UNIVERSITAT DE BARCELONA

Dissecting the cellular basis of Inflammatory Bowel Disease through Single-cell RNA sequencing

Alba Garrido Trigo

UNIVERSITAT DE BARCELONA

UNIVERSITAT DE BARCELONA

FACULTAT DE FARMÀCIA I CIÈNCIES DE L'ALIMENTACIÓ

Doctoral Thesis

Dissecting the cellular basis of Inflammatory Bowel Disease through Single-cell RNA sequencing

Alba Garrido Trigo

2023

UNIVERSITAT DE BARCELONA

FACULTAT DE FARMÀCIA I CIÈNCIES DE L'ALIMENTACIÓ

Doctoral Program in Biomedicine

(Immunology)

Institut d'Investigacions Biomèdiques August Pi I Sunyer (IDIBAPS)

Gastroenterology Department-Inflammatory Bowel Disease Group

Dissecting the cellular basis of Inflammatory Bowel Disease through Single-cell RNA sequencing

Doctoral Thesis submitted by Alba Garrido Trigo to obtain a PhD degree from the Universitat de Barcelona

Dr. Azucena Salas Martínez

Thesis Supervisor and Tutor

Dr. Marisol Veny Álvarez-Ossorio

Thesis Co-Supervisor

Alba Garrido Trigo

PhD Candidate

2023

“Imagination is more important than knowledge. Knowledge is limited. Imagination encircles the world”

A.Einstein

ABSTRACT

Crohn's disease and ulcerative colitis are chronic inflammatory bowel diseases that show perplexing heterogeneity which is poorly understood at a cellular level. We applied single-cell RNA sequencing and CosMx™ Spatial Molecular Imaging to human colon and observed profound changes in the epithelium of IBD patients with the presence of cell loss and the increase of markers linked to inflammation. Additionally, we found the highest diversity among patients in the myeloid and stromal compartment. We characterized for the first time the transcriptional signatures of intestinal neutrophils finding three different neutrophil states, including a *CXCR4+* population not found in blood. We detected, in both healthy and inflamed colon, a novel resident macrophage population (M0), that differs from the previously described M2. In addition, we found uniquely in patients a variety of classical inflammatory M1 macrophages and a previously undescribed population that we called Inflammation-Dependent Alternative (IDA) macrophages. Single-cell transcriptomic, spatial data and immunostaining uncovered two distinct populations of IDA macrophages: subepithelial IDA macrophages highly expressing *NRG1* that can act on the epithelium; and NRG*1*^{low} IDA macrophages, which were expanded within the submucosa and in Crohn's disease granulomas. *NRG1* was found to be highly upregulated in IBD, especially UC. Remarkably, we also detected expression of *NRG1* in peri-cryptal S2b fibroblasts in healthy tissue and in other fibroblast populations in IBD. Additionally, S2b fibroblasts express the neuromodulator NPY in healthy samples and was profoundly downregulated in IBD patients. Lamina propria S1 fibroblasts decrease in IBD but remaining cells show an increase in the expression of pro-inflammatory genes. Thus, the peri-cryptal and S1 fibroblasts suffer a rewiring in IBD. We also identified a population of inflammatory fibroblasts in IBD which had a unique transcriptomic profile. Remarkably, we observed high co-localization of inflammatory fibroblasts and novel IDA macrophages through CosMx SMI, suggesting cellular crosstalk.

The combination of single-cell sequencing and spatial analysis allowed us to unravel the complexity of cell populations in inflammatory bowel disease and find potential cross-talks between cell types that could help in understanding the disease and identifying new therapeutic targets.

TABLE OF CONTENTS

INTRODUCTION

SECTION 1: INFLAMMATORY BOWEL DISEASE

1. Inflammatory bowel disease classification

Crohn's disease (CD) and ulcerative colitis (UC) are chronic inflammatory bowel diseases (IBD) with a relapsing and remitting course that leads to bowel damage and disability. IBD is believed to result from the interplay between genetic susceptibility, environmental factors, and intestinal microbiota, resulting in an abnormal mucosal immune response and compromised epithelial barrier function[1]. CD may present with extremely variable phenotypes in terms of disease location (small bowel, colon or upper gastrointestinal tract), severity, and development of stenosing and penetrating complications (fistulas and abscesses). UC is also heterogeneous in terms of disease extension and severity but only involves the colon. The inflammation usually starts in the rectum and can spread and affect a portion or the entire colon. While CD can present a transmural inflammation of the gastrointestinal (GI) wall, UC only involves the colonic surface mucosa or, occasionally, the submucosa[2] (**Figure 1**).

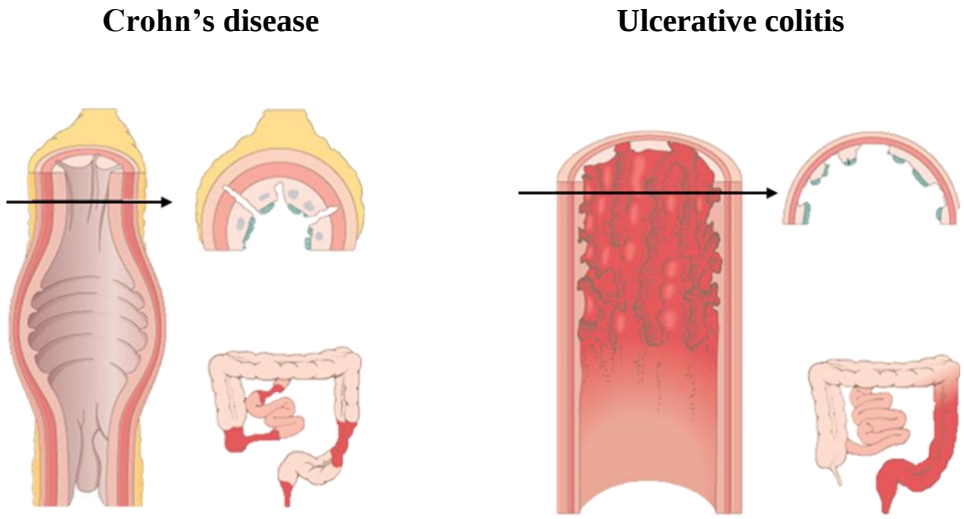


Figure 1. Scheme representing Crohn's disease and ulcerative colitis intestine showing disease extension, location, and global inflammation phenotype. Crohn's presents a transmural inflammation and in contrast, ulcerative colitis presents an inflammation only affecting the surface of the intestinal mucosa.

2. Epidemiology.

Incidence (the probability of occurrence of a given medical condition in a population within a specified period) and prevalence (the proportion of a particular population found to be affected by a medical condition at a specific time) rates for both CD and UC are found to be higher in the North of Europe and North America. In fact, IBD is more commonly found in industrialized and in westernized diet regions, suggesting that environmental factors might greatly influence IBD occurrence (**Figure 2**). Indeed, the incidence of IBD is particularly increasing in countries or areas, such as Asia or Eastern Europe, where the number of cases was relatively low hitherto[2,3]. Within Europe, the incidence of IBD is characterized by a north–south gradient, where the incidence, between 2010-2019, of CD was 6.3 per 100,000 in Northern Europe in comparison to only 3.6 per 100,000 in Southern Europe, and the incidence of UC in Northern and Southern Europe was 11.4 and 8.0 per 100,000, respectively[4,5]. Regarding age, IBD patients usually present the first episode of the disease between the ages of 20-30 years. Nevertheless, 5-15% of the patients are for the first time diagnosed between the age of 50-60. In relation to gender distribution, the incidence of IBD appears to be similar for both females and males (even though in pediatric disease, boys are more affected than girls)[6,7].

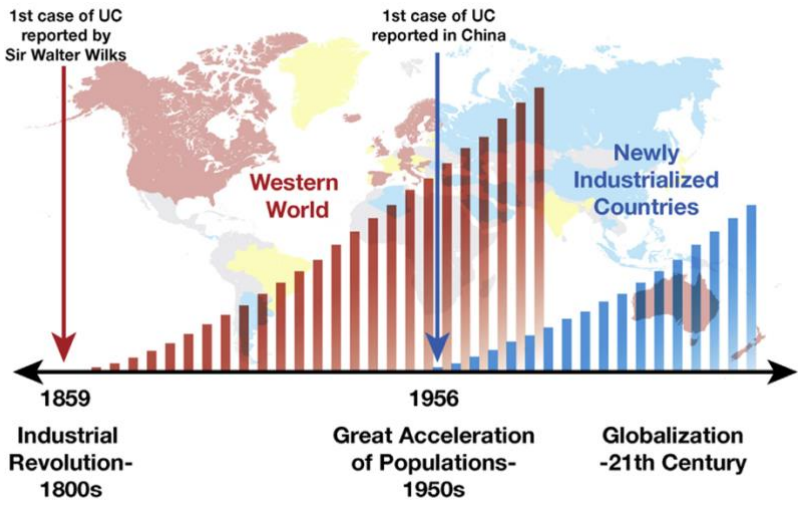


Figure 2. Increasing trend of IBD in industrialized countries. Figure from *Kaplan GG, Ng SC et al. Gastroenterology 2017*

3. Diagnosis

3.1 Symptomatology

IBD symptomatology can include hemorrhagic diarrhea, abdominal pain, tenesmus, urgency to evacuate, anorexia, and weight loss[8-10]. Nevertheless, these clinical characteristics vary depending on whether patients suffer from CD or UC. UC patients

usually experience pain in the lower left abdomen as well as diarrhea[11]. As a result, they may undergo weight loss and residual blood is found during rectal examination. On the contrary, patients suffering from CD feel pain in the lower right side of the abdomen while rectal bleeding is less common. The most common complication in CD is intestinal obstruction due to transmural inflammation, which results from intestinal wall thickening. In addition, CD patients suffer from malnutrition complications as a result of a decreased nutrient absorption[12,13]. Among 50-60% of IBD patients also suffer from extraintestinal manifestations such as arthritis, hypertrophic osteoarthropathy, vitiligo or psoriasis[11,14].

3.2 Disease Extension

As we already mentioned, disease extension significatively differ between both diseases: while UC only affects the colon[8,10] (**Fig. 3**), CD can affect the entire GI tract[2] (**Figure 4**). Around 1/3 of UC patients suffer from proctitis (**Fig 3A**), a type of intestinal inflammation that only affects the rectum. When the inflammation is extended along the colon, we can distinguish between distal colitis (**Fig 3B**) and pancolitis (**Fig 3C**). The extent of the disease and its severity correlate with clinical prognosis in UC. Thus, patients with proctitis usually have a better prognosis than those with a more extensive disease[15]. Parallelly, extensive colitis is commonly associated with a higher risk of colectomy[16] and colorectal cancer[17].

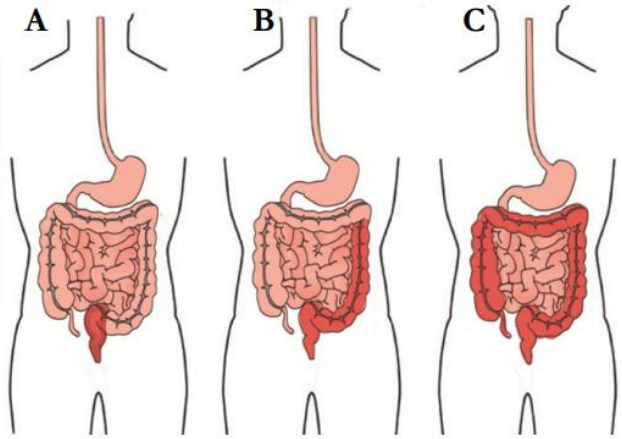


Figure 3. Disease extension in ulcerative colitis. (A) Proctitis; (B) Distal/Left colitis; (C) Pancolitis.

Regarding CD patients, according to the Montreal classification[18], intestinal inflammation might be present on the terminal ileum – L1 – (**Figure 4A**), the colon – L2 – (**Figure 4B**), both the ileum and colon – ileocolonic disease, L3 – (**Figure 4C**) or on the upper GI tract – L4 – (**Figure 4D**)[13]. Moreover, some patients might suffer from both L3 and L4 manifestations thus affecting the distal and upper GI tract [13](**Figure 4E**). The Montreal classification also describes CD according to its behavior, defining the disease as B1 when it is non-stenotic and non-penetrating; B2, for the stenotic disease and B3, when it is penetrating[19].

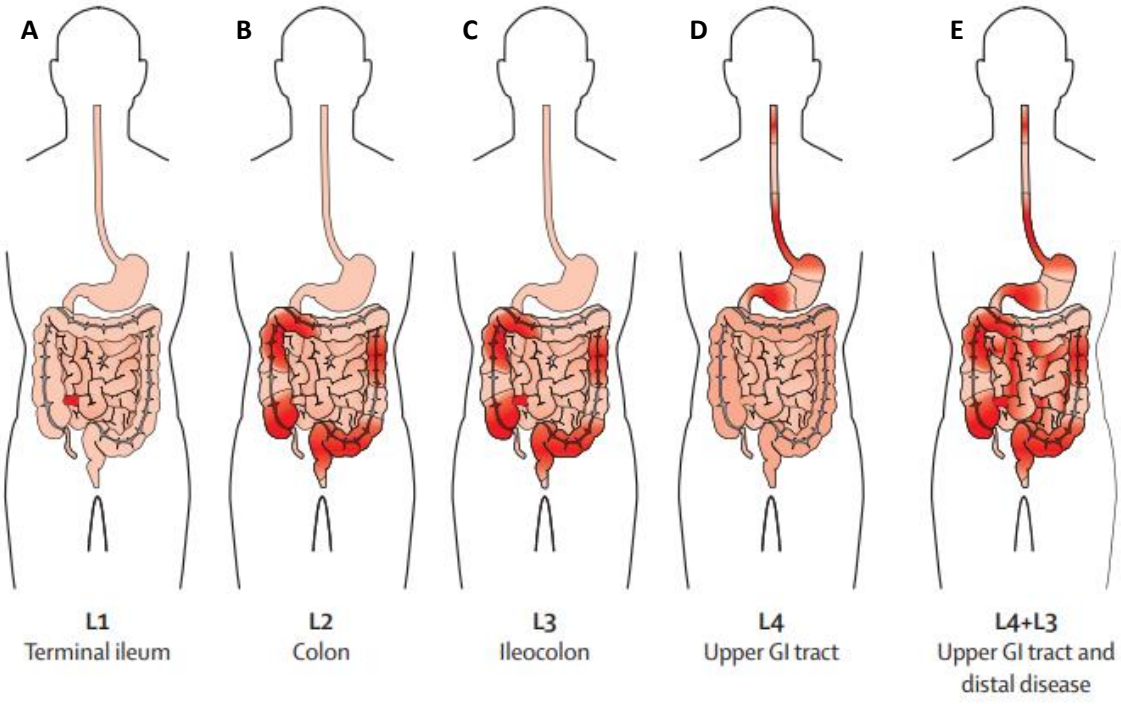


Figure 4. Disease extension in Crohn's disease by Montreal L-category. (A) Terminal Ileum inflammation "L1"; (B)Colonic affectation "L2"; (C) Ileocolonic disease "L3"; (D) Upper GI tract affectation "L4"; (E) Distal and upper GI tract inflammation "L3+L4".

3.3 Clinical activity and endoscopic indexes

In clinical practice, healthcare providers use a variety of methods to evaluate and monitor disease activity in patients with IBD. Two commonly used methods are clinical activity indexes and endoscopic indexes. Clinical activity indexes are tools that healthcare providers use to assess a patient's symptoms and overall disease activity. Endoscopic indexes, on the other hand, are tools that healthcare providers use to assess the extent and severity of inflammation in the digestive tract. These indexes involve performing an endoscopy, which is a procedure in which a flexible tube with a camera on the end is

inserted into the patient's digestive tract. The healthcare provider can then visually assess the extent and severity of inflammation in the digestive tract and assign a score based on established criteria.

- **Crohn's disease**

Clinical activity. The most widely used index of activity in clinical trials has been the validated Crohn's disease activity index (CDAI). It collects clinical data from a 7-day patient diary, including patient-reported stool patterns, average abdominal pain, general well-being, complications such as fistulas, abscesses or fevers, findings of an abdominal mass, hematocrit values and weight. It divides the clinical severity of the disease into the following categories: quiescent (CDAI <150), mild (CDAI 150-220), moderate (CDAI 220-450) and severe (CDAI >450)[20].

Endoscopic activity can be measured by the Crohn's Disease Endoscopic Index of Severity (CDEIS) in cases of ileal and/or colonic activity. The CDEIS collects data on the presence of superficial and deep ulcers, stenosis, as well as those mucosal surfaces affected by disease and ulceration in all colonic segments[21]. Daperno and colleagues presented an alternative index, the Simplified Endoscopic Severity Index for Crohn's Disease[22] (SES-CD, Table 1) which is currently wide used. This index evaluates the penetration and size of ulcerations, which are known to reflect the severity of Crohn's disease. Each segment (rectum, descending colon and sigma, transverse colon, ascending colon, and ileum) is evaluated using SES-CD and the result is all values summation. The SES-CD has excellent inter-observer agreement for all selected variables. It is also correlated to a considerable degree with the CDEIS (r = 0.92)[22].

Simple Endoscopic Score for Crohn's Disease values

Variable	**0**	**1**	**2**	**3**
Size of ulcers	None	Apthous ulcers (Ø 0.1 to 0.5 cm)	Large ulcers (Ø 0.5 to 2 cm)	Very large ulcers (Ø >2 cm)
Ulcerated surface	None	<10%	10-30%	>30%
Affected surface	Unaffected segment	<50%	50-75%	>75%

Presence of stenosis	None	Single, can be passed	Multiple, can be passed	Cannot be passed

Table 1. Definitions of Simple Endoscopic Score for Crohn's disease (SES-CD). Ø = Diameter. Table from *Daperno, M. et al Gastrointest Esdosc. 2004*

- **Ulcerative colitis**

Mayo Index is the widest used index to evaluate the severity of UC. It is a global index that includes the clinical and the endoscopic evaluation divided in four categories: deposition frequency, rectal bleeding, data obtained from colonoscopy and global assessment of physicians. The Global Mayo index goes from 0 to 12, and the endoscopic part of the index (endoscopic Mayo score[23]) has a range between 0 and 3, being 0 an inactive disease and 3 a severe activity. The different degrees of the endoscopic Mayo score are specified in Table 2.

Score	***Disease activity***	***Endoscopic features***
0	Normal or inactive	None
1	Mild	Erythema, decreased vascular pattern, mild friability
2	Moderate	Marked erythema, absent vascular pattern, friability, erosions
3	Severe	Spontaneous bleeding, ulceration

Table 2. Endoscopic Mayo score to evaluate ulcerative colitis.

4. Pathogenesis

Even though the etiology of IBD is complex and still not fully understood, several studies have suggested that IBD are multifactorial diseases that are believed to result from the interplay between genetic susceptibility, environmental factors, and chronic response to intestinal microbiota and dietary antigens, resulting in an abnormal mucosal immune response and compromised epithelial barrier function (**Fig. 5**).

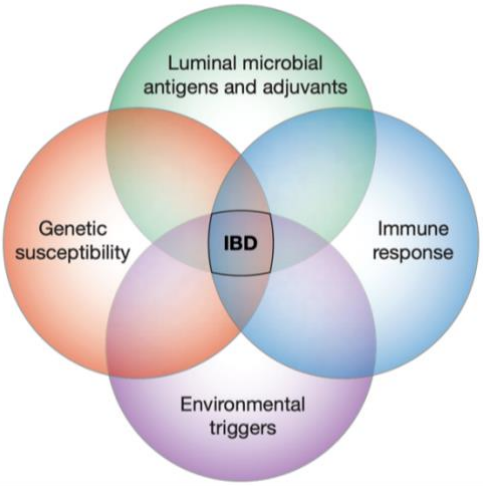


Figure 5. Interaction of various factor contributing to intestinal inflammation in IBD. Figure from *Sartor, R. Nat Rev Gastroenterol Hepatol 2016*.

4.1 Genetic Susceptibility

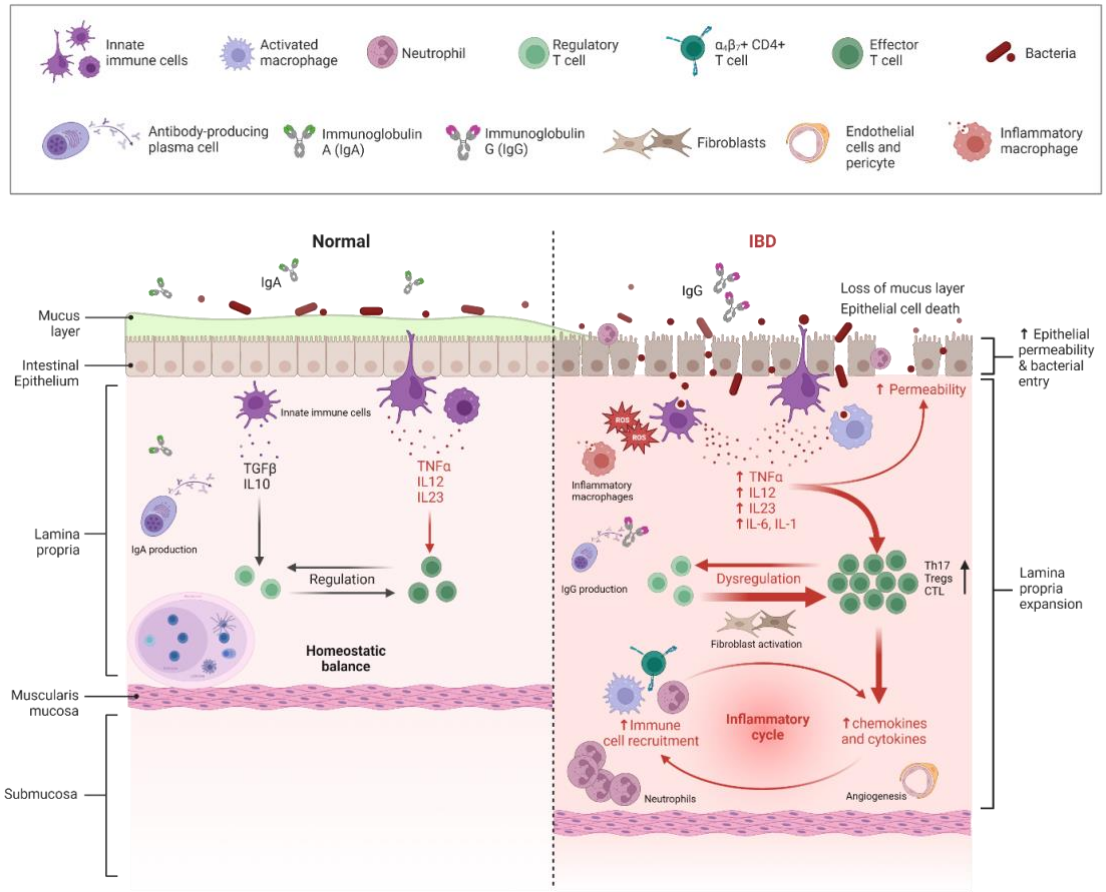
Thanks to next generation sequencing technologies, in particular Genome Wide Association Studies (GWAS) which can identify single nucleotide polymorphisms (SNPs), and the worldwide-publicly available databases, there was a great advance in the study of this field[24]. The first identified gene which appeared to be mutated in CD was NOD2 (also known as CARD15 or IBD1)[25]. Later, another study identified 163 SNPs associated to IBD (110 shared in both CD and UC, 30 specifics for CD and 23, for UC) thanks to the integration of 15 different GWAS. From the 163 identified SNPs, some genes – mainly related to autophagy (ATG16L1 or IRGM) and the immune system (interleukin (IL)-23R, JAK2 or STAT3) – have been identified to be altered in IBD patients[26]. Moreover, some of these genes have also been found to be modified in patients suffering from other autoimmune diseases, suggesting that IBD patients might present common pathways to other diseases[27]. Despite the great variety of SNPs that have been associated to IBD, altogether they only account for the 20-25% of the disease genetic susceptibility. This phenomenon does not only happen in IBD but in many other polygenic diseases.

4.2 Immune Response

The intestine is an organ highly exposed to foreign particles. The first barriers to protect the host are anatomical and chemical but underneath that, the intestinal mucosa harbors the largest compartment of the immune system in the body[28]. The immune system in the intestinal mucosa is a complex network of cells and molecules that work together to protect the body from harmful pathogens while tolerating harmless substances like food

and commensal bacteria[29]. Individuals with IBD are characterized by chronic inflammation and tissue damage, which results from dysregulation of many immune and non-immune factors.

In summary, there is a major barrier permeability due to epithelial cell death and deregulation that leads to a decrease in mucus production. Therefore, the loss of integrity on this barrier enables the intestinal luminal microorganisms to access and cross the intestinal epithelium and to interact with the immune system underneath[11,30]. Consequently, there is an increased activation of innate resident cells such as dendritic cells (DCs) and macrophages. The activation of these cells and the production of inflammatory mediators lead to the recruitment of monocytes and neutrophils from blood and activation of other resident cells such as fibroblasts and endothelium. It also leads to the secretion of inflammatory cytokines and the further activation of cells from the



adaptative immune system such as T, B and plasma cells (Figure 6). Deregulation of the

Figure 6. Disease mechanisms. Graphical representation of a comparison between a healthy colon and an inflamed colon affected by IBD. The image depicts various cellular processes and disease mechanisms that occur in the inflamed colon.

mucosal immune system (both innate and adaptive responses) has been extensively

associated with the pathogenesis of IBD[31]. In Section 2 of this Introduction, the cell types that play a role in IBD and their alterations in the disease will be further discussed.

4.3 Microbiota

In the gut a massive and diverse community of microbes circulates in proximity to the epithelial cell surface. The unhealthy balance of bacteria, as well as the abnormal communication between gut microbial communities and the host immune system at the mucosal barrier, has been identified as a defect leading to chronic intestinal inflammation. Dysbiosis, which is the global alteration of the gut microbial community, has been reported in Crohn's disease during the last 10 years[32,33]. IBD is associated with lower microbial diversity and certain species may play specialized roles in the disease. By contrast, some microbial species may act as protective commensals. As just one example, a decrease in Faecalibacterium prausnitzii, a butyrate-producing bacteria belonging to the Firmicutes phylum, is associated with Crohn's disease [33,34]and its abundance inversely correlates with endoscopic recurrence at 6 months[35]. However, currently, we do not have a clear understanding of whether the alterations in microbiota observed in individuals with IBD are a cause or consequence of the disease.

4.4 Environmental Factors

Chronic inflammatory disorders and neoplasms have become the main cause of morbidity and mortality during the last century in the Western world[36]. The increase in chronic autoimmune and inflammatory diseases (such as IBD) has been linked to the social and economic progress, as well as the increase in life expectancy that first took place in northern Europe and America, but which can currently be seen in other parts of the world (rest of Europe, Japan and South America)[37]. The "hygiene theory", or the dramatic decrease in human exposure to microbes, has been proposed as a possible contributor to this mentioned shift. This lack of contact with microbial antigens early in life affects the proper maturation of the immune system so that it would not be equipped to effectively act and thus prompting to a much more ineffective immune response[38]. Other environmental factors related to IBD include tobacco, diet, certain drugs, and stress, among others. Tobacco is the most influential environmental factor in IBD with opposite effects on UC and CD. In CD, tobacco is a risk factor that increases the risk of relapse and/or surgical resection. In UC, smoking cessation worsens the disease, suggesting a protective effect of tobacco[39]. Another environmental factor that has been linked to IBD are specific pharmacological treatments. For example, oral contraceptives have been

related to a higher risk of developing IBD, especially CD[40]. On the other hand, nonsteroidal anti-inflammatory drugs (NSAIDs) are shown to be important in causing relapse[41]. Stress also plays an important role in the disease; anxiety and depression may be important in relapse and deterioration of the disease[42]. Many other environmental factors have been linked to IBD; however, there is not sufficiently strong evidence to support a causative effect of these factors in the development of the disease. Thus, it has been accepted that all these factors might only have an accumulative effect on genetically predisposed patients[43].

5. Treatment

The aim of the currently available treatments for IBD is to induce and maintain remission. Rather than reversing pathogenic mechanisms, they intend to decrease the symptoms of the disease. Corticosteroids, immunosuppressants and biologic treatments are the most routinely used and are directed to reduce the overactivation of the immune response. Other drugs such as antibiotics might also be useful in some cases[44]. Nevertheless, when biological treatments are not effective, surgical resection may be required. Some of the most common biological drugs used in IBD are explained below:

§ Corticosteroids are usually used to induce remission in moderate or severe IBD since they act as potent anti-inflammatory drugs. They inhibit cyclooxygenases and regulate immune cells by reducing their proinflammatory capacity. Corticosteroids can be given orally (prednisolone, prednisone, budesonide) or intravenously (hydrocortisone or methylprednisolone). Liquid suppositories or enemas may also be given. Different strategies have been developed to maximize their topical effects while limiting the systemic side effects of steroids. In that sense, budesonide is a poorly absorbed corticosteroid with limited biodistribution that has therapeutic benefits with reduced systemic toxicity in CD[45] or UC[46]. These agents cannot be used as maintenance therapy due to the large number of associated side effects[47,48].

§ Immunosuppressants (including azathioprine, mercaptopurine or methotrexate) inactivate key processes, such as proliferation and survival of T lymphocytes that can promote inflammation. Except for methotrexate, this group of drugs might also have toxic effects due to their capacity to intercalate within nucleic acids[49]. Immunosuppressants are usually prescribed when patients do not respond to corticosteroids, or they are corticosteroid dependent. Specifically, in CD patients immunosuppressants are useful for

both inducing and maintaining remission [50]while in UC they are mostly indicated for maintenance[51]

§ Biologics used in IBD are monoclonal antibodies directed against different targets involved in its pathogenesis such as cytokines (TNFα, IL-12, IL-23) or α4β7 integrins, involved in lymphocyte homing to the gut. Anti-TNFα therapy is the most used biologic among IBD patients. Infliximab, adalimumab and certolizumab are the three anti-TNFα drugs that have been proved to reduce IBD-related hospitalization and the risk of surgery since they can induce mucosal healing and improve patients' life quality[52]. However, 30-50% of patients receiving this treatment fail to respond. Thus, new strategies that block alternative targets emerged. Other biologics in use include ustekinumab (which blocks the p40 subunit, common for both IL-12 and IL23 cytokines) or Vedolizumab (which blocks the α4β7 integrin[53]).

§ JAK inhibitors. Tofacitinib (Xeljanz©) and filgotinib (Jyseleca©) are synthetic small molecules that are administrated orally and have been only approved to treat UC[54,55]. Tofacitinib is a pan Janus kinase (JAK) inhibitor that has shown to induce and maintain remission in a percentage of patients with moderate to severe UC[56]. Filgotinib is a JAK inhibitor with higher affinity for JAK1. Upadacitinib is a second-generation selective JAK inhibitor targeting also the JAK1 enzyme and it is approved as a second-line agent for treating moderate to severe active rheumatoid arthritis (RA)[57] and a clinical trial (CELEST) has already been performed on CD patients[58].

Other strategies used to treat IBD are hematopoietic stem cell transplant[59] for CD patients with highly refractory and severe disease, or the mesenchymal stem-cell therapy to treat fistulizing perianal disease in CD patients[60,61].

Although there have been significant developments in the treatment of inflammatory bowel disease (IBD), the variability in therapeutic responses in both Crohn's disease (CD) and ulcerative colitis (UC) remains a perplexing observation. While some patients achieve complete remission within a few days, approximately 30-40% of patients are refractory to any treatment[62], despite the use of newly available drugs. Moreover, some IBD patients may lose their response to treatment over time, leading to worsening symptoms. This lack of response could be due to the patient's immune system developing resistance to medication or progression of the disease to a more severe form that is less responsive to treatment. Although the therapeutic armamentarium for IBD has expanded

rapidly in recent years, none of the new drugs achieve more than 20-30% remission rates at one year. Therefore, physicians are often left with the rather mundane step-up and trial-and-error approach of following IBD general guidelines[63]. In some cases, patients require surgical resection due to disease complications or uncontrolled inflammation. The most used surgical procedure is colectomy with ileostomy[64], which involves the removal of the inflamed area. Despite significant advances in our understanding of IBD and the development of new treatment options, we still lack reliable predictive factors that can accurately forecast how individual patients will respond to different treatments.

SECTION 2: THE INTESTINE IN HOMESTASIS AND INFLAMMATION

1. Anatomy of the intestine

The intestine is a tube which extends from the lower end of the stomach to the anus, the lower opening of the digestive tract. The intestine is divided into two sections called the small intestine and the large intestine. The small intestine is made up of three segments (duodenum, jejunum and ileum) and the main functions of the small intestine are digestion, absorption of food and production of gastrointestinal hormones.

The large intestine which is the focus of this thesis is made up of the cecum (a pouch) and the colon, which consists of four parts (ascending colon, transverse colon, descending colon, sigmoid colon) and the rectum. The primary functions of the colon are the reabsorption of fluids (water) and electrolytes and elimination of undigested food. All the segments of the small and large intestine are composed of the four layers characteristic of the gastrointestinal tract: mucosa (epithelium, lamina propria and muscular mucosae), the submucosa, the muscularis propria (inner circular muscle layer, intermuscular space, and outer longitudinal muscle layer), and the serosa.

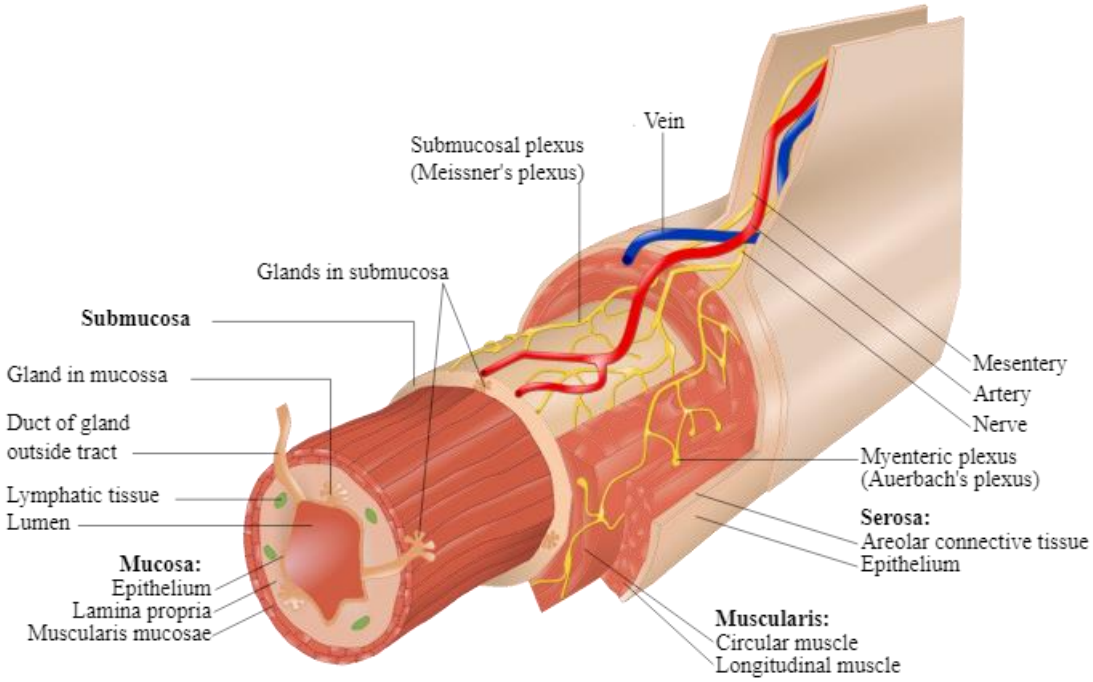


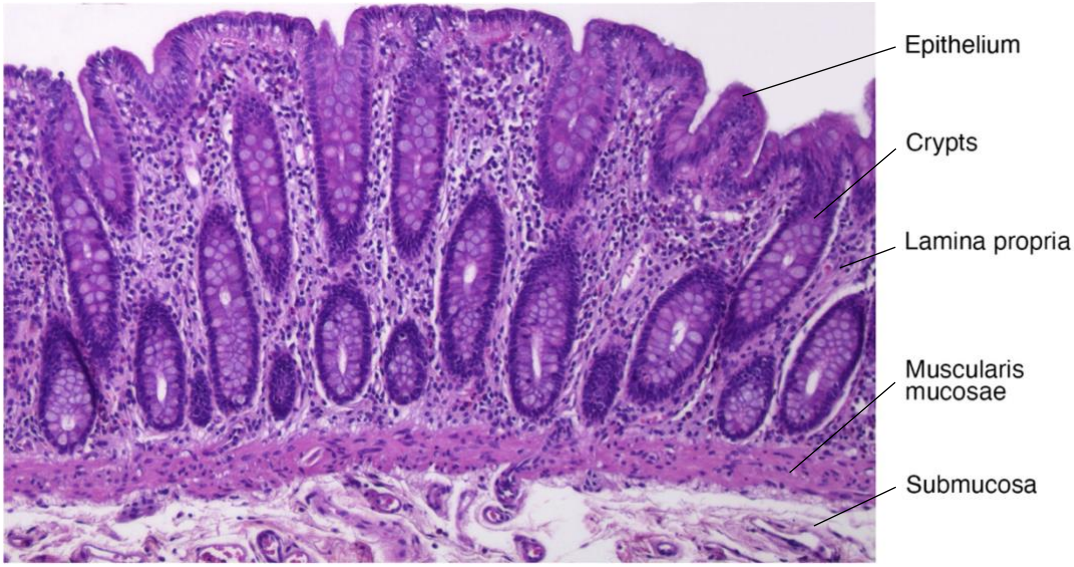
Figure 7. Anatomy of the intestine. Graphical representation of the different layers of the intestinal wall and its components.

Mucosa (or mucous membrane). The mucosa is the innermost layer and consists of three layers. The epithelium, the lamina propria and the muscularis mucosae. The first layer facing the intestinal lumen is made up of epithelial cells, which form a simple columnar epithelium attached to a basement membrane overlying the second layer, the lamina propria. The epithelial layer forms crypts of Lieberkuhn, also known as intestinal glands, are small, tubular invaginations in the lining of the small and large intestines that extend down into the submucosa and they contain various types of epithelial cells. The lamina propria is located underneath the epithelium and contains the large immune system in the body: the gut-associated lymphoid tissue (GALT) where immune responses in the colonic mucosa are mainly ensured. The GALT is a type of mucosal associated lymphoid tissue (MALT) specific of the intestine that presents lymph node-like structures. Underneath the lamina propria lays the third and deepest layer: the muscularis mucosa, a layer of smooth muscle that lies at the base of the lamina propria and it is made of stromal cells called myofibroblasts (**Figure 8**).

Submucosa. Dense-irregular connective tissue that contains Meissner's Plexus (or submucosal plexus) which are neuronal ganglia made up of neurons and glia that provides

Figure 8. Hematoxylin-Eosin staining of a healthy intestinal mucosa. Arrows show the epithelium, the crypts, lamina propria, the muscularis mucosae and the submucosa.

sensory, motor and secretory innervation to the inner layers. Also contain numerous blood and lymphatic vessels.



Muscularis propria. Consists of two orthogonal layers of smooth muscle (the inner and outer layer). The inner layer is arranged in circular rings around the tract, whereas the outer layer is arranged longitudinally. Between the two layers are the Auerbach's Plexus (or myenteric plexus) provides motor innervation of the muscularis externa and allows the peristalsis. Peristalsis is the coordinated muscular contractions and relaxations of the digestive tract that move food, liquids, and other ingested materials through the GI tract. The process of peristalsis is an involuntary and automatic function of the digestive system and involves the sequential contraction of circular and longitudinal muscles in the walls of the GI tract. Peristalsis is critical for digestion and absorption of nutrients and elimination of waste products from the body.

Serosa or adventitia. The outermost layer of the colonic wall that consists of several layers of connective tissue covered by a simple squamous epithelium, called the mesothelium, which reduces frictional forces during digestive movements.

2. Components of the colonic mucosa

The intestine's physiological function is to absorb nutrients and water, making its mucosa thin and permeable. However, this also renders it more susceptible to infections, as it serves as the largest entry point for infectious agents into the body. In this section, we

will explore the various cell types present in the colonic mucosa that play crucial roles in enabling this complex organ to function properly and their alterations in IBD.

2.1. Epithelium

2.2.1. The intestinal epithelium in homeostasis

The intestinal epithelium performs several critical functions that are essential for the proper functioning of the digestive system and for maintaining overall health. Among these functions, the epithelium absorbs nutrients from digested food and acts as a barrier to prevent harmful substances, such as toxins, bacteria, and antigens, from entering the body. Additionally, the epithelium secretes mucus that lubricates and protects the intestinal lining, as well as enzymes that help break down complex nutrients.

Moreover, the intestinal epithelium plays a vital role in the body's immune system by identifying and responding to potentially harmful substances. Furthermore, due to its high rate of turnover, the intestinal epithelium is involved in cell renewal and repair. As mentioned above, in the colon, it is formed by a single layer of different cell types that are organized into crypts, which are invaginations with a central hole.

The epithelial layer has the ability to renew every 4-5 days and repair any damage caused by injury or inflammation. Renewal relies on intestinal pluripotent stem cells that are located at the base of the crypt [65,66](**Fig. 9**). Stem cells divide and differentiate by migrating along the crypt. Eventually cells undergo spontaneous apoptosis and are released into the intestinal lumen. Stem cell proliferation and apoptosis are tightly controlled by paracrine signals, achieving a balance[66]. In this way, the intestinal epithelium is constantly renewed while the number of cells forming the epithelial layer remains stable. There are different types of differentiated cells in the healthy colonic epithelium that form the crypts, including colonocytes, goblet cells, Tuft cells, and enteroendocrine cells[66] . Also, it can be found undifferentiated cells as transit amplifying cells (TA). In the small intestine, Paneth (which secrete antimicrobial agents into the mucus) and M cells (in the epithelium overlying Peyer's patches) are also present. Differentiated cells are located at the top crypt area. The epithelial layer is supported by the basement membrane, a thin, pliable sheet-like type of extracellular matrix that provides cell and tissue support and acts as a platform for complex signaling.

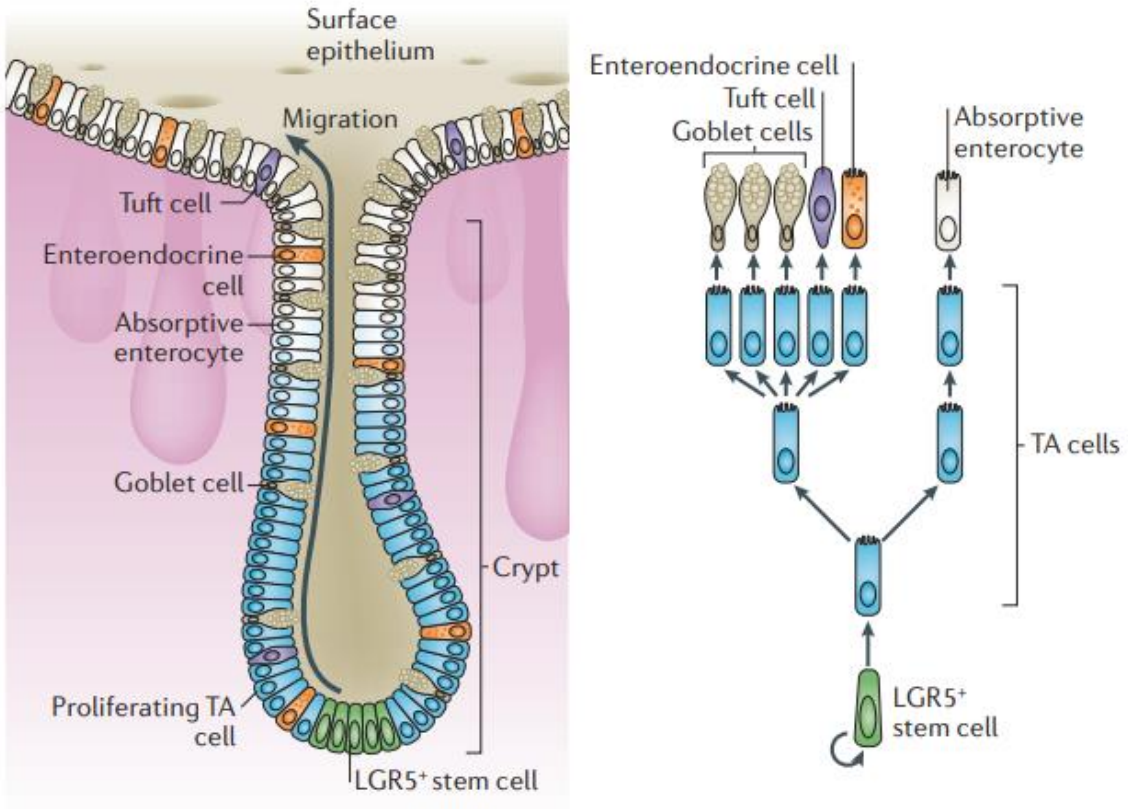


Figure 9. Epithelial self-renewal in the intestinal epithelium. In the colon LGR5+ stem cells at the crypt base generate rapidly proliferating TA cells in the lower half of the crypt (middle panel). TA cells subsequently differentiate into the mature lineages of the surface epithelium (goblet cells, enterocytes, enteroendocrine cells and tuft cells), as shown in the lineage tree on the right panel. Epithelial turnover occurs every 5–7 days. Figure from Barker, N. Nat Rev Mol Cell Biol 15, 2014.

In addition to epithelial cells, in the epithelial layer we also find intraepithelial lymphocytes (IELs)[67], which are located between epithelial cells. IELs form a heterogeneous population with cytotoxic or regulatory functions. In that way, IELs can interact with each other and with other immune cells located underneath the intestinal epithelium. IELs provide defense against pathogens, possibly through the removal of infected epithelial cells and the secretion of AMPs[68].

Importantly, above the epithelium layer there is a mucus layer formed of the mucus secreted by the epithelium and AMPs, secreted by IELs and underlying macrophages. The main role of this layer is to protect the mucosa from commensal microbes and invasive microorganisms thanks to two different types of barriers: the physical and the chemical.

The physical barrier consists of the mucus layer, the glycocalyx and the cell junctions. This barrier physically inhibits the microbial invasion of the mucosa. Specialized

epithelial cells (goblet cells) are in charge of mucus production, thus forming a viscous layer enriched in mucins: glycoproteins that form large polymers[68]. Their role is to protect the surface epithelium, lubricate and catch harmful particles. These secretory cells are more abundant in the colon compared to the small intestine. The small intestine harbors a single, tightly attached mucus layer, whereas in the colon, mucus is organized into two distinct layers: an outer loose layer (highly colonized by different microbial components) and an inner denser layer that is firmly attached to the epithelium (a feature that, together with the presence of high AMPs concentrations, prevents bacteria from reaching it)[69]. Mucus is continuously secreted, and the outer layers are sloughed off, and carried into the fecal stream[70]. The glycocalyx is a mesh of carbohydrates, glycoproteins, glycolipids, and transmembrane mucins that block bacterial invasion[71]. Cell junctions are structures specialized in cell-to-cell attachment. Altogether provide structural support and only allow the entry of certain solutes and fluids, creating a selective permeable barrier[72].

The chemical barrier is formed by AMPs such as defensins (among other anti-microbial molecules) and oxygen. Its accumulation at high concentrations within the mucosa, especially in the small intestine where Paneth cells are responsible for their production, greatly restrict the presence of potential microbial inhabitants. AMPs bind to the microbial cell membrane and induce its disruption by forming pore-shaped structures[73]. In the colon, AMPs and oxygen are produced by IECs, IELs and macrophages from the lamina propria[74].

Epithelial cell types

As mentioned before, the colonic crypt is formed by different epithelial cells:

Colonocytes are polarized differentiated absorptive epithelial cells, located at the top crypt area, responsible for the physiological function of water, electrolyte and various bacterial metabolites absorption[75]. Moreover, they are responsible for the integrity of the epithelial layer. The apical surface of colonocytes is covered with tiny microvilli increasing the cell surface for absorption and acting as sites of high digestive enzyme concentration. **Goblet cells** are simple columnar polarized differentiated cells which main function is to secrete gel-forming mucins, the main component of the mucus layer. MUC2 is the main gene responsible for that production and MUC2 KO mice present a defective mucus layer. **Intestinal Stem cells**[76] are located at the bottom of the crypt and are responsible for the rapid renewal of the entire epithelial lining. These cells give rise

to progenitor cells which subsequently differentiate into the mature cell types required for normal gut function. Maintenance of the critical absorptive function of the gut relies on the appropriate control of the stem cell pool by the local micro-environment or niche cells. **Transit-Amplifying** cells (TA) are an undifferentiated population in transition between stem cells and differentiated cells[77]. **Tuft** cells are rare, chemosensory, and secretory cells that have been proven indispensable in anti-helminthic and anti-protozoan immunity because of their production of IL-25, a critical cytokine for the initiation of type-2 immune responses again these parasites. Tuft cells are now known as important sentinels in the gastrointestinal tract as they monitor intestinal content through taste receptors[78-80]. **Enteroendocrine** cells are specialized cells that produce and release hormones in response to a number of stimuli[81]. The hormones may be released into the bloodstream to generate systemic effects or may be distributed as local messengers. They may also stimulate a nervous response. Some of these hormones have key roles in the coordination of food digestion and absorption, insulin secretion and appetite. In addition, recent studies present them as orchestrators of intestinal inflammation due to interaction with microbial metabolites and release of cytokines in response to pathogen associated molecules and peptide hormone receptors[82]. **Paneth cells** are highly specialized secretory cells and located at the bottom of small intestinal crypts in the healthy intestine and play an important role in maintaining the stability of the intestinal tract. Paneth cells are interspersed between intestinal stem cells and can be distinguished by their columnar to pyramidal shape and by the presence of eosinophilic granules in their cytoplasm. These granules contain proteins and peptides to both modulate the microbiome and mediate inflammatory response. **M cells**, as mentioned above, are only present in the healthy small intestine above PP and ILFs and its main function is sampling of luminal microbes for mucosal immune surveillance.[83]

Three years ago, a new specialized colonocyte was described due to the raise of new technologies (see Section 3). **BEST4+ OTOP2+** cells, described by Parikh *et al.*, are absorptive lineage cells present in the healthy colonic mucosa and with a pH-sensing function.

2.1.2 Intestinal Epithelium alterations in IBD

The intestinal epithelium is a critical component of the gut barrier and plays a significant role in the development and perpetuation of IBD. Dysfunctions in intestinal epithelial cells, intrinsic molecular circuits that control the homeostasis, renewal, and repair of

intestinal epithelial cells, and disturbances in epithelial barrier integrity can trigger IBD (**Fig. 10**). Dysregulation within the epithelial layer increases intestinal permeability in both UC and CD, allowing the entrance of microorganisms that disrupt the intestinal immune homeostasis and are linked to the clinical disease course.

Epithelial cell death is one of the hallmarks of IBD, especially in UC, that contributes to the decrease in the secretion of defensins, antimicrobial peptides and mucus. This cell loss also contributes to the increase in permeability and entrance of microorganisms that help in the chronification of inflammation. In addition to epithelial cell loss, it is often possible to observe cell metaplasia in the colon of patients with IBD. This process involves the transformation of one type of cell into another type, and it is frequently observed in response to chronic inflammation in the gut. For example, Paneth-like cells are observed in the colon of IBD patients, which is a cell type only present in the small intestine in healthy conditions.

The mucus layer is also essential to prevent mucosal inflammation for its critical role in microbiota sequestration[84,85]. MUC2[84] encodes for a mucin protein secreted by goblet cells, and its depletion results in a diminished mucus layer, elevated levels of pro-inflammatory cytokines, and the development of spontaneous colitis in mice. In UC, the mucus layer is weaker due to alterations in mucin glycosylation and a reduction in the number of goblet cells. Epithelial permeability increases in both CD and UC[86-88], and apoptotic mechanisms have also been found to be altered in UC patients, affecting the intestinal epithelial permeability[89]. Distortion of crypts is also observed in the mucosa of IBD patients[89], which is characterized by irregularly arranged, dilated, branched and shortened crypts, with huge inflammatory cell infiltration between them.

Besides forming a tight protective barrier, epithelial cells are actively involved in the innate immune response. Many epithelial cells have pattern recognition receptors such as Toll-like receptors (TLRs) on the cell surface and NOD-like receptors in the cytoplasm that are essential in sensing bacterial products and initiating the immune response through the pro-inflammatory transcription factor nuclear factor-κB (NF-κB) to maintain homeostasis[90]. Defects in these epithelial cell-related processes might trigger mucosal inflammation.

