

Examining the cell transcriptional distribution of Adrenomedullin and its receptors in ileal and colonic Crohn's disease

Aaliya Ashraf

123112981

Project Thesis in partial fulfilment for the degree of Master's in Bioinformatics
and Computational Biology

Supervised by Dr. Silvia Melgar, APC Microbiome Ireland

Submitted to the Department of Microbiology, University

College Cork on 01/10/2024



UCC

Coláiste na hOllscoile Corcaigh
University College Cork, Ireland



Contents

List of Abbreviation	iv
Acknowledgement	vi
Abstract.....	vii
Chapter 1. INTRODUCTION	1
1.1 Inflammatory Bowel Disease (IBD)	1
1.2 Types of IBD	1
1.2.1 Pediatric IBD.....	1
1.3 Crohn's disease	2
1.4 Epidemiology of Chron's Disease.....	2
1.5 Types of Crohn's Disease	2
1.5.1 Physiological evidence	3
1.6 Treatment Efficiency	6
1.7 Adrenomedullin	7
1.7.1 Structure and Biosynthesis of Adrenomedullin	7
1.7.2 Adrenomedullin Receptors	9
1.7.3 Adrenomedullin distribution in the gastrointestinal tract.....	10
1.7.4 Adrenomedullin and Inflammatory Bowel Disease	10
1.7.5 AM as an Anti-Inflammatory Factor.....	11
1.7.6 AM's Role in Maintaining Intestinal Epithelial Barrier Integrity	11
1.7.7 AM Improves Mucosal Healing and Re-Epithelialization	12
1.7.8 Adrenomedullin as a Therapy for IBD.....	13
1.8 Single-Cell RNA Sequencing	13
1.8.1 Pre-processing of data	14
1.8.2 Quality control	14
1.8.3 Normalization.....	15
1.8.4 Regressing out biological effects	15
1.8.5 Feature selection.....	16
1.8.6 Dimensionality reduction.....	16
1.8.7 Visualization	17
1.8.8 Downstream analysis	17
1.9 Hypothesis and Aim of the study.....	20
Chapter 2. MATERIALS AND METHODS	21
2.1 Overview of the dataset.....	21

2.1.1 Human Adult Dataset.....	21
2.1.2 Pediatric Dataset.....	22
2.2 Data Sources	22
2.3 Human Adult dataset Analysis	22
2.4 Pediatric dataset Analysis	24
Chapter 3: RESULTS.....	25
3.1 Quality metric control of intestinal datasets	25
3.2 High Variable Genes.....	25
3.4 Gene Expression Analysis of Stromal Cell Population.....	27
3.5 Immune Cell Population.....	40
3.6 Epithelial cell Population.....	46
3.7 Pediatric Cell Population.....	52
4. Discussion and Conclusion	61
5. REFERENCES	64
6. APPENDIX	74

List of Abbreviation

IBD	Inflammatory Bowel disease
CD	Crohn's disease
UC	Ulcerative Colitis
TI	Terminal ileum
PIBD	Pediatric Inflammatory Bowel disease
scRNA-seq	Single cell RNA Sequencing
AM	Adrenomedullin peptide
ADM	Adrenomedullin gene
CALCRL	Calcitonin Receptor-Like Receptor
RAMP2	Receptor Activity-Modifying Protein 2
RAMP3	Receptor Activity-Modifying Protein 3
ACKR3	Atypical Chemokine Receptor 3
CGRP	Calcitonin Gene-Related Peptide
CT	Calcitonin
CXCL11	C-X-C motif chemokine ligand 11
CXCL12	C-X-C motif chemokine ligand 12
LPS	Lipopolysaccharide
QC	Quality control
PCA	Principal Component Analysis
UMAP	Uniform Manifold Approximation and Projection
t-SNE	t-Distributed Stochastic Neighbour Embedding
DEA	Differential Expression Analysis
DEG	Differentially Expressed Gene
KEGG	Kyoto Encyclopedia of Genes and Genomes
GO	Gene Ontology
MAST	Model-based Analysis of Single-cell Transcriptomics
GLM	Generalized Linear Model
GSEA	Gene Set Enrichment Analysis
HVG	Highly Variable Gene
GI	Gastrointestinal
DC	Dendritic cells

AMPs	Antimicrobial Peptides
IECs	Intestinal Epithelial Cells
NK	Natural Killer (Cells)
TLRs	Toll-Like Receptors
JAK	Janus Kinase
PAMP	Pro-Adrenomedullin N-terminal 20 Peptide
MR-proADM	Mid-Regional Pro-Adrenomedullin
GPCRs	G-Protein Coupled Receptors
COPD	Chronic Obstructive Pulmonary Disease
UMI	Unique Molecular Identifier
KNN	K-Nearest Neighbors
ROIs	Regions of Interest
ORA	Over-Representation Analysis
VST	Variance Stabilizing Transformation
BBKNN	Batch Balanced K-Nearest Neighbors

Acknowledgement

First, I would like to express my deepest gratitude to my supervisor, Dr. Silvia Melgar, for her invaluable guidance, support and mentorship throughout the course of my thesis. I feel privileged to have received an opportunity to work on such an interesting project in her lab. I would also extend my sincere thanks to co-supervisor, Ms. Klara Martinovic, for her constant guidance, advice and critical input in my study which greatly improved my quality of work.

I am deeply thankful to Dr. Tadhg Crowley and Dr. Janaki Velmurugan for their role in providing me with right direction for my approaches in this journey. Also, to all lab mates and my friend, Glenn Ross who made their assistance which made it more manageable in challenging situations. Lastly, extending thanks to Marcus Claesson for coordinating this course and project.

A special note of gratitude for my friends and family for their encouragement, patience and understanding throughout this time.

Abstract

Crohn's Disease (CD) is a chronic inflammatory condition of the gastrointestinal (GI) tract. Ileal CD compared to colonic CD poses a higher risk for delayed diagnosis and potential complications, and with its different immune response profile, it is an important target for new treatment strategies, especially as it has been shown for different biological therapies that there is trend of lower efficacy in ileal CD compared to colonic CD.

Adrenomedullin (AM) is a small peptide hormone with various biological functions including vasodilatation, angiogenesis, anti-inflammation, and potentiation of host defences against microbes. Its potential therapeutic has been reported in preclinical models and in pilot clinical trials in patients with IBD.

This study aims to assess the hypothesis that ADM plays a differential, tissue-specific role in regulating inflammation in Crohn's disease, particularly between the terminal ileum and colon, and that its action varies between adult and pediatric patients. Human adult and pediatric single-cell RNASeq datasets were analyzed using Seurat and Scanpy packages. The expression patterns and potential roles of ADM and its key receptors (CALCRL, RAMP2, RAMP3, and ACKR3) in the stromal, immune, and epithelial cell populations were evaluated. The predominant expression of ADM and its key receptors were observed in the stromal cell population among all the cell types in both terminal ileum and colon in the adult CD data set, as well as in the pediatric dataset.

The results showed that ADM is expressed in a variety of stromal cell populations, with fibroblasts exhibiting the highest expression. In contrast, endothelial cells were the primary source of expression for CALCRL, RAMP2, and RAMP3. Only macrophages and monocytes showed significant expression of ADM in immune cells, which is probably due to an inflammatory response mechanism. ADM expression in secretory epithelial cells, especially goblet cells, raises the possibility that it maintains the homeostasis with the release of anti-microbial peptides and regulate immune response locally and preserves mucosal integrity.

To sum up, the significant expression of the receptors in endothelial cells emphasizes the role of ADM in endothelial-mediated processes in the gastrointestinal tract. This insight provides a foundation for further research into the therapeutic potential of targeting the ADM signalling pathway in gastrointestinal diseases.

Chapter 1. INTRODUCTION

1.1 Inflammatory Bowel Disease (IBD)

The gastrointestinal (GI) tract is the primary site of inflammation-related disorders known as inflammatory bowel diseases (IBDs), which are immune-mediated conditions that are chronic and characterized by flare-ups and remissions (Bouhuys et al., 2023). Multiple factors, including host genotype, environment, microbiome, and immune system, interact intricately to cause IBDs (Guan, 2019).

1.2 Types of IBD

The two main types of inflammatory bowel diseases (IBDs) are ulcerative colitis (UC) and Crohn's disease (CD). While ulcerative colitis produces chronic inflammation in mostly one part of digestive tract, primarily the colon. Crohn's disease is an IBD that causes inflammation anywhere along the lining of the digestive tract (Fakhoury et al., 2014). The symptoms of IBD can differ based on its type. Individuals who suffer from ulcerative colitis frequently have diarrhoea and lower left abdominal pain. They might thus lose weight and have blood on a rectal examination. On the other hand, compared to ulcerative colitis, patients with Crohn's disease typically experience lower right abdominal pain and less rectum bleeding. The most frequent side effect of Crohn's disease is intestinal blockage brought on by swelling, which thickens the intestinal wall. Furthermore, because of poor absorption, individuals afflicted by this disease frequently experience issues associated with malnourishment or the existence of nutritional deficiencies (Fakhoury et al., 2014).

1.2.1 Pediatric IBD

Over the past few decades, pediatric IBD (PIBD) has become more common, particularly in nations where its incidence was found to be low (Kuenzig et al., 2022). When comparing children with adults, the disease phenotype in both CD and UC varies. The most common presentation of paediatric CD is an inflammatory or non-stricturing, non-penetrating illness. In pediatric CD, stricturing and penetrating disease is comparatively rare compared to the adults. However, numerous investigations have demonstrated that, in many children, CD advances to stricturing and penetrating illness even with treatment (Sauer and Kugathasan, 2010).

Although IBD primarily affects young adults, it can strike anyone at any age, and 25% of sufferers will first show symptoms before turning 20. Adolescence is the time when IBD is most common in children; nonetheless, 20% of these children will present before the age of ten, and 5% before the age of five (Rosen et al., 2015). Most of the world is seeing an increase in the incidence of IBD, with childhood onset IBD being more common. However, onset of ulcerative colitis in the pediatric IBD is more extensive and also follows a more progressive course (Aloi et al., 2013). Several recent studies have found that in some regions of Canada but not all where there have been historically high rates of both

pediatric and adult onset IBD, there has been a rapid growth in the incidence of very early onset IBD (Sýkora et al., 2018; Kuenzig et al., 2022)

1.3 Crohn's disease

Crohn's disease is recognized as a persistent inflammatory gastrointestinal disorder, exhibiting recurrent episodes of symptom development. Additionally, it is a progressive illness that causes bowel damage and disability. The terminal ileum and proximal colon are most frequently affected, but any part of the gastrointestinal tract can be impacted. Segmental, asymmetrical, and transmural inflammation are the most common types. Most patients have an inflammatory phenotype when they are first diagnosed, but half of them will eventually develop complications (strictures, fistulas, or abscesses), which frequently requires surgery (Peyrin-Biroulet et al., 2010; Thia et al., 2010).

1.4 Epidemiology of Chron's Disease

The highest incidence of IBD has been reported in North America, northern and western Europe, and Oceania (Ng et al., 2017). The incidence of CD is 0–20.2 cases per 100,000 person-years in North America and 0.3–12.7 cases per 100,000 person-years in Europe (Ng et al., 2017). The incidence of IBD has increased worldwide since the turn of the twenty-first century, with newly industrialized countries in Asia, Africa, and South America reporting rates that are rapidly rising (Ng et al., 2017). IBD was once rare in China, but after the country became more urbanized, it became common and accounted for increased use of hospital bed (Kaplan and Ng, 2016). In China's mainland, the incidence of CD is distributed east to west and south to north (Zhao et al., 2013). Global challenges for illness prevention, health care delivery, and disease diagnosis arise from the rapid changes in CD epidemiology. In recently industrialized nations (like Asia), the rising prevalence of CD is a reflection of the impact of the Western lifestyle, especially food, urbanization, and industrialization on risk (Linares de la Cal et al., 1999). During the last decade, it was observed that IBD emerged rapidly in eastern Europe and Asia but still the western Europe has higher prevalence of IBD. Despite the number of cohorts are being low in the eastern Europe, the gap between the east and west is getting narrower with increase in the IBD in eastern countries (Mak et al., 2020).

1.5 Types of Crohn's Disease

The terminal ileum (TI) and colon (CO) are the primary sites of CD involvement. According to new research, ileal-dominant and colonic CD should be classified as distinct disease subtypes (Cleynen et al., 2016; Dulai et al., 2019). This emphasizes that it is crucial to comprehend whether and how the cellular mechanisms underlying ileal and colonic inflammation are different (Kong et al., 2023).

Nonetheless, a number of lines of evidence make it abundantly evident that the manifestation of colonic is not the same as ileal Crohn's disease. The Montreal classification of Crohn's disease already takes into account this clinically significant differentiation. In addition to phenotype and age at onset, the

classification also includes stratification by location terminal ileum: (L1), colon: (L2), ileocolonic (L3) and when disease is present in upper gastrointestinal location (L4) (Satsangi, 2006). Two thirds (57–89 percent) of patients have ileal localization, and at least one third have isolated ileal disease, according to the Vienna Classification, which divides the disease site into terminal ileum, colon, and ileum and colon (Gasche et al., 2000).

1.5.1 Physiological evidence

To better understand the differences between ileal and colonic Crohn's disease in terms of symptoms and function, it's important to first look at the normal physiological differences between the terminal ileum and the colon. By examining how disease affects the various layers, starting with the luminal surface and the epithelial layer, and then moving to the mucosa, we can gain a clearer picture of how Crohn's disease behaves differently in these two regions of the gut. (Atreya and Siegmund, 2021). The major macroscopic and histological changes is in healthy and Crohn's disease tissue is shown in Figure 1 and 2.

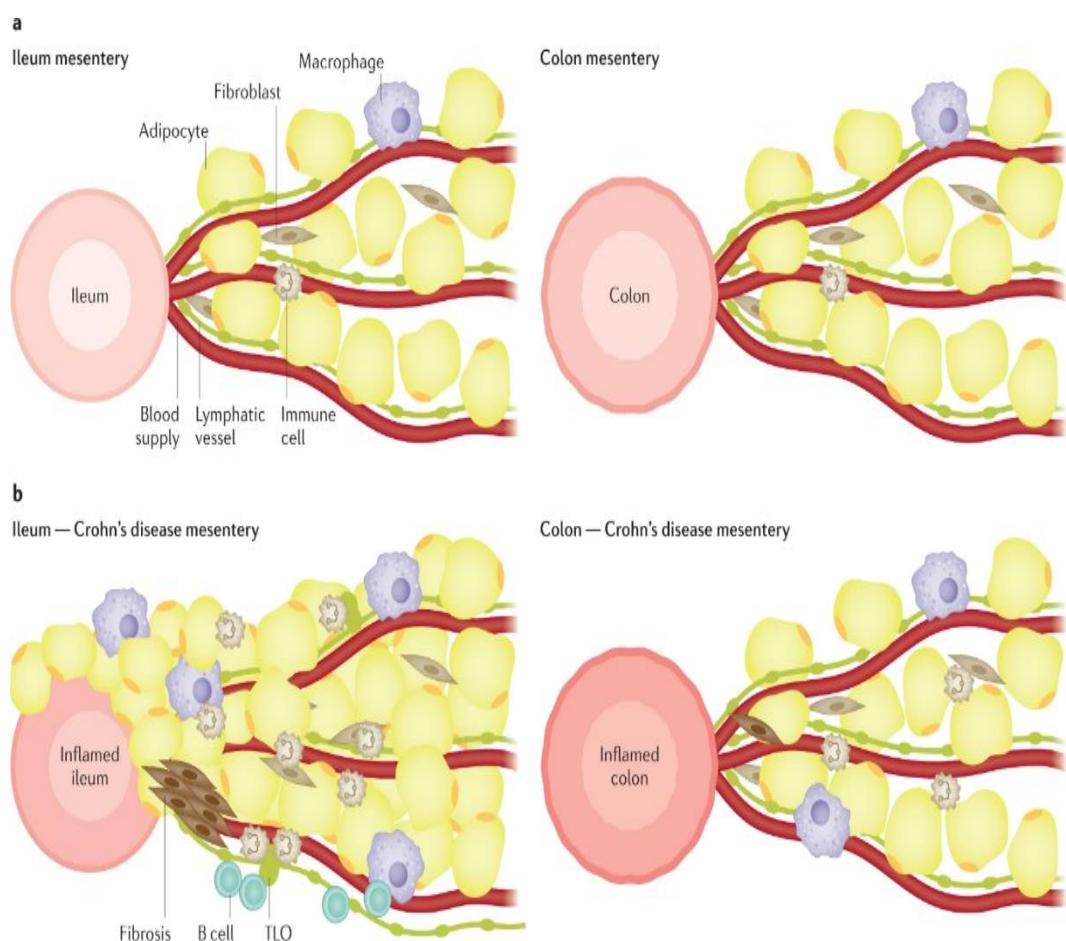


Figure 1. Mesentery structure in healthy and Crohn's disease state. a / Depicts the mesentery in a healthy state, containing immune cells, mesenchymal cells, and adipocytes. The mesentery next to the colon and the mesentery next to the ileum do not appear to differ significantly. b / The ileum's mesentery undergoes distinctive alterations in Crohn's disease, such as adipocyte hyperplasia, wrapping of the inflammatory intestinal segment (creeping fat), fibrosis, and a robust immune cell infiltration. B cells and innate lymphoid cells infiltrate the lymphatic wall, and tertiary lymphoid organs (TLOs) are located adjacent to the lymphatic vessel in creeping fat. Adipocytes are normal in size, the immune cell infiltrate is only slightly developed, and the

mesentery of the inflamed colon is only loosely attached. The colon's mesentery in Crohn's disease is similar to that of ulcerative colitis (Atreya and Siegmund, 2021).

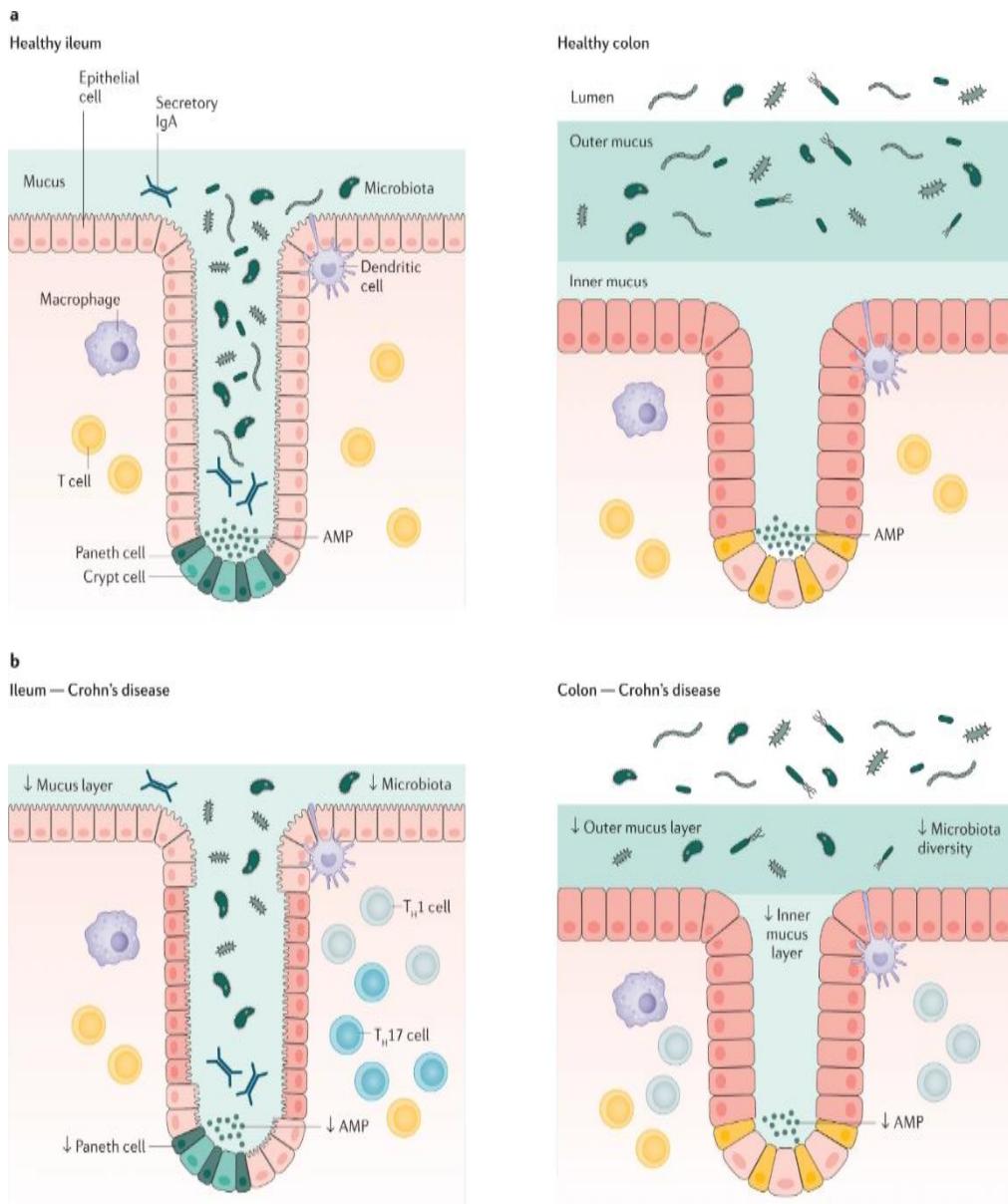


Figure 2. Distinctions in structure and immunity between the colon and ileum in a healthy state and Crohn's disease. a / The intestinal microbiota can partially penetrate the single layer of mucus that covers the ileum. Antimicrobial peptides (AMP) and IgA are present in mucus, which serves as the body's first line of defense against translocalizing bacteria. The crypt's Paneth cells facilitate the AMP's production. The mucus in a healthy colon, on the other hand, is composed of two layers, the luminal layer of which is permeable to microbiota but not the layer beneath. Again, secretory cells produce AMP, which is present in mucus as a defense mechanism. b / Crohn's disease site-specific alterations, demonstrating a reduction in the mucus layer and composition, a decline in the diversity of gut microbiota and AMP, a disruption of the epithelial barrier, and a predominance of T helper 1 (TH1) and TH17 cells in the terminal ileum's lamina propria. In the colon, epithelial cells are partially disrupted, the two mucus layers equally decrease and change in composition, the microbiota's diversity varies depending on the site, and the lamina propria is dominated by the TH1 response (Atreya and Siegmund, 2021).

There are certain differences in the cellular makeup of the epithelium in the small and large intestines. The epithelium of the small intestine differs from that of the large intestine in that it contains more M

cells and Paneth cells in the crypts (Bowcutt, 2014). The M cells are engaged in the transfer and presentation of luminal antigens to immune cells, while the Paneth cells are specialized in the secretion of anti-microbial peptides (AMPs) (Bowcutt, 2014). Due to different populations and gene expression profiles of intestinal epithelial cells (IECs), the secreted AMPs exhibit specificities in the small and large intestines (Gallo and Hooper, 2012). The epithelium of the small intestine is distinguished by the release of lysozyme/α-defensins/ phospholipase A2 (Paneth cells), as well as regenerates islet-derived protein REG3γ (Paneth cells and enterocytes) while Enterocytes secrete β-defensins and cathelicidins, which are characteristics of the epithelium of the large intestine (Gallo and Hooper, 2012). In the large intestine, the Goblet cells which are specialized in secreting mucus are more numerous than in the small intestine epithelium (Pierre et al., 2021). In addition, enteroendocrine cells (<1%) make up the IECs (Gunawardene et al., 2011). Compared to the colon, the small intestine and the rectum exhibit a higher frequency of these cells. Furthermore, the small and large intestines exhibit different morphology and hormone secretion profiles in enteroendocrine cells (Sjölund et al., 1983).

Immune cells are primarily found in the lamina propria throughout the gastrointestinal tract, with the small intestine having a higher density of these cells than the large intestine (Bowcutt, 2014). In Crohn's disease, the immune system is characterized by an imbalance between pro-inflammatory effector T cells (mainly Th1 and Th17) and regulatory T cells (Tregs). Th1 and Th17 cells, which secrete cytokines like interferon-γ, TNFα, and interleukin-17, dominate and drive inflammation, while Tregs that produce anti-inflammatory cytokines such as IL-10 and TGF-β are less effective in controlling inflammation. This dysregulation contributes to the chronic intestinal inflammation seen in Crohn's disease (Baumgart and Sandborn, 2012). Additionally, macrophages respond to the intestinal environment by producing cytokines like IL-12 and IL-23, which activate natural killer (NK) cells, further intensifying inflammation and perpetuating the immune response in the gut. This dysregulation contributes to the chronic intestinal inflammation seen in Crohn's disease (Roda et al., 2020).

In Crohn's disease, stromal cells, particularly fibroblasts and other mesenchymal populations, undergo significant alterations compared to healthy individuals. There is a diminished capacity for fibroblasts to migrate and a reduction in the number of stromal cells that support epithelial cells in patients with inflammatory bowel disease (IBD). These stromal cells actively interact with the gut microbiota, both directly through toll-like receptors (TLRs) and indirectly via microbiota-reactive memory T cells, leading to the production of pro-inflammatory factors (Barnhoorn et al., 2020). Stromal fibroblasts are central to tissue inflammation, as they produce a substantial amount of signaling molecules and growth factors. These mediators not only drive inflammation but also activate endothelial cells, which further exacerbates vascular inflammation and recruits immune cells to the inflamed tissue (Enzerink and Vaheri, 2011). The continuous interaction between fibroblasts and endothelial cells plays a vital role in maintaining the chronic inflammatory state observed in IBD.

1.6 Treatment Efficiency

Most treatments currently in use for managing CD and its symptoms are acting towards suppressing the immune response that is overly active in the intestine. Some common drugs are: Aminosalicylates, Steroids, Immunosuppressants, Antibiotics and Biologics (Gade et al., 2020). It is not uncommon for patients with Crohn's disease to undergo multiple therapies, as treatment response and drug efficacy are dependent on the location, extent, activity of the disease, and the presence of complications (Hart and Ng, 2015).

As patients who achieve mucosal healing have better outcomes, such as a lower risk of surgery, a lower relapse rate, and an improved quality of life (Bernstein et al., 2019), mucosal healing is currently the preferred treatment goal. Over the last twenty years, the management of CD has undergone a significant change due to the introduction of anti-inflammatory treatments like anti-tumour necrosis factor (anti-TNF) therapy, which includes infliximab, adalimumab, and certolizumab. These medications are now the therapy of choice and are being used earlier in the course of the disease, particularly in patients who have a high risk of the disease progressing. The need for parenteral administration and the possibility of immunogenicity are significant disadvantages of these targeted biologic therapies, despite the fact that they represent a significant advancement in the treatment of CD (Roda et al., 2020). Additional biologic therapies for CD include ustekinumab, an antagonist of IL-12 and IL-23 signaling, for induction therapy and keeping remission in patients with moderate-to-severe CD, and vedolizumab, a gut-selective monoclonal anti-integrin antibody (Torres et al., 2017).

Sphingosine-1-phosphate receptor 1 (S1PR1) agonists, such as ozanimod and etrasimod, and selective Janus kinase (JAK) inhibitors, such as filgotinib and upadacitinib, are among the many oral small molecules that are being investigated as potential CD therapies (Peyrin-Biroulet et al., 2017; D'Amico et al., 2018). Further research in CD is necessary for non-absorbable antibiotics like rifaximin, even though the antibiotic regimens that have been studied this far have not consistently shown efficacy. Furthermore, faecal microbiota transplantation and other methods of diet-based microbiota manipulation may also be beneficial and are under investigation (Roda et al., 2020).

About one-third of patients with Crohn's disease have isolated ileal disease, which has been identified in several studies as a sign of possible complications. Compared to patients with isolated colonic involvement, ileal disease location is more frequently linked to the development of a penetrating disease phenotype, increased risk of intestinal complications and increased likelihood of repeated surgeries (Atreya and Siegmund, 2021). In order to stop the progression of intestinal damage and disability in these patients, optimized anti-inflammatory therapy is crucial. Limited data points out the better therapeutic effectiveness of sulfasalazine and antibiotics in colonic patients, but enteral nutrition appears to be more effective in treating ileal and ileocolonic Crohn's disease symptoms. According to

available data, isolated ileal Crohn's disease is less effectively treated with infliximab than colonic Crohn's disease (Atreya et al., 2022). These limited number of findings have been reported not only for anti-TNF agents, but for other biologics such as vedolizumab, ustekinumab, and rizankizumab as well. In order to clarify site-specific mechanisms in ileal and colonic Crohn's disease and link them to therapeutic response, more research is required overall (Atreya and Siegmund, 2021). Additionally, specialized clinical trials that investigate the efficacy of treatments in isolated Crohn's disease patients only are required. To reflect the therapeutic response, these studies must include translational studies. Optimizing the therapy for individual patient will only be possible with a deeper understanding of the molecular mechanisms underlying both colonic and ileal Crohn's disease (Atreya et al., 2022), but novel options are required, because existing treatments have their limitations.

1.7 Adrenomedullin

Although the precise cause of IBD is still unknown, several inflammatory pathways as well as cellular and microbiota contributions have been found. As a result, novel treatments that take advantage of these disparate observations have been developed (Coskun et al., 2017).

Kitamura et al. found that human pheochromocytomas contained the endogenous vasodilatory peptide Adrenomedullin (AM) in 1993 (Kitamura et al., 1993). Following some time spent studying AM as a circulatory agonist, it was found that AM also supported immune function through its anti-inflammatory effects, activating angiogenesis, and organ preservation. AM successfully cures enteritis and stomach ulcers in animal models. It is also extensively expressed in the gastrointestinal tract's epithelia. They therefore began doing translational research on the clinical application of AM in IBD (Ashizuka et al., 2013). As an endogenous bioactive peptide, AM is regarded as reasonably safe, has a low immunogenicity, and could be used in the future to treat IBD in a novel way.

1.71 Structure and Biosynthesis of Adrenomedullin

The human adrenomedullin gene (ADM) is situated at the distal end of chromosome 11's short arm (p15.1-3) (Ishimitsu et al., 1994). AM is amidated at Tyr⁵² in the C-terminal, has a ring structure (Figure 3) with a disulfide bond between Cys¹⁶ and Cys²¹, and is 52 amino acids long (Kitamura et al., 1993). Calcitonin (CT) and the two forms of calcitonin gene-related peptide (CGRP), amylin and adrenomedullin2/intermedin, are highly conserved in both disulfide bonds and amidation, which are essential for bioactivity (Wimalawansa, 1997; Takei et al., 2004). For this reason, AM is regarded as a member of the superfamily CT/CGRP (Wimalawansa, 1997; Takei et al., 2004). Chromosome 11 contains the four exons that make up the AM gene (Ishimitsu et al., 1994). Pro-adrenomedullin (proAM), which includes amino acid residues 22–185, is created from AM, a large preprohormone (Figure 4). It is then broken down into four segments: proAM N-terminal 20 peptide (PAMP)-Gly, mid-regional pro-adrenomedullin (MR-proADM, 45–92), AM-Gly, and C-terminal proAM (adrenotensin)

(Kitamura et al., 1993). When PAMP and AM are first metabolised, they are both in inactive biological intermediate forms called C-terminally Glycine-Extended Forms. These peptides are changed to their mature bioactive forms by amidating their C-terminal glycine using an amidation enzyme; however, only a fraction of the peptides undergo this conversion (Kitamura et al., 1998).

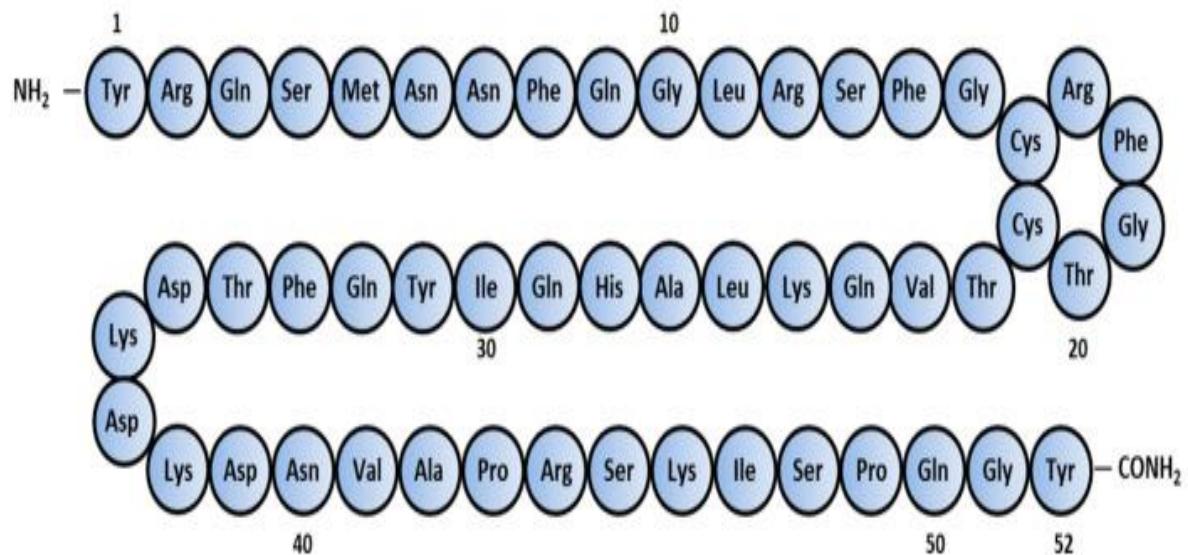


Figure 3. Structure of Adrenomedullin (ADM) (Ashizuka et al., 2021)

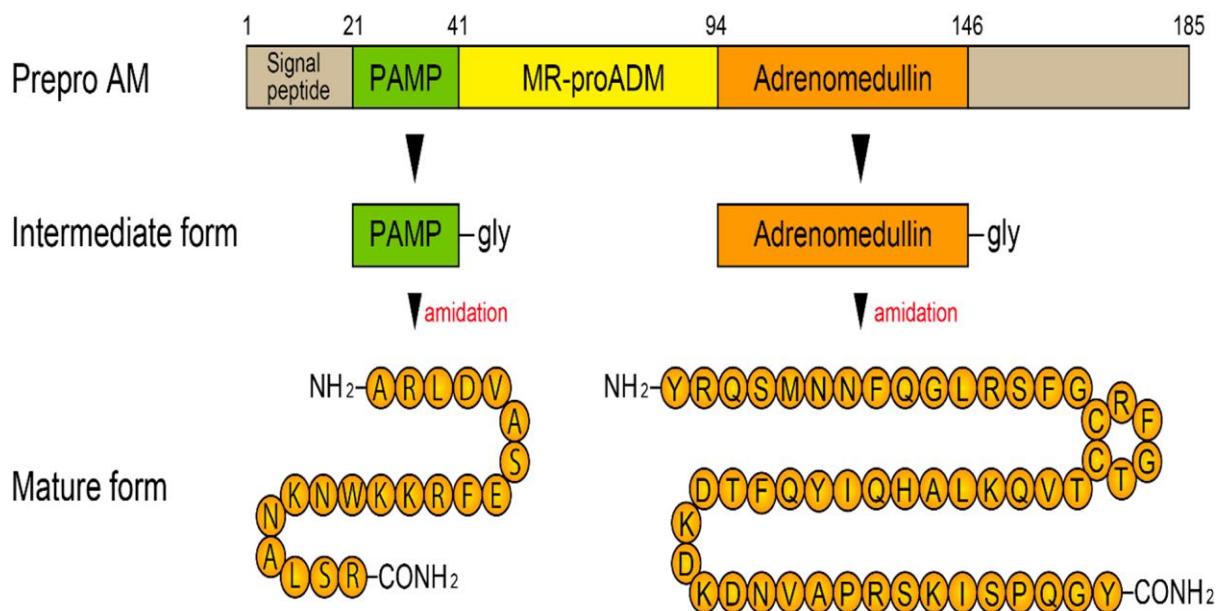


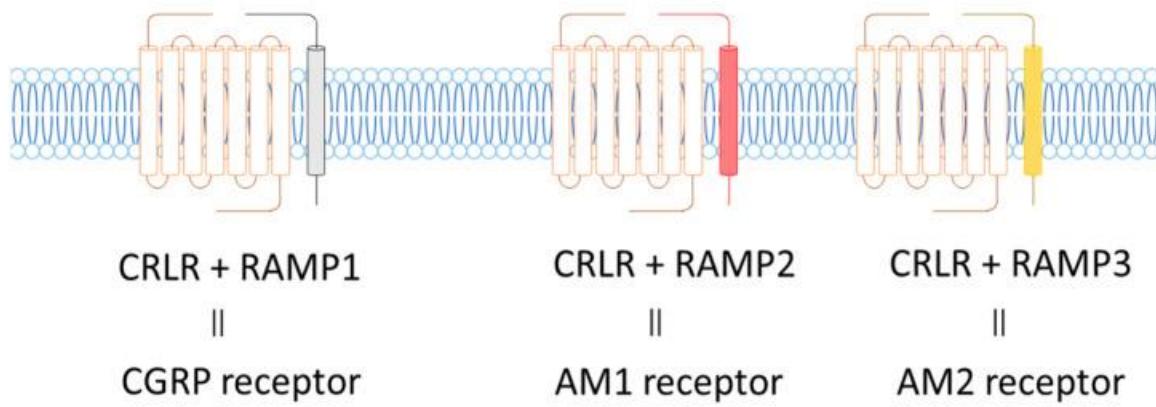
Figure 4. Structure and processing of Adrenomedullin (ADM) (Kita and Kitamura, 2022).

1.7.2 Adrenomedullin Receptors

The G protein-coupled receptors (GPCRs) and the chaperone molecules known as receptor activity modifying proteins (RAMPs) make up the receptors for the CT/CGRP family. There are three types of RAMPs (1, 2, and 3) (Figure 5) and two forms of GPCRs, calcitonin receptor (CTR) and calcitonin receptor-like receptor (CRLR), for the CT/CGRP family. As a CT receptor, CTR works on its own, but the RAMPs together with CALCR forms functional receptor to which AM binds. For instance, the amylin receptor is represented by CTR + RAMPs, the CGRP receptor by CRLR + RAMP1, the AM1 receptor by CRLR + RAMP2, and the AM2 receptor by CRLR + RAMP3. When it comes to AM1 and AM2 receptors, AM has equal affinity. Similar high affinity for the AM2 receptor is shown by AM and AM2/intermedin (Fischer et al., 2020).

It has been determined that amino acid 93 in RAMP1 is primarily responsible of its affinity for CGRP, whereas amino acid 74 in RAMP2 and RAMP3 is essential for their affinity for AM (Qi et al., 2008). RAMP2 expression is more prevalent in a physiological condition. It appears that the increase in RAMP3 expression is a mechanism to reduce AM's reactivity in conditions like pregnancy, sepsis, or heart failure where AM levels are most increased. Depending on the degree of response of AM, the balance between the RAMP2 and RAMP3 can change in different particular cell types (Gibbons et al., 2007).

Recent studies have shown that endothelial cell specific RAMP2 KO mice have proven that AM-RAMP2 system plays a major role in vascular integrity and homeostasis (Koyama et al., 2013).



CRLR = calcitonin receptor-like receptor

RAMP = receptor activity-modifying protein

Figure 5. Receptor complexes for calcitonin receptor-like receptor (CRLR) with receptor activity-modifying proteins (RAMP) (Ashizuka et al., 2021).

Although the main receptors for AM are AM1 and AM2 receptors, it has been shown that AM also has affinity for binding to atypical chemokine receptor 3 (ACKR3) (Meyrath et al., 2021). Atypical

chemokine receptors (ACKRs) are a type of decoy receptor that negatively regulates chemokine function. There are four different kinds of ACKRs reported: ACKR1, ACKR2, ACKR3, and ACKR4 (Graham et al., 2012). ACKRs may also function as AM/PAMP receptors, according to recent research (Meyrath et al., 2021). It appears that AM is the only member of the CGRP peptide superfamily that activates ACKR3 to a moderate degree (Meyrath et al., 2021). The chemokines CXCL11 and CXCL12 bind to the decoy receptor ACKR3, which then induces β -arrestin recruitment and ligand internalization (Serafin et al., 2020). Additionally, AM uses ACKR3 as a decoy receptor to cause β -arrestin recruitment (Serafin et al., 2020). Surprisingly, PAMP behaves similarly to AM in terms of ACKR3 activity. Moreover, PAMP-12, its shortened analog, exhibits higher potency toward ACKR3 than AM (Meyrath et al., 2021).

1.7.3 Adrenomedullin distribution in the gastrointestinal tract

Adrenomedullin is ubiquitously present in the human body (salivary glands, esophagus, stomach, duodenum, jejunum, ileum, cecum, colon, gall bladder, bile duct, and pancreas) and it has various biological effects (Martínez-Herrero and Martínez, 2016). It has been shown that it also has multiple functions in GI tract, e.g. in the oral cavity where it increases keratinocytes growth and acts as antimicrobial peptide (Gröschl et al., 2009), in the stomach where it regulates gastric emptying, and gastric acid release (Rossowski et al., 1997) and inhibits insulin secretion in pancreas (Martínez et al., 1996), and most importantly in the intestine where it regulates bowel movement via smooth muscle cells (Fukuda et al., 1998), regulating sugar absorption by the enterocytes (Fernández De Arcaya et al., 2005) and antimicrobial effect on both Gram-positive and Gram-negative bacteria (Zudaire et al., 2006).

Using techniques such as immunostaining, it has been shown that AM is present in all layers of the intestine, the intestinal nerves and smooth muscle cells (Fukuda et al., 1998; M. Zhou et al., 2001). However, enteroendocrine cells have been linked to the colon's predominant form of AM immunoreactivity (Mulder et al., 1996). In porcine GI tract duodenum and ileum, AM levels are four to fourteen times higher than those in other GI tissues such as large intestine, stomach and esophagus, and they are more abundant in the mucosa and submucosa (Kiyomizu et al., 2001).

