

**FINAL ENGINEERING PROJECT**

**SEARCH FOR A POTENTIALLY PREDICTIVE  
GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION  
FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED  
COLORECTAL CANCER**

**Castelli, Lucía – LU: 1133508**

Bachelor's Degree in Biotechnology

Tutors:

**Lic. Merlo, Joaquín**

**Universidad Argentina de la Empresa (UADE)**

**Functional and Molecular Glycomics Laboratory**

**Institute of Biology and Experimental Medicine – IBYME CONICET**

**Dr. Mariño, Karina**

**Universidad Argentina de la Empresa (UADE)**

**Functional and Molecular Glycomics Laboratory**

**Institute of Biology and Experimental Medicine – IBYME CONICET**

**18/07/2025**

**UADE**

**UNIVERSIDAD ARGENTINA DE LA EMPRESA**

**FACULTY OF ENGINEERING AND EXACT SCIENCES**

## ABSTRACT

Inflammatory bowel diseases (IBD), including Ulcerative Colitis (UC), are chronic inflammatory diseases that produce functional and structural abnormalities in the digestive tract. In IBD, the chronicity of inflammation increases the risk of colitis-associated colorectal cancer (CACRC). This type of neoplasia shows molecular and genetic patterns similar to those in sporadic colorectal cancer; however, both types of cancer differ in the chronology and sequence of appearance of these alterations, causing in the case of CACRC a faster malignant transformation, with multiple foci of origin. To reduce inflammation in IBD, first-line treatment employs immunosuppressive agents. However, there is debate over whether their prolonged use could promote tumor development. For this reason, the treatment of CACRC is complex, and today all efforts are dedicated to preventing malignancy through endoscopic surveillance, without biomarkers for the prediction of neoplasia in these patients.

Considering that protein glycosylation is one of the most abundant post-translational modifications, and that it mediates relevant cellular mechanisms (with a particularly important role in the immune response), the central objective of this Thesis is to identify a glycobiological signature that contributes to the development of new diagnostic and/or prevention strategies in patients with UC and at risk of developing CACRC. To this end, public transcriptomics databases of patients with intestinal inflammation (left ulcerative colitis and pancolitis), dysplasia and colitis-associated neoplasia were analyzed.

The joint analysis of these results and data from murine models allowed to delineate a signature of 10 dysregulated glycogenes in a model of CACRC with respect to one of intestinal inflammation (azoxymethane/sodium dextran sulfate, AOM-DSS): *Defa5*, *Lgals4*, *Lgals12*, *Slc26a2* (upregulated), *Icam1*, *Sell*, *Selp*, *St3gall*, *Chi3l1*, And *Chil3* (downregulated). We studied the dysregulation of four of these genes (*Defa5*, *Lgals4*, *Lgals12* and *St3gall*) by RT-qPCR in a model of DSS-induced colitis and CACRC (AOM-DSS). This approach allowed validating the altered expression of *Defa5*, and supports future research of the glycobiological signature obtained in samples from patients with CACRC, with the aim of implementing early diagnostic strategies in patients with UC.

## ABBREVIATIONS

Abreviatura	Significado
EII (IBD)	Enfermedades Inflamatorias Intestinales ( <i>Inflammatory Bowel Diseases</i> )
CU (UC)	Colitis Ulcerosa ( <i>Ulcerative Colitis</i> )
EC (CD)	Enfermedad de Crohn ( <i>Crohn's Disease</i> )
CUq (qUC)	Colitis Ulcerosa Quiescente ( <i>Quiescent Ulcerative Colitis</i> )
CU-I (L-UC)	Colitis Ulcerosa Izquierda ( <i>Left-side UC</i> )
CCR (CRC)	Cáncer Colorrectal ( <i>Colorectal Cancer</i> )
CCRAC (CACRC)	Cáncer Colorectal Asociado a Colitis ( <i>Colitis-Associated Colorectal Cancer</i> )
CCRE (sCRC)	Cáncer Colorrectal Esporádico ( <i>Sporadic Colorectal Cancer</i> )
LGD	Displasia de Bajo Grado ( <i>Low-Grade Dysplasia</i> )
HGD	Displasia de Alto Grado ( <i>High-Grade Dysplasia</i> )
AOM	Azoximetano ( <i>Azoxymethane</i> )
DSS	Dextrán Sulfato Sódico ( <i>Diocetyl Sodium Sulfosuccinate</i> )
MSI	Inestabilidad de Microsatélites ( <i>Microsatellite Instability</i> )
CIMP	Fenotipo Metilador de Islas CpG ( <i>CpG-Island Methylator Phenotype</i> )
dMMR	Reparación Deficiente de Emparejamientos Erróneos de ADN ( <i>deficient DNA MissMatch Repair</i> )
CEP	Colangitis Esclerosante Primaria
MAMPs	Patrones Moleculares Asociados a Microbios ( <i>Microbe-Associated Molecular Patterns</i> )
TLRs	Receptores tipo Toll ( <i>Toll-Like Receptors</i> )
APCs	Células Presentadoras de Antígenos ( <i>Antigen Presenting Cells</i> )
NETs	Trampas Extracelulares de Neutrófilos ( <i>Neutrophil Extracellular Traps</i> )
SLCs	Portadores de Soluto ( <i>Solute Carriers</i> )
MH	Curación de la Mucosa ( <i>Mucosal Healing</i> )
MES	Subpuntuación Endoscópica de Mayo ( <i>Mayo Endoscopic Subscore</i> )
ICI	Inhibidores de Puntos de Control Inmunitario ( <i>Immune Checkpoint Inhibitors</i> )
TMB	Carga Mutacional Tumoral ( <i>Tumor Mutational Burden</i> )
JAKs	Janus Quinasas ( <i>Janus Kinases</i> )
CDR	Dominio de Reconocimiento de Carbohidratos ( <i>Carbohydrate Recognition Domain</i> )
DEGs	Genes Diferencialmente Expresados ( <i>Differentially Expressed Genes</i> )
logFC	Logaritmo de las veces de cambio ( <i>log Fold Change</i> )
ORA	Ánalisis de Sobrerepresentación ( <i>Overrepresentation Analysis</i> )
CIB	Carga Inflamatoria Acumulada ( <i>Cummulative Inflammatory Burden</i> )

## TABLE OF CONTENTS

<b>1. INTRODUCTION .....</b>	<b>7</b>
<b>1.1. SPORADIC AND COLITIS-ASSOCIATED COLORECTAL CANCER: PATHOPHYSIOLOGY, RISK FACTORS, AND CURRENT CHALLENGES.....</b>	<b>7</b>
<b>1.2. INFLAMMATORY BOWEL DISEASES (IBD): ULCERATIVE COLITIS, CHARACTERISTICS, SYMPTOMS, DIFFERENCES WITH CROHN'S DISEASE AND CURRENT TREATMENT.....</b>	<b>13</b>
<b>1.2.1. SYMPTOMS.....</b>	<b>13</b>
<b>1.2.2. ETIOLOGY AND PATHOPHYSIOLOGY .....</b>	<b>14</b>
<b>1.2.3. CLASSIFICATION OF INFLAMMATORY BOWEL DISEASES BY EXTENT OF INFLAMMATION: ULCERATIVE PROCTITIS (E1), LEFT ULCERATIVE COLITIS (E2), AND EXTENSIVE ULCERATIVE COLITIS OR PANCOLITIS (E3) ....</b>	<b>17</b>
<b>1.2.4. MUCOSAL HEALING AND MAYO ENDOSCOPIC SUBSCORE .....</b>	<b>18</b>
<b>1.2.5. CURRENT STANDARD TREATMENT FOR ULCERATIVE COLITIS.....</b>	<b>19</b>
<b>1.2.6. GLYCOCOLOGY AS A LANGUAGE OF CELLULAR COMMUNICATION IN PHYSIOLOGICAL AND PATHOLOGICAL PHENOMENA .....</b>	<b>20</b>
<b>1.2.1. GLYCOGENES.....</b>	<b>25</b>
<b>1.3. MURINE MODELS FOR THE STUDY OF ULCERATIVE COLITIS AND COLITIS-ASSOCIATED COLORECTAL CANCER.....</b>	<b>29</b>
<b>2. OBJECTIVES .....</b>	<b>31</b>
<b>2.1. GENERAL OBJECTIVE .....</b>	<b>31</b>
<b>2.2. SPECIFIC OBJECTIVES .....</b>	<b>31</b>
<b>3. HYPOTHESIS .....</b>	<b>32</b>
<b>4. MATERIALS AND METHODS.....</b>	<b>33</b>
<b>4.1. BIOINFORMATICS ANALYSIS .....</b>	<b>33</b>
<b>4.1.1. OBTAINING DATA FROM PUBLIC COHORTS .....</b>	<b>33</b>
<b>4.1.2. QUALITY CONTROL AND PREPROCESSING OF PUBLIC COHORT DATA.....</b>	<b>40</b>

---

4.1.3. FUNCTIONAL ANALYSIS OF TRANSCRIPTOMIC DATA .....	41
4.1.4. DEVELOPMENT OF BIOINFORMATICS TOOLS .....	43
<b>4.2. TRANSCRIPTOMICS AND/OR RT-QPCR STUDIES IN EXPERIMENTAL MODELS.....</b>	<b>43</b>
4.2.1. MURINE MODELS .....	43
4.2.2. RT-QPCR.....	44
4.2.3. STATISTICAL ANALYSIS .....	46
<b>5. RESULTS.....</b>	<b>47</b>
5.1. WORKSCHEME.....	47
5.2. DEVELOPMENT OF AN INTERACTIVE APPLICATION AND AN <i>R</i> PACKAGE FOR DIFFERENTIAL EXPRESSION ANALYSIS.....	48
<b>5.3. ANALYSIS OF TRANSCRIPTIONAL PROFILES IN PUBLIC DATABASES .....</b>	<b>52</b>
5.3.1. DIFFERENTIAL GENE EXPRESSION BETWEEN PATIENTS WITH ULCERATIVE COLITIS AND HEALTHY SUBJECTS.....	52
5.3.2. ANALYSIS OF PATIENTS WITH DYSPLASIA, LEFT ULCERATIVE COLITIS (E2), AND PANCOLITIS (E3).....	65
5.3.3. ANALYSIS OF BIOPSIES OF PATIENTS WITH COLITIS-ASSOCIATED COLORECTAL CANCER AND BIOPSIES OF PATIENTS WITH ULCERATIVE COLITIS .....	74
5.3.4. STUDY OF RELEVANT GENETIC DYSREGULATION IN HUMAN PATHOLOGIES WITH TRANSCRIPTOMIC DATA FROM MURINE MODELS .....	79
<b>5.4. VALIDATION OF RESULTS BY RT-QPCR.....</b>	<b>82</b>
<b>6. CONCLUSIONS AND DISCUSSION .....</b>	<b>85</b>
6.1. DEVELOPMENT OF AN ONLINE APPLICATION AND <i>R</i> PACKAGE FOR TRANSCRIPTOMIC DATA ANALYSIS 85	
6.2. INFLAMMATION AND HEALTHY TISSUE: GLYCOCOLOGICAL DIFFERENCES .....	85
6.3. REMISSION STATUS.....	86

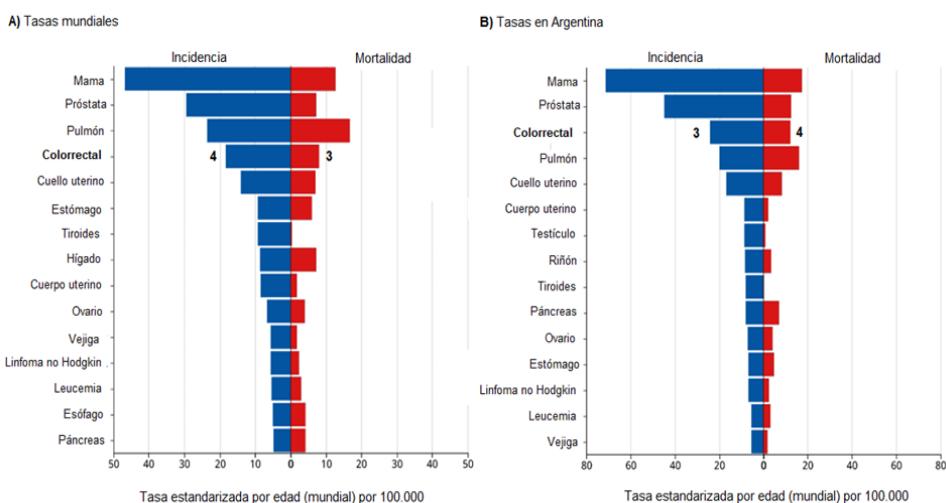
---

<b>6.4. EXTENT OF INFLAMMATION AND DYSPLASIA: MOLECULAR MECHANISMS.....</b>	<b>88</b>
<b>6.5. FROM INFLAMMATION TO MALIGNANT TRANSFORMATION: ANALYSIS OF PATIENT DATABASES AND AN EXPERIMENTAL MODEL OF COLITIS-ASSOCIATED COLORECTAL CANCER.....</b>	<b>91</b>
<b>6.6. GLYCOBIOLOGICAL SIGNATURE POTENTIALLY PREDICTIVE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER .....</b>	<b>92</b>
6.6.1. <i>DEFa5</i> .....	92
6.6.2. GALECTINS: <i>LGALS4</i> AND <i>LGALS12</i> .....	94
6.6.3. <i>ST3GAL1</i> .....	95
6.6.4. ADHESION MOLECULES: <i>SELL</i> , <i>SELP</i> , AND <i>ICAM1</i> .....	95
6.6.5. <i>CHI3L1</i> AND <i>CHIL3</i> .....	96
6.6.6. <i>SLC26A2</i> .....	98
<b>6.7. CONTRIBUTIONS AND FUTURE PERSPECTIVES .....</b>	<b>99</b>
<b>7. BIBLIOGRAPHY .....</b>	<b>100</b>
<b>8. ANNEXES.....</b>	<b>124</b>
8.1. LIST OF GLYCOBIOLOGICAL GENES OR "GLYCOGENES" .....	124
8.2. DYSREGULATED GENES IN QUIESCENT NEOPLASM/UC BUT NOT IN NEOPLASM/HEALTHY PATIENTS 127	
8.3. EXTENDED GLYCOBIOLOGICAL SIGNATURE .....	128
8.4. DETAILED REAL-TIME PCR RESULTS .....	130
8.5. APPENDIX SCRIPTS, FULL TABLES, AND DEGFIND .....	131

## 1. INTRODUCTION

### 1.1. Sporadic and colitis-associated colorectal cancer: pathophysiology, risk factors, and current challenges

Colorectal cancer (CRC) is one of the oncological diseases with the highest incidence and lethality, being the fourth most frequent type of cancer in incidence and third in the world of cancer-related deaths, while in Argentina it is the third with the highest incidence and fourth with the highest mortality (Ferlay *et al.*, 2024) (Fig. 1). Approximately 9.4% of global cancer deaths were due to CRC in 2020 (Ferlay *et al.*, 2024), and in developed countries a higher incidence but lower mortality is observed (Arnold *et al.*, 2017). In 2022, there were 8,800 deaths from CRC in Argentina, making it the fourth leading cause of death from neoplastic diseases after breast, lung, and prostate cancer (Ferlay *et al.*, 2024) (Fig. 1). Although the maximum incidence of CRC occurs in patients over 70 years of age (Papamichael *et al.*, 2015) and although the total incidence has decreased over the last 10 years due to the adoption of routine colonoscopy as an early detection test, an increase in the incidence has been observed in people under 55 years of age (Siegel *et al.*, 2018).