Furthermore, the risk of developing IBD is associated with genetic variants related to epithelial functions, mainly those affecting bacterial clearance and autophagy. Mutations

in NOD2, which is an important intracellular sensor for bacterial muramyl peptide[91], are one of the strongest risk factors for IBD. Many cellular processes, such as autophagy[92], have been found to be altered in IBD patients. For example, patients with active UC showed a decrease in the expression of activating transcription factor 4 (ATF4), an important autophagy-related protein of the intestinal mucosa[93]. In CD, the mutation of ATG16L1 and IRGM, two important autophagy-related proteins, were highly related to the pathogenesis of the disease[94,95]. Autophagy appears to be involved in the pathogenesis and progression of IBD, and some autophagy regulators have been suggested as a possible treatment for IBD, although most of them have not yet undergone clinical development.

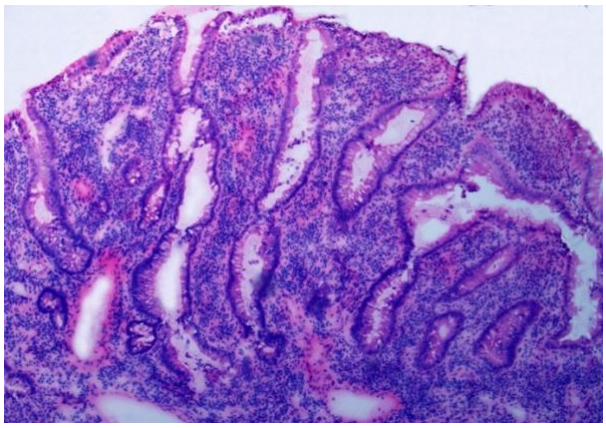


Figure 10. Epithelium in IBD. Hematoxylin-Eosin staining of CD inflamed intestinal mucosa. In addition to epithelial cell loos, crypts loss their structure and overall there is a loss of structure of the intestinal mucosa

2.2 Immune system

The immune system is composed of two main branches: the innate immune system and the adaptive immune system. The innate immune system is characterized by sensor cells that detect and recognize common patterns present in microbial agents (viruses, bacteria, fungi, and parasites) or when a tissue is damaged[96]. This branch of the immune system responds rapidly to maintain tissue homeostasis and protect the host. Cells of the **innate immunity**[97] in the colon include macrophages, dendritic cells (DCs), granulocytes (neutrophils, basophils, eosinophils and mast cells) natural killer (NK) cells, innate lymphoid cells (ILCs), and intraepithelial lymphocytes (IELs). These cell types differentiate from 2 different hematopoietic progenitor cells, the common myeloid progenitor that gives rise to the myeloid cells: granulocytes and monocytes (also to erythrocytes and platelets) and the common lymphoid progenitor that differentiate into several lymphoid cells including NK, ILCs and IELs from the innate immunity but also B and T cells from the adaptive immunity branch.

Macrophages and DCs, as well as epithelial cells, act as sentinels in the lining of the lamina propria expressing pattern recognition receptors (PRRs) such as Toll-like

receptors (TLRs) and NOD-like receptors located in the cellular membrane and the cytoplasm[98,99]. The PRRs recognize pathogen-associated molecular patterns (PAMPs) present in microorganisms and initiate an intracellular cascade that produces chemokines and cytokines, ultimately activating the transcription and production of pro-inflammatory mediators to ensure an effective innate response and elimination of the pathogen[100].

The **complement system** is a non-cellular component of the innate immune system. The complement system is made up of a variety of different plasma proteins that interact with one another to opsonize pathogens and trigger several inflammatory reactions that aid in the fight against infection. Most of the proteins and glycoproteins that constitute the complement system are present in plasma and are synthesized by hepatocytes in the liver[101]. However, in tissues like the intestine macrophages, monocytes or epithelial cells can also produce them. Several complement proteins are proteases that activate themselves despite there exist different pathways of complement activation. Overall, proteins of the complement protect against infection by opsonizing pathogens for phagocytes engulfment, act as chemoattractants to recruit more phagocytes and can damage certain pathogens by creating pores in their membranes.

Innate lymphoid cells (ILCs) are a recently discovered type of innate lymphocyte that are localized mainly in mucosa-associated tissues such as gut. ILCs can be divided into several subsets based on their cytokine production and transcription factor expression. Four major subsets of ILCs are present in the intestinal mucosa: ILC1, ILC2, ILC3 and **NK cells**. NK represent a group of CD3−CD56+ cells that can recognize and kill virus-infected cells and tumor cells. They also produce cytokines that help to activate other immune cells, including T cells and macrophages. Although NK cells are present in the gut, it remains to be defined if these cells are resident or derive from peripheral blood circulating NK cells.

DCs are an heterogenous population including cells from myeloid and lymphoid origin. Their main role is to act as specialized **antigen-presenting cells (APCs)** [102]that initiate and regulate the immune response by capturing, processing, and presenting antigens to T cells[103]. DCs are found throughout the body but are particularly abundant in areas where pathogens or foreign substances may enter, such as the skin, mucosal surfaces, and lymphoid tissues. They exist in an immature state until they encounter a foreign antigen, at which point they become activated and mature[104]. Once activated, DCs migrate to lymph nodes, where they present the antigen to T cells. This interaction activates the T

cells, which then proliferate and differentiate into effector cells that can eliminate the antigen.

Macrophages and B cells are also considered APCs. Macrophages phagocytose [105] microorganisms and B cells may specifically recognize antigens through the B-cell receptor (BCR) and both cell types will finally expose the antigens in their membrane surface through the major histocompatibility complex (MHC) (genes HLA, human leucocyte antigen). The complex MHC/antigen is presented by APCs to cells of the adaptive immune system to generate a specific response[102].

The **adaptive immunity**, on the other hand, is characterized by the ability to adapt to repeated encounters with an antigen and to generate a faster and more powerful response to the second encounter with the same antigen by generating memory. Cells of the adaptive immune response include lymphocytes such as T cells, B cells, and antibody-producing plasma cells.

T cells or lymphocytes T are produced in the bone marrow, and they mature in the thymus gland. T cells recognize and respond to specific antigens through the T cell receptor (TCR), which is unique to each T cell. Once a T cell has been activated by an antigen, it undergoes clonal expansion, producing large numbers of identical T cells that can recognize and respond to the antigen more effectively[106]. T cells are classified into several different subtypes, each with specific functions[107]. The most common types of T cells are helper T cells (CD4+) and cytotoxic T cells (CD8+). Helper T cells (Th) assist in the activation of other immune cells, such as B cells, macrophages, and other T cells, by producing cytokines and other signaling molecules. There are several subsets of CD4+ T cells that are present in the intestinal mucosa, including Th1, Th2[108], Th17, and regulatory T cells (Tregs) that have different functions. Th1 cells produce cytokines such as IFN-γ and are important for controlling intracellular pathogens, while Th2 cells produce cytokines such as IL-4[109] and are important for controlling extracellular pathogens. Th17 cells produce cytokines such as IL-17[110,111] and are important for controlling fungal and bacterial pathogens. Tregs help to maintain immune tolerance by suppressing the activity of other immune cells and are especially important in the intestinal mucosa, where they help to prevent excessive immune responses to harmless food antigens and commensal bacteria. CD8+ Cytotoxic T cells (CTL), also known as killer T cells, directly target and destroy infected or abnormal cells in the body, such as cancer cells or cells infected with viruses[112].

B cells are a type of lymphocyte that also play an important role in the adaptive immune response in the colon. They are produced in the bone marrow and can mature into a variety of different types, each with its own specific functions[113]. Apart of working as APCs, B cells differentiate into plasma cells, which are responsible for producing and releasing antibodies, which can recognize and neutralize specific antigens, including those derived from pathogens and commensal bacteria. **Plasma cells** in the intestine secrete Immunoglobulin A antibodies (IgA)[114], which is the isotype that predominates in that organ. B and plasma cells are scattered throughout the lamina propria but also present within isolated lymphoid follicles (ILFs) in the colon. ILFs are non-encapsulated small, organized clusters of immune cells (mostly T and B cells but also contain some DCs) that can be found scattered throughout the lamina propria. The encounter with antigen and differentiation to plasma cells (including generation of B memory cells) occurs in the germinal centers, which are microstructures located within the ILFs. The germinal centers contain follicular T helper cells (Tfh)[115], which are specialized T cells that help B cells in the germinal center reaction. Tfh cells are critical for the generation of high-affinity and long-lasting antibody responses[116,117].

It is important to note that not all microorganisms are pathogens, and the immune system in the intestine also interacts with dietary and commensal microbiota antigens. The balance between the generation of inflammatory and tolerogenic responses is crucial to maintain tissue homeostasis. However, in inflammatory bowel disease (IBD), the tolerance is overridden by the overreaction of innate and adaptive immunity towards common dietary and commensal microbiota antigens (see section 2.2.2 Immune system alterations in IBD below).

The results of this thesis focus on several types of myeloid cells, mainly neutrophils and macrophages. A more detailed introduction of these cell types in the healthy and inflamed colon is given in the subsection below.

2.2.1. Myeloid cells in homeostasis

- **Granulocytes**

Granulocytes are cells of the innate immune system characterized by the presence of granules in their cytoplasm. There are four specific types of granulocytes: neutrophils,

eosinophils, basophils and mast cells. In this thesis only neutrophils and eosinophils have been studied and thus are further introduced herein.

- **Neutrophils**.

Neutrophils are the most abundant cell type in human blood (40-70% of total blood cells). Neutrophils are rarely found in healthy intestinal mucosa; however they are one of the first wave of cells that cross the blood vessel wall to enter an injured or inflamed tissue. Different cytokines (i.e. IL-8, IL-17) stimulate neutrophils to leave the blood stream and enter into the tissue. Neutrophils bind to vascular endothelium through proteins called selectins (i.e. CD62L) and integrins and eventually neutrophils squeeze between endothelial cells (extravasate) that form the blood vessel[118,119]. Upon extravasation in tissue, neutrophils were believed to be eliminated fast and their lifespan to be very short. New results[120] in mice indicate that neutrophils are not immediately eliminated when they enter tissues but remain for different periods of time from hours to 1 day[121]. The main antimicrobial functions of neutrophils are phagocytosis, degranulation, and the release of nuclear material in the form of neutrophil extracellular traps (NETs). Pathogens are sequestered into NETs mixed with toxic molecules of the granules and they are effectively destroyed[122]. This results in neutrophil cell dead called "NETosis". Neutrophil's granules contain antimicrobial peptides (defensins and cathelicidins), myeloperoxidase (MPO), hydrolytic enzymes (lysozyme, sialidase, and collagenase), proteases (cathepsin G, azurocidin, and elastase), cationic phospholipase, and metal chelators (lactoferrin) that are released upon contact with microbes. Furthermore, they can produce huge amounts of reactive oxygen species and other toxic molecules that can destroy pathogens. Nicotinamide adenine dinucleotide phosphate (NADPH) oxidase, a multi-subunit enzyme complex formed on the membrane, produces reactive oxygen species in an oxidative burst upon granulocyte interaction with invading bacteria.

Heterogeneity in neutrophils has recently been reported[118]. In mice lungs, neutrophils acquired specific signature after extravasation, and it was dependent on the CXCR4-CCL12 axis[120].

- **Eosinophils**

Eosinophils are mainly present in blood, but can be found in the thymus, lower gastrointestinal tract, ovaries, uterus, spleen, and lymph nodes in healthy conditions. Eosinophils are cells with high granule content that play a role in combating pathogens

such as parasites, bacteria and viral infections. Moreover, a role in diseases like asthma or eosinophilic gastrointestinal disorders have been also attributed to them[123]. Eosinophils release hazardous inflammatory mediators that are both synthesized from scratch following cellular activation and retained in preformed vesicles[124]. Eosinophilic cationic protein, major basic protein, eosinophil protein X, eosinophil derived neuroendotoxin, and eosinophil peroxidase are the main proteins released by eosinophils[125]. These proteins exert their antimicrobial function inserting pores into the membrane of pathogens and increase smooth muscle reactivity by generating toxic oxygen radicals. Eotaxin and IL-5 are cytokines[126] that allow eosinophils to migrate to the intestine through the expression of CCR3 and IL-5 receptors.

- **Monocytes**

Monocytes are a type of myeloid cells formed in the bone marrow that circulate in the bloodstream for a short period of time before migrating into tissues where they differentiate into macrophages or dendritic cells[127-129]. Monocytes are larger than most other white blood cells and have a distinct kidney-shaped nucleus[130]. They play a key role in the immune response to infection and inflammation by phagocytosing or engulfing foreign particles such as bacteria, viruses, and dead cells. Monocytes also secrete cytokines, which are signaling molecules that activate other immune cells and coordinate the immune response[131]. Recent studies have also suggested that monocytes and macrophages in the intestinal mucosa may have additional functions beyond their classical roles in innate immunity[132]. For example, some studies have shown that macrophages in the intestinal mucosa can interact with and modulate the function of T cells. In addition, monocytes and macrophages may play important roles in maintaining tissue homeostasis[133] and repairing damage caused by inflammation or infection[134-136].

- **Macrophages**

Macrophages, first described by Élie Metchnikoff in 1883, are cells that can be found in all organs and have a dual origin. Initially, it was proposed that the presence and replenishment of macrophages in intestinal tissue was mainly due to circulating monocytes, that come from the bone-marrow (BM)[137]. That was called the monocyte "waterfall" in mice where CD14hi CCR2+ CD11Chi monocytes transit through intermediate populations, giving rise to mature CD14lo CCR2- CD11Clo macrophages [138]upon reaching the intestine through priming by local cues[139]. Nowadays, it has become

clear that most resident macrophages in the tissues, including the intestine, are seeded in tissues during development from the fetal liver and the yolk sac, and could be long-lived and self-maintaining cells[140] according to the last studies.

Macrophages in the gut are highly specialized phagocytes[105] that play a prominent role in combating pathogens by phagocytosis, presenting antigens to activate the adaptative immune response and participating in oral and microbial tolerance, thus, contributing to tissue homeostasis. In the last decade, research has uncovered the wound healing[141,142] and epithelial repair function of intestinal resident macrophages. These macrophages can be found distributed throughout the different layers of the gastrointestinal tract[143], including the lamina propria, submucosa, muscularis externa, and serosa. Macrophages are also highly plastic cells that can adapt their phenotype in response to signals from the extracellular environment, allowing them to participate in a wide range of physiological and pathological processes. Thus, each location within the tissue presents unique trophic factors, signaling molecules and physical scaffolding that modulate macrophage phenotype and function[140,144]. For instance, lamina propria macrophages help to preserve the epithelial layer integrity by phagocytosing apoptotic cells and debris. Macrophages in the submucosal plexus help in the modulation of motility and blood flow. Also, there are neuron associated macrophages in the submucosa and the muscularis[144].

Activated macrophage phenotypes have been classified into M1 and M2 polarizations, each related to inflammation or homeostasis[145,146]. M1 macrophages are considered pro-inflammatory, with robust phagocytic and cytotoxic activity and abundant release of pro-inflammatory cytokines. M2 (alternatively activated or anti-inflammatory) are phagocytic, pro-angiogenic and have a function on the resolution of inflammation, tissue repair and fibrosis[147]. Compared to M1, M2 macrophages are more functionally diverse, and several subtypes have been described in mice[148] and can be also generated *in vitro* from human blood monocytes[149] or using monocyte cell lines[150]. These phenotypes in mice have been named M2a, M2b, M2c and M2d and express different combinations of cytokines, chemokines and growth factors. Macrophages depend on different cytokines to survive such as colony stimulating factor-1 (CSF1 or M-CSF) or CSF2 (GM-CSF), expressing the corresponding receptors CSF1R and CSF2R[151,152]. M1 and M2 macrophages are susceptible to be generated in-vitro and are identified by specific set of surface markers and production of cytokines[147]. M2 macrophages are dependent of CSF1 (M-CSF) and M1 macrophages are polarized trough GM-CSF (CSF2) and usually

LPS[153,154]. This initial subdivision has been extremely useful to unravel macrophage polarization, but advances in transcriptomic and proteomic have shed light and revealed further heterogeneity within and between different tissues. For instance, a late[155] study with cutting-edge technologies have been able to uncover a novel population of CD121b+ macrophages in mice located closely to the tips of the villi thereby near apoptotic cells of the epithelium and intestinal microbes.

In the results of this thesis, macrophage populations in the healthy and inflamed colon are revisited in the light of new technologies (scRNAseq) which allowed us to identify previously undescribed states of these cells.

2.2.2. Immune system alterations in IBD

The immune cell composition of the gut is dramatically altered in IBD. One of the disease mechanisms, is the huge recruitment of immune innate cells to the intestinal mucosa. Studies have demonstrated increased levels of monocytes in the blood and inflamed tissues of patients with IBD[156], suggesting that they may be involved in the initiation and perpetuation of the inflammatory response. Monocytes in the inflamed intestinal mucosa have been also found to produce pro-inflammatory cytokines such as tumor necrosis factor-alpha (TNF-α), interleukin-1 beta (IL-1β), and IL-6, which contribute to the development of chronic inflammation in IBD. Additionally, the increased presence of M1 macrophages in IBD has been argued because of the increase of monocyte recruitment from blood. It has been discussed that inflammatory monocytes give rise to M1 macrophages, which are associated with inflammation and tissue damage. In a study by Bain and colleagues[139] using the dextran sulfate sodium (DSS)model of colitis in mice found that proinflammatory colonic macrophages were derived from inflammatory monocytes. They also observed a marked increase in proinflammatory colonic macrophages within 24 hours after commencing the DSS regimen, and this increase persisted for 5 days. Moreover, in human, a study argue that intestinal macrophages do not proliferate and that the exclusive source of intestinal macrophages are blood monocytes[157]. M1 macrophages, like monocytes, are characterized by their high expression of pro-inflammatory cytokines such as TNF-α, IL-1β, and IL-6, and their ability to generate reactive oxygen and nitrogen species, promoting tissue damage and perpetuating inflammation[158].

Interestingly, CD is characterized by the presence of granulomas, a unique feature found in approximately one-fourth of CD patients[159]. Granulomas are aggregates of macrophages that form in response to chronic inflammation and are a hallmark of CD. These macrophages fuse together to form multinucleated giant cells, which are a characteristic feature of granulomas[160]. The formation of granulomas is thought to be a result of the body's attempt to contain the inflammatory response in CD[161]. However, the role of granulomas in the progression of the disease is not fully understood. Some studies suggest that granulomas can contribute to tissue damage and fibrosis, while others propose that they may help to limit tissue damage by isolating the inflammatory response[162]. Further research is needed to fully understand the complex relationship between granulomas and the pathogenesis of CD.

In addition to the significant recruitment of monocytes to the inflamed mucosa, another hallmark of IBD is the massive infiltration of neutrophils[163]. These cells contribute to mucosal injury and debilitating disease symptoms[164]. The excessive formation of NETs in IBD can contribute to the exacerbation of inflammation and tissue damage[165,166]. However, recent studies have revealed that NETs can also play a role in maintaining tissue homeostasis and reducing bleeding in intestinal ulcers. Under specific pathological conditions such as IBD, neutrophils can cross the intestinal epithelium layer into the lumen, a process known as cryptitis[167,168]. This can contribute to the inflammatory response and lead to further tissue damage[169,170]. Interestingly, cryptitis is mostly observed in UC[171]. In IBD, higher numbers of eosinophils are also found in active patients and it is thought they have roles in tissue destruction, inflammation, diarrhea and formation of fibrosis[172].

Additionally, mast cells have also been implicated in the pathogenesis of several intestinal diseases, including IBD[173,174]. In IBD, mast cells are activated by various stimuli, such as bacteria and dietary antigens, and release pro-inflammatory mediators (TNF-a, histamin, serotonin…)[175], leading to tissue damage and inflammation. These mediators can induce various effects, including smooth muscle contraction, increased vascular permeability, and recruitment of immune cells to the site of inflammation. Mast cells can also promote epithelial cell damage and dysfunction through the release of proteases, trypytases and other mediators, leading to increased intestinal permeability and bacterial translocation.[176]

DCs are instrumental in maintaining the gut homeostasis through a balance in the induction of T effector and Tregs. However, DCs derived from CD patients decreased production of IL-10 in response to LPS compared to healthy donors[177]. In IBD, DCs are overactivated by various stimuli, such as microbial products, and increase expression of TLRs and activation/maturation marker CD40[178]. In this context, they secrete pathological relevant cytokines such as IL-6 and IL-12, which activate T cells, leading to the production of pro-inflammatory cytokines such as TNF-α and IFN-γ and ultimately favoring T effector over Treg induction.

Numerous researchers have implicated ILCs in the pathogenesis of IBD[179-181], For instance, Li *et al*[182] described increases IFN-γ-producing-CD127+ ILC1s in the inflamed gut and noted that this expansion was correlated with disease severity. Additionally, studies have shown that NK cells are present in increased numbers in the intestinal mucosa of patients with IBD, particularly in areas of active inflammation. They contribute to tissue damage in IBD through the release of cytotoxic molecules, such as perforin and granzymes, and by promoting inflammatory cytokine production. However, there is also evidence that NK cells can have a protective role in IBD by regulating the balance between pro-inflammatory and anti-inflammatory immune responses. For example, NK cells can produce the cytokine IL-22, which helps to maintain the integrity of the intestinal barrier and promote tissue repair.

In the context of IBD, the innate immune cells respond to microorganisms and epithelial cell death, leading to a significant production of inflammatory cytokines and increased antigen presentation, which in turn promotes the activation of adaptive immune cells such as T cells. For instance, DCs and M1 macrophages secrete TNF, IL-23, and IL-12, leading to the activation of the effector T cell populations Th1 and Th17, which in turn secrete inflammatory s IFNg and IL-17 respectively, perpetuating inflammation. Conversely, it has been well-documented that in IBD, there is an expansion of Tregs in the intestinal mucosa.

In IBD, there is also an increase in the number, diameter, and density of isolated lymphoid follicles (ILFs). This *de novo* structures are called tertiary lymphoid follicles and have been broadly described in other peripheral tissues and intestine under chronic inflammatory conditions[183,184]. These structures serve as organized sites of immune response induction. In this setting, the production of antibodies becomes biased towards

immunoglobulin IgG (IgG) instead of the IgA isotype[185]. This shift in isotype production is one of the main characteristics of IBD[186].

2.3. Stroma

2.3.1. **Stromal cells in homeostasis**

Stromal (or mesenchymal) cells are mesoderm-derived non-epithelial, non-hematopoietic structural cells that make up a significant portion of the colonic lamina propria. Stromal cells provide structural support of organs, synthetizing and remodeling the extracellular matrix (ECM), and regulate development, homeostasis, and repair of tissues. Several cell types conform the stromal compartment of the intestinal mucosa: fibroblasts, α smooth muscle actin (α-SMA) expressing myofibroblasts, endothelial cells, pericytes and glia[187].

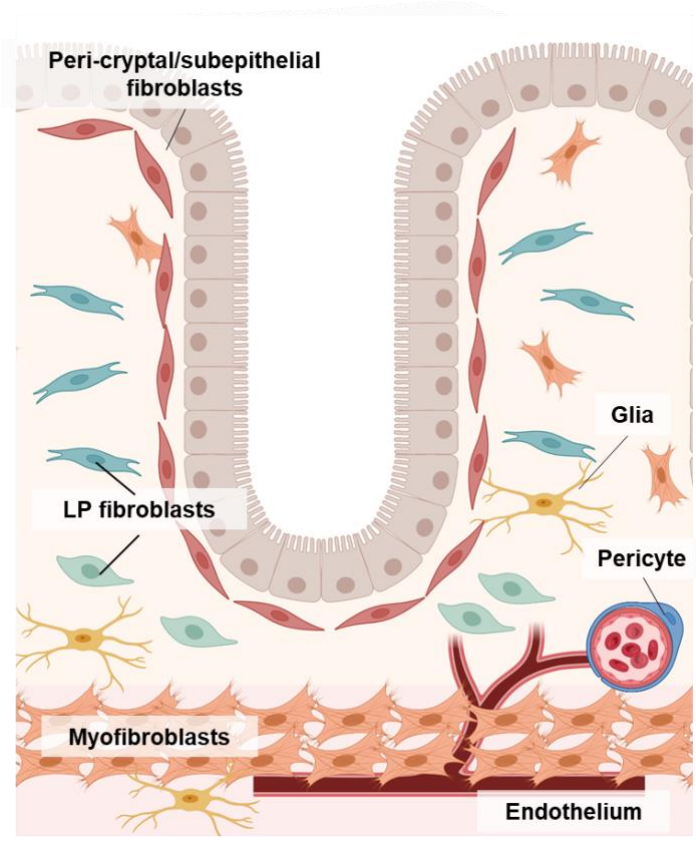


Figure 11. Stromal cell populations present in the healthy colonic lamina propria. Created with BioRender.

Endothelial cells form a single cell layer that lines all blood and lymph vessels and regulates exchanges between the bloodstream and the surrounding tissues. Endothelial cells are the major constituent of the microvasculature that is present in the lamina propria and submucosa. The presence of the endothelium in the intestinal mucosa is essential since provides a selectively permeable exchange barrier and allows recruitment and circulation of immune cells.

Myofibroblasts are contractile fusiform cells that are mostly differentiated from fibroblasts due to the expression of α-SMA[188]. Myofibroblasts are capable of speeding wound repair by contracting the edges

of the wound[189]. These cells are dispersed through the lamina propria, also next to the epithelium, and form the layer that separates the lamina propria from the submucosa. It has been previously described that myofibroblasts are activated fibroblasts[188]. Fibroblasts differentiate into myofibroblasts to participate in the repair process[190].

Pericytes are mural cells that support blood vessel structure and function, along with perivascular smooth muscle cells[191].

Glial cells are astrocyte-like cells that are the only neuron-type cell in the intestinal mucosa. Enteric glial cells maintain active signaling mechanisms with neurons. This signaling modulates gastrointestinal reflexes. In addition, bidirectional communication between glial and immune cells contributes to gastrointestinal immune homeostasis[192]. Enteric glial cells present different locations: mucosal, submucosal and in the myenteric plexus. Mucosal glia also presents a role in the maintenance of the integrity of the epithelium[193].

Fibroblasts are heterogeneous non-hematopoietic, non-epithelial and non-endothelial cell type present in all organs and tissues. Fibroblasts respond to tissue or epithelial damage by secretion of soluble mediators and ligands and play an important role in the remodeling of the extracellular matrix to support intestinal structure and integrity. In specific niches as the colonic crypts, they also support the stem cell maintenance[194] and regulate the balance between self-renewal and differentiation through gradients of specific Wnts and bone morphogenic proteins (BMP). In the establishment of immune responses against microorganisms, fibroblasts play an important role training leucocytes in secondary lymphoid organs (SLO) such as intestinal lymph nodes, ILFs or Peyer's Patches (PP) and maintaining the specialized microdomains of B and T and DCs areas in those structures[195]. Indeed, a specific fibroblast population has also been described in SLO termed fibroblastic reticular cells (FRCs) [196]which produce trophic factors and maintain the network that connects sampling areas with the T cell area, allowing the diffusion of chemokines, antigens and antibodies through high endothelial venules among other functions. Retinoic acid expressed by fibroblasts in the lamina propria of the intestine mediates the functional education of mucosal dendritic cells[117,197]. Depending on the location and local cues, fibroblasts can perform different specialized functions.

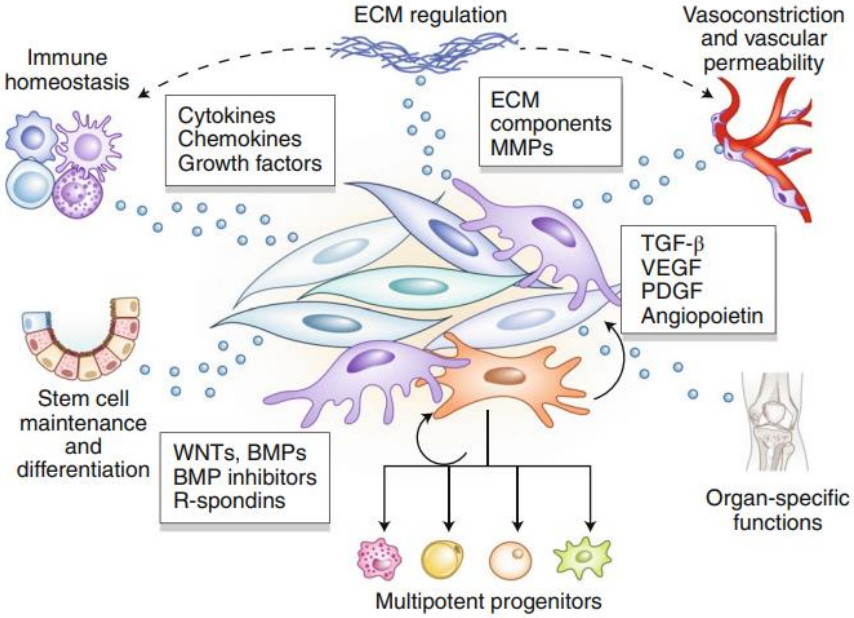


Figure 12. Stromal cell diverse functions. Stromal cells have diverse functions such as ECM regulation, stem cell maintenance and differentiation and immune homeostasis among others. Figure from *Koliaraki, V. et al Nat Immunol 2020*.

Regarding the colonic lamina propria, different subtypes of fibroblasts have been described according to location. The presence of a fibroblast sheet under the epithelium was observed more than 20 years ago and epithelial-mesenchymal interaction are thought to be critical for the formation of the basement membrane of the intestinal mucosa[198]. Other fibroblasts have been observed that were distributed along the lamina propria.

Recent studies using single-cell transcriptomics have revealed unexpected heterogeneity in the intestinal fibroblast compartment in healthy and disease conditions, specifically in UC. This heterogeneity is dependent on the distinct transcriptional expression profile and location of each subset.

In healthy colon, Kinchen *et al* confirmed the presence of a subepithelial/pericryptal fibroblast population, now called Stromal 2 (S2 fibroblasts), a nomenclature that has begun to be used after this publication. As a novelty, two S2 subtypes of fibroblasts were observed by Kinchen and colleagues: S2a and S2b. Lamina propria fibroblasts were heterogeneous and transcriptionally different from α-SMA+ myofibroblasts and were named Stromal 1 (S1) and Stromal 3 (S3) bearing different transcriptional profiles.

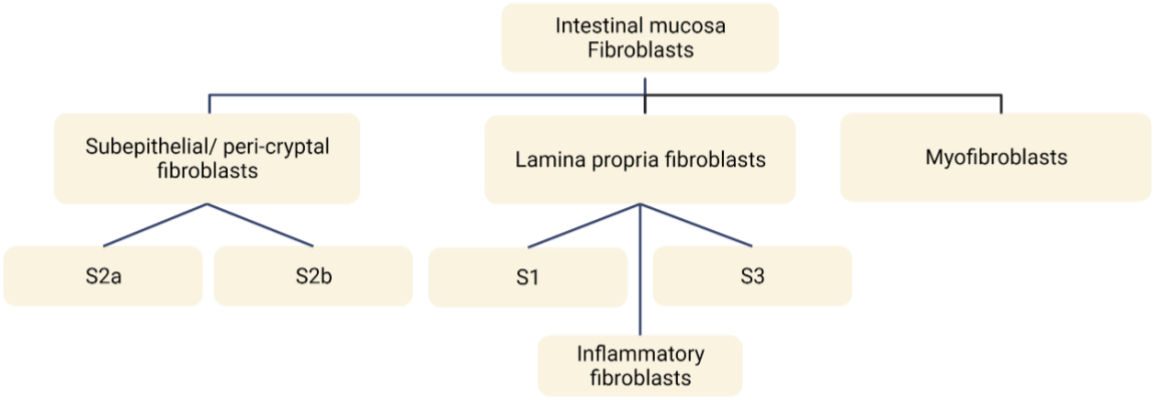


Figure 13. Scheme of new characterization of fibroblasts populations in the intestinal mucosa. Classification according *Kinchen, J. et al. Cell 2018*.

The heterogeneity observed at the single-cell level not only in the colon but in other tissues, could indicate dynamic states influenced by local cues at a given time. However, there is still lack of knowledge of how many fibroblasts subpopulations exist and their main functions, mostly in IBD.

2.3.2 Stromal alterations in IBD

While the immune cells present in the gastrointestinal tract have been the main focus of research in IBD so far, recent insights suggest that stromal cells also have a crucial role in IBD pathogenesis. These cells, which are present in the intestine, regulate both intestinal epithelial and immune cell homeostasis. As such, stromal cells are increasingly recognized as important players in the complex interplay of cellular and molecular factors that drive the development and progression of IBD[199]. Altered stromal cell behavior, such as increased fibroblast activation and collagen deposition in the intestinal mucosa of IBD patients have been observed[200]. These changes can lead to tissue fibrosis and contribute to disease progression. Intestinal fibrosis is a severe complication characterized by a range of symptoms, including disruptions in regular peristalsis and impaired nutrient absorption, as well as the excessive buildup of extracellular matrix (ECM) components. The cells accountable for such ECM deposition have been described to be fibroblasts and myofibroblasts. In addition, the formation of strictures and scar tissue can lead to thickening of the intestinal lumen, further exacerbating disease symptoms. Human colonic lamina propria fibroblasts have altered migratory capacity in IBD, as shown by in vitro studies where fibroblasts from IBD patients exhibited reduced migratory capacity

compared to healthy intestinal fibroblasts[201]. Fibroblasts from inflamed intestines in CD or UC also proliferated faster and produced more collagen in vitro than those from healthy individuals, potentially explaining the increased risk of fibrosis in IBD patients[202].

However fibroblasts are not only involved in detrimental roles in IBD. Restoration of the damaged epithelium in IBD requires fibroblast migration, collagen deposition, and controlled rebuilding of the epithelial layer.

Previously, fibroblasts were considered a homogeneous cell population, but emerging evidence indicates that fibroblasts include diverse cell types based on developmental origin, anatomic location, and function. Specifically, in UC patients Kinchen *et al.* identified the presence of a novel subset of fibroblasts which presented unique inflammatory markers. The study revealed the appearance of an activated fibroblast subset (S4) that was found to be significantly expanded in patients with ulcerative colitis. This particular subset was identified by its expression of various cytokines and molecules, including IL-6, MHC class II invariant chain (CD74), IL-33, and homeostatic cytokines such as CCL19 and CCL21, which play a role in recruiting and retaining lymphocytes specially in the organized SLO structures. Additionally, another study[203] observed a significant expansion of inflammation-associated fibroblasts in the inflamed tissues of certain IBD patients. These fibroblasts expressed IL-24, IL-11, and IL-13RA2, indicating a potential involvement in both inflammation and fibrosis. Notably, the researchers found that these fibroblasts exhibited high levels of expression of the oncostatin M receptor (OSMR), and that the expression of OSM correlated with a lack of response to anti-tumor necrosis factor (anti-TNF) therapy in patients. In a distinct study, Martin et al. detected pericytes, vascular smooth muscle cells, and two subtypes of fibroblasts among patients with CD. One of the identified fibroblast populations exhibited an activated phenotype, with high expression levels of CD90 and PDPN (podoplanin), as well as cytokines such as IL-6 and IL-11. This population also expressed chemokines that attract neutrophils (CXCL2, CXCL8, CXCL1, and CXCL5) and recruit monocytes (CCL2 and CCL7). Smillie et al also identified one fibroblast population, termed inflammation-associated fibroblasts, that was expanded in inflamed tissue of UC patients and showed enrichment for genes like IL-11, FAP, and IL-13RA2. However, the mechanism in which fibroblasts acquire an inflammatory phenotype in IBD remains understudied.

Additionally, in the context of IBD, the endothelium undergoes remarkable changes in response to inflammatory cytokines and growth factors, allowing an increased

recruitment of immune cells[204]. Moreover, there is a microvascular expansion called angiogenesis, that allows new generation of endothelium due to the inflammatory response[204]. Interestingly, Smillie *et al*[205], observed an activated cluster of endothelium expressing the molecule *ACKR1* that was expanded in CD patients.

SECTION 3: MOLECULAR TECHNOLOGIES APPLIED TO THE STUDY OF INFLAMMATORY BOWEL DISEASE

Different techniques have been used in the past years to study the intestinal mucosa and IBD. Our understanding of IBD and cell players has improved mostly due to advances in cellular and molecular technologies.

Polymerase chain rection (PCR) was the first way to study gene content of cells, allowing to amplify millions of copies of a specific segment of DNA. PCR dates to the mid 80's, when the Human Genome project started. This technology allowed the scientific community to study gene expression, specifically RNA messenger (mRNA), which codes for proteins.

In the late 80's, the **microarray** technique appeared for protein expression and in early 90's for gene expression. cDNA microarrays is a method for studying the expression profiles of protein-coding genes. The advantage of microarrays is that it involves the high-throughput detection of many small RNAs simultaneously, which is impossible to achieve using conventional PCR. Microarrays are composed of hundreds to thousands of probes bound to a solid surface allowing simultaneous detection of thousands of nucleic acid targets. Genes expressed in samples of interest are detected after fluorescent labelling and hybridization to the corresponding probes on the microarray. Although its limitations (only known probes for known genes are added) microarrays have been a widely used tool to study the transcriptome of the mucosa of IBD patients. The goal was identifying disease biomarkers and predicting response to treatment. In addition, it has been used extensively in blood to identify similar biomarkers or predictors. The use of microarrays has enabled researchers to study UC, CD, and various regions of the intestine[206,207]. In a study performed in our lab[208], we revealed through this technique that the intestinal mucosa of remitted patients was different than the inflamed mucosa of active patients or healthy controls. Moreover, we demonstrated that the intestinal mucosa of a patient in remission is different from a non-IBD control. Another study from our lab using

microarray technology showed that the transcriptomic profile of CD patients treated with anti-TNF antibodies downregulated a pool of genes related to inflammation even in the patients who fail to achieve endoscopic remission[209].

RNA sequencing (RNA-seq) is a high-throughput technique which uses next generation sequencing (NGS) meaning uses massive parallel sequencing and examines the quantity and all the sequences of RNA present in the tissue analyzed. Not using previously defined probes allows the discovery of novel RNA sequences and variants, as opposed to microarrays. In the field of IBD, RNA-seq has been also used to predict response to treatment, evaluate differences between patients with different degree of disease, and tried to distinguished CD and UC at the transcriptomic level[210,211]. We know from RNA-seq studies that hundreds (even a couple thousand) of genes can be regulated within the intestinal mucosa of both CD and UC during an active flare and after a specific treatment[212].

"Omics" techniques such as microarrays and RNA-seq have revolutionized biology and medicine but are commonly applied to heterogeneous tissues or bulk sorted cells, providing a low-resolution blended picture of the transcriptional status of the sample under study. This approach does provide information on thousands of transcripts from just a few milligrams of tissue but averages the expression of thousands or even millions of cells. Moreover, specificity on what cell type is upregulating or downregulating those genes, is lacking. Even sorted and seemingly homogeneous cell populations expression patterns are considerably heterogeneous, due to both intrinsic stochastic processes and extrinsic factors, such as the surrounding microenvironment. Indeed, cells within the same microenvironment can have different transcript profiles due to either deterministic causes (i.e. cell cycle effects) or random factors.

The most prevalent cell types seen in the intestinal mucosa in both healthy and inflammatory conditions have been identified and defined utilizing a variety of methods. In fact, immunostaining, flow cytometry, fluorescence-activated cell sorting, *in situ* hybridization (detects mRNA), and more recently mass cytometry (CyTOF) analysis show that IBD patients have significantly altered numbers and phenotypes of immune cells (including T, B lymphocytes or macrophages and dendritic cells), in addition to stromal and epithelial cells[213]. Moreover, have also allowed us to know that cells such as neutrophils and monocytes are actively recruited into the inflamed intestinal mucosa and help trigger inflammattion and activation of other cells.

However, due to the usage of described, well-established cell markers, all this data has a low resolution and it is highly biased by the current knowledge, making it difficult to discover novel subsets or comprehend how each fraction contributes to the disease. In fact, despite significant advancements in our understanding and treatment of IBD, we still have very little understanding of how each distinct cell subset contributes to disease and how it is controlled by various therapeutic approaches**.** Overall, cellular characterization remains surprisingly limited.

Despite these limitations, whole tissue transcriptomic approach helped us to understand disease pathogenesis, to identify potential targets, to discover the mechanisms of action of therapies and to segregate patients based on disease activity[214,215]. Nonetheless, whole tissue transcriptomics cannot reliably identify markers of disease progression/phenotype or predictors of response. We cannot even consistently distinguish at a molecular level colonic inflammation in UC and CD, even though these two diseases have marked micro and macroscopic unique treats. Understanding tissue composition and behavior requires high resolution techniques such as single-cell omics, developed in the last decade, that can help explain cell composition and how cells are organized and connected to build tissues in health and disease to fully grasp cell-cell tissue interactions and behaviors.

Single cell RNA-seq (scRNA-seq) is a genomic approach for the detection and quantitative analysis of mRNA molecules but differs from RNA-seq that can detect the mRNA content of each cell in a biological sample[216,217]. scRNA-seq has emerged as an indispensable tool to dissect the cellular heterogeneity of tissues and offers a great potential for *de novo* discovery. Unlike bulk RNA-seq, the crucial and first step of scRNA-seq is that tissue needs to be dissociated into a single cell suspension. This step is usually performed by enzymatic digestion but it can also be achieved by laser microdissection. The enzymatic digestion needs to be quick and effective to preserve the RNA intact, as the conditions of this step directly affect the molecular profiles of the cells (damage depends also on the cell type). Once a single cell suspension is achieved, there exist different methods to perform scRNA-seq: genomic, epigenomic and transcriptomic[218,219]. Moreover, different single-cell platforms have been developed, the most popular are C1, Chromium, and SMART-seq (all transcriptomic methods). Depending on the purpose of the study and the studied tissue, researchers may select the best option. In this thesis, Drop-seq has been used, which is a microdoplet-based

technology[220] which enables the analysis of thousands of cells per sample?(Chromium from 10x Genomics; see Methods section?).

In a recent attempt to fill the gap of human cell heterogeneity and provide a gold-standard reference for future studies, The Human Cell Atlas was born as an international effort to obtain a comprehensive census of all human cell types profiled by scRNA-seq and then locate them within tissues and in relationship to other cellular types providing an actual "atlas" of the human body at cellular and molecular resolution. The results of such an ambitious endeavor would take a few more years to be available but the applications and amount of information that will derive will probably take decades to completely understand.

We know now from recent studies in complex tissues, that scRNA-seq can reveal the existence of previously unknown cell populations and assign gene expression signatures to specific cell types dramatically increasing the resolution of the picture obtained. Having such high definition of the transcriptional contributions of each individual cell population can resolve new and known cellular subsets, reveal unique markers and identify relevant changes in biological processes within each cell type, including organ development, cell transformation, tissue regeneration, aging, chronic inflammation, infection, etc.

In the 5 past years, several studies in healthy and IBD intestinal tissue using scRNA-seq have been performed and have shed light into the cell composition of the intestine. The first study to publish data on human intestinal cell types using scRNA-seq was Kinchen et al and studied the mesenchymal content of human gut and compared it to that of UC patients, revealing a colitis associated population of fibroblasts (inflammatory fibroblasts) that expresses factors contributing to epithelial cell dysfunction and inflammation. Moreover, single-cell census of the colonic mesenchyme reveals also in healthy individuals unexpected heterogeneity. A previous study in mice focused on small intestine epithelial cells using an extensive scRNA-seq survey of over 50,000 murine cells[221]. In both studies, scRNA-seq revealed new subtypes within, up to now, well-characterized populations. Furthermore, in a subsequent study, Kinchen et al studied the epithelial compartment at the single cell level and found an undescribed colonocyte: BEST4+OTOP2+ absorptive epithelial cells that have a function in regulating pH, which has recently been linked to a proper function of some cytokines[222]. Martin et al revealed through scRNAseq a common cell module present in a group of ileal CD that did not

respond to a-TNF treatment. Smillie et al generated an atlas of UC where they also found inflammatory fibroblasts and BEST4+ epithelial cells along with other cells that also associate to unresponsiveness to a-TNF treatment.

More specifically, Corridoni *et al*[223] described the different CD8 T cells present in UC patients and Huang et al used scRNAseq together with immune phenotyping and gene risk analysis to reveal distinct and common signatures between pediatric onset CD and UC as well as high infiltration of TNF+ macrophages and a decrease of CD39+ intraepithelial lymphocytes in disease[224]. Thanks to the finding, targeting of those pathways allowed them to restore immune homeostasis in a pilot study. Moreover, a study revealed the cell composition of the different lineages of the human intestinal tract from the developing gut to the adult one[225].

In summary, many studies in the gastrointestinal tract and in other diseases such as cancer or other tissues have uncovered new cell types and interactions through scRNAseq, allowing the discovery of new and minority cell populations playing a role in disease and thus contributing to the identification of novel therapeutic targets and to the development of new therapies.

However, scRNAseq has also its limitations. First, it requires high level and specialized knowledge of computational analysis due to the amount of data generated. Moreover, although we are able to dissect cell composition of tissues, and new analysis technologies have been developed to study cell-to-cell communication and cell interactions, we still lack the spatial information which is crucial to the understanding of a tissue and near interacting partners, especially in disease. Histology and immunostaining such as immunohistochemistry give us a tool to validate new cell populations found by scRNAseq and locate them spatially. However, sometimes there is a gap between mRNA and protein expression and novel cell populations are difficult to locate. Moreover, the increasing availability of these new technologies has not been accompanied by a rapid and effective creation of antibodies. Hence, spatial technologies at single cell level are nowadays at the forefront. Spatial transcriptomics are a pool of methods for assignment of cell types (based on their content of mRNA) to their locations in the histological sections. These techniques allow to decipher both the cellular and subcellular architecture in both tissues and individual cells by translating the scRNA-seq information to a tissue section by cutting-edge computational analysis. Until now, to locate cells through its content in mRNA in a tissue was performed through different techniques of Fluorescent in situ

hybridization (FISH), where different probes can be applied to the same tissue to locate cell populations, limited by fluorochrome and probe availability. Vanguard technologies have allowed to perform an in-situ captured RNA directly in formalin-fixed tissue allowing to extract from the slide the RNA content of each region (10x Visum)[226].

However, this new technology developed in 2006 does not allow to extract and locate the RNA content in a single-cell way, but in regions, giving bulk-like results but more localized. Thus, in 2021 the Nanostring company developed high-plex imaging of RNA at subcellular resolutions: the high spatial molecular imaging CosMx SMI[227]. CosMx SMI performs multiple cycles of nucleic acid hybridization of fluorescent molecular barcodes generating super high-resolution images relying in cell segmentation in order to have single-cell resolution. Moreover, antibodies to several cellular types can be combined.

Overall, combining scRNA-seq with cutting edge spatial technologies has allowed to find and locate new cell populations and interactions that could have an important role in disease. **Figure 14**

OBJECTIVES

Despite the large advances in the field, including the development of numerous new drugs, lack of response to different lines of treatment is a frequent and major problem when managing IBD. We hypothesized that the unpredictability and heterogeneity in disease behavior is supported by the different composition and activation of intestinal cells across individuals. To understand these heterogeneity, bulk RNA sequencing of intestinal tissue has been widely used in IBD, including ulcerative colitis and Crohn's disease. While these studies significantly increased our understanding of some of the mechanisms and genes relevant for these diseases, bulk analysis lacks the resolution to capture differences between disease cell phenotypes. Therefore, the aim of this thesis was to uncover the different mechanisms leading to and maintaining an inflammatory state using single-cell technologies and provide a characterization of the intestinal mucosa in health and disease. Additionally, the determination of cell intrinsic molecular changes in patients could lead to the identification of new therapeutic targets and, potentially, seed the basis for a cell specific and fine-tuned drug development.

For that purpose, the specific objectives of this thesis were:

- To optimize single-cell RNA sequencing (scRNA-seq) to the study of intestinal biopsies from IBD patients and healthy controls.
- To describe the cellular types and their transcriptional profile present in the colonic mucosa of IBD and healthy controls.
- To determine the differences in the proportion of intestinal cell subsets and their transcriptional signatures within ulcerative colitis and Crohn's disease patient using scRNA-seq.
- Identify and characterize cellular populations subjected to heterogeneity under inflammatory conditions.
- Spatially localize cellular populations and interactions of interest using single-cell transcriptomics

MATERIALS & METHODS

Patient recruitment and sample collection

Colon samples were collected to perform scRNAseq, generate organoid cultures and analyze by flow cytometry on fresh tissue. Additional samples were fixed in formalin and paraffin embedded (FFPE) for CosMx™ SMI and tissue staining (IHC IF, ISH). For scRNAseq and flow cytometry, colon biopsies from active areas of UC and CD patients were collected during routine endoscopies performed as standard of care. Healthy controls were individuals undergoing endoscopy for colorectal cancer screening and presenting no signs of dysplasia or polyps at the time of endoscopy. For SMI analysis, surgical colon resections were obtained from non-IBD controls (patients undergoing surgery for colorectal cancer), UC and CD patients (undergoing colonic resective surgery). Organoid cultures were established exclusively from surgical samples of non-IBD controls. In non-IBD controls samples were obtained from non-tumor areas and for UC and CD patients from involved inflamed areas. Blood samples from HC, CD and UC were also collected to perform flow cytometry analysis of granulocytes. The study was approved by the Ethics Committee of Hospital Clinic Barcelona (HCB/2018/1062 and HCB/2022/0125) and the Hospital Mutua de Terrassa (CI201901). All patients signed an informed consent at the time of colonoscopy or before surgical intervention.

Figure 15. Overview of work-flow communication between gastroenterologists and scientists in IBD studies.

Human colonic cell isolation

Biopsies (n=4-6 per patient) were taken from involved areas of the colon of UC and CD patients with signs of endoscopic activity, placed immediately in cold Hank's Balanced Salt Solution (HBSS) (Gibco, MA, USA) and kept at 4ºC until processing (<1 h). Colonic biopsies from non-IBD controls were collected from the sigmoid colon and processed in the same way. Freshly collected biopsies were washed with 5mM DTT (Roche, Spain) in HBSS for 15 min and then washed in complete medium (CM) (RPMI 1640 medium

(Lonza, MD, USA) supplemented with 10% heat-inactivated fetal bovine serum (FBS) (Biosera, France), 100U/ml penicillin, 100 U/ml streptomycin and 250 ng/ml amphotericin B (Lonza), 10µg/ml gentamicin sulfate (Lonza) and 1,5mM Hepes (Lonza)) for 10 minutes. Both incubations were performed at room temperature in a platform rocker. Biopsies were chopped with a scalpel and placed into tubes containing 500 µl of Digestion Solution (CM + Liberase TM (0.5 Wünsch units/ml) (Roche, Spain) + DNase I (10 µg/mL) (Roche, Spain)) and incubated on a shaking platform for 1h at 250 RPM and 37ºC. After incubation biopsies were filtered through a 50-µm cell strainer (CellTrics, Sysmex, USA), washed with Dulbecco's Phosphate Buffered Saline (PBS; Gibco, USA) and resuspended in RPMI medium supplemented with 0.05% of Bovine serum albumin (BSA) at a concentration of ~0.5-1·10^6 cells/mL for scRNA-seq and in FACS buffer (PBS + 2% inactivated FBS (fetal bovine serum) + NaN3 0.1%) for flow cytometry analysis[228].

10x library preparation and sequencing

Following digestion, 10x Genomics 3′ mRNA single-cell method was used. Approximately 7,000 cells were loaded onto the Chromium10x Genomics platform (10x Genomics, CA, USA) to capture single cells, following the manufacturer's protocol. Generation of gel beads in emulsion (GEMs) (10x Genomics, CA, EEU), barcoding and GEM-reverse transcription was performed using the Chromium Single Cell 3′ and Chromium Single Cell V(D)J Reagent Kits (10x Genomics, CA, EEU) (user guide, no. CG000086) according to manufacturer's instructions. Full-length, barcoded cDNA was amplified by PCR to generate enough mass for library construction (Nextera® PCR

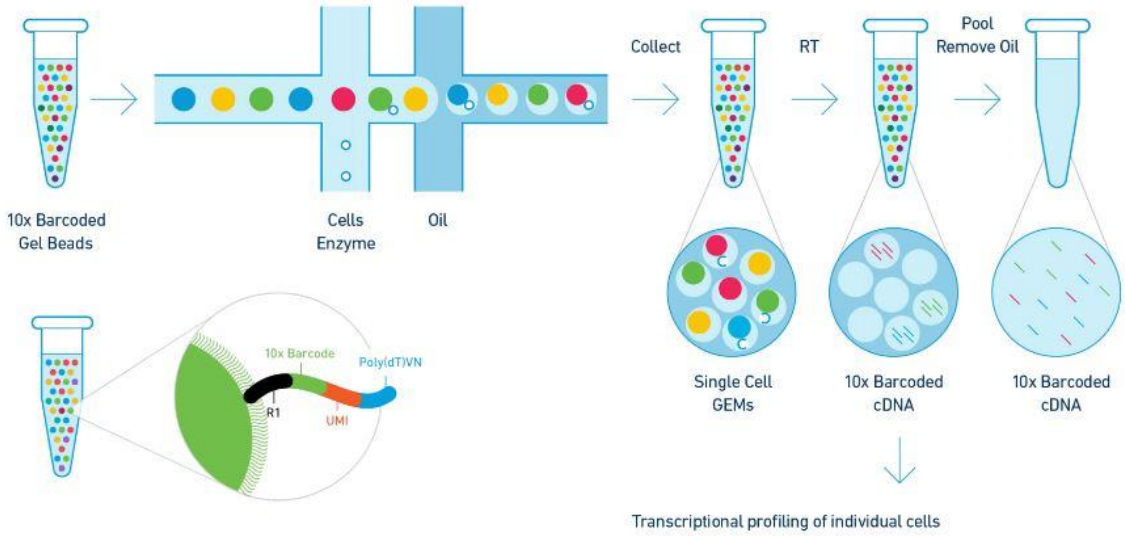


Figure 16. Chromium 10x Genomics work-flow for single-cell library preparation and sequencing

primers) (Illumina, CA, USA). Sequencing of the libraries was performed on HiSeq2500 (Illumina, CA, USA).