1.7.4 Adrenomedullin and Inflammatory Bowel Disease

Adrenomedullin is becoming known as a novel and potentially effective treatment for inflammatory digestive pathologies like IBD. This is directly tied to AM's capacity to have both systemic and localized anti-inflammatory effects. For instance, it has been shown that AM prevents peripheral blood monocytes from secreting pro-inflammatory cytokines into the medium when they transform into macrophages (Martínez-Herrero and Martínez, 2016). Not only does AM regulate immune cells, but it also suggested reduces endothelial permeability, which in turn reduces the production of inflammatory exudates (Hippenstiel et al., 2002). Additionally, as a strong angiogenic factor (Martínez, 2006), AM is important in preserving the integrity of the mucosa's microvasculature (Koyama et al., 2013), and as

previously observed, it speeds up the healing and restoration of epithelial lesions (Fukuda et al., 1999). And lastly, because of its antimicrobial properties, AM may be able to aid in the fight against the overabundance of microbiota bacteria that develops in IBD patients (Zudaire et al., 2006) (Martínez-Herrero and Martínez, 2016). There are two benefits to using AM therapeutically, first of all, the body naturally produces this hormone and secondly, exogenous administration of AM has been tried in earlier pilot and clinical trials, with negligible to no adverse effects (Troughton et al., 2001; Kataoka et al., 2010)

1.7.5 AM as an Anti-Inflammatory Factor

Peripheral blood monocytes express the ADM gene, which is rapidly up-regulated when these cells develop into macrophages (Kubo et al., 1998). Numerous studies have shown that ADM inhibits the secretion of proinflammatory cytokines (Isumi et al., 1999; Wu et al., 2003).

At the local and systemic levels, it functions as a strong anti-inflammatory factor, contributing to the development of the Th1/Th2 cytokine balance, inhibiting neutrophil infiltration into the affected area, and lowering pro-inflammatory cytokine levels (mainly – IFN- γ , IL-6, IL-10, and TNF- α) (Ashizuka et al., 2005, 2009; Talero et al., 2008). These effects might be mediated by AM's control over HIF activity; high AM levels encourage the synthesis of HIF-1 α , an endogenous mucosal protective factor that prevents inflammation (MacManus et al., 2011).

AM expression has been demonstrated to increase in a variety of cells in response to a variety of pro-inflammatory cytokines, including IL-1 α and IL-1 β , TNF- α , and TNF- β , as well as bacterial products like lipopolysaccharide (LPS) and hypoxia (Isumi et al., 1998; Tomoda et al., 2001; Poyner, 2002). Most notably, high expression of this peptide has also been shown *in vivo* in animals and humans with severe infections (Temmesfeld-Wollbrück et al., 2007). Specifically, elevated expression is noted in animals exposed to LPS as well as in sepsis and septic shock. The small intestine was found to be a significant source of AM release during polymicrobial sepsis in a rat model of cecal ligation and puncture (Zhou et al., 2001), and high expression was noted in the lung in endotoxemia (Cheung et al., 2004) as well as in acute lung injury brought on by hypoxia and LPS (Agorreta et al., 2005). Furthermore, it has been demonstrated that AM functions as a strong antimicrobial peptide against both Gram-positive and Gram-negative bacteria (Allaker et al., 2006). All these findings revealed AM has different effects depending on the disease/inflammation contexts.

1.7.6 AM's Role in Maintaining Intestinal Epithelial Barrier Integrity

Like the endothelial barrier dynamic interactions between the actin cytoskeleton, epithelial cell-cell junctions and various signalling pathways regulate the intestinal epithelial barrier (Figure 6). Acute inflammation or, in more severe cases, chronic conditions like asthma, chronic obstructive pulmonary

disease (COPD), and IBD can result from dysregulation of these well-orchestrated interactions (García-Ponce et al., 2016).

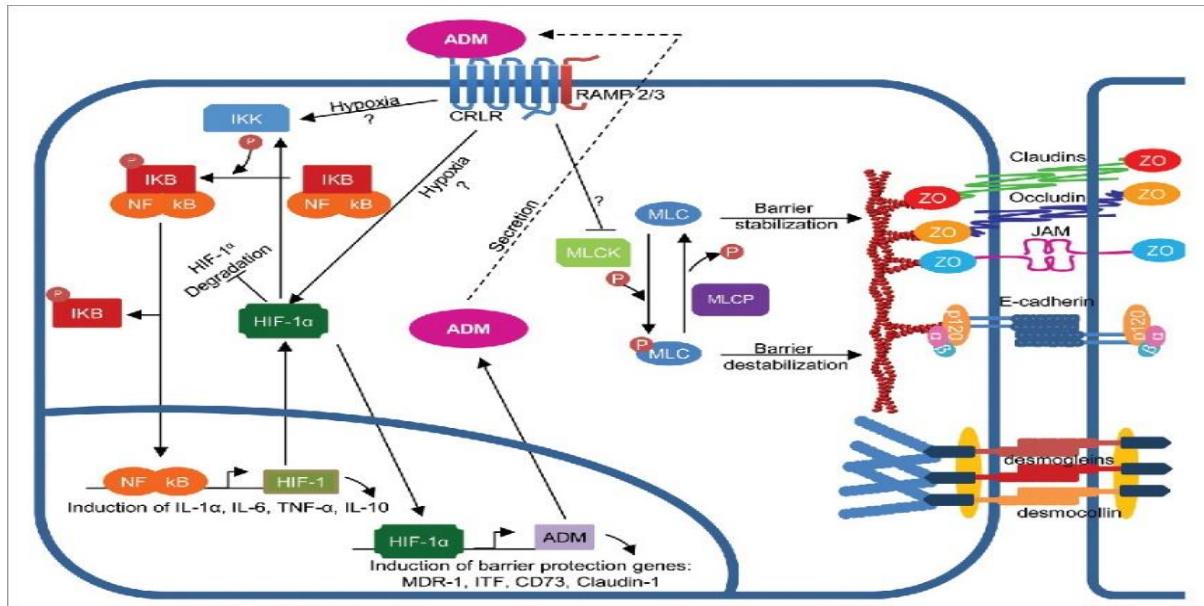


Figure 6. Signaling pathways through which adrenomedullin (ADM) regulates the epithelial barrier. Stabilization of the epithelial barrier is aided by ADM signaling. ADM triggers NF κ B, which then triggers the transcription of HIF-1 α and cytokines during hypoxia. ADM prevents HIF-1 α from degrading and permits it to move into the nucleus, which stimulates the transcription of genes that protect the barrier and more ADM, creating a positive feedback loop. Moreover, ADM suppresses MLCK activity to lower p-MLC and prevent the internalization of TJ and AJ proteins. (García-Ponce et al., 2016)

Additionally, AM plays a crucial role in maintaining the integrity of the intestinal epithelial barrier by preventing the intestinal epithelium's hyperactivation and hyperpermeability, two conditions that are indicative of IBD (Ashizuka et al., 2009). According to recent research, AM protects against intestinal epithelial barrier dysfunction by downregulating myosin light chain phosphorylation, a critical regulator of intestinal barrier function, and suppressing inflammatory cytokines (Yi et al., 2015)

1.7.7 AM Improves Mucosal Healing and Re-Epithelialization

AM facilitates and accelerates the restoration of damaged rat and human mucosa's epithelium (Fukuda et al., 1999). In the stomach, the quick restoration of epithelial integrity serves as a crucial natural defence against superficial damage. According to in vitro research, AM improves the restoration of mucosal integrity following mild to moderate gastric mucosal damage brought on by exposure to a hyperosmolar NaCl solution (Fukuda et al., 1999). Since AM is produced in the gastric mucosa's chief cells, it is likely secreted in large amounts to hasten the healing process following mucosal damage. Hashimoto et al. provided confirmation of this theory (Oksche et al., 2000) and came to the conclusion that AM levels in histocytes, smooth muscle cells, and glandular epithelia are lower during the active stage of a gastric ulcer and gradually increase during the healing and scarring stages after analysing 82 cases of the condition. These findings imply that AM overexpression might be helpful for vascular

smooth muscle cells (VSMCs) proliferation, blood flow regulation to the mucosa, and tissue repair (Oksche et al., 2000).

1.7.8 Adrenomedullin as a Therapy for IBD

At the forefront of adrenomedullin usage in treating IBD is a Japanese research group (Kita et al., 2022) which showed that AM has beneficial effects IBD experimental rodent model. They even conducted the first clinical trial in IBD patients and verified AM's efficacy in patients with steroid-resistant UC in a phase 2a clinical trial (Kita et al., 2021). One interesting example is a patient with CD who was resistant to infliximab, he was given AM medication in addition to his medication for seven days and he reached remission state (Ashizuka et al., 2019). Because AM acts through a novel mechanism that does not rely on excessive immunosuppression and promotes mucosal regeneration, it may be an effective treatment for biologic-resistant CD similar to the mentioned example (Kita et al., 2022).

1.8 Single-Cell RNA Sequencing

Since its invention more than ten years ago, RNA sequencing, or RNA-seq, has fundamentally changed our understanding of biology (Stark et al., 2019). Over this time, RNA-seq from bulk tissue has established itself as the accepted method for examining the transcriptome of various organisms. Bulk tissue RNA-seq, however, is not appropriate for identifying cell variability between individual cells or characterizing uncommon cell types (Nayak and Hasija, 2021). A growing number of studies have switched from bulk tissue experiments to single-cell RNA sequencing (scRNA-seq) as a result of these limitations and recent technological advancements (Stark et al., 2019). Owing to the enormous potential of scRNA-seq technology, a number of computational tools have been created to handle various facets of data processing. Although these tools can be used with multiple programming languages, R and Python are the most commonly utilized languages (Luecken and Theis, 2019). Among them, Seurat (Stuart et al., 2019) is a popular R package for processing scRNA-seq data, cell clustering and finding differentially expressed genes (DEGs) in one or more samples.

While there is a growing trend towards cross-environment support (Scholz et al., 2018), selecting a programming language frequently involves selecting between different analysis tools. Popular platforms with extensive analysis toolboxes and integrated environments for pipeline development are Seurat (Butler et al., 2018), Scater (McCarthy et al., 2017), and Scanpy (Wolf et al., 2018). Nevertheless, these platforms are forced to restrict themselves to tools created in the corresponding programming languages. The pipeline of scRNA-seq (single cell RNA sequencing) analysis is depicted in figure 7 which majorly includes data pre-processing, clustering and downstream analysis.

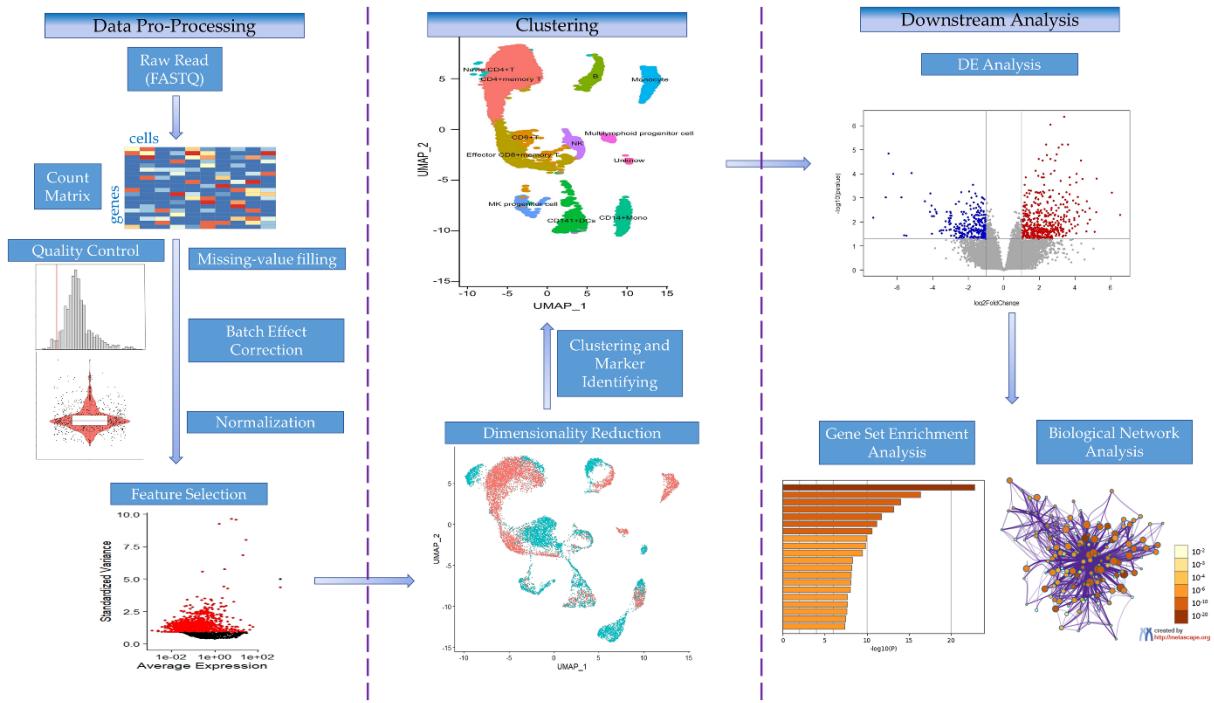


Figure 7. The figure illustrates the scRNA-seq workflow, starting with Data Pre-Processing (from raw reads to quality control and normalization). Next is Clustering, using the UMAP to group cells and identify markers. Finally, Downstream analysis includes differential expression, gene set enrichment, and biological network analysis (Xue et al., 2024).

1.8.1 Pre-processing of data

The primary distinction between preprocessing bulk RNA-seq data and scRNA-seq data is the need to handle different DNA barcodes, which help identify sequence reads and associate them with the original molecule or cell (You et al., 2021). The majority of methods work with unique molecular identifier (UMI) (Islam et al., 2014) based data. Starting with raw fastq files, typical scRNA-seq preprocessing workflows include demultiplexing, mapping, transcript quantification, and quality control (You et al., 2021). After the sequencing step, the first output of the FASTQ files is pre-processed by QC and read alignment. One of the most widely used QC tools is FASTQC and Cutadapt. FASTQ files from the sequencing machine is taken as input file and returns a quality report which is then utilized to trim the low quality reads using Cutadapt (Zhang et al., 2021).

After sequencing, the raw data can be processed into tables consisting of number of molecules or reads detected per cell. These tables represent molecular counts or read counts depending on whether unique molecular identifiers (UMIs) were used in the single cell experiment (Luecken and Theis, 2019). Pipelines for processing raw data include Cell Ranger (Zheng et al., 2017), indrops (Klein et al., 2015), SEQC (Azizi et al., 2018) or zUMIs (Parekh et al., 2018).

1.8.2 Quality control

One of the crucial phases in the preprocessing of scRNA-seq data is quality control (QC), which has the main goals of determining and eliminating low-quality cells, evaluating the quality of the data

produced from scRNA-seq experiments, and guaranteeing the data's dependability for further analysis (Hong et al., 2022)

There is a set of quantitative measure which is used to assess the quality of the scRNA-seq data generated from each cell called qc metrics. Low quality cells, doublets and batch effects can be filtered with the help of these qc metrics and hence, it benefits the researcher to take informed decisions with the dataset (Raza, 2024). To make sure all cellular barcode data correspond to viable cells before analyzing the single-cell gene expression data, quality control is performed. The three QC covariates that are typically used to perform cell QC are 1) the fraction of counts from mitochondrial genes per barcode, 2) the number of counts per barcode (count depth), and 3) the number of genes per barcode (Ilicic et al., 2016; Griffiths et al., 2018). These QC covariate distributions are inspected for outlier peaks, which are then eliminated through thresholding. These barcode outliers may indicate doublets, dying cells, or cells with ruptured membranes.

1.8.3 Normalization

Normalization is essential for scRNA-seq analysis technique development and for mitigating bias or technical noise. Many normalization techniques, from modifications of bulk sequencing techniques to completely new approaches created especially for single-cell studies, have been used during the years since scRNA-seq has been developed (Lytal et al., 2020).

These techniques often highlight the distinctions between technical noise (e.g., as a result of inaccurate measurements), and biological/medical variation, which is attributable to inherent variations between cells that are in the same biological/medical condition (Kim et al., 2015).

Apart from the previously mentioned techniques, the NormalizeData function from the Seurat (Satija et al., 2015) R package is used to obtain a baseline comparison for normalization. This global normalization technique known as Simple Norm in the plots that follow—divides the gene counts for each cell by default, multiplies by the scale factor and natural log, and then $\log(x+1)$ transforms the result to account for zero counts (Lytal et al., 2020).

1.8.4 Regressing out biological effects

Technical variability estimates should be included after normalized gene expression levels or molecular counts have been produced. This is true for any downstream analysis, but it's crucial when comparing cell-to-cell expression levels or evaluating the genetic variability of individual genes (Stegle et al., 2015).

Batch effects are technical confounders that need to be taken into account in order for the true biological signal to emerge, more like variations in sequencing depth. Biological techniques like single cell RNA-seq usually deals with batch effects, which are caused by variations in non-biological factors

such as the time of the experiment, the person conducting it, or the reagents used. Batch effects can be mistaken for genuine biological signals if they are not taken into consideration, but they can be completely avoided with careful experimental design (Baran-Gale et al., 2018).

Annotation databases have identified certain genes that are known to be involved in the cell cycle and the cell cycle has a significant impact on their expression. A common term for these gene is cell-cycle genes (Scialdone et al., 2015). Removing the cell cycle's effects on the transcriptome is the most popular biological data correction. Platforms like Scanpy and Seurat (Butler et al., 2018; Wolf et al., 2018) can adjust the data using simple methods like linear regression based on cell cycle scores. More advanced tools, such as scLVM (Buettnner et al., 2015) and f-scLVM (Buettnner et al., 2017), use more complex models for this correction. Researchers use lists of marker genes from the literature to calculate these cell cycle scores (Macosko et al., 2015).

1.8.5 Feature selection

A class of computational techniques known as feature selection aims to choose a subset of valuable features from the total feature set in the dataset. Feature selection is a useful technique for reducing feature dimension and redundancy in high-dimensional data, which can help with problems like model overfitting in downstream analysis (Yang et al., 2021).

More specifically, Seurat and PanoView first classify genes based on mean expressions into 20 bins, and then identify the most variable genes, also known as highly variable genes, or HVGs within each bin (Su et al., 2021). Feature selection is usually the first step in reducing the complexity of scRNA-seq data. This process involves keeping only the genes that show the most variation in the dataset. To do this, researchers often focus on HVGs (Brennecke et al., 2013).

Typically, between 1,000 and 5,000 HVGs are selected for further analysis, depending on the task and how complex the data is. Early findings by Klein and colleagues suggest that the exact number of HVGs chosen does not greatly affect the results of the analysis (Klein et al., 2015).

1.8.6 Dimensionality reduction

Data from scRNA-seq are sparse, noisy, and high dimensional. A crucial stage in the downstream analysis of scRNA-seq is dimension reduction. As a result, numerous techniques for dimension reduction have been created (Xiang et al., 2021).

Dedicated dimensionality reduction algorithms can further reduce the dimensions of single-cell expression matrices after feature selection. Principal component analysis (PCA) (Pearson, 1901) and diffusion maps (Coifman et al., 2005), are two common methods for reducing the complexity of data. These techniques, made popular for single-cell analysis (Haghverdi et al., 2015), work by summarizing the data into reduced dimensions.

When performing PCA, all of the data are projected into hyperdimensional space and then subjected to a linear transformation so that the first few new axes, or principle components (PCs), capture the largest variance. The loadings can be used to determine each gene's contribution to each PC, though this isn't always clear because PCA works in both positive and negative space. Additionally, the matrix of HVGs from the scRNA-Seq data may not always fit into PCA's assumptions of linear and normally distributed data (Andrews and Hemberg, 2018). PCA typically uses its top N principal components to summarize a dataset; N can be found using the permutation-test-based jackstraw method (Chung and Storey, 2015; Macosko et al., 2015) or elbow heuristics. One advantage of PCA's simple approach is that the distances in the reduced data can be understood consistently. This allows us to check how important certain factors are by comparing them with the principal components. For example, principal components can be mapped with technical nuisance covariates to see how well data correction, normalization, and quality control worked (Büttner et al., 2019) or to identify which genes are most important in the dataset (Chung and Storey, 2015).

1.8.7 Visualization

Non-linear dimensionality reduction techniques are widely used for visualization purposes. For scRNA-seq visualization, the most popular dimensionality reduction technique is the t-distributed stochastic neighbour embedding (t-SNE) (Zhou and Jin, 2020). At the expense of overall structure, t-SNE dimensions concentrate on capturing local similarity. The Uniform Approximation and Projection method (McInnes et al., 2018) and graph-based tools like SPRING (Weinreb et al., 2018) are popular substitutes for t-SNE. Perhaps the best approximation of the underlying topology is ForceAtlas2, a force-directed layout algorithm developed by SPRING and UMAP (Wolf et al., 2019). In this comparison, UMAP stands out due to its speed and scalability to a large number of cells (Becht et al., 2019). For exploratory data visualization, we therefore consider UMAP to be best practice when there are no specific biological questions. Additionally, UMAP is capable of summarizing data in multiple dimensions. While there are no known application of UMAP for data summarization, it could be a good alternative to PCA. (Luecken and Theis, 2019).

1.8.8 Downstream analysis

Following pre-processing, biological insights are extracted and the underlying biological system is described using techniques that refer to as downstream analysis.

1.8.8.1 Cluster analysis

Cluster formation is usually the initial intermediate process of a single-cell analysis. We can deduce member cell identities from clusters. Cells are grouped into clusters according to how similar their gene expression profiles are.

Notably, to prevent the introduction of artifacts, cell subpopulation identification should be done after quality control and normalization of scRNA-seq data. Methods for cell clustering can be broadly

classified into two groups based on whether or not previous information is utilized. In cases where a predetermined set of markers was employed for clustering, the techniques rely on previous data (Chen et al., 2019). As an alternative, cell populations with scRNA-seq data can be identified de novo using unsupervised clustering techniques. The four main categories of unsupervised clustering algorithms include graph-based clustering, hierarchical clustering, density-based clustering and k-means (Andrews and Hemberg, 2018).

The clustering technique used by the Scanpy and Seurat single-cell analysis platforms is the default one. It has been demonstrated to perform better than alternative clustering techniques for flow and mass cytometry data (Weber and Robinson, 2016) as well as scRNA-seq data (Freytag et al., 2018; Duò et al., 2020).

One example of typical clustering process is the implementation of the Louvain and Leiden algorithms by the popular Seurat toolkit. The standard protocol involves the several steps. First, log-transformed and normalized counts are subjected to principal component analysis. Then, the euclidean distance between the first 30 principal components of each cell pair is computed. After that 20 nearest neighbours are identified for each cell. Each cell pair is assigned a weight based on the number of neighbours they have in common. This weight is then used to define weighted edges in a network and finally the network is divided into clusters that maximize modularity (Waltman and Van Eck, 2013). The resolution parameter in the optimized modularity function lets the user choose how big the cluster partition will be. It is also possible to subcluster only specific clusters by subsetting the K-Nearest Neighbour (KNN) graph. Sub clustering like this can help the user find patterns that are solely the result of data noise, but it can also help the user know cell states within cell type clusters (Wagner et al., 2016).

1.8.8.2 Cluster annotation

Clustered data are analyzed at the gene level by identifying each cluster's gene signature. These so-called marker genes provide a biological label to the cluster and are used to characterize it. The identity of the cells within the cluster is represented by this biological label (Luecken and Theis, 2019)

Reference database information can be used to annotate clusters in two ways: either by using full gene expression profiles or by using data-derived marker genes. Differential expression (DE) testing can be used to compare two groups, the cells in one cluster and all other cells in the dataset. This helps identify marker genes that are specific to that cluster. Usually, we concentrate on genes that exhibit up-regulation within the relevant cluster. Simple statistical tests like the t-test or the Wilcoxon rank-sum test are frequently used to rank genes according to their difference in expression between two groups because marker genes are predicted to have strong differential expression effects. Marker genes are those that rank highest in the corresponding statistic test. By using enrichment tests, the Jaccard index, or other overlapping statistics to compare marker genes from the reference dataset and the dataset's marker

genes, clusters can be annotated (Luecken and Theis, 2019). To aid in cell-identity annotation, reference webtools like dropviz (Saunders et al., 2018) and mousebrain (Zeisel et al., 2018) allow users to see how marker genes in the dataset are expressed in the reference dataset.

1.8.8.3 Differential expression analysis

The main downstream analysis on scRNA-seq data is called differential expression analysis (DEA). DEA is helpful in finding biomarkers for new cell types or gene signatures that show difference in specific cells. It also serves as an input for other secondary analyses like network analysis and gene set or pathway analysis (Das et al., 2022).

Determining quantitative differences between groups or conditions through DEA is crucial to comprehend the molecular underpinnings of phenotypic variation. However, it is not immediately clear that we can just apply standard methods designed for bulk-cell data (Mou et al., 2020). Scientists can now more clearly address even the most basic question about DE thanks to the scRNA-seq technology. The reasons to why these bulk RNA-seq cannot be directly applied to scRNA-seq data are due to its unique characteristics, which include higher variability, multi-modal distribution that cannot be attributed to the zero counts, and excessive zero counts for both biological and technical reasons (Bacher and Kendziorski, 2016)

Model-based Analysis of Single-cell Transcriptomics (MAST), is one of the most widely used scRNA-seq method which uses a bimodal distribution with expression strongly different from zero or "non-detectable" to explicitly account for the dropouts, and suggests a generalized linear model (GLM) to fit the data (Finak et al., 2015). To model changes in gene expression dependent on condition and technical covariates, MAST models dropout using a hurdle model. In the previously mentioned study (Soneson and Robinson, 2018), it was the best-performing single-cell DE testing method. In a small-scale comparison on a single dataset, it performed better than bulk and single-cell methods (Vieth et al., 2017). Weighted bulk methods take longer to run than MAST, which has 10–100 times fold faster runtime (Van Den Berge et al., 2018), further limma-voom can be used which achieves a 10 fold speedup (Law et al., 2014). Despite being a bulk DE testing technique, limma-voom was demonstrated to perform on par with MAST (Luecken and Theis, 2019).

1.8.8.4 Pathway Annotation

One of the most popular methods for analysing list of genes or genome-wide regions of interest (ROIs) obtained from different high-throughput investigations is functional enrichment analysis (Dozmorov, 2017). While a plethora of tools have been created for gene-centric or epigenomic enrichment analysis, the majority are made for use with model organisms or certain domains like plants (Yi et al., 2013), fungi (Priebe et al., 2015), integrated with specific annotations, such those found in the Kyoto Encyclopaedia of Genes and Genomes (KEGG) and Gene Ontology (GO) (Nam and Kim, 2008).

Initially released in 2012, the clusterProfiler (Yu et al., 2012) library was created to compare functional profiles of different circumstances on a single level (e.g., different treatment groups) and to do over-representation analysis (ORA) (Boyle et al., 2004) using GO and KEGG for several model organisms. Since then, clusterProfiler has undergone significant development and is now capable of supporting thousands of species with the most recent gene annotations, several ontologies and pathways, user-submitted annotation data for new species, and newly emergent annotations. Supported analyses include gene set enrichment analysis (GSEA) (Subramanian et al., 2005) and ORA. A complicated experimental design that enables comparison of functional profiles of diverse conditions on different levels is supported by an extension of the comparative utility.

1.9 Hypothesis and Aim of the study

In this study, two comprehensive scRNA-Seq datasets were utilized to explore the transcriptional landscapes of human intestinal tissues (ileum and colon) across different age groups (pediatric and adults) and disease states (healthy and Crohn's disease) (Elmentait et al., 2020; Kong et al., 2023). These papers have so far explored numerous aspects of CD and the bioinformatic approaches used to analyze single-cell RNA sequencing datasets. While CD pathogenesis involves a complex mechanism and treatment strategies, but the precise role and mechanisms utilized by adrenomedullin as a therapeutic target in Crohn's disease on a local system level is still unknown.

The following analysis aims to assess the hypothesis that ADM plays a differential, tissue-specific role in regulating inflammation in Crohn's disease, particularly between the terminal ileum and colon, and that its action varies between adult and pediatric patients. Utilizing scRNA-seq analysis technique on both the human adult and pediatric dataset, ADM's role through receptor-mediated mechanisms will be investigated.

The aim for this study includes:

1. To identify the alterations in expression of adrenomedullin and its known receptors in human GI tract from patients with CD compared to normal controls in different locations of human GI tract (small intestine vs colon) and different cell fractions (epithelium, immune and stroma) with a focus on their characteristics in the ileum.
2. To assess if there is a difference in adrenomedullin and its receptor expression across different age groups (Human adult vs. Pediatric).
3. To examine ADM and its receptor expression in tissue-specific changes (non-inflamed vs. inflamed).
4. Elucidate potential signalling pathways related to ADM's effects on barrier function and healing and Crohn's disease.

Chapter 2. MATERIALS AND METHODS

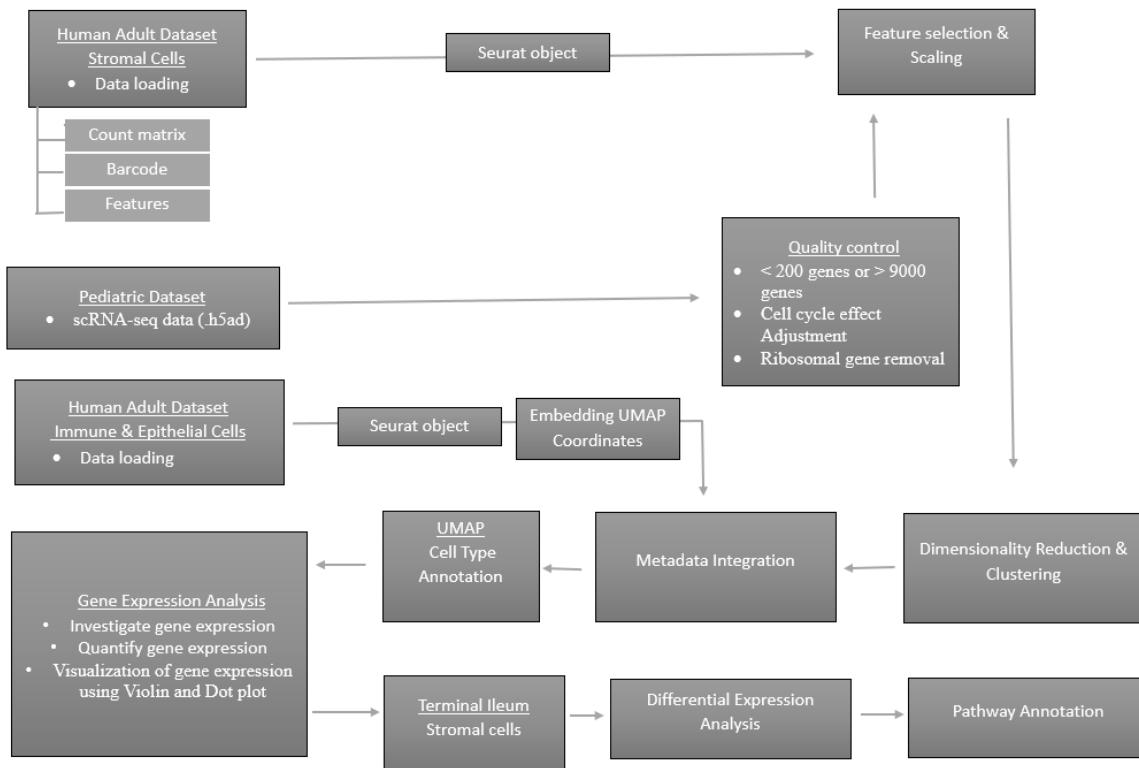


Figure 8. Flowchart illustrating the single cell RNA-Sequencing analysis pipeline for the Human Adult and Pediatric Stromal, Immune, and Epithelial Cell datasets.

2.1 Overview of the dataset

2.1.1 Human Adult Dataset

Lingjia Kong and colleagues (2023) performed a study using scRNA-Seq profiles from 720,633 cells isolated from both the terminal ileum (TI) and colon of 71 adult donors. The cohort includes both healthy individuals and Crohn's disease (CD) patients, offering a comprehensive cellular and molecular view of the inflammation process in CD. Specifically, 136 samples were collected, with 46 samples coming from Crohn's disease patients and 25 from non-IBD controls. The samples are distributed across three conditions: healthy, non-inflamed, and inflamed. For the terminal ileum, 10 healthy patients provided 13 samples, 26 non-inflamed patients contributed 49 samples, and 12 inflamed patients yielded 16 samples. For the colon, 16 healthy patients provided 32 samples, 17 non-inflamed patients contributed 20 samples, and 5 inflamed patients yielded 6 samples (Kong et al., 2023).

The tissue samples were collected with informed consent and ethical approval at Massachusetts General Hospital during the time of colonoscopy of healthy patients and endoscopy procedure of patients with Crohn's disease. The non-inflamed and inflamed condition were collected with distinction of tissue by the simple endoscopic score. The samples underwent mechanical and enzymatic dissociation to create

the single cell suspensions. The single cell suspensions from each sample were loaded on a Chromium controller (10X Genomics) and were processed using either the v2 or single-indexed v3.1 chemistry (Kong et al., 2023).

2.1.2 Pediatric Dataset

Elmentait and Ross et al. (2020) examined pediatric single-cell transcriptomic profiles derived from 22,500 cells collected from mucosal biopsies of the terminal ileum from two groups of children: healthy controls ($n = 8$, aged 4-12 years) and patients newly diagnosed with Crohn's disease (CD) ($n = 7$, aged 9-14 years). These tissue samples were obtained during colonoscopies at Addenbrooke's Hospital, Cambridge, UK, following informed consent and ethical approval. The samples underwent enzymatic dissociation to create single-cell suspensions, which were processed using the 10x Genomics Chromium 3'v2 workflow (Elmentait et al., 2020).

2.2 Data Sources

The Human Adult dataset was sourced from the Broad Single Cell Portal (SCP1884), where the normalized count matrix was utilized for analyzing the dataset. Similarly, the Pediatric dataset was obtained from the GutCellAtlas website, and the normalized file from this resource was employed for the subsequent data analysis. Figure 8 describes the methodology used to analyze the collected data sets using Seurat and Scanpy packages.

2.3 Human Adult dataset Analysis

The single-cell RNA sequencing (scRNA-seq) data for stromal cells isolated from the terminal ileum and colon was first loaded into R using the Seurat v5 package (Hao et al., 2023). The count data matrix, barcode files, and feature annotation were read and stored in appropriate formats. Seurat object was created for downstream analysis, ensuring that the counts matrix was appropriately labelled with unique gene identifiers and cell barcodes.

Quality control metrics, including the number of detected features, total RNA counts per cell, and mitochondrial gene content, were visualized using violin plots and were proceeded for further analysis. Since the data was in the normalized form, highly variable genes were identified using the variance-stabilizing transformation (VST) method, and these genes were subsequently used for principal component analysis. Then PCA was employed to reduce dimensionality, with the first 20 principal components retained for downstream analysis based on the elbow plot.

Cell metadata, including cell type annotations, library preparation details, and diagnostic conditions, was integrated into the Seurat object. Cell names were matched between the metadata and Seurat object, and the metadata was added accordingly. Using neighbourhood graph construction and Louvain algorithm with a resolution parameter of 0.5, clustering was performed to identify the cell clusters. UMAP was used to visualize the clusters in a two-dimensional space. Cell identities were set according

to cell types defined in the metadata table. The expression of genes of interest was investigated across different cell types and conditions (e.g., Healthy vs. Crohn's disease: Inflamed and Non-inflamed condition). Differential expression of markers genes (e.g., ADM, CALCRL, RAMP2, RAMP3, ACKR3) was visualized using feature plot. Violin plots were generated to compare gene expression levels across these cell types and disease conditions. These gene expression in the violin plots is in the form of normalized expression values which applies to all the violin plots accordingly. Further, results were visualized using various other plots, including bar plots to show the distribution of cell count for each cell types under different conditions, stacking bar plot indicating the proportion of cell types in each condition and dot plots to represent the expression of marker genes across different cell types and conditions with the average intensity of expression level and percentage of gene expressed to present the data in a normalized form.

The MAST v1.28.0 package (Andrew McDavid [Aut, 2017]) was used to determine the differences in gene expression between non-inflammatory, inflammatory, and healthy conditions. DEGs were calculated for three comparisons: non-inflamed vs. healthy, inflamed vs. healthy, and non-inflamed vs. inflamed. Genes with adjusted p-values below 0.05 were considered significant and were subsequently mapped to Entrez IDs for pathway analysis. A gene filtering function was implemented to identify genes expressed in a minimum fraction of cells across each disease group. The function calculated the fraction of cells with non-zero expression for each gene and retained genes expressed in at least 10% of cells per group. Functional interpretation of DEGs was conducted using the clusterProfiler v4.10.1 package (Guangchuang Yu [Aut, 2017]) for KEGG pathway enrichment analysis, which assessed both upregulated and downregulated genes in each condition. Enrichment results were visualized using ggplot2 v3.5.1 package (Wickham, 2011), where pathways were ranked by significance, showing their regulation status and gene ratio. Further in-depth pathway exploration, focusing on immune response, signal transduction, cell adhesion and other pathways that are potentially relevant for adrenomedullin signalling was performed using the pathview v1.42.0 package (Weijun Luo, 2017).

The epithelial and immune datasets were analysed in the similar manner for the downstream analysis but due to memory limitation with the large size of these datasets, the approach to analyse the dataset was different. The count data matrix, barcode files, and feature annotation were read and stored in appropriate formats. We created a Seurat object for downstream analysis, ensuring that the counts matrix was appropriately labelled with unique gene identifiers and cell barcodes. The quality control was already performed on these datasets which were visualized using violin plots to proceed further. The metadata was read and cleaned by setting appropriate column names and converting necessary columns to numeric formats. Matching cells between the metadata and the Seurat object were identified, and metadata was added to the Seurat object. This included annotations for cell types, sample types, chemical treatments, and library preparation protocols. Pre-computed UMAP coordinates were loaded

and integrated into the Seurat object. These coordinates were used for dimensionality reduction, allowing for the visualization of cell populations in a two-dimensional space using the UMAP. The gene expression across different conditions and cell types were visualized similarly to the stromal dataset.

2.4 Pediatric dataset Analysis

The single-cell RNA sequencing data from pediatric samples were obtained in the form of an .h5ad file. This data was loaded into the Scanpy (Wolf et al., 2018) environment for upstream analysis. Initial quality control metrics were assessed to evaluate the distribution of gene counts, total counts per cell, and mitochondrial gene content. These metrics were visualized using violin plots.

Cells with fewer than 200 detected genes and cells with greater than 9000 detected genes were excluded to remove low-quality cells and potential multiplets. Doublet detection was performed to identify and exclude likely doublet cells based on doublet scores. The summary statistics reported that there was no doublet in the dataset. Cell cycle effects were assessed using predefined cell cycle gene markers. Gene lists specific to the G1, S, and G2M phases were extracted, and their presence in the dataset was evaluated. To account for cell cycle effects, highly variable genes were recomputed after excluding genes associated with the cell cycle. PCA was performed on the normalized and scaled data, reducing the data to 40 principal components. To account for batch effects, the Batch Balanced K-Nearest Neighbours (BBKNN) method was applied. Clustering was conducted using the Leiden algorithm at a resolution of 0.5, followed by visualization of the clusters using UMAP. The UMAP was annotated based on the identified clusters and existing annotations. The downstream analysis was similarly applied to this dataset after annotating the clusters and genes of interest expression across different conditions and cell types were visualized using the Seurat package in a similar way as to the human adult dataset.

Chapter 3: RESULTS

3.1 Quality metric control of intestinal datasets

The scRNA-seq dataset for terminal ileum stromal cells consists of 75,695 cells and 28,923 genes. The dataset was analyzed and violin plots for the quality control metric were generated across various cell identities. Three key metrics were majorly looked for the quality control i.e., number of detected features (n_Feature_RNA) (Figure S1a), the total RNA molecule per cell (nCount_RNA) (Figure S1b) and the percentage of mitochondrial RNA content (percent.mt) (Figure S1c).

The number of detected features, representing the number of genes per cell, ranged from approximately 0 to over 6,000 across different cell identities. Most cells had feature counts between 1,000 and 4,000, indicating a high degree of cellular complexity and appropriate sequencing depth. The total RNA count per cell varied from 0 to over 15,000, with the majority of cells clustering around 5,000 to 10,000 RNA counts. Most of the cells displayed mitochondrial percentages below 5%, with majority hovering around 1-2%. The overall low mitochondrial content ensures a high-quality dataset suitable for downstream analyses.

Colon stromal single cell dataset consists of 39,433 cells and 28,663 genes and similar approach of analysis was followed for this dataset as well. These QC metrics (Figure S2 a-c) demonstrated the robustness and quality of the dataset after processing, with most cells showing appropriate levels of gene expression, RNA counts, and low mitochondrial content. This ensured that the data is of sufficient quality for further exploration, including clustering and differential gene expression analysis. Similarly, quality metric for epithelial and immune datasets were visualized using violin plots to proceed further for the downstream analysis (data not shown).