**Figure 1. Incidence and mortality rates of colorectal cancer in relation to other types of cancer with higher incidence and mortality rates.** Age-Standardized Rate (Worldwide) (ASR, *Age-Standardized Rate*) per 100,000, showing incidence and mortality for both sexes in 2022. A) Globally, colorectal cancer has the fourth highest incidence globally (ASR = 18.4) and is the third deadliest type of cancer (ASR = 8.1). B) In Argentina, colorectal cancer is the third with the highest incidence (ASR = 24.2) and the fourth most deadly (ASR = 12.2). Source: <https://gco.iarc.who.int/today/> (Ferlay *et al.*, 2024).

Although 70% of CRCs affect the colon and 30% affect the rectum (Baek & Kim, 2020), both locations are grouped together in research and clinical practice. This is because both the colon and rectum are part of the large intestine, considered a single organ because they share a similar anatomical and histological structure, with common functions such as concentration, reabsorption, transport and excretion of feces (Alzahrani *et al.*, 2021). In addition, in both cases CRC is caused by an aberrant proliferation of glandular epithelial cells (Hossain *et al.*, 2022). In particular, this type of cancer usually arises as a benign adenomatous polyp, which can evolve into an advanced adenoma with high-grade dysplasia, and then into invasive cancer with the possibility of metastasis. These different stages reflect the progressive accumulation of genetic alterations typical of the malignancy process, related to proteins that regulate the proliferation, survival, migration and invasion of tumor cells. In CRC, mutations are frequently observed in APC genes (a tumor suppressor gene that regulates the Wnt signaling pathway and whose inactivation favors abnormal cell proliferation, with 85% frequency), TP53 (a tumor suppressor gene key in the response to DNA damage and in the induction of apoptosis, with 35-55%), KRAS (involved in the transduction of signals that promote cell growth and survival, 35-45%), TGFBR2 (25-30%) and SMAD4 (10-35%) (de Freitas Junior & Morgado-Díaz, 2016; Markowitz & Bertagnolli, 2009).

CRC is classified into three types: (a) sporadic (70% of cases), caused by point mutations in individual cells; (b) hereditary (5%), which includes polyposis and non-polyposis forms, where a mutation in one allele is inherited and the disease appears when the other mutates; and (c) familial (25%), technically also hereditary, but not belonging to the forms defined for (b) (Mármol *et al.*, 2017; Ullah *et al.*, 2023). As environmental factors, obesity, low physical activity, active and passive smoking, alcohol consumption, an unhealthy diet and/or diet rich in red meat and salt, changes in the gut microbiome and reduced diversity are closely related to the incidence of colorectal cancer (Lewandowska *et al.*, 2022; Roshandel *et al.*, 2024).

Currently, four Consensus *Molecular Subtypes* (CMS) of colorectal tumors are recognized, associated with different prognosis and response to treatments (Guinney *et al.*, 2015; Ten Hoorn *et al.*, 2022): (a) CMS1 (14%) or "*MSI-like*" (*MicroSatellite Instability-like*), with mutations in genes involved in DNA repair such as *BRAF*, which generates a high mutational load, and with hypermethylation of CpG islands (*CpG-Island Methylator*

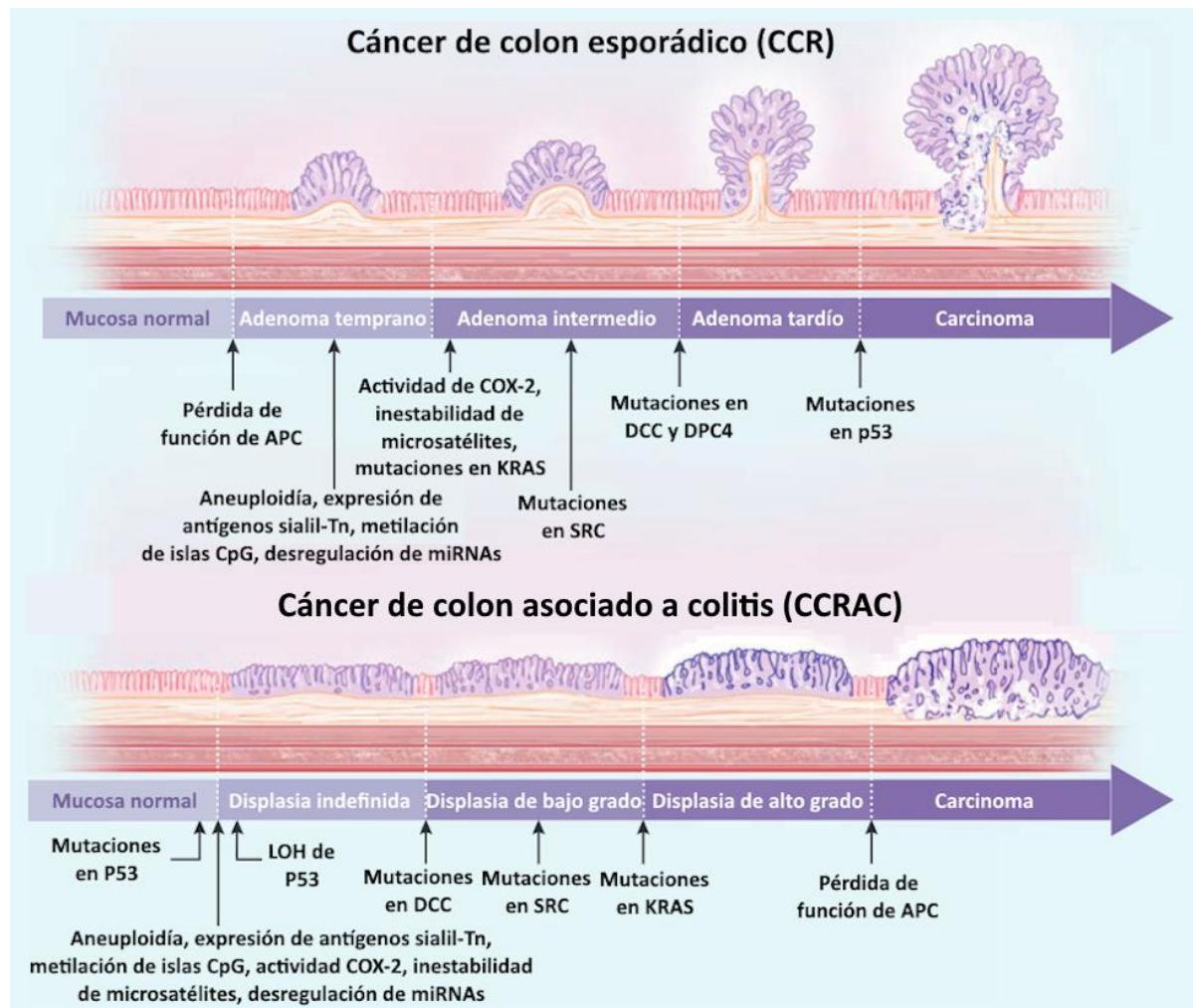
*Phenotype*, CIMP); (b) CMS2 (37%) or "canonical", with high chromosomal instability and activation of the Wnt and Myc signaling pathways, involved in tumorigenesis; (c) CMS3 (13%) or "metabolic", with dysregulation of metabolic pathways and mutated *KRAS*; and (d) CMS4 (23%) or "mesenchymal", with morphological characteristics typical of the epithelial-mesenchymal transition (EMT), a CIMP phenotype and activation of the TGF- $\beta$  pathway (Franke *et al.*, 2019).

In clinical practice, tumor mutational *burden* (TMB) is an important criterion for evaluating courses of action. A high BMR is associated with a better response to treatments with immune checkpoint *inhibitors* (ICIs), such as anti-PD-1 therapies. Patients with several types of cancer treated with them have a longer survival if their tumors have a high mutational load, because the consequent neoantigens are more easily recognizable by the immune system (Hugo *et al.*, 2016; Le *et al.*, 2015; Rizvi *et al.*, 2015; Rosenberg *et al.*, 2016; Snyder *et al.*, 2014; Van Allen *et al.*, 2015). On the other hand, the inactivation of genes involved in DNA repair (*deficient DNA MissMatch Repair*, dMMR) can generate MicroSatellite Instability (MSI). This phenomenon occurs when microsatellites, repetitive regions of non-coding DNA, change their number of repeated bases due to errors that are not corrected. For this reason, MSI indicates a malfunction of the DNA repair machinery, which favors the accumulation of mutations in tumor cells. Because of this, a high MSI promotes, like a high TMB, the generation of immunogenic neoantigens that facilitate the recognition and more effective destruction of cancer cells by the immune system (Lin *et al.*, 2020, McGranahan *et al.*, 2016; Riaz *et al.*, 2016; Schumacher & Schreiber, 2015). Therefore, both parameters (MSI and TMB) are used in the clinic as molecular diagnostic tools to stratify patients and select suitable candidates for immunotherapy.

Colitis-associated colorectal cancer (CCRAC) is a neoplasm associated with inflammatory bowel diseases (IBD), such as Ulcerative Colitis (UC) and Crohn's Disease (CD). They cause functional and structural abnormalities of the digestive tract, and the chronicity of inflammation promotes neoplastic development (Beaugerie & Itzkowitz, 2015). Although the incidence of colorectal cancer is higher in people with IBD than in the general population, within IBD, it is twenty times higher for patients with UC, while in CD it is three times higher (Lewandowska *et al.*, 2022).

CCRAC evolves through the "inflammation-dysplasia-carcinoma" sequence (Porter *et al.*, 2021), in which inflammation lays the groundwork for the development of dysplasia that, while benign, increases the risk of malignant progression by up to four times (Zhou *et al.*, 2023b). This progression can range from low-grade dysplasia, with or without high-grade dysplasia before becoming CCRAC (Zhou *et al.*, 2023b). Chronic inflammation can lead to the development of a precancerous mucosa through the production of reactive oxygen species as part of the immune response, generating oxidative damage and mutations. They end up resulting in the activation of signaling pathways that favor carcinogenesis and silence tumor suppression (Beaugerie & Itzkowitz, 2015). In addition, selective pressure is exerted in CCRAC that accelerates the evolution of malignant cells, due to the recurrent-remitting nature of inflammation and proliferative epithelial regeneration to replace damaged cells (Porter *et al.*, 2021). As for the specific form of this evolution, some evidence seems to indicate that the chronically inflamed mucosa of patients with IBD experiences "field cancerization," i.e., cancer-associated molecular alterations that accumulate over time even before there is histological evidence of dysplasia. This could explain the predisposition of the IBD mucosa to develop precancerous and multifocal cancerous changes along the gastrointestinal tract. In fact, in UC, prolonged inflammation causes widespread genetic and epigenetic changes throughout the colonic mucosa, not just isolated lesions. On the other hand, a countertheory has been proposed that is also supported by evidence and is called the "Big Bang". In this model, protumorigenic cells with multiple driver mutations arise spontaneously, followed by clonal expansion that gives rise to the tumor (Zhou *et al.*, 2023b).

CCRAC shows molecular and genetic patterns similar to those of sporadic colorectal cancer (ERCC), such as mutations in *APCs*, activation of the *KRAS* oncogene, induction of COX-2 (cyclooxygenase-2, a pro-inflammatory enzyme that promotes tumorigenesis through the synthesis of prostaglandins), and loss of function of *TP53* (Beaugerie & Itzkowitz, 2015; Zhou *et al.*, 2023b). However, both types of colorectal cancer (CRC) are distinguished by the different chronology and sequence of appearance of these alterations, causing in the case of CCRAC a more rapid malignant transformation with multiple foci of origin along the colon (Zhou *et al.*, 2023b) (**Fig. 2**).



**Figure 2. Temporal and sequential differences in molecular events that result in the development of sporadic CRC and colitis-associated CRC.** In ERCC, most sporadic tumors come from a previous lesion that progresses to adenocarcinoma, with the loss of APC function being the gateway to the presence of other alterations such as mutation, or loss of heterozygosity (LOH, *Loss Of Heterocigosity*) of *TP53*. In colitis-associated colorectal cancer, a precursor dysplastic lesion appears in the mucosa of the colon and progresses over several stages until carcinoma forms. Although the molecular alterations in CCRAC are almost the same, the timing and sequence of onset differ in both cases. For example, in CCRAC, mutations or loss of function *TP53* are seen in early stages, while loss of APC function is a late and less common event. Adapted from Beaugerie and Itzkowitz, 2015.