Single cell data analysis

Data Processing

For each sample, sequences obtained in fastq files were processed with CellRanger's count pipeline using the default parameters (10XGenomics, v. 3.1.0). This pipeline performs an alignment based on the reference genome (Gencode release 27, assembly GRCh38 p10), filtering, barcode counting, and UMI counting. The resulting filtered matrix was analyzed using R (v.4.2.0). We merged the count matrices retrieved from CellRanger using the function merge from the SeuratObject R package (version 4.0.2). At this point, doublets were assessed using the scDblFinder [229]R package and removed. Total analyzed cells: 47,600.

Initially, we analyzed healthy control and IBD samples separately to assess the similarity of cell types and samples. We processed and annotated the objects separately and we assessed for similarity by using Jaccard index and label transferring (see detailed methods in here). Samples were then pooled together in the same object. Low-quality cells were then filtered out based on mitochondrial RNA percentage and number of genes per cell. Epithelial cells required a less stringent filter of 65% of counts aligned to the mitochondrial genes for quality control. A total of 46,700 cells were considered for the analysis. Then, we logarithmically normalized, obtained the highly variable genes, and scaled the counts (default parameters) of each data set using Seurat [230](version 4.1.0). Principal component analysis (PCA) was performed[230]. Dimensionality reduction was performed by applying the Uniform Manifold Approximation and Projection (UMAP) algorithm using the optimal number of PCs. UMAP also served as a two-dimensional embedding for data visualization. Cluster analysis was performed using the Louvain clustering algorithm.

We used marker gene expression to categorize and split cells in the main five subsets to analyze them separately. EPCAM, AQP8, BEST4, MUC2, OLFM4, PLCG2, TRPM5, ZG16 for epithelial cells; CD79A, BANK1, CD19, DERL3, MS4A1, MZB1 for B and

cells; CD3D, CD3E, CD3G, CD8A, FOXP3, GZMA, GZMB, IL17A, NKG7, TRBC1 for T cells; ACTA2, ADAMDEC1, CHI3L1, COL3A1, NRXN1, PVALP, SOX6, VWF for stromal cells; AIF1, C1QA, C1QB, CD14, CMTM2, FCGR3B, LYZ, MS4A2, TPSAB1, TPSAB2 for myeloid cells. We then separated the merged data into 5 main cell types: Epithelial cells, Lymphoid cells (T cells and innate lymphoid cells), Myeloid cells (mast cells, macrophages, dendritic cells, monocytes, neutrophils, and eosinophils), Stromal cells (endothelial cells, fibroblasts, pericytes and glia), and B and Plasma cells. Immunoglobulin (IG) genes were removed from all the main cell types except B and plasma cells to reduce background noise.

Identification of cell types

Each main cell type was re-processed starting from FindVariableFeatures using the same procedure as for the whole dataset (see full code in our github page and our webpage). During the process, we further removed CD3, C1QA, DERL3, or MS4A1 expressing cells from Epithelial cells; CD3, C1QA, or EPCAM-expressing cells from the B cells; CD3, C1QA, DERL3, MS4A1 or EPCAM-expressing cells from the Stromal cells; DERL3, MS4A1, C1QA, EPCAM, or CD79A-expressing cells from the Lymphoid cells; and CD3, THY1, DERL3 or MS4A1-expressing cells from the Myeloid cells. We then systematically re-clustered each cell category. The annotation of each subcluster from the main cell type was defined by the marker genes, which were obtained by the FindAllMarkers function with the default threshold except for the min.pct parameter, which was set as 0.25, and the thresh.use, which was set to 0.25.

Batch correction

We observed batch effect by sample in the UMAPs of Lymphoid (tcells), B and plasmacells (plasmas), and Stromal cells (stroma) objects while re-processing for cell type identification. For those cell types, we ran Harmony[231] (version 0.1.0) to correct for this effect. Specifically, we used the RunHarmony function, with the optimal number of principal components found for each subset as latent space and the sample of origin as batch label. Seurat's analysis pipeline was applied again to each object, followed by UMAP generation using default settings and harmony integrated space.

Annotation of cells

Epithelial cell subset (EPCAM positive): Within the absorptive cells, we found two clusters of colonocytes sharing the expression of AQP8, FABP1 and SLC26A2, and a cluster of inflammatory colonocytes with low AQP8 expression and expressing the inflammatory markers DUOXA2, DUOX2 and LCN2. PLCG2 colonocytes express PLCG2, HES1 and ELF1, while the Laminin enterocytes present expression of the laminin gene LAMA3 and the Ephrin receptor EPHA2. Last, a cluster of BEST4 colonocytes (BEST4 and OTOP2) and Tuft cells (TRPM5, SH2D6 and POU2F3) were identified. Within the secretory cells, we found two clusters of goblet cells (MUC2, TFF3 and SPINK4), a canonical and a mature cluster expressing TFF1 and FER1L6. Secretory progenitor cells express RETNLB, CLCA4 and HEPACAM2. A cluster of Paneth-like cells was identified in the inflamed samples expressing the defensin genes DEFA5 and DEFA6 and REG3A. Enteroendocrine cells were identified by CHGA, CHGB and NEUROD1. Cells with high expression of ribosomal genes (RPS19, RPS18, RPL35) clustered together and were annotated as Epithelium Ribhi. While LGR5+ stem cells did not form a cluster, we identified cycling transit-amplifying cells by the expression of MKI67, TOP2A and PCNA.

Stromal cell subset: fibroblasts were annotated based on the nomenclature of Kinchen et al.[232] Three clusters of S1 lamina propria fibroblasts were identified, all sharing the expression of ADAMDEC1, CP and FABP4. Submucosal S3 fibroblasts express OGN, CCDC80 and GREM1. Peri-cryptal S2 (SOX6, F3) fibroblasts subclustered into the bottom crypt S2a (VSTM2) and the top-crypt S2b (NPY, NRG1) clusters. Moreover, we also found a cluster of Rib^{hi} fibroblasts (RPL7A, RPL28), a cluster of MT fibroblasts (with high expression of mitochondrial genes) and an IER fibroblasts cluster, characterized by the expression of immediate-early response genes (FOS, FOSB, IRF1). Fibroblastic reticular cells, FRCs, were identified by the expression of CCL19 and CCL21, and myofibroblasts by the expression of SOSTDC1, ACTG2 and MYH11. Last, inflammatory fibroblasts (also known as S4) were identified by the expression of IL11, CXCL5, FAP, INHBA and IL24, among others.

Endothelial cells express VWF, PECAM1 and PLVAP. In inflammation, a cluster of activated endothelium was identified expressing ACKR1 and the selectins SELE and SELP. We further identified a small lymphatic endothelium cluster expressing MRC1, MMRN1 and LYVE1.

Finally, pericytes were characterized by the expression of NOTCH3 and RGS5; and the markers NRXN1, S100B and CDH19 identified a cluster of glial cells.

B and plasma cell subset: Plasma cells (DERL3, MZB1, XBP1) could be subdivided based on whether they are IgA, IgG or IgM producers, and whether they are plasmablasts, PB, ($PRDM1^{low}$) or fully differentiated plasma cells, PC ($PRDM1^{high}$)[233]. We found some clusters expressing in an exclusive manner IGLC3 and IGLC2, which we classified as a Lambda cluster, and a plasma cell cluster expressing IGLL5, IGLC7 and IGLL1, annotated as PC IGLL5. Moreover, two clusters of IgA plasma cells expressing the heat-shock genes HSPA1B, HSPA1 and HSPB1 were annotated as PC IgA heat shock. As in other subsets, cells expressing immediate-early genes (FOS, JUN and FOSB) were annotated as PC IER.

Unlike plasma cells which loss its expression through differentiation, B cells were identified by the expression of CD20 (MS4A1). Within the B cell lineage, we identified a cluster of memory B cells (ITGAX, FGR and PDCD1), naive B cells (IGHD and FCER2) and germinal center (GC) B cells (SUGCT, MME and HRK). Besides, B cells with high expression of ribosomal genes were annotated as B cell Ribhi. Last, two clusters of cells in a proliferative state were identified, annotated as cycling cells (TUBB, TOP2A, MKI67) and including a mixture of B and plasma cells.

T cell subset: T cells were determined by the expression of T cell receptor component CD3D, CD3E and CD3G. CD8 T cells had expression of CD8A, CD8B, or both, and CD4 cells were $CD4^{low}$. Withing the CD4 cells, CD4 naïve cells were identified by the expression of CCR7, LEF1 and SELL. CD4 $ANXA1^{234}$ cluster express ANXA1, IL7R and GPR186, and could represent a resident phenotype. Tregs were identified by the expression of FOXP3, IL2RA, CXCR6 and BATF, among others. T follicular helper cells, ThF, express CXCL13 and MAGEH1. While we did not observe a canonical Th17 population, the CCL20 cluster represent a heterogeneous population with inflammatory related genes that probably include Th17 cells (CCL20, RORA, KLRB1, ETPD1)[235]. S1PR1 T cells correspond to a circulating phenotype (Central Memory Cells) and express CCR4, KLF2 and S1PR1 among others. Within the CD8 cells, CD8 cytotoxic effector cells were identified by the expression of GZMK and KLRG1, while CD8 TRM lack expression of GZMK but express other GZM genes, KLRB1, SPRY1 and IGTA1. CD8 FGFBP2 corresponds to cytotoxic cells described previously in ulcerative colitis[223], which expresses FGFBP2, FCGR3A, S1PR5 along with cytotoxic genes (GZMH,

NKG7). The MAIT cluster is $CD8A^+$ and were identified by the expression of NRC3. DN TNF is a cluster of double negative T cells ($CD8^-CD4^-$) expressing immediate-early response genes (FOS, JUN and FOSB) phenotype and the inflammatory cytokine TNF. DN EOMES is a cytotoxic effector DN cluster (CD8-CD4-) characterized by the expression of the transcription factor EOMES and GZMK. GD IEL correspond to a cluster of intraepithelial lymphocytes with expression of the T receptor genes TRDC and TRGC1, which is further identified by the expression of FCER1G, KLRC2, CD160, KLRD1, KLRC1, ITGA1 and KIR2DL4[236]. Regarding the innate lymphocytes, we observed a cluster of ILC3 (LIF, KIT and PCDH9) and NK (expressing KLRF1 and lacking expression of CD3G, CD3D, CD8A and CD4).

Myeloid cell subset: The myeloid cell subset includes macrophages, monocytes, dendritic cells, mast cells eosinophils and neutrophils. Macrophages were identified by the expression of CD68, CD14 and complement coding-genes (C1QA, C1QB). While M2 and M2.2 macrophages express CD209, CD163L1 and FOLR2, M0 macrophages lack the expression of M2 markers but express macrophage marker genes (CD68, C1QA, C1QB and so on). The IDA subset was identified by the expression EGF ligands (NRG1 AREG, EREG and HBEGF). M1 ACOD1 and M1 CXCL5 shared the expression of TNIP3, IL1B, INHBA, IL6, VCAN, CD300E but express ACOD1 and CXCL5, respectively, in an exclusive manner[147]. Unlike M1, inflammatory monocytes were identified by higher expression of VCAN, CD300E and CD14, and the expression of FCN1 and S100A9. We identified two clusters of mast cells (TPSB2 and TPSAB1), one of them expressing LTC4S. We further identified eosinophils by the expression of CLC, IL4 and IL13. All neutrophil clusters were identified by the expression of PROK2, CMTM2, CXCL8, FCRG3B, AQP9, S100A8 and S100A9. Finally, two clusters of dendritic cells were identified (LAMP3). On one side, DCs CD1c express CD1c, TRL10 and TCTN3, while DCs CCL22 were identified by the expression of CCL22 and CCL19.

Differential abundance testing

We tested for differences in cell abundances between myeloid colonic cells from IBD patients and healthy controls using the *miloR* package (v 1.4.0)[237]. Specifically, we performed differential abundance testing of healthy patients and both inflammatory chronic diseases, CD and UC, separately. To do so, the dataset was subsetted to include

myeloid cells only (677 healthy cells, 1935 CD cells and 1161 UC myeloid cells). For each analysis, a k-nearest neighbors (KNN) graph was constructed using the graph slot from the adjacency matrix of the previously processed *Seurat* object and cells were assigned to neighborhoods (*k=20, d=30*). To leverage the variation in the number of cells between UC, CD and healthy samples, the cells belonging to each sample in each neighborhood were counted and the Euclidean distances between single cells in a neighborhood were calculated. Differential abundance testing in a generalized linear model framework was then performed between IBD samples and healthy controls with default parameters. In addition, to check if the differences in abundances are particularly strong for a certain cell type, we assigned to each neighborhood a specific annotation label considering most cells belonging to that neighborhood. For those neighborhoods where less than 80% of cells shared the most abundant label, a "mixed" label was assigned. Neighborhoods were grouped using the *groupNhoods* function with *max.lfc.delta=5, overlap=3* for CD and healthy samples, and *max.lfc.delta=3, overlap=2* for UC and healthy samples.

Trajectory analysis

We applied the Monocle 2 algorithm (v 2.24.1)[238] to order myeloid cells in pseudotime to indicate their developmental trajectories. Such pseudotime analysis is a measure of progress through biological processes based on their transcriptional similarities. With that aim, we selected all distinct macrophages and monocytes and created a *CellDataSet* object using a negative binomial model. We ran Monocle 2 using the 2000 most highly variable genes selected with *Seurat* (v 4.1.0) and default parameters of Monocle after DDRTree dimension reduction and cell ordering. To visualize the ordered cells in the trajectory, we used the *plot_cell_trajectory* function to plot the minimum spanning tree on the cells. The starting point of the trajectory was set manually to M0 macrophages.

Publicly available scRNAseq datasests

Data from Smillie et al.[205] and Martin et al.[239] were downloaded and reanalyzed following the same protocol used for our samples. We separated in silico the macrophage (CD14, CD163, CSF1R, and CD68) compartment to annotate them. In addition to manual annotation, we used the MatchScore2 package[240] which compares the marker gene list with a reference dataset (in this case our dataset) and assigns a value (Jaccard index) to

each comparison. High Jaccard indexes indicate high similarity between clusters. See detailed code here.

Gene-sets from microarray analysis of stimulated macrophages were obtained from the analysis of GSE121825, GSE155719, GSE156921, GSE161774 and GSE94608 datasets[241-244]. Using these gene-sets, we calculated the average expression levels of each cluster on our dataset at single cell level, using the function *AddModuleScore* of the *Seurat* R Package. See full code here.

CosMx™ Spatial Molecular Imaging (SMI)

CosMx™ technology [CosMx™ Human Universal Cell Characterization RNA Panel (1000-plex) + 20 custom genes; Nanostring, USA] was applied to 9 FFPE samples: 3 non IBD healthy controls, 3 CD and 3 UC with a mean of 19 fields of view (FoVs) per sample. The mean area explored was 11,53 mm^2 per sample. The following cell surface markers were used for morphology visualization: B2M/CD298, PanCK, CD45, CD3 antibodies and DAPI.

CosMx™ Spatial Molecular Imager (SMI) sample preparation

FFPE tissue sections were prepared for CosMx™ SMI profiling as described in He et al.[227] Briefly, five-micron tissue sections mounted on VWR Superfrost Plus Micro slides (cat# 48311-703) were baked overnight at 60°C, followed by preparation for in-situ hybridization (ISH) on the Leica Bond RXm system by deparaffinization and heat-induced epitope retrieval (HIER) at 100°C for 15 minutes using ER2 epitope retrieval buffer (Leica Biosystems product, Tris/EDTA-based, pH 9.0). After HIER, tissue sections were digested with 3 µg/ml Proteinase K diluted in ACD Protease Plus at 40°C for 30 minutes.

Tissue sections were then washed twice with diethyl pyrocarbonate (DEPC)-treated water (DEPC H2O) and incubated in 0.00075% fiducials (Bangs Laboratory) in 2X saline sodium citrate, 0.001% Tween-20 (SSCT solution) for 5 min at room temperature in the dark. Excess fiducials were rinsed from the slides with 1X PBS, then tissue sections were fixed with 10% neutral buffered formalin (NBF) for 5 min at room temperature. Fixed samples were rinsed twice with Tris-glycine buffer (0.1M glycine, 0.1M Tris-base in DEPC H2O) and once with 1X PBS for 5 min each before blocking with 100 mM N-

succinimidyl (acetylthio) acetate (NHS-acetate, Thermo Fisher Scientific) in NHS-acetate buffer (0.1M NaP, 0.1% Tween pH 8 in DEPC H2O) for 15 min at room temperature. The sections were then rinsed with 2X saline sodium citrate (SSC) for 5 min and an Adhesive SecureSeal Hybridization Chamber (Grace Bio-Labs) was placed over the tissue.

NanoString® ISH probes were prepared by incubation at 95°C for 2 min and placed on ice, and the ISH probe mix (1nM 980 plex ISH probe, 10nM Attenuation probes, 1nM SMI-0006 custom, 1X Buffer R, 0.1 U/µL SUPERase•In™ [Thermo Fisher Scientific] in DEPC H2O) was pipetted into the hybridization chamber. The chamber was sealed to prevent evaporation, and hybridization was performed at 37°C overnight. Tissue sections were washed twice in 50% formamide (VWR) in 2X SSC at 37°C for 25 min, washed twice with 2X SSC for 2 min at room temperature, and blocked with 100 mM NHS-acetate in the dark for 15 min. In preparation for loading onto the CosMx SMI instrument, a custom-made flow cell was affixed to the slide.

CosMx Spatial Molecular Imager (SMI) instrument run

RNA target readout on the CosMx SMI instrument was performed as described in He et al. Briefly, the assembled flow cell was loaded onto the instrument and Reporter Wash Buffer was flowed to remove air bubbles. A preview scan of the entire flow cell was taken, and 15-25 fields of view (FoVs) were placed on the tissue to match regions of interest identified by H&E staining of an adjacent serial section. RNA readout began by flowing 100 µl of Reporter Pool 1 into the flow cell and incubation for 15 min. Reporter Wash Buffer (1 mL) was flowed to wash unbound reporter probes, and Imaging Buffer was added to the flow cell for imaging. Nine Z-stack images (0.8 µm step size) for each FoV were acquired, and photocleavable linkers on the fluorophores of the reporter probes were released by UV illumination and washed with Strip Wash buffer. The fluidic and imaging procedure was repeated for the 16 reporter pools, and the 16 rounds of reporter hybridization-imaging were repeated multiple times to increase RNA detection sensitivity.

After RNA readout, the tissue samples were incubated with a 4-fluorophore-conjugated antibody cocktail against CD298/B2M (488 nm), PanCK (532 nm), CD45 (594 nm), and CD3 (647 nm) proteins and DAPI stain in the CosMx™ SMI instrument for 2 h. After

unbound antibodies and DAPI stain were washed with Reporter Wash Buffer, Imaging Buffer was added to the flow cell and nine Z-stack images for the 5 channels (4 antibodies and DAPI) were captured.

Segmentation and quality control

Images were segmented to obtain cell boundaries, assign transcripts at the cell-level, and obtain a transcript by cell count matrix[227]. Cells with an average negative control count greater than 0.5 and less than 20 detected features were filtered out. After quality control, a mean of 95,3% of cells across samples and FoVs was retained, corresponding to ~46,160 cells per sample on average.

Preprocessing and feature selection

Following library size normalization and log-transformation using logNormCounts (scater package)[245], highly variable genes (HVGs) were identified with modelGeneVar (scran package)[246], blocking by sample (i.e., variance modelling is performed per sample and statistics combined across samples), and selecting for genes with a positive biological variance component. For comparability with the lowest-coverage sample, per-sample scaling normalization with multiBatchNorm from the batchelor package was applied[247].

Integration and dimension reduction

Next, Principal Component Analysis (PCA) was computed on highly variable genes (HVGs). Upon inspection of the corresponding elbow plot (% variance explained vs. # of PCs), we selected the first 30 PCs as input for integration using *harmony*[231], and for dimensionality reduction via uniform manifold approximation and projection (UMAP)

Cell type annotation

To identify subpopulation markers, we ran *findMarkers* (*scran* package)[246], blocking by sample, considering HVGs only, and testing for positive log-fold changes. The 100 top ranked genes were selected and passed to *SingleR*[248] along with the reference scRNA-seq dataset's count matrix and subpopulation assignments, for label transfer. To obtain higher-resolution annotations, we grouped cells into 5 biological compartments according to their lower-resolution label (namely: epithelial, myeloid, plasma, stroma, and T cells)

and, for each compartment separately, reperformed marker detection and label transfer as described above.

Co-localization analysis

To quantify how pairs of cell types co-localize in a given sample and FoV, we computed, for every sample, FoV, and pair of cell types, two-dimensional kernel density estimates (KDEs) of cell coordinates using *kde*. Within each sample-FoV, estimation was performed over a rectangular window according to the boundary coordinates of cells from a given pair of types, and the Pearson correlation coefficient between KDEs was computed. Only comparisons with more than 10 cells from both types were taken into consideration.

Data and code availability

Expression profiles and raw single-cell RNA sequencing data of all cell compartments are publicly available in GEO: GSE214695 (reviewer token: **qrovseekbbwjjkj)**. Analyzed data can be explored in our interactive webpage: https://servidor2-ciberehd.upc.es/external/garrido/app/ (username **reviewer**, password **qrovseekbbwjjkj**). Full code of the analysis is in our github page https://gitfront.io/r/user-1871750/StAFev2VqZh6/ibd-bcn-single-cell/ and our main webpage https://servidor2-ciberehd.upc.es/garrido.html

CosMx SMI data analysis pipeline available at: https://gitfront.io/r/user-1871750/qH1qRCB32SDH/CosMx-SMI/. The CosMx SMI data will be made available upon publication.

Granulocytes isolation from blood

Human granulocytes were isolated from blood by a double density gradient. In brief, diluted blood in PBS (1/2) was layered over Lymphoprep™ (1.077 g/ml) that was layered over denser Polymorphprep™ (1.113 g/ml). The double gradient was then centrifuged at 500g for 30 min obtaining 2 separated cell layers. The lower layer containing the granulocytes was collected, washed and red blood cells were lysed using a commercial lysis buffer (BioLegend). Purity of granulocytes achieved with this method is >95%.

Flow cytometry

Blood granulocytes and single-cell suspensions obtained from digested biopsies of heathy donors and IBD patients were stained with the following antibodies CD66b PE-Cy7, CD16 PerCP, CD62L APC-Cy7, CD69 BV421, CD193 FITC, CD63 FITC (BioLegend) and CXCR4 PE (R&D systems) and Zombie Aqua Fixable Viability Kit (BioLegend) for death cells. Cells were fixed using BD Stabilizing Fixative [BD], acquired using a BD FACSCanto II flow cytometer (BD) and analyzed with FlowJO software (BD).

Organoid cultures from human colonic crypts

Epithelial 3D organoid cultures were generated using surgical samples from adult healthy donors, as previously described[249]. Ex vivo Organoids embedded in Matrigel (BD Biosciences, CA, USA) were passaged for further expansion approximately every 5-6

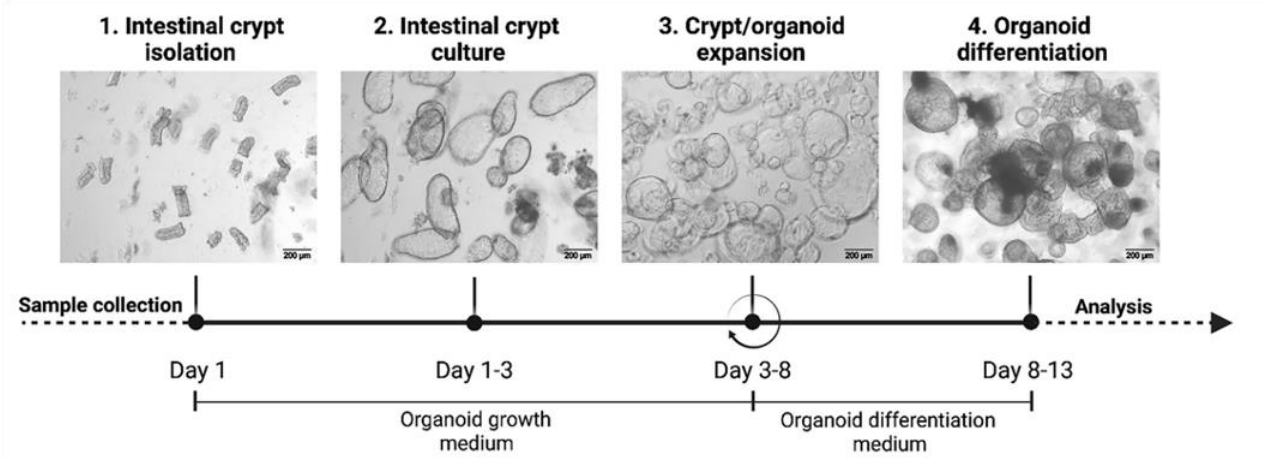


Figure 17. Schematic representation of the experimental workflow from sample collection to 3D organoids generation and differentiation. Running the first round of expansion (i.e., from crypts to organoids and their differentiation) takes 13 days on average. Step 3 can be repeated several times (↻) depending on the experimental plan. Once generated, both stem cell-enriched and differentiated cultures can be used for several downstream applications. Figure from *I, Dotti; PlosOne 2022*

days. After 1-3 passages, organoid samples were stimulated with neuregulin 1 (R&D, Germany) at 1 ug/mL for 48h and harvested for RNA extraction.

RNA isolation and quantitative RT-PCR

Organoid samples were harvested in Trizol (Ambion, Thermo Fisher Scientific, USA) and RNA was isolated using GeneJET RNA Cleanup and Concentration Micro Kit (Thermo Fisher Scientific, USA). Biopsies were placed in RNAlater RNA Stabilization Reagent (QIAGEN, Hilden, Germany) and stored at –80ºC until RNA isolation. RNA

was isolated using an RNeasy Kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions and sequenced for bulk RNA-seq. Total RNA from organoid samples were transcribed to cDNA using reverse transcriptase (High Capacity cDNA Archive RT kit, Applied Biosystems, Carlsbad, CA, USA) Quantitative real-time PCR (qPCR) was performed in an ABI PRISM 7500 Fast RT-PCR System (Applied Biosystems) using predesigned TaqMan Assays (Applied Biosystems) (OLFM4, LRG5 and MKI67).

Bulk RNA-seq

Barcoded RNAseq libraries were prepared from total RNA using a TruSeq stranded mRNA kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. Libraries were subjected to paired-end sequencing (50 bp) on a HighSeq-3000 platform (Illumina, CA, USA) at the Translational Medicine and Genomics group (Boehringer-Ingelheim GmbH & Co, Biberach, Germany). Quality filtering and adapter trimming was performed using Skewer version 2.2.8 Reads were mapped against the human reference genome using the STAR aligner version 2.5.2a. The genome used was GRCh38.p10, and gene annotation was based on Gencode version 27 (EMBL-EBI, Hinxton, UK). Read counts per gene were obtained using RSEM version 1.2.31 and the Ensembl GTF annotation file (EMBL-EBI, Hinxton, UK). Analyses were performed using the R (version 3.2.3) statistics package. HC and IBD patient samples used are described in Supplementary Table 1.

Statistical analysis

Graphs of bulk RNA-seq, qPCR organoid and flow cytometry results and statistical analysis was performed using Graphpad Prism 9.0 (Graphpad Software, CA, USA). Differences between groups in the bulk RNA-seq data were tested using Ordinary one-way ANOVA test and correcting for multiple comparisons by controlling the false discovery rate [two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli]. Graphs show median. For organoid results, One-sample t test was performed. Data is shown as fold change (FC) relative to the vehicle treated condition. Bars represent mean and standard deviation (SD). Differences between groups in flow cytometry data were tested using the non-parametric Kruskal–Wallis test and correcting for multiple comparisons by controlling the false discovery rate [two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli]. Graphs indicate median and range.

Immunohistochemistry (IHQ) and immunofluorescence (IHF)

Immunostaining of tissue sections of healthy donors and IBD patients was performed using commercially available antibodies (anti-rabbit CD68 (Sigma, MO, USA), anti-mouse CD209 (Santa Cruz Biotechnologies, TX, USA), anti-rabbit OLFM4 (Cell Signaling, MA, USA), anti-rabbit MPO (Sigma, MO, USA), anti-mouse MBP (Sigma, MO, USA)). Deparaffinization, rehydration and epitope retrieval of the samples was performed using PT Link (Agilent, CA, USA) using Envision Flex Target Retrieval Solution Low pH (Dako, Germany). Blocking of the samples was performed using horse or goat serum (10-20%) (Vector, NY, USA) in a PBS + 0,5% BSA solution. Secondary antibodies were used for IHQ (goat anti-rabbit and horse anti-mouse (Vector, NY, USA) and IHF (goat anti-rabbit AF488 and goat anti-mouse Cy3). For IHQ the immunoperoxidase detection system was used (Vector, NY, USA). DAPI (Thermo Fisher Scientific, USA) counterstaining was performed on IHF samples. Image acquisition was performed on a Nikon Ti microscope (Japan) using Nis-Elements Basic Research Software. Immunofluorescence composite was performed using ImageJ software.

Single molecule RNA *in situ* hybridization (ISH)

ISH was carried out on FFPE tissues fixed in 10% neutral buffered formalin for 48 hours. RNA probes for NRG1 (cat#311181) and OLFM4 (cat#311041) and RNAscope 2.5 HD assay – Red kit (Biotechne, CO, USA) were used according to the manufacturer's instructions. Pre-warmed probes were added to the slides and incubated in the HybEZ oven (Biotechne, CO, USA) for 2 hours at 40ºC. After a 6-step signal amplification, for double IHQ and ISH samples, IHQ protocol was followed. Tissue sections were counter-stained with Gill's hematoxylin. Slides were mounted with ECOmount mounting medium (Biocare medical, CA, USA) and photographed using Nikon Ti microscope (Japan) and Nis-Elements Basic Research Software

RESULTS

SECTION 1: Cellular colonic map of health and Inflammatory Bowel Disease at the single cell level

The intestinal mucosa is a complex environment comprised of a diverse array of cell types, both distinct and related. To better understand the cellular landscape in health and disease, we employed single-cell RNA sequencing (scRNA-seq) technology. This approach allowed us to decipher the cellular composition of the intestinal mucosa with unprecedented resolution and to unravel the heterogeneity of cell populations within this tissue.

Integration of single-cell RNA sequencing and spatial molecular imaging analysis provides a map of healthy and inflamed colon

We applied scRNA-seq analysis in colonic biopsies of HC (n=6), CD (n=6) and UC (n=6) active patients (**Fig. 18**; **Annex Table 1**) getting a total of 46,700 cells and identified 79 different clusters, whose proportions varied significantly between groups (**Fig 19** and **Fig 20a**). We grouped the 79 clusters in five different compartments classified as epithelium, stroma, B and plasma cells, T cells and myeloid cells. Each compartment presented a distinct transcriptomic profile (**Fig 19**). We isolated each compartment in *silico* to achieve higher resolution on cell populations and explore them in depth. Analysis of differentially expressed genes (DEGs) detected cluster-specific marker genes with adjusted p-values (Methods). DEGs for each cluster were used to identify cell types and annotate subpopulations. Within the different cell compartments, we found some cell populations that were only present in the IBD samples that will be discussed in this thesis. Overall,

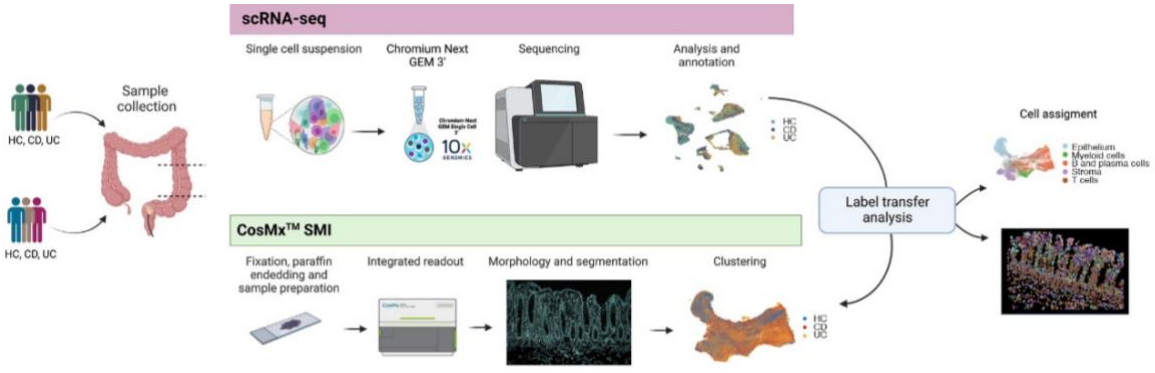


Figure 18. Integration of single cell RNA sequencing (scRNA-seq) and Spatial Molecular Imaging (SMI) provides a map of healthy and inflammatory bowel disease (IBD) colonic biopsies. Overview of the study design for scRNA-seq, CosMx™ SMI and label transfer from scRNA-seq annotations to the SMI dataset. Two cohorts of Colonic samples including active Crohn's disease (CD), active ulcerative colitis (UC) and healthy controls (HC) were processed by scRNA-seq (n=18 samples) and CosMx™ SMI (n=9 samples). See Annex Table 1 for details

by proportion analysis, we observed a decrease of the epithelium compartment and an increase in the myeloid and B and plasma cell compartment in the IBD samples compared to HC.

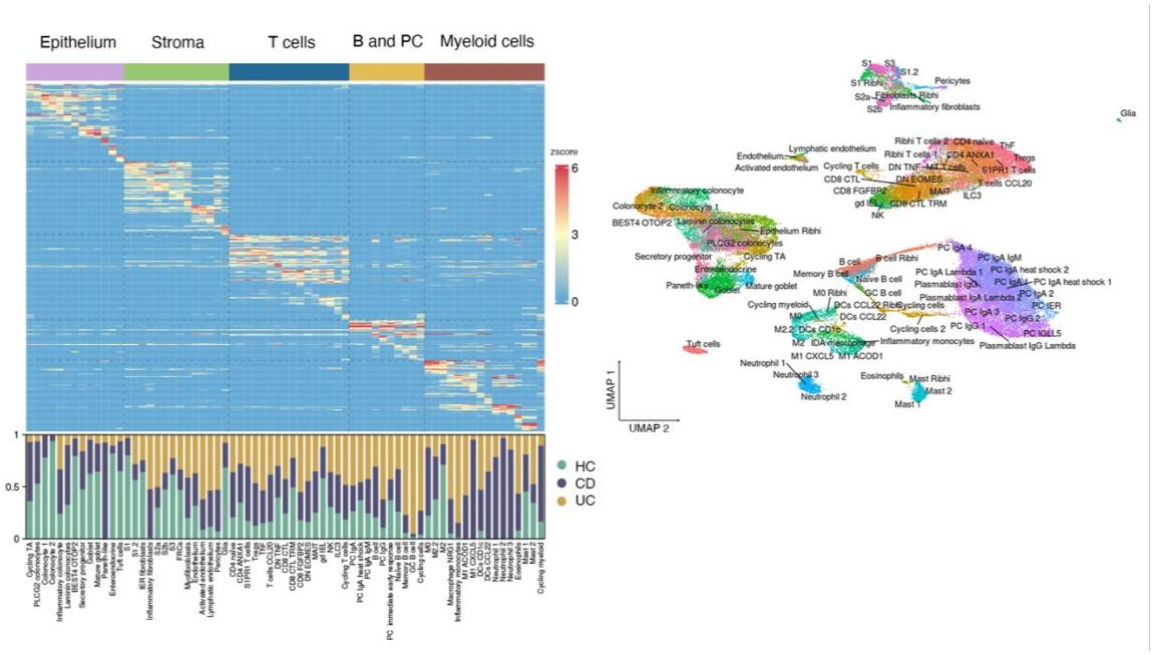


Figure 19. Five different compartments found in the intestinal mucosa. Heatmap of top marker genes discriminating the different cell subsets (epithelium, stroma, T cells, B and Plasma cells, and myeloid cells) and, below, barplots representing the proportions of each cell type resolved by scRNA-seq for HC, CD and UC. On the right, UMAP showing annotation of all cell types identified by scRNA-seq.

In order to validate scRNA-seq cell populations, an additional cohort of formalin-fixed paraffin-embedded (FFPE) colonic samples (**Fig. 1**) was processed using CosMx Spatial Molecular Imaging (SMI; NanoString Technologies)[227]. Scanner for some areas of the tissue called Fields of View (FoVs) and immunofluorescence staining of pan-cell markers (CD45, PanCK, CD3) were performed on all samples (**Annex Fig 1**). Selected FoVs were processed by CosMx SMI using a multiplex panel of 1,000 genes. Annotation of cells was performed by label transfer based on scRNA-seq clusters, using the 100 top-ranked markers and count matrix (**Methods**), which assigned a unique label to each cell (**Annex Fig. 1**) and allowed us to find the cell populations of interest in the tissue. Similar to the scRNA-seq analysis, SMI data showed a decrease in epithelial cell populations and an increase in myeloid and B and plasma cells in IBD samples. Compared to scRNA-seq, the SMI cohort exhibited a higher proportion of stromal cells and a lower proportion of T cells across all individuals. This could be explained by the nature of the samples (**Fig**

3). Most of the samples processed by SMI were surgical pieces, in contrast to the pinch biopsies used in the scRNA-seq experiments. Surgical pieces are bigger and usually contain deeper layers of the intestinal wall, where stromal cells predominate. However, almost all scRNA-seq populations were able to be located in the SMI images.

Thus, by integrating scRNA-seq and CosMx SMI from human colonic samples, we have generated the first spatially resolved map of healthy and diseased colon at single-cell spatial resolution.

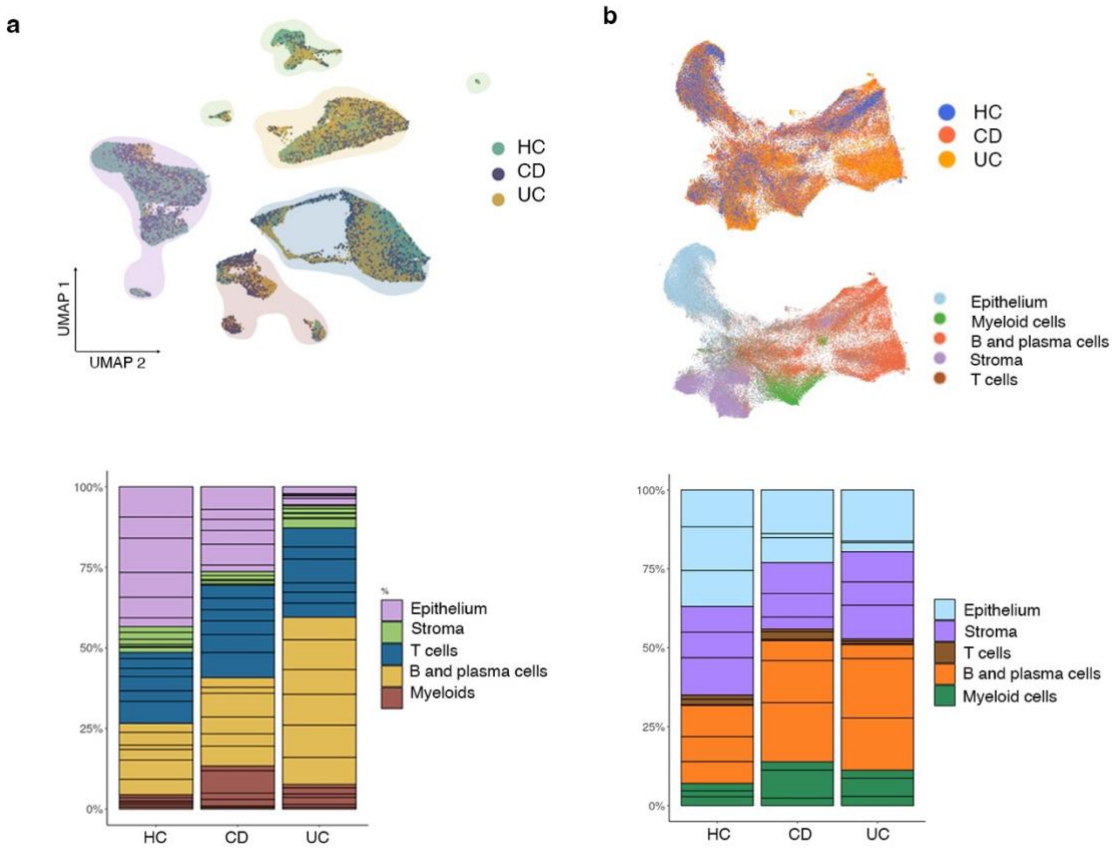


Figure 20. Cell subsets change in proportion between health and disease. a, UMAP and barplots of scRNA-seq data. Cells in UMAP are colored by group origin (HC, CD and UC) and clusters are shaded by cell subset (epithelium, stroma, T cells, B and Plasma cells, and myeloid cells). Barplots show the proportions of each cell subset in HC, CD and UC. **b**, UMAP and barplots of CosMx^{TM} SMI data. Top UMAP shows cells colored by group (HC, CD and UC) while bottom UMAP is colored by cell subset (epithelium, stroma, T cells, B and Plasma cells and myeloid cells). Barplots show the proportions of each cell subset in HC, CD and UC.

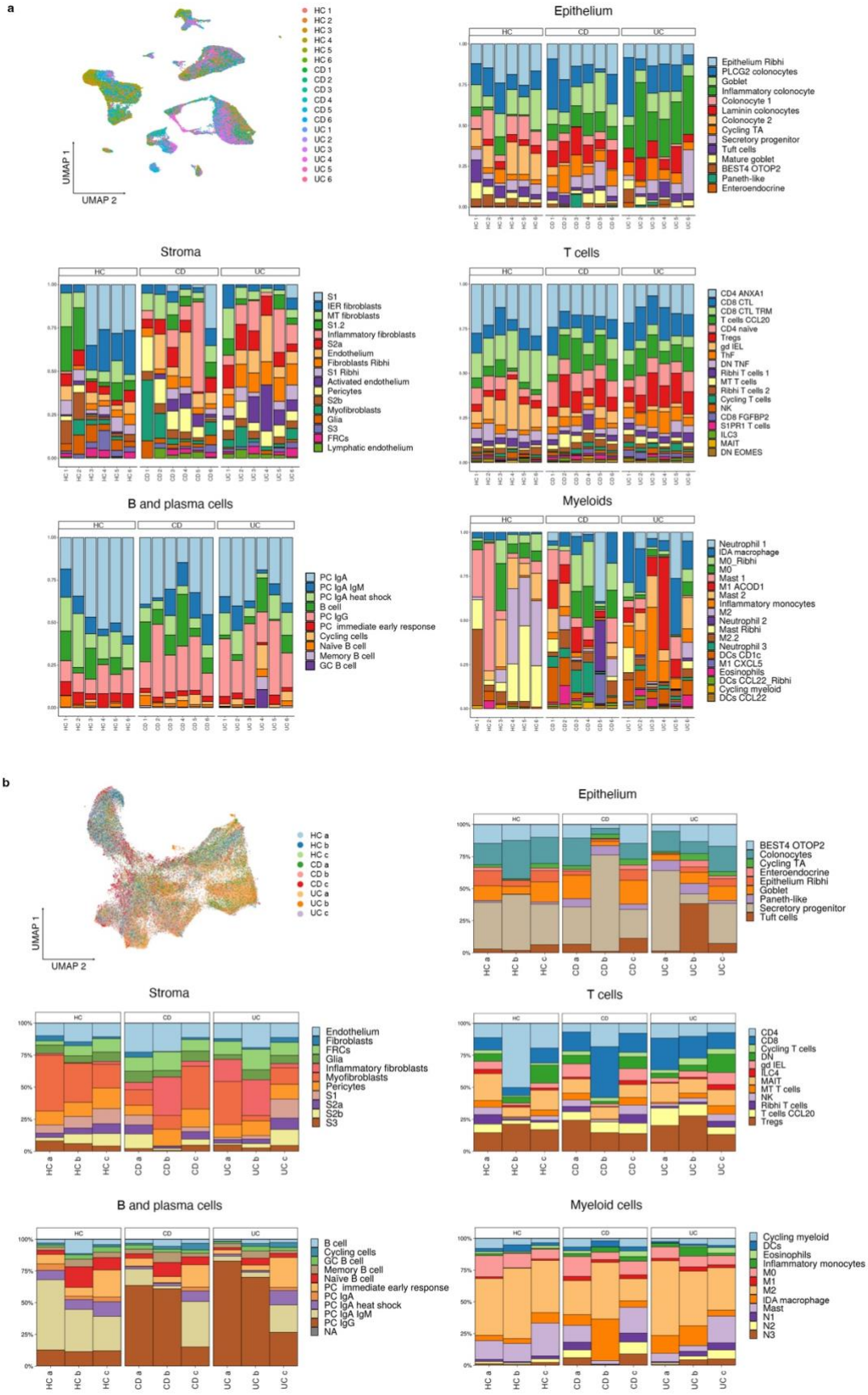


Figure 21. Single-cell RNA-seq (scRNA-seq) and CosMx™ Spatial Molecular Imaging (SMI) cell clusters. a, UMAP representation of a sample (n=18) distribution across subsets analyzed by scRNA-seq. Barplots describe the proportions of each cell type within each cell subset (epithelium, stroma, B cells and

plasma cells, T cells and myeloid cells) in healthy controls (HC, n=6), active CD (n=6) and active UC (n=6) patients using scRNA-seq data. b, UMAP representation of samples analyzed by CosMx™ SMI (n=9). Cells in CosMx™ SMI were annotated by label-transfer of scRNA-seq differentially expressed genes per cell cluster. Barplots describe the proportions of each cell type within each cell subset (epithelium, stroma, B cells and plasma cells, T cells and myeloid cells) in HC (n=3), active CD (n=3) and UC (n=3) patients using CosMx™ SMI data.

SECTION 2: The epithelium in health and disease

Our analysis revealed that the epithelium underwent the most significant changes in cell proportions between healthy and IBD tissue, as shown in Figure 3. Motivated by this finding, we first focused our investigation on this cell compartment to gain a deeper understanding of the specific alterations present in IBD patients.

Intestinal epithelium suffers a rewiring in inflammatory bowel disease

The most abundant cell type identified within healthy colonic mucosa are epithelial cells. Using scRNA-seq, we were able to find all epithelial cell populations expected in the healthy controls. Cluster analysis revealed 14 different epithelial cell types/states, the majority of which have been previously described in healthy and UC colon[205,250] . These included secretory cells: goblet cells (*TFF3*, *ITLN1*, *ZG16*) and mature goblet (*TFF1*), tuft cells (*SH2D6*, *TRPM5*, *IL17RB*), enteroendocrine cells (*CHGA*, *GCG*, *NEUROD1*) and secretory progenitors (*SPINK1*, *LEFTY1*), and several types of colonocytes (**Fig. 22**), the differentiated ones characterized by the expression of *AQP8*. Besides fully differentiated colonocytes, with our analysis we identified new clusters that we annotated as Laminin colonocytes, and PLCG2 colonocytes, which function remain unknown. Moreover, recently described BEST4+ colonocytes were also observed in our cohort. We could also distinguish the presence of a dividing cell population (characterized by expression of *MKI67*, *TOP2A*, *CENPF*, *NUSAP1*...) different from a stem subset without expression of canonical stem cell markers (*LGR5*) that we annotated as cycling transit-amplifying (TA). However, in this dataset, we were not able to get stem cells.

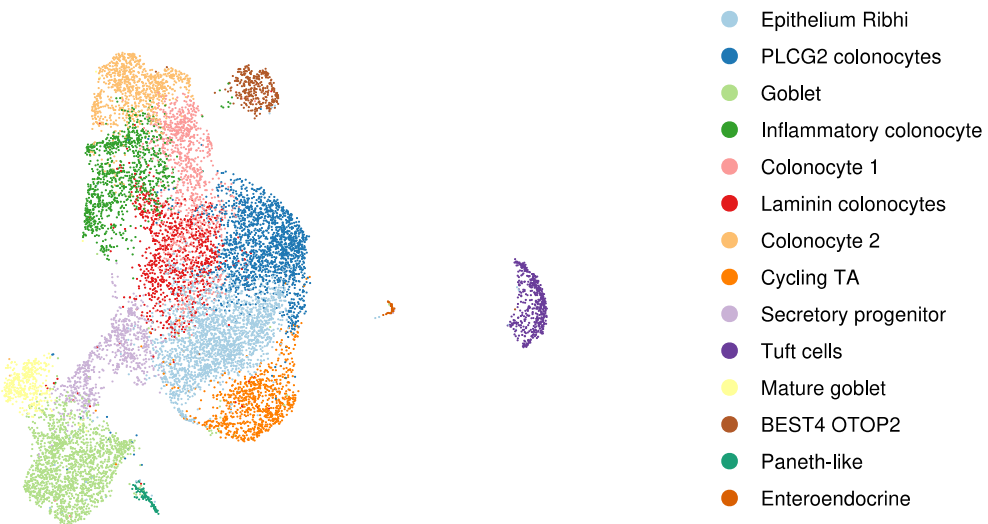


Figure 22. Colonic epithelial cell populations. UMAP representing the merged epithelial cell populations of HC, CD and UC

By using scRNA-seq, we confirmed the occurrence of epithelial cell loss in patients with IBD. All 14 epithelial cell types identified in HC were present in CD patients, although for some of the clusters, the relative abundance was partially decreased (**Fig. 23**). This epithelial reduction, however, was more pronounced in UC patients in all clusters finding a high decrease in cell numbers, especially in differentiated subsets like differentiated colonocytes and goblet cells.

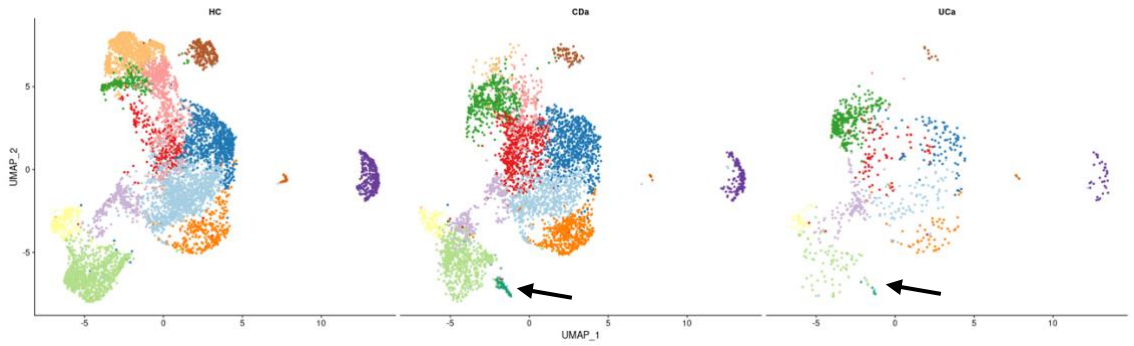


Figure 23. Colonic epithelial cell populations in health and disease. UMAP representing the cell populations present in each group: HC, CD and UC. Different abundances are observed between health and disease. Great loss of epithelial cells are observed in UC. Arrows show Paneth-like cells.

Interestingly, we detected a small cluster of cells expressing the Paneth-cell markers *DEFA5* and *DEFA6*, in addition to other epithelium-restricted antimicrobial genes *REG1A*, *REG1B* and *REG3A* only in IBD mucosa (**Fig. 23 and 24**). The presence of Paneth cell metaplasia in the colon has been long described in the literature as a feature of chronic IBD. Here we found this population only in one UC and one CD patient, which suggests these are rare-occurring cells (Arrows **Fig. 23**).