3.2 High Variable Genes

In the next step high variable genes across cells were identified which was essential to prioritize genes for downstream analysis. These genes are observed in the scatter plot (Figure 9) representing the

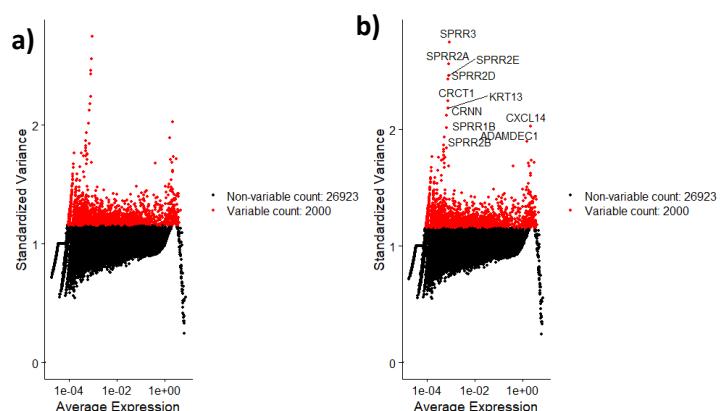
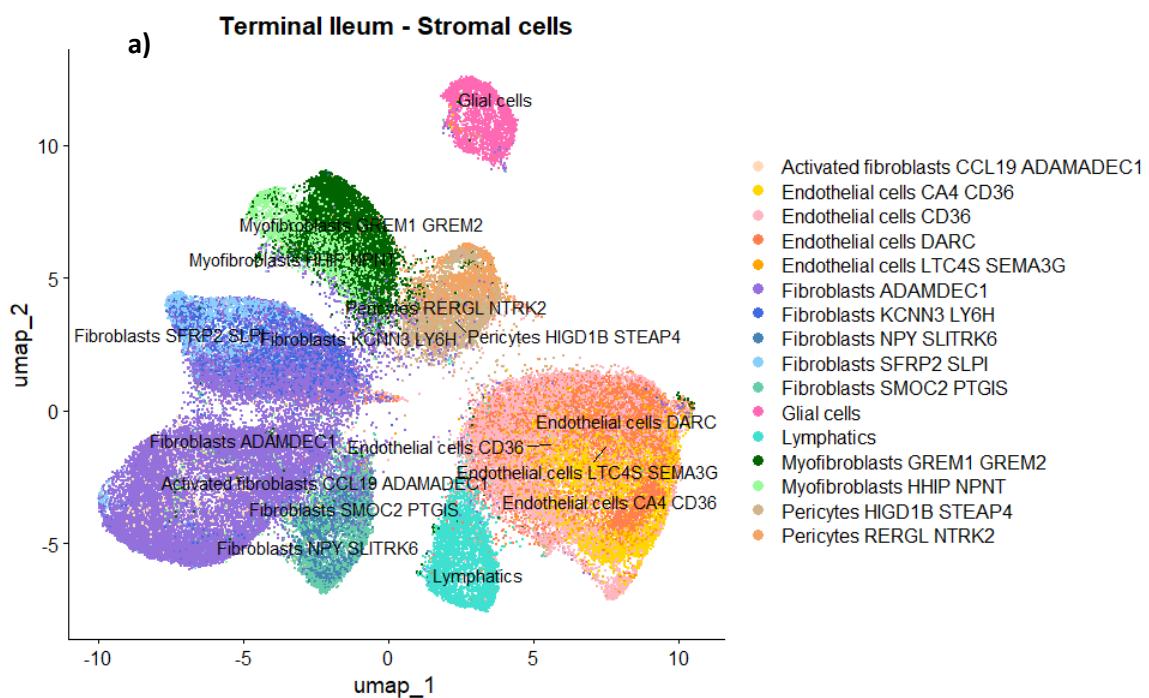


Figure 9. Illustrates the scatter plot depicting the relationship between the standardized variance and the average expression levels for all genes for terminal ileum stromal cells. a) variable genes highlighted in red and non-variable highlighted in black. b) highlighting the top genes in among the high variable genes in the stromal cell population

relationship between standardized variance and average expression for all genes. The genes highlighted as highly variable (in red on the plot) exhibit higher standardized variance for a given average expression level, making them critical candidates for further analysis to understand their role in tissue homeostasis, immune responses, or inflammatory processes in the terminal ileum. Genes such as SPRR3, SPRR1B, and ADAMDEC1 were among the most variable genes, indicating their potential significance in the regulation of stromal cell function and inflammatory responses. A total of 2000 genes were identified as highly variable genes which represents the most informative gene in the dataset.

3.3 UMAP Visualisation of Stromal Cell Population in Terminal Ileum & Colon

UMAP was employed to visualize the cell population clusters in the stromal cells covering both the site i.e., terminal ileum and colon from their individual dataset. As shown in Figure 10, distinct cell populations were identified within each tissue type, highlighting the complex cellular environments in these regions. Cells from both sites formed well-separated clusters, indicating clear distinction in gene expression of the clusters. Stromal cells clustered into various endothelial and fibroblast subsets, pericytes and glial cells with unique marker expression in the terminal ileum (Figure 10a). For colon (Figure 10b), similar cell type clusters were identified using marker genes. However, stromal cycling cells and inflammatory fibroblast were identified as unique clusters in the colon dataset suggesting different biology in the region. Likewise endothelial cells CA4CD36 were identified as a unique cluster in the terminal ileum dataset



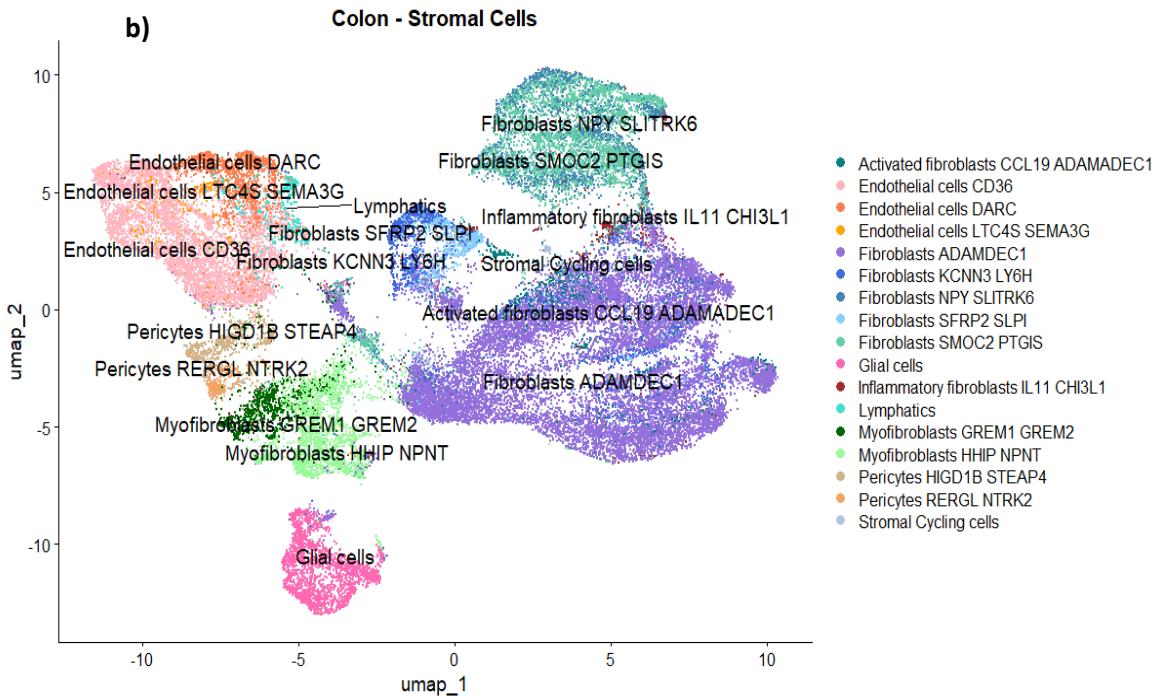


Figure 10. **Stromal Cell clusters.** a) Uniform Manifold Approximation Plot (UMAP) displaying the annotated clusters into different cell types for Terminal ileum dataset. b) UMAP displaying the annotated clusters into different cell types for colon dataset.

3.4 Gene Expression Analysis of Stromal Cell Population

After clustering, the expression of our marker genes of interest (ADM, CALCRL, RAMP2, RAMP3, ACKR3) were investigated among the cell subsets in the terminal ileum stromal cell population. It was found that adrenomedullin gene (ADM) is widely expressed across multiple cell clusters as shown in UMAP feature plot (Figure 11a), predominantly in fibroblast subtypes and some in endothelial cell population. The adrenomedullin gene receptor CALCRL (Figure 11b) and co-receptors, RAMP2 (Figure 11c) and RAMP3 (Figure 11d) which are essential in ADM signalling has notable expression mainly in endothelial cells followed by some in fibroblasts, except for RAMP3 which is mainly associated to endothelial cells. However, the expression of atypical chemokine receptor 3 (ACKR3) (Figure 10e) which acts as a decoy receptor for ADM is predominantly expressed in the fibroblast population. Colon stromal dataset showed the similar gene expression results from the feature scatter plot (Figure S3 a-e). The immune and epithelial datasets were also analyzed using the similar approach (see sections 3.5 and 3.6).

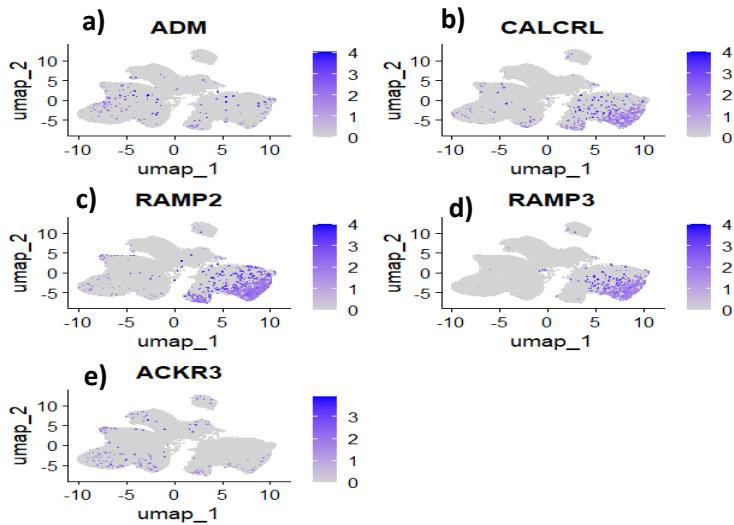


Figure 11. Feature plot showing the expression of adrenomedullin and its receptor marker genes among the terminal ileum stromal cell population. Subplots represent expression levels of (a) ADM, (b) CALCRL, (c) RAMP2, (d) RAMP3, and (e) ACKR3. Each subplot shows UMAP projections with colour intensity reflecting the expression level of each gene, with darker shades indicating higher expression.

Now to perform in-depth expression analysis, violin plots were generated to analyze the distribution of expression of the genes of interest among different cell types regardless of disease state. ADM expressing cells (Figure 12a) were observed in all the fibroblast cell subpopulation, particularly higher expression was seen in the fibroblast ADAMDEC1 with maximum cells expressing towards higher normalized expression values between 2.5 - ~4 followed by uniform expression in fibroblast SFRP2 SLPI. The endothelial cell subset also showed some basal level of ADM expression while glial and pericytes displayed minimal to no expression.

The highest number of cells expressing CALCRL, RAMP2, and RAMP3 is observed in the endothelial cell subsets. CALCRL (Figure 12b) and RAMP2 (Figure S4a) showed maximum number of cells expressing these genes in all the endothelial cells including arterial, venous and lymphatic followed by few cells expressing in fibroblast ADAMDEC1, while RAMP3 (Figure S4b) displayed minimal to no expression in lymphatic endothelial cells. It was observed that cells expressing ACKR3 (Figure S4c) were mainly confined to fibroblast subpopulations, with highest number of cells expressing in fibroblast ADAMDEC1 which goes in line with ADM expression. Similar trend was observed for the expression of ADM and its receptors in colon stromal cells. However, in the colon, minimal to no cell expressing CALCRL and RAMP3 was observed in the lymphatic endothelial cells (Figure S5a-e). Epithelial and Immune datasets were analyzed in a similar manner (see section 3.5 & 3.6).

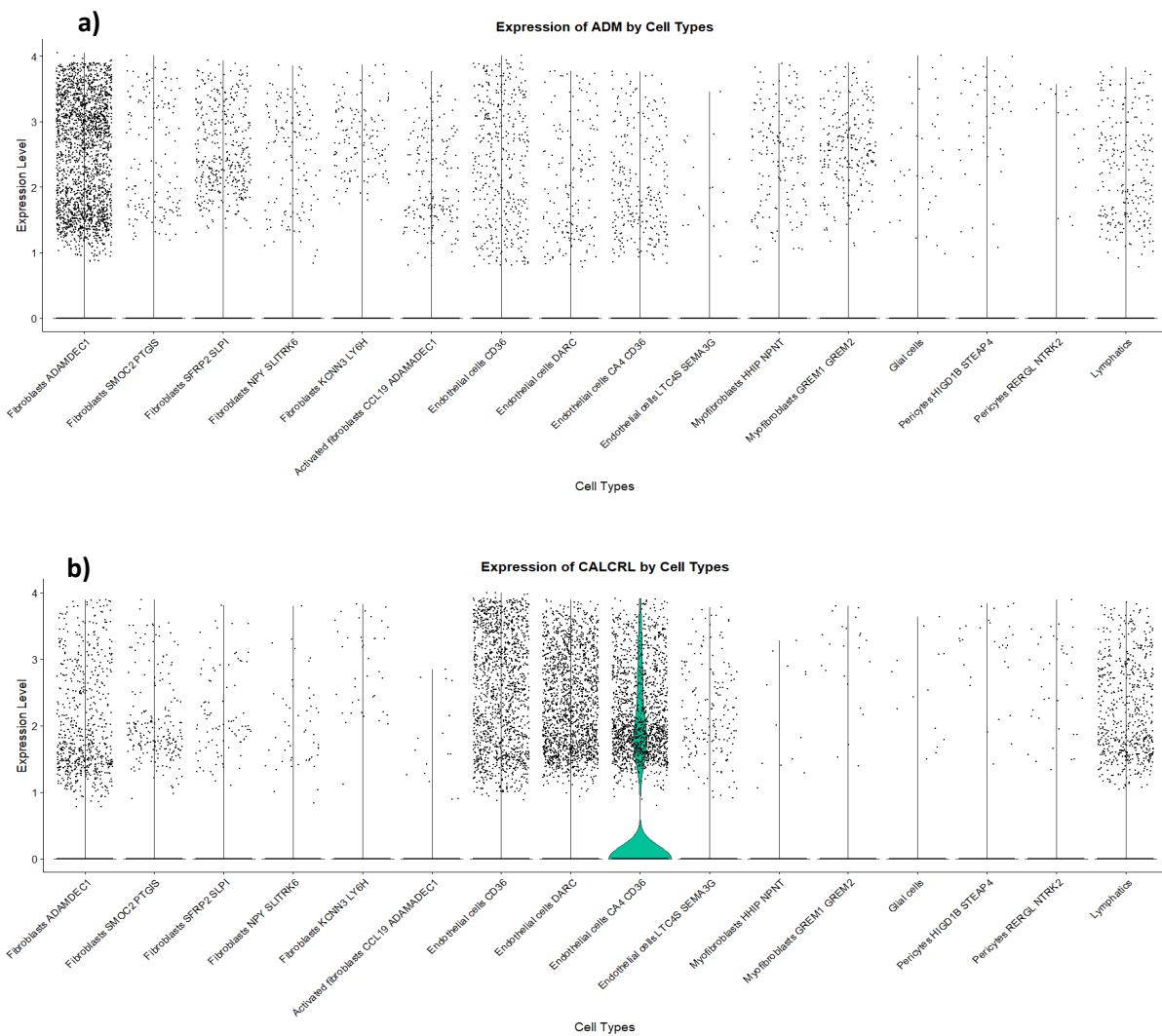


Figure 12. Terminal ileum stromal cells. Violin plots showing the expression marker genes across various cell types. a) ADM
b) CALCRL

Comparative analysis to identify the expression level of ADM and its associated receptors (CALCRL, RAMP2, RAMP3, ACKR3) across three conditions: Healthy (Heal), Non-Inflamed (NonI) and Inflamed (Inf) with relevance to Crohn's disease was performed. These conditions will help to evaluate the potential role of ADM signalling to regulate disease progression.

As shown in the violin plots, ADM and its receptors tend to have higher number of cells expressing this gene in the non-inflamed (NonI) condition followed by inflamed (Inf) and healthy (Heal) conditions. For ADM (Figure. 13a), its expression in the inflamed condition appeared to be in two subpopulations with one showing a higher expression level and the other a lower expression. CALCRL, RAMP2 and RAMP3 (Figure. 13b-d) exhibited similar and dense distribution of cells across all levels of expressions in both the disease conditions while healthy showed the moderate expression level. For ACKR3 (Figure. 13e), both the inflamed and noninflamed condition had maximum number of cells depicting low

expression level while the healthy condition showed to have moderate expression level for the majority of cells.

In the colon dataset, the inflamed condition had the least number of cells exhibiting expression for adrenomedullin gene and its receptors. In contrast, the healthy and noninflamed condition presented nearly similar number of cells expressing ADM (Figure S6a), RAMP2 (Figure S6c) and RAMP3 (Figure S6d), while CALCRL (Figure S6b) and ACKR3 (Figure S6e) exhibited the highest number of cells expressing these genes in noninflamed condition. Similar approach was applied to the immune and epithelial dataset in terminal ileum and colon (see section 3.5 & 3.6).

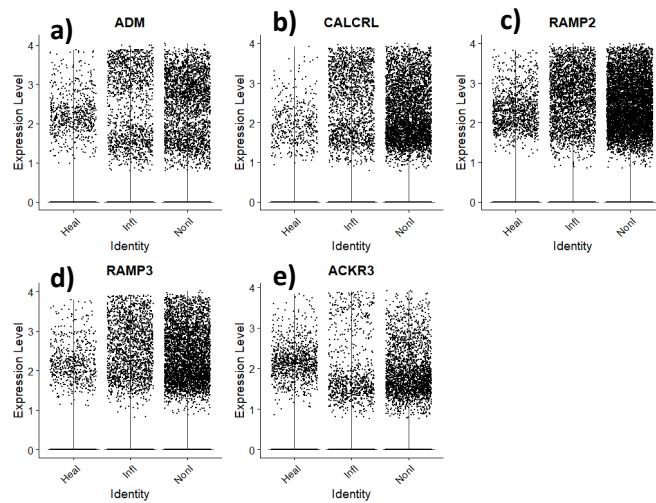


Figure 13. Violin plot representing gene expression across different conditions (Healthy, Inflamed, Non-inflamed) in Terminal Ileum stromal cells. a) ADM b) CALCRL c) RAMP2 d) RAMP3 e) ACKR3

To further explore the expression patterns of adrenomedullin gene (ADM) and its key receptors (CALCRL, RAMP2, RAMP3, ACKR3), violin plots were generated to assess the distribution of these markers across the different stromal cell types when comparing healthy, non-inflamed, and inflamed conditions.

In the terminal ileum dataset, ADM (Figure 14a) again showed broad distribution of cells expressing in fibroblast ADAMDEC1 in all the conditions followed by fibroblast SFRP2 SLPI primarily in the healthy condition as shown by the wide violin plot. It is clearly observed that inflamed condition has two distinct subpopulations with one showing low expression and the other exhibiting higher level of expression. Endothelial cells subset showed some basal uniform to low distribution of cells expressing ADM with maximum expression in non-inflamed condition and minimal in healthy. The narrow distribution highlighted a smaller subset of endothelial cells contributing to ADM expression. Overall, there was increasing trend of cells expressing ADM in non-inflamed condition followed by inflamed and healthy condition as described earlier.

Endothelial cells showed the most prominent and widespread expression of CALCRL (Figure 14b) and RAMP2, RAMP3 genes (Figure S7a-b) across almost all three conditions in the terminal ileum stromal cell population. This was evident from the wide and tall violin plots across these cell types. CALCRL expression was prominent in endothelial cell subsets, with the highest expression in non-inflamed and inflamed conditions. The healthy condition (yellow) exhibited no expression in endothelial CD36 for CALCRL and in lymphatics endothelial for both CALCRL and RAMP3. In RAMP2 and RAMP3 healthy condition, endothelial cells showed a broad distribution, with a significant portion of the cells showing moderate to high expression, which peaks around 3 to 6 normalized expressions values. In the non-inflamed condition (cyan), the distribution was also wide but slightly more concentrated at higher expression levels, which might indicate that receptors expression is elevated in these cells. For the inflamed condition, the expression in endothelial cells remained strong, but the violin plot was slightly narrower than in healthy and non-inflamed conditions. This suggested that although receptors are still expressed during inflammation, its distribution might be more tightly regulated. Myofibroblasts, pericytes, and glial cells showed little to no expression of receptors across all conditions, emphasizing that receptors are highly specific to arterial, venous and lymphatic endothelial cells. ACKR3 expression (Figure S7c) was highly specific to fibroblast populations across the different conditions. The fibroblast ADAMDEC1 cell type displayed the highest levels of ACKR3 expressing cells among all fibroblast subtypes with being dominant in non-inflamed condition. This fibroblast SFRP2 SLPI showed a moderate level of ACKR3 expression in healthy condition, though the number of cells expressing high levels is few compared to ADAMDEC1 fibroblasts. Overall, the numbers of cells expressing all the receptors were higher in non-inflamed followed by inflamed and healthy for most cell types.

In the colon dataset, ADM showed distribution of cells mainly in fibroblast subtypes (Figure S8a) however it was highly distributed here compared to terminal ileum. Fibroblast ADAMDEC1 dominated the population here similar to terminal ileum but with nearly similar number of cells expressing uniformly in both healthy and non-inflamed condition. The inflamed condition for ADM exhibited the varied expression level in the wide spread of violin plot however the number of cells expressing in this condition is less. Overall, for most fibroblasts cell subtypes, healthy and non-inflamed condition showed tall and narrow violin plots with cells showing varied and high expression level.

Colonic endothelial cells such as CD36 and DARC endothelial cells exhibited the highest number of cells expressing CALCRL (Figure S8b) across all conditions similar to terminal ileum but here mature arterial endothelial cell (LCT4S SEMA3G) also showed varied expression level of cells in inflamed disease condition. RAMP2 and RAMP3 (Figure S8c, S8d) displayed the similar expression to terminal ileum dataset. ACKR3 (Figure S8e) also showed the similar dominance in number of cells expressing in fibroblast ADAMDEC1 however all other stromal subtypes displayed more varied expression level in inflamed condition.

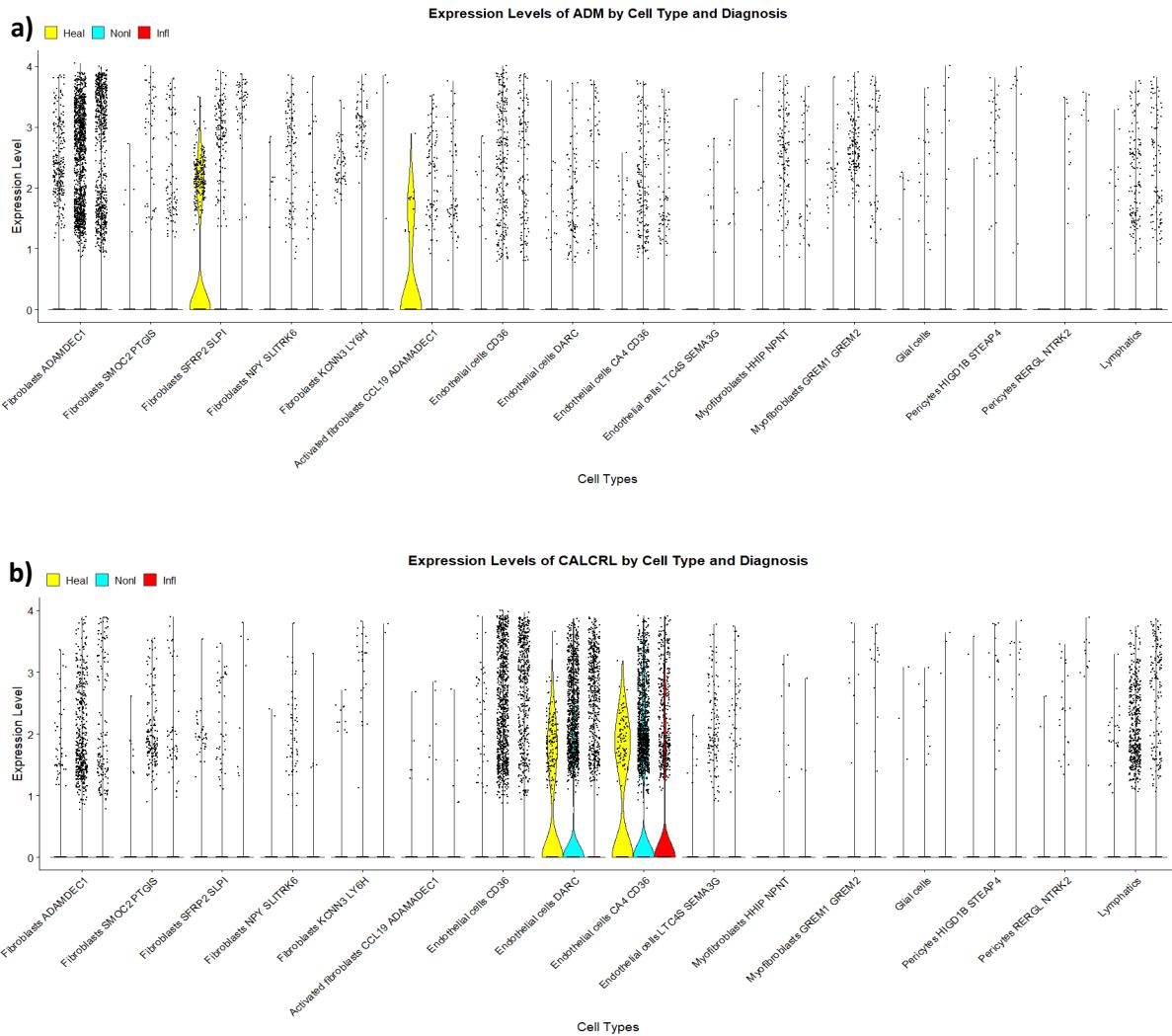


Figure 14. Terminal ileum stromal cells. Violin plots showing the expression levels of marker across various cell types and conditions (Healthy -yellow, Non-inflamed -cyan, Inflamed – red). a) ADM b) CALCR.

To account for the higher number of cells expressing the marker genes in the non-inflamed condition, the total number of cells across all cell types in each condition was visualized using bar plot. As shown in the Figure 15, fibroblast ADAMDEC1 accounted for the highest number of cells in both the colon and terminal ileum datasets. Notably, a higher number of cells count was observed in the non-inflamed condition in both the datasets.

Specifically, in the terminal ileum (Figure 15a), the non-inflamed condition displayed over 13,000 fibroblast ADAMDEC1 cells, which is more than five times higher than in the healthy condition and approximately twice as many as in the inflamed state. Across most stromal cell types whether pericytes, myofibroblasts, glial cells, or endothelial cells the overall pattern remained consistent: non-inflamed condition displayed the most cells followed by inflamed, with healthy tissue showing the smallest cell count.

Similar to the terminal ileum dataset, the non-inflamed condition (green) in the colon dataset (Figure 15b) shows the highest number of cells across all stromal cell types. Fibroblast ADAMDEC1 again dominates the highest cell count in the dataset with nearly 9,000 non-inflamed cells. However, here the healthy condition (red) has the second highest cell counts in the stromal cell population exceeding the inflamed condition for all cell types which is due to larger number of samples in the healthy condition.

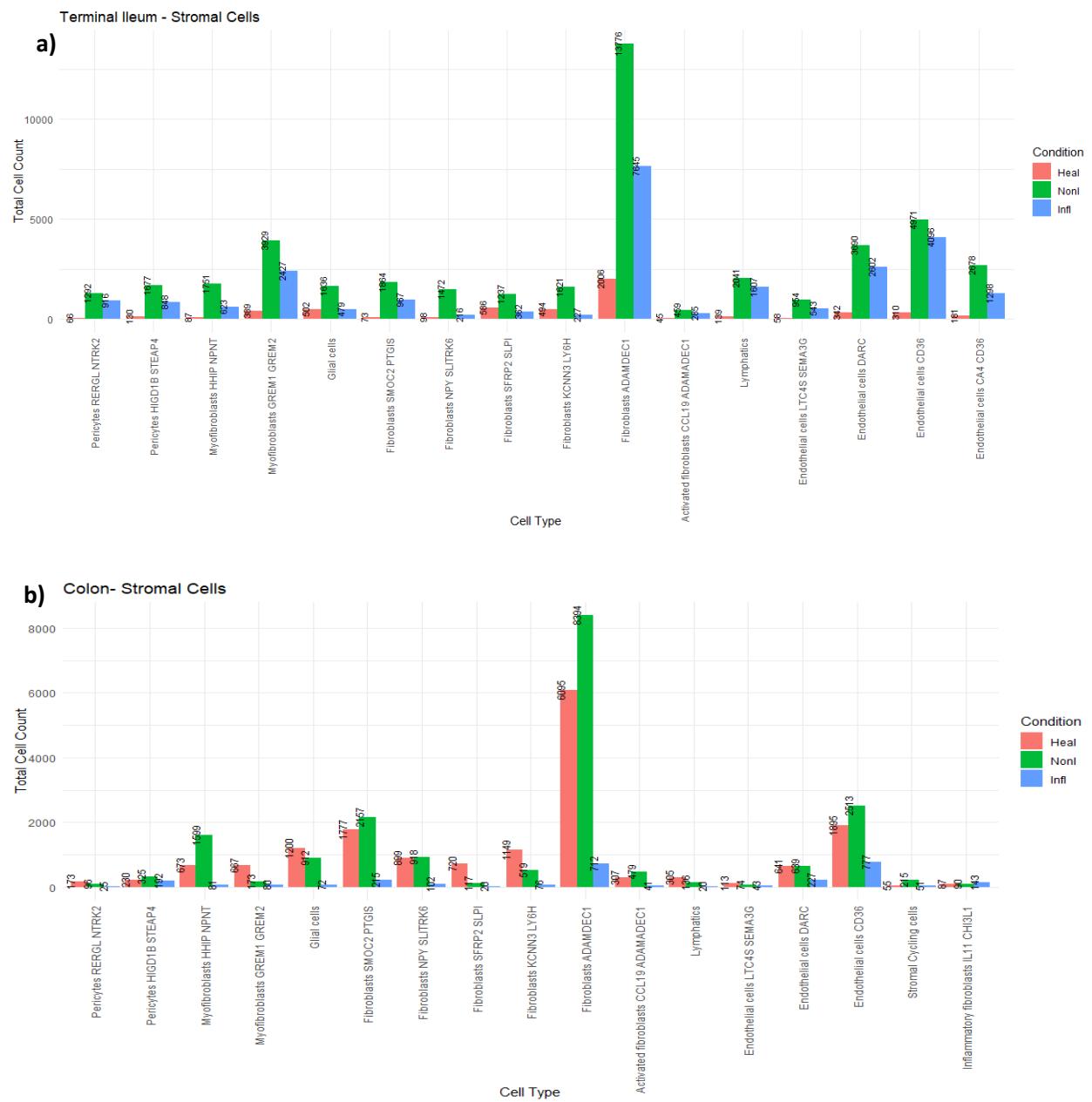


Figure 15. Stromal Cell Count Distribution Bar plot representing the cell count across different conditions (health-red, non-inflamed-green, inflamed-blue) a) in the terminal ileum. b) in the colon.

Because of the difference in cell counts between the different disease states i.e. the numbers were generally higher for the non-inflamed conditions and some specific cell types, the data was normalized

as percentage in the dot plot to ensure accurate representation of gene expression patterns. The dot plot visualized the percentage of cells expressing each gene (dot size) and the intensity of average expression scaled (dot colour). Alongside, a proportional plot was included to observe these changes in cell composition and account for the differences in cell numbers within the dataset and, ensuring that cell count variation is not influencing the observed expression patterns.

The gene expression profiles of ADM, CALCRL, RAMP2, RAMP3, and ACKR3 were consistent between the terminal ileum and the colon with adrenomedullin gene being highly expressed in fibroblast subtypes and receptors in endothelial subtypes (Figure 16 and 17).

ADM showed highest expression in fibroblast SFRP2 SLPI and activated fibroblast in the dot plot (Figure 16b) although from proportional plot fibroblast ADAMDEC1 constitutes the highest proportion of the terminal ileum dataset (Figure 16a) which is why normalization was indeed required. The healthy condition in these subtypes exhibited the highest expression followed by inflamed and non-inflamed but the intensity of expression increased slightly from healthy to inflamed. The endothelial cell populations showed nearly a similar basal expression in all conditions. However, the major expression of ADM is mainly in fibroblast cell population subtypes.

A decreasing trend (Figure 16b) in the proportion of cells expressing adrenomedullin key receptors was observed across conditions, particularly as we moved from healthy to non-inflammatory and then inflammatory condition. Despite the reduction in the proportion of cells expressing these receptors, the intensity of expression remained largely unchanged. This indicated that although fewer cells expressed the genes in inflammatory conditions, the average expression level remained consistent. CALCRL and RAMP2 showed expression in all venous, arterial and lymphatic endothelial but RAMP3 had prominent expression seen in arterial and venous endothelial cells. ACKR3 displayed highest percentage of expression in the healthy condition for most fibroblast subtypes with a decreasing trend but with the exception of fibroblast NPY SLTRK6 which showed an increase of percentage expression in inflamed condition from non-inflamed as well as maximum intensity of expression is in this condition. The proportional plot showed stepwise increase in the endothelial population while fibroblasts exhibited a stepwise decrease for most of its subtypes.

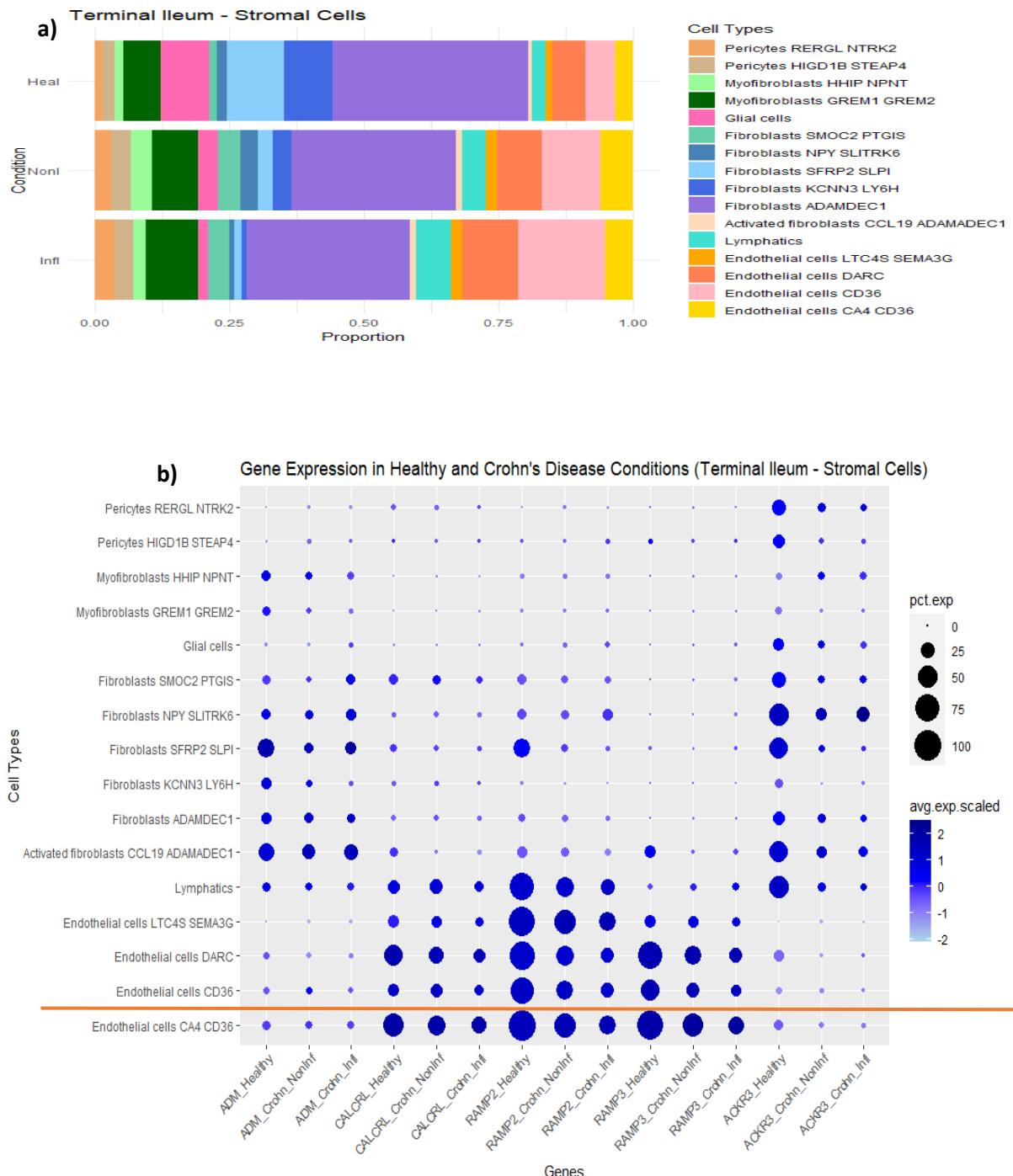
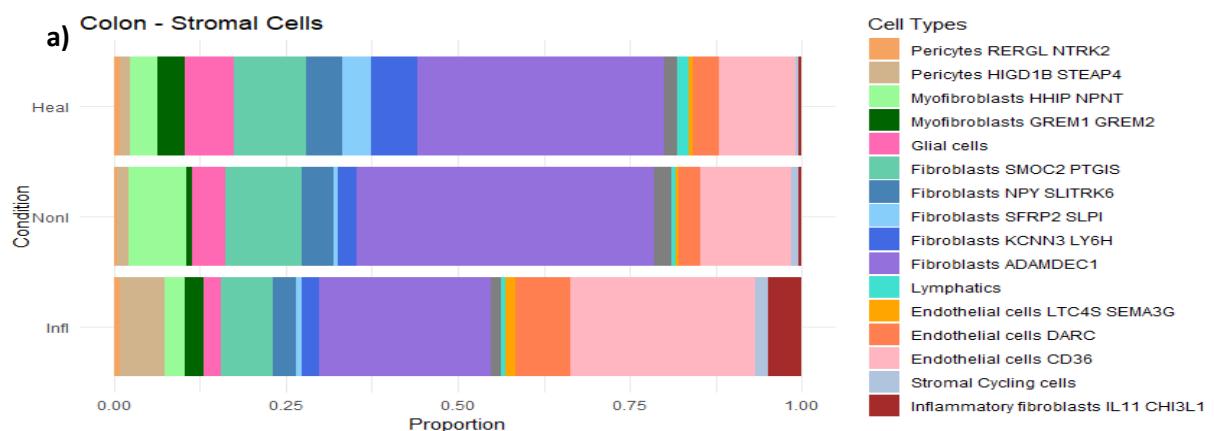


Figure 16. Terminal ileum stromal cells a) Proportional plot showing the distribution cell types across healthy, non-inflamed, and inflamed conditions. b) Dot plot showing the expression levels of ADM, CALCRL, RAMP2, RAMP3, ACKR3 across cell types under healthy, non-inflamed, and inflamed conditions. Dot size represents the percentage of cells expressing the gene, while the colour intensity indicates the average scaled expression level. The orange line separates the unique cell types present in the terminal ileum while the cell populations above the line are the shared populations with the colon.

In the colon dataset, it was seen that an opposite increasing trend was followed with ADM, CALCRL and ACKR3 (Figure 17b) being highly expressed in the inflamed condition compared with terminal ileum followed by healthy and non-inflamed condition. However, it was observed that ADM expression is more prominent in fibroblast from colon samples similar to terminal ileum dataset. CALCRL and RAMP2 was highly expressed in all endothelial subtypes while RAMP3 showed prominent expression

only in the venous endothelial cells (DARC, CD36). RAMP2 showed similar expression level and percentage in all the three conditions with less variability in CD36 and LTC4S SEMA3G. RAMP3 exhibited increased percentage of expression in inflamed condition followed by healthy and non-inflamed and uniform expression intensity in DARC endothelial cells while CD36 showed both uniform percentage and intensity here. ACKR3 goes in line with the ADM expression with maximum expression in inflamed condition of fibroblast subtypes and highest expression intensity in healthy stromal cycling cells and pericytes non-inflamed condition. The proportional plot is visualized using stack bar plot and it indicated change in cell type composition based on disease conditions. Strongest difference observed between healthy and inflamed condition where pericytes and endothelial cells are more abundant, while many decrease is in numbers of fibroblast ADAMDEC1 in inflamed samples (Figure 17a).