The risk of developing CCRAC in a patient with IBD is increased if they also have primary sclerosing cholangitis (PSC), a condition of chronic inflammation of the bile ducts that occurs in about 10% of IBD patients (Zhou *et al.*, 2023b). PSC increases the risk of developing CCRAC by 5 to 9 times compared to IBD patients who do not have this disease (Zhou *et al.*, 2023b). In addition, PSC is more common in patients with UC than with CD (Le Berre *et al.*, 2023). On the other hand, microbiota-host interactions play an important role in both IBD and

CCRAC (Kang & Martin, 2017; Li *et al.*, 2023). In experimental models of CCRAC with mice genetically modified such that they are susceptible to inflammation or cancer, no tumor development was observed when the mice were germ-free or treated with antibiotics (Beaugerie & Itzkowitz, 2015).

First-line treatment for patients with IBD employs immunomodulatory agents. However, it has been proposed that its prolonged use could promote tumor development (Axelrad *et al.*, 2016). The use of thiopurines increases the risk of tumor development, but this risk decreases if the drug is stopped: in the case of biological therapies (e.g., anti-TNF), conflicting results were observed. The reason behind this paradoxical behavior is not yet fully understood, but one theory suggests that even residual levels of inflammation after IBD treatment could be enough to drive tumor development (Zhou *et al.*, 2023b). In this sense, the treatment of patients with CCRAC becomes difficult, and currently, although controlling intestinal inflammation is essential in IBD, the focus is on preventing malignancy through endoscopic surveillance. Since the risk of developing CCRAC increases with the extent and duration of inflammation, surveillance begins today around eight years post-diagnosis; however, some patients develop CCRAC before eight years of IBD, including ~13% who already have cancer at the time of initial diagnosis (Zhou *et al.*, 2023b). Patients with right-sided tumors require particular attention, as they tend to have a worse prognosis and are more prone to metastasis than left-sided tumors (Adebayo *et al.*, 2023).

If CCRAC is found early, treatment is usually surgical removal plus chemotherapy, plus radiation therapy if the tumor cannot be completely removed. Unfortunately, chemotherapy can lead to unwanted side effects such as anemia or neutropenia (a low neutrophil or red blood cell count, respectively), mucositis, vomiting, fatigue, hematologic disorders, and liver toxicity (Adebayo *et al.*, 2023). In particular, oxaliplatin generates peripheral neurotoxicity in more than 85% of patients, a serious but reversible clinical problem that limits the dose and could lead to discontinuation of treatment (Cheng *et al.*, 2023). In this sense, and although there are important differences between CCRAC and sporadic CRC (including that it tends to present tumors of higher malignant grade, less differentiated and more likely to be multifocal and synchronous, especially in advanced stages), current treatments for patients with colorectal cancer are the same regardless of the mechanisms of generation. These treatment options include surgery,

chemotherapy, radiation therapy, immunotherapy, and targeted treatments, such as vascular endothelial growth factor (VEGF) inhibitors (Herrera-Gómez *et al.*, 2022). Compared with sporadic tumor surgeries, surgeries in patients with CRCC are usually more extensive (such as total proctocolectomy) rather than segmental resections, even when the tumor location is similar. This is due to the multifocal nature of neoplasm in IBD (Holubar *et al.*, 2021). Notably, the use of ICIs, while effective in counteracting tumor immunosuppression, can cause *immune-mediated colitis* (ImC) in the colon, an autoimmune response due to the nature of the treatment. The presence of IBD predisposes the patient to develop BMI but does not seem to worsen survival (Giesler *et al.*, 2025). In turn, a 2020 study in a murine model of CCRAC concluded that anti-PD1 treatment did not reduce tumor proliferation or even enhance it, without intensifying intestinal inflammation (Collard *et al.*, 2020).

Currently, clinical research focuses on identifying subsets of CRC patients who respond effectively to different treatments and on developing combinations of clinical strategies to treat patients with different genetic makeups (Gandini *et al.*, 2023). The treatment of patients with CCRAC is complex, and today all efforts are dedicated to preventing malignancy through endoscopic surveillance, but there are no biomarkers for the prediction of neoplasia in these patients.

## **1.2. Inflammatory bowel diseases (IBD): ulcerative colitis, characteristics, symptoms, differences with Crohn's disease and current treatment**

### **1.2.1. Symptoms**

Ulcerative colitis is defined by a chronic inflammation of the colonic mucosa, which begins from the rectum and extends proximally continuously, symmetrically and circumferentially. Colonoscopy frequently reveals a loss of typical vascular pattern, granularity, friability, and ulceration, as well as cryptic abscesses and inflammatory polyps (also called pseudopolyps), which eventually shorten the colon. This condition usually has a relapsing and remitting course (Lynch & Hsu, 2023; Zhou *et al.*, 2023b; Sambuelli *et al.*, 2019). The most common symptoms include rectal bleeding, fecal incontinence or urge to defecate, diarrhea, mucous secretions, and abdominal pains similar to cramps, especially before and relieved by

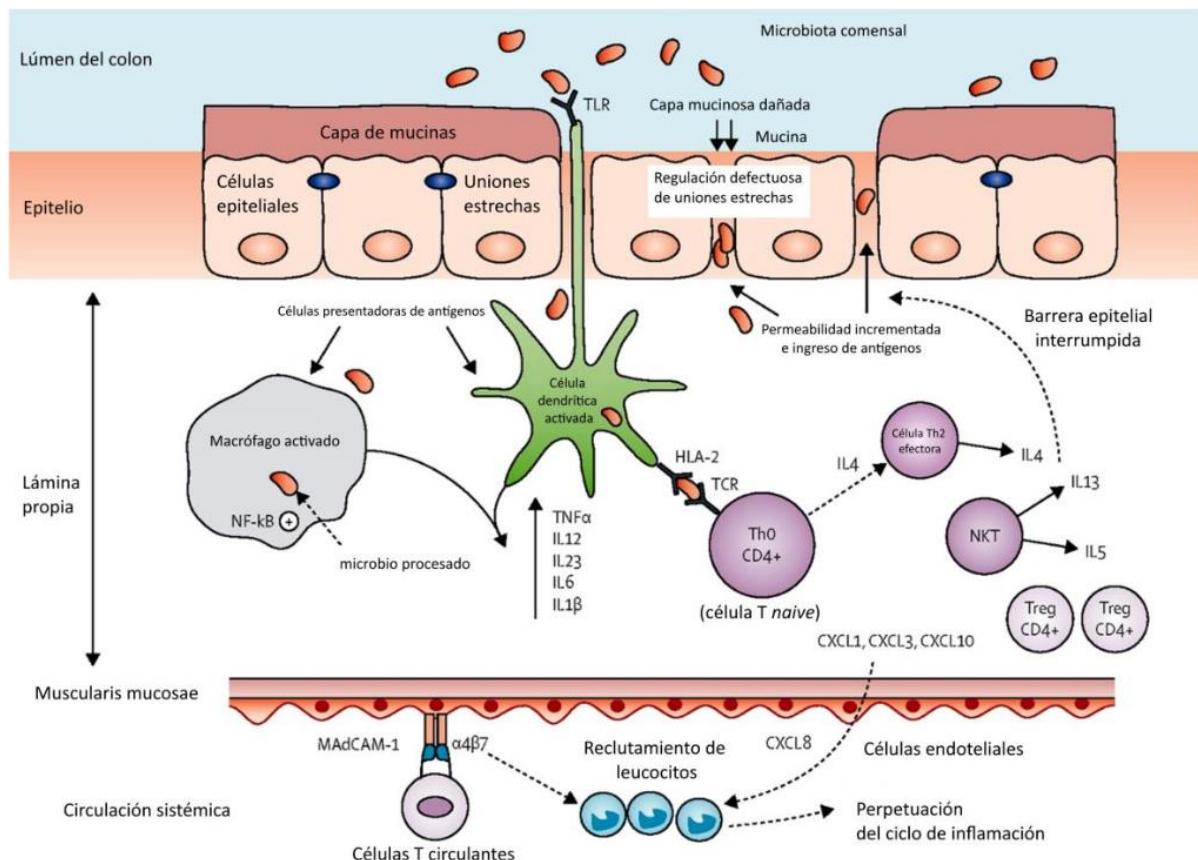
bowel movements. In addition, 20-35% of patients may present with extraintestinal manifestations such as peripheral arthritis, and more commonly in ulcerative colitis than in Crohn's disease, primary sclerosing cholangitis, and pyoderma gangrenosum (Le Berre *et al.*, 2023). It can also cause systemic symptoms such as fatigue, nausea or vomiting, weight loss, and fever. Because many of these symptoms are not specific to UC or IBD in general, diagnosis should be based on the combination of clinical, serological, endoscopic, histopathological, and radiological features (Gecse & Vermeire, 2018). In the long term, UC can significantly reduce quality of life and lead to hospitalization and even colectomy, total or partial removal of the colon (Boal Carvalho & Cotter, 2017).

### **1.2.2. Etiology and pathophysiology**

The pathogenesis of the disease is characterized by a defective intestinal epithelial barrier with less mucus secretion, essential for lubricating the passage of food and protecting the epithelium against pathogens. Mucus is composed mostly of mucins (highly glycosylated proteins), mainly mucin 2 (MUC2), which it homopolymerizes by means of disulfide (S-S) bonds at its cysteine-rich N and C ends, forming a flexible network (Grondin *et al.*, 2020; Le Berre *et al.*, 2023). In UC there is a lower production of mucin (partly due to the reduction in the number of calcitine cells that secrete them), and, in addition, their sulfation is altered, modifying their properties (Klinken *et al.*, 1999). These changes, added to the damage to the epithelial barrier, increase permeability, increasing the entry of antigens from the lumen in a feedback *loop* that promotes inflammation (Ordás *et al.*, 2012). Because many bacteria that colonize the gut are associated with *mucus*, mucins are critical in the establishment of the microbiota (Joja *et al.*, 2025).

The integrity of the epithelial barrier is further compromised by the overexpression of claudins, pore-forming proteins at the apical tight junctions involved in paracellular communication (Le Berre *et al.*, 2023). The intestine also produces defensins, peptides initially characterized by their antibacterial activity, but today it is known that they also play a highly relevant role in the repair of the epithelium and the maintenance of the epithelial barrier (Fu *et al.*, 2023). Although an aberrant innate and adaptive immune response against microbial

antigens has been proposed as a mechanism in UC (**Fig. 3**), epithelial injury in this disease is becoming increasingly relevant (Wan *et al.*, 2022).



**Figure 3. Pathophysiology of ulcerative colitis.** The loss of tight joints and the reduction of *Mucus* they increase the permeability of the epithelial, allowing antigen-presenting cells (APCs) to recognize the commensal microbiota in the lamina propria. This activates pathways such as NF- $\kappa$ B, stimulating the production of pro-inflammatory cytokines and leukocyte recruitment. CPAs also induce the differentiation of naïve CD4+ T cells into effector TH2 cells that release IL-4. In addition, NKT cells produce IL-13 that further damage the epithelium. Circulating T lymphocytes with integrin- $\alpha$ 4 $\beta$ 7 adhere to mucosal vascular adhesion molecule 1 (MAdCAM-1) of blood vessel endothelial cells, with increased expression in the inflamed intestine. Adapted from Ordás *et al.*, 2012.

The alteration in the composition and diversity of the gut microbiota, understood as an imbalance between commensal and (potentially) pathogenic microorganisms, is called dysbiosis, and its connection with IBD is widely studied (Santana *et al.*, 2022). In the gut of UC patients, there is a decline in species such as *Roseburia hominis* and *Faecalibacterium prausnitzii*, commensal bacteria that produce butyrate, a key short-chain fatty acid as an energy source for colonocytes with antimicrobial and anti-inflammatory activity (Le Berre *et al.*, 2023; Schulthess *et al.*, 2019). Butyrate deficiency favors anaerobic metabolic processes, which

generate less energy than aerobic ones, resulting in oxidative stress and damage to the epithelium. In addition, anaerobic processes favor the expansion of facultative anaerobes, which release toxins that cause more damage to the epithelium and intestinal mucosa. In addition, in active UC there are fewer bacteria of the genus *Bacteroides* the greater the inflammation, but in patients in remission their number increases; this suggests their possible immunomodulatory role, possibly related to the suppression of inflammatory responses mediated by type 1 (TH1) and 2 (TH2) T helper cells (Nishihara *et al.*, 2021; Popov *et al.*, 2021). Conversely, *Listeria monocytogenes*, which induces a TH1-type immune response, has been shown to be increased in patients with IBD (Santana *et al.*, 2022).

In UC, the loss of tight epithelial junctions and the reduction of *mucus* allows bacteria to cross the intestinal barrier, aberrantly activating *antigen presenting cells* (APCs), i.e. macrophages and dendritic cells, present in the lamina propria (Kałužna *et al.*, 2022; Ordás *et al.*, 2012). This occurs by recognizing *Microbe-Associated Molecular Patterns* (MAMPs) by *Toll-like receptors* (TLRs), which triggers signaling cascades such as NF- $\kappa$ B, which trigger the secretion of pro-inflammatory cytokines, such as IL-1 $\beta$ , IL-6, IL-18, and TNF- $\alpha$  (De Souza & Fiocchi, 2016; Neurath, 2014; Ordás *et al.*, 2012). This results in an activation of the adaptive response of a less tolerogenic profile, with regulatory T cells (TREGs) with reduced immunoregulatory activity. In addition, APCs promote the migration and degranulation of neutrophils, cells of the innate response that make up the majority of the immune infiltrate in the inflamed intestine. These cells release *Neutrophil Extracellular Traps* (NETs), cross-linked chromatin structures, and antimicrobial substances to immobilize and neutralize pathogens, which stimulate the production of proinflammatory cytokines. In the face of excessive and prolonged activation, neutrophils can significantly damage tissue (Leppkes & Neurath, 2020).