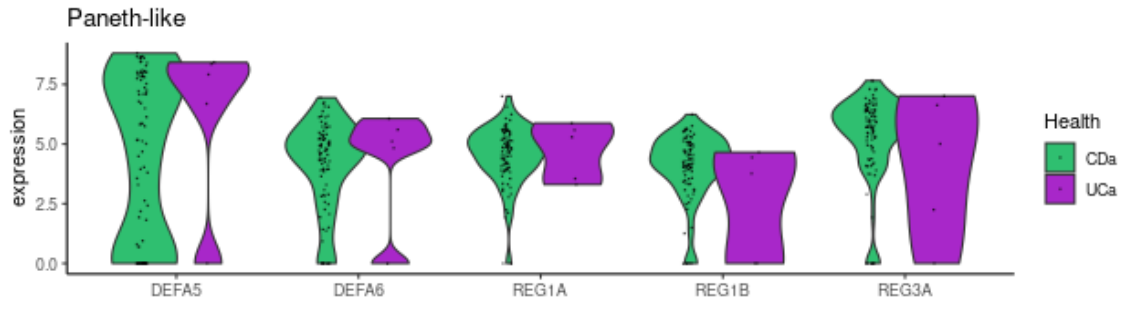


Figure 24. Paneth cell-like cluster is observed in the colon of IBD patients. Specific cell markers of Paneth cells that are not present in the healthy mucosa but appear in IBD patients as a feature of the disease.

Another change that we observed in the epithelium of UC and CD was the marked increase of inflammatory genes. A cluster of colonocytes, Inflammatory colonocytes, expressing defense mechanisms genes including *LCN2*, encoding for lipocalin, and inflammation related genes *DUOX2*, *DUOXA2*, and *NOS2* (**Fig. 25**) was expanded in IBD. Moreover, this population does not express the canonical colonocyte marker *AQP8* (**Fig. 26**). This upregulation of inflammatory related genes in IBD samples was observed not only within the cluster of inflammatory colonocytes, but also within other enterocytes including cycling TA, laminin colonocytes, PLCG2 colonocytes, and BEST4 OTOP2 or secretory cells such as goblet, goblet mature and secretory progenitors (**Fig.25**). These results highlight the presence of an inflammatory signature in the remaining epithelium of IBD patients.

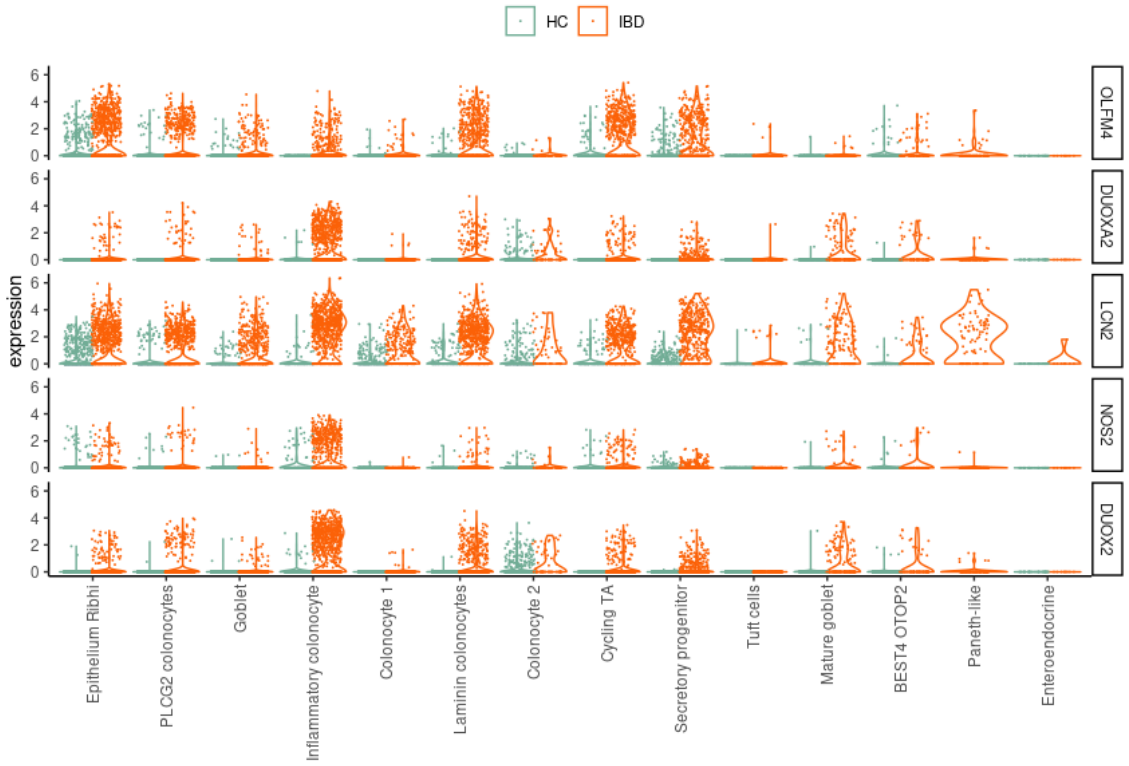


Figure 25. Inflammatory epithelial genes are present in the epithelium of IBD. Violin plots showing the expression (y-axis) of inflammatory markers in epithelial clusters (x-axis) from HC and IBD colons The cluster of inflammatory colonocytes presented high expression of inflammation-related epithelial genes such as *LCN2*, *DUOX2*, *DUOXA2*, *NOS2*. Additionally, other clusters expressed *LCN2* in inflammation.

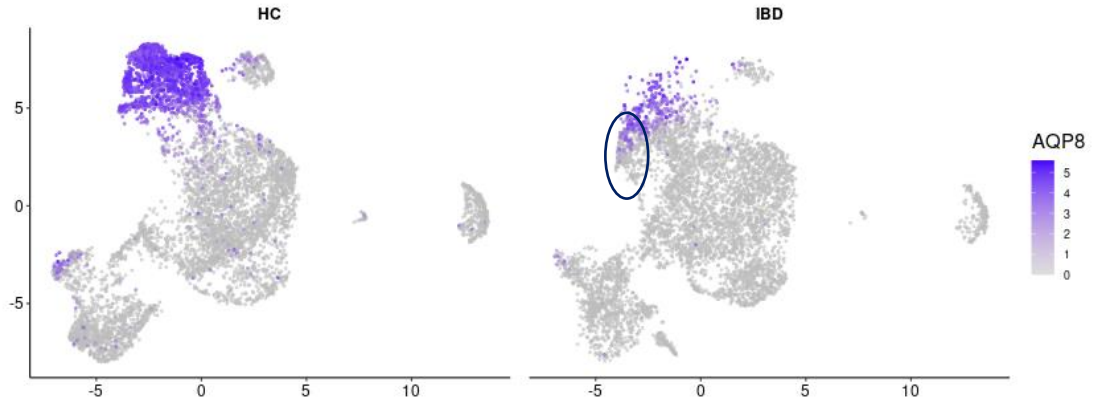


Figure 26. Epithelial expression of *AQP8* in scRNA-seq data. UMAP representing *AQP8* expression in the epithelium of HC and IBD, which is restricted to colonocytes. However, inflammatory colonocytes only present in IBD (shown by a circle) loos expression of *AQP8*.

Besides inflammatory changes in the epithelium, scRNA-seq data revealed the exclusively epithelial origin of *OLFM4*. Olfactomedin-4 has emerged as a stem cell marker in the human and mouse intestine, as observed to be present at the bottom of the intestinal crypt. A wide range of functions have been given to *OLFM4*, among them, an important role in mucosal defense. In our scRNA-seq data of healthy colon, *OLFM4* is restricted to a progenitor cell type (annotated as Epithelium Ribhi, **Figure 23**) located at the bottom of the crypt that is different from the proliferating *MKI67* expressing cluster (cycling TA) (**Fig. 27**). Importantly and in agreement with previous data (REFS) *OLFM4*

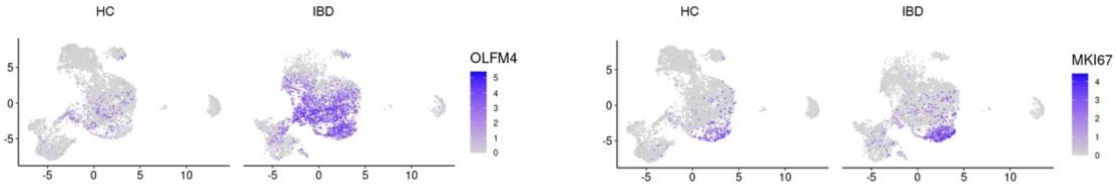


Figure 27. *OLFM4* and *MKI67* expression in the intestinal epithelium. UMAP showing the expression of *OLFM4* and *MKI67* in the epithelial compartment from healthy control (HC) and IBD tissues (UC and CD) by scRNA-seq.

is markedly upregulated in IBD patients, with several progenitor-like cell types including transit TA, Ribhi and secretory progenitor clusters highly expressing it (**Fig. 28**). According to our data these are the main populations present in the epithelium of inflamed samples. Since we did not identify a stem cell cluster in our dataset, we may not rule out the expression of *OLFM4* in those cells. However, we did observe expression of *OLFM4* in other cell types different from the stem cells as already mentioned.

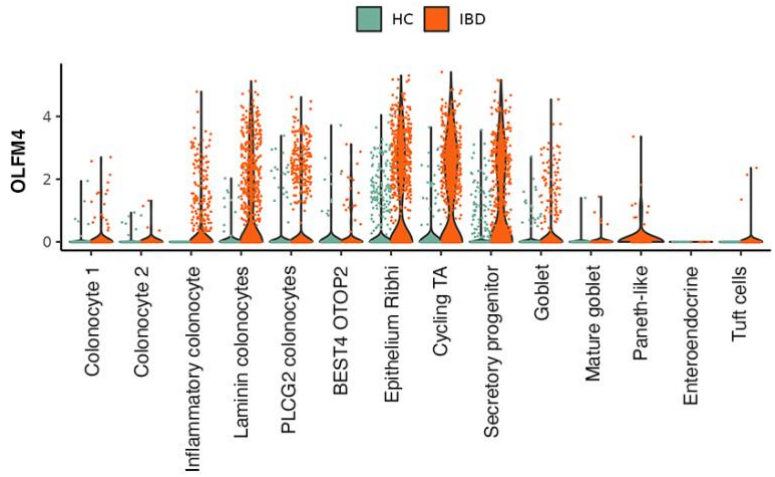


Figure 28. Expression of *OLFM4* in the colonic epithelium. Violin plot showing *OLFM4* expression (y-axis) in all epithelial cell subsets (x-axis) in HC (green) and IBD (orange) samples by scRNA-seq data.

We next sought to understand how much of these epithelial-restricted changes could be measured by bulk biopsy RNA analysis. To do that we took advantage of a cohort of HC, CD, and UC patients, for which colonic RNA was sequenced using bulk RNA-seq analysis. In agreement with scRNA-seq data, Paneth cell markers *DEF5* and *REG1A*, as well as inflammatory genes *DUOX2* and *LCN2*, and the TA cell marker *OLFM4* were significantly upregulated in both CD and UC colonic biopsies confirming scRNA-seq results. (**Fig. 29**).

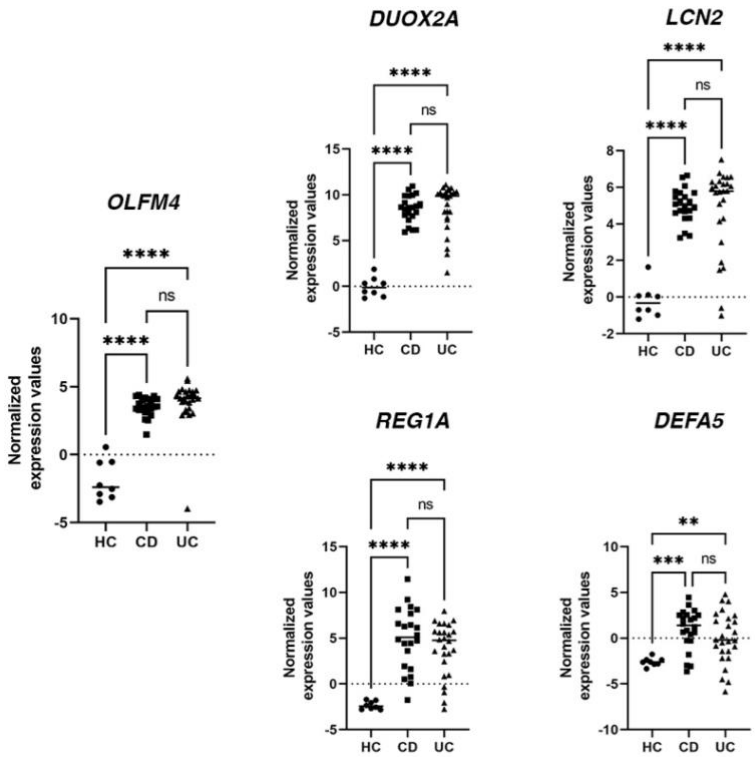


Figure 29. Bulk RNA-seq expression of inflammatory-related genes. Expression of specific inflammatory epithelial cell markers in HC (n=8), active CD (n=22) and active UC (n=26) patients using bulk biopsy RNA-seq data. p<0,05(*), p<0,01 (**), p<0,001(***), p<0,0001(****), ns: not significant.

Immunohistochemistry (IHQ) and in situ hybridization (ISH) analyses revealed that, in addition to being increased, OLFM4 expression in IBD is distributed extensively throughout the crypt, as opposed to being restricted to the base of the crypt as observed in healthy controls (HC). (**Fig. 30**).

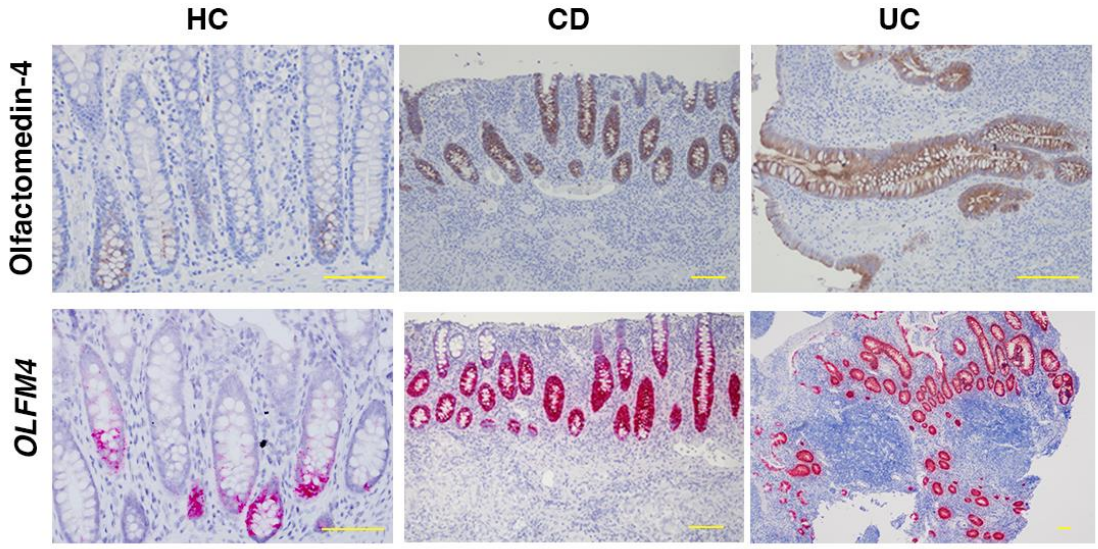


Figure 30. ***OLFM4* RNA and protein expression in colonic tissue**. Olfactomedin 4 immunostaining and *OLFM4 in situ* hybridization in HC, active CD and active UC colon. Increased staining along the crypt is observed in the IBD samples. Scale bars= 100 µm

Using CosMx SMI and label transferring, we were able to find all epithelial cell populations observed in scRNA-seq and corroborate the changes in *OLFM4* expression in the epithelium of IBD patients (**Fig. 31**). We also observed the upregulation of anti-microbial mechanisms such as the expression of defensins (*DEFA5*), lipocalins (*LCN2*) and enzymes involved in producing reactive oxygen species (*DUOXA2*) in the epithelium of IBD samples (**Fig. 32**).

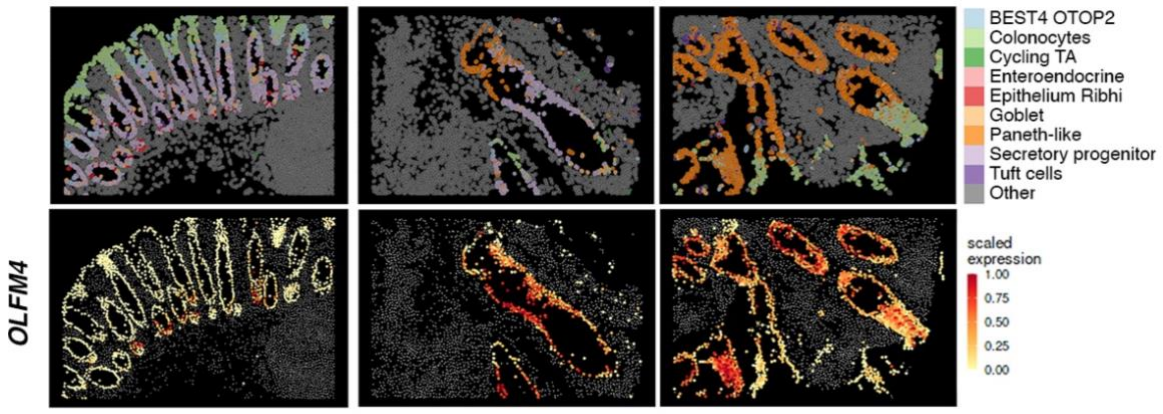


Figure 31. CosMx™ SMI analysis of epithelial cell localization in healthy and IBD tissue. CosMx™ SMI visualization and localization of the different epithelial cell subsets described by scRNA-seq, from left to right in a HC and two UC representative Fields of View (FoVs) and (iv) mean expression of *OLFM4* in each of those cells analyzed by CosMx™ SMI

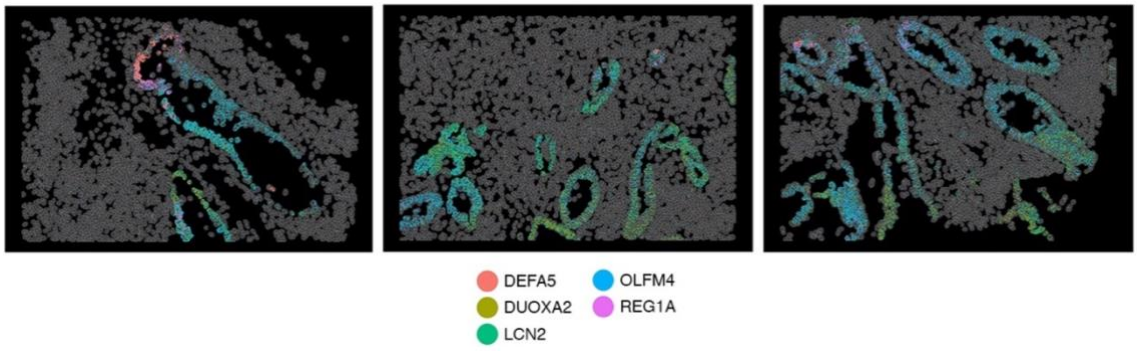


Figure 32. CosMx™ SMI analysis of epithelial genes in IBD tissue. Expression of *DEFA5*, *LCN2*, *DUOX2A*, *REG1A*, and *OLFM4* within epithelium of representative FoVs of a UC patient analyzed by CosMx™ SMI. Dots represents mRNA molecules.

In summary, our data confirm the epithelial changes associated with both UC and CD, including epithelial cell loss and a decrease in *AQP8+* crypt top enterocytes. We also observed pronounced upregulation of antimicrobial and pro-inflammatory genes in the epithelial cells present in IBD, as well as a new inflammatory colonocyte population with higher expression of these genes. In addition, we discovered a previously unknown laminin 3-expressing epithelial cell population in the colonic mucosa, and a Paneth-like cell population only present in IBD samples. Interestingly, while UC appears to be associated with more significant epithelial cell loss, the transcriptional signatures of this compartment do not differ significantly between UC and CD.

SECTION 3: Myeloid cells heterogeneity in inflammatory bowel disease

To gain a more comprehensive understanding of the compartments within our dataset, we conducted a thorough analysis using the Morisita-Horn similarity index (**Fig. 33**). Our goal was to identify variations in these compartments across patients and health status. The results of our analysis revealed that the myeloid and stromal compartments exhibited the most significant differences in cluster proportions within patient groups, suggesting that these cell groups may be strongly influenced by patient-specific factors, which could contribute to patient-to-patient heterogeneity. Given the significance of these findings, we delved further into studying the myeloid and stromal compartments seeking to understand the underlying factors contributing to their heterogeneity. By doing so, we aimed to shed light on the role of these cell types in disease progression and to identify potential therapeutic targets to mitigate the observed heterogeneity.

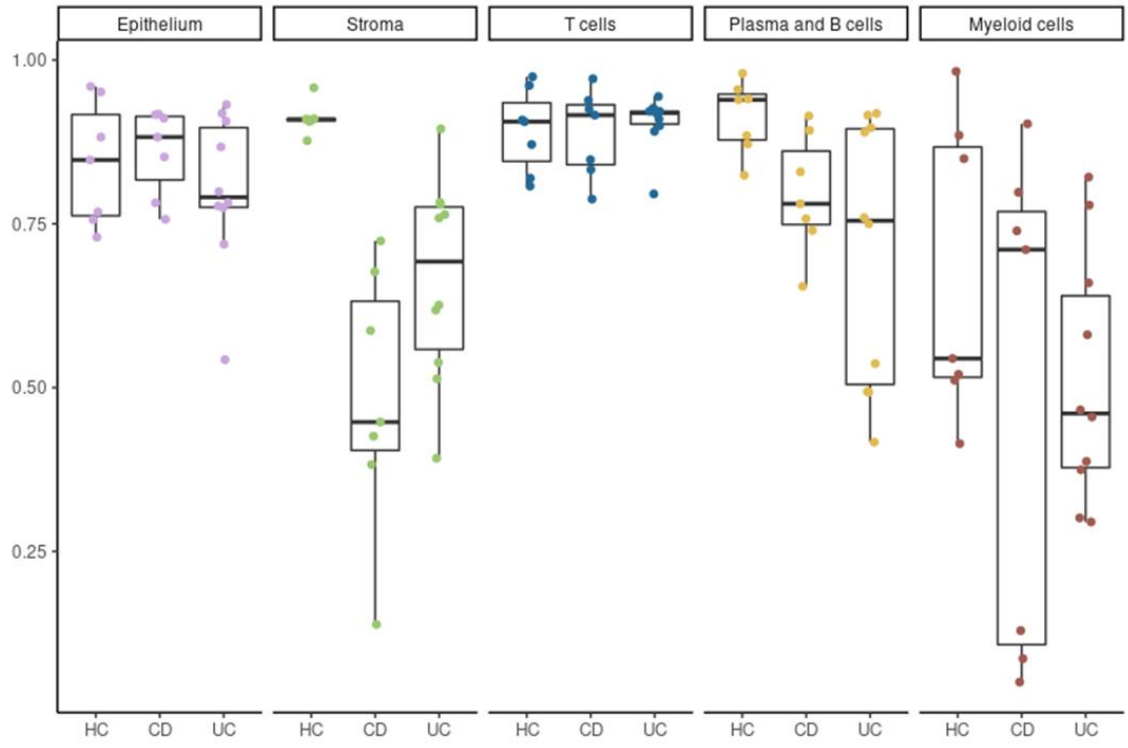


Figure 33. Morisita-Index analysis of scRNA-seq main subsets. Morisita-Index analysis of dispersion between samples within each group (HC, CD, UC) for each cell subset (epithelium, stroma, B cells and plasma cells, T cells and myeloid cells) based on scRNA-seq data.

Regarding the myeloid compartment, pooled HC, CD and UC scRNA-seq data identified different myeloid clusters including several macrophage subtypes, dendritic cells,

inflammatory monocytes, mast cells, neutrophils, and eosinophils, whose proportions changed in both UC and CD samples compared to controls (**Fig. 34**). To perform a differential abundance test without relying on cell clustering we applied Milo, a tool which relies on k-nearest neighbor graphs, and we confirmed the differential abundance between groups in the myeloid compartment (**Fig. 35**). The changes that we observed in this analysis were the presence of M1 macrophages and neutrophils only in the IBD patients as well as inflammatory monocytes and a newly identified macrophage population, Inflammation-dependent alternative (IDA) macrophage.

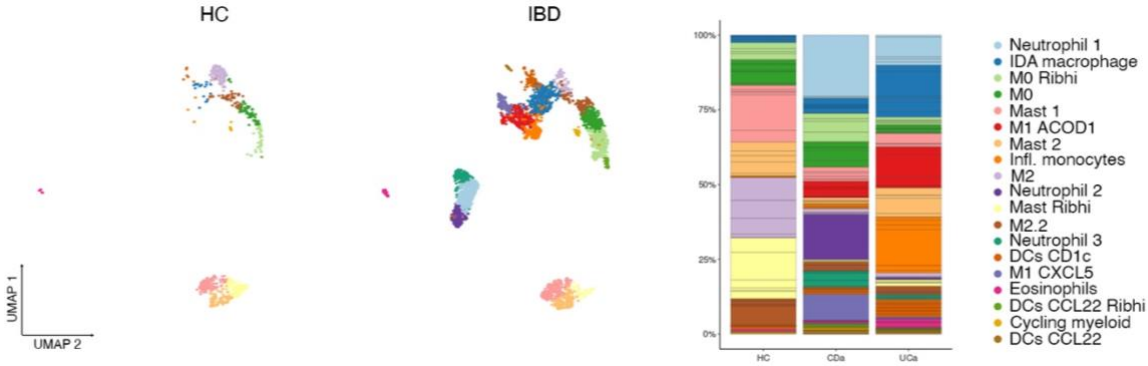


Figure 34. **Analysis of myeloid cell subsets in healthy and inflamed colonic mucosa**. UMAP representation of scRNA-seq data showing the different myeloid clusters (macrophages, mast cells, dendritic cells, neutrophils and eosinophils) present in healthy controls (HC, n=6) and IBD colonic samples (CD n= 6 and UC n=6); myeloid cell subset proportions across healthy and IBD samples, x-axis represents patient groups and y-axis the percentages of each myeloid subset (ii).

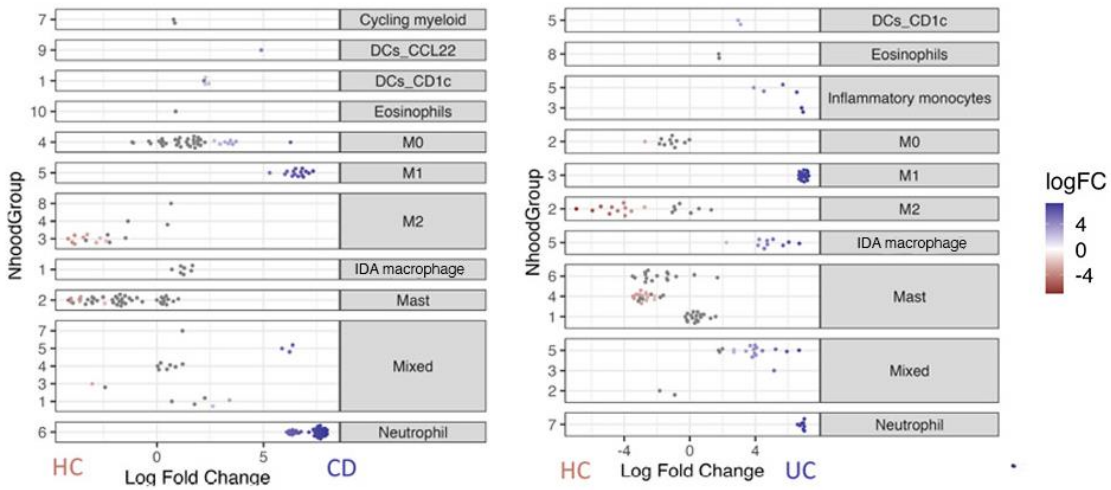


Figure 35. independent proportion analysis reveals differences in abundances between HC and IBD. Cell type enrichment analysis using the differential abundance test Milo. Panels show the cell subsets enriched in HC vs CD (left) or in HC vs UC (right). Annotation of each nhood group based on our analysis (panel a) is shown on the right for each comparison. The logarithm of the fold-change comparing CD or UC to HC is represented on the horizontal axis. The nhood group number is represented in the vertical axis for each analysis.

Three distinct neutrophil phenotypes are present in inflammatory bowel disease

Exploring the myeloid compartment in our scRNA-seq data, we found diverse populations of intestinal granulocytes: mast cells, eosinophils, and neutrophils (**Fig. 34**). We identified three clusters of mast cells, one of them was very enriched for ribosomal genes (Mast Ribhi) which is a cell state of high activation, and the other two clusters were different regarding the expression of LTC4S (**Annex Table 2**). We additionally identified one small cluster of eosinophils that were identified by the expression of *CLC*, *MS4A3*, *CCR3* and the "Th2" cytokines *IL4* and *IL13* markers (**Annex Table 2**). Our analysis revealed that eosinophils were significantly increased in samples from active UC patients,

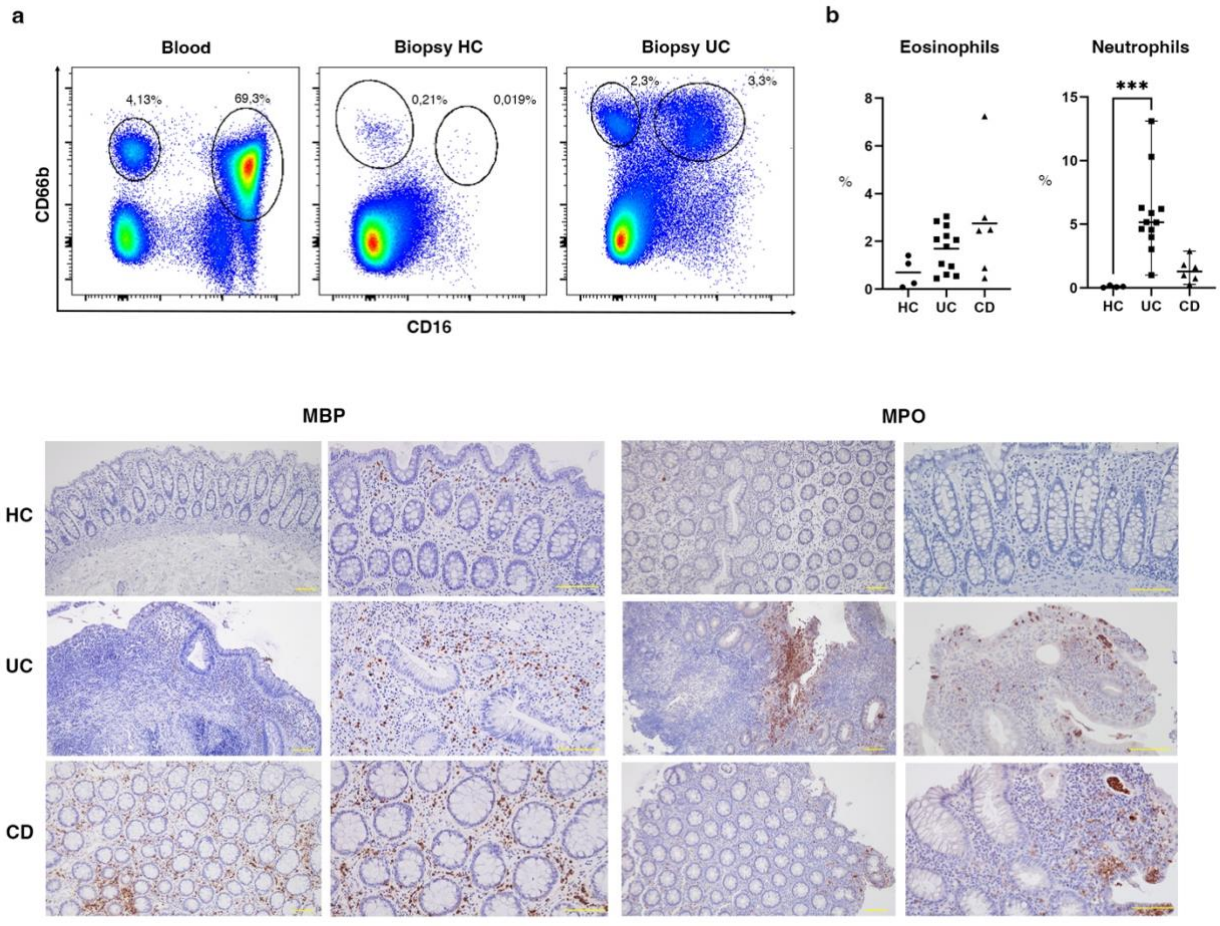


Figure 36. **Analysis of neutrophils and eosinophils in inflammatory bowel disease (IBD). a**, Flow cytometry gating strategy to detect eosinophils and neutrophils in IBD blood and colonic biopsies. Numbers represent the percentages of neutrophils ($CD66b^+$ $CD16b^+$) and eosinophils ($CD66b^+$ $CD16b^-$) in the sample displayed as representative of all samples analyzed. **b**, Percentage of neutrophils ($CD66b^+$ $CD16b^+$) and eosinophils ($CD66b^+$ $CD16b^-$) in colonic biopsies from healthy controls (HC, n=4), active CD (n=6) and active UC (n=12) colonic samples analyzed by flow cytometry. **c**, Immunostaining for myeloperoxidase (MPO, marker neutrophils) and myelin basic protein (MBP, marker of eosinophils) in representative HC, active UC and active CD colonic tissues (scale bar = 100 µm).

while only a small number of eosinophils were detected in a single healthy intestinal sample. To confirm these findings, we conducted flow cytometry using CD66b+CD16- (**Fig.36a** and **36b**) and performed immunohistochemistry with MBP staining (**Fig. 36c**). Consistent with our previous results, we found that eosinophil levels were higher in biopsies obtained from IBD compared to healthy controls although this increase was not significant by flow cytometry in our cohort. Supporting their increase in UC and CD, the expression of most eosinophil markers was significantly increased in bulk biopsy RNA-seq (**Fig. 37a**)

As mentioned, CCR3 (CD193) is a known specific marker of eosinophils that we observed by scRNA-seq. Additionally, we validated CD193 by flow cytometry. All eosinophils in blood express this marker (97%) and a big majority of colonic eosinophils also express it on its membrane surface (47%).

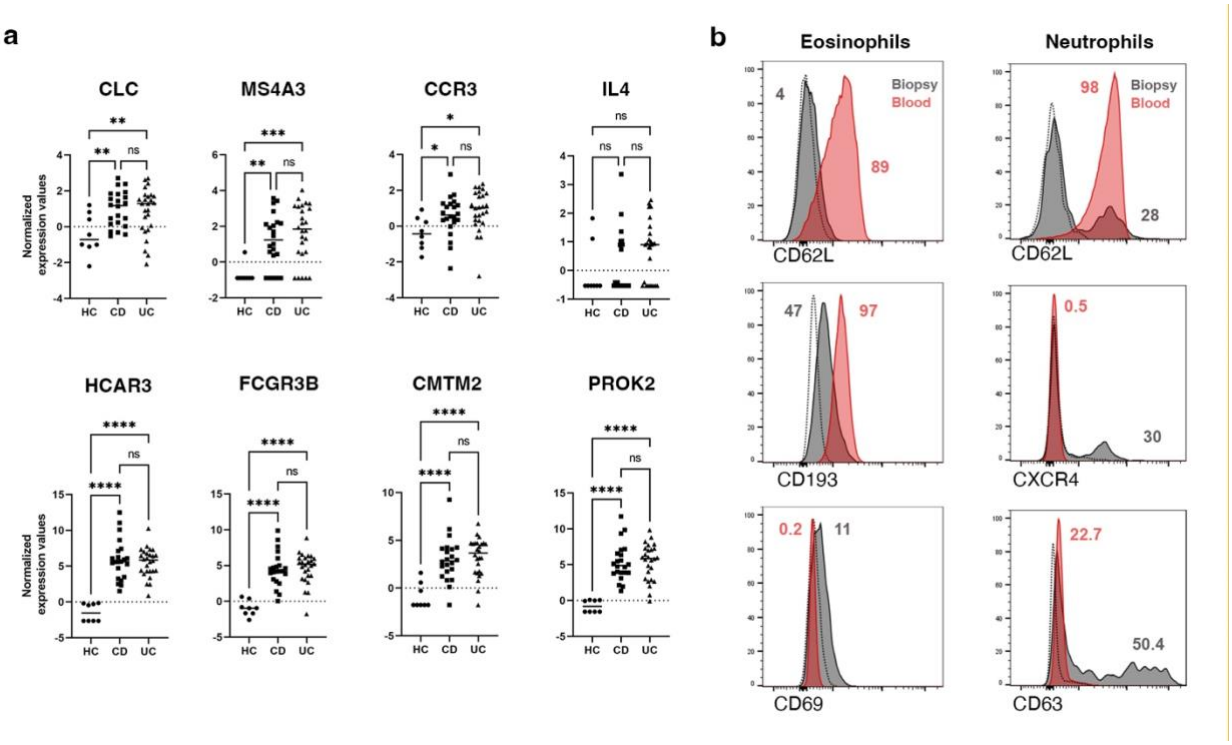


Figure 37. Protein and mRNA analysis of neutrophils and eosinophils in inflammatory bowel disease (IBD) a, Protein expression detected by flow cytometry in neutrophils (one representative sample is shown from12 colonic and 6 blood samples) and eosinophils (one representative sample shown from 15 colonic and 5 blood samples) from blood and biopsies of IBD patients. Numbers show the percentage of positive cells for the protein in blood (red) and biopsy (grey)Histogram for the corresponding isotype control in the biopsy sample is shown as a dashed line. **b**, Bulk colonic biopsy RNA-seq expression of neutrophil and eosinophil-specific markers in HC (n=8), active CD (n=22) and UC (n=26) patients. p<0,05(*), p<0,01 (**), p<0,001(***), p<0,0001(****), ns: not significant.

We detected neutrophils in inflamed CD and UC samples while these cells were completely absent from healthy mucosa as we observed by scRNA-seq (**Fig.34**), flow cytometry (**Fig. 36a**) and immunohistochemistry by MPO staining (**Fig. 36c**). To our knowledge, this is the first dataset to detect colonic neutrophils using the 10x Genomics platform (**Methods**). A crucial step in identifying these cells was applying a lower filter of genes/cell in the analysis, as neutrophils typically have lower mRNA content (**Fig. 38**).

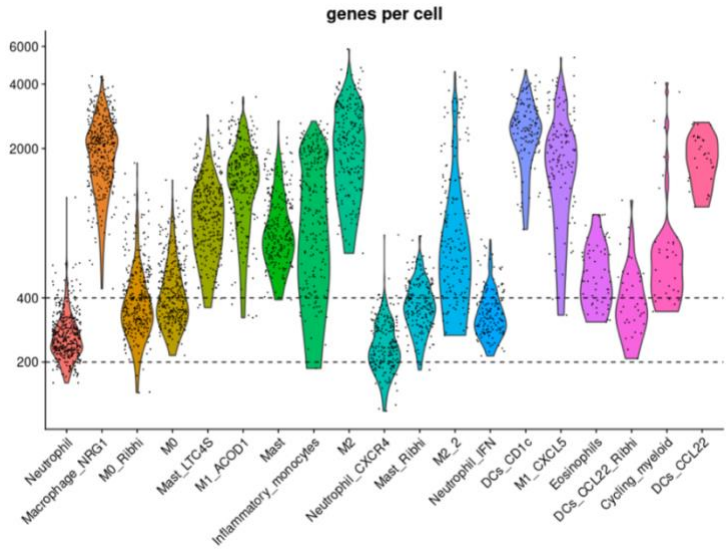


Figure 38. Neutrophils have a very low mRNA content. Violin plot representing the number of genes per cell of each population of the myeloid compartment.

Interestingly, neutrophils segregated in 3 different clusters. The 3 clusters presented general common markers of neutrophils being the most specific: *HCAR3*, *FCGR3B* (CD16b), *CMTM2* and *PROK2. FCGR3B* (CD16b) and *CMTM2* are known markers of neutrophils in other settings. We confirmed the relevance of these neutrophil-specific markers in the inflamed mucosa by analyzing their presence in bulk RNA-seq data from whole IBD biopsies. All 4 neutrophil markers *HCAR3*, *FCGR3B* (CD16b), *CMTM2* and *PROK2* appeared highly increased both in active ulcerative colitis and Crohn's disease

compared with healthy intestinal mucosa (**Fig. 37a**). The 3 unique neutrophil states were annotated as N1, N2 and N3 and their relative abundance varied on individual patients and disease type (**Fig. 39a**). The 3 different clusters of neutrophils, N1, N2 and N3 appeared in the majority of analyzed IBD patients but with variable abundances depending on individual patients and disease type. Remarkably, between our patients we found one single UC patient with no neutrophils at all (despite presenting active inflammation) and one single CD patient with 16.5 times more neutrophils than the mean of neutrophils present in the rest of CD patients (**Fig. 39b**).

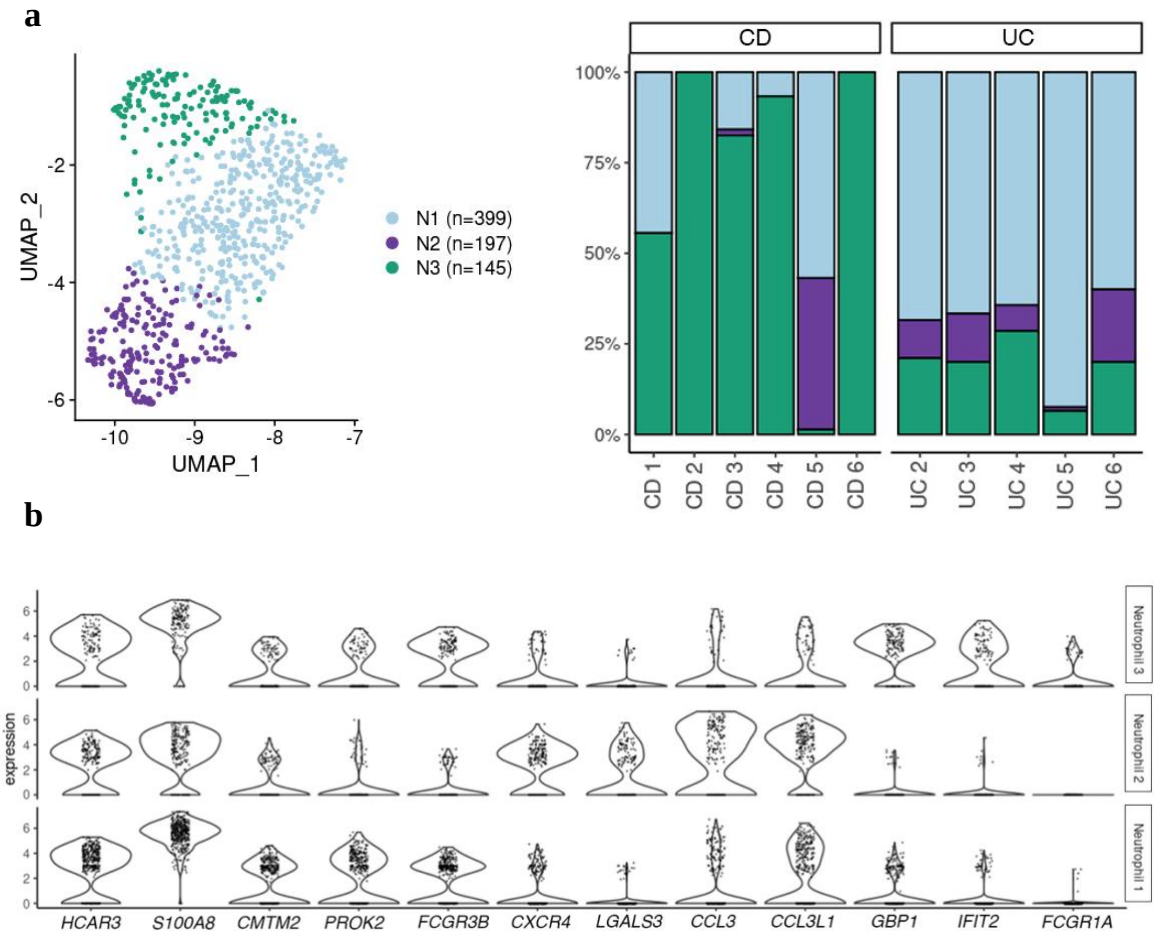


Figure 39. **Analysis of the heterogeneity of neutrophil populations in inflammatory bowel disease (IBD) colonic mucosa**. **a**,UMAP showing the three neutrophil subsets/states (N1, N2, N3) observed in IBD samples by scRNA-seq analysis. Barplot representing the proportions of each neutrophil subset across health and IBD. **b**, Violin plots visualizing the expression (x-axis) of marker genes common and specific for all three neutrophil populations (y-axis).

N1 represents the majority of neutrophils in our cohort of UC patients (58.6% out of total neutrophils in each patient) but not in CD patients (20.6% of total Neutrophils). The expression profile of N1 subcluster was reminiscent of blood neutrophils and we could

not identify a marker or group of them that uniquely identifies this subcluster. To confirm the similarity of colonic N1 neutrophils to the circulating ones, we compared our neutrophils to public scRNA-seq datasets from peripheral neutrophils. N1 and N3 neutrophils showed the highest similarity to both bone marrow mature neutrophils (BM matureN)[251] and to umbilical cord blood neutrophils (UCB)[252] (**Fig. 40**). In contrast, N2 neutrophils showed little overlap with peripheral neutrophils.

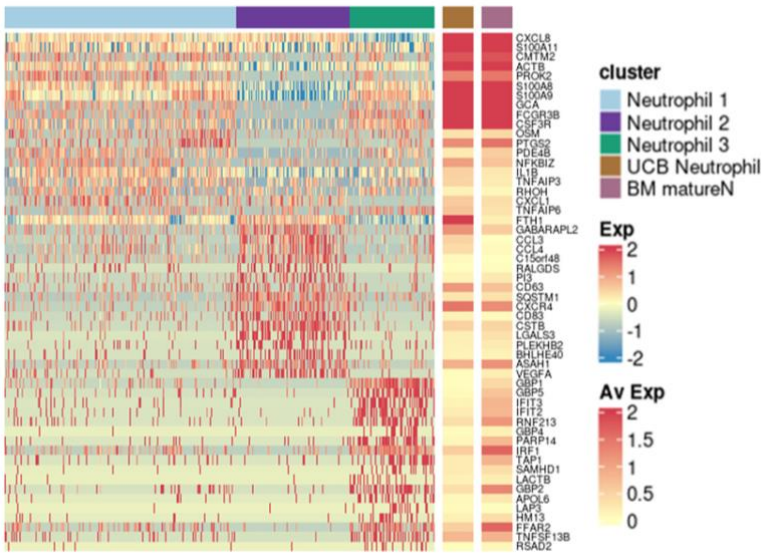


Figure 40. Intestinal and blood neutrophil's profile. Heat map showing the average normalized and scaled expression of differentially expressed genes in all three neutrophil subsets. Average expression of these genes on neutrophils from cord blood and bone marrow mature neutrophils is shown on the far right (Xie, X., *et al* 2021, Zhao, Y., *et al* 2019)

N2 is the least abundant subcluster of neutrophils both in UC and CD (8.7% in UC and 7.2% in CD out of total Neutrophils). N2 cluster is characterized by a lower expression of the main neutrophil markers i.e. *S100A8*, *SA100A9*, *FCGR3B*, *CMTM2*, *TNFAIP6*, *AQP9* and *CSF3R*. In opposition these N2 neutrophils presented a higher expression of a set of chemokines and chemokine-receptors: *CXCR4*, *CCL3*/*CCL4*, *CCL4L2*, *CCL3L1*, *CXCL2*, *CXCL8*, *CCRL2*, suggesting distinct tissue localization. They also presented increased expression of genes suggestive of different activation states or functions: CD63, *CD83*, *FTH1*, *VEGFA*. Further, we wanted to validate by protein the presence of the N2 neutrophils in the tissue of IBD patients, so we stained with antibodies for CXCR4 and CD63 N2 markers. Protein expression of these N2 markers was confirmed for 10% and 61% of tissue neutrophils compared to 0.26% and 14% of blood neutrophils, respectively (**Fig. 37b**), demonstrating the presence of these neutrophils in colonic IBD tissue.

Finally, N3 is the most abundant population of neutrophils in CD patients (72% out of total neutrophils in each patient) but not in UC (16%). N3 neutrophils characterized by

increased expression of interferon induced genes, i.e. GBP proteins (*GBP1*, *GBP2*, *GBP4*, *GBP5*), and *IRF1*, *IFIT2*, *IFIT3*, *RSAD2*, *MX1*, *MX2*.

Moreover, all 3 neutrophil subsets were found spatially using CosMx SMI and showed scattered distribution throughout inflamed lamina propria, with predominant localization in crypt abscesses and ulcerated areas (**Fig. 41**).

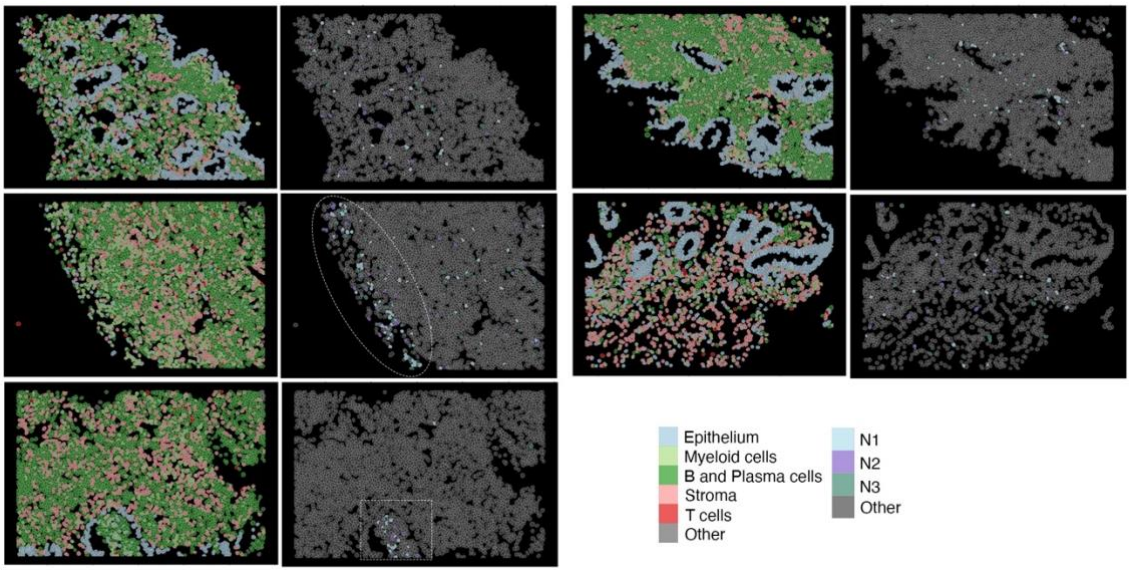
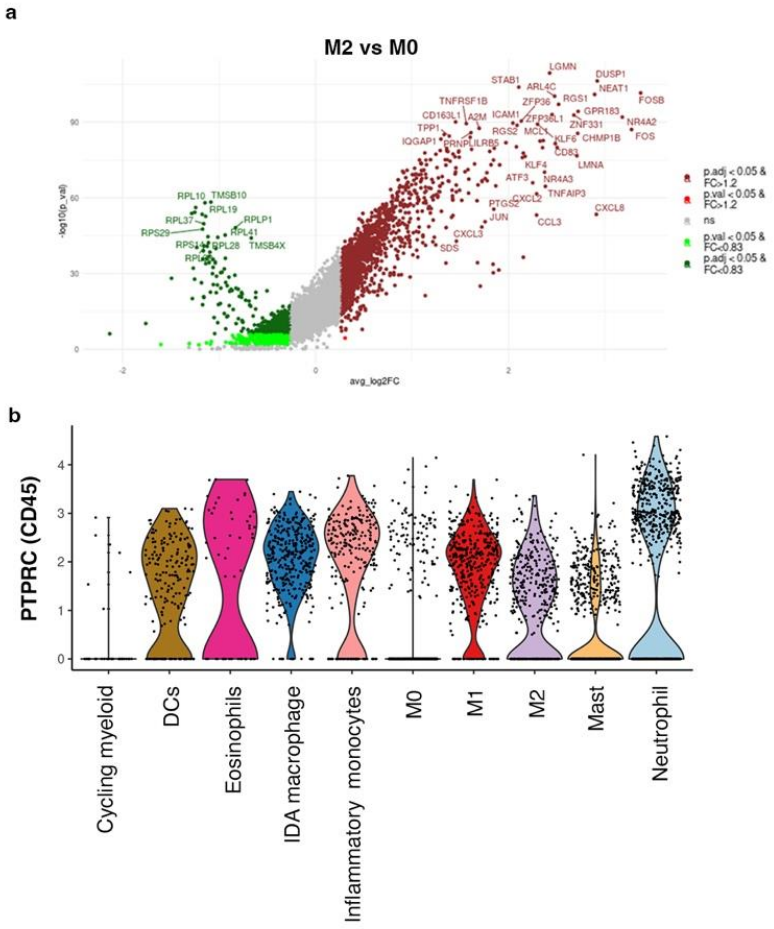


Figure 41. Spatial distribution of neutrophils by CosMx SMI analysis. Representative CosMx SMI images of IBD inflamed tissue showing the spatial location of N1, N2 and N3 neutrophil subsets. Circle shows the surface of an ulcer, and a square shape is used to indicate a crypt abscess.