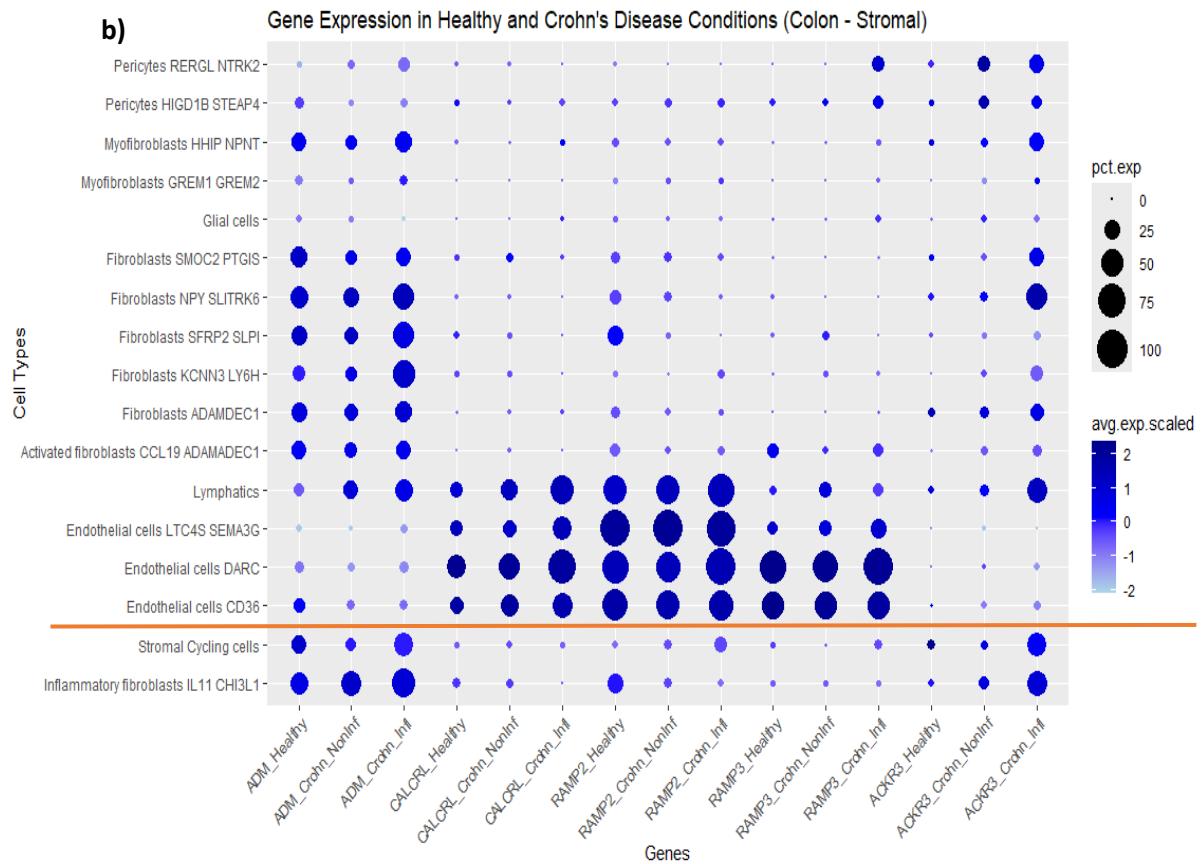


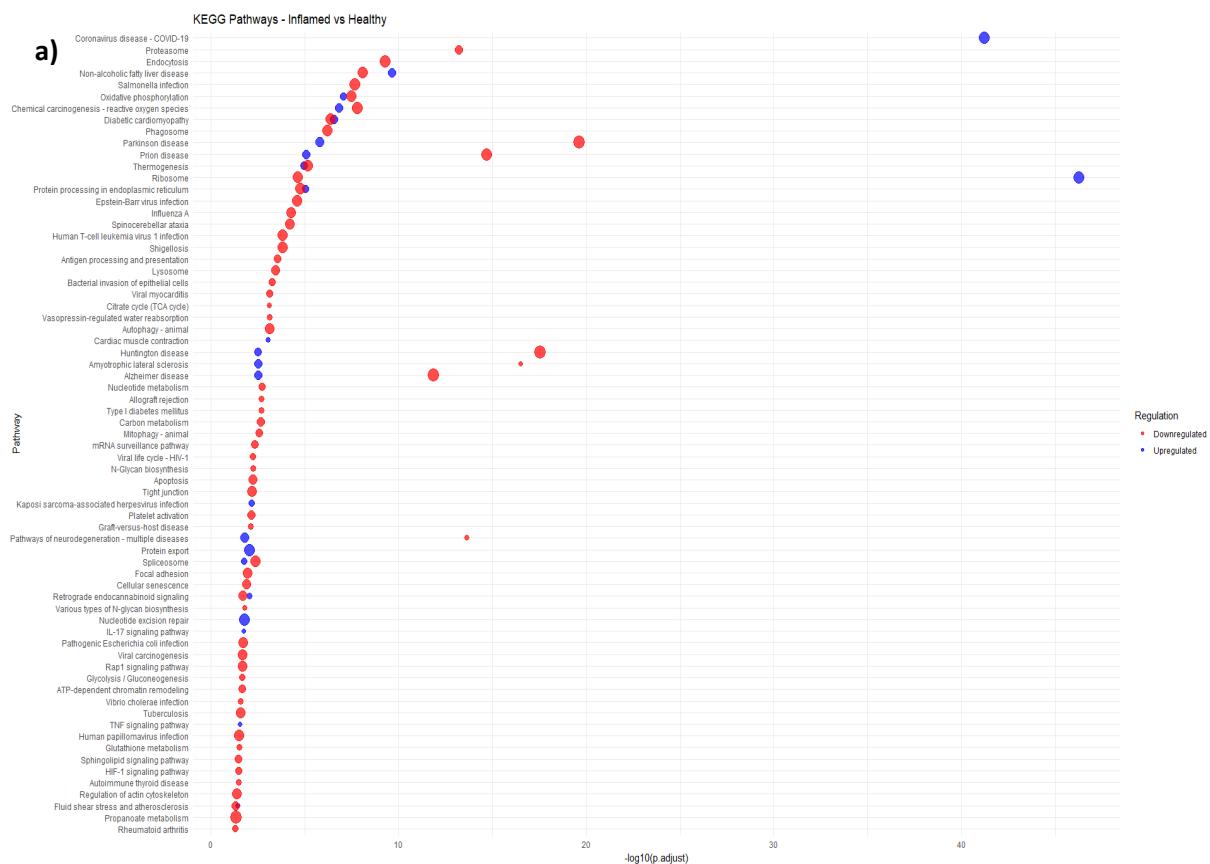
Figure 17. Colon stromal cells a) Proportional plot showing the distribution cell types across healthy, non-inflamed, and inflamed conditions. b) Dot plot showing the expression levels of marker genes across cell types under healthy, non-inflamed, and inflamed conditions. Dot size represents the percentage of cells expressing the gene, while the colour intensity indicates the average scaled expression level. The orange line separates the unique cell type present in the colon while the cell populations above the line are the shared populations with the terminal ileum.

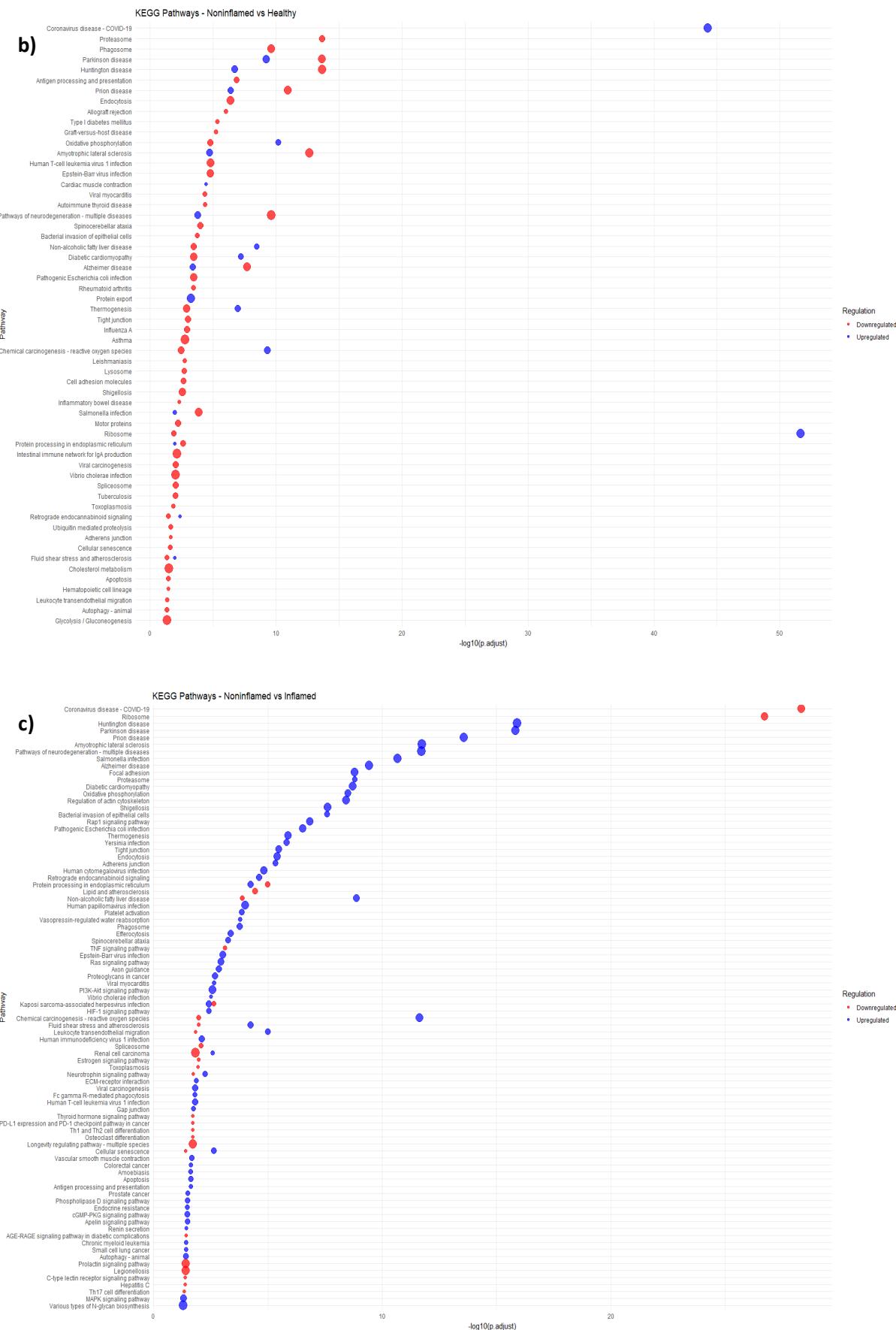
Next, differentially expressed genes (DEGs) were identified within the terminal ileum stromal cell populations, focusing on non-inflamed vs. healthy, inflamed vs. healthy, and non-inflamed vs. inflamed comparisons. The main focus was on endothelial and fibroblast subtypes, particularly because of their higher levels of ADM expression and its key receptors.

The dot plots visually represented KEGG pathway enrichment across these conditions (Figure 18 a-c), where the size of the dots indicates the number of genes involved in the pathways, and the colour distinguishes between upregulated (blue) and downregulated (red) pathways. Larger and more significant pathways are shown by a higher p-value.

This analysis revealed changes in expression of genes of multiple pathways including several pathways that we consider interesting in contexts of adrenomedullin signalling based on current knowledge in the literature. Some of these included: Focal Adhesion, Cell Adhesion Molecules, Adherens Junctions, HIF-1 Signalling, Rap1-Signaling, Ras Signalling, PI3K-Akt Signalling, MAPK Signalling, and Vascular Smooth Muscle Contraction (Brain and Grant, 2004; García-Ponce et al., 2016).

Among these, the Vascular Smooth Muscle Contraction pathway (Figure 18d) emerged as a major point of interest, being the only signalling pathway present in the KEGG database (Brain and Grant, 2004) that was previously described for ADM. From the KEGG analysis, this pathway was found to be upregulated when comparing the non-inflamed vs inflamed condition (Figure 18c). Analyzing the Vascular Smooth Muscle Contraction pathway revealed upregulation of key genes, such as CRLR, Gs, AC and PKA, involved in initiating vasodilation, a process influenced by ADM.





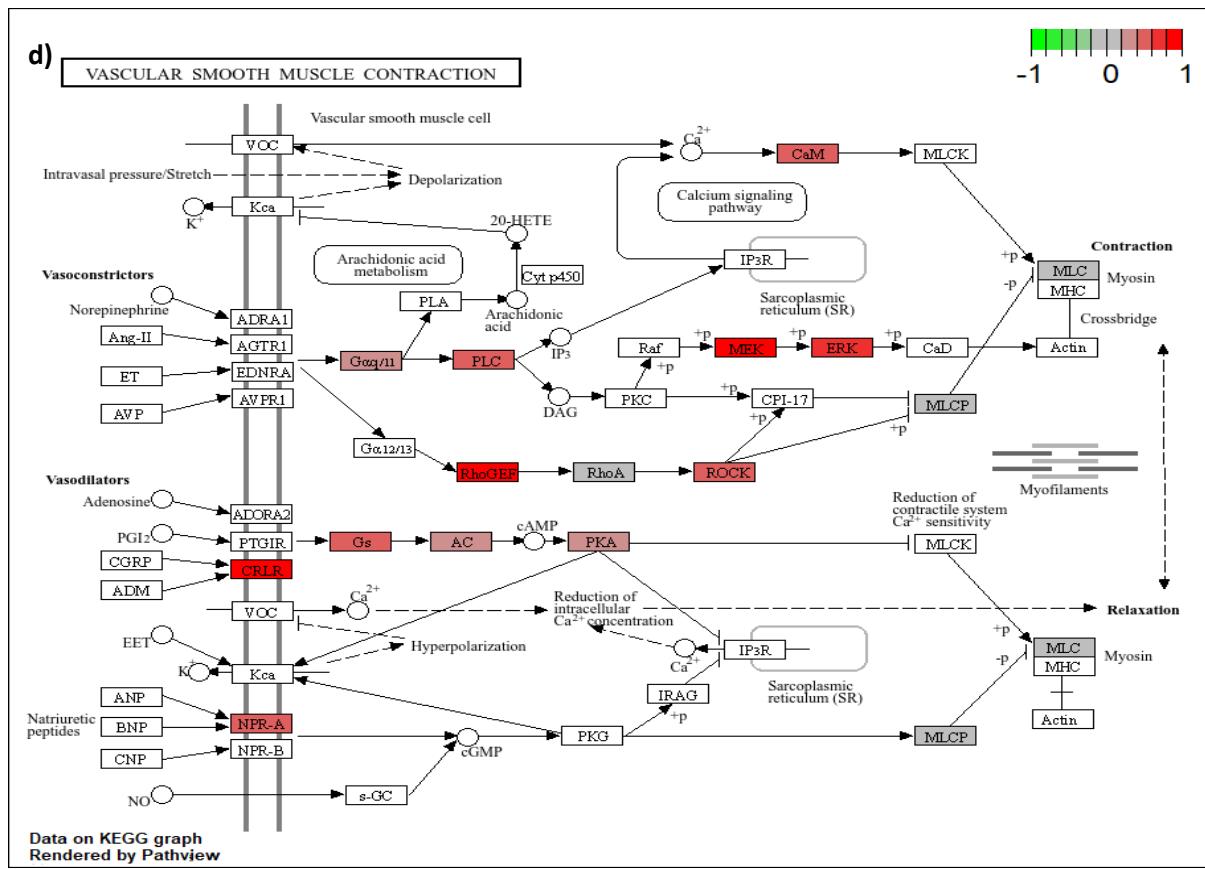


Figure 18. KEGG Pathway Analysis of Endothelial CA4 CD36 Cells from Terminal Ileum Stromal Cells. (a) KEGG pathways enriched in the inflamed vs. healthy condition. The dot size represents the number of genes involved in each pathway, and the colour indicates upregulation (blue) or downregulation (red). (b) KEGG pathways enriched in the non-inflamed vs. healthy condition. (c) KEGG pathways enriched in the non-inflamed vs. inflamed condition. (d) Vascular smooth muscle contraction pathway highlighting upregulated genes in the non-inflamed condition.

3.5 Immune Cell Population

A similar bioinformatic analysis was performed in the adult immune cell population. The immune cell dataset comprised of total 201,072 cells and 28,923 genes for the terminal ileum and 152,509 cells and 28,663 genes for the colon. UMAP was used to visualize the cell population clusters in the immune cells across the datasets. As shown in UMAP, distinct cell populations were identified, highlighting the complex immune landscape present in the dataset. The cells formed well-separated clusters of major immune cell populations including T cell subsets, B cells, monocytes, macrophages, plasma cells, and dendritic cells (DC). These cell subsets were annotated based on the expression of their canonical marker genes from original paper for this human adult dataset.

Both the terminal ileum and colon datasets revealed similar immune cell types, but some differences were observed. For example, the terminal ileum (Figure 19a) included some unique populations such as Cycling cells, IELs ID3 ENTPD1, Monocytes HBB, neutrophils and some specific macrophages subtype. In contrast, the colon dataset (Figure 19b) contained the unique populations of Immune cycling cells, Macrophages metallothionein and NK-like cells AD3 ENTPD1. These distinctions point to slight variances in the immune response and profile between the two tissue sites.

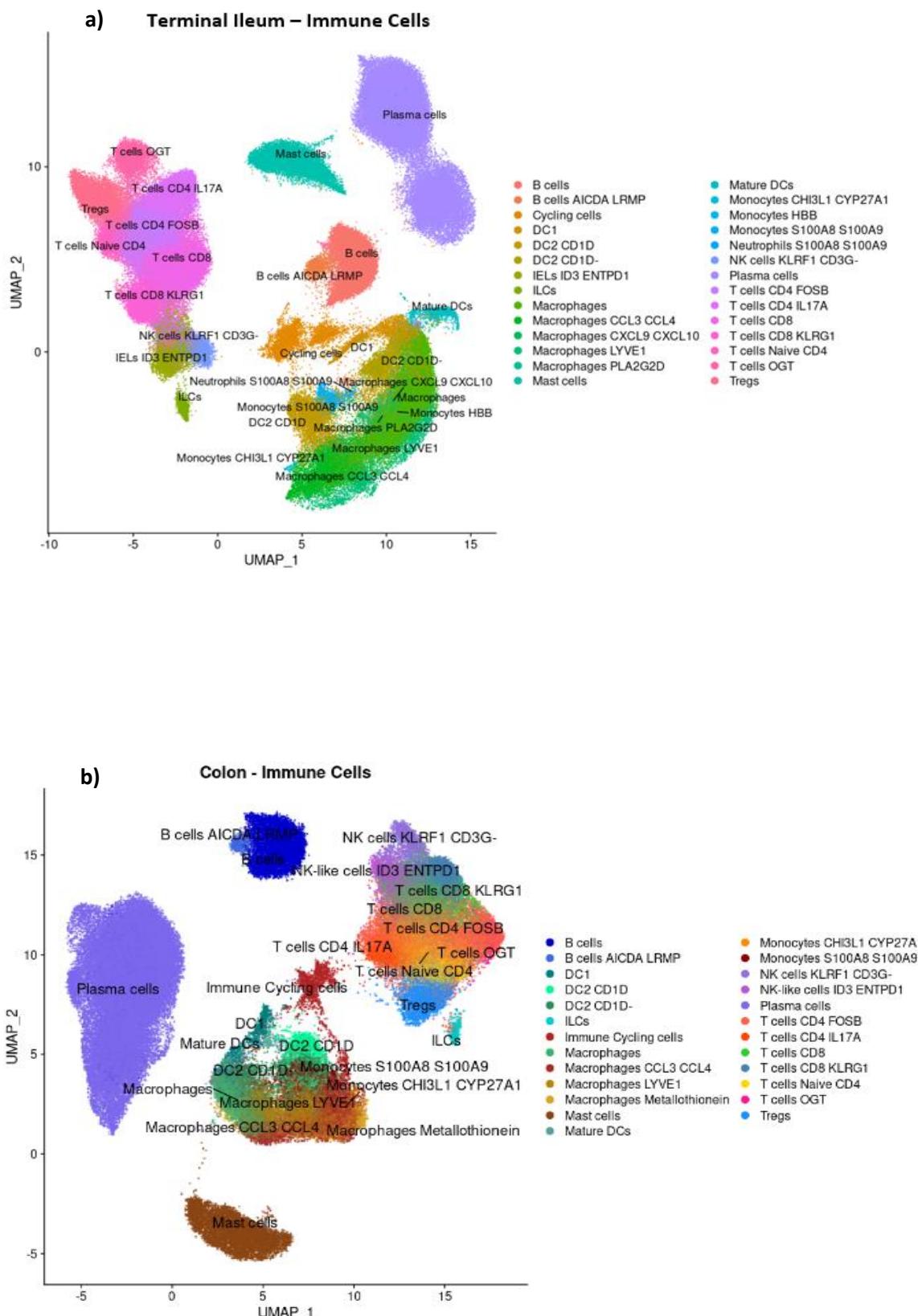


Figure 19. **Immune Cell clusters.** a) UMAP displaying the annotated clusters into different cell types for Terminal ileum dataset. b) UMAP displaying the annotated clusters into different cell types for colon dataset.

The expression pattern in the terminal ileum was visualized using violin plots. ADM (Figure S9a) displayed prominent expression in plasma cells, macrophages CCL3/CCL4, DC2 CD1D and monocytes with comparatively higher number of cells expressing in the non-inflamed condition. In the colon dataset ADM was similarly expressed (Figure S10a) in plasma cells, macrophages CCL3 and DC2 CD1D but with higher number of cells in both healthy and non-inflamed condition. Notably, macrophages exhibited increased cell expressing ADM in the non-inflamed condition while monocytes showed higher cell expressing in the inflamed condition.

CALCRL (Figure S9b) displayed higher number of cells expressing in plasma cells, dendritic cells, macrophages CCL3 in the non-inflamed condition of the terminal ileum. In the colon dataset (Figure S10b) CALCRL expressing cells were in plasma cells, with nearly similar cell expression in the healthy and non-inflamed condition while dendritic cells showed increased cell expression in the non-inflamed condition.

RAMP2 (Figure S9c) showed elevated number of cell expressing in the non-inflamed condition for plasma cells, macrophages, macrophages CCL3/CCL4 and dendritic DC2 CD1D in the terminal ileum dataset. In the colon dataset, RAMP2 displayed high number of cells expressing (Figure S10c) in plasma cells particularly in both the healthy and non-inflamed condition.

RAMP3 (Figure S10d) exhibited limited number of cells expressing in plasma cells in the colon dataset, with most cell expressing in the healthy condition, while almost no expression of RAMP3 was observed in immune cells of the terminal ileum (Figure S9d).

ACKR3 (Figure S9e) was expressed in few cell types including T-cell subtypes, plasma cells, B cells and dendritic cells, with higher cells expressing in the non-inflamed condition of the terminal ileum. The colon dataset also showed some similar expression of ACKR3 marker gene to terminal ileum (Figure S10e).

To accurately assess the biological significance of gene expression patterns, it was essential to account for the total cell counts across different conditions. This ensured that observed differences in gene expression are not skewed by variations in cell population sizes but reflect true biological effects.

Cell counts was estimated using the bar plot (Figure 20a) and the analysis for the terminal ileum dataset revealed that while healthy and inflamed condition showed only slight difference in cell number, the non-inflamed condition displayed a significant increase. Among all the cell types, plasma cells had the highest cell count with 28,818 cell counts followed by T-cells and macrophages.

In the colon dataset (Figure 20b), healthy and non-inflamed condition shared similar high cell count, whereas the inflamed condition exhibited the lowest cell count across all the conditions. Plasma cells

again accounted for highest cell counts, with 28,381 cells in non-inflamed condition and 26,139 cells in healthy condition.

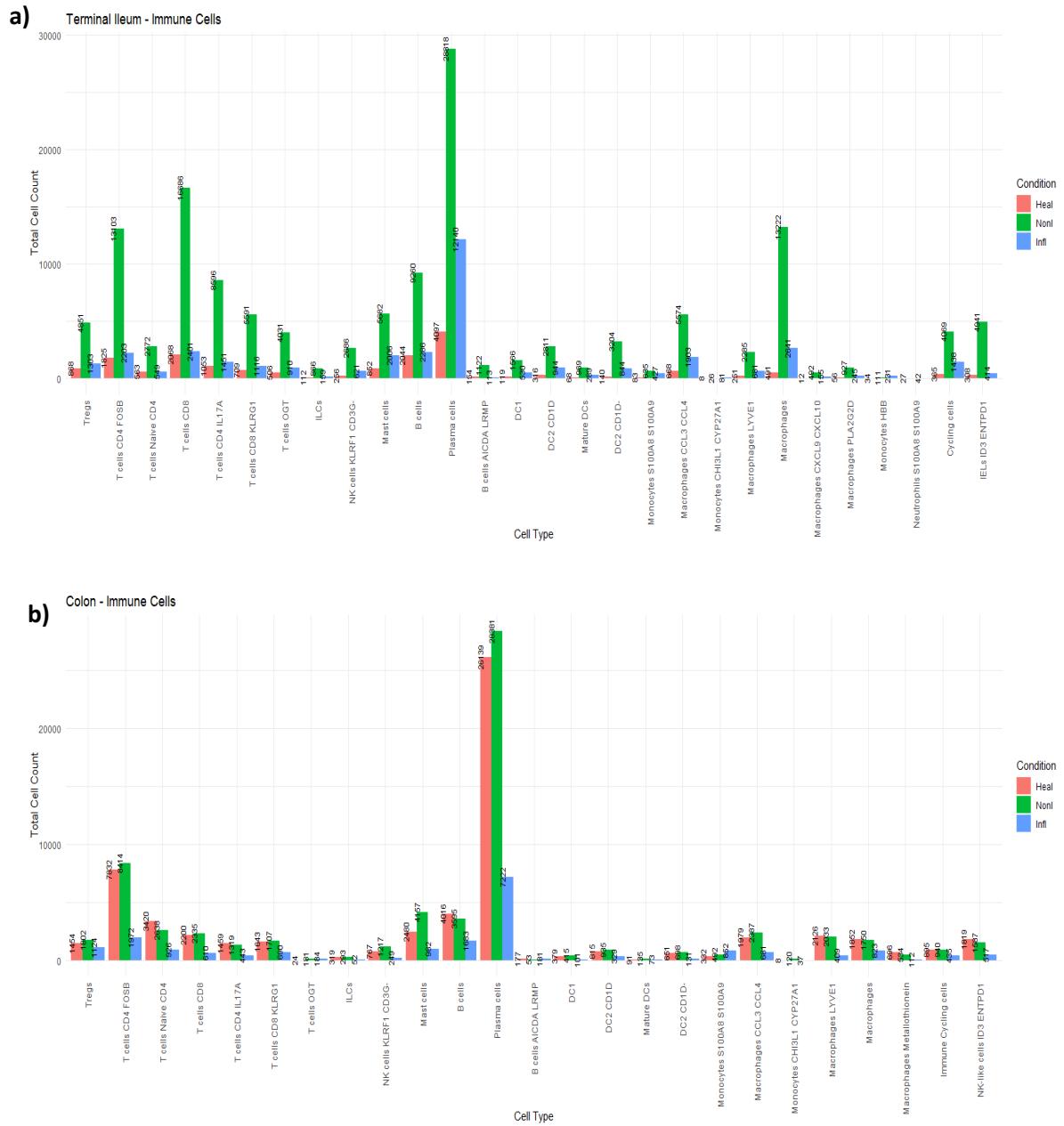
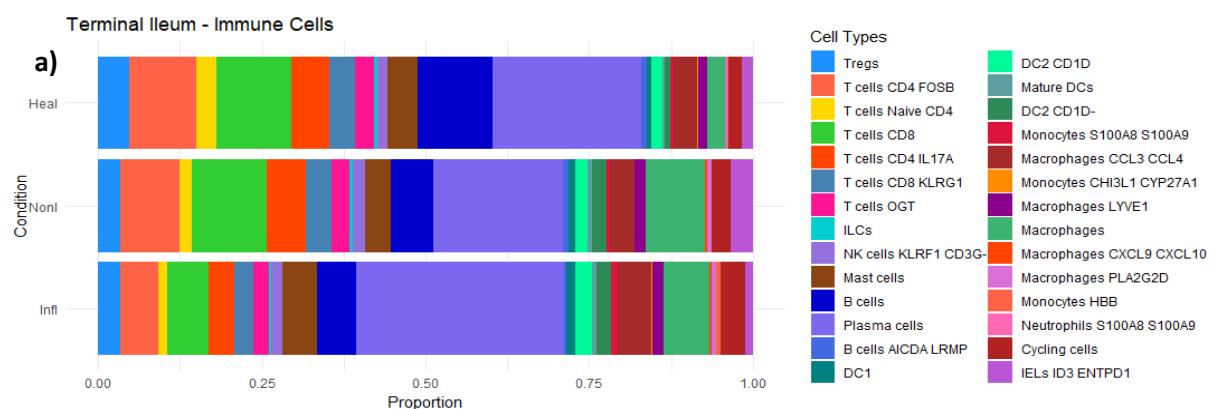


Figure 20. Immune Cell Count Distribution. Bar plot representing the cell count across different conditions (healthy-red, non-inflamed-green, inflamed-blue) a) in the terminal ileum. b) in the colon.

Given the high number of cells observed in the non-inflamed condition for the terminal ileum and a comparatively lower cell count in the inflamed condition for the colon, the data was normalized by percentage to ensure accurate representation of gene expression patterns. Alongside, a proportional plot (Figure 21a, 22a) was included to account for the differences in cell abundance within the dataset, ensuring that cell count variation is not influencing the observed expression patterns.

In both the terminal ileum and colon, ADM expression was predominantly found in macrophages and monocytes. The terminal ileum (Figure 21b) showed maximum expression in the monocytes CH3L1 CYP27A1 with similar percentages expression across conditions but with maximum intensity observed in the healthy condition. Monocytes S100A8 S100A9 exhibited higher intensity of ADM expression in the non-inflamed condition while macrophages CCL3 CCL4 and neutrophils showed increased ADM expression intensity in the inflamed condition. In the colon dataset (Figure 22b), ADM displayed an increasing trend in expression percentage across macrophages and monocytes subtypes with an exception of macrophages metallothionein which showed maximum expression in the non-inflamed condition in both percentage and intensity. ADM was limited expressed in plasma cells and dendritic cells where its cell percentage showed a decreasing trend in plasma cells and an increasing trend in dendritic cells.

CALCRL displayed a decreasing trend in percentage expression in dendritic cell subtypes in the terminal ileum while colon exhibited increasing trend, with the highest intensity observed in the mature DCs. RAMP2 was mainly expressed in macrophages and monocytes and plasma cells in the terminal ileum, with a decreasing trend of expression in the macrophages and monocytes. However, the colon dataset showed RAMP2 expression exclusively in the plasma cells, with an inverse trend of intensity expression. RAMP3 exhibited no expression immune cell population of both the terminal ileum and colon. ACKR3 showed some expression in B cells in the terminal ileum dataset and almost no expression in the colon dataset. From the proportion plot (Figure 21a) in the terminal ileum dataset, population difference presented a decrease in the T cells subtypes, a stepwise increase in plasma cells and most of macrophage subtypes.



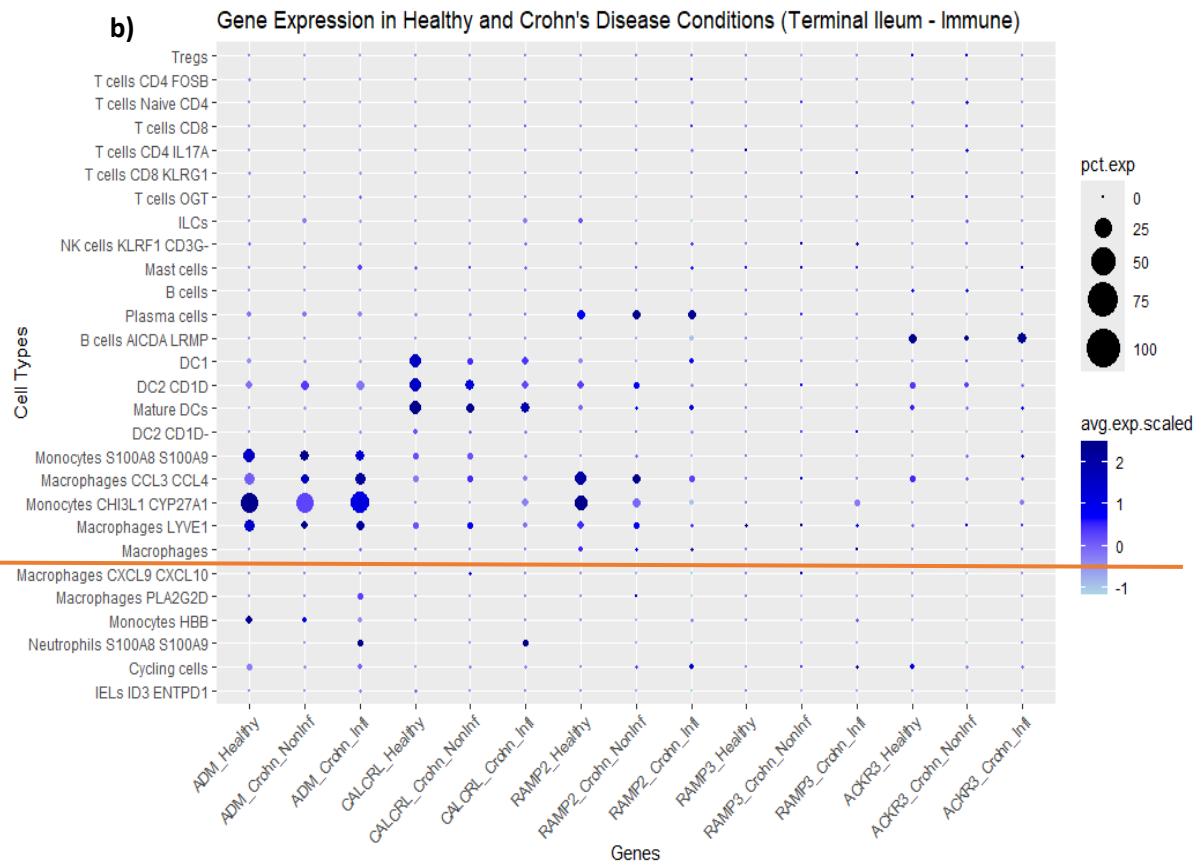
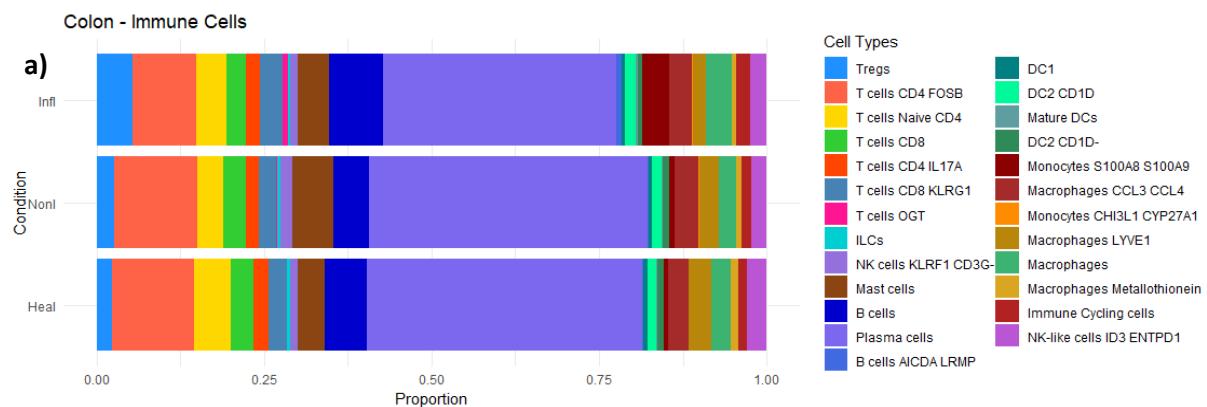


Figure 21. Immune Cell Composition and Gene Expression in Terminal Ileum. a) Proportional plot showing the distribution of cell types across healthy, non-inflamed, and inflamed conditions. b) Dot plot showing the expression levels of marker genes across immune cell types under healthy, non-inflamed, and inflamed conditions. Dot size represents the percentage of cells expressing the gene, while the colour intensity indicates the average scaled expression level. The orange line separates the unique cell population present in terminal ileum while the cell populations above the line are the shared populations with the colon.



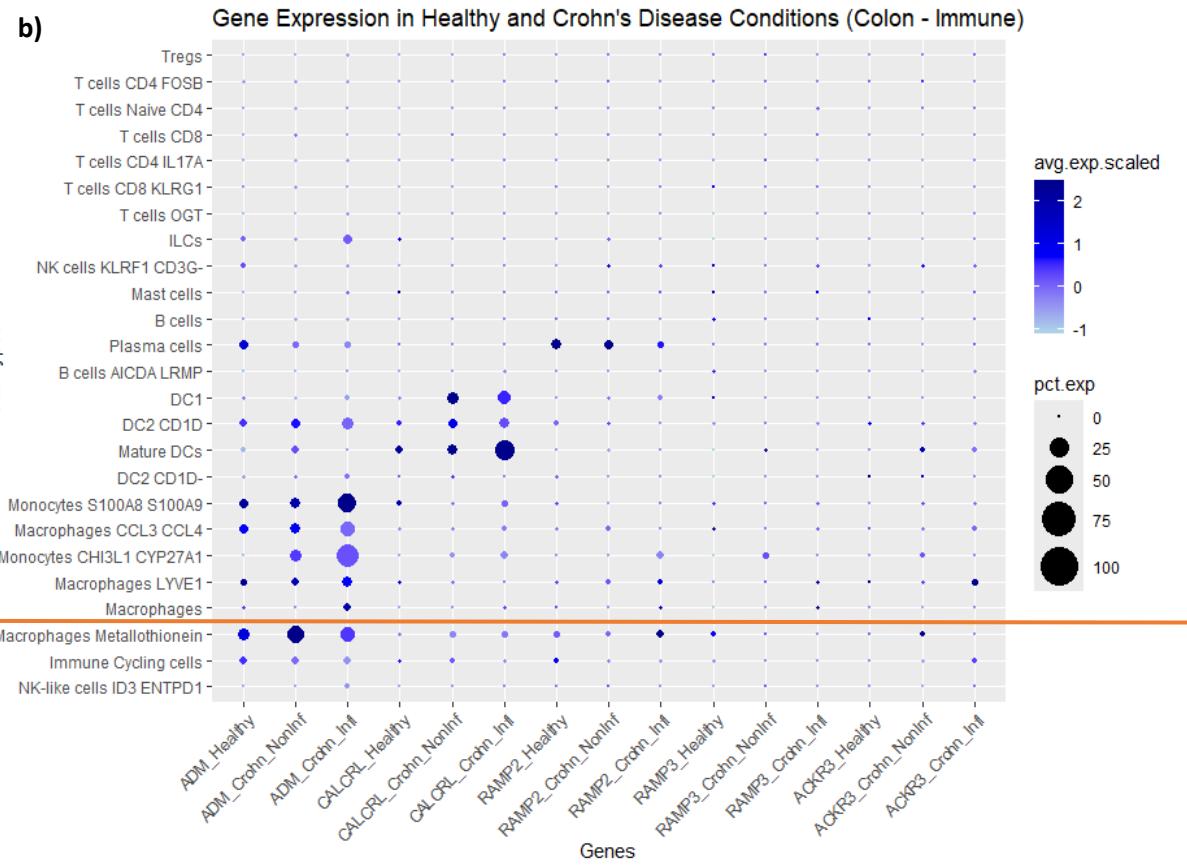


Figure 22. Immune Cell Composition and Gene Expression in Colon. a) Proportional plot showing the distribution of cell types across healthy, non-inflamed, and inflamed conditions b) Dot plot showing the expression levels of marker genes across immune cell types under healthy, non-inflamed, and inflamed conditions. Dot size represents the percentage of cells expressing the gene, while the colour intensity indicates the average scaled expression level. The orange line separates the unique cell population present in colon while the cell populations above the line are the shared populations with terminal ileum.

3.6 Epithelial cell Population

Finally, a bioinformatic analysis was performed in the adult epithelial cell population. The epithelial cell dataset comprised of a total of 154,136 cells and 28,923 genes for the terminal ileum and 97,788 cells and 28,663 genes for the colon. The epithelial that lines the gut is the first point of contact with an external environment. The epithelial dataset clustered into well-separated clusters with major cell population consisting of enterocytes, goblet cells, tuft cells, stem cells, and Paneth cells for both the datasets. These cell subsets were annotated based on expression of canonical marker genes from original paper for this human adult dataset.

Both the terminal ileum and colon revealed similar epithelial cell types, however, some distinctions between the datasets were observed. In the terminal ileum (Figure 23a), there was some unique populations such as stem cells OLM4 GSTA1, epithelial cells METTL12 MAFB, epithelial HBB HBA, enterocytes TMIGD1 MEP1A GSTA1, enterochromaffin cells and L cells. In contrast, the colon dataset (Figure 23b) revealed specific unique clusters of enterocytes CA1 CA2 CA4 and enteroendocrine cells. The difference in cell types between the site are mainly due to their distinct role in digestion and

absorption, as well as varying microbial environments they encounter (Kong et al., 2018). The terminal ileum more focuses on nutrient absorption, while colon is responsible for water absorption and immune defence (Kiela and Ghishan, 2016).

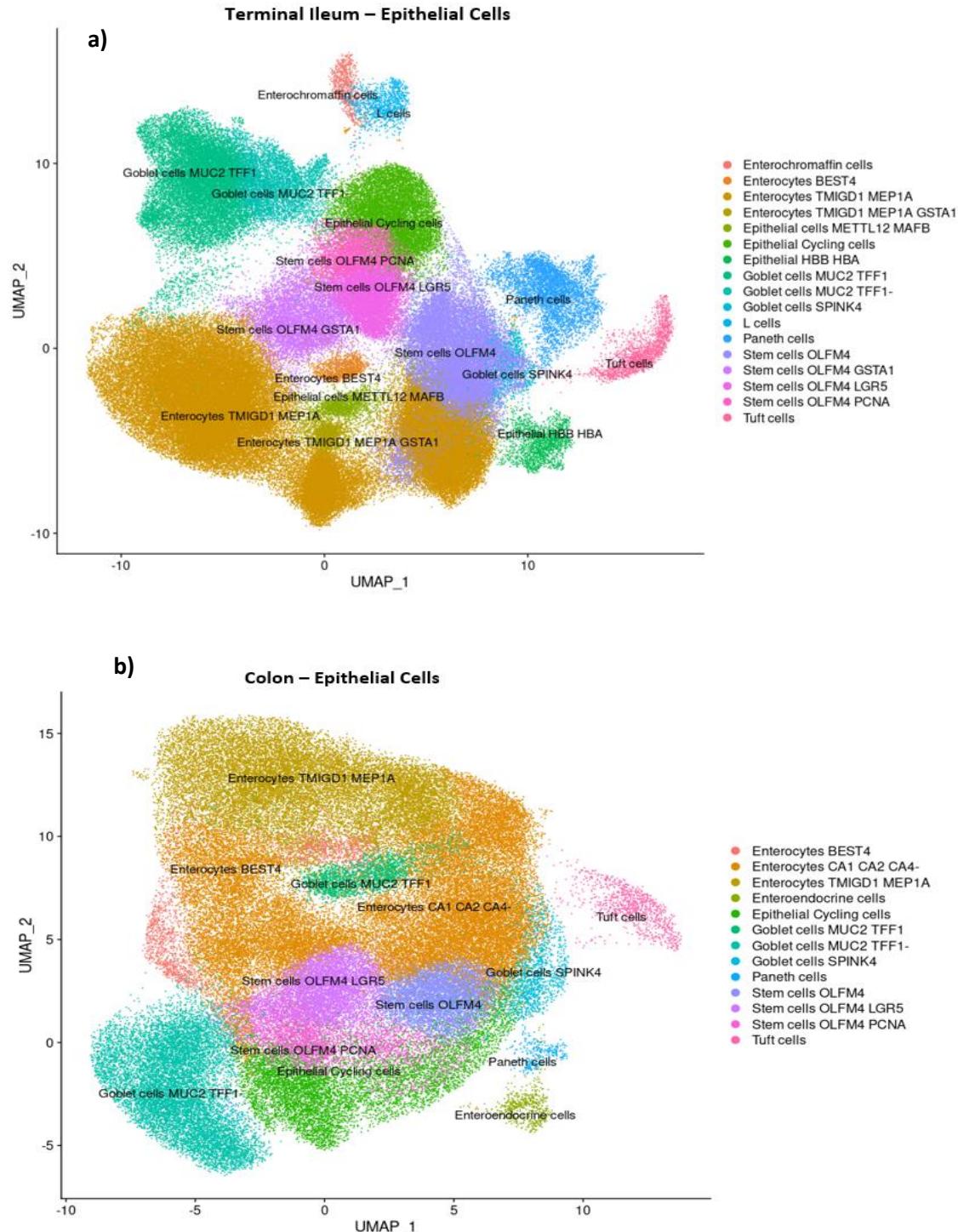


Figure 23. **Epithelial Cell clusters.** a) UMAP displaying the annotated clusters into different cell types for Terminal ileum dataset. b) UMAP displaying the annotated clusters into different cell types for colon dataset.

