Cooksley C, Lee TJ, Jeong B, Kao S, et al. Zinc-depletion
512 associates with tissue eosinophilia and collagen depletion in chronic rhinosinusitis. *Rhinology*.
513 2020;58(5):451-9.
- 514 49. Devirgiliis C, Zalewski PD, Perozzi G, Murgia C. Zinc fluxes and zinc transporter genes in
515 chronic diseases. *Mutat Res-Fund Mol M*. 2007;622(1-2):84-93.
- 516 50. Jenkins D, Seth R, Kummer JA, Scott BB, Hawkey CJ, Robins RA. Differential levels of
517 granzyme B, regulatory cytokines, and apoptosis in Crohn's disease and ulcerative colitis at first
518 presentation. *J Pathol*. 2000;190(2):184-9.
- 519 51. Kotsafti A, D'Incà R, Scarpa M, Fassan M, Angriman I, Mescoli C, et al. Weak Cytotoxic T
520 Cells Activation Predicts Low-Grade Dysplasia Persistence in Ulcerative Colitis. *Clin Transl*
521 *Gastroen*. 2019;10.
- 522 52. Gu D, Cao T, Yi S, Li X, Liu Y. Transcription suppression of GABARAP mediated by
523 lncRNA XIST-EZH2 interaction triggers caspase-11-dependent inflammatory injury in ulcerative
524 colitis. *Immunobiology*. 2024;229(3):152796.
- 525 53. Xu M, Kong Y, Chen N, Peng W, Zi R, Jiang M, et al. Identification of Immune-Related
526 Gene Signature and Prediction of CeRNA Network in Active Ulcerative Colitis. *Front Immunol*.
527 2022;13:855645.
- 528 54. Tamarit-Sebastian S, Ferrer-Soler FM, Lucendo AJ. Current options and investigational drugs
529 for the treatment of eosinophilic esophagitis. *Expert Opin Investig Drugs*. 2022;31(2):193-210.
- 530 55. Votto M, De Filippo M, Lenti MV, Rossi CM, Di Sabatino A, Marseglia GL, et al. Diet
531 Therapy in Eosinophilic Esophagitis. Focus on a Personalized Approach. *Front Pediatr*.
532 2021;9:820192.
- 533