Novel populations of resident and inflammatory macrophages are found in the colonic mucosa

Due to the high heterogeneity found in the macrophage compartment, we decided to explore further this cell type. In healthy colon, resident macrophages (expressing *C1QA*, *HLA-DRB1* and *SELENOP*, among others) were found in different transcriptional states (**Annex Fig. 2**). These included macrophages expressing well-described M2-specific markers (i.e., *CD163L1*, *CD209*, *FOLR2*), annotated as M2 and M2.2, and hereafter referred to as M2. We annotated the other two clusters present in HC as M0 (M0 and M0_Rib^{hi}), as they lacked M2 markers but highly expressed all other macrophage-specific genes (**Fig. 42a** and Supplementary Table 4), while showing low expression of *PTPRC* (CD45) (**Fig. 42b**).



Besides M0 and M2, as mentioned before, samples from UC and CD additionally contained inflammatory monocytes and activated macrophage clusters that we annotated as M1 ACOD1, M1 CXCL5 (pooled as M1) and IDA macrophages (**Fig. 34** and **Fig. 43**).

Figure 42. **Analysis of myeloid cell subsets in healthy and inflamed colonic mucosa** . **a,** Volcano plot of differentially expressed genes (DEGs) in M2 compared to M0 macrophages (Supplementary Table 4 contains the complete list of genes). Genes upregulated in M2 macrophages are shown in dark (UUP, p value<0.05) or light red (UP, nominal p value<0.05). Genes downregulated in M2 are shown in dark (DDW, FDR<0.05) or light green (DW, p value<0.05). **b**, Violin plot visualization of *PTPRC* (CD45) expression in myeloid populations by scRNA-seq.

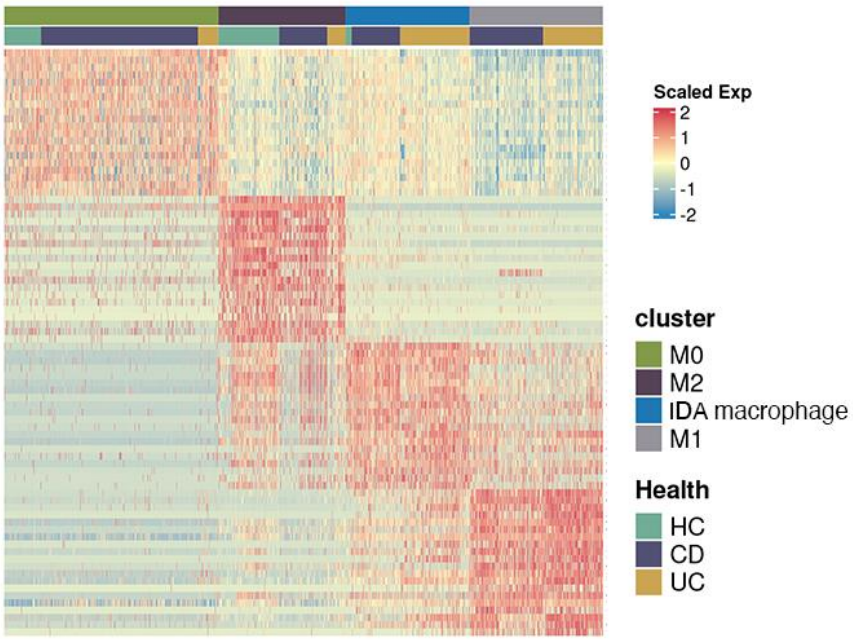


Figure 43. Transcriptional profile of macrophages present in health and disease. Heat map showing the average expression of DEGs for M0, M2, M1 and IDA macrophages in HC, CD and UC.

We were able to find all myeloid subsets in tissue sections analyzed by CosMx SMI, showing marked differences in abundance and spatial distribution depending on the patient and/or disease type (**Fig. 44**).

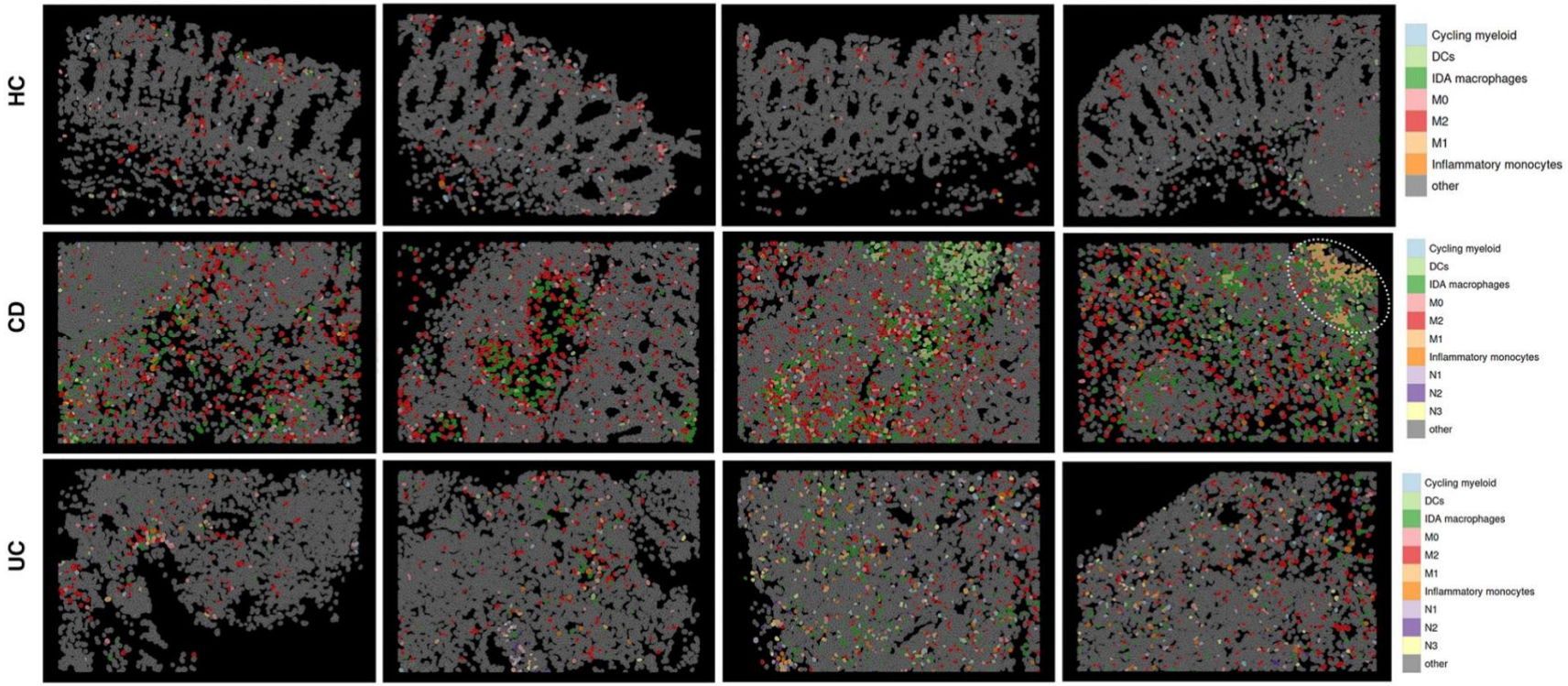


Figure 44. CosMx SMI identifies all myeloid cell populations. CosMx^{TM} SMI images showing spatial distribution of the different myeloid cell populations in representative Fields of View (FoVs) of colonic tissue of two HC, one inflamed CD and two inflamed UC patients. White dotted circle indicates ulcerated area with abundant M1 macrophages.

M0 and M2 represent two independent states

Until now, M0 macrophages have not been formally described in the human intestine. Thus, we first compared our monocyte/macrophage signatures to publicly available data from HC, UC[205] and CD terminal ileum[239] (**Fig. 45a**) and found, in both cohorts, populations of macrophages that resembled the M0 and M2 subsets (Jaccard indexes=0.3) corroborating our own data (**Fig. 45b**).

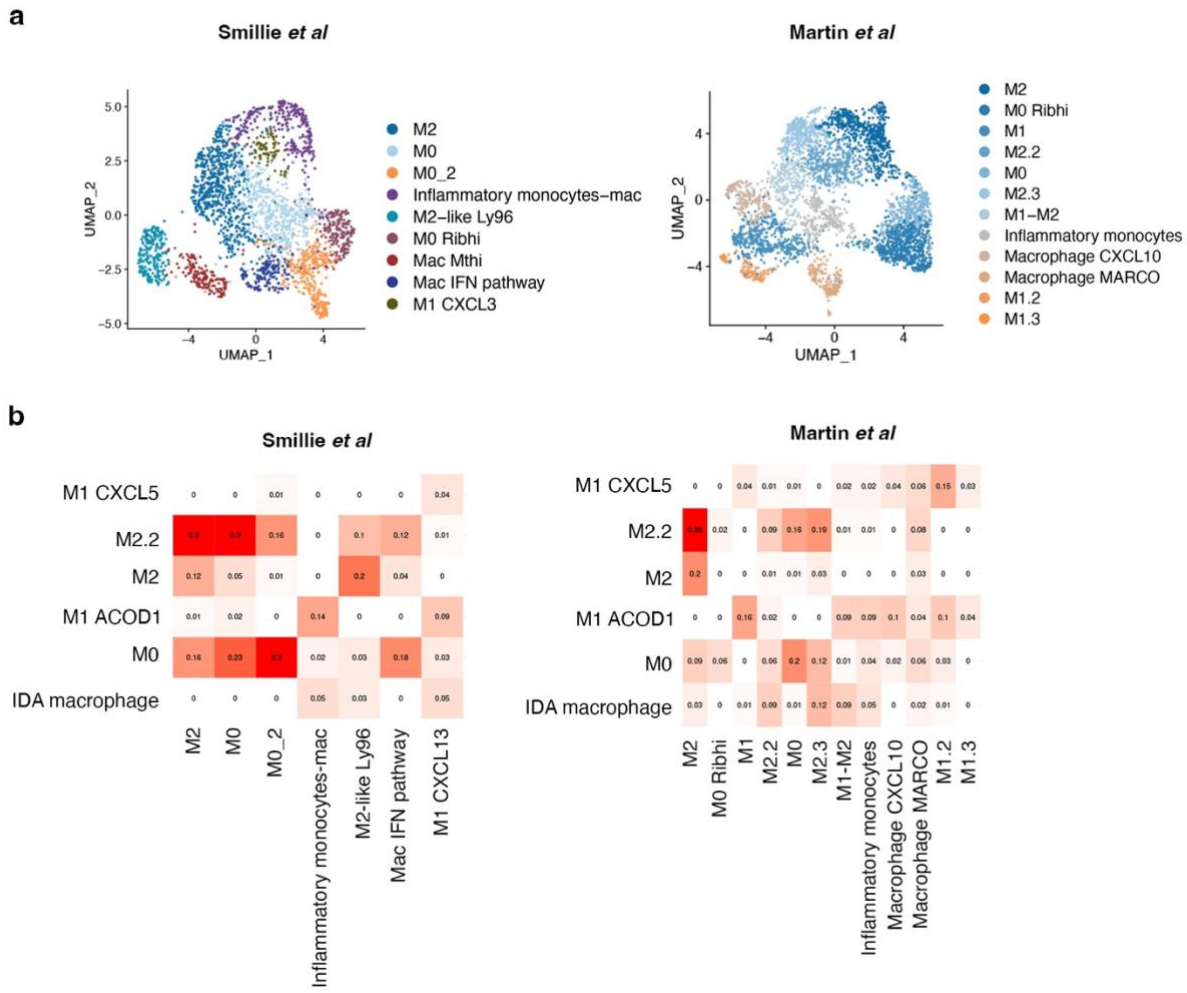


Figure 45. Intestinal Macrophages across studies. a, UMAP representation of Smillie *et al*- and Martin *et al*- macrophage subsets isolated *in silico*. **b**, Jaccard similarity index between scRNA-seq data of Smillie *et al*- and Martin *et al*- macrophages and our scRNA-seq macrophage populations

In agreement with the scRNA-seq data, we could visualize both $CD209^{+}CD68^{+}$ (M2) and $CD209^{-}$ $CD68^{+}$ (M0) cells using immunostaining in healthy colonic lamina propria, mostly localizing below the apical epithelium (**Fig. 46**) and also present throughout the lamina propria. Likewise, M0 and M2 cells were identified by CosMx SMI analysis (**Fig.**

47), confirming the dual identity of these resident macrophages. Interestingly, in the immunofluorescence staining seems that M0 macrophages are located at the most top crypt area, near the lumen, in the colon.

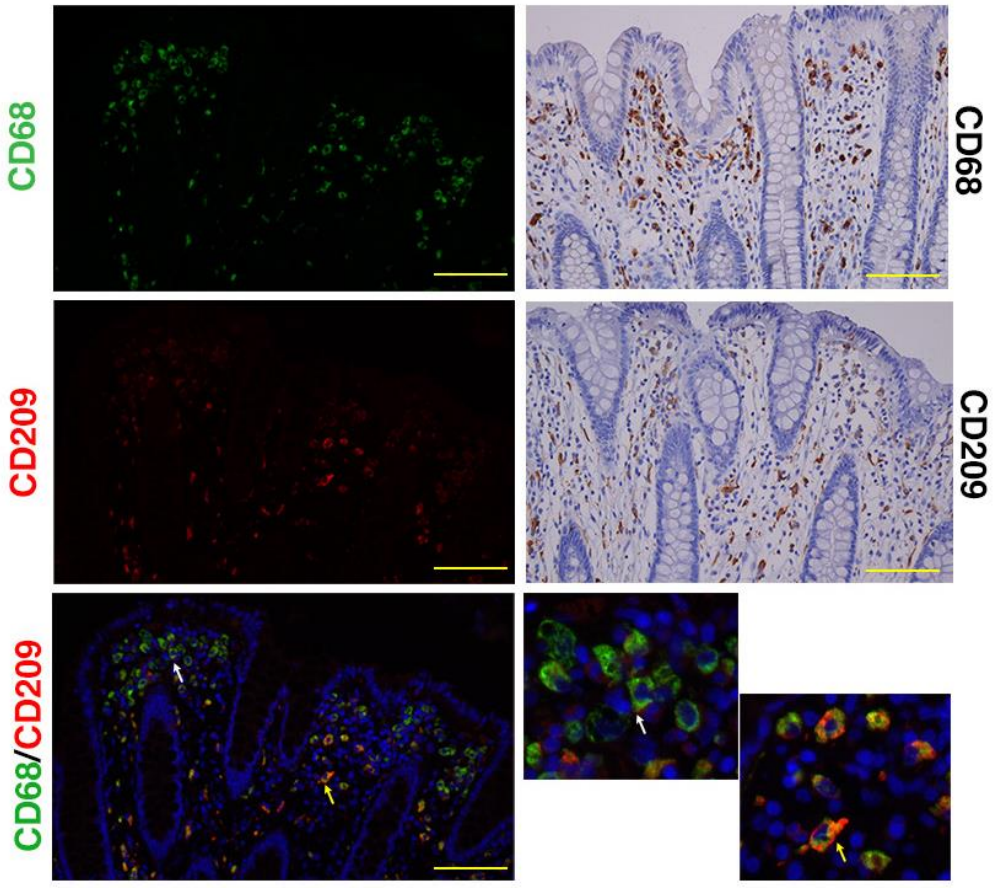


Figure 46. **Two different resident intestinal macrophages are present in the healthy colon**. On the left and right insets, double immunofluorescence showing M2 ($CD209^+$ in red and $CD68^+$ in green) and M0 ($CD209^-$ and $CD68^+$) subsets in a representative healthy colonic tissue. White and yellow arrows indicate M0 and M2 macrophages, respectively. Right two top panels: immunohistochemistry showing the distribution of CD68 and CD209 markers in healthy colonic tissue (Scale bar 20 µm)

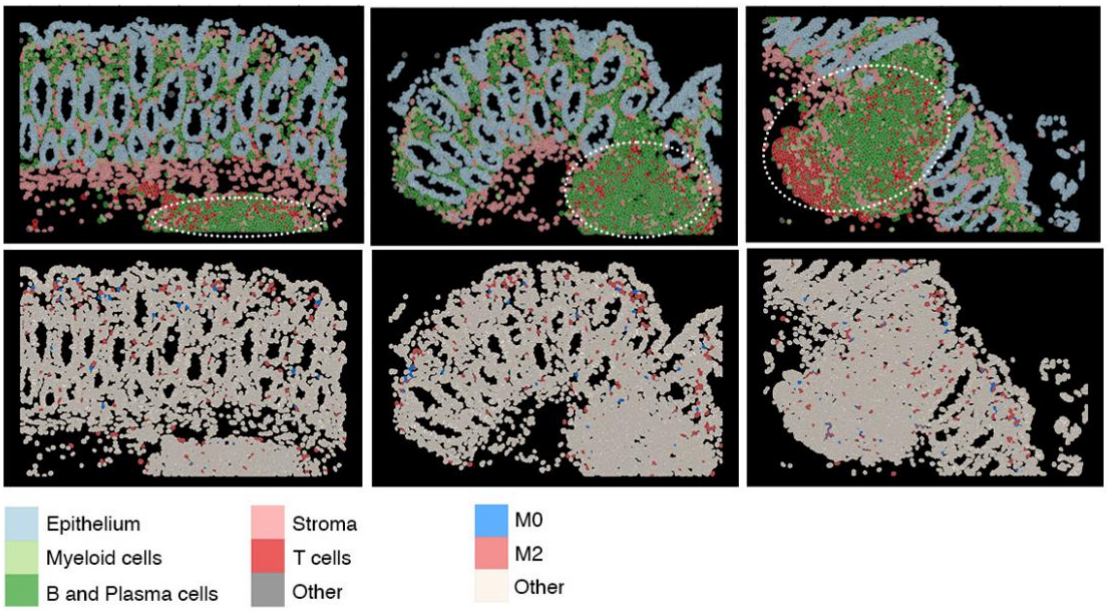


Figure 47. **Two different resident intestinal macrophages are present in the healthy colon**. CosMx™ SMI images of a representative healthy colonic mucosa with lymphoid follicles (highlighted by dotted circles). Upper panels show the cellular localization of the five cell subsets (epithelial, stroma, T cells B and plasma cells, and myeloid cells) and lower panels M0 and M2 resident macrophages location.

To understand their phylogenetic origin and their relation to other previously described macrophage subsets, we mapped our dataset to a recently published human monocyte-macrophage database containing data from 41 studies on several organs and diseases (MoMac-VERSE)[253]. M0, M2.2 and M2 mapped to independent macrophage clusters within the MoMac-Verse dataset, supporting the hypothesis that they do represent two unique states (**Annex Fig. 2a**). Interestingly, M2.2 mapped near the M0 cluster.

Indeed, comparison of M0 and M2 macrophages in our dataset to *in vitro* monocyte-derived macrophages from published datasets[244] showed high similarity between intestinal M2 and *in vitro* M-CSF monocyte-derived macrophages (**Fig. 48**), while none or little overlapping with M0 macrophages was observed under these same conditions.

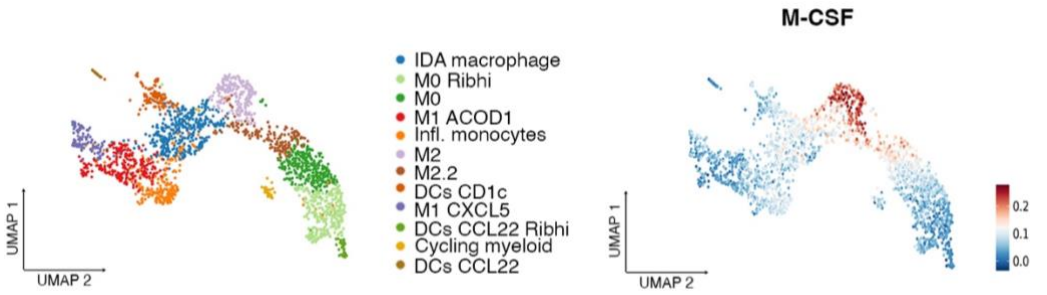


Figure 48. Resident M2 macrophage present a signature similar to M-CSF stimulated blood-derived macrophages . a, UMAP showing monocyte, macrophage and dendritic cell clusters identified using scRNA-seq analysis of HC, CD and UC colonic biopsies (left panel). Representation of overlapping signatures (average expression score) of upregulated genes in *in vitro* M-CSF derived macrophages (Cuevas VD *et al*., 2022) on our macrophage cell UMAP (right panel).

Next, we decided to perform a trajectory analysis to interrogate macrophage populations in order to see if there is a biological process of differentiation among them. Interestingly, this analysis of our data (**Fig. 49**) and the two public datasets analyzed (**Annex Fig. 2b**) suggests separate pseudo-time states for M0 and M2 clusters. In our data, M2.2 cluster, which express M2 markers, appear close to M0, suggesting they may represent a transitional state between the two resident compartments and suggesting M0 as the origin of M2 macrophages.

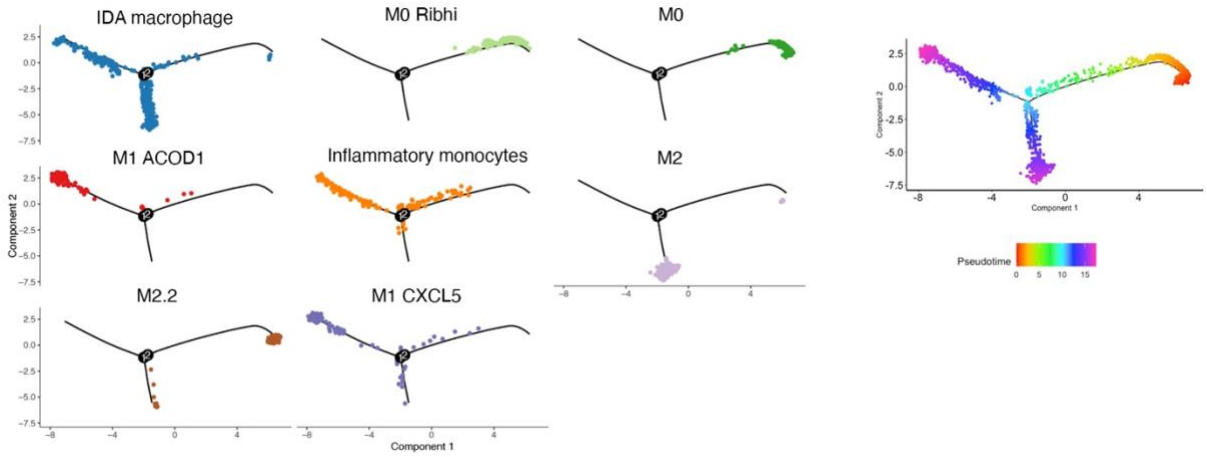


Figure 49. **Trajectory analysis revealed the origin of intestinal macrophages**. Pseudo-time trajectory analysis of monocytes and macrophage populations in our scRNA-seq data from pooled HC, CD and UC.

Overall, we conclude that in the healthy colon, resident macrophages are found in at least two states. M2 macrophages could potentially originate from circulating monocytes

exposed to M-CSF in tissues, while the origin of M0 macrophages remains unknown, but we hypothesize that could be a resident phenotype and also the origin of M2 macrophages in the tissue.

Inflammatory macrophages show highly heterogeneous signatures among patients

Compared to HC, IBD patients showed a marked increase in the total number and transcriptional heterogeneity of the macrophage population (**Fig. 34** and **Fig. 35**). Apart from of M0 and M2, we found inflammatory/activated macrophages in at least three different states: two transcriptionally different M1 populations (M1 ACOD1 and M1 CXCL5) and a newly identified IDA macrophage cluster, in addition to a population of inflammatory monocytes. M1 clusters had in common some inflammatory markers that define them as inflammatory macrophages like *TNIP3*, *CLEC5A* and *INHBA*, but they present differences at the transcriptomic level, representing two different M1 phenotypes and suggesting potential different functions. M1 ACOD1 uniquely expressed ACOD1, IDO1, and interferon-response genes *GBP2*, *GBP5*, *GBP1* and *IFITM3*. In contrast, M1 CXCL5 highly expressed *CXCL5* and *CCL8*, which code for neutrophil recruiter cytokines and matrix metalloproteinases *MMP9*, *MMP12* and *MMP10*. Inflammatory monocytes highly expressed *CD300E*, *VCAN*, *FCN1* and *LYZ* among other inflammatory genes such as *S100A9* calprotectin subunit. IDA macrophages presented a different signature from M1 and will be discussed in the next section of this thesis.

In contrast to M0 and M2 macrophages, the similarities between inflammatory macrophages in our cohort and those found in other intestinal datasets[205,239] were weaker, suggesting that activated macrophages may be found in highly patient/context-dependent states (Jaccard Index <=0.14 Smillie et al, and Jaccard Index <=0.16 Martin et al.; **Figure 45b**). We also compared our signatures and mapped them to MoMAc VERSE data set. Our inflammatory macrophages were mapped together with inflammatory monocytes and macrophages in that data set (**Annex Fig. 2a**).

As with M2 and M-CSF-derived macrophages, intestinal M1 CXCL5 cells showed high similarity to the *in vitro* GM-CSF-derived macrophages, while the signature of M1 ACOD1 was shared by both M-CSF and GM-CSF-derived macrophages stimulated with LPS[244] (**Fig. 50**). Remarkably, our analysis revealed that IDA macrophages exhibited a striking similarity to M-CSF-derived macrophages that had been treated with serotonin (5-HT). These findings suggest that 5-HT may play an important role in modulating the phenotype and function of IDA macrophages in the context of intestinal inflammation[241].

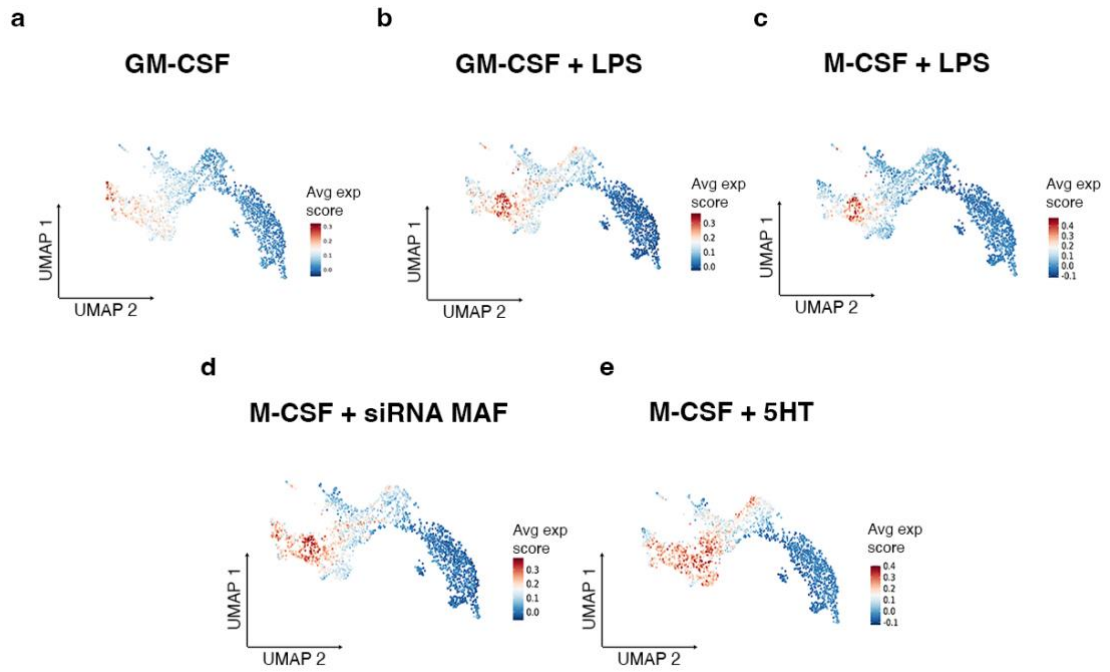


Figure 50. In vitro blood-derived macrophages signatures compared to our scRNA-seq data. Representation of overlapping signatures (average expression score) of *in vitro* **a)** GM-CSF-derived macrophages (Cuevas VD *et al*., 2022), **b)** GM-CSF-derived macrophages stimulated with LPS (Cuevas VD *et al*., 2022), **c)** M-CSF-derived macrophages stimulated with LPS (Cuevas VD *et al*., 2022), **d)** upregulated genes of M-CSF-derived macrophages inhibited with MAF siRNA (Vega MA *et al*, 2020) and **e)** upregulated genes of M-CSF-derived macrophages stimulated with 5-HT (serotonin) (Nieto C *et al*, 2020; Domínguez-Soto Á et al, 2017) in our macrophage cell UMAP

Based on trajectory analysis, we observed M1 subsets populated a different branch to those of M2/M0 subsets in all 3 datasets (**Fig. 49 and Annex Fig. 2b**), with inflammatory monocytes exclusively transitioning towards the fully activated M1 state. IDA macrophages instead appear to contain a heterogeneous population divided between the M1 and the M2 branches, suggesting they may represent a transitional state between those subsets. Analysis of overlapping markers between M1, M2 and IDA macrophages reveals that the

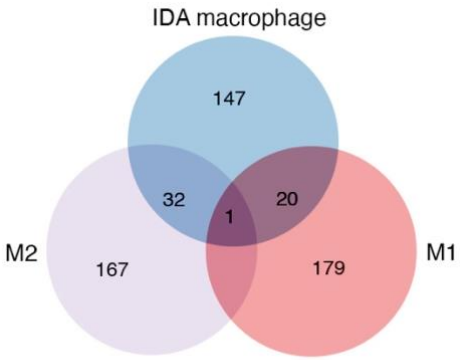


Figure 51. **Inflammation-dependent alternative (IDA) macrophages show a unique signature compared to M2 and M1 macrophages**. Venn diagram showing the overlap between the top 200 markers of each macrophage population (IDA, M2 (M2 & M2.2) and M1 (M1 ACOD1 & M1 CXCL5) macrophages in the scRNA-seq cohort.

latter share about 16% and 10% of its top 200 marker genes with M2 and M1, respectively (**Fig. 51**).

Overall, we show that in the context of inflammation, macrophages can adopt diverse transcriptional signatures, with high heterogeneity between patients. Our data also suggests that intestinal macrophages could originate from monocytes activated under different stimuli including GM-CSF, GM-CSF+LPS, M-CSF+LPS or M-CSF+5-HT, highlighting the importance of the microenvironment in modulating their phenotypes.

Inflammation-dependent alternative macrophages are expanded in IBD and express neuregulin 1

As mentioned before, we observed in our data a new macrophage profile, IDA macrophage, that was different from the M1 inflammatory macrophages, but appear only in the IBD samples, especially in UC. Moreover, only in vitro 5-HT stimulated monocytes were lightly similar to this profile. The markers that we observed in IDA macrophages include epidermal growth factor receptor (*EGFR*) family ligands like AREG and *HBEGF*

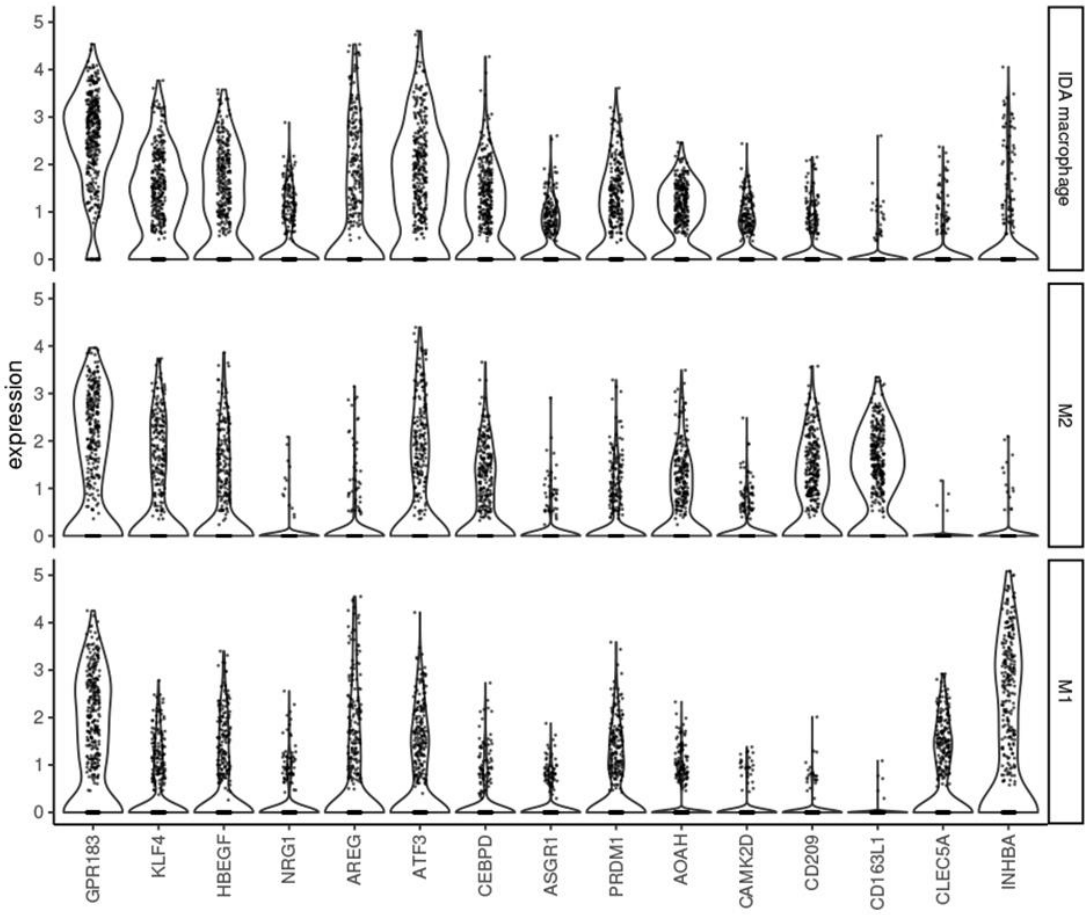


Figure 52. **Neuregulin 1 expression and function in colonic mucosa**. Violin plot showing the expression of marker genes of IDA, M2 and M1 macrophages from pooled HC, CD and UC scRNA-seq data

and specifically *NRG1*, while showing lower expression of M1 and M2 canonical markers (**Fig. 52**).

Since this macrophage subset that expresses *NRG1* was only present in IBD, we next wanted to check the expression of these gene in the bulk RNA-seq data. In agreement with scRNA-seq data (**Fig. 53a**), *NRG1* was significantly increased in bulk RNA-seq analysis of UC colonic mucosa compared to HC and CD (**Fig. 53b**), suggesting the

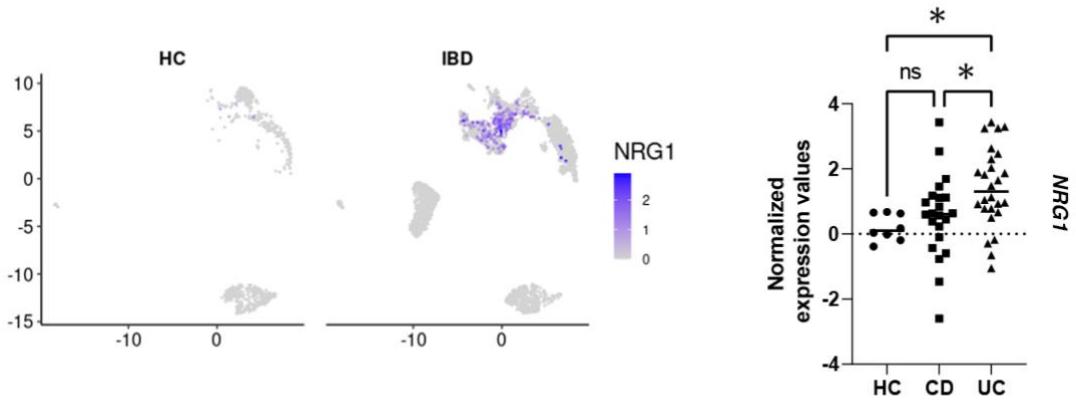


Figure 53. **Neuregulin 1 expression and function in colonic mucosa. a**, UMAP showing *NRG1* expression in the myeloid compartment of HC and IBD (CD and UC) data. **b**, *NRG1* expression from bulk biopsy RNA-seq data in HC (n=8), and active CD (n=22) and UC (n=26) patients. Ordinary one-way ANOVA corrected (p<0,05(*), ns>0,01)

potential involvement of neuregulin 1 in these patients. To validate the upregulation of NRG1 in IBD, we took advantage of *in situ* hybridization to check *NRG1* mRNA in the tissue. Moreover, using a double *in situ* hybridization for NRG1 and immunohistochemistry for pan-macrophage marker CD68, we confirmed its expression on abundant $CD68^+$ macrophages in IBD, demonstrating its existence (**Fig. 54**). Interestingly, *NRG1* expression in HC was also detected but it was limited to a specific population underlying the surface epithelium, with no colocalization within $CD68^+$ macrophages (**Fig. 54**). Further results on this population are unraveled in Chapter 4.

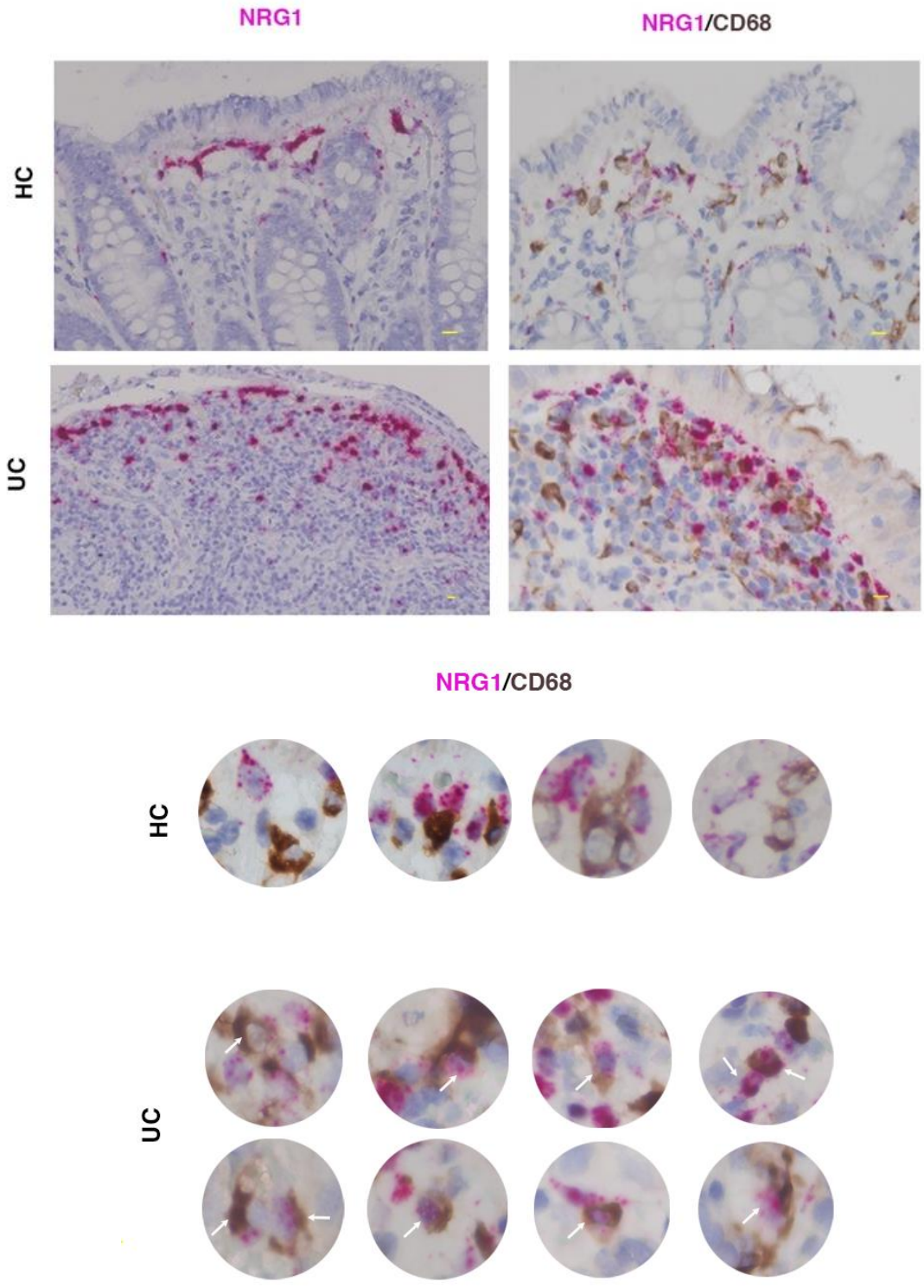


Figure 54. Presence of NRG1+ cells in HC and UC tissue. **a**, Single *in situ* hybridization of *NRG1* and double with immunohistochemistry for CD68 in HC and active UC tissue. **b**, Zoom in of double *in situ* hybridization of *NRG1* with immunohistochemistry for CD68 in HC and active UC tissue. $CD68^{+}$ $NRG1^{-}$, $CD68^{+}$ $NRG1^{+}$ and $CD68^{-}NRG1^{+}$ cells are shown. White arrows show $NRG1^{+}$ $CD68^{+}$ cells in UC patients (Scale bar= 10 µm).

Distinct tissue distribution of IDA macrophages is observed in UC and CD

We next aimed to locate IDA macrophages in the tissue using CosMx SMI analysis. Surprisingly, we localized abundant IDA macrophages scattered throughout the inflamed (UC and CD) colon, representing the most expanded inflammation-dependent macrophage state (**Fig. 55a**), while M1 macrophages were less abundant in the lamina propria and submucosa, but predominated within surface ulcers (**Fig. 44**). Of note, in one CD patient we found abundant granulomas (**Fig. 55b**, **56**, **57** and **58**). Interestingly, IDA macrophages, together with some M2, and a few M0 and M1 macrophages, were the predominant macrophage state within granulomas (**Fig. 55b**), which were surrounded by diverse lymphoid subsets (**Fig. 56**). This pattern was consistent across all 9 granulomas analyzed within this patient. In contrast, a lymphoid follicle also within the submucosa of this CD patient contained B cells and T cells that were the major constituents, and macrophages where predominantly located outside the follicle.

In agreement with the CosMx SMI results, immunostaining showed low and scattered staining of the M2 (CD209) markers within $CD68^{+}$ cells in granuloma (**Fig. 55b, 57 and 58**). Confirming M2 are not the major constituents of granulomas. Compared to lamina propria macrophages, *NRG1* expression within the granulomas was low (**Fig. 58**).

Thus, IDA macrophages abundantly present in the inflamed colon display differential NRG1 expression depending on their tissue location. While $NRG1^{hi}$ IDA macrophages localize to the most apical subepithelial compartment of the mucosa, $NRG1^{low}$ alternatively activated macrophages accumulate within granulomas in CD and in the submucosa of both UC and CD patients, suggesting independent functions.

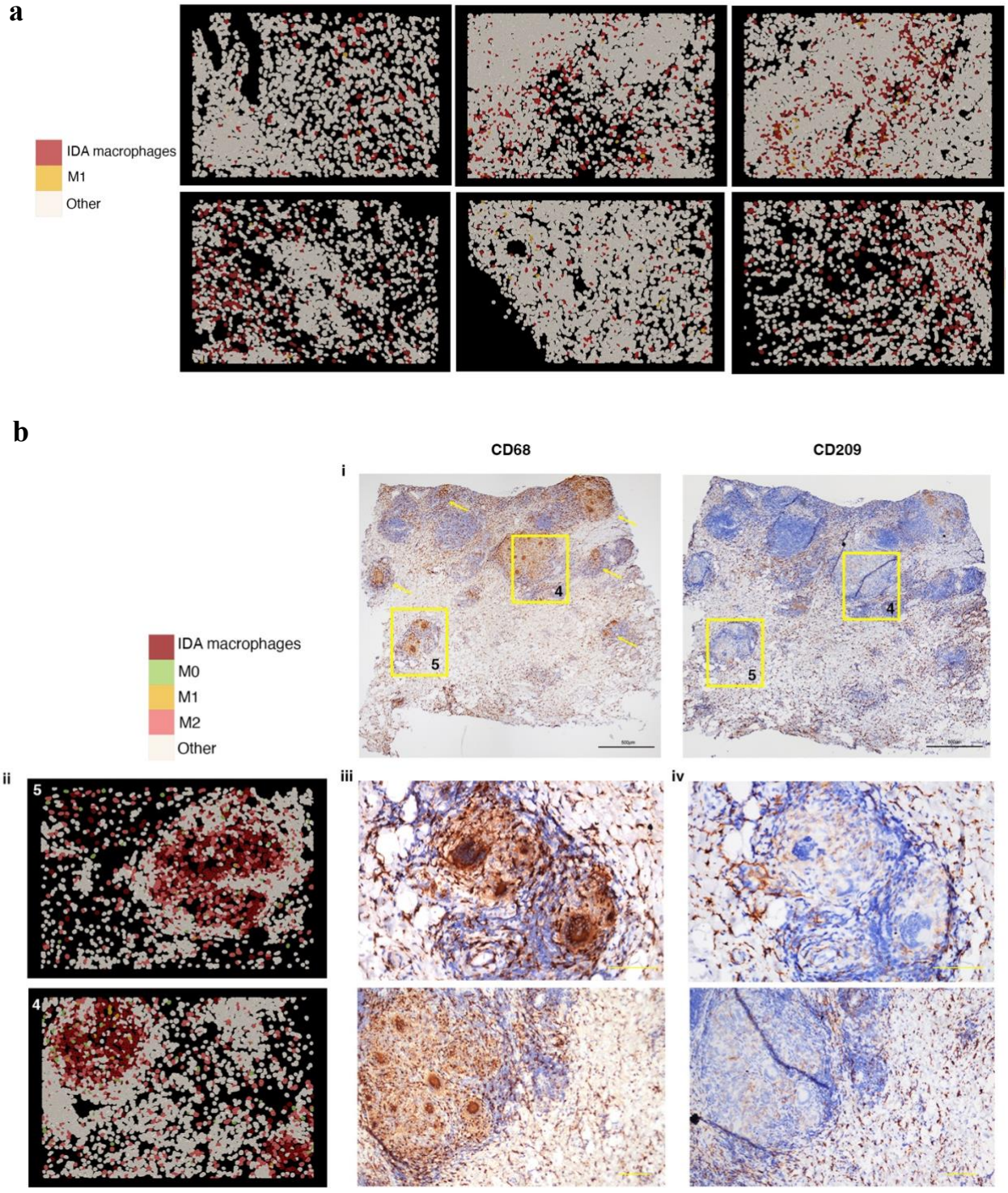


Figure 55. **Inflammation-Dependent Alternative (IDA) macrophages are widely distributed in ulcerative colitis (UC) and present in Crohn's disease (CD) granulomas. a,** CosMx™ SMI distribution of IDA and M1 macrophages in IBD colonic samples. **b**, CD colonic sample with multiple granulomas (CD patient). (i) Field of Views (FoVs) 4 and 5 from this tissue section are indicated by squares and other granulomas found in the same sample by yellow arrows. (ii) Macrophages within granulomas are shown by CosMx™ SMI in FoVs 4 and 5 and protein expression of (iii) CD68 and (iv) CD209 is shown by immunohistochemistry on the same tissue sections (scale bars= 100 µm)

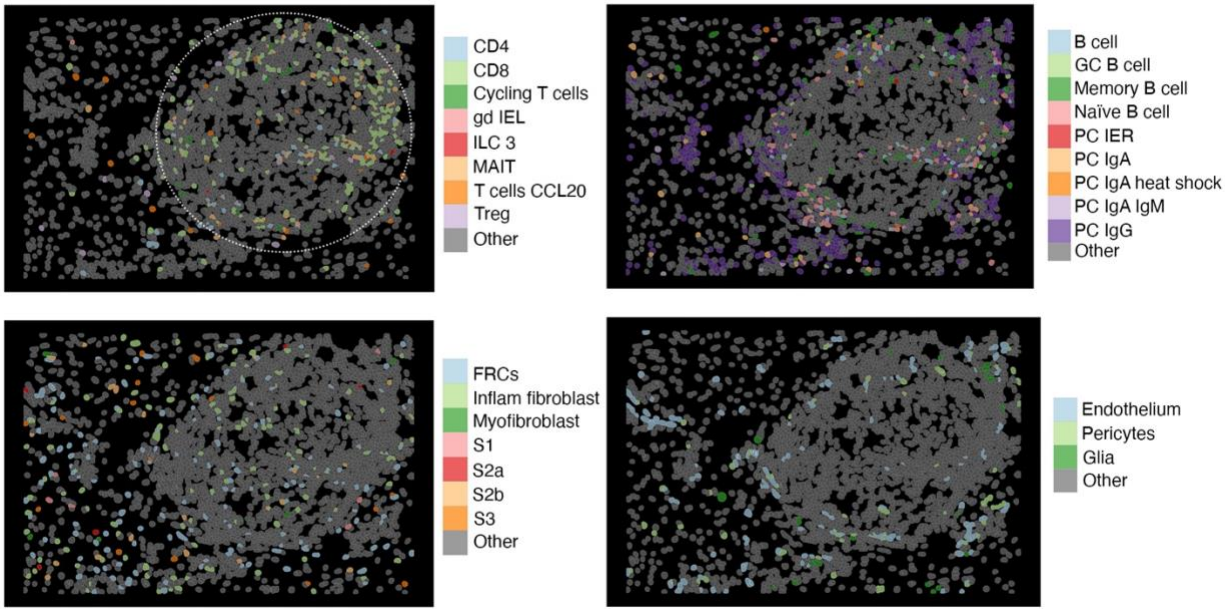


Figure 56. Other cellular components present in CD granuloma's. CosMx^{TM} SMI images showing diverse cell types (stroma cells, T cells, B and plasma cells) present in the granulomas and surrounding area of the analyzed sample. Granuloma is shown by a dotted circle

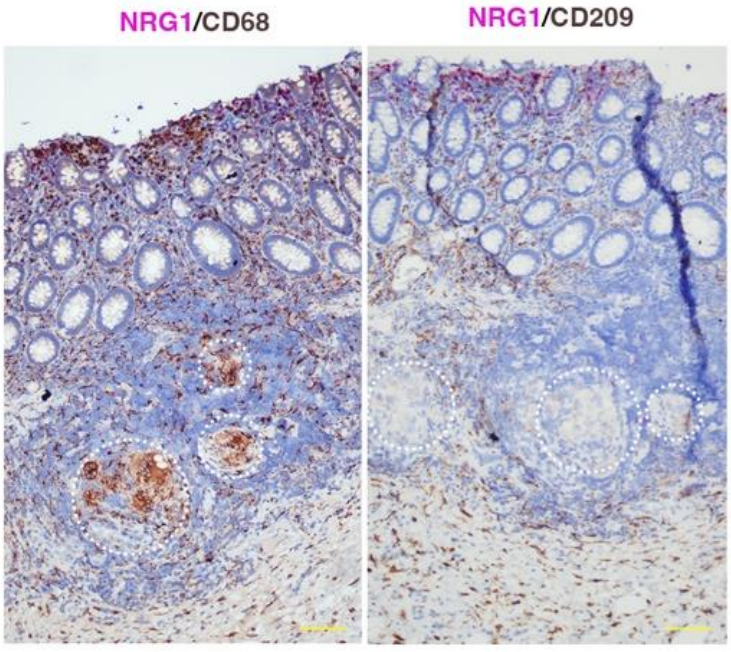


Figure 57. NRG1 localizes near the epithelium and granulomas contain very few CD209+ cells. Double *NRG1 in situ* hybridization and CD68 or CD209 immunostaining in tissue sections from the CD patient (CD b) containing abundant granulomas. *In situ* hybridization of *NRG1* shows an increasing gradient of expression towards the apical mucosa. Granulomas are indicated by dotted circles (scale bars 100 µm

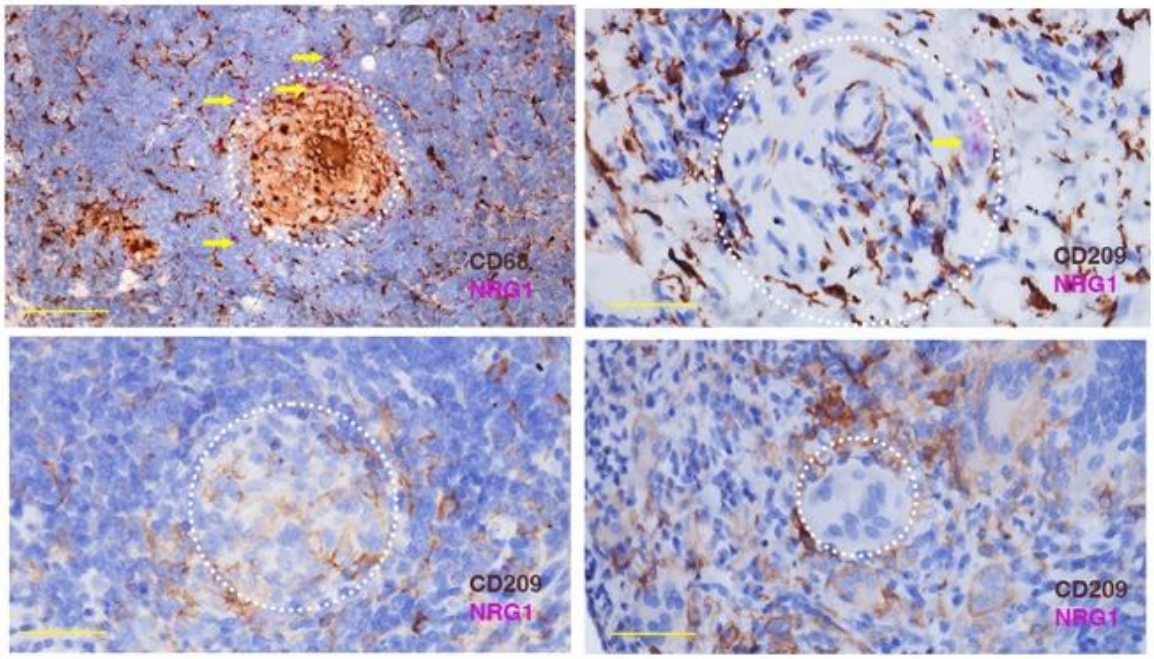


Figure 58. **Low presence of NRG1 in the granulomas of CD that contains IDA macrophages**. Magnified pictures of representative granulomas of the same CD tissue stained for *NRG1* using *in situ* hybridization and CD68 or CD209 immunostaining. Granulomas are indicated by dotted circles and *NRG1* positive cells are shown by arrows (scale bars 100 µm).