The expression of marker genes was visualized using the violin plots for both the terminal ileum and colon datasets. Among these marker genes, ADM showed more prominent expressed in colon dataset due to higher degree of specialization in immune response. In the terminal ileum, ADM (Figure S11a) was majorly expressed in stem cell subtypes, epithelial cycling cells, enterocytes (TMIGD1 MEP1A) and goblet cell subtypes with a higher number of cells expressing in non-inflamed condition. In contrast, in the colon dataset, ADM (Figure S12a) exhibited expression mainly in goblet cell subtypes, where non-inflamed condition presented the highest number of expressing cells. Enterocytes subtypes displayed similar number of cells expressing ADM across healthy and non-inflamed followed by inflamed condition. Epithelial cycling cells and stem cell subtypes showed higher number of cells expressing the gene in the healthy condition in the colon. Interestingly, Paneth cells exhibited varied gene expression in the healthy condition.

CALCRL (Figure S11b, S12b) exhibited little number of cells expressing the gene in epithelial cells for the both the dataset. RAMP2 displayed some cells expressing which were similar in the terminal ileum (Figure S11c) and colon dataset in stem cells subtypes, goblet cell subtypes and enterocytes (Figure S12c). RAMP3 exhibited little to no cell expressing the gene in both datasets (Figure S11d, S12d). ACKR3 (Figure S11e) showed major cells expressing in the enterocytes (TMGD1 MEP1A), particularly in the non-inflamed condition for the terminal ileum dataset. However, in the colon dataset, ACKR3 showed no significant cell expression (Figure S12e).

Cell counts was estimated using the bar plot (Figure 24a) and the analysis revealed that the non-inflamed condition displayed a significant increase in the numbers for the terminal ileum dataset. Among all the cell types, enterocytes TMIGD1 MEP1A accounted for the highest cell count with 42,810 cells in the non-inflamed condition followed by healthy and inflamed condition. The stem cells OLFM4 and goblet cells MUC2 TFF1 followed.

In the colon dataset (Figure 24b), healthy and non-inflamed condition shared nearly similar high cell count, whereas the inflamed condition contributed the lowest cells across all the conditions for most cell types. Enterocytes CA1 CA2 CA4 accounted for the highest cell count with nearly 14000 in both healthy and non-inflamed condition in this dataset followed by enterocytes and goblet subtypes.

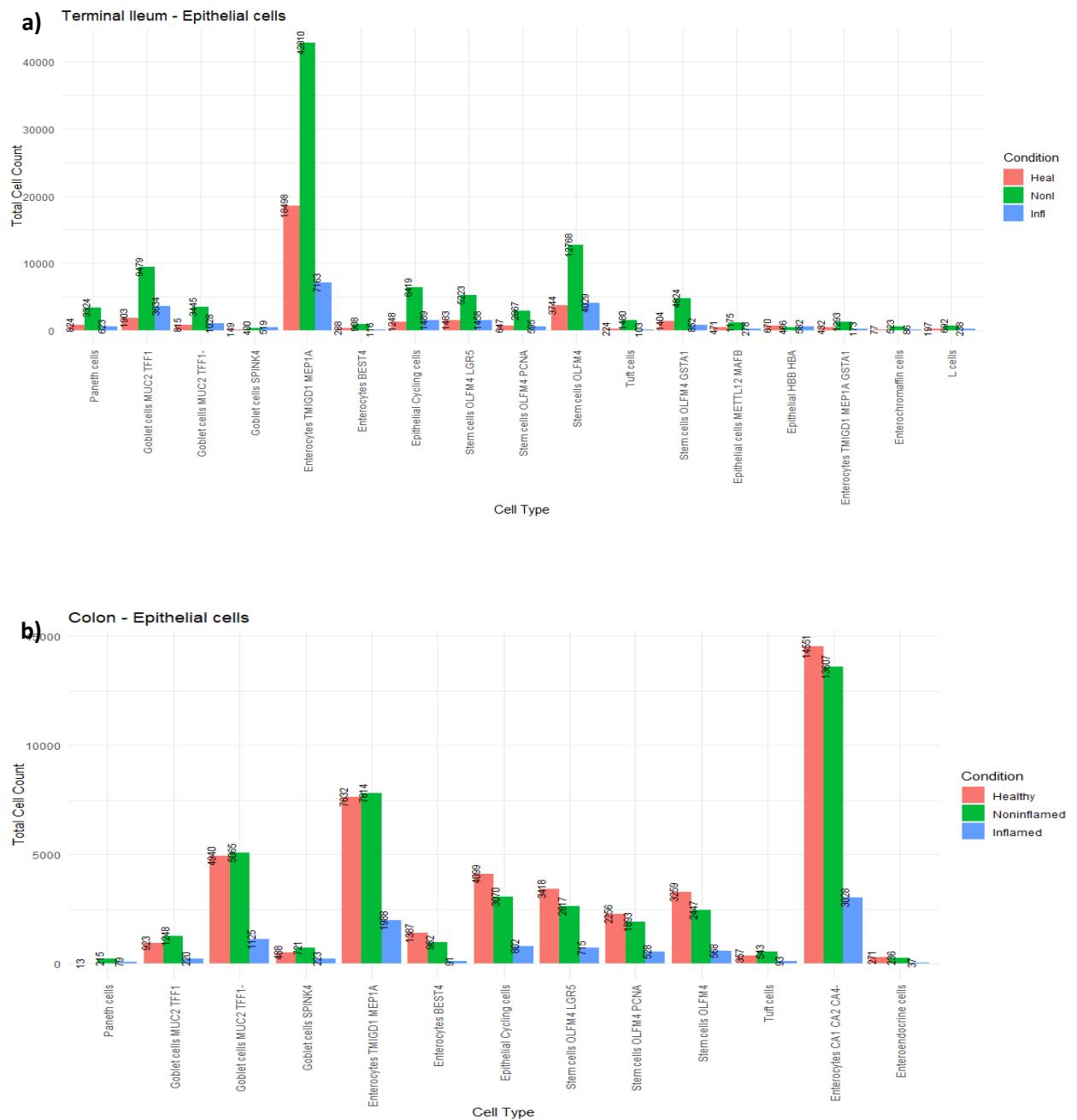


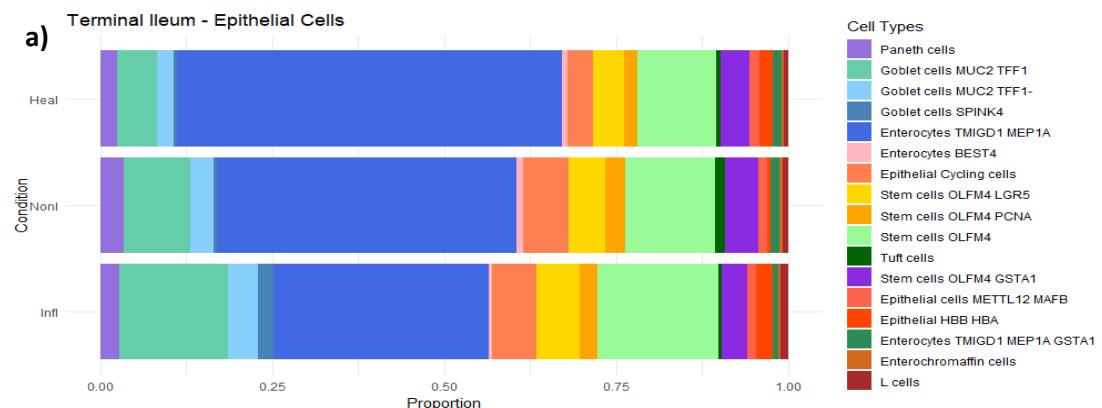
Figure 24. Epithelial Cell Count Distribution. Bar plot representing the cell count across different conditions (healthy-red, non-inflamed-green, inflamed-blue) a) in the terminal ileum. b) in the colon.

Given the high number of cells observed in one of the conditions, the data was normalized by percentage to ensure accurate representation of gene expression patterns. Alongside, a proportional plot of terminal ileum (Figure 25a, 26a) was included to account for the differences in cell numbers within the dataset, ensuring that cell count variation is not influencing the observed expression patterns.

In the dot plot (Figure 25b), ADM showed most prominent expression in goblet cells MUC TFF1 with highest percentage and maximum intensity in the non-inflamed condition for terminal ileum. Epithelial

cells METTL12 MAFB exhibited an increasing trend in percentage, with maximum intensity in the inflamed condition, while enterochromaffin cell showed a decreasing trend in percentage and strong intensity in the healthy condition. CALCRL, RAMP2 and RAMP3 did not show any expression in the dataset while ACKR3 displayed very little expression in the enterocytes subtypes. In the proportional plot, most cell types stayed with similar abundance except for increase in the goblet cells MUC2 TFF1 and stem cells OLFM4 and a stepwise decrease in the enterocytes TMIGD1 MEP1A.

In the colon dataset, ADM displayed more expression in the epithelial cells (Figure 26b), particularly in goblet cell (MUC2 TFF1) and enterocytes (TMIGD1 MEP1A) with a noticeable increasing trend in both. Goblet cells showed maximum intensity in the non-inflamed condition while enterocytes exhibited maximum intensity in the inflamed condition. Paneth cell displayed highest percentage and maximum intensity in the healthy condition. There was a consistent expression of ADM in most of epithelial cell types in colon. There was no significant expression for other markers in the dataset. From the proportion plot, the least difference in cell population between the healthy and diseased condition was observed in most of the cell types of colon dataset.



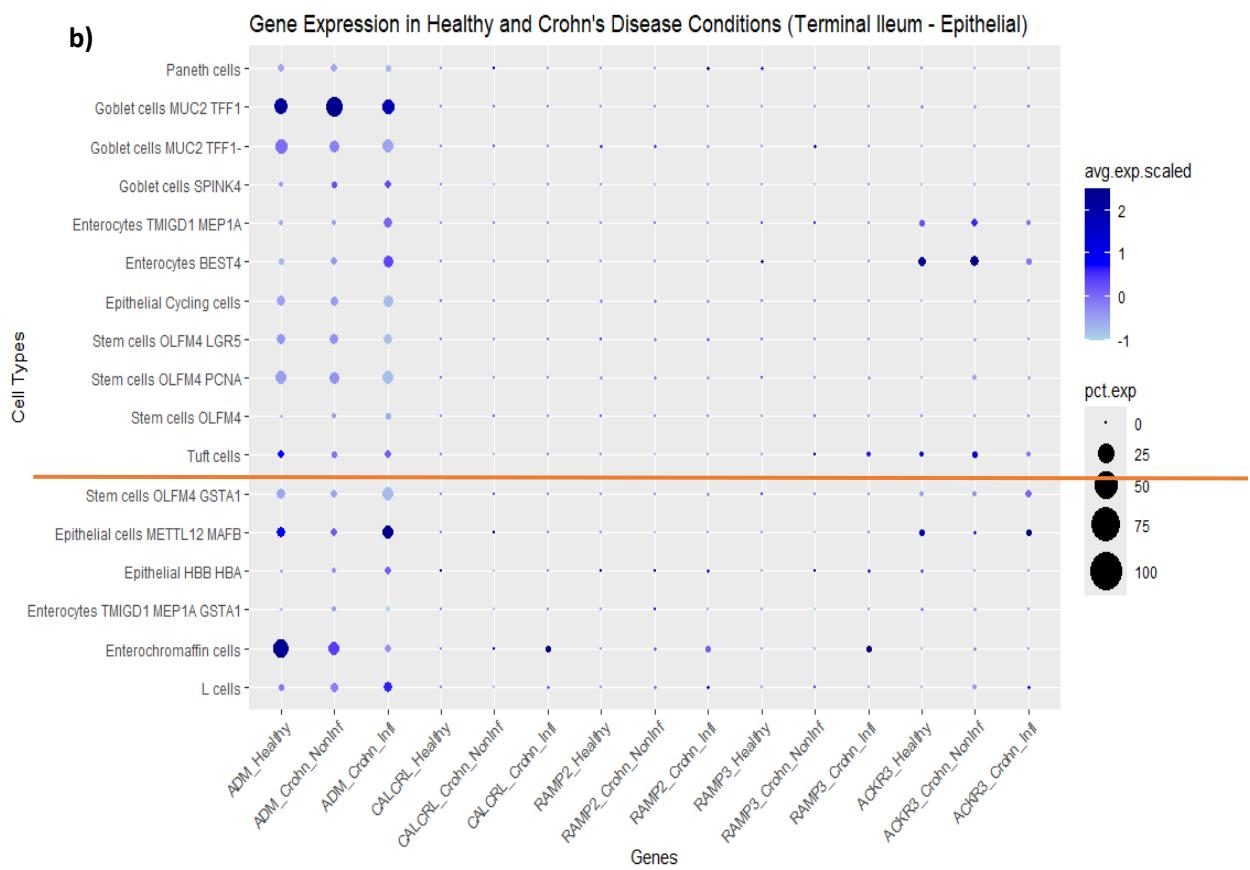
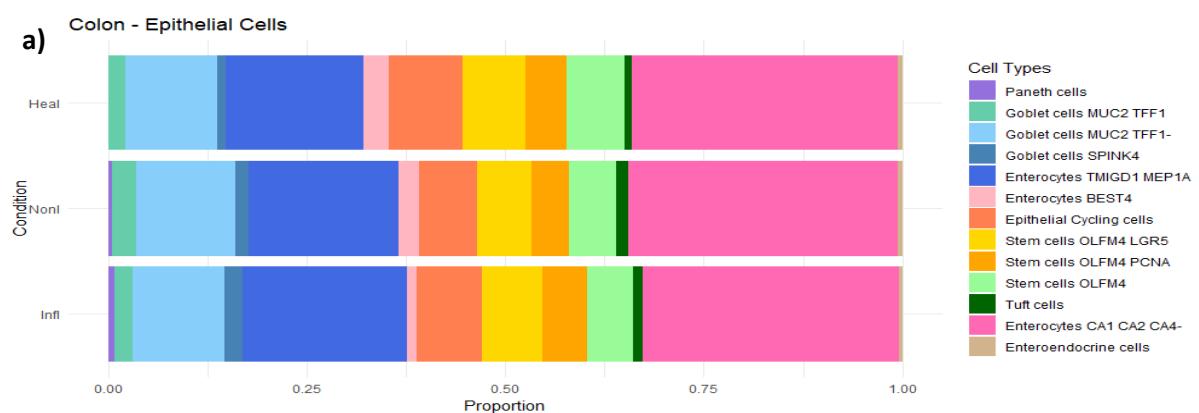


Figure 25. Epithelial Cell Composition and Gene Expression in Terminal Ileum. a) Proportional plot showing the distribution of cell types across healthy, non-inflamed, and inflamed conditions b) Dot plot showing the expression levels of levels of genes ADM, CALCRL, RAMP2, RAMP3, ACKR3 across cell types under healthy, non-inflamed, and inflamed conditions. Dot size represents the percentage of cells expressing the gene, while the colour intensity indicates the average scaled expression level. The orange line separates the unique cell type present in the terminal ileum while the cell populations above the line are the shared populations with colon.



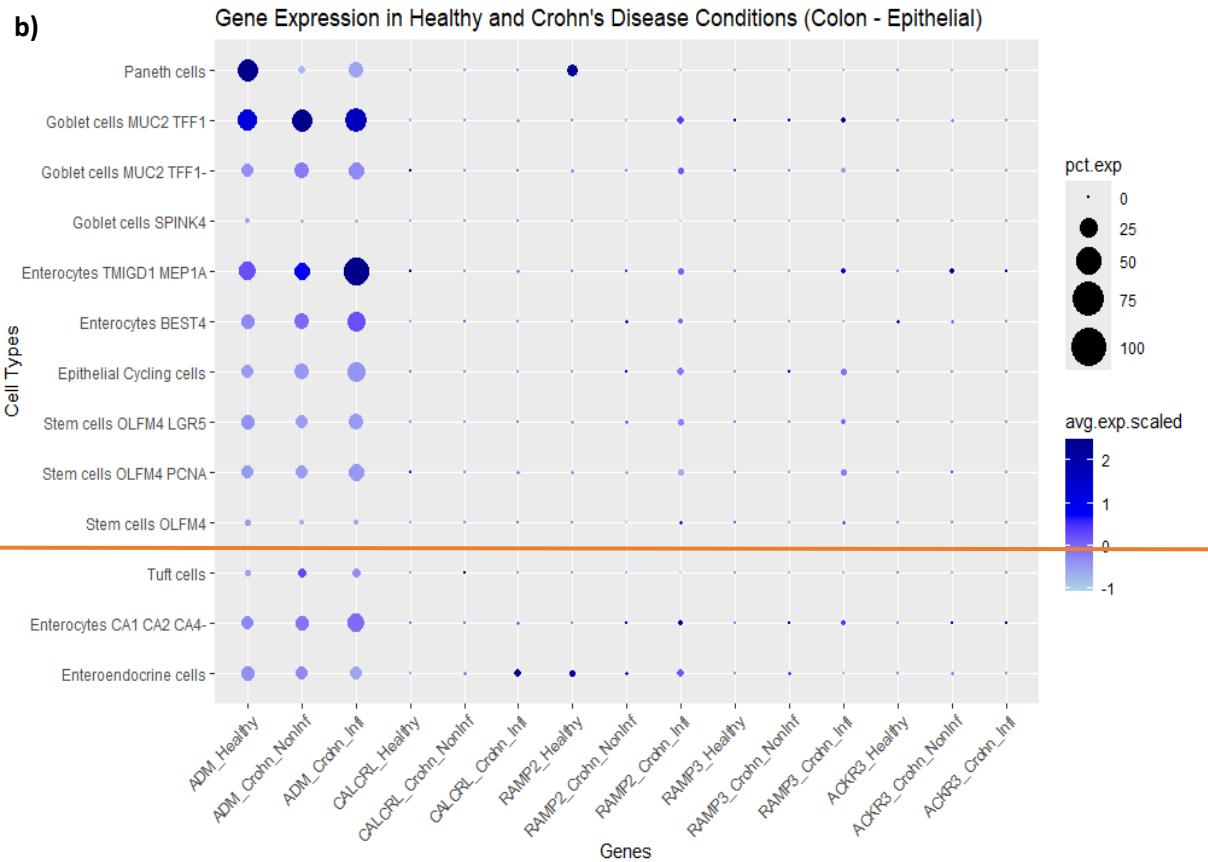


Figure 26. Epithelial Cell Composition and Gene Expression in Colon. a) Proportional plot showing the distribution of cell types across healthy, non-inflamed, and inflamed conditions b) Dot plot showing the expression levels of genes ADM, CALCRL, RAMP2, RAMP3, ACKR3 across cell types under healthy, non-inflamed, and inflamed conditions. Dot size represents the percentage of cells expressing the gene, while the colour intensity indicates the average scaled expression level. The orange line separated the unique cell types present in the colon while the cell populations above the line are the shared populations with terminal ileum.

3.7 Pediatric Cell Population

The pediatric dataset consisted of total 22,500 cells with 26,657 genes from terminal ileum. The dataset formed well-separated clusters with different cell types based on clustering with the same marker genes that were used in original paper for this pediatric dataset. Analysis distinguishes 41 different cell types that can be separated in epithelial, stromal and immune cells. The UMAP (Figure 27) was used to visualize these 41 cell types including some crypt, enterocytes subtypes, goblet, paneth, tuft as epithelial cells, fibroblast subtypes and endothelial subtypes as stromal cells and B cells, plasma cells, dendritic cells, T cells, monocytes and macrophages as immune cell population.

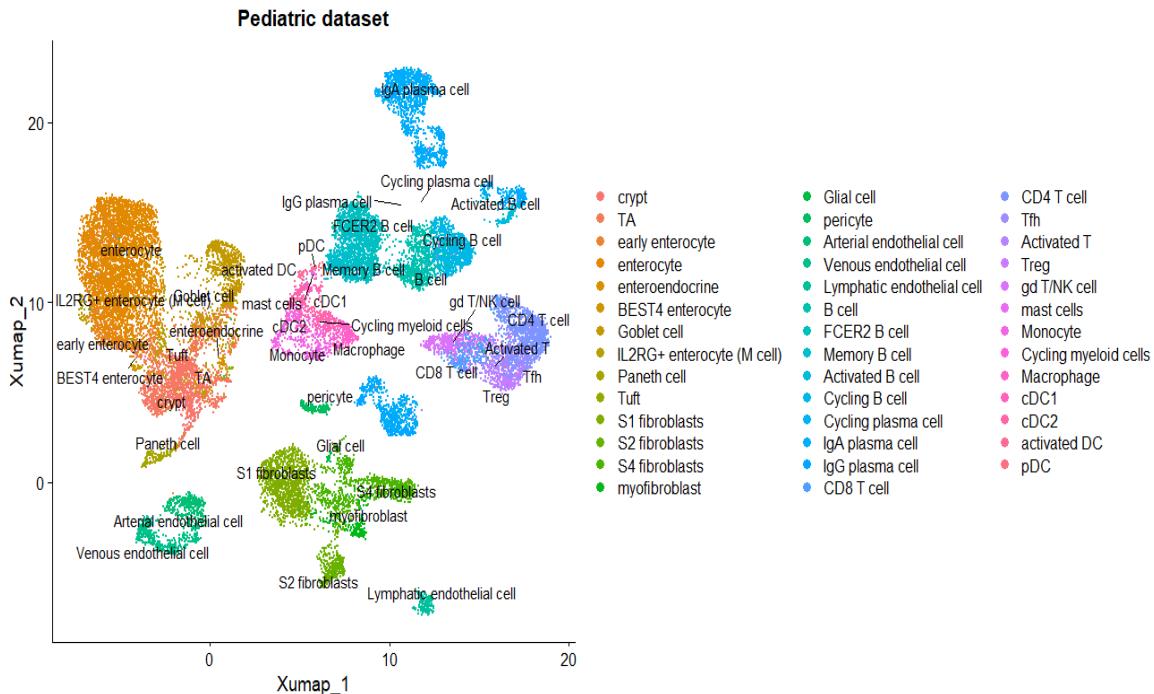


Figure 27.. Pediatric dataset. UMAP displaying the annotated clusters into different cell types.

Feature scatter plot was used to visualize the genes of interest expression in pediatric cell population. It was found that ADM (Figure 28a) was widely distributed in the fibroblast subtypes and some epithelial cell population. The receptors CALCRL, RAMP2 and RAMP3 (Figure 28b-d) showed notable expression in endothelial subtypes similar to human adult dataset while RAMP3 showed no expression in lymphatic endothelial. ACKR3 (Figure 28e) again, just like the adult dataset goes in line with expression in fibroblasts similar to ADM.

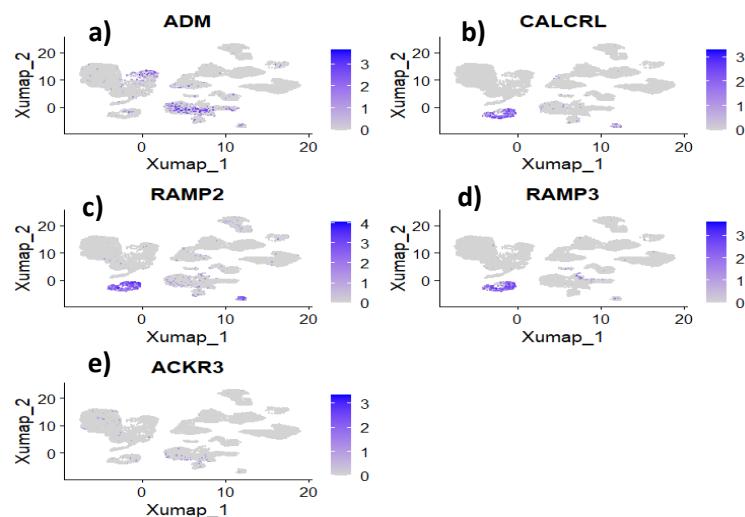
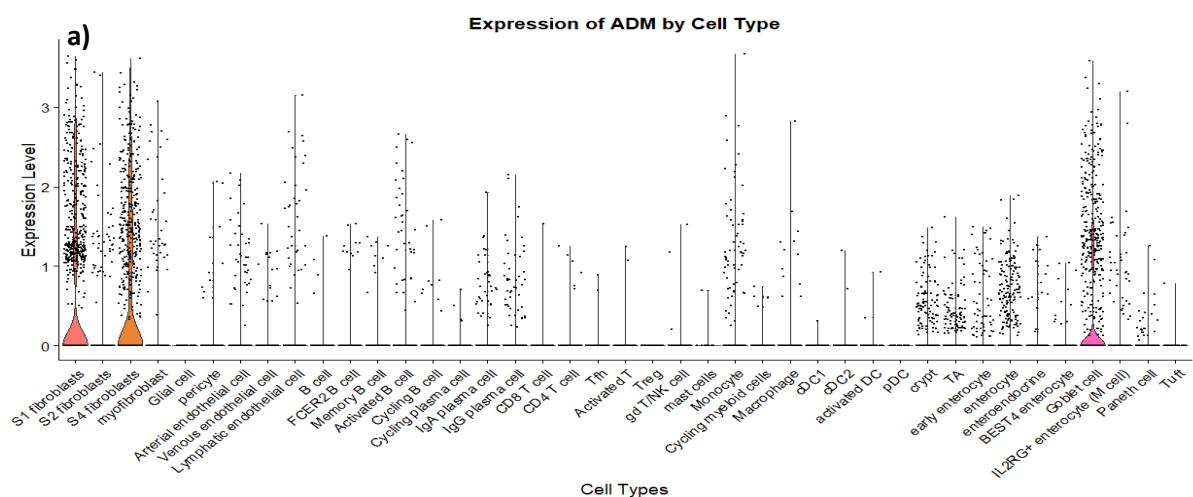


Figure 28. Feature plot showing the expression of marker genes in the pediatric dataset. Subplots represent expression levels of (a) ADM, (b) CALCRL, (c) RAMP2, (d) RAMP3, and (e) ACKR3. Each subplot shows UMAP projections with colour intensity reflecting the expression level of each gene, with darker shades indicating higher expression.

To analyze the dataset in depth for the expression level, violin plots were generated to visualize the distribution across different cell types. ADM (Figure 29a) was observed significantly in S1 and S4 fibroblast in the stromal cell population and mainly in goblet cells for epithelial cell population which aligns well with the human adult dataset as Fibroblast ADAMDEC1 has similar markers to S1 and S4 fibroblast from pediatric dataset.

Receptors were mainly observed in the endothelial subtypes. CALCRL (Figure 29b) showed to have the most varied expression in venous and arterial endothelial cells with some expression in S1 fibroblast. RAMP2 (Figure S13a) showed cells expressing highest normalized expression values in the arterial endothelial cells but the number of cells expressing were similar in arterial and venous endothelial cells. Here, it was observed that there was some expression in fibroblast subtypes specifically S1 being the dominant among them as well as plasma cells. RAMP3 (Figure S13b) exhibited expression only in two endothelial subtypes namely arterial and venous and some expression in pericyte and S4 fibroblast. ACKR3 (Figure S13c) displayed some cell expressing mainly in S1 fibroblast and enterocytes.



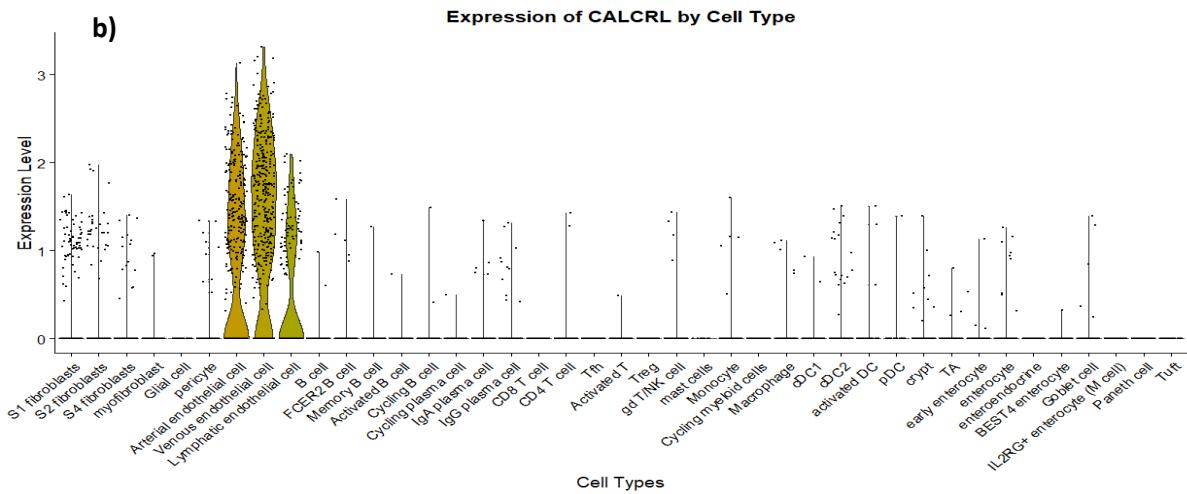


Figure 29. *Pediatric dataset*. Violin plots showing the expression marker genes across various cell types. a) ADM b) CALCRL

Comparative analysis was conducted to evaluate the expression pattern of marker genes across normal control and Crohn's disease condition using violin plots. As shown in the violin plots ADM (Figure 30a) exhibited reduced numbers of expressing cells in Crohn's disease with majority of cells displaying low expression. CALCRL and RAMP3 (Figure 30b, 30d) showed a significant shift in cell expression pattern in disease condition with overall uniform expression. RAMP2 (Figure 30c) displayed higher number of cells in Crohn disease while majority of cells in healthy condition demonstrated lower expression. ACKR3 (Figure 30e) maintained a consistent expression profile across both normal control and Crohn's disease, showing no significant changes between two states.

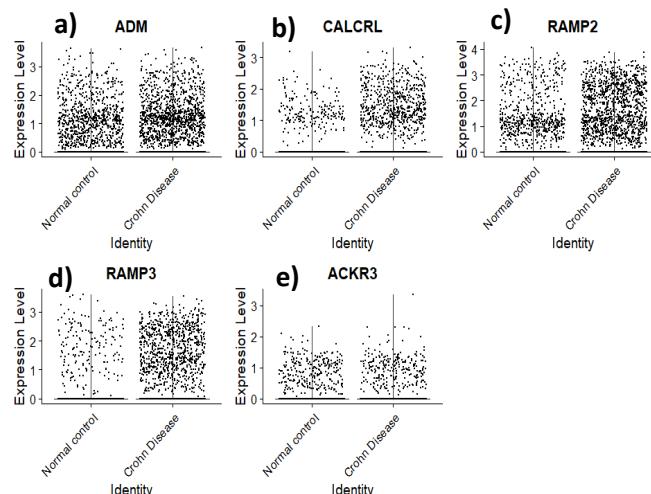


Figure 30. Violin plot representing gene expression across different condition (Normal control, Crohn's disease) in Pediatric dataset. a) ADM b) CALCRL c) RAMP2 d) RAMP3 e) ACKR3

To further explore understanding of the expression patterns of ADM and key receptors, violin plots were generated to assess the distribution across various cell types in both the healthy and disease conditions. ADM (Figure 31a) displayed similar cell expression in both conditions for S1 fibroblast.

However, S4 fibroblast, lymphatic endothelial cell, M cell showed higher cell expression in the diseased condition. In contrast, B cells and goblet exhibited higher expression in normal control.

Endothelial subtypes (including arterial, venous and lymphatic) exhibited increased expression of the receptors CALCRL, RAMP2, RAMP3 (Figure 31b, S14a-c) in the Crohn's disease condition compared to the normal control. Fibroblasts, particularly in the normal control, showed higher expression of RAMP2. Cell expressing ACKR3 (Figure S14e) remained similar in both the conditions across most cell types.

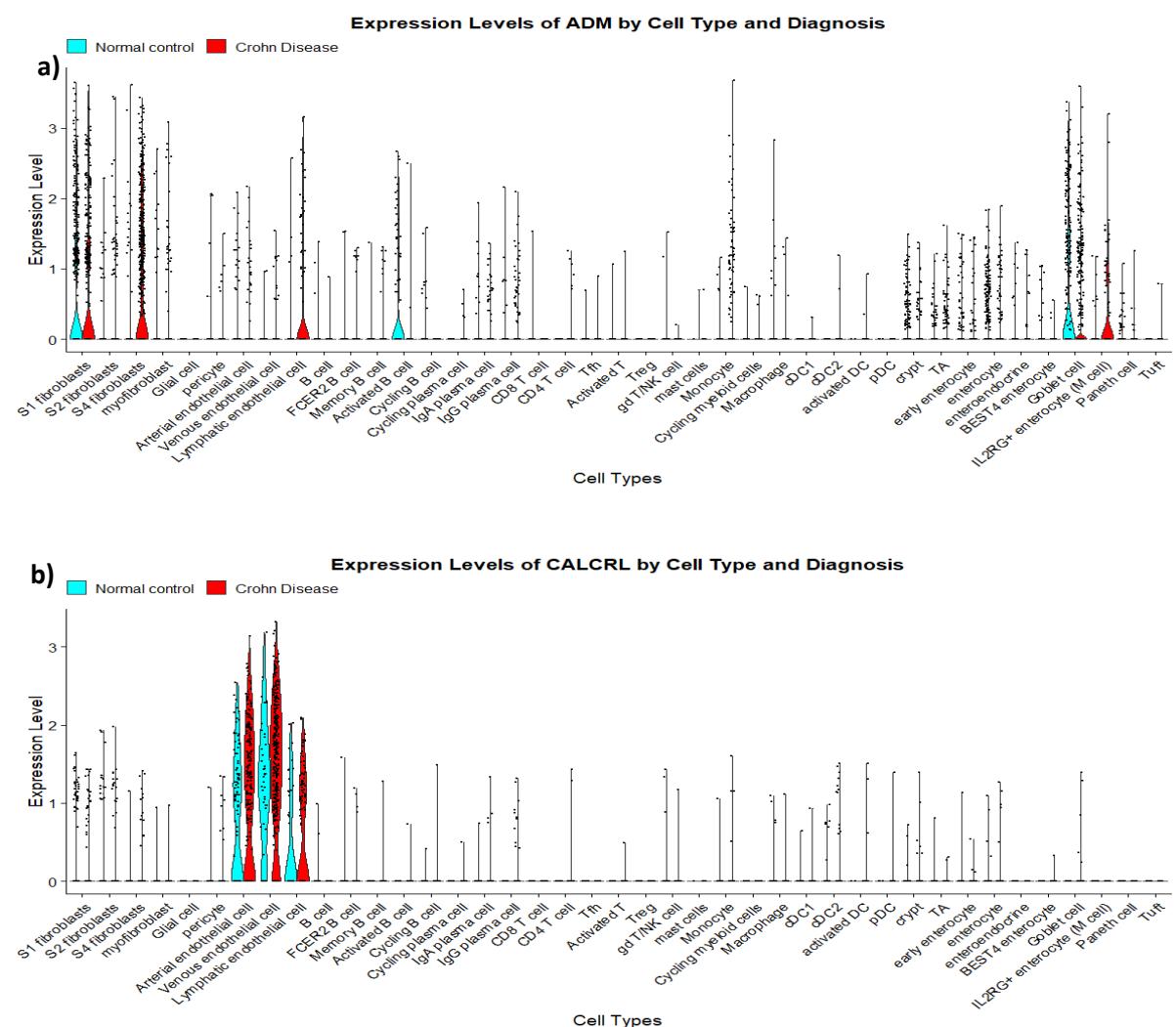


Figure 31. Pediatric dataset. Violin plots showing the expression levels of marker across various cell types and conditions (Healthy -cyan, Crohn disease – red). a) ADM b) CALCRL

The dataset was statistically analyzed to assess the cell count (Figure 32) across different cell types for both the normal control and Crohn's disease conditions. This ensured the observed differences are due to biological variance rather than disparities in the numbers.

S4 fibroblast, pericytes, endothelial subtypes, plasma cells, monocytes and goblet cell exhibited higher cell count in the Crohn's disease. In contrast, B cells, dendritic cells and enterocytes showed greater number of cells in normal control. Some cell types, such as S1 fibroblasts displayed nearly similar cell count in both the conditions.

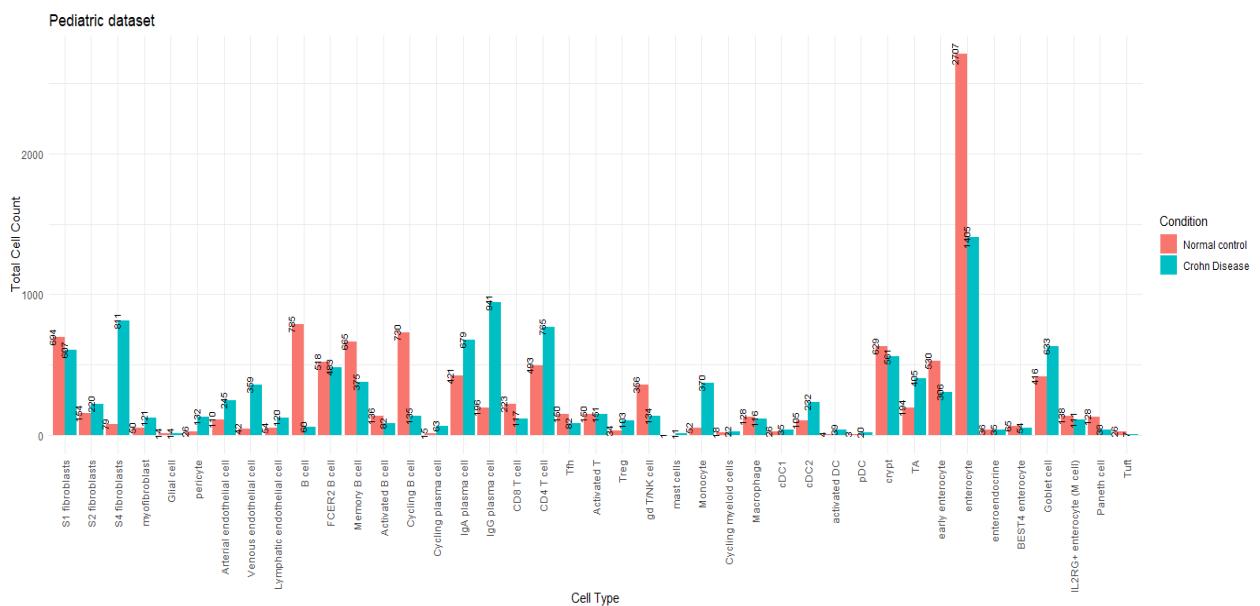


Figure 32. Pediatric Cell Count Distribution. Bar plot representing the cell count across different conditions (Normal control -red, Crohn disease – blue).

To address the disparities in cell counts, the data was normalized and expressed as a percentage in the dot plot (Figure 33d). Additionally, a proportion plot (Figures 33a-c) was included to verify that variations in cell proportions did not influence the expression patterns seen in the dot plot.

From the stromal proportional plot (Figure 33a), there was noticeable increase in the S4 fibroblast disease condition while S2, S1 showed significant decrease in the disease condition. The endothelial subtypes remained consistent except the venous endothelial which shifted drastically in the disease condition. The immune cell population (Figure 33b) showed an increase in the plasma cells, monocytes in the diseased condition while decreased in the B cells. For epithelial cell population (Figure 33c) it was observed that there was a decrease in the enterocyte, early enterocytes with the most significant change however, crypt, TA and goblet cells increased in the disease condition.

ADM showed the highest expression across fibroblast subtypes, consistent with findings in the human adult dataset. Specifically, S1 fibroblasts and S4 fibroblasts showed the most notable expression in both healthy and disease conditions that share same markers as of fibroblast ADAMDEC1. While S1

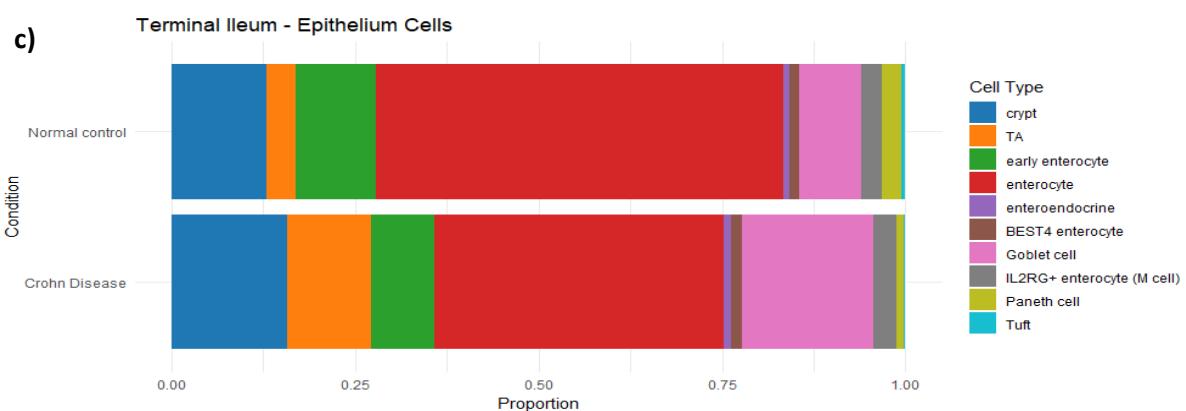
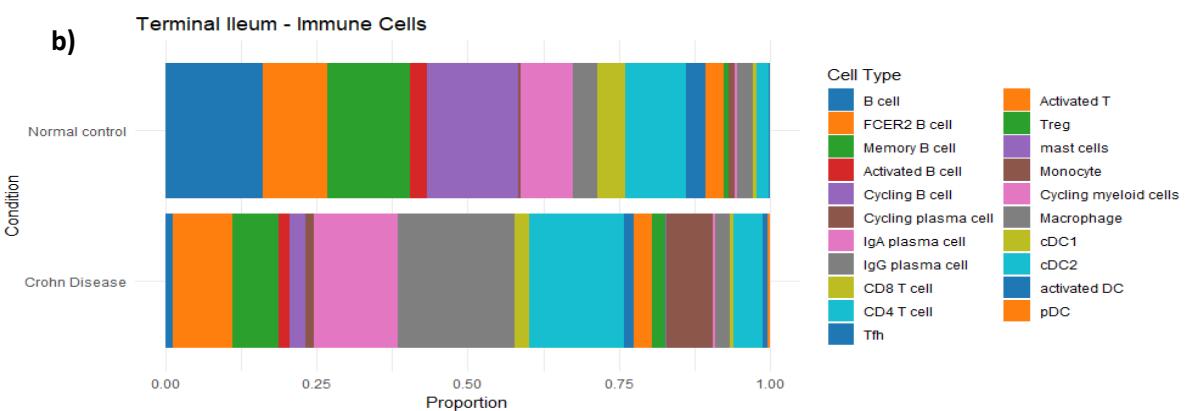
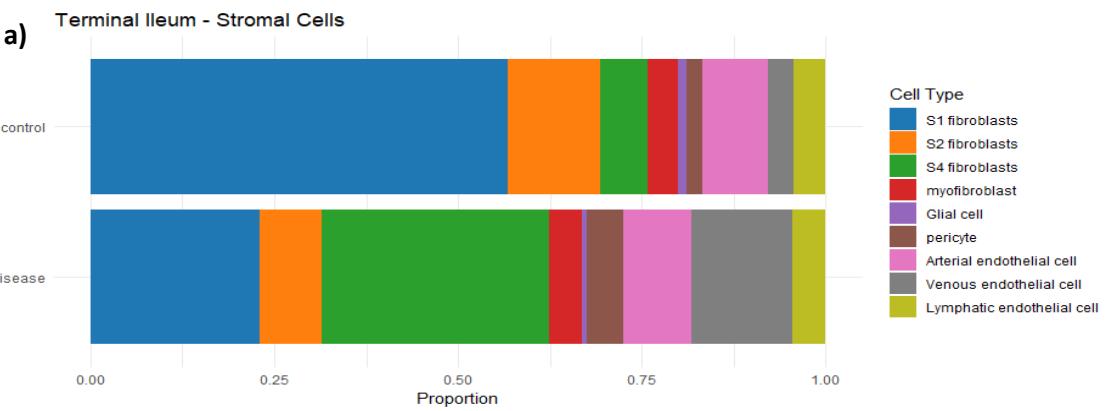
fibroblasts maintained similar percentage expression between healthy and disease conditions, S2 and S4 fibroblasts showed an increasing trend in the percentage expression. Goblet cells also exhibited high ADM percentage expression and intensity, particularly in the healthy condition but monocyte and M cell showed a notable increased expression in the Crohn's disease condition which contrasts with findings in human adult dataset. Lymphatic endothelial exhibited some expression of ADM in the disease condition but arterial and venous did not show any expression. Overall, ADM expression remains primarily in the fibroblast subtypes and goblet cells, consistent with the human adult dataset.