In UC there is also an atypical response of helper T cells (TH). For example, TH2 lymphocytes in CU are overactivated, contributing to chronic inflammation, mediating humoral response (e.g., by secretion of IL-4 and IL-10), and negatively affecting intestinal integrity by inducing apoptosis, inhibiting regeneration, and increasing claudin-2 expression through IL-13 secretion (Kałužna *et al.*, 2022; Ordás *et al.*, 2012). Finally, TH9 lymphocytes are also increased in CU. They secrete IL-9, an interleukin involved in the allergic and parasite response, mast cell and eosinophil activation, neutrophil infiltrate, and epithelial permeability by influencing

the expression of tight junction-forming proteins. Also of relevance in CU are TH17 cells, which differentiate from CD4+ T cells in the presence of IL-6, TGF- $\beta$ , IL-21 and IL-23. These lymphocytes, which are increased in CU, secrete cytokines such as IL-21, IL-22, IL-12, TNF- $\alpha$ , and IL-17A. The latter stimulate monocytes, epithelial and endothelial cells to release chemokines (CXCL1, CXCL2, CXCL5, CXCL8), which attract lymphocytes and neutrophils to the inflamed tissue (Kałužna *et al.*, 2022).

The atypical response of TH cells and the imbalance of TREG cells cause the dysregulated production of several key cytokines and have determined the therapeutic approach of ulcerative colitis towards anticytokine drugs. These cytokines include TNF, interferon gamma (TH1), IL-5, IL-6, IL-13 (TH2), IL-17, IL-21, and IL-22 (TH17), and have been implicated in various ways in the pathophysiology of ulcerative colitis (Kałužna *et al.*, 2022). Proinflammatory cytokines also upregulate the expression of mucosal vascular adhesion 1 (MadCAM-1) in the vascular endothelium of mucosal blood vessels, promoting the extravasation of leukocytes to tissue, through interaction with integrin- $\alpha 4\beta 7$  (Ordás *et al.*, 2012). Meanwhile, a 2022 study identified that intestinal and circulating B cells in UC were reduced in diversity and level of maturation (Uzzan *et al.*, 2022).

### **1.2.3. Classification of inflammatory bowel diseases by extent of inflammation: ulcerative proctitis (E1), left ulcerative colitis (E2), and extensive ulcerative colitis or pancolitis (E3)**

UC can be defined by the extent of colorectal inflammation at the radiographic, endoscopic, or histological level. The three UC subgroups defined by extension are:

1. Ulcerative proctitis (E1): Involvement limited to the rectum (i.e., the proximal extent of inflammation is distal to the rectosigmoid junction).
2. Left UC (E2) (also known as distal UC): Involvement limited to the portion of the colon and rectus distal to the splenic angle.
3. Extensive UC (E3), also known as pancolitis: involvement that extends proximal to the splenic angle.

This three-level classification system assists in the determination of medical treatment and prognosis: the extent of colitis influences disease activity or severity, medication use, and rates of hospitalization and colectomy, as well as the risk of developing colorectal cancer (Silverberg *et al.*, 2005). In particular, the cumulative risk of CRC after 30 years of disease ranges from 0 to 12% in patients with ulcerative proctitis, and from 4 to 47% in those with extensive disease or pancolitis (Silverberg *et al.*, 2005).

Approximately 20% to 40% of patients with UC have disease in all or most of the colon; these patients, with extensive ulcerative colitis, have the worst prognosis, accounting for the majority of deaths associated with UC (Cuffari *et al.*, 2005). These patients also have the least favorable response to treatment, a more aggressive disease course, and are more likely to suffer extraintestinal complications (such as arthropathy and uveitis), and require colectomy more frequently than those with left-sided disease; eventually 40% will even require a total colectomy (Cuffari *et al.*, 2005). In addition, and of particular interest for this Thesis, patients with extensive UC have the highest risk of developing dysplasia and colorectal cancer (Cuffari *et al.*, 2005; Munkholm, 2003). The case of a patient with pancolitis presents the specialist with the most difficult therapeutic decisions regarding the timing and use of certain medications compared to total colectomy. Disease extent and early age at diagnosis are important and independent determinants of an increased relative risk compared to the general population (Cuffari *et al.*, 2005).

#### **1.2.4. Mucosal healing and Mayo endoscopic subscore**

In the past, treatment of UC focused on symptom management and maintenance of clinical remission. Currently, however, the focus has shifted to mucosal healing (MH), because it is associated with a lower risk of recurrence, hospitalizations, and colectomy. The importance of MH has been known since 1966, when Wright and Truelove concluded that patients with MH had clinical remission more frequently after one year (40 vs. 18%) (Boal Carvalho & Cotter, 2017; Viscido *et al.*, 2022; Wright and Truelove, 1966). These researchers established the first criterion for evaluating endoscopic activity, but over the years others were developed, such as the Baron scale, the Rachmilewitz endoscopic index and the St. Mark index, but the most commonly adopted criterion in the clinic is the *Mayo Endoscopic Score* (MES) (Boal

Carvalho and Cotter, 2017). It consists of four components, each with its own score; the sum of these gives a total score between 0 and 12 (**see Table I**) (Schroeder *et al.*, 1987).

**TABLE I. Scoring system for the evaluation of ulcerative colitis activity.** Adapted from Schroeder *et al.*, 1987.

Frecuencia de las deposiciones
0 = Frecuencia normal para el paciente
1 = 1-2 más deposiciones de lo normal
2 = 3-4 más deposiciones de lo normal
3 = 5+ deposiciones más de lo normal
Sangrado rectal
0 = Nulo
1 = Vetas de sangre con heces menos de la mitad del tiempo
2 = Sangrado obvio con la mayoría de las deposiciones
3 = Paso únicamente de sangre
Hallazgos en rectosigmoidoscopia flexible
0 = Enfermedad normal o inactiva
1 = Enfermedad leve (eritema, patrón vascular disminuido, friabilidad leve)
2 = Enfermedad moderada (eritema marcado, patrón vascular ausente, friabilidad, erosiones)
3 = Enfermedad grave (sangrado espontáneo, ulceración)
Evaluación global del médico
0 = Normal
1 = Enfermedad leve
2 = Enfermedad moderada
3 = Enfermedad severa

In the clinic, mucosal healing is measured through endoscopic *healing* (EH). In this sense, MH is usually defined as a normal mucosa in endoscopic examinations (MES 0) or with mild erythema and friability (MES 1), although the first score is associated with better long-term results. MH is considered the ultimate goal of treatment, and many drugs in current use can induce and maintain it in most patients, but anti-TNF- $\alpha$  has the best results in moderate and severe UC (Boal Carvalho & Cotter, 2017; Viscido *et al.*, 2022).

### **1.2.5. Current Standard Treatment for Ulcerative Colitis**

The selection of treatments for UC depends on the severity and extent of the disease. In cases of mild to moderate UC, the first-line treatment to induce clinical remission is mesalazine rectally (with suppositories) in proctitis, or combined oral plus rectal in left (E2) or extensive (E3) colitis. If there is no response, corticosteroids, such as prednisone or beclomethasone, are used. Maintenance of remission is with mesalazine, and in proctitis in the form of rectal or oral; in left or extensive UC it is oral or topical. In moderate to severe UC, induction is initiated with

oral corticosteroids and maintenance is performed with immunosuppressants (thiopurines) or with biologics (anti-TNF such as infliximab, anti-integrin  $\alpha 4\beta 7$  such as vedolizumab, anti-IL12/23 p40 subunit such as ustekinumab), or small molecules (JAK kinase inhibitors such as tofacitinib for JAK2/3 or upadacitinib for JAK1, which block the pro-inflammatory signaling pathway JAK-STAT, or modulators of the S1P receptor, a G-protein-coupled receptor, by sequestering lymphocytes in lymph nodes, such as ozanimod). In severe cases or with poor pharmacokinetics, it is recommended to combine biologics with immunosuppressants. In severe acute colitis, the patient should be hospitalized and treated with intravenous corticosteroids, hydration, heparin to prevent thrombosis, and correction of anemia and electrolyte abnormalities. If there is no improvement within 3 days, cyclosporine or infliximab is given, and if there is no response within 4–7 days, colectomy is indicated (Cai *et al.*, 2021; Le Berre *et al.*, 2023).

### **1.2.6. Glycobiology as a language of cellular communication in physiological and pathological phenomena**

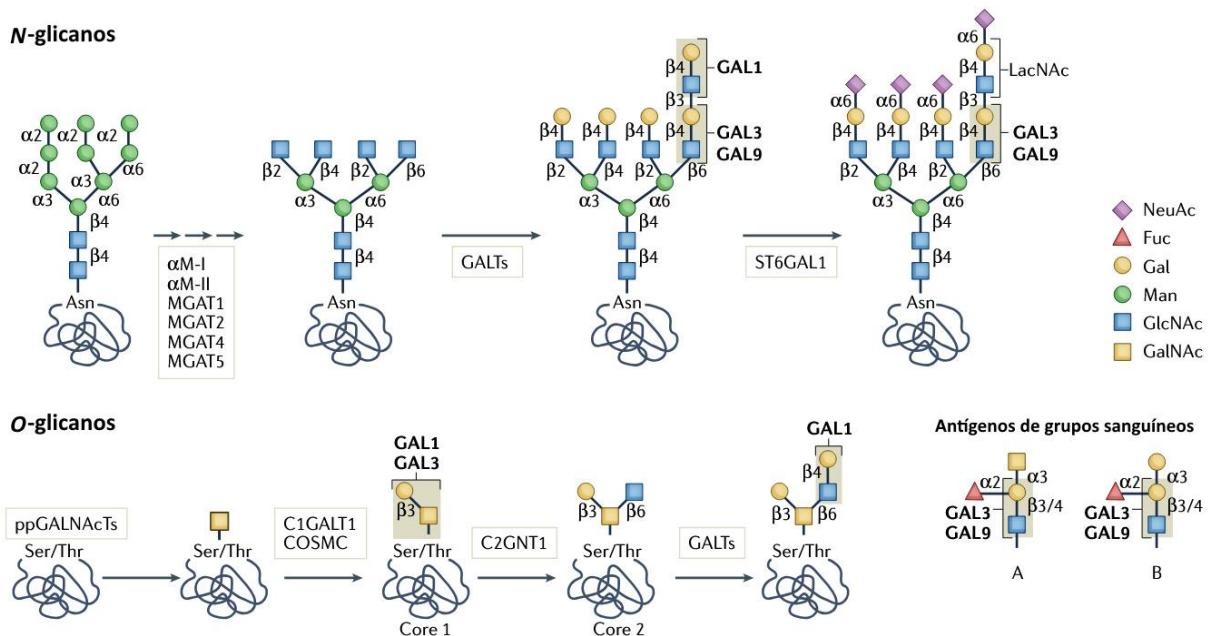
Glycobiology studies the structure, biosynthesis and function of glycans. These oligosaccharides or glycans, bound to lipids and proteins, form glycoconjugates that can be found both on the cell surface and secreted (e.g., as part of the extracellular matrix). Glycans are structurally diverse, and are involved in key cellular functions, establishing a language of communication between the cell and its environment (Varki, 2017). The heterogeneity of glycoconjugates is the result of their biosynthesis without a template; it involves a sequence of sequential enzymatic reactions mediated by glycosyltransferases and glycosidases, which in some cases are competitive (Varki *et al.*, 2022).

Glycosylation is one of the most common post-translational modifications, generating a great diversity of glycoforms (**Fig. 4**) that affect protein folding, stability, and recognition (Varki, 2017). Depending on their binding to protein, glycans can be classified into *N-glycans* (bound to the amino group of an asparagine residue, within the Asn-X-Ser/Thr consensus motif), and *O-glycans*, which bind to oxygen from serine or threonine residues. While *N-glycans* can be classified into high-mannose, hybrid or complex structures, the *most studied O-*

# UADE

## SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía

glycans are the so-called mucin-like ones, which are bound to the protein by an N-acetylgalactosamine (GalNAc) and present structural diversity (Reily et al., 2019).



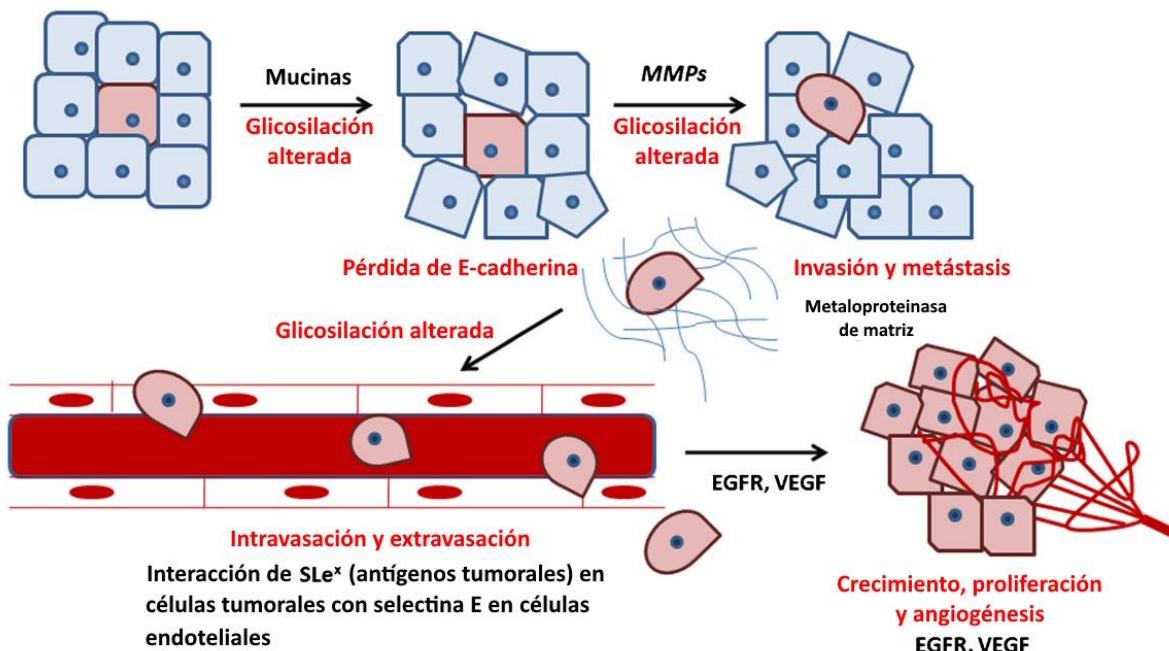
**Figure 4. Abbreviated biosynthesis of N- and O-glycans, and their affinity for different lectins.** The enzymes  $\alpha(1\text{-}2)$ -mannosidase I and II ( $\alpha M\text{-I}$  y  $\alpha M\text{-II}$ ) hydrolyze the precursor of N-glycans. N-acetylglucosaminyltransferases *MGAT1*, *MGAT2*, *MGAT4* y *MGAT5* they generate antennae, and galactosyltransferases (*GALTs*) generate structures of N-acetyllactosamine, essential for the specific binding of galectins. The incorporation of sialic acid  $\alpha(2,6)$ , mediated in N-glycans by *ST6GAL1*, can block the binding of Galectin-1 (*GAL1*), which recognizes only N-Terminal acetyllactosamines. On the other hand, mucin-like O-glycosylation is initiated by peptide N-acetylgalactosaminyltransferases (ppGALNAcTs). Then, the synthesis of the Core 1 begins with the action of *C1GALT1*, assisted by the chaperone *COSMC*. This core is recognized for *GAL1* and by *GAL3*, the latter with greater affinity. This core it can also be branched, forming a core 2 (*C2GNT1*), or sialized by *ST3GAL1*. In addition, *GAL3* y *GAL9* show affinity for blood group oligosaccharides A and B, present in various glycoconjugates. Adapted from Mariño *et al.*, 2023.