SECTION 4: Fibroblasts suffer a rewiring in IBD

The mesenchymal compartment comprises a group of heterogeneous cell types that are essential for maintaining tissue architecture and promoting vascularization and regeneration upon injury. In agreement with a previous study[232], in our scRNA-seq data the stromal compartment in the colon showed high heterogeneity including several types of fibroblasts populations present in the healthy and the inflamed colonic mucosa (**Fig. 59**). Indeed, the stromal compartment showed a wide range in the Morisita-Horn Similarity Index (**Fig. 33**) confirming a great heterogeneity of this population comparable to the already discussed for myeloid cells.

In the stromal compartment we were able to identify all expected cell types. 17 clusters were observed, finding endothelial and activated endothelial cells, lymphatic endothelium, pericytes, glial cells, myofibroblasts and different types of fibroblasts. Lamina propria S1 fibroblasts already described by Kinchen *et al.* were found in two different clusters S1 and S1.2. Regarding pericryptal fibroblasts, we observed the two described clusters S2a and S2b. We also observed S3 fibroblasts and also a small cluster of Fibroblastic Reticular Cells (FRCs). Interestingly, a cluster of inflammatory fibroblasts was also found, and only present in IBD samples (**Fig. 59**).

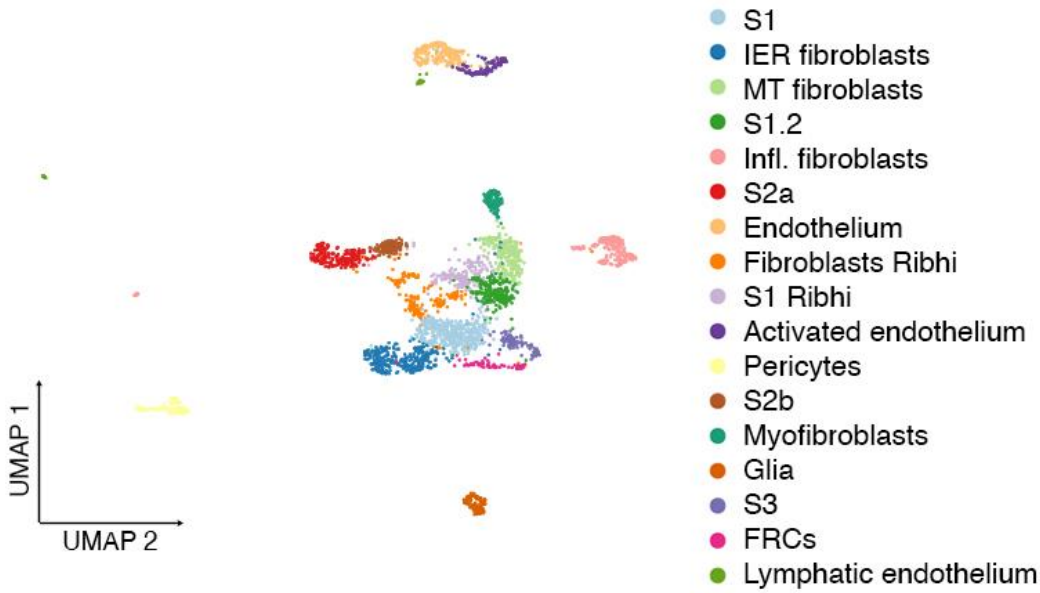


Figure 59. Stromal cell populations analyzed by single-cell RNA-seq (scRNA-seq). UMAP of stromal clusters observed by scRNA-seq cohort samples

Moreover, using SMI, we were able to locate all stromal cell populations in healthy and IBD tissue (**Fig. 60**).

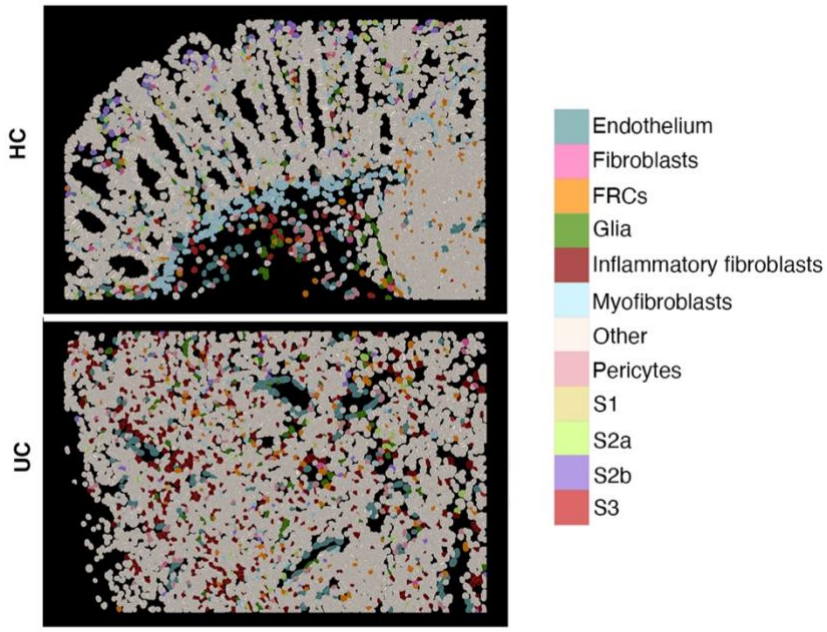


Figure 60. Stromal cell populations analyzed by Spatial Molecular Imaging (SMI). Spatial localization of stromal cells observed by CosMx™ SMI in representative healthy and inflamed UC colonic mucosa-

S2b peri-cryptal fibroblasts express NRG1 and it is highly upregulated in the fibroblast compartment in IBD

S2 peri-cryptal fibroblasts, also known as telocytes (Degirmenci et al., 2018; Shoshkes-Carmel et al., 2018), express SOX6, F3, AGT, COL4A6, COL4A5, BMP7, and PTX3. These fibroblasts consist of two clusters of cells, S2a and S2b that are located near the epithelium along the crypt, and we detected them in all patients and healthy controls with no difference between groups. While they play an essential role in epithelial development and support (Kaestner et al., 1997; McCarthy et al., 2020)

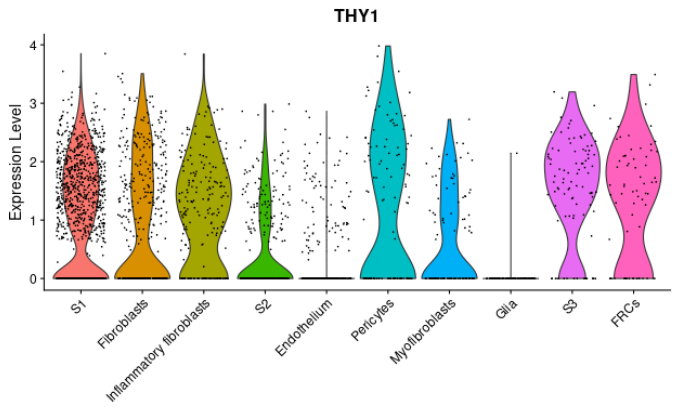


Figure 61. Pan-fibroblast marker THY1 has diminished expression in S2 fibroblasts. Violin plot visualization of THY1 expression in stromal populations by scRNA-seq. S1 and S2 clusters have been pooled together.

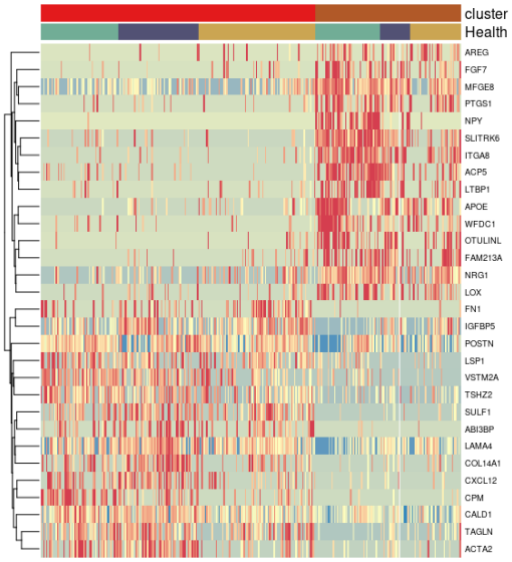


Fig. 62. S2a and S2b fibroblasts present a different transcriptomic profile. Heatmap of top marker genes discriminating S2a and S2b fibroblast subpopulations in scRNA-seq data.

(Fawkner-Corbett et al, 2020), they show significantly reduced expression of THY1 (pan-marker or fibroblasts) compared to other fibroblast clusters (**Fig. 61)**. Using single-cell RNA sequencing, we observed that S2a and S2b have distinct transcriptomic profiles (**Fig. 62**) and SMI showed different spatial localization (**Fig. 63**) suggesting probable different functions. S2b fibroblasts were located at the top of the crypt area next to differentiated cells, while S2a were located at the middle or bottom of the crypt (**Fig. 63**)

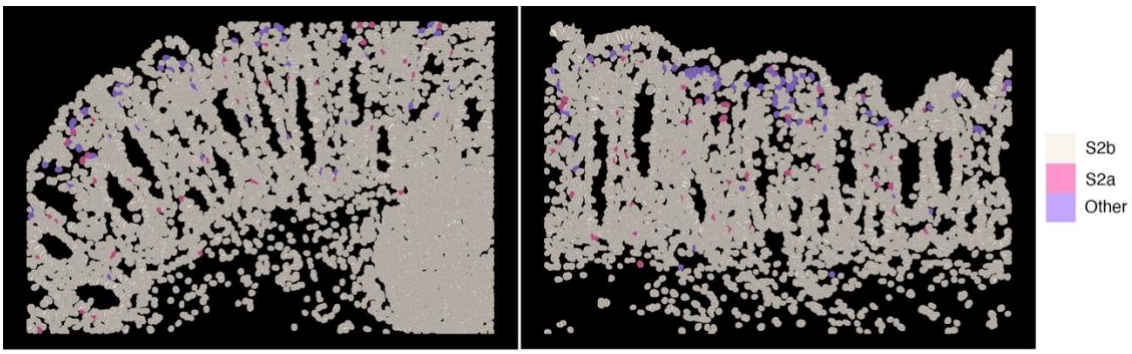


Fig 63. Stromal cell populations analyzed by and Spatial Molecular Imaging (SMI). S2a and S2b differential spatial distribution showed by CosMx™ SMI in 2 representative FoVs from healthy colon.

To explore the differences between these two fibroblast subtypes, we performed differential gene expression analysis and found that some genes were uniquely expressed in S2b (**Fig. 64**).

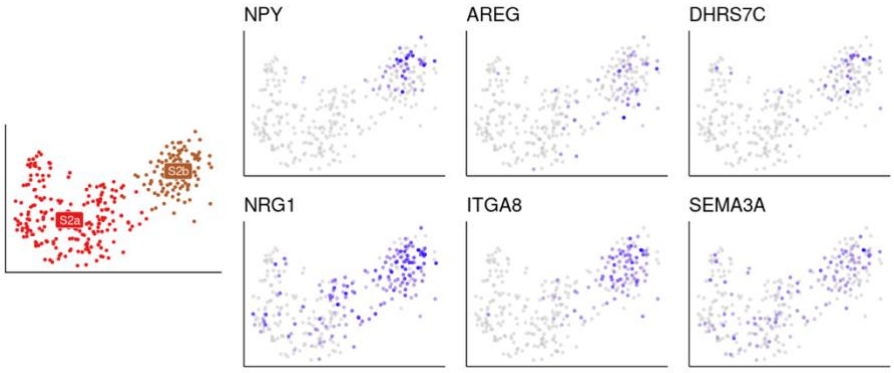


Figure 64. Specific genes expressed in S2b fibroblast subtype. On the left, UMAP showing both S2 cells. On the right, UMAP representation of genes that we found to be specific of S2b fibroblasts.

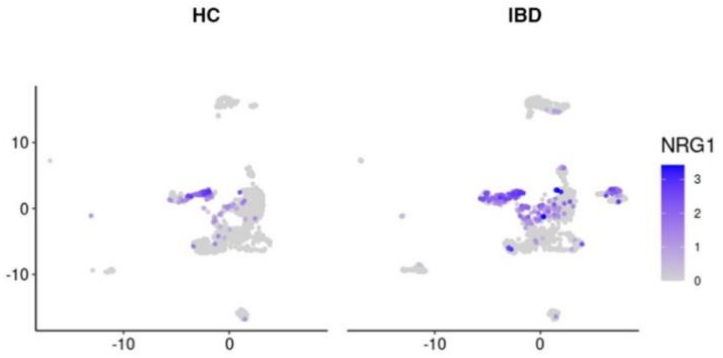


Figure 65. NRG1 expression in the stromal compartment. UMAP representation of the expression of NRG1 in the stromal compartment in HC and IBD cells.

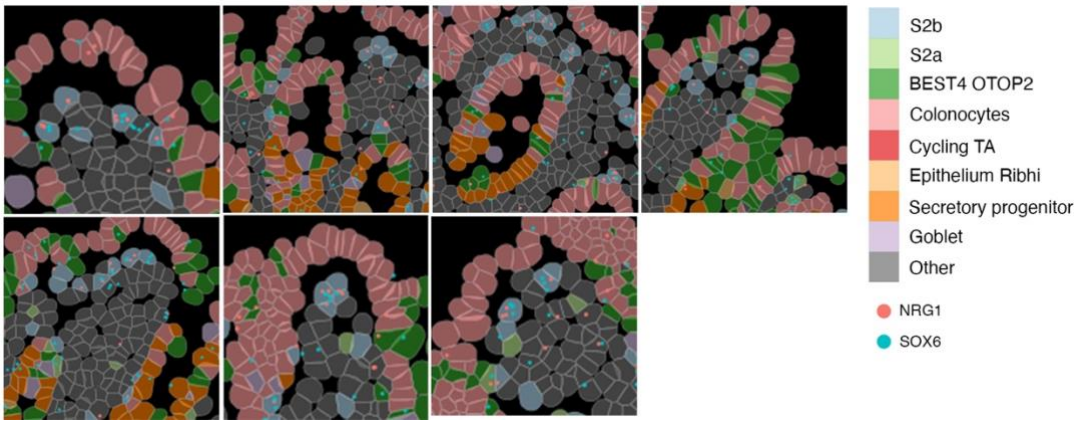


Figure 66. S2b fibroblasts express NRG1 in healthy tissue. Representative images of healthy tissue are shown using CosMx SMI, with S2a fibroblasts depicted in light green and S2b fibroblasts in light blue, in close proximity to the epithelium. Blue dots denote the S2 pericryptal marker SOX6, while red dots represent NRG1 mRNA molecules. Notably, both SOX6 and NRG1 dots were detected within S2b cells

Notably, NRG1 was exclusively expressed in S2b cells in healthy tissue, while its expression increased in other fibroblasts, particularly S2b and inflammatory fibroblasts, in IBD (**Fig. 65**). Thus, the upregulation of NRG1 in IBD patients that we observed by bulk RNA-seq analysis (**Fig. 53b**) is not exclusively coming from the new IDA macrophage but also from the fibroblast compartment as these results highlights **(Fig. 67**).

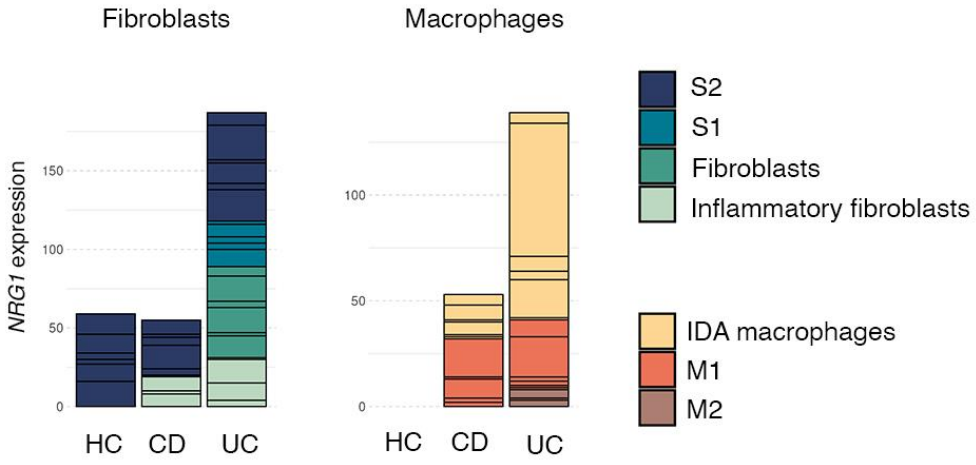


Figure 67. In healthy tissue, NRG1 is primarily derived from S2b fibroblasts. In the context of IBD, however, both IDA macrophages and the fibroblast compartment may contribute to NRG1 expression. Barplots showing *NRG1* expression in the fibroblast and macrophage compartments in HC and IBD.

Overall, our findings suggest that S2 peri-cryptal fibroblasts, particularly S2b, play a crucial role in epithelial development and support, and their unique gene expression profile may have implications for IBD pathogenesis.

Neuregulin 1 secreted by macrophages and fibroblasts acts on the epithelium increasing the expression of OLFM4

Neuregulin 1 is a member of the epidermal growth factor family that binds to ErbB receptors[254] present in the epithelium. In humans, it promotes epithelial differentiation to secretory lineages[255], while in mice, it enhances stem cell proliferation and regeneration[256]. Given that the epithelium expresses receptors for NRG1 (**Fig. 68**), we

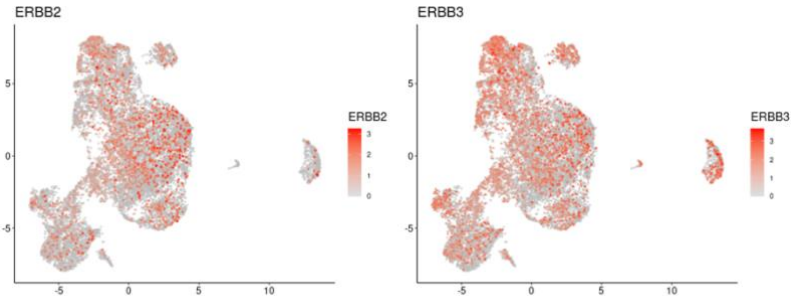


Figure 68. Erbb receptors are expressed in the intestinal epithelium. UMAP showing the expression of *ERBB2* and *ERBB3* in the epithelial compartment of pooled data from HC and IBD samples.

sought to investigate its function further. Using intestinal epithelial stem cell-enriched organoids, we stimulated them with Neuregulin 1 for 48 hours and observed a significant decrease in the stem cell marker LGR5 and an upregulation of OLFM4, without any changes in the proliferation marker MKI67 (**Fig. 69**). OLFM4 expression is mainly restricted to a progenitor cell type in healthy tissues, while it is dramatically increased in ulcerative colitis (UC) and Crohn's disease (CD), as shown in Chapter 2 of this thesis. Our findings suggest that NRG1 may increase OLFM4 expression in the epithelium of patients with IBD.

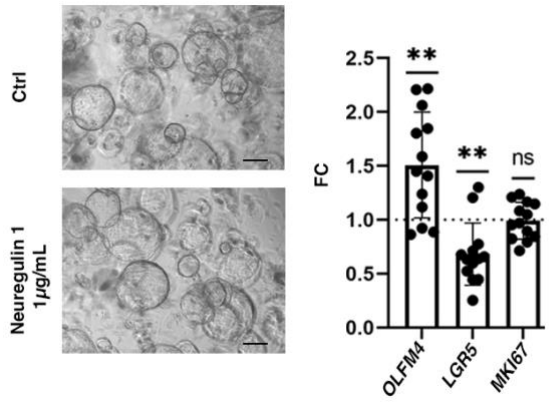


Figure 69. Neuregulin 1 increases OLFM4 expression and decreases stem cell marker LGR5 in epithelial organoids. Representative pictures of human-derived epithelial organoids under vehicle (Ctrl) or Neuregulin 1 (1 mg/mL) for 48h. Scale bars= 100 µm. mRNA expression of OLFM4, LGR5 and MKI67 was determined by RT-qPCR (n=13). Data is shown as fold change (FC) relative to the vehicle treated condition. Bars represent mean ± standard deviation (SD). p<0,05(*), p<0,01 (**), p<0,001(***), ns: not significant

Overall, we demonstrate that NRG1 is overexpressed in IBD, particularly in UC patients, by IDA macrophages and S2b fibroblasts. In addition to other functions, Neuregulin 1 promotes the expansion of the transit-amplifying epithelial compartment, which could play a role in the regeneration of the epithelium.

S2b fibroblasts lose expression of the neuropeptide NPY in ulcerative colitis

In our study we found that S2b cells from healthy controls present, in addition to the previously described markers and *NRG1*, a very specific expression of the neuropeptide *NPY* (**Fig. 64**). However, in S2b fibroblasts from IBD patients,

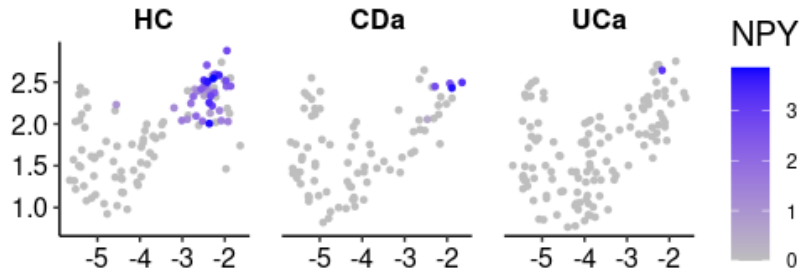


Figure 70. NPY expression in S2 fibroblasts is only present in healthy tissue. UMAP representation of NPY expression in scRNA-seq data in S2 fibroblasts in HC, UC and CD.

we observed a dramatic down modulation of NPY expression (**Fig. 70**). Consistent with our findings, NPY transcription was significantly downregulated in bulk RNA-seq data from UC, as well as in CD, although to a significantly different degree (**Figure 71**). Conversely, S2b fibroblasts were not decreased in sc-RNAseq data, indicating that the decrease in NPY expression is not due to the loss of S2b cells in IBD patients as it is also corroborated by proportion analysis **(Fig 21 and Fig. 75**). IHQ staining confirmed our results showing higher NPY expression in cells surrounding the crypts in healthy donors but none in IBD (**Fig. 72**).

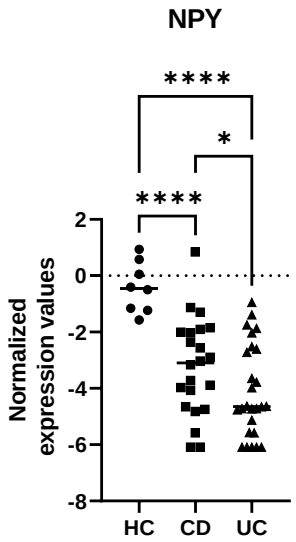


Figure 71. NPY expression is highly downregulated in IBD in bulk RNA-seq data. *NPY* expression by bulk RNA-seq of biopsy samples from HC (n=8), active CD (n=22) and active UC (n=26) patients. p<0,05(*), p<0,01 (**), p<0,001(***), p<0,0001(****), ns: not significant.

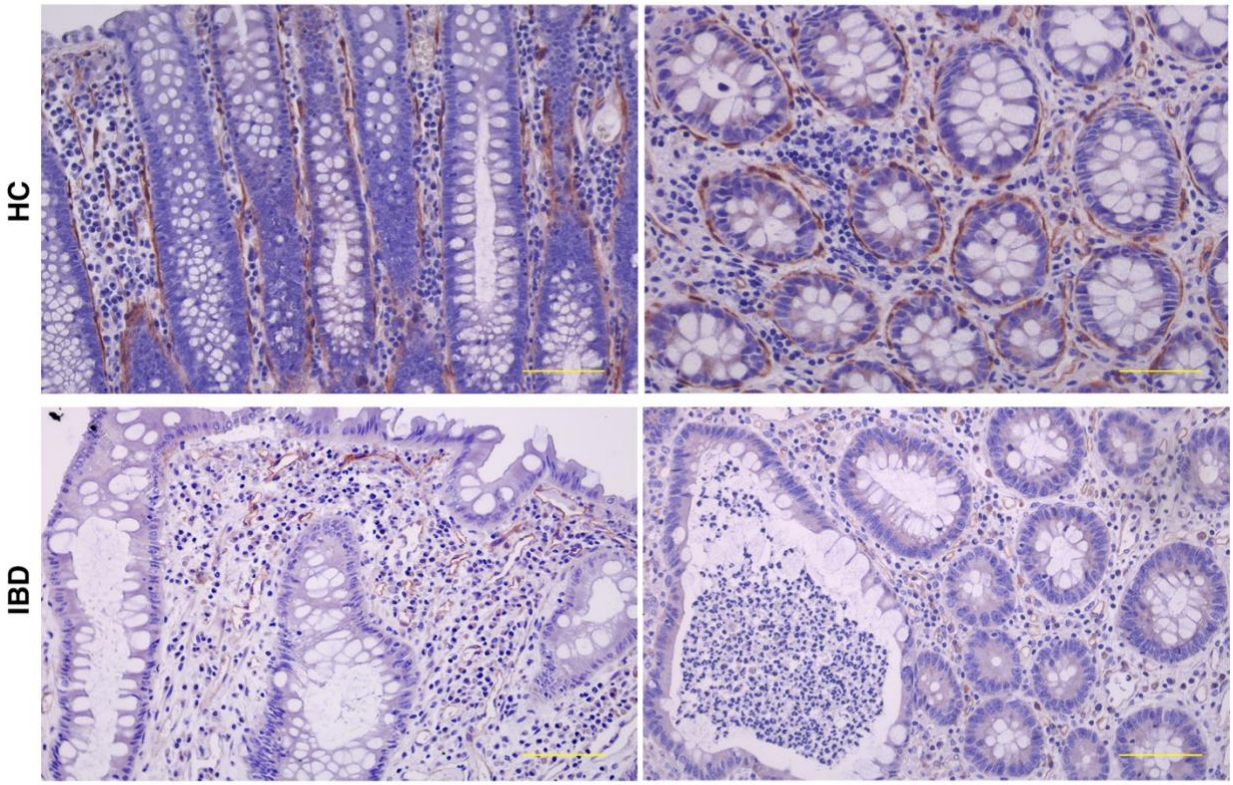


Figure 72. NPY expression is observed in peri-cryptal fibroblasts in healthy and IBD tissue. Representative pictures of protein expression of NPY in representative HC and IBD samples

Inflammatory fibroblasts are present in ulcerative colitis and Crohn's disease.

In our data, a cluster of inflammatory fibroblasts was also observed in colonic CD, besides in UC as described by Kinchen et al. These fibroblasts were characterized by the expression of CHI3L1, IL11, CXCL5, IL24, CA12, WNT2, MMP3, COL7A1, INHBA and FAP (**Fig. 73a**). Remarkably, the increase of inflammatory fibroblasts in IBD was confirmed by the up-regulation of all these genes in bulk RNAseq analysis of colonic biopsies. (**Fig. 73b**).

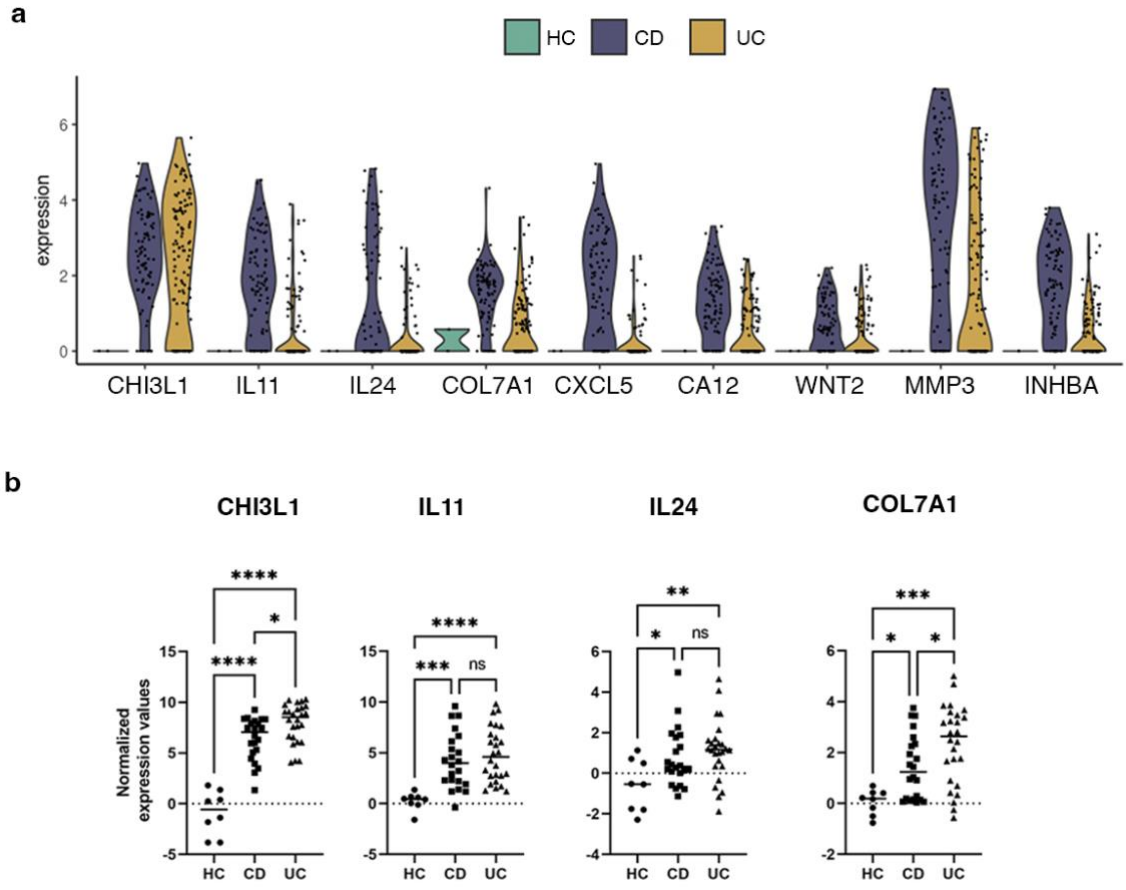


Figure 73. Inflammatory fibroblasts are present in the mucosa if both CD and UC. Violin plots showing expression (y-axis) of marker genes (x-axis) of inflammatory fibroblasts in HC, active CD and active UC determined by scRNA-seq. e, Expression of markers of inflammatory fibroblasts in HC (n=8), and active CD (n=22) and UC (n=26) patients in bulk biopsy RNA-seq data. p<0,05(*), p<0,01 (**), p<0,001(***), p<0,0001(****), ns: not significant.

Additionally, we observed increased frequencies of inflammatory endothelial cells characterized by the expression of *ACKR1*, which also highly expresses adhesion molecules (*SELE*, *MAdCAM1*) (**Figure 59 and Annex Table X**). We also observed a very small cluster of lymphatic endothelial cells (*LYVE1*), and pericytes (*RGS5*, *NOTCH3*) which were increased in IBD patients(**Figure 59 and Annex Table X**). Bulk

RNA-seq also demonstrated the increase in inflammatory endothelium, lymphatic endothelium, and pericytes (**Fig. 74**).

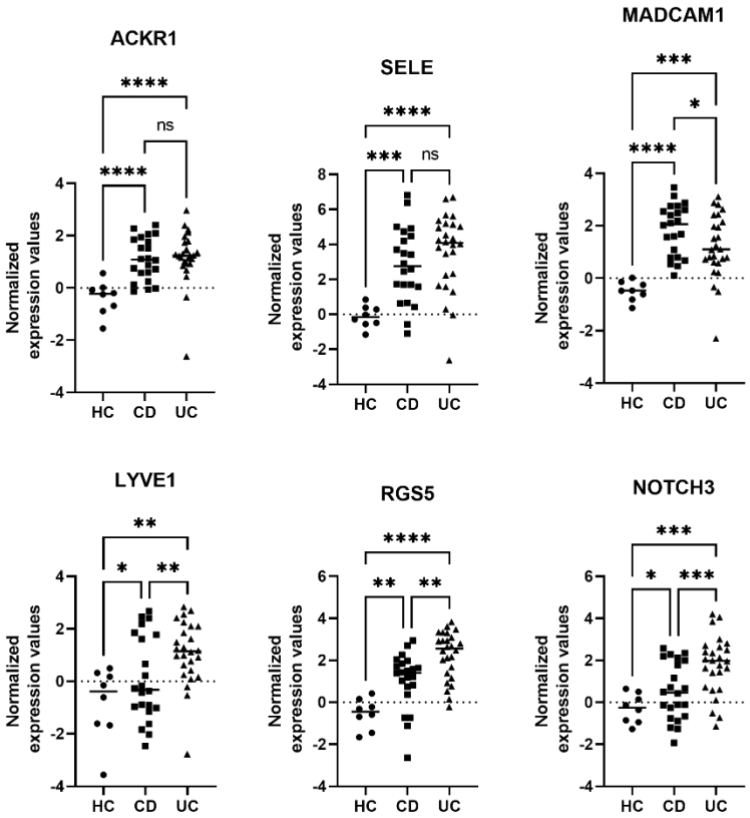


Figure 74. Activated and lymphatic endothelium and pericytes are increased in IBD. Bulk colonic biopsy RNA-seq expression of activated endothelium, lymphatic endothelium and pericytes-specific markers in HC (n=8), active CD (n=22) and UC (n=26) patients. p<0,05(*), p<0,01 (**), p<0,001(***), p<0,0001(****), ns: not significant.

S1 fibroblasts are highly decreased in inflammatory bowel disease and present an inflammatory phenotype

Lamina propria fibroblasts are classified into two subtypes according to a new study[232], S1 and S3 fibroblasts. In our dataset, we identified two clusters of S1 fibroblasts. Interestingly, we observed a significant decrease in the proportion of all S1 fibroblasts in IBD samples in contrast to the increase of inflammatory fibroblasts (**Fig. 75**). Additionally, we noted transcriptional changes in the S1 fibroblasts present in the IBD mucosa. DEG analysis revealed 493 genes that were significantly regulated in all S1 cells between HC and IBD patients (**Fig. 76 and Fig. 77**). S1 cells from IBD patients exhibited downregulation of canonical markers such as ADAMDEC1 and ABCA8, as well as decreased expression of FABP4, CCL8, CFH, ITIH5, CP, GSN, FABP5, CCL2 and FN1 in UC (**Fig. 53** and **Annex Table**). Conversely, S1 fibroblasts in IBD showed an increase in expression of POSTN, S100A6, PDPN, and CHI3L1, which closely resembled the

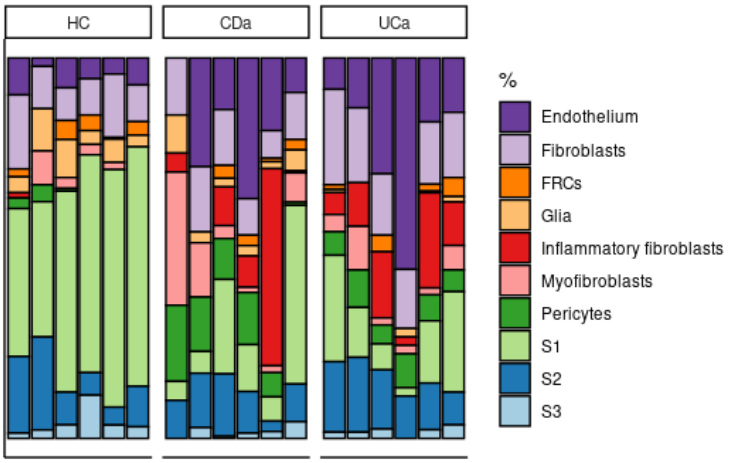


Fig. 75. Stromal cell proportions. Barplot showing pooled frequencies of stromal cell populations in HC, CD and UC. S1 fibroblasts are observed to be highly decreased in IBD samples, while S2 fibroblasts do no present any change in proportion between the groups. An increase in endothelium is also observed in IBD samples.

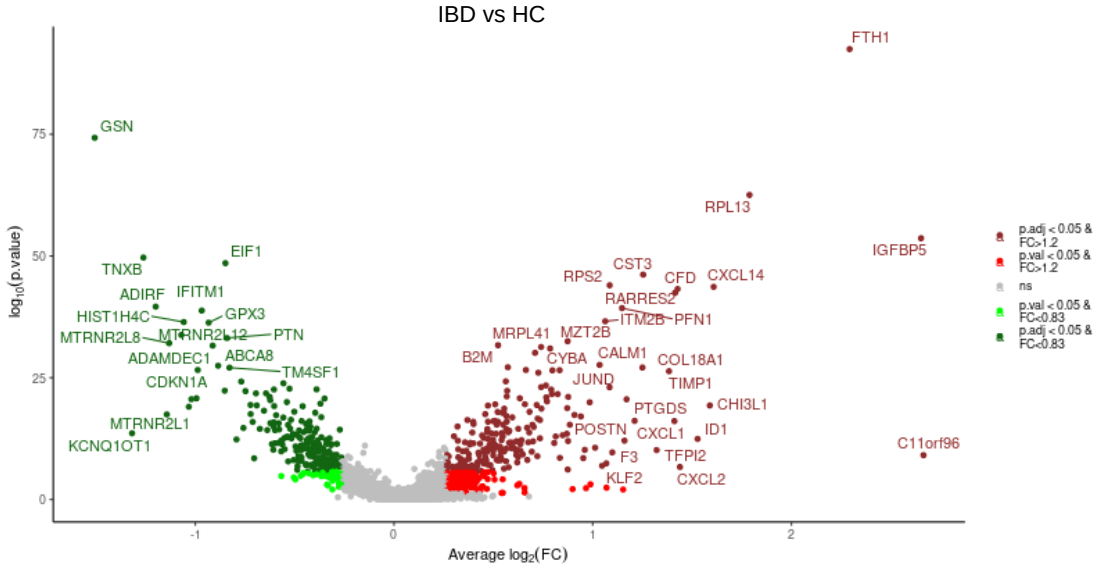


Figure 76. S1 fibroblasts change their transcriptional profile in IBD. Volcano plot of differentially expressed genes (DEGs) in S1 between IBD and HC (Annex Table contains the complete list of genes). Genes upregulated in IBD are shown in dark (UUP, p value<0.05) or light red (UP, nominal p value<0.05). Genes downregulated in IBD are shown in dark (DDW, FDR<0.05) or light green (DW, p value<0.05)

inflammatory fibroblast signature. We also observed upregulation of inflammatory cytokines such as CXCL1, CXCL2, and CXCL6 in IBD S1 cells, which are chemoattractant for neutrophils and granulocytes in general. Interestingly, we also found upregulation of the peri-cryptal marker F3.

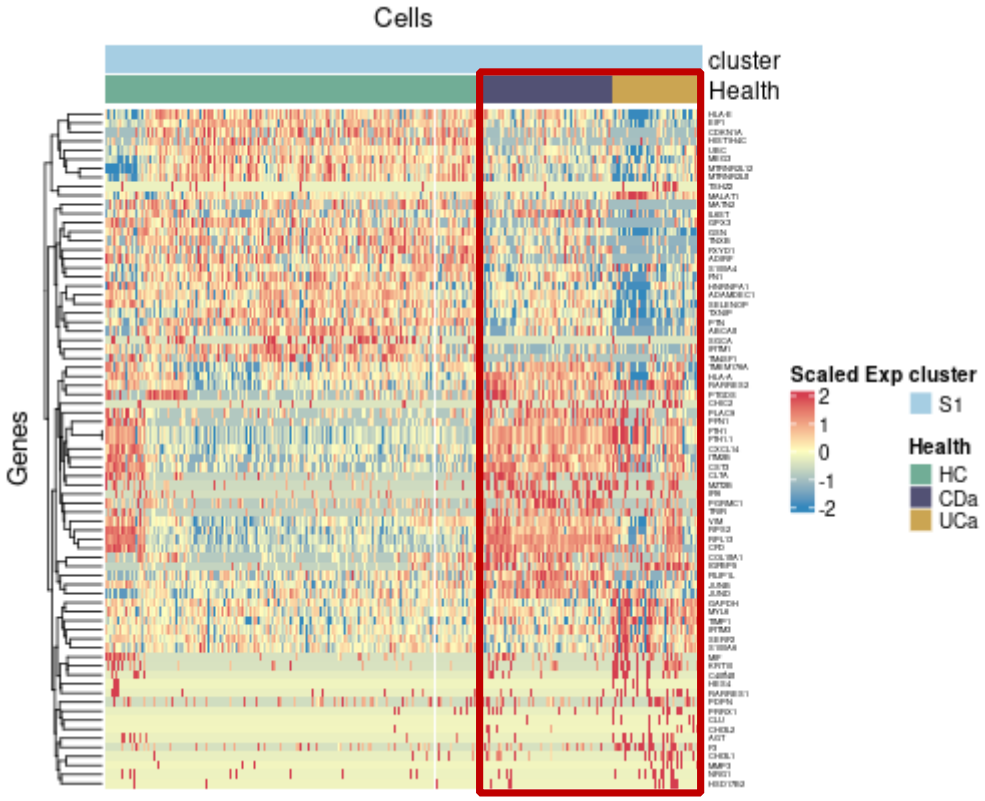


Figure 77. S1 fibroblasts present different signature between HC and IBD. Heat map showing the average expression of DEGs for pooled S1 fibroblasts in HC, CD and UC

Bulk RNAseq analysis supported these results, showing downregulation of ABCA8, FABP4, and ADAMDEC1 S1 markers in active IBD samples (**Fig. 78**). This down modulation of genes in bulk RNA-seq may be attributed to the significant changes in gene expression within the S1 subset in UC and CD, in addition to the reduction in the abundance of S1 fibroblasts in UC.

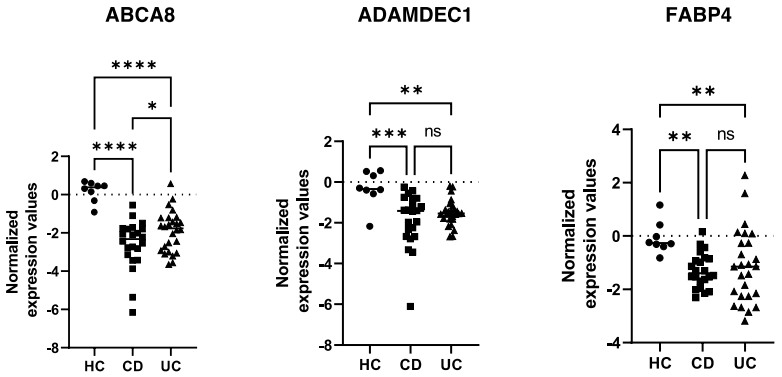


Fig. 78. Expression of S1 fibroblast markers are downmodulated in IBD. ABCA8, ADAMDEC1, and FABP4 expression by bulk RNA-seq of biopsy samples from HC (n=8), active CD (n=22) and active UC (n=26) patients. p<0,05(*), p<0,01 (**), p<0,001(***), p<0,0001(****), ns: not significant

Interestingly, SMI analysis in healthy tissue did not reveal any inflammatory fibroblast, but we confirmed the high abundance of these cells in IBD tissues, scattered throughout the intestinal mucosa. In contrast, abundant S1 cells were observed in healthy tissue, but they were highly decreased in IBD samples confirming scRNA-seq and bulk RNA-seq results. (**Fig 79.**).

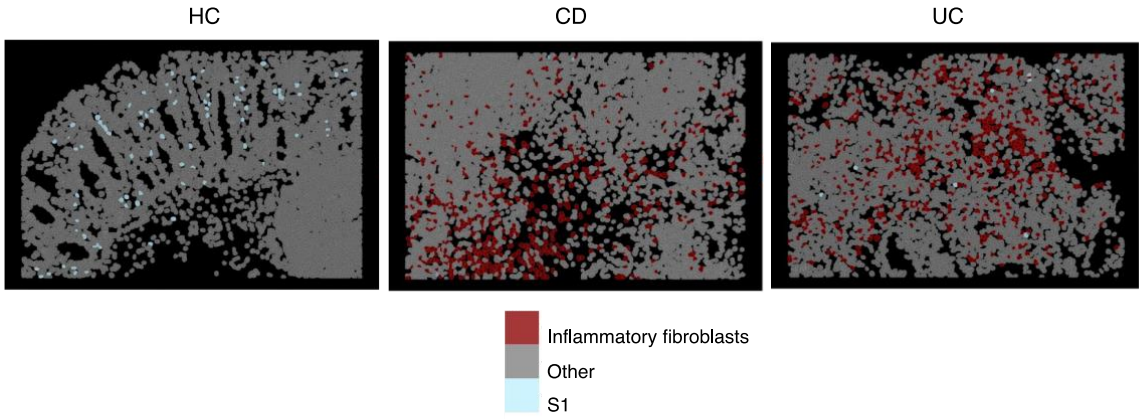


Fig 79. Spatial visualization of inflammatory fibroblasts and S1. CosMx SMI visualization of inflammatory fibroblasts and S1 in representative FoVs of a HC, CD and UC samples. S1 cells are mainly observes in the healthy tissue, while inflammatory fibroblasts predominate in the IBD tissues.

Based on these results, our next objective was to investigate whether S1 fibroblasts can differentiate into the inflammatory fibroblast phenotype within the context of the IBD microenvironment. To accomplish this, we utilized trajectory analysis (**Fig. 80**), a technique that utilizes scRNA-seq data to study the developmental or differentiation pathways of cells. In the trajectory analysis, we found that there were three distinct branches. The first branch mainly consisted of S1 cells, the second branch contained S2 fibroblasts and myofibroblasts, and the third branch was mostly made up of inflammatory fibroblasts (**Fig. 80**).

What is interesting about this analysis is that S1 cells were observed to follow a trajectory path that led them to the end of the third branch, where the inflammatory fibroblasts were located. This suggests that there may be a cellular transformation process happening in which S1 fibroblasts are undergoing changes to become inflammatory fibroblasts, as other results in this chapter pointed out. Furthermore, some S2 fibroblasts were found also to be in close proximity to the inflammatory fibroblasts.

These findings suggests that inflammatory fibroblasts could be the result of a polarization of S1 fibroblasts towards a more inflammatory phenotype.

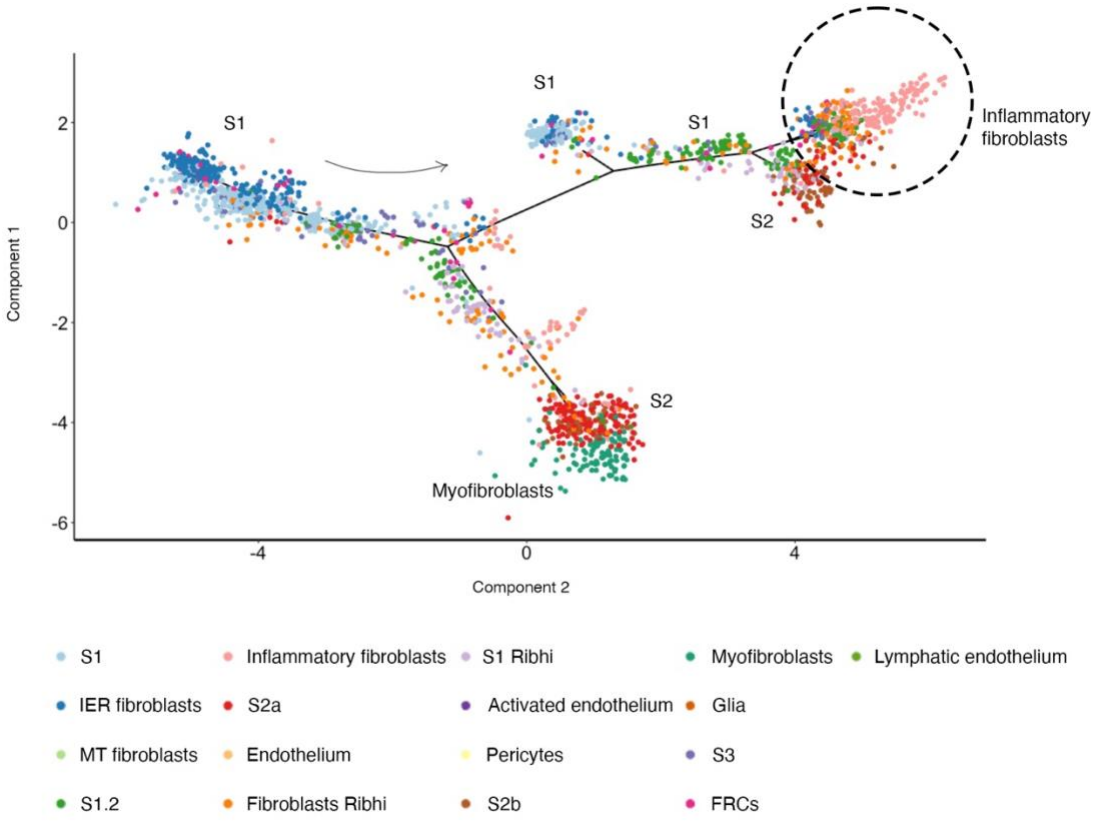


Figure 80. Trajectory analysis shows different distribution of fibroblast populations. Pseudo-time trajectory analysis of the stromal compartment of our scRNA-seq data of pooled HC, CD and UC individuals. S1 fibroblasts are located at the beginning of the trajectory and S2 and myofibroblasts, and inflammatory fibroblasts, in segregated branches.

Overall, our findings suggest that there is a rewiring of the lamina propria fibroblast compartment in IBD, potentially resulting from the interplay between the plasticity observed among fibroblast subtypes and the specific inflammatory milieu in each patient.