CALCRL showed most expression in all endothelial subtypes, with lower expression in lymphatic endothelial cells similar to adult human dataset. There was very faint expression in fibroblasts and no expression in B cells, dendritic cells and monocytes which all goes in line with the human adult dataset.

RAMP2 displayed strong expression across all endothelial subtypes and the percentage and intensity remained uniform in both the condition which is unlike to the decreasing trend observed in human adult endothelial subtypes. Fibroblast exhibited some expression of RAMP2 in the healthy condition but not in disease state, which align well with the human adult fibroblast expression. Immune cells including plasma cells, monocytes and macrophages displayed very faint expression which follows the expression in human adult immune cells, with higher intensity. There was no expression of RAMP2 in epithelial cells in both the pediatric and adult dataset.

RAMP3 showed highest expression in venous endothelial cells, followed by arterial but no expression in lymphatics consistent with expression pattern in the human adult endothelial subtypes. It also exhibited some expression in fibroblast S4 in the healthy condition and pericytes in crohn's disease, similar to expression in human adult where RAMP3 expressed in Activated fibroblasts CCL19 ADAMDEC1.

ACKR3 showed higher expression in fibroblast subtypes, lymphatics endothelial cells and pericytes which is cell type dependent whether expression was higher or lower in Crohn's disease.



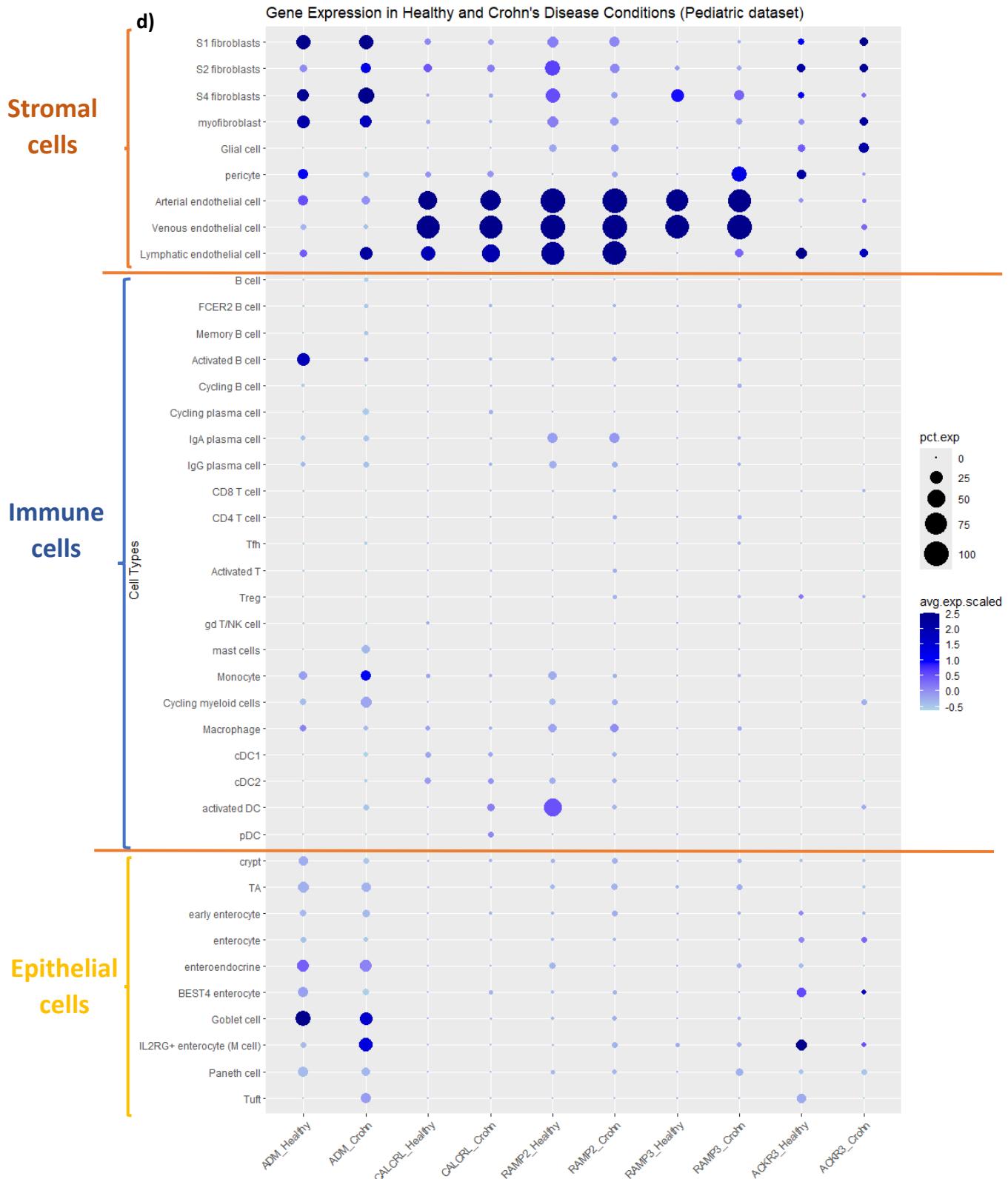


Figure 33. Pediatric Cell Composition and Gene Expression. a) Proportional plot showing the distribution of stromal cell types across healthy and Crohn's disease condition. b) Proportional plot showing the distribution of immune cell types. c) Proportional plot showing the distribution of epithelial cell types. d) Dot plot showing the expression levels of marker genes ADM, CALCRL, RAMP2, RAMP3, ACKR3 across stromal, immune and epithelial cell types under healthy and Crohn's disease condition. Dot size represents the percentage of cells expressing the gene, while the colour intensity indicates the average scaled expression level. The orange line denotes the separation of all cell types into stromal, immune and epithelial cell population.

4. Discussion and Conclusion

Discussion

The aim for this study was to explore the expression patterns and potential roles of Adrenomedullin and its key receptors (CALCRL, RAMP2, RAMP3, and ACKR3) in both the dataset consisting of stromal, immune, and epithelial cell populations between the different locations, age group, tissue specific changes and its potential signalling pathways. This was achieved using scRNA-seq analysis technique on both the datasets. Despite the site difference (terminal ileum vs colon) and age group (human adult vs pediatric), the overall expression remained nearly consistent. We can only say that the age group did not majorly affect the expression patterns, but nothing can be said about the mechanism of action of the disease.

The predominant expression of ADM and its key receptors were observed in the stromal cell population among all the cell types in both terminal ileum and colon in human adult dataset and we observed similar consistency with the pediatric dataset. The results showed that ADM is expressed in a variety of stromal cell populations, with fibroblasts exhibiting the highest expression. In contrast, endothelial cells were the primary source of expression for CALCRL, RAMP2, and RAMP3. This supports the hypothesis that the endothelium layer mediates the effects of ADM, especially in vascular regulation (Spoto et al., 2024).

Pathway analysis in the terminal ileum stromal cells, revealed ADM's involvement in the vascular smooth muscle contraction pathway, which was of particular interest because it was upregulated in the non-inflammatory state. This analysis revealed a significant interaction between ADM and vasodilatory pathways, despite the lack of a direct correlation between ADM and the onset or remission of Crohn's disease. This discovery could suggest that ADM modulates disease indirectly by influencing vascular responses, particularly by activating downstream effectors like CRLR to encourage vasodilation.

Fibroblast expression of ADM in the gut is particularly interesting due to its close proximity to the endothelial cell population. This spatial relationship suggests a potential paracrine effect, where fibroblast-derived ADM could influence endothelial cell function (Enzerink and Vaheri, 2011). Based on the high expression of RAMP2, RAMP3 and CALCRL in endothelial cells, it suggests that ADM primarily regulates vascular tone, permeability, and homeostasis. From the KEGG pathway result, these receptor complexes allowed ADM's vasodilatory action, one of the function, to occur crucial for preserving vascular integrity in times of stress or inflammation (Iring et al., 2019). By strengthening vascular barrier function, decreasing vascular permeability, and encouraging vascular remodelling, ADM may have a protective effect in inflammatory conditions such as Crohn's disease, where endothelial dysfunction contributes to the progression of the disease (Tanaka et al., 2016). This could be a critical factor in tissue repair and inflammation regulation in Crohn's disease.

Only macrophages and monocytes showed significant expression of ADM in immune cells, which is probably due to an inflammatory response mechanism. This suggests that while ADM may affect inflammation through vascular or paracrine mechanisms, it does not play a major role in immune modulation (Lu et al., 2024). Monocytes and macrophages are pivotal players in the innate immune response, with a well-established role in both maintaining gut homeostasis and contributing to inflammation during disease states (Roda et al., 2020). The increased ADM expression in monocytes and macrophages under inflamed conditions is supported by previous studies (Wong et al., 2005). On the other hand, endothelial cells are essential for recruiting immune cells to inflammatory sites. Activation of endothelial cells by inflammatory cues, like cytokines (e.g., IL-1, and TNF- α), and on their surface they express adhesion molecules such as integrins and selectins. These molecules aid in leukocyte adhesion, rolling, and transmigration (e.g., monocytes, neutrophils, and other blood cell types) into the tissue that is inflamed (Pober and Sessa, 2007). Dendritic cells, on the other hand, bridge innate and adaptive immunity (Roda et al., 2020) to maintain homeostasis. The immune cell increased under inflammatory conditions align with the key role in innate immune responses, while dendritic cells contribute to maintaining homeostasis.

ADM expression in secretory epithelial cells, especially goblet cells, raises the possibility that it maintains the homeostasis with the release of anti-microbial peptides and regulate immune response locally and preserves mucosal integrity (Raya Tonetti et al., 2024). The primary interface between the host and microorganisms, the intestinal epithelial cells act as a chemical and physical barrier and are essential for fighting off microbial invasion. The expression of ADM in all these cells such as goblet cells, paneth cells, and enterocytes, indicates the production of antimicrobial proteins (AMPs), which are key innate immune effectors that control the gut microbiota by either direct elimination or inhibition of microorganisms especially in the colon (Ra and Bang, 2024). By attracting different immune cells to the infection site, they enhance the immune response. This serves to preserve homeostasis, and any dysregulation in AMPs can affect the balance of microbiota, which in turn can lead to a number of intestinal disorders, including IBD (Mukherjee and Hooper, 2015).

The receptor ACKR3, whose expression patterns resembled those of ADM, could serve as a scavenging agent. However, given its established higher affinity for the Pro-Adrenomedullin N-terminal peptide (PAMP) and lower to AM and also that adrenomedullin peptide may not be translated from the gene, it may have a wider regulatory role in the bioavailability of ADM (Meyrath et al., 2021). Although ACKR3 is best known for being a CXCL12 scavenger receptor, this study's findings co-expresses with ADM indicate that it may also be involved in controlling the bioavailability of ADM. Proadrenomedullin N-terminal 20 peptide (PAMP), which has a stronger affinity for ACKR3, is another protein encoded by the ADM gene (Meyrath et al., 2021). Whether ACKR3 binds to ADM directly is still unknown. Since ADM is not always generated in its active state, there may be more going on in

the relationship between ACKR3 and ADM. ACKR3 may interact with the translated form of ADM or bind to PAMP and carry out an antimicrobial action (Sigmund et al., 2023). In conclusion, while ACKR3 is known to act as a scavenging molecule, its exact interaction with ADM remains to be fully understood.

Interestingly, a notable observation was the higher expression of ADM and its key receptors in non-inflamed tissues of the terminal ileum stromal cells. This may indicate that ADM plays a regulatory role in resolving inflammation or preventing its recurrence. Given ADM's role in maintaining vascular integrity and immune homeostasis, its expression in non-inflamed conditions could suggest a protective function against chronic inflammation. However, variations in sample numbers across conditions, particularly in the adult datasets, may influence this observation. The pediatric dataset, which had more balanced sample numbers across conditions, provides more robust support for this hypothesis. But for this dataset, we do not know the inflammation condition of the patient which is an important aspect and has distinct behaviour in the adult dataset. Data normalization was essential in revealing the actual expression patterns of ADM and its key receptors across varying sample sizes, ensuring that ADM's potential regulatory role is accurately reflected.

Conclusion

The results of this study highlight the roles of stromal, endothelial, and immune cells in ADM's regulatory effects, which improves our understanding of ADM signalling in Crohn's disease. To sum up, the significant expression of RAMP2, RAMP3, and CALCRL in endothelial cells for stromal cell population emphasizes the role of ADM in endothelial-mediated processes in the gastrointestinal tract. This insight provides a foundation for further research into the therapeutic potential of targeting the ADM in gastrointestinal disease.

Limitations and Future Perspective:

This study's limitations include the reliance on single-cell RNA sequencing, which captures transcriptional activity but lacks protein-level validation. Future research should focus on validating these findings through proteomic analyses to confirm the roles of ADM and its receptors at the protein level.

In summary, ADM's diverse role in the intestinal microenvironment is highlighted by its signalling through endothelial cells as well as possible paracrine effects in epithelial and immune cells. According to this research, ADM may have a significant role in Crohn's disease tissue repair and inflammation resolution, and its potential therapeutic uses are worth investigating.

5. REFERENCES

- Agorreta J, Zulueta JJ, Montuenga LM, Garayoa M (2005) Adrenomedullin expression in a rat model of acute lung injury induced by hypoxia and LPS. *American Journal of Physiology-Lung Cellular and Molecular Physiology* 288:L536–L545.
- Allaker RP, Grosvenor PW, McAnerney DC, Sheehan BE, Srikanta BH, Pell K, Kapas S (2006) Mechanisms of adrenomedullin antimicrobial action. *Peptides* 27:661–666.
- Aloi M, D’Arcangelo G, Pofi F, Vassallo F, Rizzo V, Nuti F, Di Nardo G, Pierdomenico M, Viola F, Cucchiara S (2013) Presenting features and disease course of pediatric ulcerative colitis. *Journal of Crohn’s and Colitis* 7:e509–e515.
- Andrew McDavid [Aut C (2017) MAST. Available at: <https://bioconductor.org/packages/MAST> [Accessed September 22, 2024].
- Andrews TS, Hemberg M (2018) Identifying cell populations with scRNASeq. *Molecular Aspects of Medicine* 59:114–122.
- Ashizuka S, Inagaki-Ohara K, Kuwasako K, Kato J, Inatsu H, Kitamura K (2009) Adrenomedullin treatment reduces intestinal inflammation and maintains epithelial barrier function in mice administered dextran sulphate sodium. *Microbiology and Immunology* 53:573–581.
- Ashizuka S, Inatsu H, Inagaki-Ohara K, Kita T, Kitamura K (2013) Adrenomedullin as a Potential Therapeutic Agent for Inflammatory Bowel Disease. *CPPS* 14:246–255.
- Ashizuka S, Ishikawa N, Kato J, Yamaga J, Inatsu H, Eto T, Kitamura K (2005) Effect of adrenomedullin administration on acetic acid-induced colitis in rats. *Peptides* 26:2610–2615.
- Ashizuka S, Kita T, Inatsu H, Kitamura K (2021) Adrenomedullin: A Novel Therapeutic for the Treatment of Inflammatory Bowel Disease. *Biomedicines* 9:1068.
- Ashizuka S, Kuroishi N, Nakashima K, Inatsu H, Kita T, Kitamura K (2019) Adrenomedullin: A Novel Therapy for Intractable Crohn’s Disease with a Loss of Response to Infliximab. *Intern Med* 58:1573–1576.
- Atreya R, Bojarski C, Kühl AA, Trajanoski Z, Neurath MF, Siegmund B (2022) Ileal and colonic Crohn’s disease: Does location makes a difference in therapy efficacy? *Current Research in Pharmacology and Drug Discovery* 3:100097.
- Atreya R, Siegmund B (2021) Location is important: differentiation between ileal and colonic Crohn’s disease. *Nat Rev Gastroenterol Hepatol* 18:544–558.
- Azizi E, Carr AJ, Plitas G, Cornish AE, Konopacki C, Prabhakaran S, Nainys J, Wu K, Kiseliovas V, Setty M, Choi K, Fromme RM, Dao P, McKenney PT, Wasti RC, Kadaveru K, Mazutis L, Rudensky AY, Pe’er D (2018) Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. *Cell* 174:1293–1308.e36.
- Bacher R, Kendziorski C (2016) Design and computational analysis of single-cell RNA-sequencing experiments. *Genome Biol* 17:63.
- Baran-Gale J, Chandra T, Kirschner K (2018) Experimental design for single-cell RNA sequencing. *Briefings in Functional Genomics* 17:233–239.
- Barnhoorn MC, Hakuno SK, Bruckner RS, Rogler G, Hawinkels LJAC, Scharl M (2020) Stromal Cells in the Pathogenesis of Inflammatory Bowel Disease. *Journal of Crohn’s and Colitis* 14:995–1009.
- Baumgart DC, Sandborn WJ (2012) Crohn’s disease. *The Lancet* 380:1590–1605.
- Becht E, McInnes L, Healy J, Dutertre C-A, Kwok IWH, Ng LG, Ginhoux F, Newell EW (2019) Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol* 37:38–44.
- Bernstein CN, Hitchon CA, Walld R, Bolton JM, Sareen J, Walker JR, Graff LA, Patten SB, Singer A, Lix LM, El-Gabalawy R, Katz A, Fisk JD, Marrie RA, CIHR Team in Defining the Burden and

- Managing the Effects of Psychiatric Comorbidity in Chronic Immunoinflammatory Disease (2019) Increased Burden of Psychiatric Disorders in Inflammatory Bowel Disease. *Inflammatory Bowel Diseases* 25:360–368.
- Bouhuys M, Lexmond WS, Van Rheenen PF (2023) Pediatric Inflammatory Bowel Disease. *Pediatrics* 151:e2022058037.
- Bowcutt R (2014) Heterogeneity across the murine small and large intestine. *WJG* 20:15216.
- Boyle EI, Weng S, Gollub J, Jin H, Botstein D, Cherry JM, Sherlock G (2004) GO::TermFinder—open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* 20:3710–3715.
- Brain SD, Grant AD (2004) Vascular Actions of Calcitonin Gene-Related Peptide and Adrenomedullin. *Physiological Reviews* 84:903–934.
- Brennecke P, Anders S, Kim JK, Kołodziejczyk AA, Zhang X, Proserpio V, Baying B, Benes V, Teichmann SA, Marioni JC, Heisler MG (2013) Accounting for technical noise in single-cell RNA-seq experiments. *Nat Methods* 10:1093–1095.
- Buettner F, Natarajan KN, Casale FP, Proserpio V, Scialdone A, Theis FJ, Teichmann SA, Marioni JC, Stegle O (2015) Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat Biotechnol* 33:155–160.
- Buettner F, Pratanwanich N, McCarthy DJ, Marioni JC, Stegle O (2017) f-scLVM: scalable and versatile factor analysis for single-cell RNA-seq. *Genome Biol* 18:212.
- Butler A, Hoffman P, Smibert P, Papalexi E, Satija R (2018) Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 36:411–420.
- Büttner M, Miao Z, Wolf FA, Teichmann SA, Theis FJ (2019) A test metric for assessing single-cell RNA-seq batch correction. *Nat Methods* 16:43–49.
- Chen G, Ning B, Shi T (2019) Single-Cell RNA-Seq Technologies and Related Computational Data Analysis. *Front Genet* 10:317.
- Cheung B, Hwang I, Li C, O W, Tsang K, Leung R, Kumana C, Tang F (2004) Increased adrenomedullin expression in lungs in endotoxaemia. *Journal of Endocrinology* 181:339–345.
- Chung NC, Storey JD (2015) Statistical significance of variables driving systematic variation in high-dimensional data. *Bioinformatics* 31:545–554.
- Cleynen I et al. (2016) Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. *The Lancet* 387:156–167.
- Coifman RR, Lafon S, Lee AB, Maggioni M, Nadler B, Warner F, Zucker SW (2005) Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proc Natl Acad Sci USA* 102:7426–7431.
- Coskun M, Vermeire S, Nielsen OH (2017) Novel Targeted Therapies for Inflammatory Bowel Disease. *Trends in Pharmacological Sciences* 38:127–142.
- D'Amico F, Fiorino G, Furfaro F, Allocca M, Danese S (2018) Janus kinase inhibitors for the treatment of inflammatory bowel diseases: developments from phase I and phase II clinical trials. *Expert Opinion on Investigational Drugs* 27:595–599.
- Das S, Rai A, Rai SN (2022) Differential Expression Analysis of Single-Cell RNA-Seq Data: Current Statistical Approaches and Outstanding Challenges. *Entropy* 24:995.
- Dozmorov MG (2017) Epigenomic annotation-based interpretation of genomic data: from enrichment analysis to machine learning. *Valencia A, ed. Bioinformatics* 33:3323–3330.
- Dulai PS, Singh S, Vande Casteele N, Boland BS, Rivera-Nieves J, Ernst PB, Eckmann L, Barrett KE, Chang JT, Sandborn WJ (2019) Should We Divide Crohn's Disease Into Ileum-Dominant and Isolated Colonic Diseases? *Clinical Gastroenterology and Hepatology* 17:2634–2643.
- Duò A, Robinson MD, Soneson C (2020) A systematic performance evaluation of clustering methods for single-cell RNA-seq data. *F1000Res* 7:1141.

- Elmentait R, Ross ADB, Roberts K, James KR, Ortmann D, Gomes T, Nayak K, Tuck L, Pritchard S, Bayraktar OA, Heuschkel R, Vallier L, Teichmann SA, Zilbauer M (2020) Single-Cell Sequencing of Developing Human Gut Reveals Transcriptional Links to Childhood Crohn's Disease. *Developmental Cell* 55:771-783.e5.
- Enzerink A, Vaheri A (2011) Fibroblast activation in vascular inflammation. *Journal of Thrombosis and Haemostasis* 9:619–626.
- Fakhoury M, Al-Salami H, Negruj R, Mooradian A (2014) Inflammatory bowel disease: clinical aspects and treatments. *JIR*:113.
- Fernández De Arcaya I, Lostao MP, Martínez A, Berjón A, Barber A (2005) Effect of adrenomedullin and proadrenomedullin N-terminal 20 peptide on sugar transport in the rat intestine. *Regulatory Peptides* 129:147–154.
- Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, Slichter CK, Miller HW, McElrath MJ, Prlic M, Linsley PS, Gottardo R (2015) MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol* 16:278.
- Fischer J-P, Els-Heindl S, Beck-Sickinger AG (2020) Adrenomedullin – Current perspective on a peptide hormone with significant therapeutic potential. *Peptides* 131:170347.
- Freytag S, Tian L, Lönnstedt I, Ng M, Bahlo M (2018) Comparison of clustering tools in R for medium-sized 10x Genomics single-cell RNA-sequencing data. *F1000Res* 7:1297.
- Fukuda K, Tsukada H, Onomura M, Saito T, Kodama M, Nakamura H, Taniguchi T, Tominaga M, Hosokawa M, Seino Y (1998) Effect of adrenomedullin on ion transport and muscle contraction in rat distal colon. *Peptides* 19:1043–1047.
- Fukuda K, Tsukada H, Oya M, Onomura M, Kodama M, Nakamura H, Hosokawa M, Seino Y (1999) Adrenomedullin promotes epithelial restitution of rat and human gastric mucosa in vitro. *Peptides* 20:127–132.
- Gade AK, Douthit NT, Townsley E (2020) Medical Management of Crohn's Disease. *Cureus* Available at: <https://www.cureus.com/articles/30854-medical-management-of-crohns-disease> [Accessed September 30, 2024].
- Gallo RL, Hooper LV (2012) Epithelial antimicrobial defence of the skin and intestine. *Nat Rev Immunol* 12:503–516.
- García-Ponce A, Chávez Paredes S, Castro Ochoa KF, Schnoor M (2016) Regulation of endothelial and epithelial barrier functions by peptide hormones of the adrenomedullin family. *Tissue Barriers* 4:e1228439.
- Gasche C, Scholmerich J, Brynskov J, D'Haens G, Hanauer SB, Irvine JE, Jewell DP, Rachmilewitz D, Sachar DB, Sandborn WJ, Sutherland LR (2000) A Simple Classification of Crohn's Disease: Report of the Working Party for the World Congresses of Gastroenterology, Vienna 1998: Inflammatory Bowel Diseases 6:8–15.
- Gibbons C, Dackor R, Dunworth W, Fritz-Six K, Caron KM (2007) Receptor Activity-Modifying Proteins: RAMPing up Adrenomedullin Signaling. *Molecular Endocrinology* 21:783–796.
- Graham GJ, Locati M, Mantovani A, Rot A, Thelen M (2012) The biochemistry and biology of the atypical chemokine receptors. *Immunology Letters* 145:30–38.
- Griffiths JA, Scialdone A, Marioni JC (2018) Using single-cell genomics to understand developmental processes and cell fate decisions. *Molecular Systems Biology* 14:e8046.
- Gröschl M, Wendler O, Topf H-G, Bohlender J, Köhler H (2009) Significance of salivary adrenomedullin in the maintenance of oral health: Stimulation of oral cell proliferation and antibacterial properties. *Regulatory Peptides* 154:16–22.
- Guan Q (2019) A Comprehensive Review and Update on the Pathogenesis of Inflammatory Bowel Disease. *Journal of Immunology Research* 2019:1–16.

- Guangchuang Yu [Aut C (2017) clusterProfiler. Available at: <https://bioconductor.org/packages/clusterProfiler> [Accessed September 22, 2024].
- Gunawardene AR, Corfe BM, Staton CA (2011) Classification and functions of enteroendocrine cells of the lower gastrointestinal tract: Classification and functions of colorectal enteroendocrine cells. *International Journal of Experimental Pathology* 92:219–231.
- Haghverdi L, Buettner F, Theis FJ (2015) Diffusion maps for high-dimensional single-cell analysis of differentiation data. *Bioinformatics* 31:2989–2998.
- Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, Srivastava A, Molla G, Madad S, Fernandez-Granda C, Satija R (2023) Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol* Available at: <https://www.nature.com/articles/s41587-023-01767-y> [Accessed September 22, 2024].
- Hart AL, Ng SC (2015) Crohn's disease. *Medicine* 43:282–290.
- Hippenstiel S, Witzenrath M, Schmeck B, Hocke A, Krisp M, Krüll M, Seybold J, Seeger W, Rascher W, Schütte H, Suttorp N (2002) Adrenomedullin Reduces Endothelial Hyperpermeability. *Circulation Research* 91:618–625.
- Hong R, Koga Y, Bandyadka S, Leshchyk A, Wang Y, Akavoor V, Cao X, Sarfraz I, Wang Z, Alabdullatif S, Jansen F, Yajima M, Johnson WE, Campbell JD (2022) Comprehensive generation, visualization, and reporting of quality control metrics for single-cell RNA sequencing data. *Nat Commun* 13:1688.
- Ilicic T, Kim JK, Kolodziejczyk AA, Bagger FO, McCarthy DJ, Marioni JC, Teichmann SA (2016) Classification of low quality cells from single-cell RNA-seq data. *Genome Biol* 17:29.
- Iring A, Jin Y-J, Albarrán-Juárez J, Siragusa M, Wang S, Dancs PT, Nakayama A, Tonack S, Chen M, Künne C, Sokol AM, Günther S, Martínez A, Fleming I, Wettschureck N, Graumann J, Weinstein LS, Offermanns S (2019) Shear stress–induced endothelial adrenomedullin signaling regulates vascular tone and blood pressure. *Journal of Clinical Investigation* 129:2775–2791.
- Ishimitsu T, Kojima M, Kangawa K, Hino J, Matsuoka H, Kitamura K, Eto T, Matsuo H (1994) Genomic Structure of Human Adrenomedullin Gene. *Biochemical and Biophysical Research Communications* 203:631–639.
- Islam S, Zeisel A, Joost S, La Manno G, Zajac P, Kasper M, Lönnberg P, Linnarsson S (2014) Quantitative single-cell RNA-seq with unique molecular identifiers. *Nat Methods* 11:163–166.
- Isumi Y, Kubo A, Katafuchi T, Kangawa K, Minamino N (1999) Adrenomedullin suppresses interleukin-1 β -induced tumor necrosis factor- α production in Swiss 3T3 cells. *FEBS Letters* 463:110–114.
- Isumi Y, Shoji H, Sugo S, Tochimoto T, Yoshioka M, Kangawa K, Matsuo H, Minamino N (1998) Regulation of Adrenomedullin Production in Rat Endothelial Cells ¹. *Endocrinology* 139:838–846.
- Kaplan GG, Ng SC (2016) Globalisation of inflammatory bowel disease: perspectives from the evolution of inflammatory bowel disease in the UK and China. *The Lancet Gastroenterology & Hepatology* 1:307–316.
- Kataoka Y, Miyazaki S, Yasuda S, Nagaya N, Noguchi T, Yamada N, Morii I, Kawamura A, Doi K, Miyatake K, Tomoike H, Kangawa K (2010) The First Clinical Pilot Study of Intravenous Adrenomedullin Administration in Patients With Acute Myocardial Infarction. *Journal of Cardiovascular Pharmacology* 56:413–419.
- Kiela PR, Ghishan FK (2016) Physiology of Intestinal Absorption and Secretion. *Best Practice & Research Clinical Gastroenterology* 30:145–159.
- Kim JK, Kolodziejczyk AA, Ilicic T, Teichmann SA, Marioni JC (2015) Characterizing noise structure in single-cell RNA-seq distinguishes genuine from technical stochastic allelic expression. *Nat Commun* 6:8687.

- Kita T et al. (2021) Adrenomedullin for steroid-resistant ulcerative colitis: a randomized, double-blind, placebo-controlled phase-2a clinical trial. *J Gastroenterol* 56:147–157.
- Kita T et al. (2022) Adrenomedullin for biologic-resistant Crohn's disease: A randomized, double-blind, placebo-controlled phase 2a clinical trial. *J of Gastro and Hepatol* 37:2051–2059.
- Kitamura K, Kangawa K, Kawamoto M, Ichiki Y, Nakamura S, Matsuo H, Eto T (1993) Adrenomedullin: A Novel Hypotensive Peptide Isolated from Human Pheochromocytoma. *Biochemical and Biophysical Research Communications* 192:553–560.
- Kitamura K, Kato J, Kawamoto M, Tanaka M, Chino N, Kangawa K, Eto T (1998) The Intermediate Form of Glycine-Extended Adrenomedullin Is the Major Circulating Molecular Form in Human Plasma. *Biochemical and Biophysical Research Communications* 244:551–555.
- Kiyomizu A, Kitamura K, Kawamoto M, Eto T (2001) Distribution and molecular forms of adrenomedullin and proadrenomedullin N-terminal 20 peptide in the porcine gastrointestinal tract. *Journal of Gastroenterology* 36:18–23.
- Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, Peshkin L, Weitz DA, Kirschner MW (2015) Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells. *Cell* 161:1187–1201.
- Kong L, Pokatayev V, Lefkovich A, Carter GT, Creasey EA, Krishna C, Subramanian S, Kocher B, Ashenberg O, Lau H, Ananthakrishnan AN, Graham DB, Deguine J, Xavier RJ (2023) The landscape of immune dysregulation in Crohn's disease revealed through single-cell transcriptomic profiling in the ileum and colon. *Immunity* 56:444-458.e5.
- Kong S, Zhang YH, Zhang W (2018) Regulation of Intestinal Epithelial Cells Properties and Functions by Amino Acids. *BioMed Research International* 2018:1–10.
- Koyama T et al. (2013) Vascular Endothelial Adrenomedullin-RAMP2 System Is Essential for Vascular Integrity and Organ Homeostasis. *Circulation* 127:842–853.
- Kubo A, Minamino N, Isumi Y, Katafuchi T, Kangawa K, Dohi K, Matsuo H (1998) Production of Adrenomedullin in Macrophage Cell Line and Peritoneal Macrophage. *Journal of Biological Chemistry* 273:16730–16738.
- Kuenzig ME et al. (2022) Twenty-first Century Trends in the Global Epidemiology of Pediatric-Onset Inflammatory Bowel Disease: Systematic Review. *Gastroenterology* 162:1147-1159.e4.
- Law CW, Chen Y, Shi W, Smyth GK (2014) voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* 15:R29.
- Linares de la Cal JA, Cantón C, Hermida C, Pérez-Miranda M, Maté-Jiménez J (1999) Estimated incidence of inflammatory bowel disease in Argentina and Panama (1987-1993). *Rev Esp Enferm Dig* 91:277–286. PMID: 10348927.
- Lu H, Suo Z, Lin J, Cong Y, Liu Z (2024) Monocyte-macrophages modulate intestinal homeostasis in inflammatory bowel disease. *Biomark Res* 12:76.
- Luecken MD, Theis FJ (2019) Current best practices in single-cell RNA-seq analysis: a tutorial. *Molecular Systems Biology* 15:e8746.
- Lyal N, Ran D, An L (2020) Normalization Methods on Single-Cell RNA-seq Data: An Empirical Survey. *Front Genet* 11:41.
- MacManus CF, Campbell EL, Keely S, Burgess A, Kominsky DJ, Colgan SP (2011) Anti-inflammatory actions of adrenomedullin through fine tuning of HIF stabilization. *FASEB j* 25:1856–1864.
- Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, Trombetta JJ, Weitz DA, Sanes JR, Shalek AK, Regev A, McCarroll SA (2015) Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* 161:1202–1214.

- Mak WY, Zhao M, Ng SC, Burisch J (2020) The epidemiology of inflammatory bowel disease: East meets west. *J of Gastro and Hepatol* 35:380–389.
- Martínez A (2006) A new family of angiogenic factors. *Cancer Letters* 236:157–163.
- Martínez A, Weaver C, López J, Bhathena SJ, Elsasser TH, Miller MJ, Moody TW, Unsworth EJ, Cuttitta F (1996) Regulation of insulin secretion and blood glucose metabolism by adrenomedullin. *Endocrinology* 137:2626–2632.
- Martínez-Herrero S, Martínez A (2016) Adrenomedullin regulates intestinal physiology and pathophysiology. *Domestic Animal Endocrinology* 56:S66–S83.
- McCarthy DJ, Campbell KR, Lun ATL, Wills QF (2017) Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R Hofacker I, ed. *Bioinformatics* 33:1179–1186.
- McInnes L, Healy J, Melville J (2018) UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. Available at: <https://arxiv.org/abs/1802.03426> [Accessed September 22, 2024].
- Meyrath M, Palmer CB, Reynders N, Vanderplasschen A, Ollert M, Bouvier M, Szpakowska M, Chevigné A (2021) Proadrenomedullin N-Terminal 20 Peptides (PAMPs) Are Agonists of the Chemokine Scavenger Receptor ACKR3/CXCR7. *ACS Pharmacol Transl Sci* 4:813–823.
- Mou T, Deng W, Gu F, Pawitan Y, Vu TN (2020) Reproducibility of Methods to Detect Differentially Expressed Genes from Single-Cell RNA Sequencing. *Front Genet* 10:1331.
- Mukherjee S, Hooper LV (2015) Antimicrobial Defense of the Intestine. *Immunity* 42:28–39.
- Mulder H, Ahrén B, Karlsson S, Sundler F (1996) Adrenomedullin: localization in the gastrointestinal tract and effects on insulin secretion. *Regulatory Peptides* 62:107–112.
- Nam D, Kim S-Y (2008) Gene-set approach for expression pattern analysis. *Briefings in Bioinformatics* 9:189–197.
- Nayak R, Hasija Y (2021) A hitchhiker’s guide to single-cell transcriptomics and data analysis pipelines. *Genomics* 113:606–619.
- Ng SC, Shi HY, Hamidi N, Underwood FE, Tang W, Benchimol EI, Panaccione R, Ghosh S, Wu JCY, Chan FKL, Sung JJY, Kaplan GG (2017) Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *The Lancet* 390:2769–2778.
- Oksche A, Boese G, Horstmeyer A, Papsdorf G, Furkert J, Beyermann M, Bienert M, Rosenthal W (2000) Evidence for Downregulation of the Endothelin-B-Receptor By the Use of Fluorescent Endothelin-1 and a Fusion Protein Consisting of the Endothelin-B-Receptor and the Green Fluorescent Protein: *Journal of Cardiovascular Pharmacology* 36:S44–S47.
- Parekh S, Ziegenhain C, Vieth B, Enard W, Hellmann I (2018) zUMIs - A fast and flexible pipeline to process RNA sequencing data with UMIs. *GigaScience* 7 Available at: <https://academic.oup.com/gigascience/article/doi/10.1093/gigascience/giy059/5005022> [Accessed September 22, 2024].
- Pearson K (1901) LIII. *On lines and planes of closest fit to systems of points in space*. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2:559–572.
- Peyrin-Biroulet L, Christopher R, Behan D, Lassen C (2017) Modulation of sphingosine-1-phosphate in inflammatory bowel disease. *Autoimmunity Reviews* 16:495–503.
- Peyrin-Biroulet L, Loftus EV, Colombel J-F, Sandborn WJ (2010) The Natural History of Adult Crohn’s Disease in Population-Based Cohorts. *American Journal of Gastroenterology* 105:289–297.
- Pierre N, Salée C, Vieujean S, Bequet E, Merli A, Siegmund B, Meuwis M, Louis E (2021) Review article: distinctions between ileal and colonic Crohn’s disease: from physiology to pathology. *Aliment Pharmacol Ther* 54:779–791.

- Pober JS, Sessa WC (2007) Evolving functions of endothelial cells in inflammation. *Nat Rev Immunol* 7:803–815.
- Poyner DR (2002) International Union of Pharmacology. XXXII. The Mammalian Calcitonin Gene-Related Peptides, Adrenomedullin, Amylin, and Calcitonin Receptors. *Pharmacological Reviews* 54:233–246.
- Priebe S, Kreisel C, Horn F, Guthke R, Linde J (2015) FungiFun2: a comprehensive online resource for systematic analysis of gene lists from fungal species. *Bioinformatics* 31:445–446.
- Qi T, Christopoulos G, Bailey RJ, Christopoulos A, Sexton PM, Hay DL (2008) Identification of N-Terminal Receptor Activity-Modifying Protein Residues Important for Calcitonin Gene-Related Peptide, Adrenomedullin, and Amylin Receptor Function. *Mol Pharmacol* 74:1059–1071.
- Ra YE, Bang Y-J (2024) Balancing Act of the Intestinal Antimicrobial Proteins on Gut Microbiota and Health. *J Microbiol* 62:167–179.
- Raya Tonetti F, Eguileor A, Llorente C (2024) Goblet cells: guardians of gut immunity and their role in gastrointestinal diseases. *egastro* 2:e100098.
- Raza K (2024) Preprocessing and Quality Control. In: Machine Learning in Single-Cell RNA-seq Data Analysis, pp 17–30 SpringerBriefs in Applied Sciences and Technology. Singapore: Springer Nature Singapore. Available at: https://link.springer.com/10.1007/978-981-97-6703-8_2 [Accessed September 27, 2024].
- Roda G, Chien Ng S, Kotze PG, Argollo M, Panaccione R, Spinelli A, Kaser A, Peyrin-Biroulet L, Danese S (2020) Crohn's disease. *Nat Rev Dis Primers* 6:22.
- Rosen MJ, Dhawan A, Saeed SA (2015) Inflammatory Bowel Disease in Children and Adolescents. *JAMA Pediatr* 169:1053.
- Rossowski WJ, Jiang N-Y, Coy DH (1997) Adrenomedullin, amylin, calcitonin gene-related peptide and their fragments are potent inhibitors of gastric acid secretion in rats. *European Journal of Pharmacology* 336:51–63.
- Satija R, Farrell JA, Gennert D, Schier AF, Regev A (2015) Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol* 33:495–502.
- Satsangi J (2006) The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. *Gut* 55:749–753.
- Sauer CG, Kugathasan S (2010) Pediatric Inflammatory Bowel Disease: Highlighting Pediatric Differences in IBD. *Medical Clinics of North America* 94:35–52.
- Saunders A, Macosko EZ, Wysoker A, Goldman M, Krienen FM, De Rivera H, Bien E, Baum M, Bortolin L, Wang S, Goeva A, Nemesh J, Kamitaki N, Brumbaugh S, Kulp D, McCarroll SA (2018) Molecular Diversity and Specializations among the Cells of the Adult Mouse Brain. *Cell* 174:1015–1030.e16.
- Scholz CJ et al. (2018) FASTGenomics: An analytical ecosystem for single-cell RNA sequencing data. Available at: <http://biorxiv.org/lookup/doi/10.1101/272476> [Accessed September 22, 2024].
- Scialdone A, Natarajan KN, Saraiva LR, Proserpio V, Teichmann SA, Stegle O, Marioni JC, Buettner F (2015) Computational assignment of cell-cycle stage from single-cell transcriptome data. *Methods* 85:54–61.
- Serafin DS, Harris NR, Nielsen NR, Mackie DI, Caron KM (2020) Dawn of a New RAMPage. *Trends in Pharmacological Sciences* 41:249–265.
- Sigmund EC, Bauer A, Jakobs BD, Tatliadim H, Tacconi C, Thelen M, Legler DF, Halin C (2023) Reassessing the adrenomedullin scavenging function of ACKR3 in lymphatic endothelial cells Lupo G, ed. *PLoS ONE* 18:e0285597.
- Sjölund K, Sandén G, Håkanson R, Sundler F (1983) Endocrine Cells in Human Intestine: An Immunocytochemical Study. *Gastroenterology* 85:1120–1130.