Thanks to the advancement of new technologies that allowed a better characterization of the glycome—the repertoire of glycans expressed at the cellular and/or tissue level—it was shown that it is highly dynamic, varying according to the nutritional status of the cell, its stage of development, or its degree of activation and/or differentiation, in the case of immune cells (Mariño *et al.*, 2011; Varki *et al.*, 2017).

The biological information encoded by the glycome is interpreted by proteins called lectins, which comprise different families of recognition of specific carbohydrates (Pinho & Rabinovich, 2025). These interactions play a very important role in cell-cell communication, especially at the immunological level, where the lectin-glycan axis can be exploited by the

tumor to promote its progression and evade the immune system (Cagnoni *et al.*, 2016; Girotti *et al.*, 2020; Mariño *et al.*, 2023).

In neoplastic diseases, the presence of aberrant glycosidic structures in tumor cells and in circulating glycoproteins in serum has been demonstrated (Girotti *et al.*, 2020; Munkley & Elliott, 2016; Vajaria & Patel, 2017) (**Fig. 5**). For this reason, multiple aberrant glycoconjugates serve as biomarkers for cancer diagnosis and prognosis (Vajaria & Patel, 2017). For example, overexpression of the *Sialil Lewis X* tumor antigen (SLex) in cancer cells can induce E-selectin expression in endothelial cells, facilitating migration and metastasis (Vajaria and Patel, 2017). In addition, aberrant glycosylation of mucins can favor the development of several types of tumors (Kufe, 2009; Radziejewska, 2024). In particular, mucin 1 (MUC1), normally located in the apical membrane of epithelial cells, may undergo altered glycosylation initiated by O-glycosylation of GALNT6 (N-acetylgalactosamintransferase 6) polypeptide and followed by additional modifications. This induces conformational changes that relocate MUC1 to the basolateral membrane, where it participates in the formation of a signaling complex together with galectin-3 (**Section 1.3.1**, of the **Introduction**). This interaction promotes the proteolytic degradation of E-cadherin, a key event in the epithelial-mesenchymal transition, a process where epithelial cells acquire migratory and invasive characteristics, typical of metastasis (Vajaria & Patel, 2017; Zhang *et al.*, 2014). The participation of galectin-3 in this signaling pathway is consistent with the fact that at high levels of expression it is associated with greater vascular invasion and worse survival in cancer patients (Song *et al.*, 2020). In addition, decreased levels of E-cadherin results in: (1) an increase in the expression of matrix metalloproteases, such as MMP2 and MMP9, which participate in metastasis through degradation of the extracellular matrix (Fonseca *et al.*, 2016) and (2) an increase in proliferation and angiogenesis due to the activation of epidermal growth factor (EGFR) (Vajaria and Patel, 2017) and consequently, vascular endothelial growth factor (VEGF) (Hung *et al.*, 2016; Vajaria & Patel, 2017).



**Figure 5. The role of altered glycosylation in metastasis.** Cancer progression is associated with changes in the glycosylation of cell surface proteins and mucins. These changes are involved in the loss of intercellular adhesion and metastatic spread of tumor cells. Metalloproteinases (MMPs) are involved in the degradation of the physical barrier and the extracellular matrix. During metastasis, tumor antigens such as SLe<sup>x</sup> interact with selectin in endothelial cells, which is also involved in growth, proliferation, and angiogenesis, a process mediated by EGFR and VEGF. Adapted from Vajaria and Patel, 2017.

In both IBD and colorectal cancer, intestinal glycans have been shown to be altered (Leite-Gomes *et al.*, 2022). As an example, in IBD, reduced sialylation and galactosylation in circulating immunoglobulin G was demonstrated, an alteration that is repeated in other inflammatory diseases and that was proposed as a possible diagnostic and/or prognostic tool (Šimurina *et al.*, 2018).

In UC and CRC, mucin O-glycans (which are key to epithelial barrier integrity and gut flora establishment, Bergstrom *et al.*, 2017) are altered. These truncated O-glycans include the Thomsen-Friedenreich (T), Thomsen-nouveau (Tn) antigens, and their sialylated forms (ST and STn) (Leite-Gomes *et al.*, 2022). In active UC, STn is increased in secreted mucins (Kudelka *et al.*, 2020; Larsson *et al.*, 2011), with high levels associated with an increased risk of neoplasia (Leite-Gomes *et al.*, 2022). In CRC, tumor MUC1 has high levels of Tn, contributing, via lectin-mediated signaling, to the suppression of the immune response of dendritic cells (Saeland *et al.*, 2007). Finally, the STn antigen promotes tumor growth and immune evasion, increasing the

activity of TREG FOXP3+ immunomodulatory cells and affecting the maturation of dendritic cells (Carrascal et al., 2014; Leite-Gomes et al., 2022). In comparison, O-glycans in non-malignant epithelial cells are mainly of the core 3 type, biosynthesized by the enzyme  $\beta$ -1,3-N-acetylglucosaminyltransferase (Core 3 synthase, C3GnT). In murine models of sodium azoxymethane-dextran sulfate-induced CCRAC (AOM-DSS), the elimination of C3GnT produced not only the loss of core-3 O-glycans, but also a reduction in the expression of mucin 2 (Muc2), causing greater intestinal permeability and susceptibility to colitis and tumor development (An et al., 2007).

Sialylation is also affected in IBD and CRC. On the one hand, in IBD, an increase in sialil Lewis A (sLea) antigens is observed in the inflamed mucosa of patients and in models of colitis induced by sodium dextran sulfate (DSS) (Kelm et al., 2020). In addition, in a model of intestinal inflammation by administration of 2,4,5-trinitrobenzylsulfonic acid (TNBS), our laboratory has described an altered and inflammation-dependent glycoma in activated CD8+ T lymphocytes; by decreasing sialylation  $\alpha$ (2,6), these cells (key mediators of inflammation) become more sensitive to the pro-apoptotic effect of Galectin-1 (Morosi et al., 2021). Alterations in sialylation can also negatively affect the gut microbiota, aggravating the disease (Huang et al., 2015). On the other hand, in CRC, sLea is also increased in tumor cells and in the serum of patients with a worse prognosis and metastases (Leite-Gomes et al., 2022). Finally, CRC cells with high levels of sialylation are recognized by immune cells through their interaction with Siglecs (sialic acid-binding immunoglobulin-like lectins); this interaction inhibits antigen-presenting activity and, consequently, the effector functions of CD8+ T cells, promoting immune evasion (Chandrasekar et al., 2023).

N-glycosylation is also affected in CU and CCR, especially by the synthesis of complex branching N-glycans. Among them, branched N-glycans  $\beta$ (1,6) stand out, whose formation is catalyzed by the enzyme  $\beta$ 1,6-N-acetylglucosaminyltransferase-V (GnT-V), encoded by the *MGAT5* gene. In active UC, the reduction of these glycans in CD4+ T cells (a consequence of a decrease in *MGAT5* expression) lowers their activation threshold, generating excessive immune responses (Leite-Gomes et al., 2022). However, supplementation with N-acetylglucosamine (GlcNAc, orally) can restore proper glycosylation and reduce inflammation

(Dias *et al.*, 2018). On the other hand, in the context of CRC, the correlation of complex N-glycans with aggressiveness and poor prognosis is widely documented. Even in the early stages of CRC, *MGAT5* is overexpressed, favoring tumor invasion and angiogenesis (Pérez *et al.*, 2020). The aberrant expression of branched N-glycans  $\beta(1,6)$  stimulates an immunosuppressive tumor microenvironment, so the reversal of this phenotype increases the immune response to the tumor and is proposed as a therapy (Silva *et al.*, 2020).

Fucosylation is another altered modification in IBD and CRC. In patients with Crohn's disease, *FUT2* deficiency (a gene encoding  $\alpha$ -1,2-fucosyltransferase) is associated with the dysbiosis characteristic of the disease (Rausch *et al.*, 2011). On the other hand, *FUT8* (encoding  $\alpha(1,6)$ -fucosyltransferase) is increased in T cells from DSS-induced colitis models and in the inflamed mucosa of patients with IBD; the absence of this enzyme reduces the severity of colitis, by decreasing the responses of TH1 and TH2 cells (Fujii *et al.*, 2016). In particular, CD4+ T cells from *Fut8<sup>-/-</sup>* mice have been shown to exhibit lower production of proinflammatory cytokines, suggesting that fucosylation of TCRs is necessary for T cell signaling and induction of colitis *in vivo* (Fujii *et al.*, 2016). On the contrary, in CRC, *FUT8* is overexpressed and associated with a worse prognosis, making it a potential biomarker for advanced oncological stages (Shan *et al.*, 2019). In addition, mutations in guanosine-dysphosphate-mannose-4,6-dehydratase (GMDS, a key enzyme for fucose synthesis), lead to a loss of fucosylation in CRC tumor cells, allowing evasion of NK-cell-mediated apoptosis (Chandrasekar *et al.*, 2023).

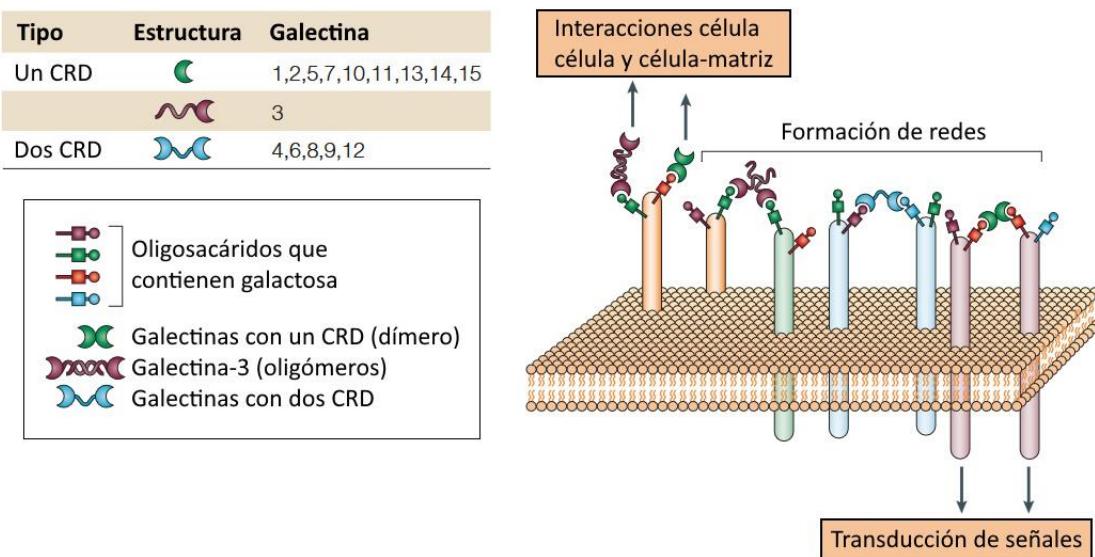
### 1.2.1. Glycogenes

"Glycogenes" are defined as all those genes that participate in the generation and structural modification of glycans (directly or indirectly), and those that encode recognition proteins (lectins), that is, all those that are involved in the glycobiology of a physiological or pathological process.

Glycosyltransferases and glycosidases are enzymes that catalyze the formation or hydrolysis (respectively) of glycosidic bonds in the synthesis of glycoconjugates. The expression profiles of these enzymes give rise to a range of glycans that can be cell-dependent

or tissue-dependent, and their altered expression in IBD pathology, generates changes in the glycooma, which alters cell functionality, intestinal barrier properties, glycan-lectin and host-microbe interactions, contributing to the pathogenesis of IBD; among the most studied glycosyltransferases is *ST3GAL1* (ST3) b-galactoside  $\alpha$ -2,3-σιαλψλτρανσφερασε 1) ρεσπονσιβλε φορ τηε  $\alpha$ (2,3) sialylation of *O-glycans* (Kudelka *et al.*, 2020).