SECTION 5: Fibroblasts and macrophage crosstalk

Inflammatory fibroblasts co-localize with Inflammation-Dependent Alternative macrophages in inflammatory bowel disease

Given the abundant number and heterogeneity in distribution patterns of IDA macrophages in UC and CD patients, we leveraged the multiplexed spatial data to identify the cell types that were most frequently found in their proximity. IDA macrophages tended to localize near to other macrophage subsets (M0, M2 and M1), some stromal cells, and T cells, particularly $CD8^+$ T cells, Tregs and T cells CCL20 (**Fig. 81**). Interestingly, within the stromal compartment, IDA macrophages presented high spatial correlation with inflammatory fibroblasts in both UC and CD (**Fig. 82**)

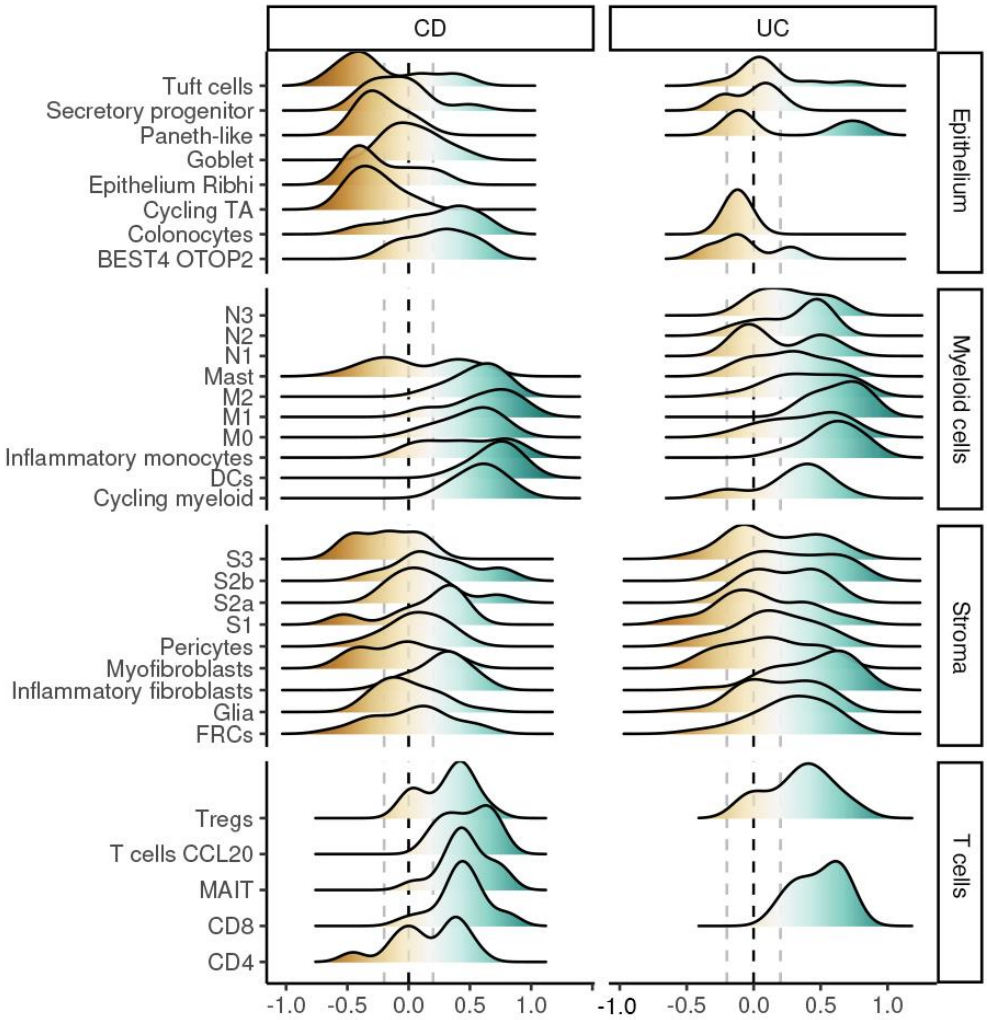


Figure 81. Inflammation-Dependent Associated (IDA) macrophages co-localize with inflammatory fibroblasts. Ridge plot of co-localization analysis of IDA macrophages and epithelial, other myeloid, stromal and T lymphocytes by CosMx™ SMI. Correlation for cell positions was calculated per cell type (0 indicates no correlation, >1 indicates co-localization with 1 being cells sharing the same position; <1 indicates negative correlation between the indicated cell types).

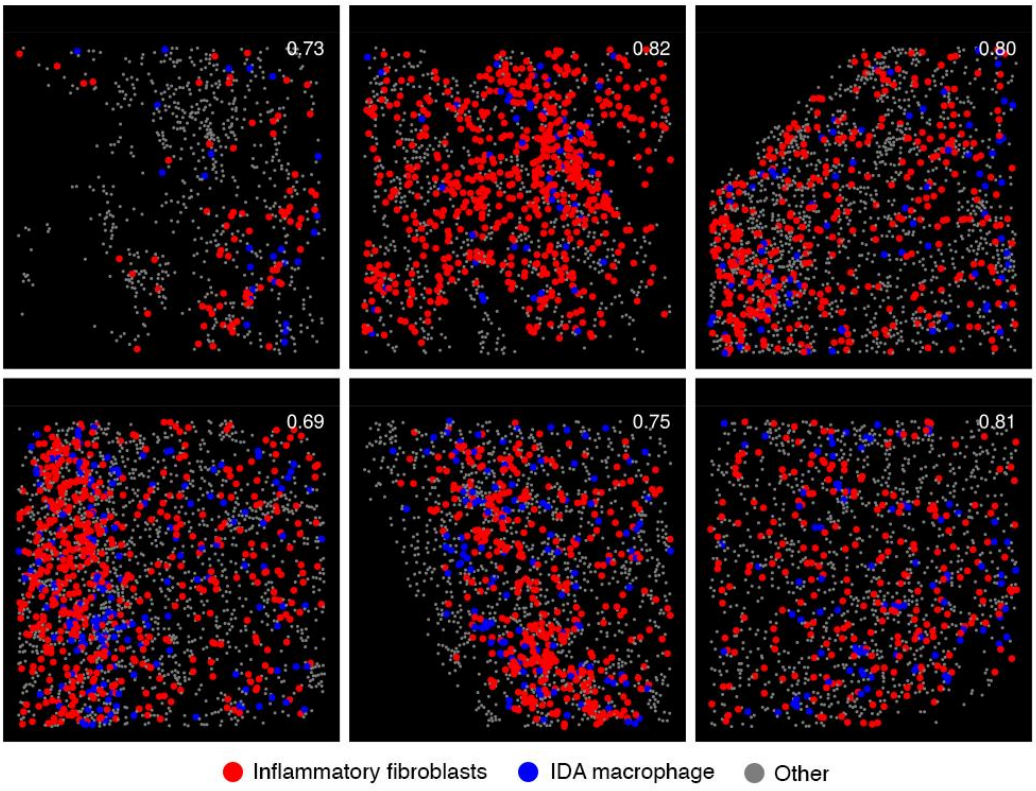


Figure 82. Inflammation-Dependent Associated (IDA) macrophages co-localize with inflammatory fibroblasts. Representative Fields of View (FoVs) of co-localization analysis between IDA macrophages and inflammatory fibroblasts in inflamed UC tissue. Co-localization scores are indicated in white for each FoV.

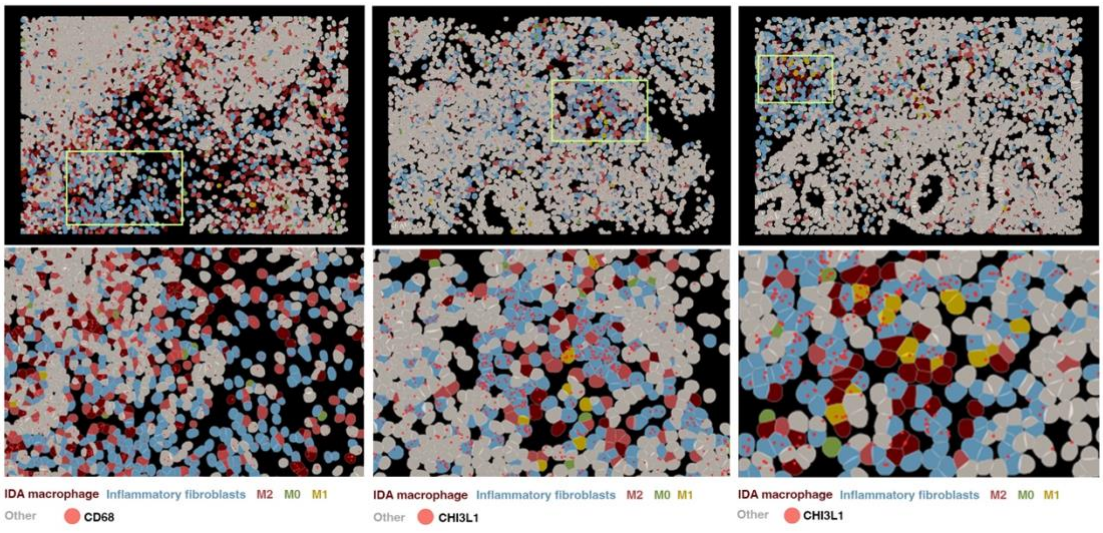
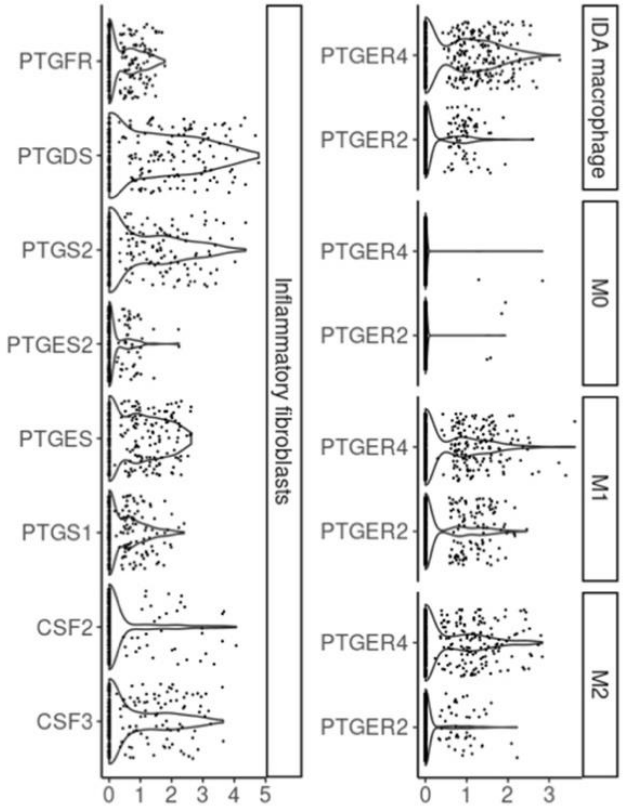


Figure 83. Inflammation-Dependent Associated (IDA) macrophages co-localize with inflammatory fibroblasts. Representative FoVs of IBD inflamed tissues containing IDA macrophages and inflammatory fibroblasts. Expression of CD68 (macrophages) or CHI3L1 (inflammatory fibroblasts) is shown as red dots. Each dot represents a single mRNA molecule.



Importantly, inflammatory fibroblasts expressed *CSF2* and *CSF3*, encoding for GM-CSF and G-CSF, respectively, and prostaglandin-producing enzymes *PTGS1*, *PTGES* and *PTGS2* (**Fig. 84**). In fact, a recent study[257] showed that prostaglandins are produced by activated fibroblasts and drive the differentiation of IDA-like macrophages expressing *HBEGF* and *EREG*, but not *NRG1*, in the synovium of rheumatoid arthritis patients.

Figure 84. Expression of ligan-receptors in macrophages and inflammatory fibroblasts. Violin plot visualizing scRNAseq-based expression (y-axis) of prostaglandin-related genes in inflammatory fibroblasts, IDA macrophage, M2 (M2 & M2.2) and M1 (M1 ACOD1 & M1 CXCL5) in pooled data of HC, CD and UC. Expression of CSF3 and CSF2 in inflammatory fibroblasts has been also included.

Thus, we argue that a crosstalk between inflammatory fibroblasts and macrophages may take place in IBD via specific ligand-receptor interactions.

After identifying IDA macrophages as the primary component of CD granulomas, we investigated whether these cells co-localize with inflammatory fibroblasts within these structures. Interestingly, we observed some inflammatory fibroblasts within the granulomas themselves, but we found a particularly high number of these cells in the areas adjacent to the granulomas (**Fig. 85**).

Overall, the combined use of novel scRNA-seq and single-cell spatial technologies has enables us to elucidate cellular interactions that are highly relevant in the pathophysiology

of IBD. These findings could pave the way for the identification of new therapeutic targets in the treatment of this debilitating condition.

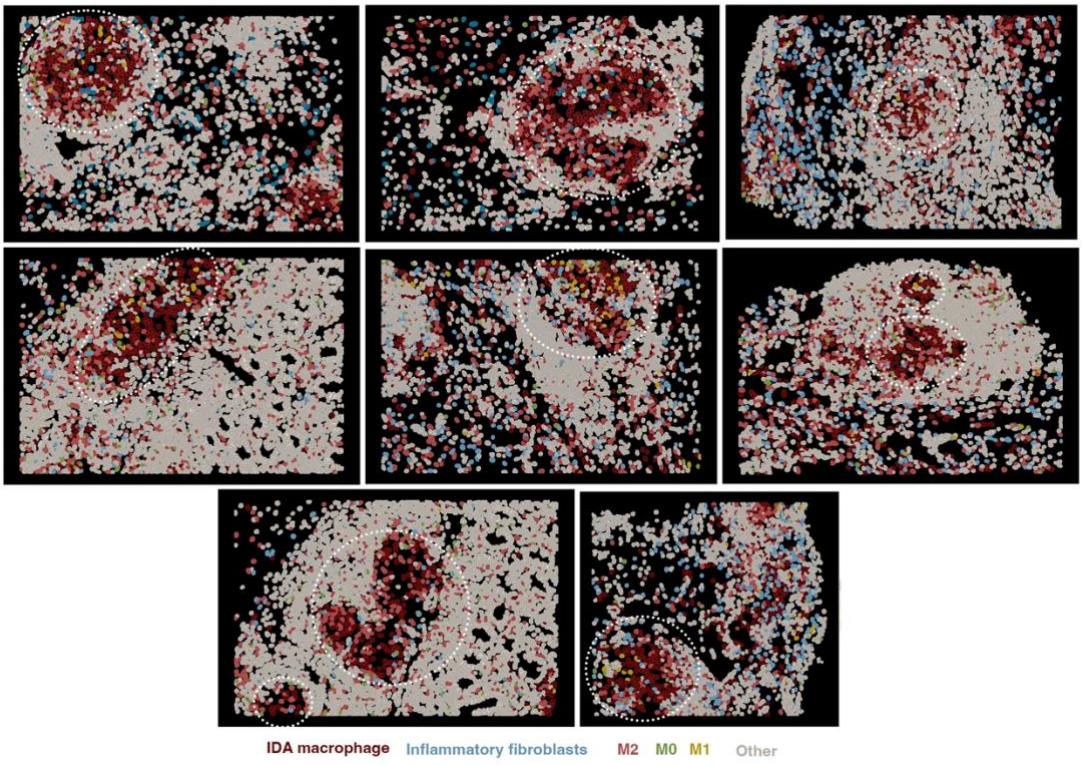


Figure 85. Spatial visualization of inflammatory fibroblasts in inflammatory bowel disease (IBD). a, CosMx SMI visualization of inflammatory fibroblasts within granulomas and the surrounding areas in the CD patient that contained multiple granulomas (CD b). Granulomas are indicated by dotted circles. As shown in Figure 4, granulomas contain abundant IDA macrophages. Inflammatory fibroblasts (light blue) can be found within granulomas and in other adjacent areas.

DISCUSSION

ScRNA-seq has boosted the resolution at which complex tissues, including the inflamed intestine, can be studied[205,223,232,239,250]. Nonetheless, available scRNA-seq datasets lack information on tissue distribution and spatially relevant cell-to-cell interactions. To fill this critical gap, highly multiplexed spatial technologies are rapidly evolving.[258] This thesis is the first to provide combined scRNA-seq data with spatial transcriptomics at single-cell resolution of healthy and IBD colonic tissue to start unraveling patient-dependent disease mechanisms.

Our study has generated a comprehensive map of the colonic epithelial cells in both health and disease, shedding light on the changes that occur in these cells under IBD inflammatory conditions. Overall, we observed a significant loss of epithelial cells in IBD, which is more closely associated with UC. Moreover, there is also transcriptional change in epithelial populations in IBD compared HC. The results of this thesis are consistent with a growing body of research that has investigated the role of the epithelium in IBD pathogenesis[250]. Previous studies had already reported a reduction in the number of epithelial cells in IBD patients[259-261], particularly in the colonocyte and goblet cell subsets. For instance, a study by Haberman et al. demonstrated a reduction in the number of colonocytes and goblet cells in both CD and UC patients. Similarly, a study by Wang *et al*. (2019) reported a reduction in the number of differentiated colonocytes and goblet cells in UC patients, as well as an increase in the number of secretory cells. These results match our findings, since the cell populations that remain in the epithelium of IBD samples were the most undifferentiated ones (Cycling TA and Epithelium Ribhi). Cycling TA cluster represents transit amplifying cells that are in a proliferative state as they express proliferative markers such as MKI67 and TOP2A. Epithelium Ribhi is a cluster of cells that contains high expression of ribosomal genes, that in the literature has been attributed to highly active cells. The presence of these cells (Cycling TA and Epithelium Ribhi) suggests the potential ability to restore the epithelial barrier.

Besides the loss of epithelium observed in patients, some new populations appear in the mucosa of IBD patients, suggesting a rewiring of the epithelium in these specific inflammatory conditions. Notably, we identified a novel unique cell type, a cluster of inflammatory colonocytes, which is exclusively present in IBD and presented expression of *LCN2*, *DUOXA2*, *DUOX2* and *NOS2*. The increase of some of these genes was already described in the epithelium of IBD patients [262], cancer [263]and in response to microorganisms[71,264]. We deciphered that the main source of its upregulation is a specific

type of colonocyte only present in IBD. Moreover, our analysis revealed that other epithelial cell types upregulate also *LCN2* in IBD conditions, which is a disease activity marker in IBD. Increase of *DUOXA2* has been associated with epithelial TLR4 activation in mice. Furthermore, the attachment to intestinal epithelial cells of Citrobacter rodentium, an enteric pathogen, upregulates the expression of *Nos2*, *Duoxa2* and *Duox2*, which are involved in the production of reactive oxygen species, facilitating Th17 cell differentiation in the colonic lamina propria[265]. Interestingly, inflammatory colonocytes become the most abundant epithelial cell in UC patients. These results indicate a global shift in the transcriptome of epithelial cells towards a more inflammatory phenotype Our study also confirms the potential role of Paneth cells in IBD[266] pathogenesis. Paneth cells are primarily found in the small intestine, where they play a crucial role in host defense by producing antimicrobial peptides[267]. However, through scRNA-seq, we observed Paneth-like cells in the colon in IBD patients. The presence of Paneth cells metaplasia in IBD has been extensively reported and this observation is consistent with previous studies, such as a study by Bevins et al. (2009), which demonstrated the presence of Paneth cell metaplasia in UC patients. In our study, we provide the transcriptomic profile of Paneth-like cells, which were more abundant in CD. Moreover, though CosMx SMI, we localized the presence of Paneth-like cells in the bottom of the crypt in inflamed samples.

Additionally, we observed the increase expression of Olfactomedin-4 (*OLFM4)* in the epithelium of IBD samples, and to a higher extent, in UC patients, in agreement with others[268]. Like us, several studies have reported an upregulation of *OLFM4* expression in inflamed samples of IBD patients. *OLFM4* has been described as a stem cell marker in the human and murine small intestine[269] and has been shown to play a crucial role in mucosal defense. However, we show using scRNAseq analysis that *OLFM4* expression in the healthy colon is restricted to the Secretory progenitor and Epithelial Ribhi clusters, with only residual expression within the proliferating epithelial subset (Cycling TA). Additionally, by immunostaining techniques we observed that *OLFM4* in the healthy controls is restricted to the bottom of the crypt. A limitation of this study is that our dataset does not contain stem cells and thus, we are not able to corroborate or rule out *OLFM4* expression in these cells. Previous scRNA-seq data set showed *OLFM4* expression in clusters of undifferentiated secretory epithelium[250], undifferentiated and mitotic undifferentiated (equivalent to our cycling TA)[224] or in a stem and TA subset[205]. Thus,

we and others demonstrate that transcriptional *OLFM4 expression*, while restricted to precursor populations, is not exclusively expressed by classical stem cells (*LGR5*+). Additionally, as shown by protein and mRNA expression, *OLFM4* is not limited to the bottom of the crypt but rather expanded in IBD samples. Indeed, through SMI we could confirm our findings. *OLFM4* is localized at the bottom of the crypt area in healthy samples and expressed though the entire crypt in IBD samples in undifferentiated epithelial cells. Thus, OLFM4 upregulation may serve as a marker of the epithelial intestinal inflammation in IBD, especially in UC.

Besides the palpable loss of epithelium, we observed a high heterogeneity of myeloid cell populations in the IBD samples. Beyond macrophages, recent scRNA-seq studies have explored the diversity of blood neutrophils[270,271]. These short-lived cells, originally thought to exist in fixed states, have more recently been shown to be heterogenous and transcriptionally dynamic, adopting multiple transcriptional states depending on their maturation stage. Ours is the first report to provide scRNA-seq data on intestinal neutrophils. With only 741 neutrophils sequenced, our analysis was able to identify 3 different states of neutrophils in the inflamed colon which corroborates the heterogeneity of this cell type. N1 and N3 were very similar to peripheral neutrophils, including the expression of maturation marker CXCR2[270], which is absent in N2 neutrophils. It has been recently published that neutrophils bearing an IFN signature are present in the peripheral blood of WT mice and human healthy donors representing 7% of the human circulating neutrophils (Wigerblad et al JI 2022). Interestingly, IFN signature and maturation status in blood neutrophils is increased in HC females compared to males (Gupta et al. PNAS 2020).

However neutrophils presenting the IFN signature are expanded in a pletora of different pathologies as proved by many previous studies reporting the apparition or increase of this population in autoimmune pathologies like Systemic lupus erythematosus (SLE) (Mistry et al. PNAS 2019) or VEXAS (vacuoles, E1 enzyme, X-linked, Autoinflammatory, Somatic syndrome), infections[272,273] and cancer, including patients after receiving HSCT treatment[271]. To our knowledge there are no published studies reporting the increase of blood neutrophils with an IFN signature in IBD but the data discussed here suggests that this population might already be present in the periphery of IBD patients further migrating into the inflamed colon.

Remarkably, CXCR4 neutrophils (N2), which also expressed VEGFA, were not found in peripheral datasets. CXCR4, which is essential for bone marrow retention of immature neutrophils, has been identified in mice to mark a subset of pro-angiogenic neutrophils found both in lung and intestine[120], and to be expressed by neutrophils in inflamed tissues[274,275]. Given that neutrophils are not present in the healthy human colon, these data lead to hypothesize that neutrophils are not only an heterogenous circulating population but that they are also plastic cells that can dramatically alter their transcriptional profile to adapt to the gut inflammatory environment. Unfortunately, up to now there is no scRNAseq data available on circulating IBD neutrophils which would shed light on this topic. Further analysis is required to fully understand the origin and function of these neutrophils in the gut.

Macrophages are well-known for their tissue plasticity. The origin, phenotype, and function of intestinal macrophages, however, continues to be a subject of debate[144,276]. Nonetheless, resident macrophages are known to display heterogenous functions[277,278]. In the context of IBD, activated subsets have been described as having a proinflammatory function[147,279]. Cell classification, however, relies on surface markers that may not be consistently used across studies, thus making standardization challenging. ScRNA-seq provides instead unbiased whole transcriptome profiles of cell types, independently of prior knowledge of marker expression. Using unsorted cells, we discovered at least two unique resident macrophage states (M0 and M2) in healthy colonic mucosa. Both subsets were still present in active patients, together with a variety of activated inflammatory macrophages. Remarkably, the profiles of M0 and M2 macrophages were also consistently found in two independent datasets[205,239] and localized by CosMx SMI to the intestinal lamina propria. To our understanding, M0 macrophage has not been observed before in the intestinal mucosa. We argue that M0 population has gone unnoticed due to its low expression of the canonical hematopoietic marker CD45 (*PTPRC*) and the lack of specific markers. The transcriptomic profile of M0 macrophages and their location at the top crypt area suggestive of a role in antigen processing and presentation (*HLA-DPB1*, *HLA-DPA1*, *HLA-DQA1*), antimicrobial activity (*LYZ*) and phagocytosis and autophagy processes (*C1QA*, *C1QB*, *C1QC*). Additionally, we observed two distinct clusters of M2 populations: “M2” highly express M2-canonical markers (*CD209*, *CD163L1*…) while “M2.2” present low expression of those genes. According to trajectory analysis, M2.2 was closer to M0 cells, potentially representing an intermediate state between the fully

differentiated M2 and the M0 population. Thus, M0 might be a long-lived resident macrophage that may differentiate into other phenotypes such as M2.

We observed that the transcriptional signatures of inflammation-associated macrophages varied markedly between patients and datasets. We argue that, compared to canonically differentiated resident macrophages (M0 and M2), inflammatory macrophages adapt their phenotype to a variety of patient-dependent microenvironments. Based on published data from *in vitro* differentiated macrophages and pseudo-time analysis of scRNA-seq signatures, we propose that both infiltrating monocytes and resident M2 macrophages give rise to different phenotypes of activated macrophages. Multiplexed spatial analysis confirmed the diversity in the macrophage populations, and importantly showed that most inflammation-dependent macrophages do not display the full characteristic M1-signature exhibiting instead an alternative activation pattern characterized by the expression of EGFR ligands, *NRG1* and *HBEGF*, and the C-type lectin receptors *CLEC10A* and *ASGR1*. While an M1 signature can be reproduced *in vitro* by exposure of blood monocytes to GM-CSF, GM-CSF/LPS or M-CSF/LPS, the origin of IDA macrophages remains incompletely understood. We found that an endogenously produced factor, serotonin, which is highly abundant in the gut, can induce a signature on M-CSF-derived macrophages (M2-like) that resembles that of the IDA subset found in colonic IBD. Previous studies have shown that serotonin, primarily produced by enterochromaffin cells, modulates macrophage cytokine secretion and its phenotype *in vitro* and that M2 macrophages, in contrast to M1 macrophages, express the serotonin receptors *HTR1D*, *HTR2B* and *HTR7*[33]. Serotonin (5-HT) is elevated in patients with UC and CD[280], however, our data does not prove this to be a relevant mechanism in patients, so it does not rule out the existence of other signals, including fibroblast-derived prostaglandins[257], that could drive this alternative activation to IDA macrophages. Interestingly, IDA macrophages represented the most abundant activated phenotype in the CosMx SMI analysis, while M1 were restricted to ulcerated areas. Despite the wide assumption that blood monocytes are responsible for macrophage replenishment in the gut, a more recently study has challenged this assumption and showed that different macrophage populations reside in the murine gut with none or little dependency on monocyte replenishment[140]. Thus, we argue that M1 could be the result of blood-monocytes and M2 could give rise to IDA macrophages in specific inflammatory conditions. To our knowledge, macrophages expressing neuregulin 1 have only been described in a murine

model of myocardial infarction[281] and suggested to prevent the progression of fibrosis in mouse hearts. The functional role of IDA macrophages in our patients, thus, remains unclear. However, according to our data, the presence of this macrophage phenotype seems more abundant in UC. We observed *NRG1*^{hi} IDA macrophages to localize to the most apical side of the mucosa in UC. In contrast, *NRG1*^{low} IDA macrophages were found within granulomas of a CD patient and in the submucosa of inflamed patients, suggesting IDA macrophages may play different roles depending on their environment. To our knowledge, the phenotype of macrophages within granulomas[279] have not been yet deciphered, and we suggest they do not have the typical M1 profile.

Our study reveals that *NRG1*, an EGFR ligand that promotes epithelial cell growth and survival, is not only produced by IDA macrophages, which are mainly found in UC, but also by the mesenchymal compartment in IBD. Our investigation also uncovered the heterogeneity of the S2 fibroblast compartment, which comprises distinct subtypes: S2a and S2b, with different transcriptomic profiles and spatial locations. Specifically, we observed that S2b fibroblasts are located in the top crypt area, adjacent to more differentiated epithelial cells. Our results show that *NRG1* expression is upregulated in IBD, particularly in UC samples. While healthy tissue primarily exhibits *NRG1* expression in the S2b population, in IBD conditions, other fibroblasts, such as S2a and inflammatory fibroblasts, also express *NRG1*. In response to intestinal damage, *NRG1* expression is induced and contributes to epithelial regeneration in mice following irradiation-induced injury by binding and robustly activating ErbB2 and ErbB3 receptors (Jardé *et al*.). Additionally, Jardé and collaborators observed a strong *NRG1* signal detected in the vicinity of regenerating crypts post-injury, probable accounting for S2b fibroblasts or IDA macrophages. Thus, IDA macrophages and S2b fibroblasts located near the epithelium have the potential to synergize in their functions and play a crucial role in epithelial regeneration.

Our study revealed the interesting finding that NRG1 promotes the upregulation of *OLFM4* in stem organoids, a result that is consistent with previous findings by Holloway *et al* in intestinal enteroids, and diminished expression of stem cell marker *LGR5*. Furthermore, Jardé *et al* also demonstrated that the loss of *NRG1* significantly impaired the presence of *OLFM4*+ cells in mice, while NRG1 treatment increased the expression of *OLFM4* in the epithelium. *Olfm4* deletion significantly enhanced intestinal-crypt proliferation and inflammation in azoxymethane/ dextran sodium sulfate (AOM/DSS)-

induced colitis mice model[282]. *Olfm4*-deficient mice exhibited significant weight loss and lower survival compared to wild type (WT) mice. The colon in *Olfm4*-deficient mice demonstrated severe inflammation and mucosal damage while only mild to moderate inflammation was observed in WT mice[283]. Thus, *OLFM4* has an important role in the regulation of intestinal inflammation and crypt proliferation induced by AOM/DSS. In addition, the *in vitro* model of intestinal damage in small intestinal organoids conducted by Montenegro-Miranda *et al*, showed that *OLFM4* increased during the regeneration phase, but not *LGR5*, suggesting that *NRG1* plays a crucial role in regulating *OLFM4* expression and promoting intestinal epithelial regeneration independently of *LGR5*+ stem cells.

In summary, there exists a correlation between the increase of *NRG1* and *OLFM4*[284] that could be part of the mechanism of the intestinal mucosa to regenerate damaged epithelium, one that could be leveraged in IBD to boost mucosal healing. Overall, these results suggest that *NRG1* plays a crucial role in regulating *OLFM4* expression and promoting intestinal epithelial regeneration without relying on *LGR5*+ stem cells.

However, numerous signals can induce the upregulation of OLFM4 in myeloid and epithelial cells (Cancer Metastasis Rev (2016) 35:201–212) including wnt/b-catenin, bacterial-notch, NF-KB signals, and IL-22 in epithelial cell lines (doi:10.1016/j.crohns.2011.09.013). Notch signaling is activated in the intestinal epithelial cells of patients with IBD, and in a recent study, NOTCH and TNF signaling was shown to promote OLFM4 cytoplasmic accumulation in intestinal epithelium (Kuno R et al)

OLFM4 has also been found to be secreted into the mucus in active IBD[268] and thus apparently involved in defense.[285] *OLFM4* plays important roles in innate immunity against bacterial infection, gastrointestinal inflammation, and cancer[284]. We argue that the increase in *OLFM4* secretion in inflammatory conditions such as IBD is also a defense mechanism against the loss of mucus and the entrance of higher number of microbial agents. Further confirming this, Kinchen et al showed increased expression of *OLFM4* in organoids after IL-6 stimulation. Thus, *OLFM4* increase is triggered by inflammation and can potentially perform different functions in immune regulation, defense, and regeneration.

In addition to the expression of *NRG1*, we discovered the specific expression of Neuropeptide-Y (*NPY*)[286] in S2b fibroblasts from healthy individuals. *NPY* is a peptide from the NPY family, which comprises three peptides: NPY, peptide YY (*PYY*), and pancreatic polypeptide (*PP*)[287-289]that act as hormones and/or neurotransmitters/neuromodulators[290]. *NPY* is expressed in multiple neuronal systems of the brain, from the medullary brainstem to the cerebral cortex, and in enteric neurons including secretomotor and inhibitory motoneurons, while PYY and PP are localized in enteroendocrine cells in the ileum, colon and rectum[291,292]. The expression of this neuropeptide in S2b peri-cryptal fibroblasts is highly surprising, and interestingly we observed *NPY* to be down-modulated in IBD patients in scRNA-seq and bulk RNA-seq data. *NPY* exerts a wide range of functions in the nervous system preferentially through NPY1R (Y1) and NPY2R (Y2) receptors[293], which are members of the G protein-coupled receptor superfamily. It plays a significant role in modulating various physiological processes such as anxiety, appetite, blood pressure, and nociception[286]. In the gut, *NPY* is predominantly produced in sympathetic nerves and is released in conjunction with norepinephrine. This co-release results in colonic relaxation.

Y1 and Y2 receptor antagonism as well as PYY and NPY knockout modify colonic ion transport and motility[294,295]. The pertinent findings indicate that, both in the human and mouse colonic mucosa, PYY and NPY exert a tonic antisecretory effect mediated by epithelial Y1 and neural Y2 receptors[296]. Colonic transit is tonically accelerated by Y1 receptor stimulation and tonically inhibited by Y2 receptor stimulation[294]. NPY inhibits gastrointestinal motility as well as electrolyte and water secretion, and its inhibitory action on intestinal Cl^- secretion is seen in all regions of the intestine[297].

NPY knockout mice or antagonists of Y1 receptor showed resistance to the induction of colitis. However, evidence in humans suggests the downmodulation of NPY at systemic level in IBD, enlightening some discrepancies between mouse and human. Interestingly, our data shows that localized NPY in colonic S2b fibroblasts is dramatically downregulated in active UC and CD patients. Interestingly, NPY1R is present in the intestinal mucosa and primarily expressed by epithelial cells, specifically BEST4+, enteroendocrine cells and contractile pericytes[298] (data not shown). We propose stromal NPY to play a role in the maintenance of the electrolyte function and pH of the epithelium. Thus, downregulation of the expression of NPY in peri-cryptal S2 fibroblasts might affect the colonic ion transport in IBD patients affecting motility and water secretion. It is

important to note that the functions attributed to NPY in the intestine are described for neurons and neuronal function and the exact role that NPY may have in S2b fibroblasts is not known for certain and further investigation will be needed.

Overall, despite we did not observe a change in the proportions of S2 fibroblasts (in contrast to what Kinchen *et al* described) we did observe a significant deregulation in S2b fibroblasts, including a substantial increase in NRG1 expression and complete loss of NPY expression. Therefore, these observations suggest dysregulation in the fibroblasts surrounding the crypts, which play an important role in the epithelium, in both forms of IBD.

In addition to changes in S2b fibroblasts, we also observed significant alterations in other types of fibroblasts in patients with IBD. Specifically, S1 fibroblasts showed a notable decrease in abundance in IBD samples and exhibited an inflammatory profile compared to healthy individuals. These cells are present throughout the lamina propria and are abundant in healthy tissue, as observed by scRNA-seq and CosMx SMI. However, in IBD samples, they show an increased expression of inflammatory cytokines, suggesting that they have undergone rewiring. Trajectory analysis revealed that S1 cells from IBD patients were close to inflammatory fibroblasts, indicating that they may be the source of these cells in the inflamed mucosa. S2 fibroblasts also showed potential for polarization towards a more inflammatory phenotype, as some of these cells were found in close proximity to inflammatory fibroblasts in the trajectory analysis. These findings suggest that fibroblasts are not static populations but rather plastic ones that can polarize under specific conditions.

Inflammatory fibroblasts represent a disease-specific fibroblast subset characterized by the expression of multiple cytokines including profibrotic IL11, IL24, IL8, IL6, TGFβ1, and tissue remodeling metallopeptidases, making them attractive therapeutic targets to treat inflammation and potentially, fibrosis, a common and difficult-to-treat complication of chronic intestinal inflammation. Inflammatory fibroblasts have been observed in other[205,232] IBD datasets as well as in various types of cancer, where they are known as cancer-associated fibroblasts (CAFs)[299-302]. Interestingly, the marker genes for inflammatory fibroblasts in IBD datasets exhibit some degree of variability. For instance, inflammatory fibroblasts from Kinchen *et al* revealed that their closest homology was with FRCs. This suggests that there may be heterogeneity within the population of

inflammatory fibroblasts. However, probably due to the small number of cells in our cohort, we were unable to observe this heterogeneity.

Emphasizing their interaction with macrophages, inflammatory fibroblasts express *CSF2* (GM-CSF), which promotes macrophage activation, while activated macrophages can produce mediators (i.e., OSM, IL6, TNF, etc) that can drive fibroblast activation. Furthermore, besides its role on epithelial regeneration, EGFR (NRG1 receptor) signaling is a robust regulator of fibroblast motility[303] and may be involved in cartilage and bone destruction in rheumatoid arthritis[304]. In accordance with our results, Martin and colleagues found a strong correlation in a subset of CD inflamed ileum patients between the presence of activated fibroblasts and other inflammatory cells such as macrophages, dendritic cells, T cells, plasma cells, and endothelial cells. The researchers also observed that the inflammatory macrophages were always in close proximity to PDPN-positive fibroblasts (inflammatory phenotype fibroblasts). Based on both scRNA-seq and SMI, we hypothesize that the interaction between M2 or IDA macrophages and inflammatory fibroblasts could play a role in disease pathophysiology. While there is abundant data in the literature to support the interaction of macrophages and fibroblasts, particularly in the context of cancer and fibrosis[304], little is known about their crosstalk in the context of chronic inflammation. We argue that inflammatory fibroblasts could potentially polarize M2 macrophages to IDA phenotype due to its proximity in the tissue and the receptor-ligands that they express.

Despite the important information that can be drawn from our datasets, there are a few limitations to our study that must be considered. First, the number of total individuals/samples analyzed, especially given the high heterogeneity observed, is too low for us to explore the relationship between the identified signatures and disease behavior. Nonetheless, preliminary analysis of additional datasets, including 74 total patients generated in our group, confirms the presence of heterogeneous resident and inflammatory macrophages, as well as neutrophil subsets across patients. In addition, the 1000-gene SMI panel used in our study, while sufficiently large to cover most cell types, lacked important markers that may have limited our accuracy when assigning cell identities. This may be especially true for cell subsets sharing most of their transcriptomic signature (i.e., N1, N2 and N3 neutrophils).

Overall, we provide evidence to support high patient-dependent heterogeneity within the myeloid compartment in both UC and colonic CD. It is important to note that single-cell

analysis not only allows us to discover new cell populations but also sheds light on the diverse cellular states that can exist within the same cell type. We argue that intestinal macrophages, which sense changes in the microenvironment, could act as reliable indicators of patient-specific molecular patterns and thus, promising targets. Furthermore, we show that by combining scRNA-seq with SMI, cell subsets can be assigned to likely interacting partners, thus providing crucial niche information. This spatial resolution will be essential in understanding cellular function, and to faithfully link biologically relevant interactions to specific cell types.

CONCLUSIONS

- Single-cell RNA sequencing (scRNA-seq) technology allowed us to decipher the cellular composition of the intestinal mucosa with unprecedented resolution and to unravel the heterogeneity of cell populations within the colon in health and IBD.
- The combined use of CosMx SMI and scRNAseq allowed to find and localize new cellular populations and understand their connections.

Epithelium compartment

- The epithelium undergoes the most significant changes in cell proportions between healthy and IBD tissue. Epithelial cell loss occurs in patients with IBD especially affecting differentiated cell types. The loss is more pronounced in UC compared to CD patients.
- The remaining epithelium in IBD patients is shifted towards a pro-inflammatory gene signature. The most abundant epithelial cell type in UC, a novel cluster of colonocytes only present in IBD, together with other undifferentiated cell types are responsible for the increased expression of inflammatory genes.
- OLFM4 is markedly upregulated in IBD patients and highly expressed by several progenitor-like cell types along the crypt. OLFM4 expression is not exclusive of the canonical stem cell subset.

Myeloid compartment

- For the first time by scRNAseq, we were able to identify neutrophils in the inflamed colon. Our analysis revealed 3 different states suggesting potential functional specializations and adding evidence for the heterogeneity of neutrophils especially in tissues.
- We identified two colonic resident macrophages, M2 and the novel M0 which lacked expression of M2 canonical markers but presented a signature with antigen presentation and phagocytic capacities. According to trajectory analysis, M2 macrophages could generate not only from monocytes but also from resident M0 macrophages.
- Inflammatory macrophages appeared only in the IBD colon displaying great heterogeneity. Remarkably, we described a novel Inflammation-Dependent Alternative (IDA) macrophages that do not express M1 canonical markers but expressEGF-like ligands and specifically *NRG1*. *NRG1* is highly upregulated in IBD and plays a role on the regeneration of the epithelium.
- Spatial CosMx analysis revealed that IDA macrophages with low or no expression of NRG1 are the main components of CD granulomas.

Stromal compartment

- Fibroblasts, especially S2b which suffer a transcriptional shift in patients, are also contributors of the increase of NRG1 in IBD.
- S1 fibroblasts were decreased in IBD and shifted towards a more pro-inflammatory phenotype. An additional population of inflammatory fibroblasts were only present in the mucosa of IBD patients. Trajectory analysis suggested that S1 fibroblasts could shift and polarize into inflammatory fibroblasts.
- CosMx analysis revealed the high co-localization of IDA macrophages and inflammatory fibroblasts, suggesting a cross-talk between these two cell types in IBD.

REFERENCES

1. Sartor, R.B. Mechanisms of disease: pathogenesis of Crohn's disease and ulcerative colitis. *Nature clinical practice. Gastroenterology & hepatology* **3**, 390-407 (2006).
2. Weimers, P. & Munkholm, P. The Natural History of IBD: Lessons Learned. *Curr Treat Options Gastroenterol* **16**, 101-111 (2018).
3. Burisch, J. & Munkholm, P. The epidemiology of inflammatory bowel disease. *Scand J Gastroenterol* **50**, 942-951 (2015).
4. Mak, W.Y., Zhao, M., Ng, S.C. & Burisch, J. The epidemiology of inflammatory bowel disease: East meets west. *J Gastroenterol Hepatol* **35**, 380-389 (2020).
5. Collaborators, G.B.D.I.B.D. The global, regional, and national burden of inflammatory bowel disease in 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Gastroenterol Hepatol* **5**, 17-30 (2020).
6. Cosnes, J., Gower-Rousseau, C., Seksik, P. & Cortot, A. Epidemiology and natural history of inflammatory bowel diseases. *Gastroenterology* **140**, 1785-1794 (2011).
7. Gupta, N., *et al.* Gender differences in presentation and course of disease in pediatric patients with Crohn disease. *Pediatrics* **120**, e1418-1425 (2007).
8. Danese, S. & Fiocchi, C. Ulcerative colitis. *N Engl J Med* **365**, 1713-1725 (2011).
9. Low, D., Nguyen, D.D. & Mizoguchi, E. Animal models of ulcerative colitis and their application in drug research. *Drug Des Devel Ther* **7**, 1341-1357 (2013).
10. Ordas, I., Eckmann, L., Talamini, M., Baumgart, D.C. & Sandborn, W.J. Ulcerative colitis. *Lancet* **380**, 1606-1619 (2012).
11. Nikolaus, S. & Schreiber, S. Diagnostics of inflammatory bowel disease. *Gastroenterology* **133**, 1670-1689 (2007).
12. Fakhoury, M., Negrulj, R., Mooranian, A. & Al-Salami, H. Inflammatory bowel disease: clinical aspects and treatments. *J Inflamm Res* **7**, 113-120 (2014).
13. Baumgart, D.C. & Sandborn, W.J. Crohn's disease. *Lancet* **380**, 1590-1605 (2012).
14. Levine, J.S. & Burakoff, R. Extraintestinal manifestations of inflammatory bowel disease. *Gastroenterology & hepatology* **7**, 235-241 (2011).
15. Etchevers, M.J., *et al.* Risk factors and characteristics of extent progression in ulcerative colitis. *Inflamm Bowel Dis* **15**, 1320-1325 (2009).
16. Hoie, O., *et al.* Low colectomy rates in ulcerative colitis in an unselected European cohort followed for 10 years. *Gastroenterology* **132**, 507-515 (2007).
17. Ekbom, A., Helmick, C., Zack, M. & Adami, H.O. Ulcerative colitis and colorectal cancer. A population-based study. *N Engl J Med* **323**, 1228-1233 (1990).
18. Satsangi, J., Silverberg, M.S., Vermeire, S. & Colombel, J.F. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. *Gut* **55**, 749-753 (2006).
19. Freeman, H.J. Application of the Montreal classification for Crohn's disease to a single clinician database of 1015 patients. *Can J Gastroenterol* **21**, 363-366 (2007).
20. Van Assche, G., *et al.* The second European evidence-based Consensus on the diagnosis and management of Crohn's disease: Definitions and diagnosis. *Journal of Crohn's & colitis* **4**, 7-27 (2010).
21. Mary, J.Y. & Modigliani, R. Development and validation of an endoscopic index of the severity for Crohn's disease: a prospective multicentre study. Groupe d'Etudes Therapeutiques des Affections Inflammatoires du Tube Digestif (GETAID). *Gut* **30**, 983-989 (1989).
22. Daperno, M., *et al.* Development and validation of a new, simplified endoscopic activity score for Crohn's disease: the SES-CD. *Gastrointest Endosc* **60**, 505-512 (2004).
23. Dave, M. & Loftus, E.V., Jr. Mucosal healing in inflammatory bowel disease-a true paradigm of success? *Gastroenterology & hepatology* **8**, 29-38 (2012).
24. Duerr, R.H. Genome-wide association studies herald a new era of rapid discoveries in inflammatory bowel disease research. *Gastroenterology* **132**, 2045-2049 (2007).
25. Hugot, J.P., *et al.* Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* **411**, 599-603 (2001).
26. Zhang, Y.Z. & Li, Y.Y. Inflammatory bowel disease: pathogenesis. *World journal of gastroenterology : WJG* **20**, 91-99 (2014).

27. Xavier, R.J. & Podolsky, D.K. Unravelling the pathogenesis of inflammatory bowel disease. *Nature* **448**, 427-434 (2007).
28. Okumura, R. & Takeda, K. Maintenance of gut homeostasis by the mucosal immune system. *Proc Jpn Acad Ser B Phys Biol Sci* **92**, 423-435 (2016).
29. Rooks, M.G. & Garrett, W.S. Gut microbiota, metabolites and host immunity. *Nature reviews. Immunology* **16**, 341-352 (2016).
30. Hisamatsu, T., *et al.* Immune aspects of the pathogenesis of inflammatory bowel disease. *Pharmacol Ther* **137**, 283-297 (2013).
31. Silva, F.A., Rodrigues, B.L., Ayrizono, M.L. & Leal, R.F. The Immunological Basis of Inflammatory Bowel Disease. *Gastroenterology research and practice* **2016**, 2097274 (2016).
32. Frank, D.N., *et al.* Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci U S A* **104**, 13780-13785 (2007).
33. Martinez-Medina, M., Aldeguer, X., Gonzalez-Huix, F., Acero, D. & Garcia-Gil, L.J. Abnormal microbiota composition in the ileocolonic mucosa of Crohn's disease patients as revealed by polymerase chain reaction-denaturing gradient gel electrophoresis. *Inflamm Bowel Dis* **12**, 1136-1145 (2006).
34. Pascal, V., *et al.* A microbial signature for Crohn's disease. *Gut* **66**, 813-822 (2017).
35. Sokol, H., *et al.* Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci U S A* **105**, 16731-16736 (2008).
36. Cohen, M.L. Changing patterns of infectious disease. *Nature* **406**, 762-767 (2000).
37. Ouyang, Q., *et al.* The emergence of inflammatory bowel disease in the Asian Pacific region. *Curr Opin Gastroenterol* **21**, 408-413 (2005).
38. Bach, J.F. The effect of infections on susceptibility to autoimmune and allergic diseases. *N Engl J Med* **347**, 911-920 (2002).
39. Thomas, G.A., Rhodes, J. & Green, J.T. Inflammatory bowel disease and smoking--a review. *Am J Gastroenterol* **93**, 144-149 (1998).
40. Cornish, J.A., *et al.* The risk of oral contraceptives in the etiology of inflammatory bowel disease: a meta-analysis. *Am J Gastroenterol* **103**, 2394-2400 (2008).
41. Kaufmann, H.J. & Taubin, H.L. Nonsteroidal anti-inflammatory drugs activate quiescent inflammatory bowel disease. *Ann.Intern.Med.* **107**, 513-516 (1987).
42. Bitton, A., *et al.* Predicting relapse in Crohn's disease: a biopsychosocial model. *Gut* **57**, 1386-1392 (2008).
43. Danese, S. & Fiocchi, C. Etiopathogenesis of inflammatory bowel diseases. *World journal of gastroenterology : WJG* **12**, 4807-4812 (2006).
44. Travis, S.P., *et al.* European evidence based consensus on the diagnosis and management of Crohn's disease: current management. *Gut* **55 Suppl 1**, i16-35 (2006).
45. Kuenzig, M.E., *et al.* Budesonide for the Induction and Maintenance of Remission in Crohn's Disease: Systematic Review and Meta-Analysis for the Cochrane Collaboration. *J Can Assoc Gastroenterol* **1**, 159-173 (2018).
46. Faubion, W.A., Jr., Loftus, E.V., Jr., Harmsen, W.S., Zinsmeister, A.R. & Sandborn, W.J. The natural history of corticosteroid therapy for inflammatory bowel disease: a population-based study. *Gastroenterology* **121**, 255-260 (2001).
47. Dignass, A., *et al.* The second European evidence-based Consensus on the diagnosis and management of Crohn's disease: Current management. *Journal of Crohn's & colitis* **4**, 28-62 (2010).
48. Lichtenstein, G.R., Hanauer, S.B., Sandborn, W.J. & Practice Parameters Committee of American College of, G. Management of Crohn's disease in adults. *Am J Gastroenterol* **104**, 465-483; quiz 464, 484 (2009).
49. Warner, B., *et al.* A practical guide to thiopurine prescribing and monitoring in IBD. *Frontline Gastroenterol* **9**, 10-15 (2018).

50. Prefontaine, E., Sutherland, L.R., Macdonald, J.K. & Cepoiu, M. Azathioprine or 6-mercaptopurine for maintenance of remission in Crohn's disease. *Cochrane Database Syst Rev*, CD000067 (2009).
51. Gisbert, J.P., Linares, P.M., McNicholl, A.G., Mate, J. & Gomollon, F. Meta-analysis: the efficacy of azathioprine and mercaptopurine in ulcerative colitis. *Aliment Pharmacol Ther* **30**, 126-137 (2009).
52. Peyrin-Biroulet, L. & Lemann, M. Review article: remission rates achievable by current therapies for inflammatory bowel disease. *Aliment Pharmacol Ther* **33**, 870-879 (2011).
53. Ko, Y., Paramsothy, S., Yau, Y. & Leong, R.W. Superior treatment persistence with ustekinumab in Crohn's disease and vedolizumab in ulcerative colitis compared with anti-TNF biological agents: real-world registry data from the Persistence Australian National IBD Cohort (PANIC) study. *Aliment Pharmacol Ther* **54**, 292-301 (2021).
54. Salas, A., *et al.* JAK-STAT pathway targeting for the treatment of inflammatory bowel disease. *Nat Rev Gastroenterol Hepatol* **17**, 323-337 (2020).
55. Garrido-Trigo, A. & Salas, A. Molecular structure and function of Janus kinases: implications for the development of inhibitors. *Journal of Crohn's & colitis* (2019).
56. Panaccione, R., *et al.* Characterization of Creatine Kinase Levels in Tofacitinib-Treated Patients with Ulcerative Colitis: Results from Clinical Trials. *Dig Dis Sci* **66**, 2732-2743 (2021).
57. Rocha, C.M., *et al.* Current jakinibs for the treatment of rheumatoid arthritis: a systematic review. *Inflammopharmacology* **29**, 595-615 (2021).
58. Aguilar, D., *et al.* Randomized Controlled Trial Substudy of Cell-specific Mechanisms of Janus Kinase 1 Inhibition With Upadacitinib in the Crohn's Disease Intestinal Mucosa: Analysis From the CELEST Study. *Inflamm Bowel Dis* **27**, 1999-2009 (2021).
59. Corraliza, A.M., *et al.* Differences in Peripheral and Tissue Immune Cell Populations Following Haematopoietic Stem Cell Transplantation in Crohn's Disease Patients. *Journal of Crohn's and Colitis*, jjy203-jjy203 (2018).
60. Carvello, M., Lightner, A., Yamamoto, T., Kotze, P.G. & Spinelli, A. Mesenchymal Stem Cells for Perianal Crohn's Disease. *Cells* **8**(2019).
61. Snowden, J.A., *et al.* Autologous stem cell transplantation in refractory Crohn's disease - low intensity therapy evaluation (ASTIClite): study protocols for a multicentre, randomised controlled trial and observational follow up study. *BMC gastroenterology* **19**, 82 (2019).
62. Billioud, V., Sandborn, W.J. & Peyrin-Biroulet, L. Loss of response and need for adalimumab dose intensification in Crohn's disease: a systematic review. *Am J Gastroenterol* **106**, 674-684 (2011).
63. Lamb, C.A., *et al.* British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. *Gut* **68**, s1-s106 (2019).
64. Hwang, J.M. & Varma, M.G. Surgery for inflammatory bowel disease. *World journal of gastroenterology : WJG* **14**, 2678-2690 (2008).
65. Barker, N. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nature reviews. Molecular cell biology* **15**, 19-33 (2014).
66. van der Flier, L.G. & Clevers, H. Stem cells, self-renewal, and differentiation in the intestinal epithelium. *Annu Rev Physiol* **71**, 241-260 (2009).
67. Olivares-Villagomez, D. & Van Kaer, L. Intestinal Intraepithelial Lymphocytes: Sentinels of the Mucosal Barrier. *Trends in immunology* **39**, 264-275 (2018).
68. Kim, Y.S. & Ho, S.B. Intestinal goblet cells and mucins in health and disease: recent insights and progress. *Current gastroenterology reports* **12**, 319-330 (2010).
69. Maynard, C.L., Elson, C.O., Hatton, R.D. & Weaver, C.T. Reciprocal interactions of the intestinal microbiota and immune system. *Nature* **489**, 231-241 (2012).
70. Johansson, M.E., Larsson, J.M. & Hansson, G.C. The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. *Proc Natl Acad Sci U S A* **108 Suppl 1**, 4659-4665 (2011).
71. Okumura, R. & Takeda, K. Roles of intestinal epithelial cells in the maintenance of gut homeostasis. *Exp Mol Med* **49**, e338 (2017).