- Soneson C, Robinson MD (2018) Bias, robustness and scalability in single-cell differential expression analysis. *Nat Methods* 15:255–261.
- Spoto S, Basili S, Cangemi R, Yuste JR, Lucena F, Romiti GF, Raparelli V, Argemi J, D'Avanzo G, Locorriere L, Masini F, Calarco R, Testorio G, Spiezio S, Ciccozzi M, Angeletti S (2024) A Focus on the Pathophysiology of Adrenomedullin Expression: Endothelitis and Organ Damage in Severe Viral and Bacterial Infections. *Cells* 13:892.
- Stark R, Grzelak M, Hadfield J (2019) RNA sequencing: the teenage years. *Nat Rev Genet* 20:631–656.
- Stegle O, Teichmann SA, Marioni JC (2015) Computational and analytical challenges in single-cell transcriptomics. *Nat Rev Genet* 16:133–145.
- Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, Hao Y, Stoeckius M, Smibert P, Satija R (2019) Comprehensive Integration of Single-Cell Data. *Cell* 177:1888–1902.e21.
- Su K, Yu T, Wu H (2021) Accurate feature selection improves single-cell RNA-seq cell clustering. *Briefings in Bioinformatics* 22:bbab034.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 102:15545–15550.
- Sýkora J, Pomahačová R, Kreslová M, Cvalínová D, Štých P, Schwarz J (2018) Current global trends in the incidence of pediatric-onset inflammatory bowel disease. *WJG* 24:2741–2763.
- Takei Y, Inoue K, Ogoshi M, Kawahara T, Bannai H, Miyano S (2004) Identification of novel adrenomedullin in mammals: a potent cardiovascular and renal regulator. *FEBS Letters* 556:53–58.
- Talero E, Sánchez-Fidalgo S, De La Lastra CA, Illanes M, Calvo JR, Motilva V (2008) Acute and chronic responses associated with adrenomedullin administration in experimental colitis. *Peptides* 29:2001–2012.
- Tanaka M, Koyama T, Sakurai T, Kamiyoshi A, Ichikawa-Shindo Y, Kawate H, Liu T, Xian X, Imai A, Zhai L, Hirabayashi K, Owa S, Yamauchi A, Igarashi K, Taniguchi S, Shindo T (2016) The endothelial adrenomedullin-RAMP2 system regulates vascular integrity and suppresses tumour metastasis. *Cardiovasc Res* 111:398–409.
- Temmesfeld-Wollbrück B, Hocke A, Suttorp N, Hippenstiel S (2007) Adrenomedullin and endothelial barrier function. *Thromb Haemost* 98:944–951.
- Thia KT, Sandborn WJ, Harmsen WS, Zinsmeister AR, Loftus EV (2010) Risk Factors Associated With Progression to Intestinal Complications of Crohn's Disease in a Population-Based Cohort. *Gastroenterology* 139:1147–1155.
- Tomoda Y, Isumi Y, Katafuchi T, Minamino N (2001) Regulation of adrenomedullin secretion from cultured cells. *Peptides* 22:1783–1794.
- Torres J, Mehandru S, Colombel J-F, Peyrin-Biroulet L (2017) Crohn's disease. *The Lancet* 389:1741–1755.
- Troughton RW, Frampton CM, Lewis LK, Yandle TG, Richards AM, Nicholls MG (2001) Differing thresholds for modulatory effects of adrenomedullin infusion on haemodynamic and hormone responses to angiotensin II and adrenocorticotropic hormone in healthy volunteers. *Clinical Science* 101:103–109.
- Van Den Berge K, Perraudeau F, Soneson C, Love MI, Rissó D, Vert J-P, Robinson MD, Dudoit S, Clement L (2018) Observation weights unlock bulk RNA-seq tools for zero inflation and single-cell applications. *Genome Biol* 19:24.
- Vieth B, Ziegenhain C, Parekh S, Enard W, Hellmann I (2017) powsimR: power analysis for bulk and single cell RNA-seq experiments Hofacker I, ed. *Bioinformatics* 33:3486–3488.

- Wagner A, Regev A, Yosef N (2016) Revealing the vectors of cellular identity with single-cell genomics. *Nat Biotechnol* 34:1145–1160.
- Waltman L, Van Eck NJ (2013) A smart local moving algorithm for large-scale modularity-based community detection. *Eur Phys J B* 86:471.
- Weber LM, Robinson MD (2016) Comparison of clustering methods for high-dimensional single-cell flow and mass cytometry data. *Cytometry Pt A* 89:1084–1096.
- Weijun Luo (2017) pathview. Available at: <https://bioconductor.org/packages/pathview> [Accessed September 22, 2024].
- Weinreb C, Wolock S, Klein AM (2018) SPRING: a kinetic interface for visualizing high dimensional single-cell expression data Berger B, ed. *Bioinformatics* 34:1246–1248.
- Wickham H (2011) ggplot2. *WIREs Computational Stats* 3:180–185.
- Wimalawansa SJ (1997) Amylin, Calcitonin Gene-Related Peptide, Calcitonin, and Adrenomedullin: A Peptide Superfamily. *Crit Rev Neurobiol* 11:167–239.
- Wolf FA, Angerer P, Theis FJ (2018) SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* 19:15.
- Wolf FA, Hamey FK, Plass M, Solana J, Dahlin JS, Göttgens B, Rajewsky N, Simon L, Theis FJ (2019) PAGA: graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol* 20:59.
- Wong LYF, Cheung BMY, Li Y-Y, Tang F (2005) Adrenomedullin Is Both Proinflammatory and Antiinflammatory: Its Effects on Gene Expression and Secretion of Cytokines and Macrophage Migration Inhibitory Factor in NR8383 Macrophage Cell Line. *Endocrinology* 146:1321–1327.
- Wu R, Zhou M, Wang P (2003) Adrenomedullin and adrenomedullin binding protein-1 downregulate TNF- α in macrophage cell line and rat Kupffer cells. *Regulatory Peptides* 112:19–26.
- Xiang R, Wang W, Yang L, Wang S, Xu C, Chen X (2021) A Comparison for Dimensionality Reduction Methods of Single-Cell RNA-seq Data. *Front Genet* 12:646936.
- Xue J, Zhou X, Yang J, Niu A (2024) Comparative study on differential expression analysis methods for single-cell RNA sequencing data with small biological replicates: Based on single-cell transcriptional data of PBMCs from COVID-19 severe patients Karakulah AS, ed. *PLoS ONE* 19:e0299358.
- Yang P, Huang H, Liu C (2021) Feature selection revisited in the single-cell era. *Genome Biol* 22:321.
- Yi X, Du Z, Su Z (2013) PlantGSEA: a gene set enrichment analysis toolkit for plant community. *Nucleic Acids Research* 41:W98–W103.
- Yi Z, Fan H, Liu X, Tang Q, Zuo D, Yang J (2015) Adrenomedullin improves intestinal epithelial barrier function by downregulating myosin light chain phosphorylation in ulcerative colitis rats. *Molecular Medicine Reports* 12:3615–3620.
- You Y, Tian L, Su S, Dong X, Jabbari JS, Hickey PF, Ritchie ME (2021) Benchmarking UMI-based single-cell RNA-seq preprocessing workflows. *Genome Biol* 22:339.
- Yu G, Wang L-G, Han Y, He Q-Y (2012) clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology* 16:284–287.
- Zeisel A et al. (2018) Molecular Architecture of the Mouse Nervous System. *Cell* 174:999–1014.e22.
- Zhang Z, Cui F, Wang C, Zhao L, Zou Q (2021) Goals and approaches for each processing step for single-cell RNA sequencing data. *Briefings in Bioinformatics* 22:bbaa314.
- Zhao J et al. (2013) First Prospective, Population-Based Inflammatory Bowel Disease Incidence Study in Mainland of China: The Emergence of “Western” Disease. *Inflammatory Bowel Diseases*:1.

- Zheng GXY et al. (2017) Massively parallel digital transcriptional profiling of single cells. *Nat Commun* 8:14049.
- Zhou B, Jin W (2020) Visualization of Single Cell RNA-Seq Data Using t-SNE in R. In: *Stem Cell Transcriptional Networks* (Kidder BL, ed), pp 159–167 *Methods in Molecular Biology*. New York, NY: Springer US. Available at: http://link.springer.com/10.1007/978-1-0716-0301-7_8 [Accessed September 22, 2024].
- Zhou M, Chaudry IH, Wang P (2001) The small intestine is an important source of adrenomedullin release during polymicrobial sepsis. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 281:R654–R660.
- Zudaire E, Portal-Núñez S, Cuttitta F (2006) The central role of adrenomedullin in host defense. *Journal of Leukocyte Biology* 80:237–244.

6. APPENDIX

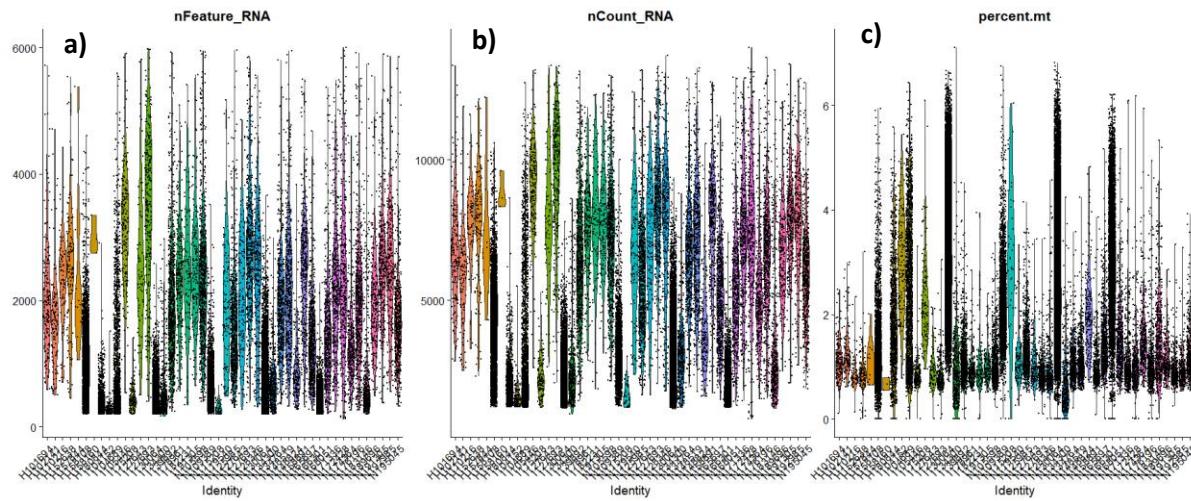


Figure S1. Quality control metrics for terminal ileum stromal cells. The three box plots show (a) the number of detected features per cell (nFeature_RNA), (b) the total RNA counts per cell (nCount_RNA), and (c) the percentage of mitochondrial gene expression (percent.mt) across different cell types.

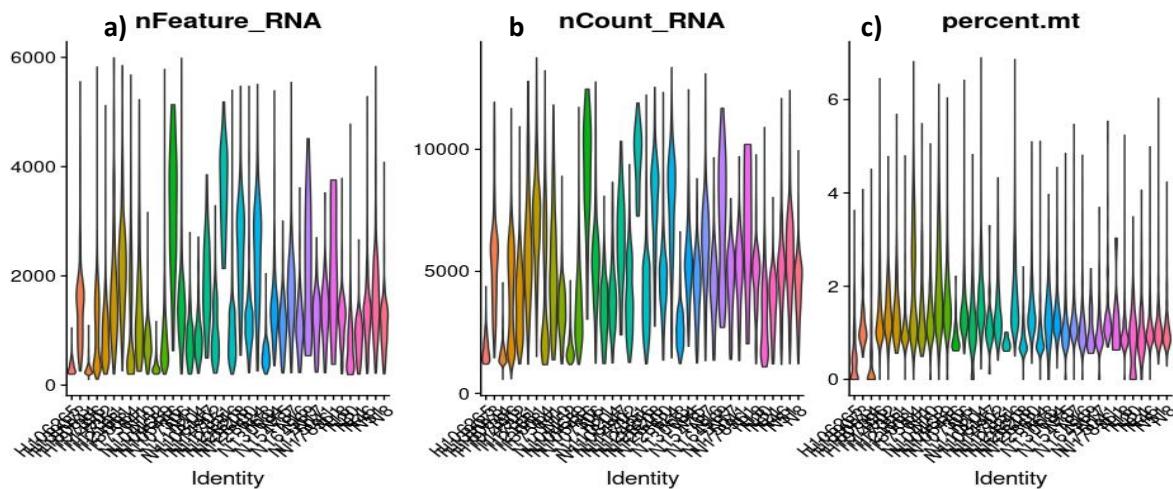


Figure S2. Quality control metrics for colon stromal cells. The three box plots show (a) the number of detected features per cell (nFeature_RNA), (b) the total RNA counts per cell (nCount_RNA), and (c) the percentage of mitochondrial gene expression (percent.mt) across different cell types.

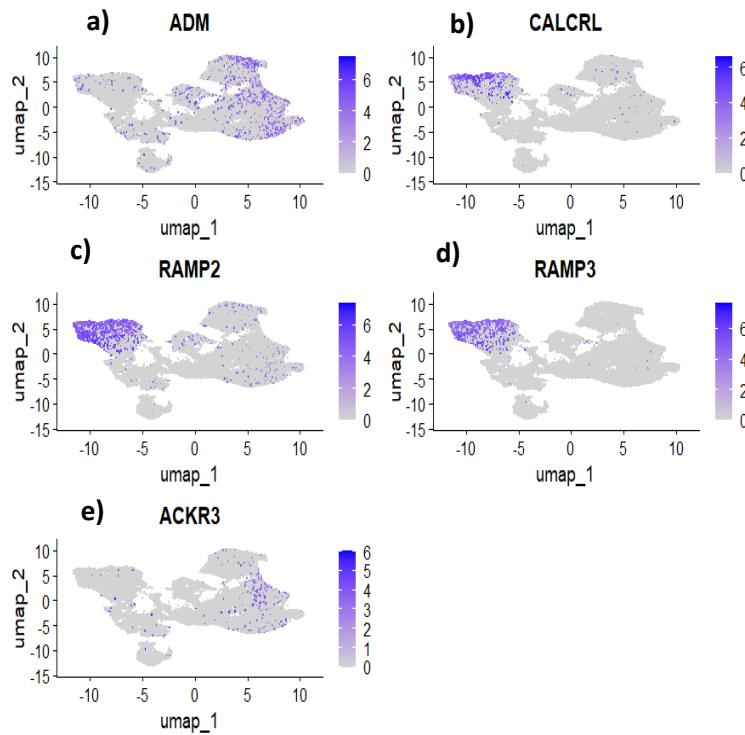
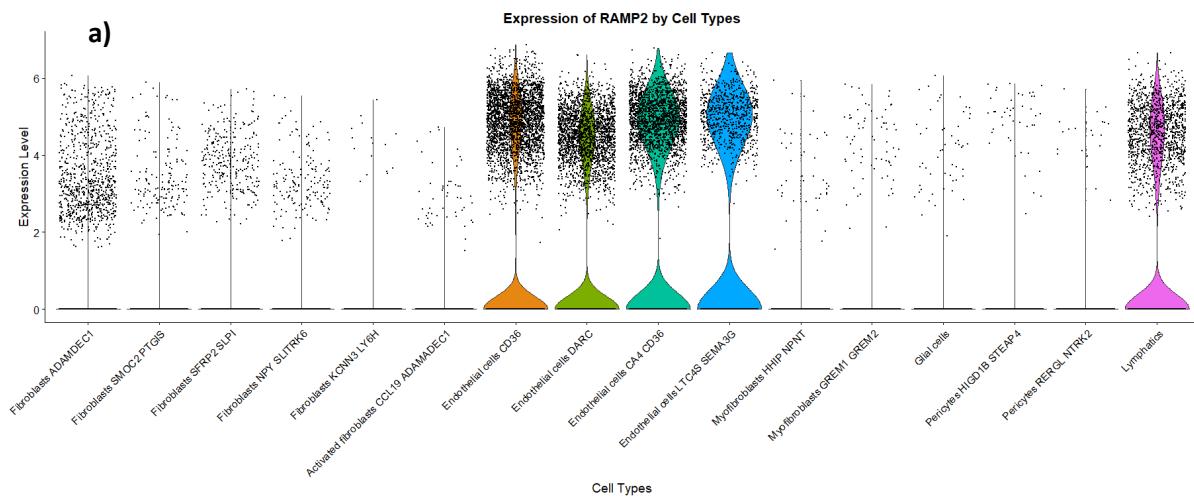


Figure S3. Feature plot showing the expression of marker genes among the colon stromal cell population. Subplots represent expression levels of (a) ADM, (b) CALCRL, (c) RAMP2, (d) RAMP3, and (e) ACKR3. Each subplot shows UMAP projections with colour intensity reflecting the expression level of each gene, with darker shades indicating higher expression.



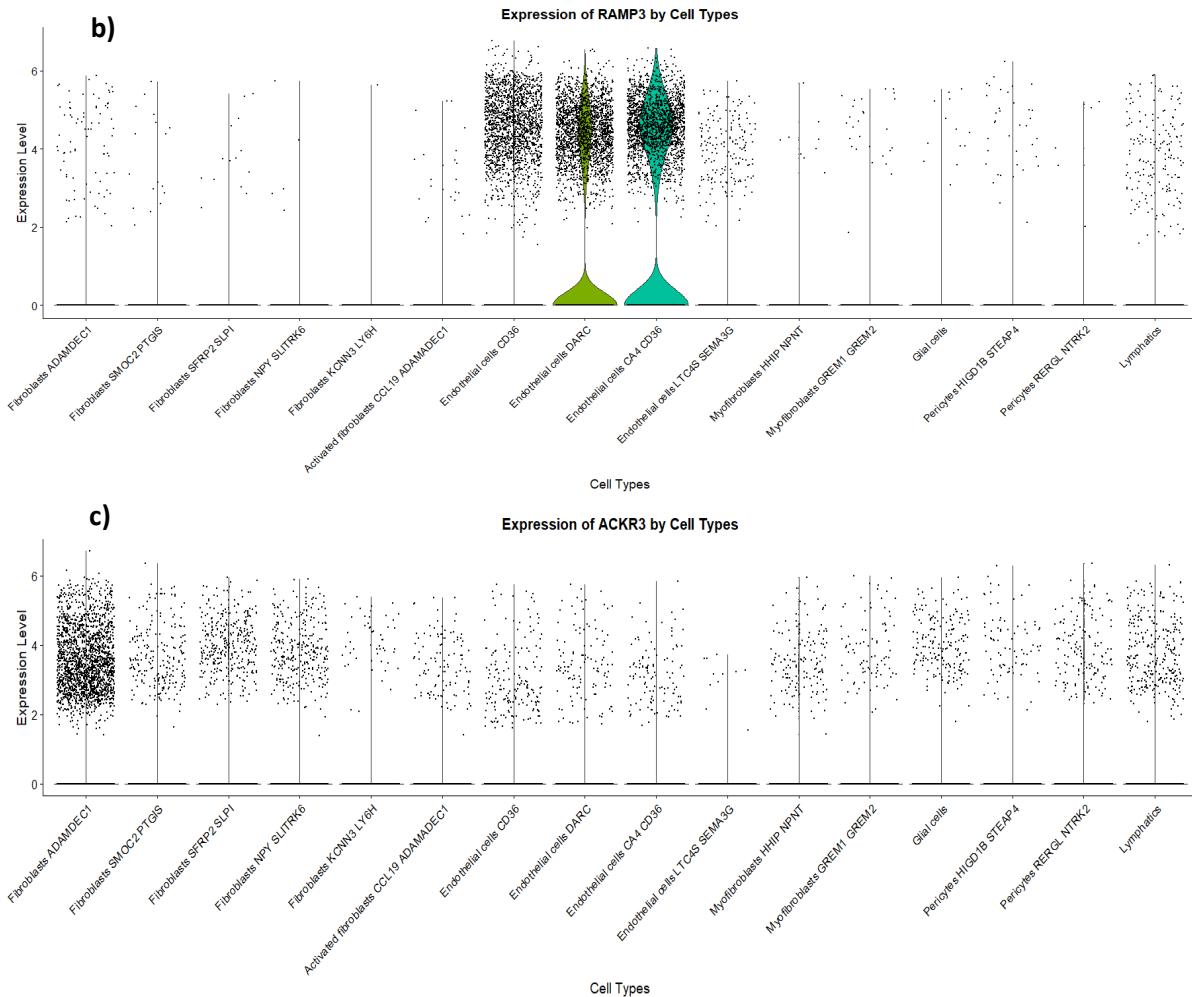
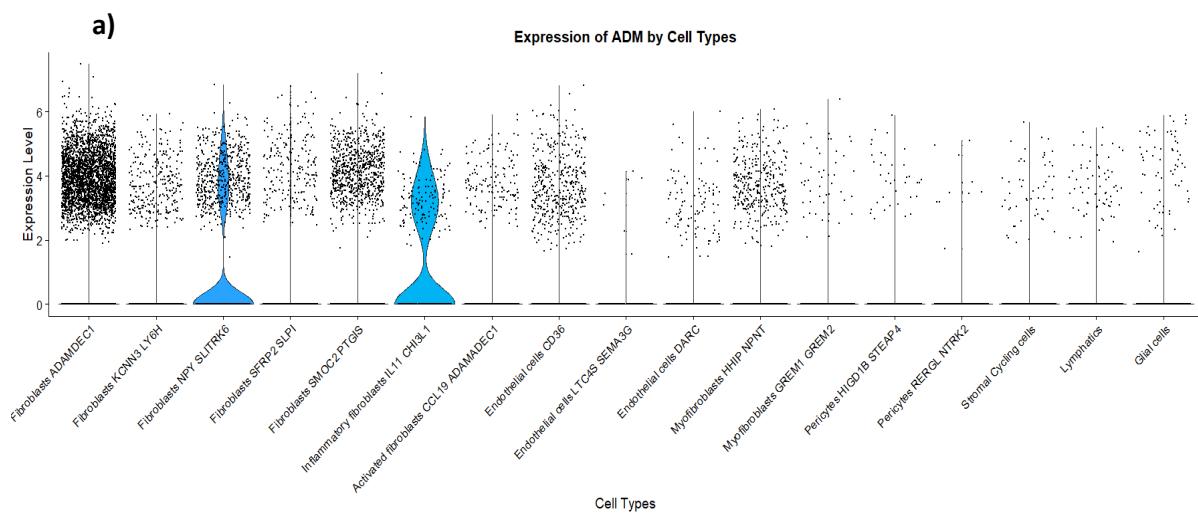
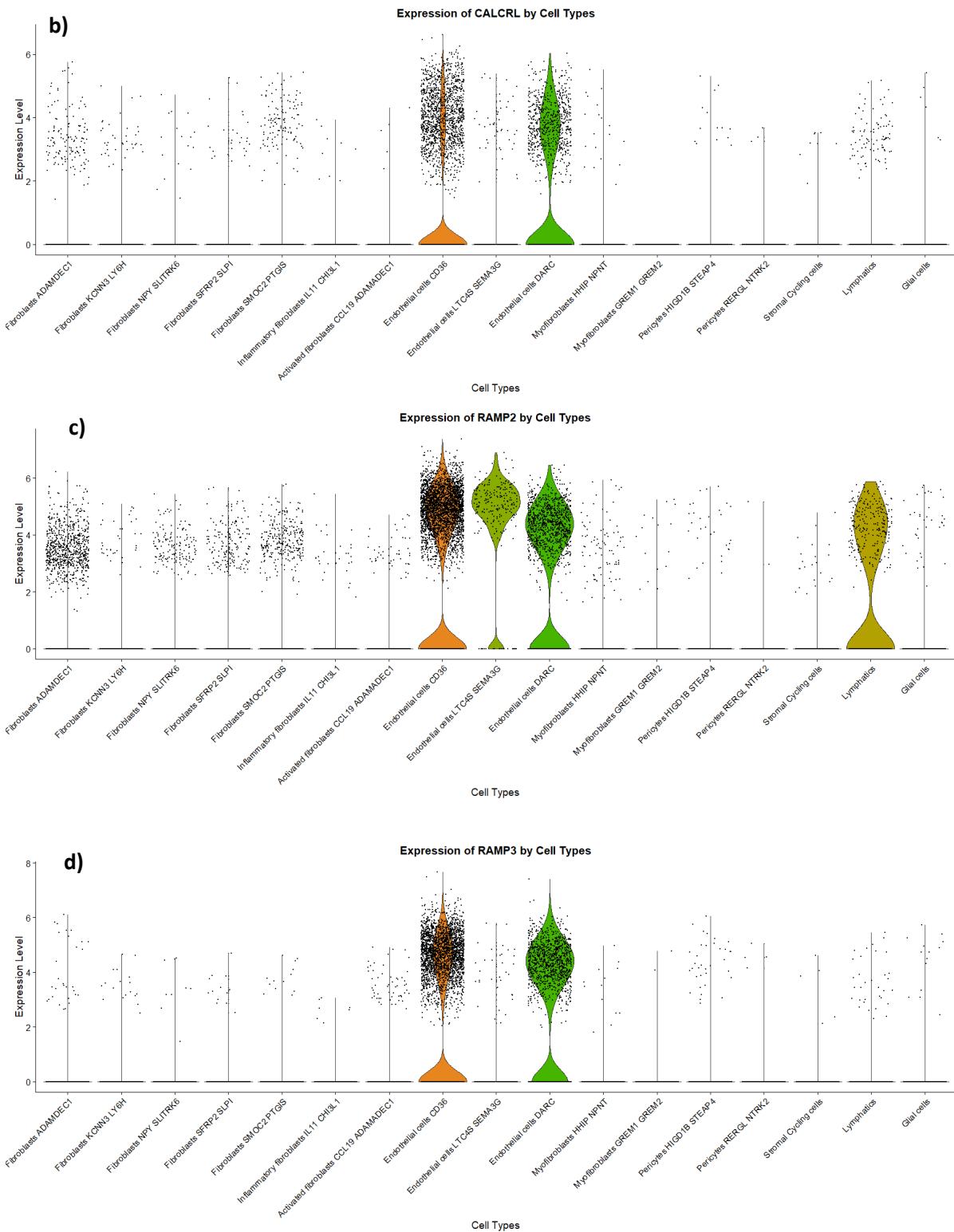


Figure S4 Terminal ileum stromal cells. Violin plots showing the expression marker genes across various cell types. a) RAMP2 b) RAMP3 c) ACKR3





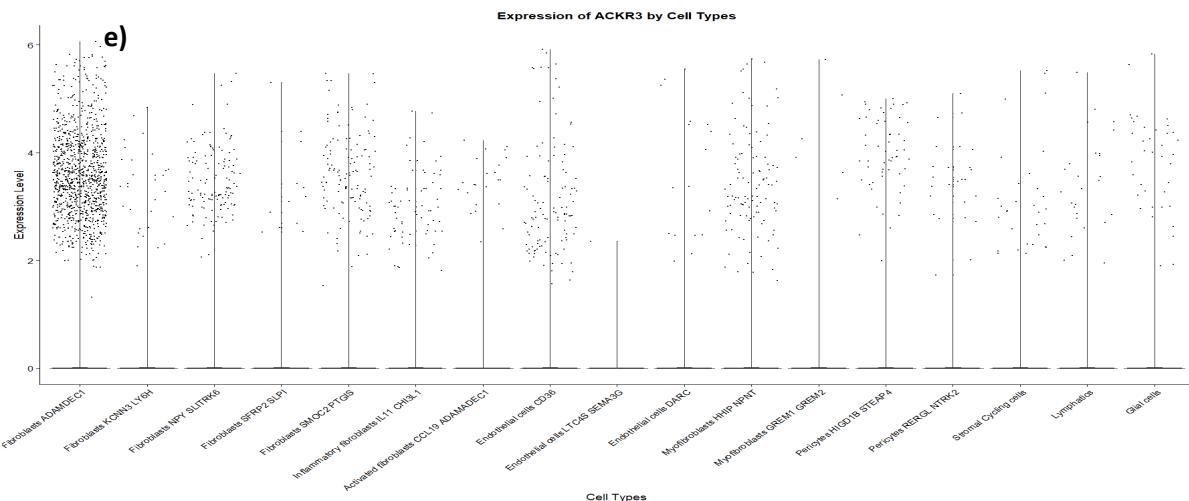


Figure S5. Violin plots showing the expression marker genes in colon stromal cells across various cell types a) ADM b) CALCRL c) RAMP2 d) RAMP3 e) ACKR3

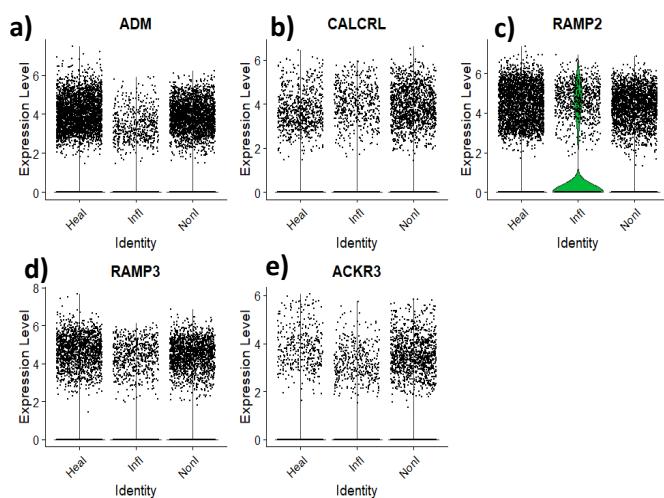
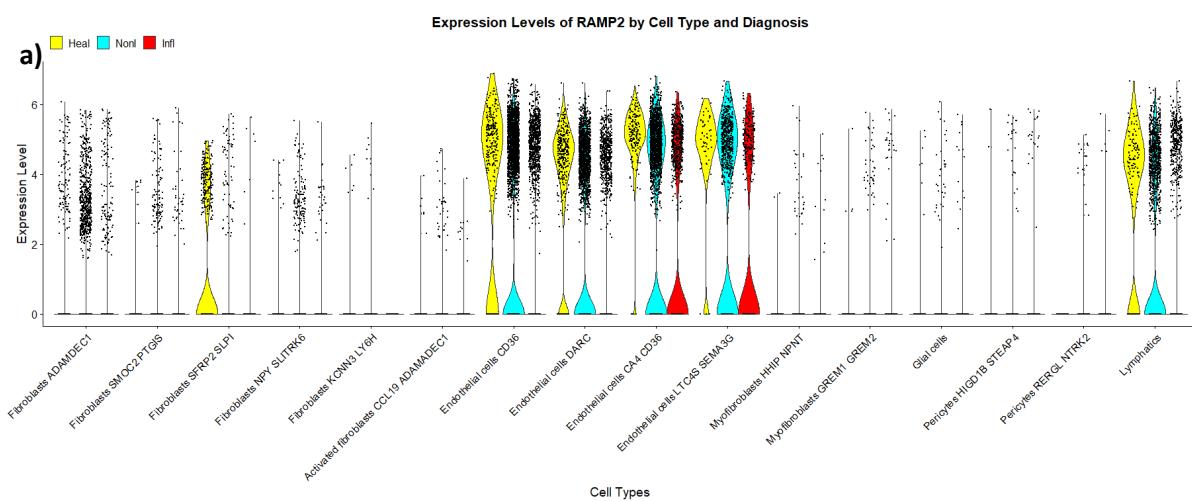


Figure S6. Violin plot representing gene expression across different conditions (Healthy, Non-Inflamed, Inflamed) in Colon stromal cells. a) ADM b) CALCRL c) RAMP2 d) RAMP3 e) ACKR3



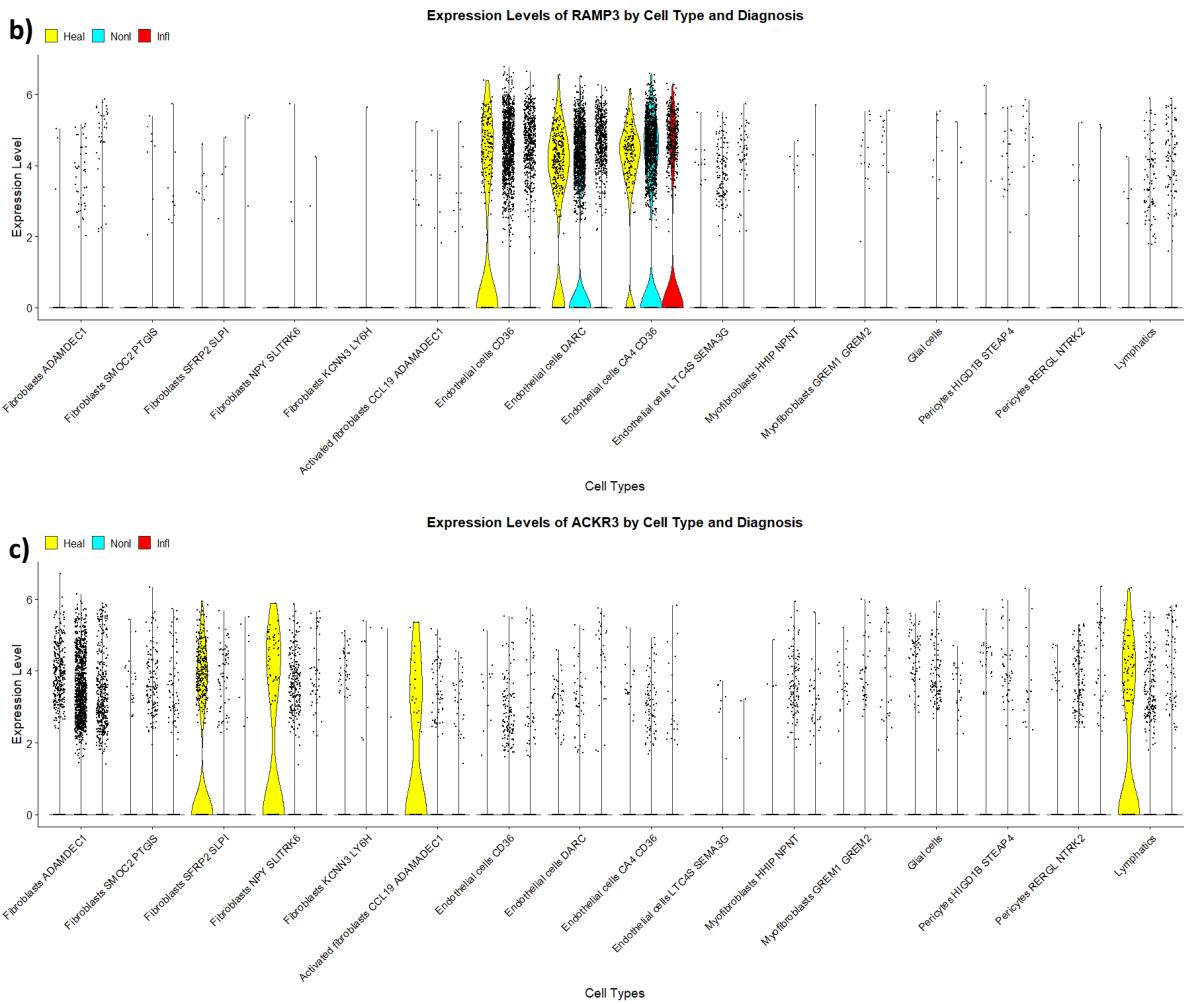
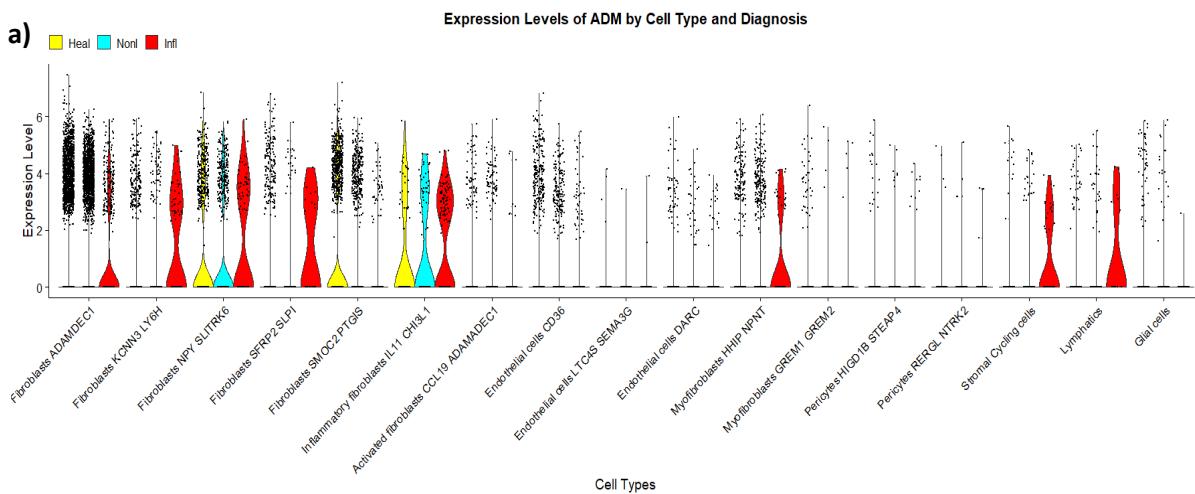
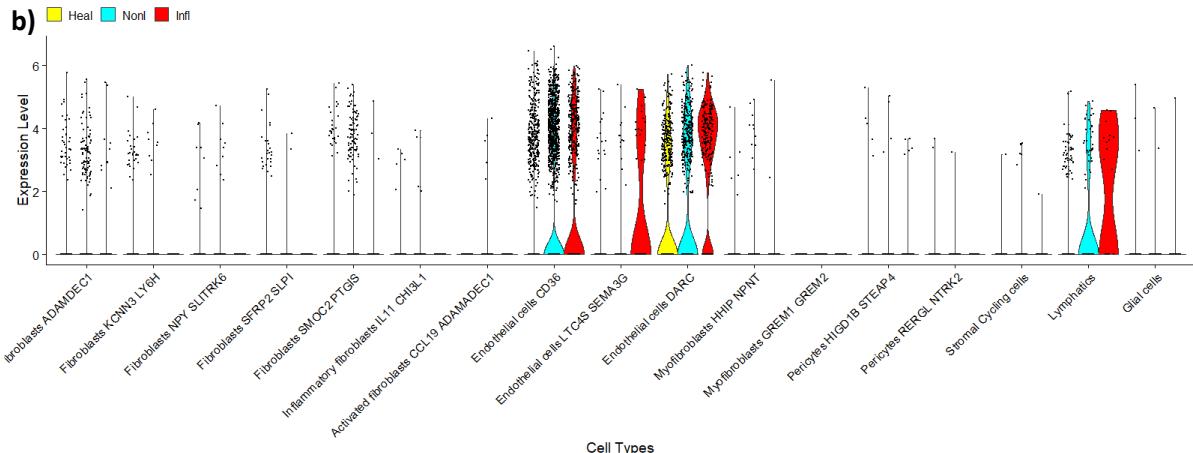


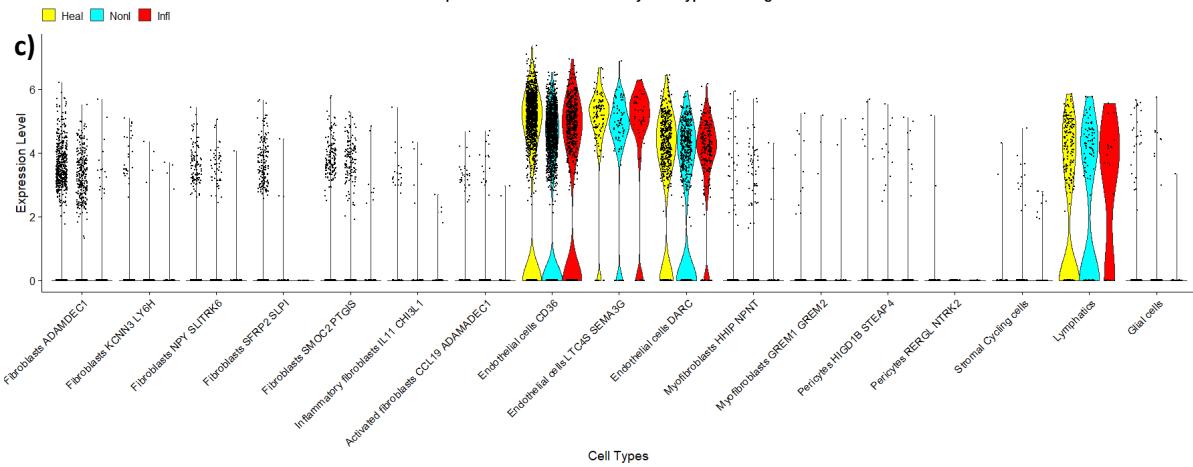
Figure S7. Terminal ileum stromal cells. Violin plots showing the expression levels of marker across various cell types and conditions (Healthy -yellow, Non-inflamed -cyan, Inflamed – red). a) RAMP2 b) RAMP3 c) ACKR3