There are several families of proteins with immunological relevance that are responsible for mediating the recognition of cell surface carbohydrates, "translating" the glycooma's signals into biological and immunological processes. These include galectins, a family of lectins with key roles in modulating immune, inflammatory and cell signaling processes, characterized by conserved *carbohydrate recognition domains (CRDs)*, which are responsible for binding to their glycosidic ligands (Johannes *et al.*, 2018). So far, 15 galectins have been identified in mammals. These can be subdivided into three groups: prototypes, which contain a single CRD (galectin-1, -2, -5, -7, -10, -11, -13, -14 and -15) and form homodimers; tandem types, which have two different CRDs in tandem, connected by a connector peptide (galectin-4, -6, -8, -9 and -12); and galectin-3, which is chimera-type, which features unusual tandem repeats of short segments rich in proline and glycine fused to its CRD. Many galectins are bivalent or multivalent with respect to their carbohydrate-binding activity. This modular organization determines their ability to recognize specific glycan patterns and establish cross-linking networks with glycoconjugates in the extracellular matrix or on the cell surface, e.g., with transmembrane proteins, triggering a signaling cascade (**Fig. 6**).



**Figure 6. The family of the galectins.** Many galectins are bivalent or multivalent with respect to their carbohydrate-binding activity: prototype galectins (a single CRD) are dimers; tandem galectins have two covalently bonded CRDs; and finally, galectin-3, which forms oligomers. Although this schematic shows binding to only two glycoproteins, galectins can group multiple multivalent glycoconjugates, forming network-like structures. They can also join two cells of the same type or different types, and connect cells with proteins from the extracellular matrix. For simplicity, the saccharides recognized by the galectins are represented here as disaccharides, although they may be oligosaccharides. Also, the different colors reflect that different galectins recognize different sets of oligosaccharides. Adapted from Liu and Rabinovich, 2005.

Galectins have been studied in their biological and immunological functions in chronic inflammatory diseases (Toscano *et al.*, 2018) and in particular, in IBD. While the role of Galectin-1 as an anti-inflammatory agent is clear (Morosi *et al.*, 2021; Papa-Gobbi *et al.*, 2021), Galectin-4 is one of the least studied lectins with the most discussed role in these diseases, with literature showing that it exacerbates intestinal inflammation (Hokama *et al.*, 2008), and acts as a tumor suppressor in CRC (Satelli *et al.*, 2011).

Other carbohydrate-binding proteins relevant in the intestinal inflammatory process are defensins, small proteins with a positive charge and essential for the host's defense against bacteria, viruses and fungi (Fu *et al.*, 2023). These peptides interact with glycans on microbial surfaces, altering their membranes and inhibiting bacterial growth; in turn, they are able to interact with glycans on immune cell receptors (e.g.,  $\beta$ -Defensin-2 in interaction with TLR-4 in dendritic cells), and regulate cytokine production (Biragyn *et al.*, 2002). Roles of defensins in modulating the tumor microenvironment have also been reported (Sabit *et al.*, 2024), with some

presenting dysregulation in expression such as Defensin Alpha 5 (*DEFA5*), which was described as a molecule with tumor suppressor activity in colon cancer (Qiao *et al.*, 2021).

Meanwhile, selectins, carbohydrate-binding glycoproteins found in endothelial cells (E-selectin), leukocytes (L-selectin), and platelets (P-selectin, also expressed in endothelial cells), participate in the transport of cells of the innate immune system, T lymphocytes, and platelets, in processes of chronic and acute inflammation (Smith & Bertozzi, 2021b; 2021a). The absence of selectins or their ligands has serious consequences, causing recurrent bacterial infections and persistent disease (Ley, 2003).

Notably, and outside the field of immunomodulatory lectins, solute carrier proteins (SLCs) are key in cellular function, as they allow the movement of substances in both directions across membranes, both plasma and organelles. They facilitate the transport of ions, nutrients, neurotransmitters, hormones and nucleotide sugars; its dysfunction or alteration in expression has been associated with a range of pathologies (Schlessinger *et al.*, 2023). In particular, dysregulation of SLC26A2 expression has been demonstrated in IBD (Heneghan *et al.*, 2010), a sulfate transporter, which could affect intestinal mucin sulfation (Corfield, 2018) and metalloprotease activity (Ra *et al.*, 2009), endopeptidases that participate in tissue remodeling and immunity by degrading the extracellular matrix and modulating the activity of cytokines, receptors, and adhesion molecules. As an example, metalloproteinase 9 (*MMP-9*) is elevated in inflamed tissue from patients with IBD and in models of colitis, where it is associated with greater severity of active UC and increased proinflammatory cytokines (Liu *et al.*, 2013).

Finally, chitinases are enzymes that hydrolyze chitin. In this context, CHI3L1 (chitinase-like protein 1 3) is a pseudochitinase that is expressed in a variety of cells including synovial cells, chondrocytes, intestinal epithelial cells, macrophages, and neutrophils. Although CHI3L1 does not have catalytic activity as a hydrolase, it is a lectin capable of binding to ligands such as chitin, heparin, hyaluronan, and collagen (Deutschmann *et al.*, 2019). Its expression is induced in various pathologies including enteritis, pneumonia, asthma, arthritis (Fan *et al.*, 2024), and of particular interest for this Thesis, in epithelial cells and macrophages during intestinal inflammation (Mizoguchi, 2006) has generated interest in recent years in the context of IBD. In particular, it has been described that CHI3L1 regulates bacterial colonization in the

intestinal mucus, through peptidoglycans. This interaction between CHI3L1 and bacteria is beneficial for colonization, particularly for gram-positive bacteria such as *Lactobacillus*. In addition, a deficiency of Chi3l1 leads to an imbalance in the gut microbiota, exacerbating sodium dextran sulfate-induced colitis (DSS) (Chen *et al.*, 2024; Deutschmann *et al.*, 2019). In experimental models of colitis-associated colorectal cancer (AOM-DSS), *knockout* (KO) animals for Chi3l1 develop severe colitis, but with a lower incidence of tumors (Low *et al.*, 2015). The expression of this lectin is increased during inflammation and to a greater extent in the chronic phase, with an essential role in the proliferation of intestinal epithelial cells. The expression of CHI3L1 is increased in several types of tumor, which has led to postulate it as a link between chronic inflammation and cancer (Fan *et al.*, 2024).

### **1.3. Murine models for the study of ulcerative colitis and colitis-associated colorectal cancer**

To study the pathogenesis and possible treatments for IBD, several experimental models have been developed. Possibly, the most commonly used is colitis induced by sodium dextran sulphate (DSS), a negatively charged sulphated polysaccharide with anticoagulant action, which is administered in drinking water. This is an easy-to-implement, reliable, and adaptable system that allows inducing acute, chronic, or recurrent colitis by varying doses and times of DSS administration (Eichele & Kharbanda, 2017). The most widely accepted proposed mechanism of action is direct damage to the intestinal epithelium, followed by the entry of bacteria and other lumen antigens into the underlying tissues, causing an inflammatory response (Yang & Merlin, 2024).

The initial inflammatory response (acute phase) is mainly caused by innate immunity, with activation of macrophages and neutrophils, secretion of cytokines, chemokines, and reactive oxygen species, which activate the complement system. After one day of treatment with DSS, the colon has reduced expression of tight-junction proteins, such as claudins and Zonula occludens-1 (ZO-1), and upregulation of pro-inflammatory cytokines secreted by TH1 cells, such as TNF- $\alpha$ , IFN- $\gamma$ , IL-1 $\beta$ , and IL-12, whose secretion progressively increases with treatment. On the other hand, and since the experimental model cannot recapitulate the cycles of inflammation and remission observed in patients with UC, the induction of chronic colitis

requires the administration of DSS in a cyclical or continuous manner in the long term, and is characterized by cytokines of TH2 cells, such as IL-4, IL-6 and IL-10 (Yang and Merlin, 2024).

The clinical and histopathological signs of the intestinal inflammation model recapitulate several aspects of human IBD, although the degree of agreement varies depending on factors such as the mouse strain, gut microbiota, diet, and the molecular size of the DSS used. Cytokine expression is related to the level of barrier integrity and clinical manifestations, making this model a useful tool for investigating intestinal inflammation. Some of the symptoms caused in the acute inflammation stage include diarrhea, weight loss, blood in the stool, anemia, and eventually death. In the chronic stage, clinical manifestations do not usually reflect the severity of intestinal inflammation (Perše & Cerar, 2012; Yang & Merlin, 2024). Histologically, damage to epithelial cells is observed, with a decrease in mucins, neutrophil infiltrate in the lamina propria, and also, in the chronic stage, disruption of the architecture of the intestinal crypts and deep mucosal lymphocytosis (Yang & Merlin, 2024).

Another advantage of the DSS model of colitis induction is the possibility of coupling inflammation to the use of azoxymethane (a mutagen) to generate the colitis-associated colorectal cancer model (AOM-DSS). Tumor development in this model can occur as early as 10 weeks, and histopathology of induced tumors recapitulates key facets of CCRAC, such as distal tumors and invasive adenocarcinomas (Parang *et al.*, 2016). The application of the AOM-DSS model has been crucial in unraveling the pathogenesis of CCRAC, from the role of signaling pathways (e.g., Toll 4, IKK $\beta$ , and IL-6 receptors) to the influence of the microbiota. Therefore, the AOM-DSS model is an interesting tool for the study of the pathogenesis of colitis-associated colorectal cancer (Parang *et al.*, 2016).

## **2. OBJECTIVES**

### **2.1. General objective**

This study aims to identify glycobiological signatures that contribute to the development of new diagnostic and/or prevention strategies in patients with ulcerative colitis and at risk of colitis-associated colorectal cancer.

### **2.2. Specific objectives**

**A. Develop and validate a glycogene signature based on samples from patients with ulcerative colitis (UC).** Glycobiological signatures of relevance in the malignancy process will be searched through transcriptomics studies in public databases of patients with ulcerative colitis with or without neoplasm. The glycobiologic axis in both conditions (UC with or without neoplasia) will be compared with transcriptomics of healthy subjects (without IBD).

**B. To comparatively analyze the glycogene signature obtained in UC with and without neoplasia in experimental models of colitis and CCRAC.** To this end, the signatures obtained and their validity in databases of experimental models of colitis and colitis-associated colorectal cancer (DSS/AOM-DSS) will be analyzed.

**C. To analyze the role of glycogene signature *in vivo* by RT-qPCR.** Based on the results obtained previously, the dysregulation of target genes in experimental models of colitis and colitis-associated colorectal cancer, already established in the laboratory (DSS/AOM-DSS), will be studied.

### **3. HYPOTHESIS**

Glycogenes (genes directly or indirectly involved in the glycosylation process, genes that code for lectins) could be dysregulated in the process of neoplastic development associated with intestinal inflammation, so their study could allow establishing an expression signature that allows predicting the development of neoplasia in these patients.

## 4. MATERIALS AND METHODS

### 4.1.Bioinformatics analysis

#### 4.1.1. Obtaining data from public cohorts

Information from public databases (Bjerrum *et al.*, 2014; Pekow *et al.*, 2013; Tang *et al.*, 2012; Watanabe *et al.*, 2007), containing the digitized and preprocessed information resulting from *high-throughput* experiments of IBD patients and in the AOM-DSS model, obtained from the *Gene Expression Omnibus* (GEO) (see Tables II and III; Fig. 7).

TABLE II. Public datasets analysed in this final project

Código GSE	Primer autor	Organismo	Plataforma	N	Grupos biológicos	Digital Object Identifier (DOI)
GSE224758	Digby-Bell	<i>Homo sapiens</i>	RNA-seq: Illumina HiSeq 2500 ( <i>Homo sapiens</i> )	22	Sanos (6), CU activa (16)	Sin paper asociado
GSE37283	Pekow	<i>Homo sapiens</i>	Microarray: Affymetrix HT HG-U133+ PM Array Plate	20	Sanos (5), CU quiescente (4), CU con Neoplasia (11)	10.1097/MIB.0b013e3182802bac
GSE47908	Bjerrum	<i>Homo sapiens</i>	Microarray: Affymetrix Human Genome U133 Plus 2.0 Array	60	Sanos (15), CU Izquierda (20), Pancolitis (19), Displasia (6)	10.1097/MIB.0000000000235
GSE3629	Watanabe	<i>Homo sapiens</i>	Microarray: Affymetrix Human Genome U133 Plus 2.0 Array	115	Pancolitis (43), CU con CCRAC (10)	10.1158/1078-0432.CCR-06-0753
GSE31106	Tang	<i>Mus musculus</i>	Microarray: Affymetrix Mouse Genome 430 2.0 Array	18	Control (3), y modelos AOM/DSS para: Inflamado (3), Displasia (9), Adenocarcinoma (3)	10.1093/carcin/bgs183

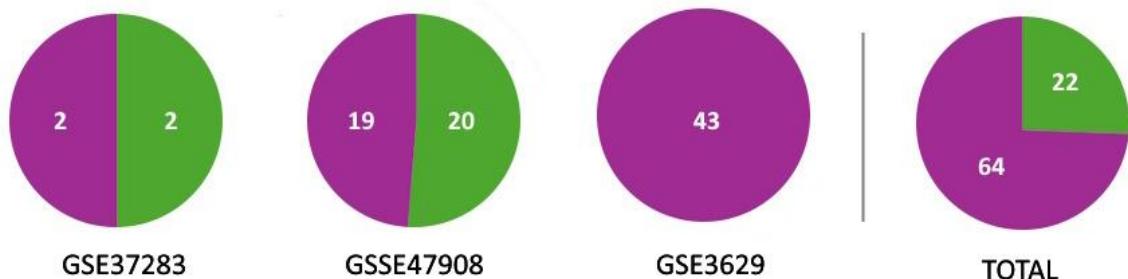
**UADE** SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía

**TABLE III. Summary Table of Metadata of the patients analyzed.** MONTH = Mayo endoscopic score, PSC = primary sclerosing cholangitis, UC = ulcerative colitis, CUq = quiescent ulcerative colitis, CRC = colorectal cancer (associated with colitis), LGD = low-grade dysplasia, HGD = high-grade dysplasia, Neo = Neoplasia.