72. Turner, J.R. Intestinal mucosal barrier function in health and disease. *Nature reviews. Immunology* **9**, 799-809 (2009).
73. Brogden, K.A. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat Rev Microbiol* **3**, 238-250 (2005).
74. Ramasundara, M., Leach, S.T., Lemberg, D.A. & Day, A.S. Defensins and inflammation: the role of defensins in inflammatory bowel disease. *J Gastroenterol Hepatol* **24**, 202-208 (2009).
75. Ali, A., Tan, H. & Kaiko, G.E. Role of the Intestinal Epithelium and Its Interaction With the Microbiota in Food Allergy. *Frontiers in immunology* **11**, 604054 (2020).
76. Umar, S. Intestinal stem cells. *Current gastroenterology reports* **12**, 340-348 (2010).
77. Hsu, Y.C., Li, L. & Fuchs, E. Transit-amplifying cells orchestrate stem cell activity and tissue regeneration. *Cell* **157**, 935-949 (2014).
78. Gerbe, F., Legraverend, C. & Jay, P. The intestinal epithelium tuft cells: specification and function. *Cellular and molecular life sciences : CMLS* **69**, 2907-2917 (2012).
79. Hendel, S.K., *et al.* Tuft Cells and Their Role in Intestinal Diseases. *Frontiers in immunology* **13**, 822867 (2022).
80. Schneider, C., O'Leary, C.E. & Locksley, R.M. Regulation of immune responses by tuft cells. *Nature reviews. Immunology* **19**, 584-593 (2019).
81. Gribble, F.M. & Reimann, F. Function and mechanisms of enteroendocrine cells and gut hormones in metabolism. *Nat Rev Endocrinol* **15**, 226-237 (2019).
82. Worthington, J.J., Reimann, F. & Gribble, F.M. Enteroendocrine cells-sensory sentinels of the intestinal environment and orchestrators of mucosal immunity. *Mucosal immunology* **11**, 3-20 (2018).
83. Dillon, A. & Lo, D.D. M Cells: Intelligent Engineering of Mucosal Immune Surveillance. *Frontiers in immunology* **10**, 1499 (2019).
84. Van der Sluis, M., *et al.* Muc2-deficient mice spontaneously develop colitis, indicating that MUC2 is critical for colonic protection. *Gastroenterology* **131**, 117-129 (2006).
85. Boltin, D., Perets, T.T., Vilkin, A. & Niv, Y. Mucin function in inflammatory bowel disease: an update. *J Clin Gastroenterol* **47**, 106-111 (2013).
86. Pullan, R.D., *et al.* Thickness of adherent mucus gel on colonic mucosa in humans and its relevance to colitis. *Gut* **35**, 353-359 (1994).
87. Larsson, J.M., *et al.* Altered O-glycosylation profile of MUC2 mucin occurs in active ulcerative colitis and is associated with increased inflammation. *Inflamm Bowel Dis* **17**, 2299-2307 (2011).
88. Gersemann, M., *et al.* Differences in goblet cell differentiation between Crohn's disease and ulcerative colitis. *Differentiation* **77**, 84-94 (2009).
89. Heller, F., *et al.* Interleukin-13 is the key effector Th2 cytokine in ulcerative colitis that affects epithelial tight junctions, apoptosis, and cell restitution. *Gastroenterology* **129**, 550-564 (2005).
90. Khor, B., Gardet, A. & Xavier, R.J. Genetics and pathogenesis of inflammatory bowel disease. *Nature* **474**, 307-317 (2011).
91. Dickson, I. Crohn's disease: Impaired bacterial clearance in IBD. *Nat Rev Gastroenterol Hepatol* **13**, 251 (2016).
92. Shao, B.Z., *et al.* The Role of Autophagy in Inflammatory Bowel Disease. *Front Physiol* **12**, 621132 (2021).
93. Hu, X., *et al.* ATF4 Deficiency Promotes Intestinal Inflammation in Mice by Reducing Uptake of Glutamine and Expression of Antimicrobial Peptides. *Gastroenterology* **156**, 1098-1111 (2019).
94. Tysk, C., Lindberg, E., Jarnerot, G. & Floderus-Myrhed, B. Ulcerative colitis and Crohn's disease in an unselected population of monozygotic and dizygotic twins. A study of heritability and the influence of smoking. *Gut* **29**, 990-996 (1988).
95. Hampe, J., *et al.* A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nature genetics* **39**, 207-211 (2007).

96. Neurath, M.F. Cytokines in inflammatory bowel disease. *Nature reviews. Immunology* **14**, 329-342 (2014).
97. Blander, J.M. & Sander, L.E. Beyond pattern recognition: five immune checkpoints for scaling the microbial threat. *Nature reviews. Immunology* **12**, 215-225 (2012).
98. Akira, S. & Takeda, K. Toll-like receptor signalling. *Nature reviews. Immunology* **4**, 499-511 (2004).
99. Kawai, T. & Akira, S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nature immunology* **11**, 373-384 (2010).
100. Takeuchi, O. & Akira, S. Pattern recognition receptors and inflammation. *Cell* **140**, 805-820 (2010).
101. Lubbers, R., van Essen, M.F., van Kooten, C. & Trouw, L.A. Production of complement components by cells of the immune system. *Clin Exp Immunol* **188**, 183-194 (2017).
102. Roche, P.A. & Furuta, K. The ins and outs of MHC class II-mediated antigen processing and presentation. *Nature reviews. Immunology* **15**, 203-216 (2015).
103. Steinman, R.M. Lasker Basic Medical Research Award. Dendritic cells: versatile controllers of the immune system. *Nat Med* **13**, 1155-1159 (2007).
104. Steinman, R.M. & Banchereau, J. Taking dendritic cells into medicine. *Nature* **449**, 419-426 (2007).
105. Aderem, A. & Underhill, D.M. Mechanisms of phagocytosis in macrophages. *Annu Rev Immunol* **17**, 593-623 (1999).
106. Kumar, B.V., Connors, T.J. & Farber, D.L. Human T Cell Development, Localization, and Function throughout Life. *Immunity* **48**, 202-213 (2018).
107. Murphy, K.M. & Reiner, S.L. The lineage decisions of helper T cells. *Nature reviews. Immunology* **2**, 933-944 (2002).
108. Mosmann, T.R. & Coffman, R.L. TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. *Annu Rev Immunol* **7**, 145-173 (1989).
109. Paul, W.E. Interleukin-4: a prototypic immunoregulatory lymphokine. *Blood* **77**, 1859-1870 (1991).
110. Korn, T., Bettelli, E., Oukka, M. & Kuchroo, V.K. IL-17 and Th17 Cells. *Annu Rev Immunol* **27**, 485-517 (2009).
111. Weaver, C.T., Hatton, R.D., Mangan, P.R. & Harrington, L.E. IL-17 family cytokines and the expanding diversity of effector T cell lineages. *Annu Rev Immunol* **25**, 821-852 (2007).
112. Weigelin, B. & Friedl, P. T cell-mediated additive cytotoxicity - death by multiple bullets. *Trends Cancer* **8**, 980-987 (2022).
113. Spencer, J. & Sollid, L.M. The human intestinal B-cell response. *Mucosal immunology* **9**, 1113-1124 (2016).
114. Mora, J.R. & von Andrian, U.H. Differentiation and homing of IgA-secreting cells. *Mucosal immunology* **1**, 96-109 (2008).
115. Gao, X., *et al.* T follicular helper 17 (Tfh17) cells are superior for immunological memory maintenance. *Elife* **12**(2023).
116. Nguyen, Q. & Borrow, P. Human T follicular regulatory cells with helper and regulatory lineage origins. *Nature reviews. Immunology* (2023).
117. Zhang, Y., *et al.* Plasma cell output from germinal centers is regulated by signals from Tfh and stromal cells. *J Exp Med* **215**, 1227-1243 (2018).
118. Silvestre-Roig, C., Hidalgo, A. & Soehnlein, O. Neutrophil heterogeneity: implications for homeostasis and pathogenesis. *Blood* **127**, 2173-2181 (2016).
119. Kolaczkowska, E. & Kubes, P. Neutrophil recruitment and function in health and inflammation. *Nature reviews. Immunology* **13**, 159-175 (2013).
120. Ballesteros, I., *et al.* Co-option of Neutrophil Fates by Tissue Environments. *Cell* **183**, 1282-1297 e1218 (2020).
121. Brannigan, A.E., *et al.* Neutrophil apoptosis is delayed in patients with inflammatory bowel disease. *Shock* **13**, 361-366 (2000).
122. Brinkmann, V., *et al.* Neutrophil extracellular traps kill bacteria. *Science* **303**, 1532-1535 (2004).

123. Wechsler, M.E., *et al.* Eosinophils in Health and Disease: A State-of-the-Art Review. *Mayo Clin Proc* **96**, 2694-2707 (2021).
124. Bandeira-Melo, C. & Weller, P.F. Mechanisms of eosinophil cytokine release. *Mem Inst Oswaldo Cruz* **100 Suppl 1**, 73-81 (2005).
125. Davoine, F. & Lacy, P. Eosinophil cytokines, chemokines, and growth factors: emerging roles in immunity. *Frontiers in immunology* **5**, 570 (2014).
126. Mould, A.W., Matthaei, K.I., Young, I.G. & Foster, P.S. Relationship between interleukin-5 and eotaxin in regulating blood and tissue eosinophilia in mice. *J Clin Invest* **99**, 1064-1071 (1997).
127. Geissmann, F., *et al.* Development of monocytes, macrophages, and dendritic cells. *Science* **327**, 656-661 (2010).
128. Jakubzick, C., *et al.* Minimal differentiation of classical monocytes as they survey steady-state tissues and transport antigen to lymph nodes. *Immunity* **39**, 599-610 (2013).
129. Rivollier, A., He, J., Kole, A., Valatas, V. & Kelsall, B.L. Inflammation switches the differentiation program of Ly6Chi monocytes from antiinflammatory macrophages to inflammatory dendritic cells in the colon. *J Exp Med* **209**, 139-155 (2012).
130. Ziegler-Heitbrock, L. Monocyte subsets in man and other species. *Cell Immunol* **289**, 135-139 (2014).
131. Ma, W.T., Gao, F., Gu, K. & Chen, D.K. The Role of Monocytes and Macrophages in Autoimmune Diseases: A Comprehensive Review. *Frontiers in immunology* **10**, 1140 (2019).
132. Mitchell, A.J., Roediger, B. & Weninger, W. Monocyte homeostasis and the plasticity of inflammatory monocytes. *Cell Immunol* **291**, 22-31 (2014).
133. Ginhoux, F. & Jung, S. Monocytes and macrophages: developmental pathways and tissue homeostasis. *Nature reviews. Immunology* **14**, 392-404 (2014).
134. Zhao, Y., Zou, W., Du, J. & Zhao, Y. The origins and homeostasis of monocytes and tissue-resident macrophages in physiological situation. *J Cell Physiol* **233**, 6425-6439 (2018).
135. Groeneweg, L., Hidalgo, A. & N, A.G. Emerging roles of infiltrating granulocytes and monocytes in homeostasis. *Cellular and molecular life sciences : CMLS* **77**, 3823-3830 (2020).
136. Robinson, A., Han, C.Z., Glass, C.K. & Pollard, J.W. Monocyte Regulation in Homeostasis and Malignancy. *Trends in immunology* **42**, 104-119 (2021).
137. Bain, C.C. & Schridde, A. Origin, Differentiation, and Function of Intestinal Macrophages. *Frontiers in immunology* **9**, 2733 (2018).
138. Bernardo, D., *et al.* Human intestinal pro-inflammatory CD11c(high)CCR2(+)CX3CR1(+) macrophages, but not their tolerogenic CD11c(-)CCR2(-)CX3CR1(-) counterparts, are expanded in inflammatory bowel disease. *Mucosal immunology* **11**, 1114-1126 (2018).
139. Bain, C.C., *et al.* Resident and pro-inflammatory macrophages in the colon represent alternative context-dependent fates of the same Ly6Chi monocyte precursors. *Mucosal immunology* **6**, 498-510 (2013).
140. De Schepper, S., *et al.* Self-Maintaining Gut Macrophages Are Essential for Intestinal Homeostasis. *Cell* **176**, 676 (2019).
141. Bain, C.C. & Mowat, A.M. Macrophages in intestinal homeostasis and inflammation. *Immunol Rev* **260**, 102-117 (2014).
142. Joeris, T., Muller-Luda, K., Agace, W.W. & Mowat, A.M. Diversity and functions of intestinal mononuclear phagocytes. *Mucosal immunology* **10**, 845-864 (2017).
143. Viola, M.F. & Boeckxstaens, G. Intestinal resident macrophages: Multitaskers of the gut. *Neurogastroenterology and motility : the official journal of the European Gastrointestinal Motility Society* **32**, e13843 (2020).
144. Shaw, T.N., *et al.* Tissue-resident macrophages in the intestine are long lived and defined by Tim-4 and CD4 expression. *J Exp Med* **215**, 1507-1518 (2018).

145. Isidro, R.A. & Appleyard, C.B. Colonic macrophage polarization in homeostasis, inflammation, and cancer. *American journal of physiology. Gastrointestinal and liver physiology* **311**, G59-73 (2016).
146. Italiani, P. & Boraschi, D. From Monocytes to M1/M2 Macrophages: Phenotypical vs. Functional Differentiation. *Frontiers in immunology* **5**, 514 (2014).
147. Gonzalez-Dominguez, E., *et al.* CD163L1 and CLEC5A discriminate subsets of human resident and inflammatory macrophages in vivo. *J Leukoc Biol* **98**, 453-466 (2015).
148. Gu, H., *et al.* NLRP3 activation in tumor-associated macrophages enhances lung metastasis of pancreatic ductal adenocarcinoma. *Transl Lung Cancer Res* **11**, 858-868 (2022).
149. Yao, Y., Xu, X.H. & Jin, L. Macrophage Polarization in Physiological and Pathological Pregnancy. *Frontiers in immunology* **10**, 792 (2019).
150. Li, P., *et al.* Proteomic characterization of four subtypes of M2 macrophages derived from human THP-1 cells. *J Zhejiang Univ Sci B* **23**, 407-422 (2022).
151. Hamilton, J.A. & Achuthan, A. Colony stimulating factors and myeloid cell biology in health and disease. *Trends in immunology* **34**, 81-89 (2013).
152. Sehgal, A., *et al.* The role of CSF1R-dependent macrophages in control of the intestinal stem-cell niche. *Nat Commun* **9**, 1272 (2018).
153. Jaguin, M., Houlbert, N., Fardel, O. & Lecureur, V. Polarization profiles of human M-CSF-generated macrophages and comparison of M1-markers in classically activated macrophages from GM-CSF and M-CSF origin. *Cell Immunol* **281**, 51-61 (2013).
154. Draijer, C., Penke, L.R.K. & Peters-Golden, M. Distinctive Effects of GM-CSF and M-CSF on Proliferation and Polarization of Two Major Pulmonary Macrophage Populations. *J Immunol* **202**, 2700-2709 (2019).
155. Kang, B., *et al.* Commensal microbiota drive the functional diversification of colon macrophages. *Mucosal immunology* **13**, 216-229 (2020).
156. Koch, S., Kucharzik, T., Heidemann, J., Nusrat, A. & Luegering, A. Investigating the role of proinflammatory CD16+ monocytes in the pathogenesis of inflammatory bowel disease. *Clin Exp Immunol* **161**, 332-341 (2010).
157. Smythies, L.E., *et al.* Mucosal IL-8 and TGF-beta recruit blood monocytes: evidence for cross-talk between the lamina propria stroma and myeloid cells. *J Leukoc Biol* **80**, 492-499 (2006).
158. Atri, C., Guerfali, F.Z. & Laouini, D. Role of Human Macrophage Polarization in Inflammation during Infectious Diseases. *Int J Mol Sci* **19**(2018).
159. Heresbach, D., *et al.* [Frequency and prognostic value of epithelioid granuloma in inflammatory bowel disease]. *Gastroenterol Clin Biol* **23**, 1376-1387 (1999).
160. Heresbach, D., *et al.* Frequency and significance of granulomas in a cohort of incident cases of Crohn's disease. *Gut* **54**, 215-222 (2005).
161. Molnar, T., Tiszlavicz, L., Gyulai, C., Nagy, F. & Lonovics, J. Clinical significance of granuloma in Crohn's disease. *World journal of gastroenterology : WJG* **11**, 3118-3121 (2005).
162. Pierik, M., *et al.* Epithelioid granulomas, pattern recognition receptors, and phenotypes of Crohn's disease. *Gut* **54**, 223-227 (2005).
163. Bevilacqua, M.P., Stengelin, S., Gimbrone, M.A., Jr. & Seed, B. Endothelial leukocyte adhesion molecule 1: an inducible receptor for neutrophils related to complement regulatory proteins and lectins. *Science* **243**, 1160-1165 (1989).
164. Weiss, S.J. Tissue destruction by neutrophils. *N Engl J Med* **320**, 365-376 (1989).
165. Dinallo, V., *et al.* Neutrophil Extracellular Traps Sustain Inflammatory Signals in Ulcerative Colitis. *Journal of Crohn's & colitis* **13**, 772-784 (2019).
166. Grisham, M.B. & Granger, D.N. Neutrophil-mediated mucosal injury. Role of reactive oxygen metabolites. *Dig Dis Sci* **33**, 6S-15S (1988).
167. Sinagra, E., *et al.* Focal active colitis as a predictor of inflammatory bowel disease: results from a single-center experience. *J Biol Regul Homeost Agents* **31**, 1119-1125 (2017).
168. Wightman, H.R. Active focal colitis. *Hum Pathol* **29**, 887-888 (1998).

169. Greenson, J.K., Stern, R.A., Carpenter, S.L. & Barnett, J.L. The clinical significance of focal active colitis. *Hum Pathol* **28**, 729-733 (1997).
170. Sokol, H., *et al.* Crypt abscess-associated microbiota in inflammatory bowel disease and acute self-limited colitis. *World journal of gastroenterology : WJG* **16**, 583-587 (2010).
171. DeRoche, T.C., Xiao, S.Y. & Liu, X. Histological evaluation in ulcerative colitis. *Gastroenterol Rep (Oxf)* **2**, 178-192 (2014).
172. Lampinen, M., *et al.* Eosinophil granulocytes are activated during the remission phase of ulcerative colitis. *Gut* **54**, 1714-1720 (2005).
173. Bischoff, S.C., *et al.* Quantitative assessment of intestinal eosinophils and mast cells in inflammatory bowel disease. *Histopathology* **28**, 1-13 (1996).
174. Bischoff, S.C. Mast cells in gastrointestinal disorders. *Eur J Pharmacol* **778**, 139-145 (2016).
175. Sellge, G. & Bischoff, S.C. Isolation, culture, and characterization of intestinal mast cells. *Methods Mol Biol* **315**, 123-138 (2006).
176. Krystel-Whittemore, M., Dileepan, K.N. & Wood, J.G. Mast Cell: A Multi-Functional Master Cell. *Frontiers in immunology* **6**, 620 (2015).
177. Correa, I., *et al.* Defective IL-10 production in severe phenotypes of Crohn's disease. *J.Leukoc.Biol.* **85**, 896-903 (2009).
178. Hart, A.L., *et al.* Characteristics of intestinal dendritic cells in inflammatory bowel diseases. *Gastroenterology* **129**, 50-65 (2005).
179. Geremia, A. & Arancibia-Carcamo, C.V. Innate Lymphoid Cells in Intestinal Inflammation. *Frontiers in immunology* **8**, 1296 (2017).
180. Bostick, J.W. & Zhou, L. Innate lymphoid cells in intestinal immunity and inflammation. *Cellular and molecular life sciences : CMLS* **73**, 237-252 (2016).
181. Peters, C.P., Mjosberg, J.M., Bernink, J.H. & Spits, H. Innate lymphoid cells in inflammatory bowel diseases. *Immunol Lett* **172**, 124-131 (2016).
182. Li, J. & Glover, S.C. Innate Lymphoid Cells in Inflammatory Bowel Disease. *Arch Immunol Ther Exp (Warsz)* **66**, 415-421 (2018).
183. McNamee, E.N. & Rivera-Nieves, J. Ectopic Tertiary Lymphoid Tissue in Inflammatory Bowel Disease: Protective or Provocateur? *Frontiers in immunology* **7**, 308 (2016).
184. Donaldson, D.S., Bradford, B.M., Artis, D. & Mabbott, N.A. Reciprocal regulation of lymphoid tissue development in the large intestine by IL-25 and IL-23. *Mucosal immunology* **8**, 582-595 (2015).
185. Castro-Dopico, T. & Clatworthy, M.R. Mucosal IgG in inflammatory bowel disease - a question of (sub)class? *Gut microbes* **12**, 1-9 (2020).
186. Ruthlein, J., Ibe, M., Burghardt, W., Mossner, J. & Auer, I.O. Immunoglobulin G (IgG), IgG1, and IgG2 determinations from endoscopic biopsy specimens in control, Crohn's disease, and ulcerative colitis subjects. *Gut* **33**, 507-512 (1992).
187. Powell, D.W., Pinchuk, I.V., Saada, J.I., Chen, X. & Mifflin, R.C. Mesenchymal cells of the intestinal lamina propria. *Annu Rev Physiol* **73**, 213-237 (2011).
188. Baum, J. & Duffy, H.S. Fibroblasts and myofibroblasts: what are we talking about? *J Cardiovasc Pharmacol* **57**, 376-379 (2011).
189. Tai, Y., *et al.* Myofibroblasts: Function, Formation, and Scope of Molecular Therapies for Skin Fibrosis. *Biomolecules* **11**(2021).
190. Phan, S.H. Biology of fibroblasts and myofibroblasts. *Proc Am Thorac Soc* **5**, 334-337 (2008).
191. Armulik, A., Genove, G. & Betsholtz, C. Pericytes: developmental, physiological, and pathological perspectives, problems, and promises. *Developmental cell* **21**, 193-215 (2011).
192. Seguella, L. & Gulbransen, B.D. Enteric glial biology, intercellular signalling and roles in gastrointestinal disease. *Nat Rev Gastroenterol Hepatol* **18**, 571-587 (2021).
193. Progatzky, F. & Pachnis, V. The role of enteric glia in intestinal immunity. *Current opinion in immunology* **77**, 102183 (2022).

194. Stzepourginski, I., *et al.* CD34+ mesenchymal cells are a major component of the intestinal stem cells niche at homeostasis and after injury. *Proc Natl Acad Sci U S A* **114**, E506-E513 (2017).
195. Acton, S.E., Onder, L., Novkovic, M., Martinez, V.G. & Ludewig, B. Communication, construction, and fluid control: lymphoid organ fibroblastic reticular cell and conduit networks. *Trends in immunology* **42**, 782-794 (2021).
196. Fletcher, A.L., Acton, S.E. & Knoblich, K. Lymph node fibroblastic reticular cells in health and disease. *Nature reviews. Immunology* **15**, 350-361 (2015).
197. Vicente-Suarez, I., *et al.* Unique lamina propria stromal cells imprint the functional phenotype of mucosal dendritic cells. *Mucosal immunology* **8**, 141-151 (2015).
198. Tentaku, A., *et al.* Proximal deposition of collagen IV by fibroblasts contributes to basement membrane formation by colon epithelial cells in vitro. *The FEBS journal* **289**, 7466-7485 (2022).
199. Barnhoorn, M.C., *et al.* Stromal Cells in the Pathogenesis of Inflammatory Bowel Disease. *Journal of Crohn's & colitis* **14**, 995-1009 (2020).
200. Wu, X., *et al.* Cellular and Molecular Mechanisms of Intestinal Fibrosis. *Gut Liver* (2023).
201. Leeb, S.N., *et al.* Reduced migration of fibroblasts in inflammatory bowel disease: role of inflammatory mediators and focal adhesion kinase. *Gastroenterology* **125**, 1341-1354 (2003).
202. Lawrance, I.C., Maxwell, L. & Doe, W. Altered response of intestinal mucosal fibroblasts to profibrogenic cytokines in inflammatory bowel disease. *Inflamm Bowel Dis* **7**, 226-236 (2001).
203. West, N.R., *et al.* Oncostatin M drives intestinal inflammation and predicts response to tumor necrosis factor-neutralizing therapy in patients with inflammatory bowel disease. *Nat Med* **23**, 579-589 (2017).
204. Cromer, W.E., Mathis, J.M., Granger, D.N., Chaitanya, G.V. & Alexander, J.S. Role of the endothelium in inflammatory bowel diseases. *World journal of gastroenterology : WJG* **17**, 578-593 (2011).
205. Smillie, C.S., *et al.* Intra- and Inter-cellular Rewiring of the Human Colon during Ulcerative Colitis. *Cell* **178**, 714-730 e722 (2019).
206. Abraham, C. & Cho, J.H. Inflammatory bowel disease. *N Engl J Med* **361**, 2066-2078 (2009).
207. Lawrance, I.C., Fiocchi, C. & Chakravarti, S. Ulcerative colitis and Crohn's disease: distinctive gene expression profiles and novel susceptibility candidate genes. *Hum.Mol.Genet.* **10**, 445-456 (2001).
208. Planell, N., *et al.* Transcriptional analysis of the intestinal mucosa of patients with ulcerative colitis in remission reveals lasting epithelial cell alterations. *Gut* **62**, 967-976 (2013).
209. Leal, R.F., *et al.* Identification of inflammatory mediators in patients with Crohn's disease unresponsive to anti-TNFalpha therapy. *Gut* **64**, 233-242 (2015).
210. Haberman, Y., *et al.* Pediatric Crohn disease patients exhibit specific ileal transcriptome and microbiome signature. *J Clin Invest* **124**, 3617-3633 (2014).
211. Cardinale, C.J., *et al.* Transcriptome profiling of human ulcerative colitis mucosa reveals altered expression of pathways enriched in genetic susceptibility loci. *PloS one* **9**, e96153 (2014).
212. Visvanathan, S., *et al.* Selective IL-23 inhibition by risankizumab modulates the molecular profile in the colon and ileum of patients with active Crohn's disease: Results from a randomised phase II biopsy substudy. *Journal of Crohn's & colitis* (2018).
213. Rubin, S.J.S., *et al.* Mass cytometry reveals systemic and local immune signatures that distinguish inflammatory bowel diseases. *Nat Commun* **10**, 2686 (2019).
214. Verstockt, B., *et al.* Distinct transcriptional signatures in purified circulating immune cells drive heterogeneity in disease location in IBD. *BMJ Open Gastroenterol* **10**(2023).

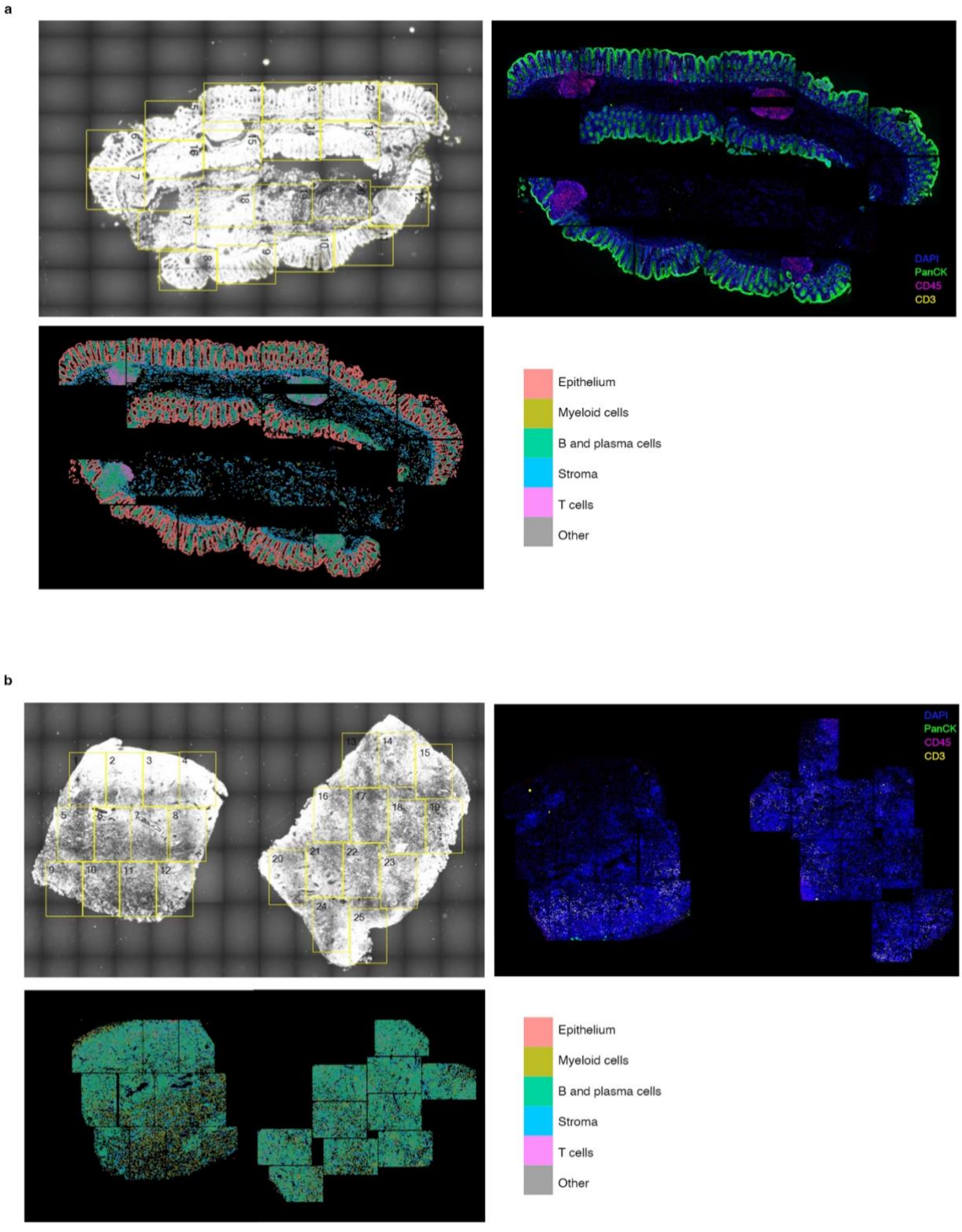
215. Burczynski, M.E., *et al.* Molecular classification of Crohn's disease and ulcerative colitis patients using transcriptional profiles in peripheral blood mononuclear cells. *J Mol Diagn* **8**, 51-61 (2006).
216. Sandberg, R. Entering the era of single-cell transcriptomics in biology and medicine. *Nature methods* **11**, 22-24 (2014).
217. Hedlund, E. & Deng, Q. Single-cell RNA sequencing: Technical advancements and biological applications. *Mol Aspects Med* **59**, 36-46 (2018).
218. Haque, A., Engel, J., Teichmann, S.A. & Lonnberg, T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Med* **9**, 75 (2017).
219. Lee, J., Hyeon, D.Y. & Hwang, D. Single-cell multiomics: technologies and data analysis methods. *Exp Mol Med* **52**, 1428-1442 (2020).
220. Kashima, Y., *et al.* Single-cell sequencing techniques from individual to multiomics analyses. *Exp Mol Med* **52**, 1419-1427 (2020).
221. Haber, A.L., *et al.* A single-cell survey of the small intestinal epithelium. *Nature* **551**, 333-339 (2017).
222. Gaggero, S., *et al.* IL-2 is inactivated by the acidic pH environment of tumors enabling engineering of a pH-selective mutein. *Sci Immunol* **7**, eade5686 (2022).
223. Corridoni, D., *et al.* Single-cell atlas of colonic CD8(+) T cells in ulcerative colitis. *Nat Med* **26**, 1480-1490 (2020).
224. Huang, B., *et al.* Mucosal Profiling of Pediatric-Onset Colitis and IBD Reveals Common Pathogenics and Therapeutic Pathways. *Cell* **179**, 1160-1176 e1124 (2019).
225. Elmentaite, R., *et al.* Cells of the human intestinal tract mapped across space and time. *Nature* **597**, 250-255 (2021).
226. Fawkner-Corbett, D., *et al.* Spatiotemporal analysis of human intestinal development at single-cell resolution. *Cell* **184**, 810-826 e823 (2021).
227. He, S., *et al.* High-plex imaging of RNA and proteins at subcellular resolution in fixed tissue by spatial molecular imaging. *Nat Biotechnol* **40**, 1794-1806 (2022).
228. Veny, M., *et al.* Dissecting Common and Unique Effects of Anti-alpha4beta7 and Anti-Tumor Necrosis Factor Treatment in Ulcerative Colitis. *Journal of Crohn's & colitis* **15**, 441-452 (2021).
229. Germain, P.L., Lun, A., Garcia Meixide, C., Macnair, W. & Robinson, M.D. Doublet identification in single-cell sequencing data using scDblFinder. *F1000Res* **10**, 979 (2021).
230. Hao, Y., *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587 e3529 (2021).
231. Korsunsky, I., *et al.* Fast, sensitive and accurate integration of single-cell data with Harmony. *Nature methods* **16**, 1289-1296 (2019).
232. Kinchen, J., *et al.* Structural Remodeling of the Human Colonic Mesenchyme in Inflammatory Bowel Disease. *Cell* **175**, 372-386 e317 (2018).
233. Shaffer, A.L., *et al.* Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. *Immunity* **17**, 51-62 (2002).
234. D'Acquisto, F., *et al.* Annexin-1 modulates T-cell activation and differentiation. *Blood* **109**, 1095-1102 (2007).
235. van der Leun, A.M., Thommen, D.S. & Schumacher, T.N. CD8(+) T cell states in human cancer: insights from single-cell analysis. *Nature reviews. Cancer* **20**, 218-232 (2020).
236. Jaeger, N., *et al.* Single-cell analyses of Crohn's disease tissues reveal intestinal intraepithelial T cells heterogeneity and altered subset distributions. *Nat Commun* **12**, 1921 (2021).
237. Dann, E., Henderson, N.C., Teichmann, S.A., Morgan, M.D. & Marioni, J.C. Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nat Biotechnol* **40**, 245-253 (2022).
238. Trapnell, C., *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* **32**, 381-386 (2014).

239. Martin, J.C., *et al.* Single-Cell Analysis of Crohn's Disease Lesions Identifies a Pathogenic Cellular Module Associated with Resistance to Anti-TNF Therapy. *Cell* **178**, 1493-1508 e1420 (2019).
240. Mereu, E., *et al.* Benchmarking single-cell RNA-sequencing protocols for cell atlas projects. *Nat Biotechnol* **38**, 747-755 (2020).
241. Nieto, C., *et al.* Serotonin (5-HT) Shapes the Macrophage Gene Profile through the 5-HT(2B)-Dependent Activation of the Aryl Hydrocarbon Receptor. *J Immunol* **204**, 2808-2817 (2020).
242. Dominguez-Soto, A., *et al.* Serotonin drives the acquisition of a profibrotic and anti-inflammatory gene profile through the 5-HT7R-PKA signaling axis. *Sci Rep* **7**, 14761 (2017).
243. Vega, M.A., *et al.* MAFB and MAF Transcription Factors as Macrophage Checkpoints for COVID-19 Severity. *Frontiers in immunology* **11**, 603507 (2020).
244. Cuevas, V.D., *et al.* The Gene Signature of Activated M-CSF-Primed Human Monocyte-Derived Macrophages Is IL-10-Dependent. *J Innate Immun* **14**, 243-256 (2022).
245. McCarthy, D.J., Campbell, K.R., Lun, A.T. & Wills, Q.F. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* **33**, 1179-1186 (2017).
246. Lun, A.T., McCarthy, D.J. & Marioni, J.C. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Res* **5**, 2122 (2016).
247. Haghverdi, L., Lun, A.T.L., Morgan, M.D. & Marioni, J.C. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat Biotechnol* **36**, 421-427 (2018).
248. Aran, D., *et al.* Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nature immunology* **20**, 163-172 (2019).
249. Dotti, I., Mayorgas, A. & Salas, A. Generation of human colon organoids from healthy and inflammatory bowel disease mucosa. *PloS one* **17**, e0276195 (2022).
250. Parikh, K., *et al.* Colonic epithelial cell diversity in health and inflammatory bowel disease. *Nature* **567**, 49-55 (2019).
251. Xie, X., *et al.* Single-cell transcriptomic landscape of human blood cells. *Natl Sci Rev* **8**, nwaa180 (2021).
252. Zhao, Y., *et al.* Single-cell transcriptomic landscape of nucleated cells in umbilical cord blood. *Gigascience* **8**(2019).
253. Mulder, K., *et al.* Cross-tissue single-cell landscape of human monocytes and macrophages in health and disease. *Immunity* **54**, 1883-1900 e1885 (2021).
254. Buonanno, A. & Fischbach, G.D. Neuregulin and ErbB receptor signaling pathways in the nervous system. *Curr Opin Neurobiol* **11**, 287-296 (2001).
255. Holloway, E.M., *et al.* Mapping Development of the Human Intestinal Niche at Single-Cell Resolution. *Cell stem cell* **28**, 568-580 e564 (2021).
256. Jarde, T., *et al.* Mesenchymal Niche-Derived Neuregulin-1 Drives Intestinal Stem Cell Proliferation and Regeneration of Damaged Epithelium. *Cell stem cell* **27**, 646-662 e647 (2020).
257. Kuo, D., *et al.* HBEGF(+) macrophages in rheumatoid arthritis induce fibroblast invasiveness. *Science translational medicine* **11**(2019).
258. Moffitt, J.R., Lundberg, E. & Heyn, H. The emerging landscape of spatial profiling technologies. *Nature reviews. Genetics* **23**, 741-759 (2022).
259. Marchiando, A.M., Graham, W.V. & Turner, J.R. Epithelial barriers in homeostasis and disease. *Annual review of pathology* **5**, 119-144 (2010).
260. Coskun, M. Intestinal epithelium in inflammatory bowel disease. *Front Med (Lausanne)* **1**, 24 (2014).
261. Gitter, A.H., Wullstein, F., Fromm, M. & Schulzke, J.D. Epithelial barrier defects in ulcerative colitis: characterization and quantification by electrophysiological imaging. *Gastroenterology* **121**, 1320-1328 (2001).

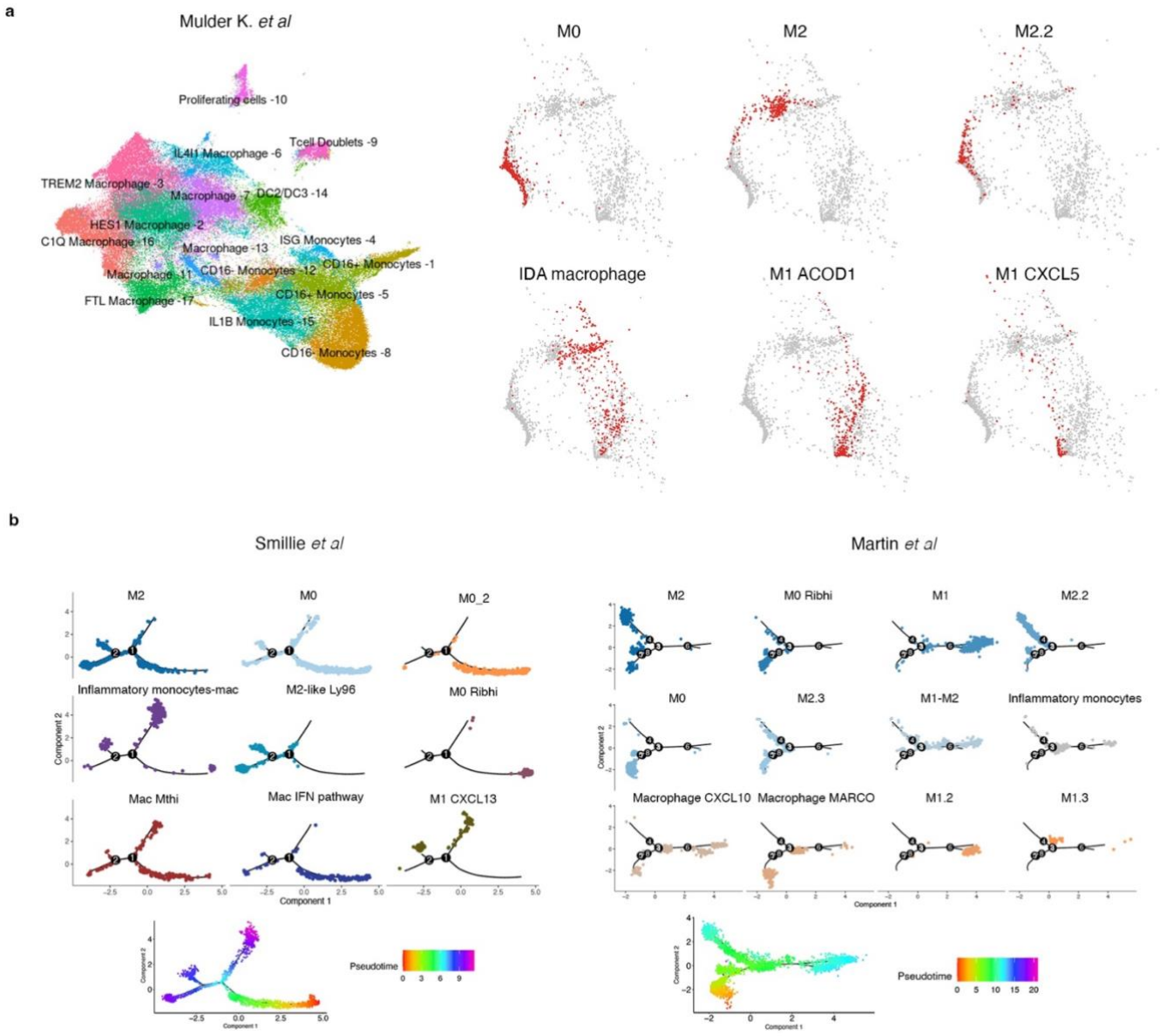
262. Grasberger, H., *et al.* Increased Expression of DUOX2 Is an Epithelial Response to Mucosal Dysbiosis Required for Immune Homeostasis in Mouse Intestine. *Gastroenterology* **149**, 1849-1859 (2015).
263. Burgueno, J.F., *et al.* Epithelial TLR4 Signaling Activates DUOX2 to Induce Microbiota-Driven Tumorigenesis. *Gastroenterology* **160**, 797-808 e796 (2021).
264. Allen, J.M., *et al.* Psychological stress disrupts intestinal epithelial cell function and mucosal integrity through microbe and host-directed processes. *Gut microbes* **14**, 2035661 (2022).
265. Atarashi, K., *et al.* Th17 Cell Induction by Adhesion of Microbes to Intestinal Epithelial Cells. *Cell* **163**, 367-380 (2015).
266. Elphick, D.A. & Mahida, Y.R. Paneth cells: their role in innate immunity and inflammatory disease. *Gut* **54**, 1802-1809 (2005).
267. Ayabe, T., *et al.* Secretion of microbicidal alpha-defensins by intestinal Paneth cells in response to bacteria. *Nature immunology* **1**, 113-118 (2000).
268. Gersemann, M., *et al.* Olfactomedin-4 is a glycoprotein secreted into mucus in active IBD. *Journal of Crohn's & colitis* **6**, 425-434 (2012).
269. van der Flier, L.G., Haegebarth, A., Stange, D.E., van de Wetering, M. & Clevers, H. OLFM4 is a robust marker for stem cells in human intestine and marks a subset of colorectal cancer cells. *Gastroenterology* **137**, 15-17 (2009).
270. Wigerblad, G., *et al.* Single-Cell Analysis Reveals the Range of Transcriptional States of Circulating Human Neutrophils. *J Immunol* **209**, 772-782 (2022).
271. Montaldo, E., *et al.* Cellular and transcriptional dynamics of human neutrophils at steady state and upon stress. *Nature immunology* **23**, 1470-1483 (2022).
272. Sturge, C.R., *et al.* TLR-independent neutrophil-derived IFN-gamma is important for host resistance to intracellular pathogens. *Proc Natl Acad Sci U S A* **110**, 10711-10716 (2013).
273. Rocha, B.C., *et al.* Type I Interferon Transcriptional Signature in Neutrophils and Low-Density Granulocytes Are Associated with Tissue Damage in Malaria. *Cell reports* **13**, 2829-2841 (2015).
274. Isles, H.M., *et al.* The CXCL12/CXCR4 Signaling Axis Retains Neutrophils at Inflammatory Sites in Zebrafish. *Frontiers in immunology* **10**, 1784 (2019).
275. Hartl, D., *et al.* Infiltrated neutrophils acquire novel chemokine receptor expression and chemokine responsiveness in chronic inflammatory lung diseases. *J Immunol* **181**, 8053-8067 (2008).
276. Bujko, A., *et al.* Transcriptional and functional profiling defines human small intestinal macrophage subsets. *J Exp Med* **215**, 441-458 (2018).
277. Delfini, M., Stakenborg, N., Viola, M.F. & Boeckxstaens, G. Macrophages in the gut: Masters in multitasking. *Immunity* **55**, 1530-1548 (2022).
278. Smythies, L.E., *et al.* Human intestinal macrophages display profound inflammatory anergy despite avid phagocytic and bacteriocidal activity. *J Clin Invest* **115**, 66-75 (2005).
279. Dharmasiri, S., *et al.* Human Intestinal Macrophages Are Involved in the Pathology of Both Ulcerative Colitis and Crohn Disease. *Inflamm Bowel Dis* **27**, 1641-1652 (2021).
280. Manzella, C.R., *et al.* Serum Serotonin Differentiates Between Disease Activity States in Crohn's Patients. *Inflamm Bowel Dis* **26**, 1607-1618 (2020).
281. Shiraishi, M., Yamaguchi, A. & Suzuki, K. Nrg1/ErbB signaling-mediated regulation of fibrosis after myocardial infarction. *FASEB J* **36**, e22150 (2022).
282. Liu, W., *et al.* Olfactomedin 4 deletion induces colon adenocarcinoma in Apc(Min/+) mice. *Oncogene* **35**, 5237-5247 (2016).
283. Amirbeagi, F., *et al.* Olfactomedin-4 autoantibodies give unusual c-ANCA staining patterns with reactivity to a subpopulation of neutrophils. *J Leukoc Biol* **97**, 181-189 (2015).
284. Liu, W. & Rodgers, G.P. Olfactomedin 4 expression and functions in innate immunity, inflammation, and cancer. *Cancer Metastasis Rev* **35**, 201-212 (2016).
285. Liu, W., *et al.* Olfactomedin 4 down-regulates innate immunity against Helicobacter pylori infection. *Proc Natl Acad Sci U S A* **107**, 11056-11061 (2010).

286. Chandrasekharan, B., Nezami, B.G. & Srinivasan, S. Emerging neuropeptide targets in inflammation: NPY and VIP. *American journal of physiology. Gastrointestinal and liver physiology* **304**, G949-957 (2013).
287. Adrian, T.E., *et al.* Human distribution and release of a putative new gut hormone, peptide YY. *Gastroenterology* **89**, 1070-1077 (1985).
288. Allen, J.M. & Bloom, S.R. Neuropeptide Y: a putative neurotransmitter. *Neurochem Int* **8**, 1-8 (1986).
289. Tatemoto, K., Carlquist, M. & Mutt, V. Neuropeptide Y--a novel brain peptide with structural similarities to peptide YY and pancreatic polypeptide. *Nature* **296**, 659-660 (1982).
290. Tatemoto, K., Lundberg, J.M., Jornvall, H. & Mutt, V. Neuropeptide K: isolation, structure and biological activities of a novel brain tachykinin. *Biochem Biophys Res Commun* **128**, 947-953 (1985).
291. Vona-Davis, L.C. & McFadden, D.W. NPY family of hormones: clinical relevance and potential use in gastrointestinal disease. *Curr Top Med Chem* **7**, 1710-1720 (2007).
292. Holzer, P., Reichmann, F. & Farzi, A. Neuropeptide Y, peptide YY and pancreatic polypeptide in the gut-brain axis. *Neuropeptides* **46**, 261-274 (2012).
293. Wahlestedt, C., *et al.* Neuropeptide Y receptor subtypes, Y1 and Y2. *Ann N Y Acad Sci* **611**, 7-26 (1990).
294. Tough, I.R., *et al.* Endogenous peptide YY and neuropeptide Y inhibit colonic ion transport, contractility and transit differentially via Y(1) and Y(2) receptors. *Br J Pharmacol* **164**, 471-484 (2011).
295. Cox, H.M. & Tough, I.R. Neuropeptide Y, Y1, Y2 and Y4 receptors mediate Y agonist responses in isolated human colon mucosa. *Br J Pharmacol* **135**, 1505-1512 (2002).
296. Fujimiya, M., *et al.* Neuropeptide Y induces fasted pattern of duodenal motility via Y(2) receptors in conscious fed rats. *American journal of physiology. Gastrointestinal and liver physiology* **278**, G32-38 (2000).
297. Saria, A. & Beubler, E. Neuropeptide Y (NPY) and peptide YY (PYY) inhibit prostaglandin E2-induced intestinal fluid and electrolyte secretion in the rat jejunum in vivo. *Eur J Pharmacol* **119**, 47-52 (1985).
298. Burclaff, J., *et al.* A Proximal-to-Distal Survey of Healthy Adult Human Small Intestine and Colon Epithelium by Single-Cell Transcriptomics. *Cell Mol Gastroenterol Hepatol* **13**, 1554-1589 (2022).
299. Pelon, F., *et al.* Cancer-associated fibroblast heterogeneity in axillary lymph nodes drives metastases in breast cancer through complementary mechanisms. *Nat Commun* **11**, 404 (2020).
300. Kieffer, Y., *et al.* Single-Cell Analysis Reveals Fibroblast Clusters Linked to Immunotherapy Resistance in Cancer. *Cancer Discov* **10**, 1330-1351 (2020).
301. Costa, A., *et al.* Fibroblast Heterogeneity and Immunosuppressive Environment in Human Breast Cancer. *Cancer cell* **33**, 463-479 e410 (2018).
302. Bonneau, C., *et al.* A subset of activated fibroblasts is associated with distant relapse in early luminal breast cancer. *Breast cancer research : BCR* **22**, 76 (2020).
303. Ware, M.F., Wells, A. & Lauffenburger, D.A. Epidermal growth factor alters fibroblast migration speed and directional persistence reciprocally and in a matrix-dependent manner. *Journal of cell science* **111 (Pt 16)**, 2423-2432 (1998).
304. Buechler, M.B., Fu, W. & Turley, S.J. Fibroblast-macrophage reciprocal interactions in health, fibrosis, and cancer. *Immunity* **54**, 903-915 (2021).

ANNEX



Annex Figure 1. CosMx™ Spatial Molecular Imaging (SMI) of healthy and inflammatory bowel disease (IBD) colonic tissue. Panels show tissue scanner (i), protein staining (ii) and cell type composition (iii) of one representative a, healthy control and b, ulcerative colitis patient, respectively.



Annex Figure 2. Intestinal Macrophages across studies. a, Projection of our intestinal macrophage clusters found in our study into the MoMAc-VERSE UMAP data (Mulder *et al.2021*). **b**, Pseudo-time trajectory analysis of Smillie *et al*- and Martin *et al- in silico* isolated colonic and ileal macrophages, respectively.