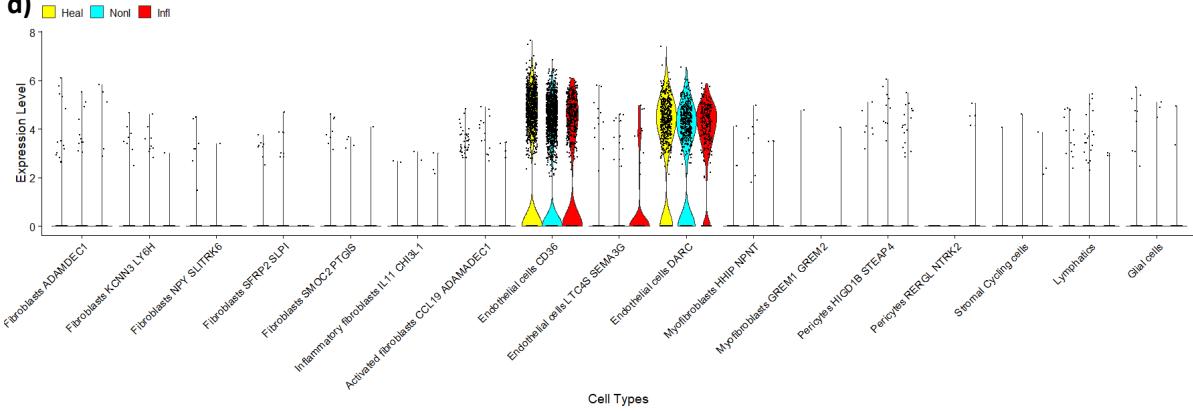
Expression Levels of CALCRL by Cell Type and Diagnosis



Expression Levels of RAMP2 by Cell Type and Diagnosis



Expression Levels of RAMP3 by Cell Type and Diagnosis



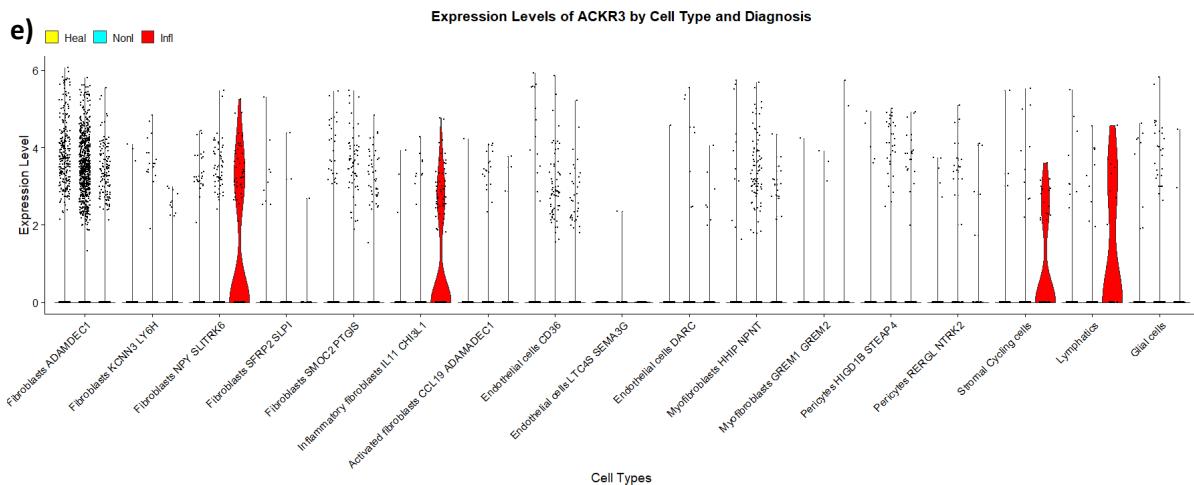
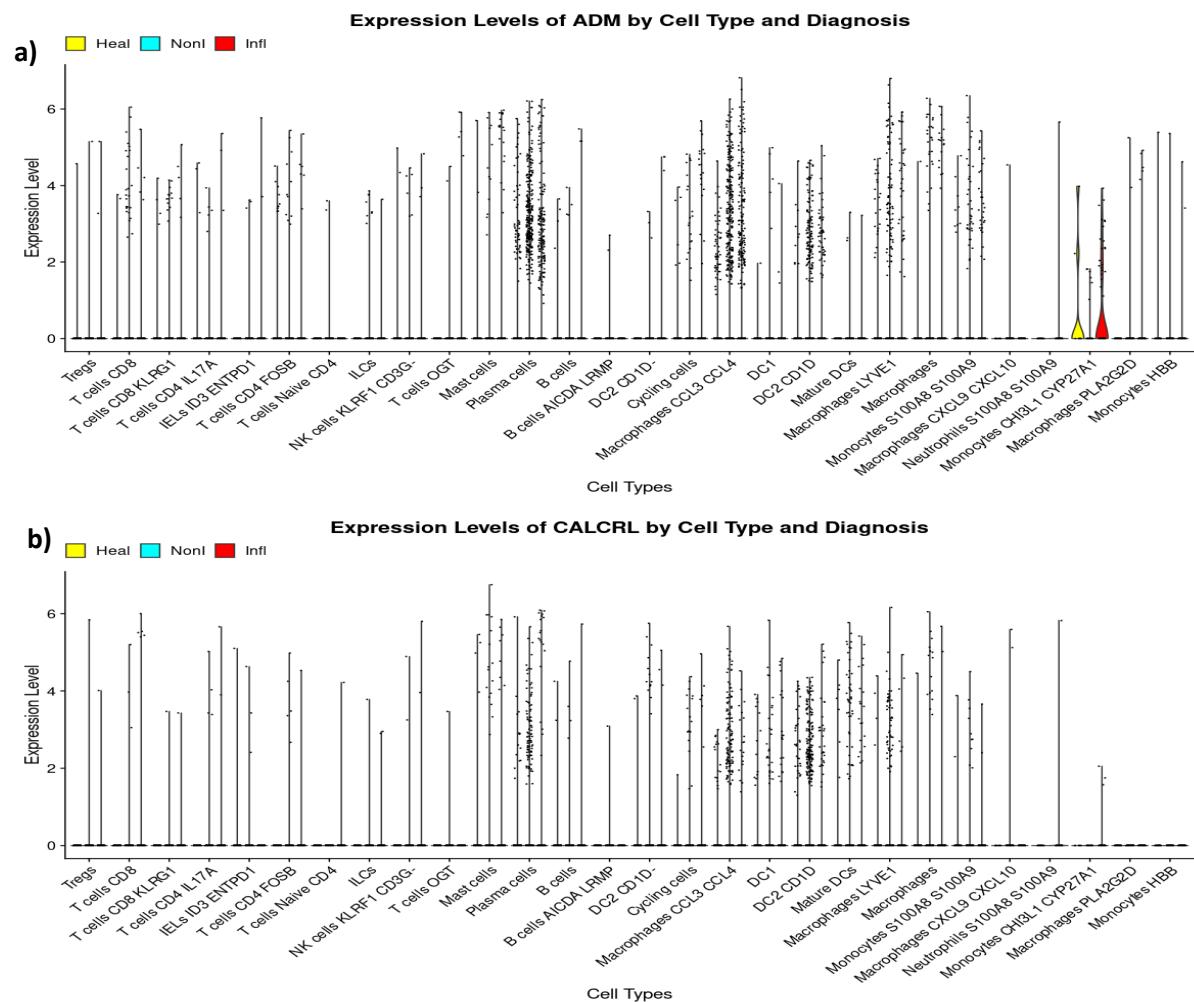


Figure S8. Colon stromal cells. Violin plots showing the expression levels of marker across various cell types and conditions (Healthy -yellow, Non-inflamed -cyan, Inflamed – red). a) ADM b) CALCRL c) RAMP2 d) RAMP3 e) ACKR3



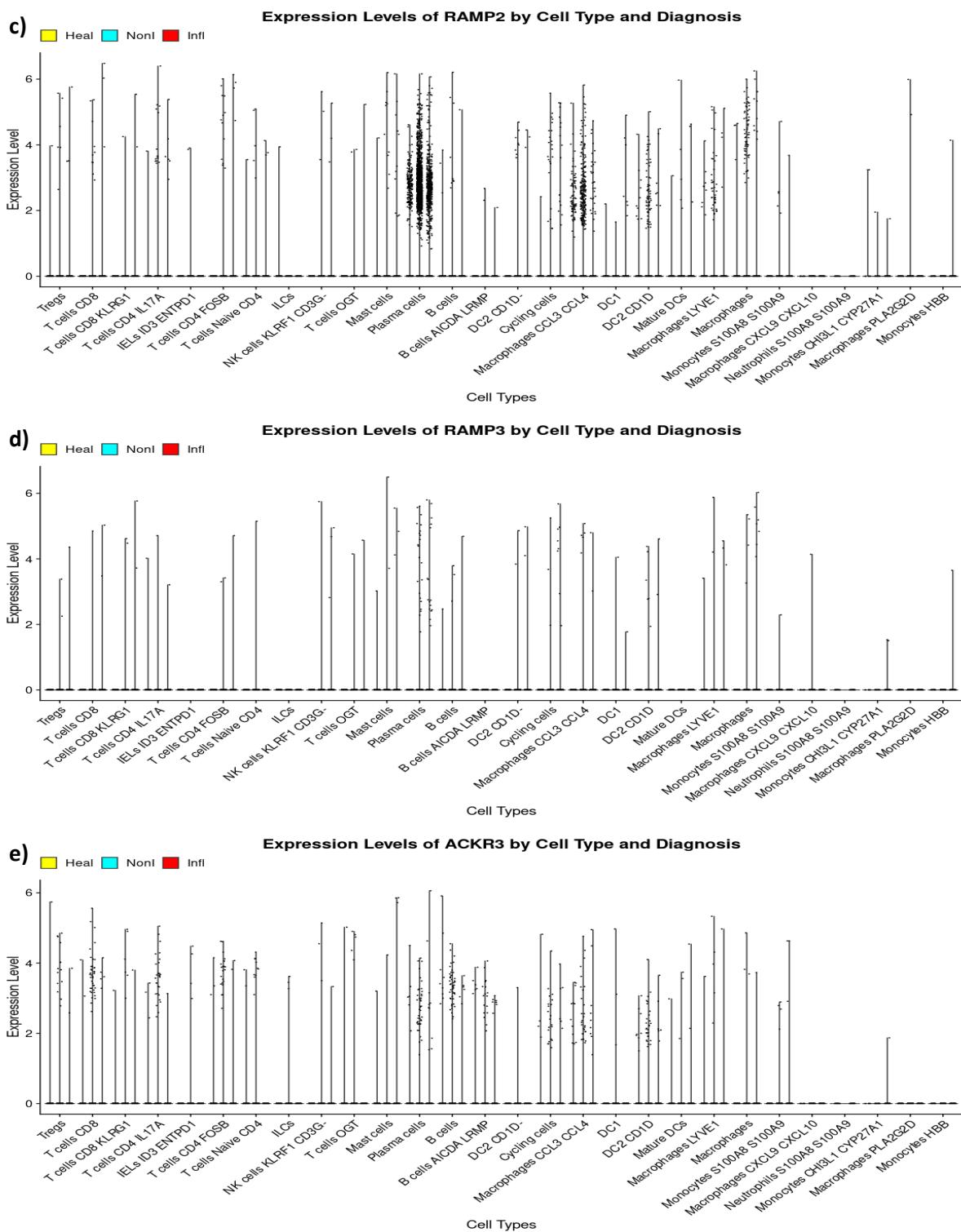
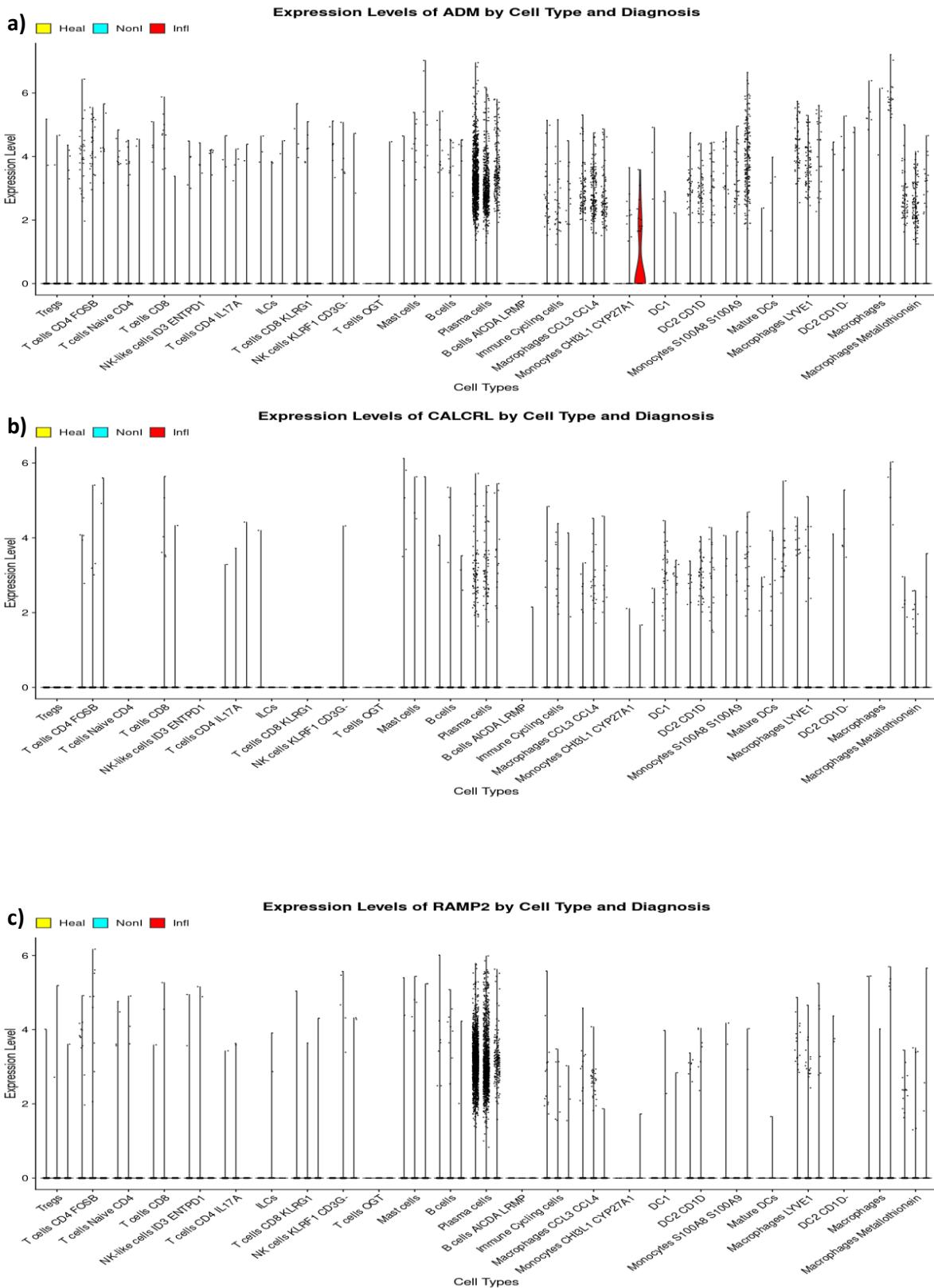


Figure S9. Terminal ileum immune cells. Violin plots showing the expression levels of marker across various cell types and conditions (Healthy -yellow, Non-inflamed -cyan, Inflamed – red) a) ADM b) CALCRL c) RAMP2 d) RAMP3 e) ACKR3



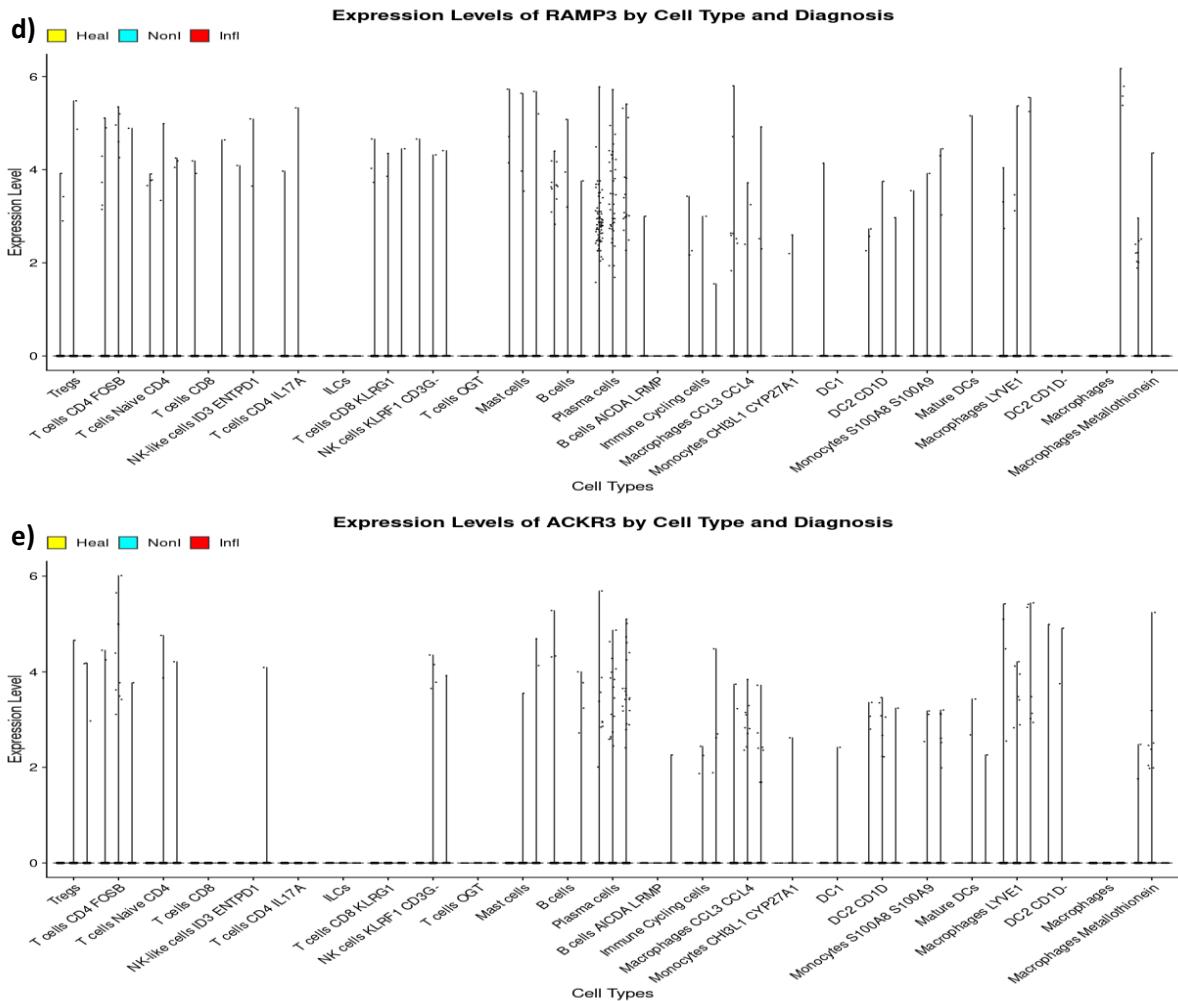
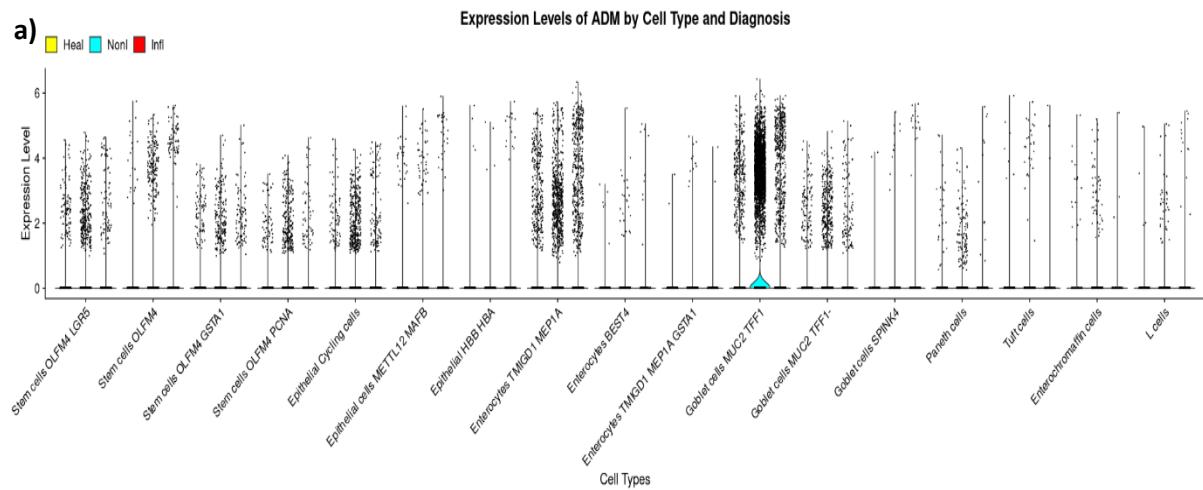
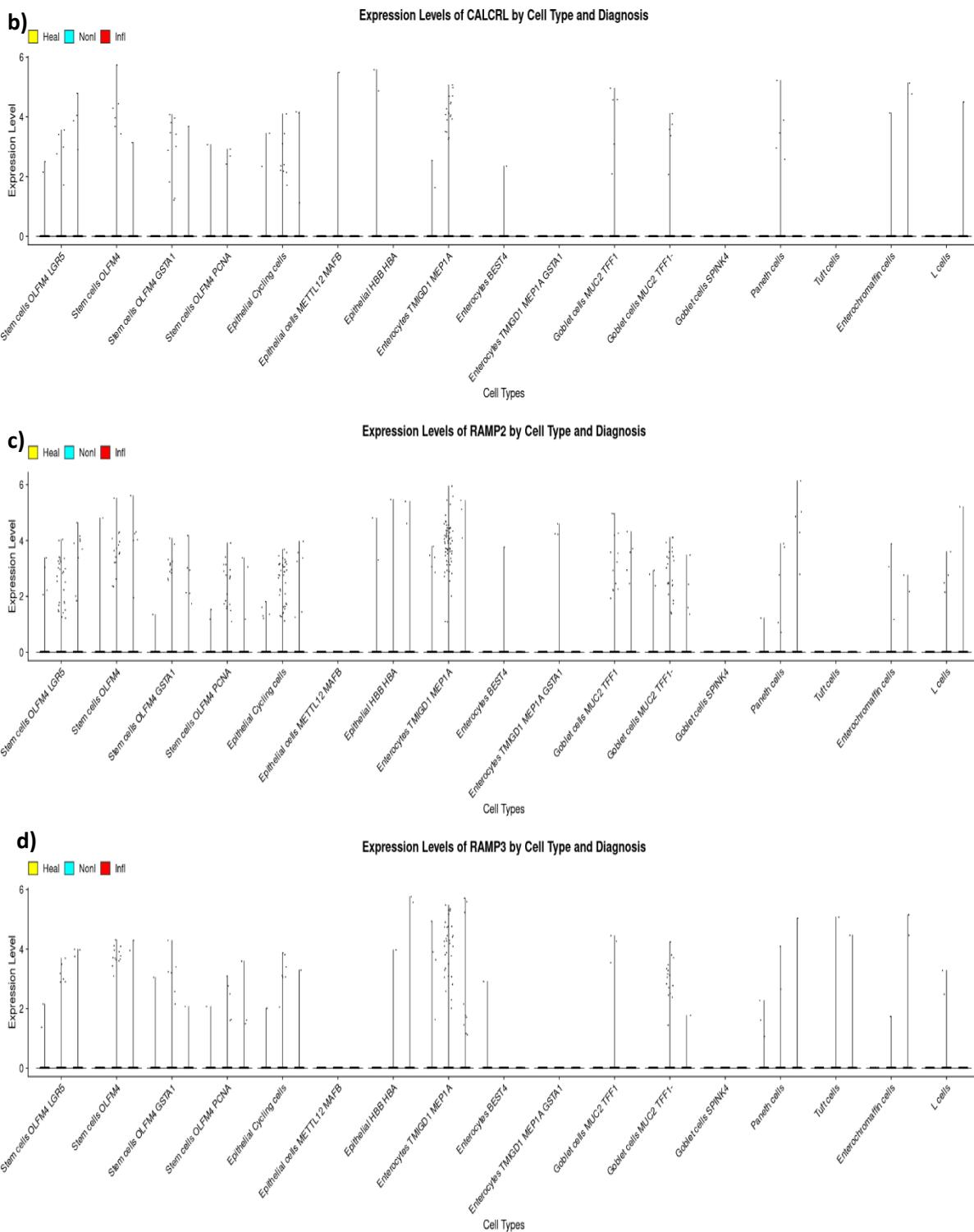


Figure S10. Colon immune cells. Violin plots showing the expression levels of marker across various cell types and conditions (Healthy -yellow, Non-inflamed -cyan, Inflamed – red). a) ADM b) CALCRL c) RAMP2 d) RAMP3 e) ACKR3





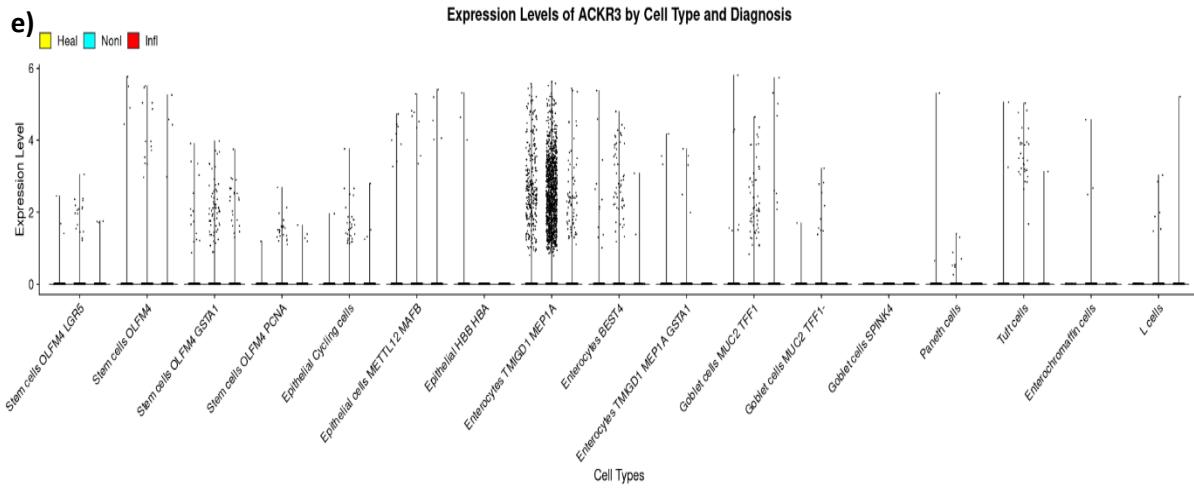
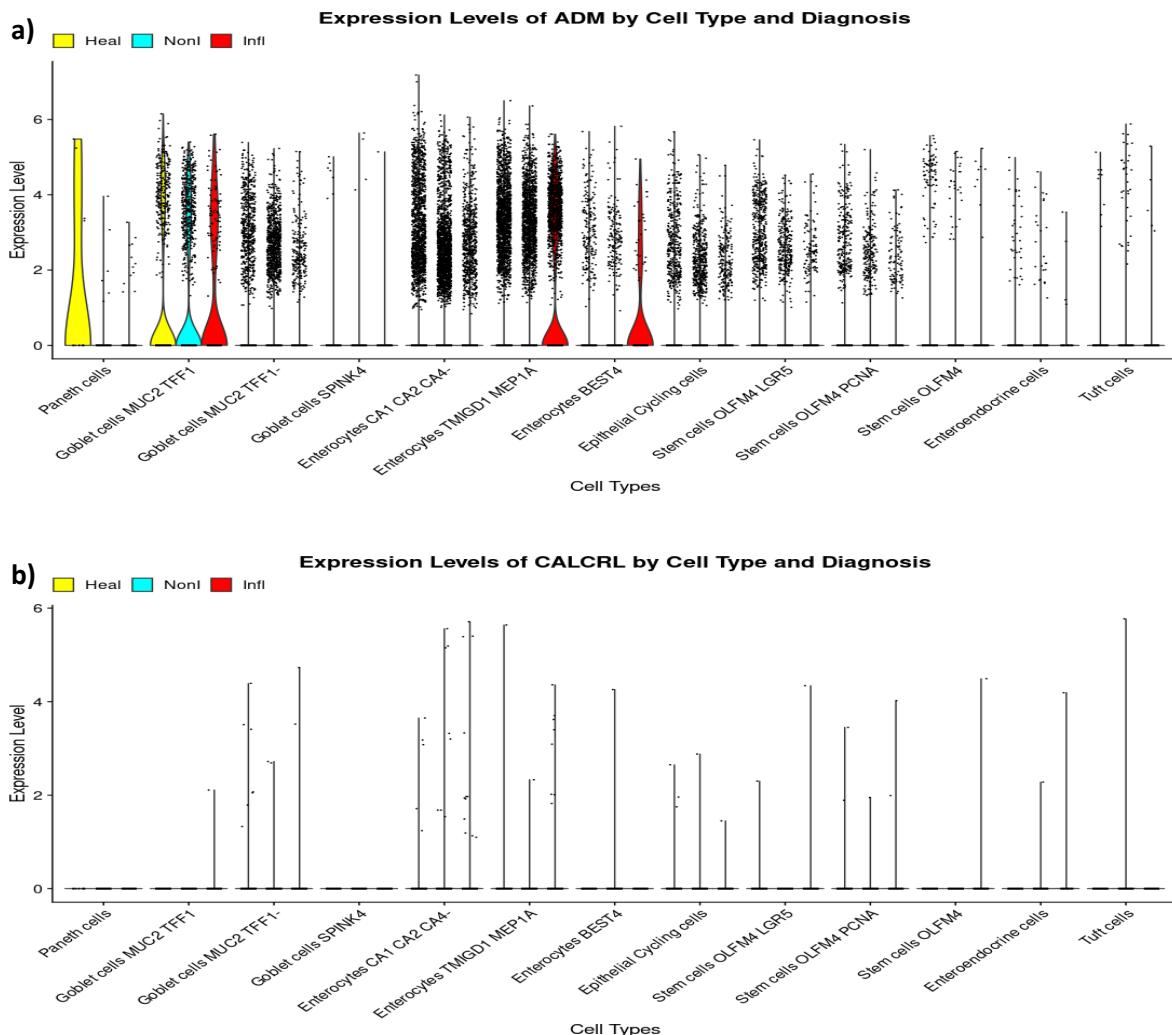


Figure S11. Terminal ileum epithelial cells. Violin plots showing the expression levels of marker across various cell types and conditions (Healthy -yellow, Non-inflamed -cyan, Inflamed – red) a) ADM b) CALCRL c) RAMP2 d) RAMP3 e) ACKR3



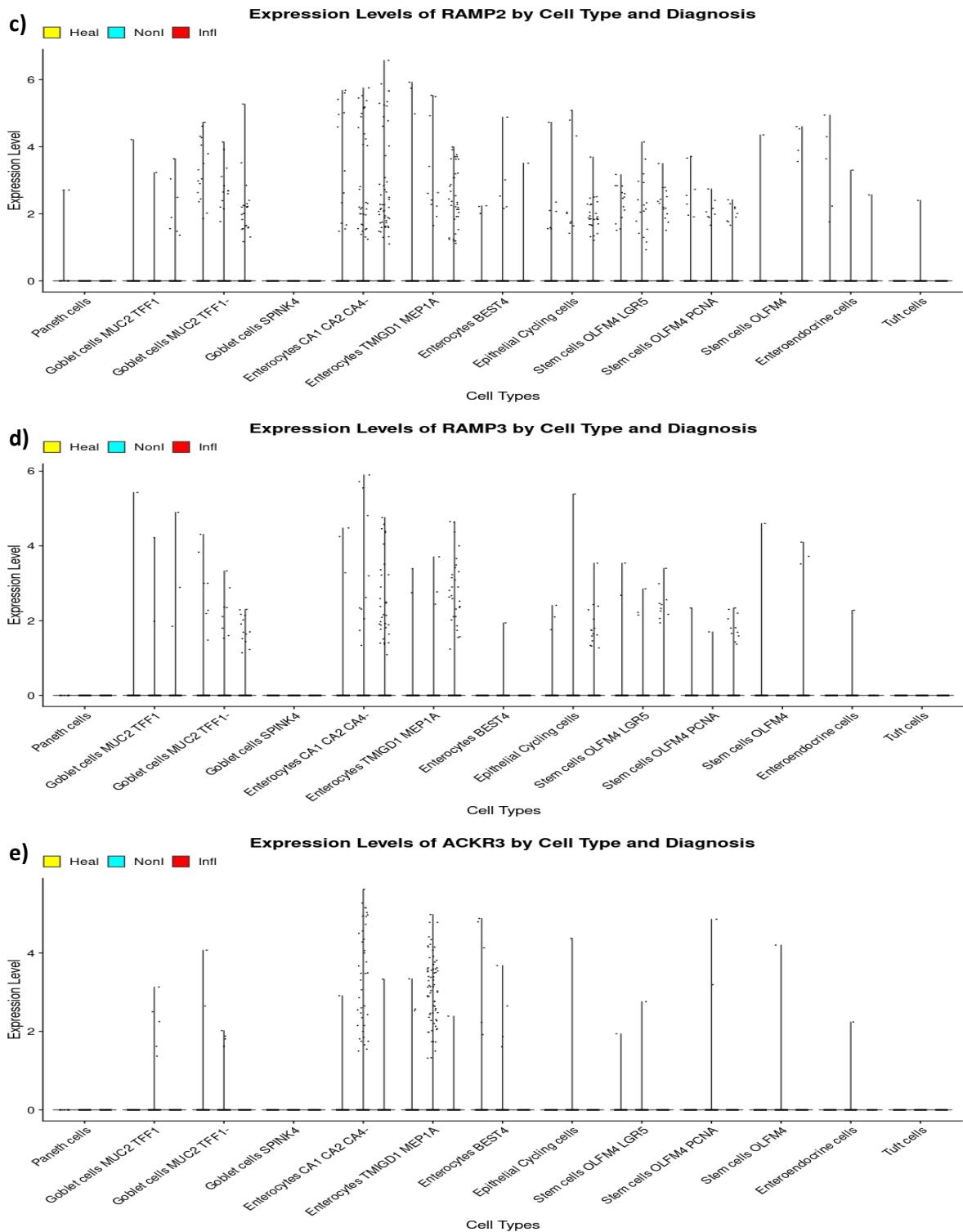


Figure S12. **Colon epithelial cells.** Violin plots showing the expression levels of marker across various cell types and conditions (Healthy -yellow, Non-inflamed -cyan, Inflamed – red). a) ADM b) CALCRL c) RAMP2 d) RAMP3 e) ACKR3

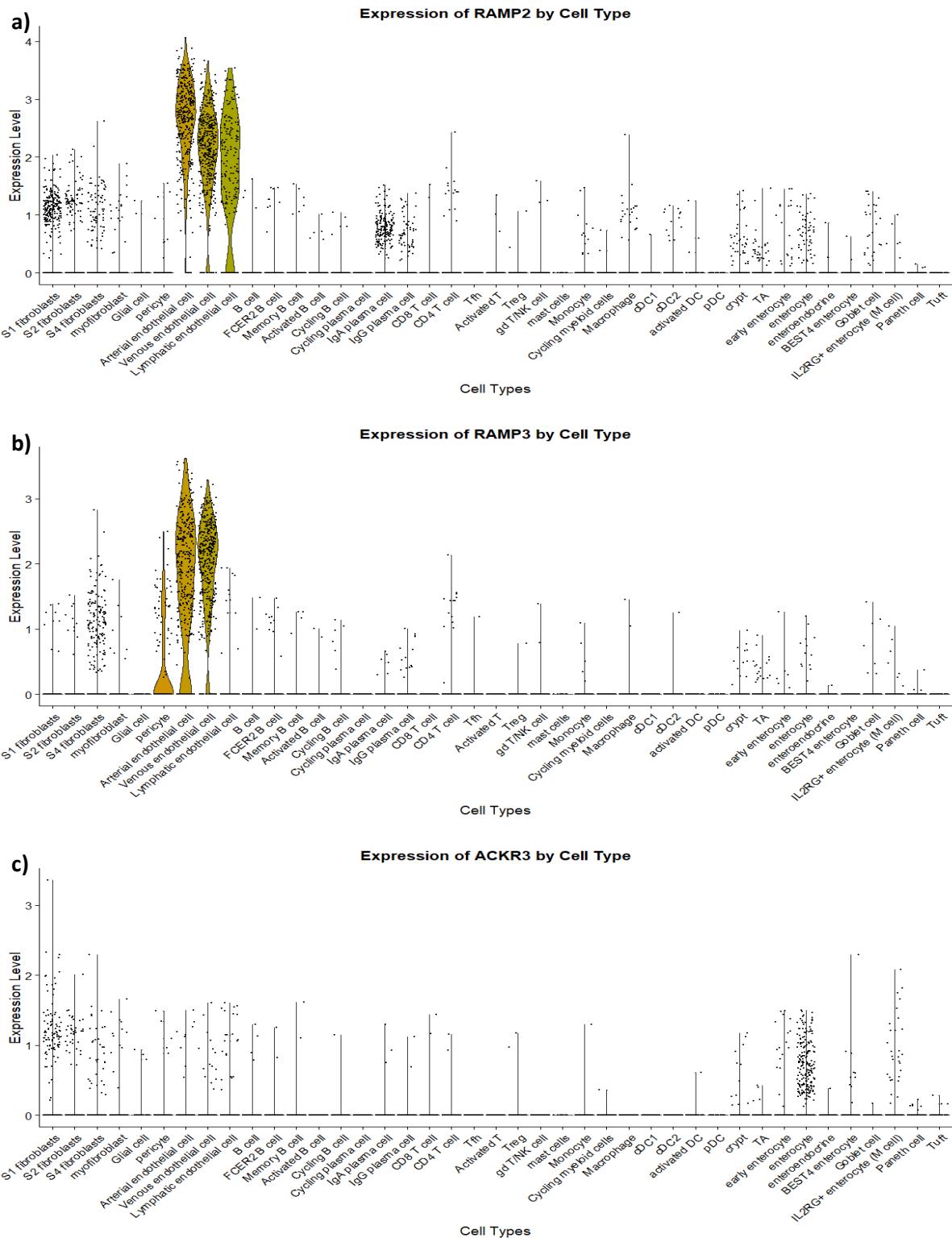


Figure S13. Pediatric dataset. Violin plots showing the expression marker genes across various cell types. a) RAMP2 b) RAMP3 c) ACKR3

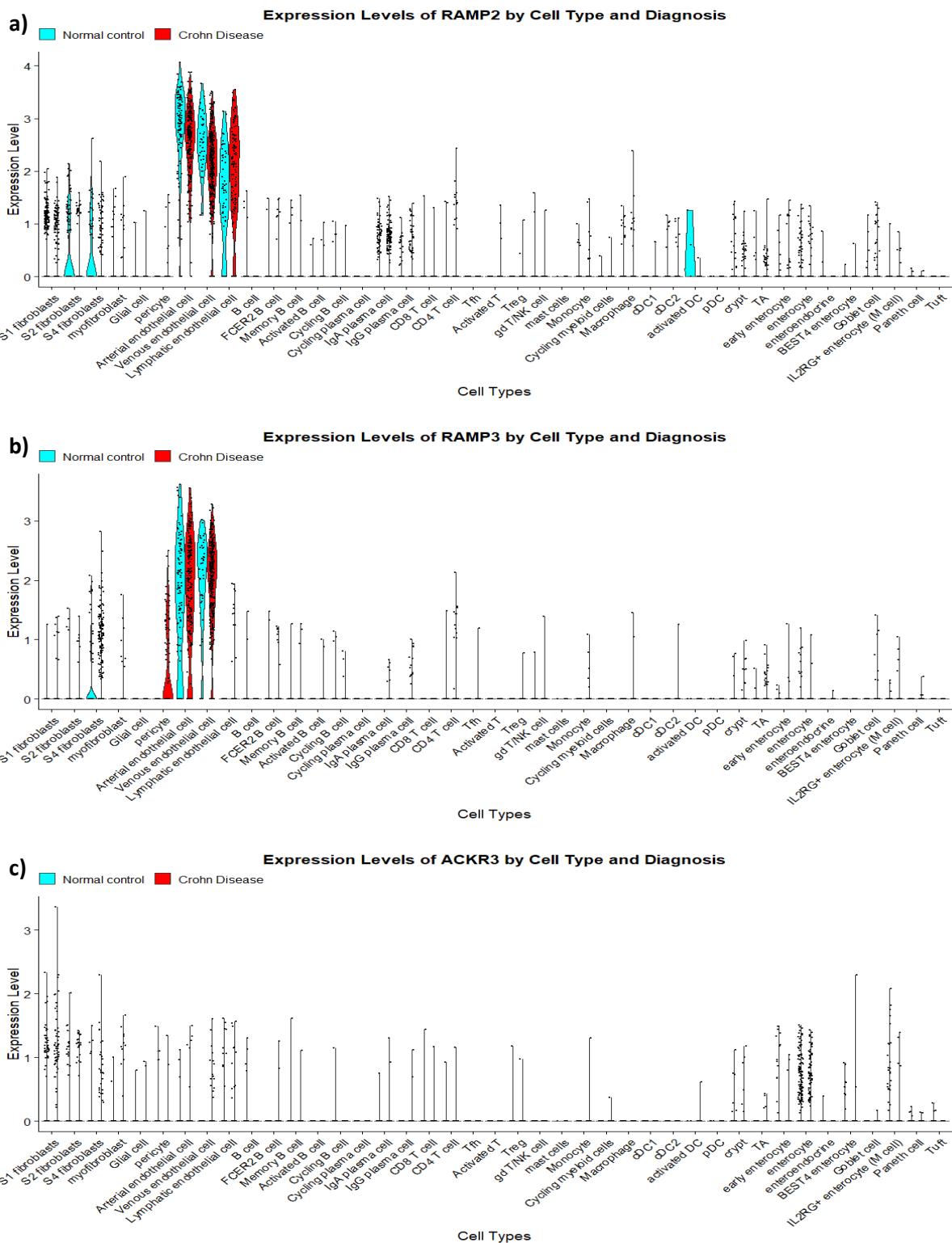


Figure S14. Pediatric dataset. Violin plots showing the expression levels of marker across various cell types and conditions (Healthy -cyan, Crohn disease – red). a) RAMP2 b) RAMP3 c) ACKR3