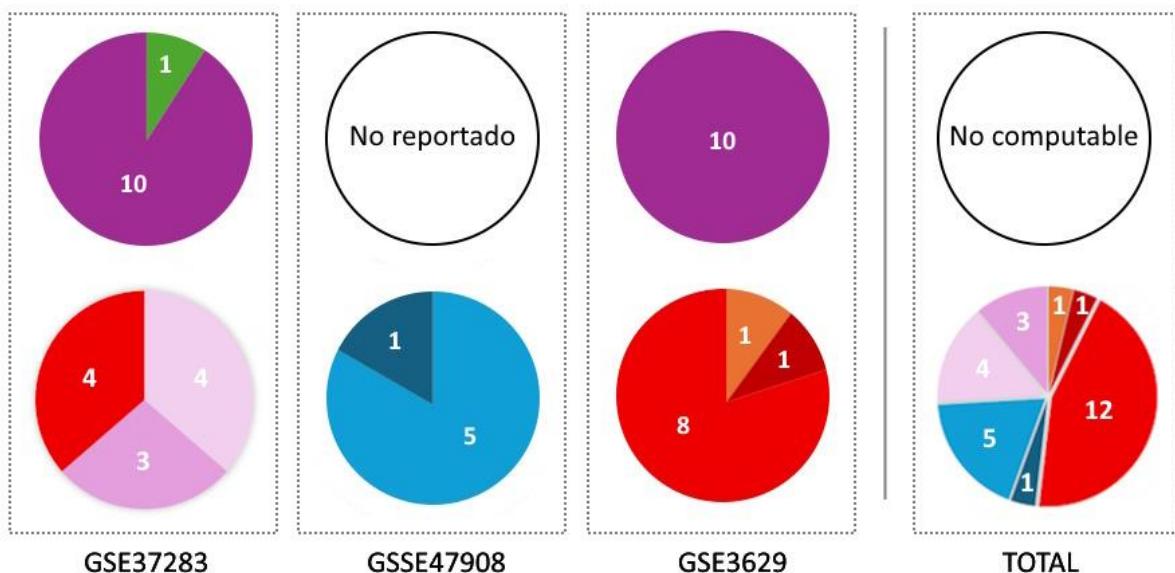
Left UC corresponds to the E2 classification, and pancolitis to E3.

Características		GSE37283 (Pekow)			GSE47908 (Bjerrum)			GSE3629 (Watanabe)	
		Control	CUq	CU+Neo	Control	CU	Displasia	CU	CU+CCR
Pacientes	N	5	4	11	15	39	6	43	10
	Sexo (M/F)	5/0	4/0	11/0	4/11	16/23	4/2	27/16	7/3
	Edad promedio	46	48.3	46.9	52	37	49	47.1	51.9
	Duración de la enfermedad (años)		20, todos > 7	13.5, todos > 7		26 < 10, 13 > 10	2 < 10, 4 > 10	14, todos > 7	16, todos > 7
	Edad de diagnóstico		-	-		-	-	32.6	35.9
	Fuma/no fuma	-	-	-	2/13	3/36	0/6	-	-
Clasificación de la enfermedad	CU Izquierda		2	1		20	-	0	0
	Pancolitis		2	10		19	-	43	10
	LGD			4			5		1
	HGD			3			1		1
	Adenocarcinoma			4			0		8
Localización de lesiones	Colon ascendente			2			0	0	-
	Colon transversal			3			0	0	-
	Colon descendiente			1			0	0	-
	Colon sigmoideo			3			0	0	1
	Recto			2			0	0	-
	Colon sigmoideo y transversal			0			0	0	1
Grado de inflamación	Ninguna		-	-		-	-	11	3
	Leve		-	-		-	-	20	4
	Moderada		-	-		-	-	11	3
	Severa		-	-		-	-	1	0
Otras características clínicas	Mayo score		-	-		6	1.5	-	-
	MES		-	-		2	0.5	-	-
	CEP		-	-		-		0	0
Medicación	Mesalamina		4	8		40	2	38	8
	Corticosteroides		-	-		-	-	5	1
	Inmunomoduladores		2	3		-	-	-	-
	Anti-TNF		0	1		-	-	-	-
	Esteroides sistémicos		-	-		9	0	-	-
	Esteroides tópicos		-	-		4	0	-	-
	Tiopurinas		-	-		13	2	-	-
	Infliximab		-	-		6	0	-	-
	Ninguna		-	-		2	2	-	-

**Pacientes con EII únicamente**



**Pacientes con EII + displasia o cáncer**



**Referencias**

- |                     |       |            |                    |
|---------------------|-------|------------|--------------------|
| ● Pancolitis (E3)   | ● LGD | ● CU + LGD | ● CU + CCRAC + LGD |
| ● CU Izquierda (E2) | ● HGD | ● CU + HGD | ● CU + CCRAC + HGD |
|                     |       |            | ● CU + CCRAC       |

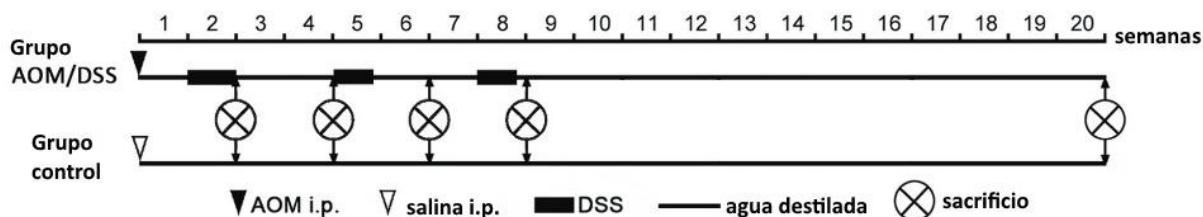
**Figure 7.** Venn diagrams describing the composition of samples from patients with IBD, with or without associated dysplasia or neoplasia. The diagrams within the dotted rectangles correspond to the same patient groups. CCRAC = colitis-associated colorectal cancer, LGD = low-grade dysplasia, HGD = high-grade dysplasia, UC = ulcerative colitis.

Regarding the origin of the samples, it should be noted that those of patients with neoplasia were taken, in all cases, from non-neoplastic mucosa (Pekow *et al.*, 2013; Watanabe *et al.*, 2007). The authors looked for what is known as the "field effect," i.e., elucidating potential mechanisms of carcinogenesis in areas without active inflammation but with remote dysplasia, looking for dysregulations that can serve as more specific biomarkers of dysplasia in those at higher risk (i.e., those who have already had dysplasia). In contrast, in the experimental model, samples were taken from the same tumor masses (Tang *et al.*, 2012). The details of the samples according to each dataset are as follows:

- GSE224758: Colonic biopsies were obtained endoscopically from the rectosigmoid area of 6 healthy donors without mucosal inflammation, both histologically and endoscopically, and 16 patients with UC with evidence of active disease.
- GSE37283: The samples are of non-dysplastic mucosa from patients with ulcerative colitis with remote neoplastic lesions, and no biopsy presented a neoplastic lesion. Patients with a previous clinical diagnosis of UC confirmed by an expert gastrointestinal pathologist, a disease duration > 7 years, and disease extension > 20 cm proximal to the anal margin were included. On histological review, all patients had IBD with quiescent disease (no active inflammation) or minimal inflammatory activity in the rectosigmoid colon. 4 mucosal biopsies were obtained from each endoscopy, 10 to 20 cm proximal to the anal margin. The samples obtained in colectomies were by biopsies of the tissue of the rectosigmoid colon immediately after colon removal. Of the patients with neoplastic lesions, 4 had low-grade dysplasia (LGD), 3 had high-grade dysplasia (HGD), and 4 had adenocarcinomas. However, researchers do not note in the metadata which condition each tissue corresponds to, so they are noted and analyzed under the name of "neoplasm." It was reported that 2 neoplastic lesions were found in the ascending colon, 3 in the transverse colon, 1 in the descending colon, 3 in the sigmoid colon, and 2 in the rectum, although it is not specified which is the case of each patient. This data group is the only one that reports the ethnicity of the participants: the healthy controls are 2 African American and 3 Caucasian, the patients with UC are 3 Caucasian and 1 unknown, the patients with neoplastic lesions are 1 African American, 6 Caucasian, 1 Asian and 3 unknown (Pekow *et al.*, 2013).

- GSE47908: The samples are mucosal colonic biopsies (10-20 mg each) that were obtained endoscopically from the descending colon and come from an area without ulceration, but generally inflamed (with the exception of the group without IBD and three of six samples from patients with dysplasia). All dysplastic specimens originated on the left side of the colon and presented with low-grade dysplasia, except one, with high-grade dysplasia (Bjerrum *et al.*, 2014).
- GSE3629: Specimens were collected from specimens surgically resected or during colonoscopy and always correspond to non-neoplastic rectal mucosa; in fact, they were only included when microscopic examination verified the absence of neoplastic cells. All lesions were located in the large intestine and no cases showed metastatic disease in the cancer cases. All patients with pancolitis had it for more than 7 years. Among the groups analyzed, there were no significant differences in age at biopsy, age at diagnosis of UC, sex, duration of disease, extent of UC, presence of anti-inflammatory medication, presence of primary sclerosing cholangitis, and disease activity. The database includes patients with sporadic colorectal cancer, without their metadata, which the researchers do not incorporate into their published analyses. Of the patients with neoplastic lesions, one had HGD in the sigmoid colon, and another LGD in the sigmoid and transverse colon (Watanabe *et al.*, 2007).
- GSE31106: The experimental design (**Fig. 8**) consists of 6-week-old male ICR mice, with a single intraperitoneal injection of AOM (10 mg/kg) on day 1, followed by three cycles of DSS administration (cycle 1: 2%, days 8~14; cycle 2: 1.5%, days 29~33; cycle 3: 1.5%, days 50~54) in drinking water. Instead, the control group drank distilled water and had a single intraperitoneal injection of saline (10 mg/kg) on day 1. Colorectal tissues were collected on days 14, 28, 42, 56, and 140 (at the end of the 2nd, 4th, 6th, 8th, and 20th weeks) from the AOM-DSS group and on day 14 from the control group. The pathological changes of each sample were identified under the microscope: the samples of the 2nd week corresponded to inflamed mucous membranes; from the 4th week, to low-grade dysplasias; from the 6th, to high-grade dysplasias; from the 8th, to high-grade dysplasias with active inflammation; from the 20th, to carcinomas; and the control sample, to normal mucous membranes. For the collection of samples, the large

intestine was removed, opened longitudinally along the main axis and washed with saline. The samples taken in the 2nd week corresponded to acute inflammation. The masses, the inflamed and normal colonic mucosa, were cut and placed in 5 volumes of serum (Tang *et al.*, 2012). Adenocarcinomas were diagnosed based on the histological criteria of Boivin *et al.* (2003): Adenocarcinoma is defined as a malignant neoplasm of the glandular epithelium composed of tubular and/or villous structures that penetrate through the mucosal muscle lamina, and is categorized according to its degree of differentiation (well-differentiated, moderately differentiated, or poorly differentiated), and its histological type: (a) tubular/tubulovillous/villous adenocarcinoma; (b) mucinous adenocarcinoma, with 50% of the tumor composed of extracellular mucin, and signet ring cells may be present; (c) signet ring adenocarcinoma, with 50% of the tumor composed of signet ring cells; and (d) undifferentiated, uniform, or pleomorphic carcinoma (Boivin *et al.*, 2003).



**Figure 8. Experimental design for AOM-DSS model sample collection.** The samples collected at the end of the second week corresponded to inflamed mucous membranes; the fourth week, to low-grade dysplasias; the sixth, to high-grade dysplasias; the eighth, to high-grade dysplasias with active inflammation; week 20, carcinomas; and the control sample, to normal mucous membranes. AOM i.p. = Intraperitoneal AOM. Adapted from Tang *et al.*, 2012.

In addition, results of two differential expression analyses of UC with respect to healthy patients (without IBD) were incorporated, namely: (1) 463 differentially expressed genes (*p*-adjusted value < 0.05,  $\log FC \geq 1.5$ ) from the bioinformatics dissertation of Román Gabriel Lanzillotta (LU1049699) (Lanzillotta, 2022), which were obtained from the *microarray data* of the GSE73661 cohort (Arijs *et al.*, 2018), and (2) 65 differentially expressed genes from Zhang's analysis (Zhang *et al.*, 2023), which were obtained from RNA-seq experiments of three integrated cohorts: GSE87473, GSE92415, and GSE206285 (see Table IV).

TABLE IV. Cohorts analyzed by (Zhang *et al.*, 2023)

**UADE** SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía

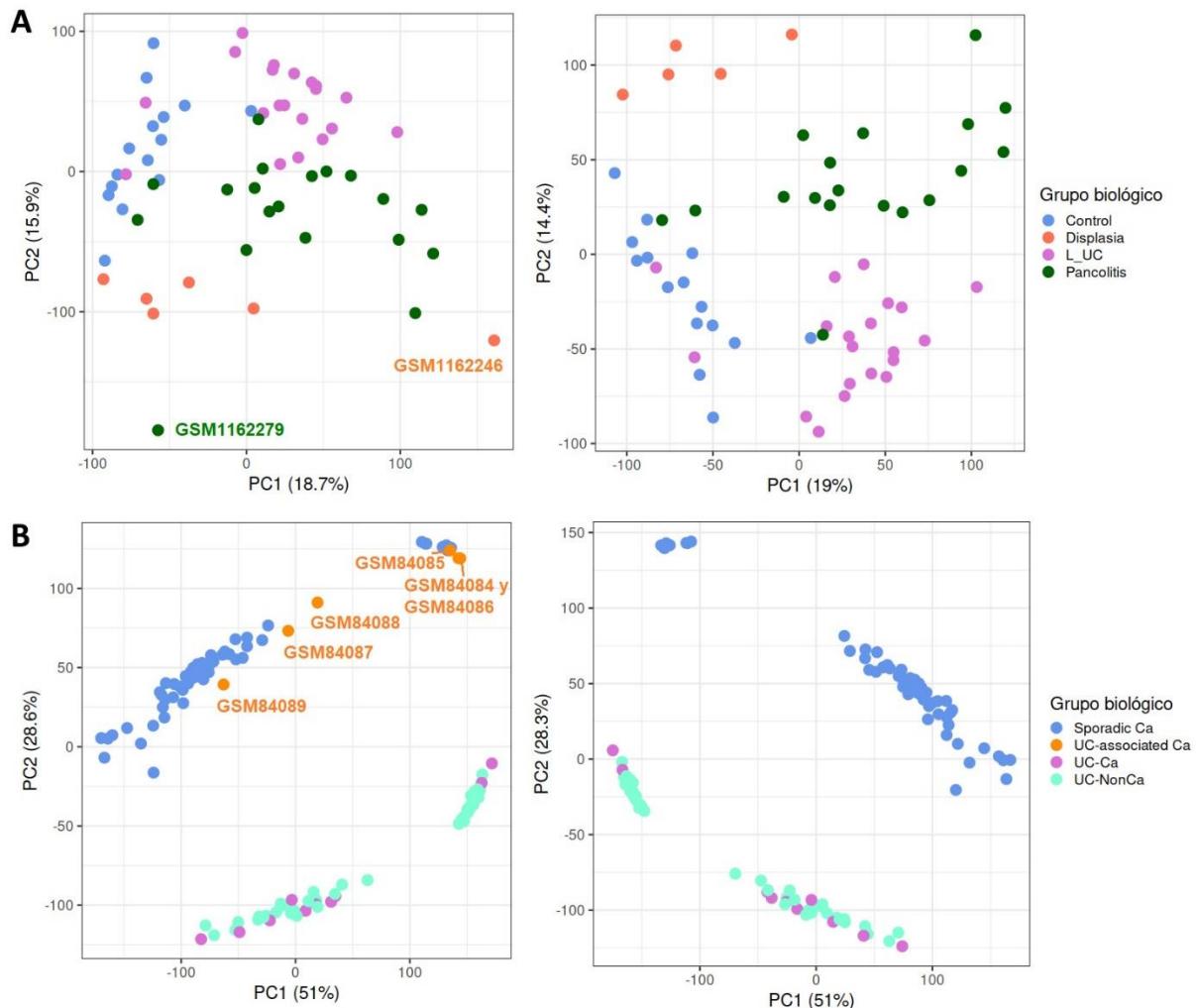
<i>Código de acceso</i>	<i>Muestras</i>	<i>Pacientes sanos</i>	<i>Colitis Ulcerosa</i>
<b>GSE87473</b>	127	21	106
<b>GSE92415</b>	108	21	87
<b>GSE206285</b>	568	18	550

#### 4.1.2. Quality control and preprocessing of public cohort data

For the microarray datasets, pre-QA and normalization was applied using *limma package functions* (Ritchie *et al.*, 2015). In cases where you worked directly with crude fluorescence files (.CEL), the expression signals were normalized using the *rma* function of the *affy packet* (Gautier *et al.*, 2004), which performs background preprocessing, probe effect adjustment, and quantile normalization, returning the expression values in *log2 scale*. In datasets that have already been processed or when the CEL files were not available, the *normalizeBetweenArrays* function of *limma* was used to ensure comparability between samples. Additionally, when relevant according to the experimental design, low-expression genes were filtered using a minimum threshold: only those genes with expression values greater than 1 were conserved in all samples of each phenotype analyzed.

For the RNA-seq dataset, we worked from a normalized array of beads using the *TMM (Trimmed Mean of M-values)* method with the *calcNormFactors* function of the *edgeR package* (Chen *et al.*, 2025). This normalization adjusts for differences in library size between samples. Subsequently, low-expression genes were filtered out, preserving only those with *cpm (counts per million)* > 2 in all samples from each group. In addition, the *voom* function of the *limma package was used* to adjust the mean-variance relationship in the genes, reducing the differences between the highly and the least expressed, which contributes to improving the robustness and reliability of the statistical analysis.

The principal component analysis was performed with the *prcomp* function of the *stats package* (R Core Team, 2025), and was used to evaluate the consistency between the transcriptomic profiles of the samples. After this analysis, we decided to eliminate the *outlier* samples: GSM1162246 (cohort GSE47908); and GSM84084, GSM84085, GSM84086, GSM84087, GSM84088, and GSM84089 (cohort GSE3629), which in our analysis showed an expression profile more similar to sporadic cancer than to ulcerative colitis (**Fig. 9**).



**Figure 9. Principal Component Analysis.** A) Pre-Sample Disposal (L) and Post-GSE47908 PCA (R) Outlier (GSM1162279). B) PCA of GSE3629 before (left) and after removal of samples that showed an expression profile more similar to sporadic cancer than to ulcerative colitis.

#### 4.1.3. Functional analysis of transcriptomic data

The count of the analyzed readings was carried out to identify differentially expressed genes (SDRs) using the *limma* and *edgeR* packages, as appropriate to the experimental design. Lists of significant genes were generated with the result of the *toptable* function of the *limma package*, with its p-value adjusted by FDR according to the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995), and considering a cut-off value of p-value = 0.05.

The totally or differentially expressed genes, as appropriate, were subjected to gene cluster enrichment analyses to identify dysfunctional biological, molecular, and cellular

component processes. A *Gene Set Enrichment Analysis (GSEA)* was performed using the *GSEABase* package using the *GSEA function of the homonymous package* (Subramanian *et al.*, 2005) and for each cohort, a *Gene Set Variation (GSVA)* was performed with the *GSVA function* of the GSVA package (Hänelmann *et al.*, 2013), to enrich gene sets in individual samples by finding subgroups of patients. For both the *GSEA* and GSVA methods, the p-value was adjusted by FDR according to the Benjamini-Hochberg (1995) procedure, and considering a cut-off value of p-value = 0.05. To filter out terms of higher relevance, *GSEA* results filter terms containing between 5 and 500 genes, and for *GSVA results*, terms less than 500 genes. The following *Molecular Signatures Database (MSigDB) databases* (Liberzon *et al.*, 2011):

- From the *C2 collection (Curated Gene sets)*, the *CP Reactome subset* (Canonical Pathway Gene Sets Derived from the Reactome Pathways Database), and
- From the *C5 collection (Ontology Gene sets)*, the *subset* derived from the *Gene Ontology (GO)* resource that contains *BP*, *CC*, and *MF components*.

To analyze overrepresentation of subsets of genes (ORA), *the gost function of the gprofiler2 package was used*, which allows the corresponding hypergeometric analysis to determine which sets of genes defined a priori are more present (overrepresented) in a subset of genes of interest than would be expected by chance (Huang *et al.*, 2009). As input genes, those from relevant *clusters* that emerged from the integration of the different cohorts were used, and they were with known sources of functional information to detect significantly enriched terms. Four databases of transcriptional programs were used: *GO: MF*, *GO: BP*, *GO: CC* and *Reactome*.

Other graphs to visualize results and develop bioinformatics tools were made in *R*, with functions of the *ggplot2 package* (Wickham, 2016). The *Venn heatmaps* and diagrams were made with the *pheatmap* (Kolde, 2025) and *ggvenn* (Yan, 2021) functions, of the homonymous packages, respectively. In turn, the *upset* function of the *UpSetR package* was used (Conway *et al.*, 2017) to efficiently represent multiple sets of genes.

The so-called "glycogenes" in the present research are 965 human genes and their corresponding murine orthologs. The starting human genes are 877 manually cured and previously collected glycogenes, 6 alpha defensins and 82 beta defensins (**Annex 8.1**). The list

of 877 genes was selected using different sources, including the Affymetrix glycoV4 chip (Consortium of Functional Glycomics, Core E, GEO:GPL11097) and GlycoMaple (Huang *et al.*, 2021). From here, genes involved in glycosylation were obtained, such as glycosyltransferases and glycosidases, enzymes involved in the metabolism of nucleotides and sugars, transporters and genes relevant to the function of the endoplasmic reticulum and the Golgi apparatus. Genes that were described as relevant in the glycosylation of intestinal epithelial cells were also included (Kudelka *et al.*, 2020) and glycobiological genes related to the activation of the transcription factor XBP1, which is key in intestinal inflammation and can remodel the structural patterns of N-glycans (Dewal *et al.*, 2015; Wong *et al.*, 2018).

For the conversion of genes from humans to their murine orthologs, the *Ensembl* database was used through the *biomaRt package* of R (Durinck *et al.*, 2009). The connection to the *mmusculus\_gene\_ensembl* dataset and the *with\_hsapiens\_homolog filter*, which restricts the search to genes with known human counterparts, were established. When multiple possible human orthologues exist for the mouse gene (e.g., *Defa5*), a several-to-one binding is performed, resulting in several rows for each matching pair.

#### **4.1.4. Development of bioinformatics tools**

The development of an interactive web application was done using the *Shiny ()*, a free and open-source tool to create interactive figures that can be published online directly from R. The online publication was carried out through the Shiny Apps platform (Sievert, 2025), which provides compatible servers. Genetic information is presented directly from the National *Center for Biotechnology Information* (NCBI) website (National Center for Biotechnology Information, 2025).

### **4.2. Transcriptomics and/or RT-qPCR studies in experimental models**

#### **4.2.1. Murine models**

RNA from 8-12 week old C57Bl/6 mice from 3 experimental groups was analyzed:

1. *Control mice*: 3 samples from the intestine of females, without treatment, frozen in Trizol in September 2023, were used. These animals had the region of the Galectin-4

gene flanked with a Loxp region, but their phenotype was wild (*VillinCre<sup>wt</sup>/wtLgals4<sup>flox/phlox</sup>*).

2. *Mice with acute colitis induced by dextran sodium sulfate (DSS)*: 3 samples from the intestine of female *VillinCre<sup>wt</sup>/wtLgals4<sup>flox/phlox</sup>* treated with DSS 3% were used in drinking water for 5 days and then 5 days in regular water. The total duration of the experiment was 10 days. The intestine samples, inflamed or not, were frozen in Trizol in February 2023.
3. *Mice with colitis-associated colorectal cancer (azoxymethane/DSS)*: 3 samples of female mice injected intraperitoneally with azoxymethane (AOM, 12.5 mg/kg) and administered 5 days later with DSS 2.5% in drinking water 2.5% for 5 days, followed by 14 days of rest with regular water. This cycle is repeated three more times, for a total duration of 12 weeks. The tumor samples obtained at the end point were frozen in Trizol in January 2024.

#### **4.2.2. RT-qPCR**

RNA was extracted from cell suspensions of murine gut samples in TRIzol™ following the manufacturer's procedures (Invitrogen). Briefly, chloroform (Sigma-Aldrich) was used to separate RNA in the aqueous phase. RNA was then precipitated using isopropanol (Sigma-Aldrich) and washed with 75% V/V ethanol (Sigma-Aldrich). The ethanol was then allowed to evaporate at room temperature and the RNA was resuspended in ultrapure RNase-free water and quantified using a NanoDrop (Thermo) device by spectrophotometry. Concentration was analyzed and purity measured according to the 260/280 and 260/230 absorbance ratios. All disposable materials used were RNase-free or sterilized to avoid any type of contamination or degradation, while non-disposable items were exposed to RNaseZap (Invitrogen). Complementary DNA (cDNA) was obtained by reverse transcription using 1 to 2 µg of RNA per sample. Briefly, the RNA was treated with DNase I (Thermo Scientific) to remove contaminating DNA for 15 minutes at room temperature, cutting off the reaction by adding 1 µL of 25mM EDTA and incubating at 65°C for 15 minutes. Then, reverse transcription was carried out using *random primers* and SuperScript IV (Thermo Scientific) reverse transcriptase.

For qPCR, the commercial SsoAdvanced Universal SYBR Green Supermix (Bio-Rad) kit was used, which already contains the Taq enzyme, the SYBR Green dye, dNTPs, and all the necessary cofactors to carry out the PCR. Briefly, the reaction mixture containing the commercial mixture and the messenger-specific *primers* pairs of the desired gene, which were previously designed in the laboratory (**see Table V**), was prepared. The total reaction volume was 20 $\mu$ L, including 1 $\mu$ L of cDNA and 19 $\mu$ L of SYBR mixture and the *specific primers*. For each pair of *primers*, a standard curve was used using a mixture of the cDNAs of all the samples. The curve consisted of 5 points (2 $\mu$ L, 1 $\mu$ L, 0.5 $\mu$ L, 0.1 $\mu$ L and 0.01 $\mu$ L of mixture, and a negative control without tempering) and allowed quantifying the relative expression of each gene in the samples. Each sample was run in duplicate. It was performed on a CFX96 Real Time (Bio-Rad) thermal cycler using CFX Manager software. Amplification began with an activation phase at 95°C for 30 seconds, followed by denaturation at 95°C for 10 seconds. Then, the alignment and extension stage was at 60°C for 30 seconds. The denaturation and alignment/extension steps were repeated in a total of 38 cycles. Finally, a *Melt Curve Analysis* is performed with a temperature range between 65°C and 95°C, increasing in intervals of 0.5°C every 5 seconds. The data were analyzed using the CFX Maestro software, verifying that the efficiency of each pair of *primers* was between 90%-110%. The relative expression of each gene was calculated within the software using the standard curve generated in each case. The data from each sample were normalized to the geometric mean of expression of the control genes.

**TABLE V. Primers used in qPCR.**

Gen	Nombre	Sentido	Secuencia
<i>B2m</i>	Beta-2-Microglobulina	FWD	ACCGTCTACTGGGATCGAGA
		REV	TGCTATTCTTCTGCGTCAT
<i>Defa5</i>	Defensina Alfa 5	FWD	TTGTCCTCCTCTGCCCTTG
		REV	ATGAAGAGCAGACCCTTGTG G
<i>Lgals4</i>	Galectina-4	FWD	ATGGTCACCCATCTGCAAGT
		REV	AAGCTGGAATAGTCATGGCTCC
<i>Lgals12</i>	Galectina-12	FWD	AACTGACGAACACCTGGACC
		REV	TGCCATAAGGAATCACCGGG
<i>St3gal1</i>	ST3 beta-galactósido alfa-2,3-sialiltransferasa 1	FWD	TCCTACAAC TGCACACGCTCG
		REV	TGTTTCGCCTGGTGCCTGG

#### 4.2.3. Statistical analysis

All data analyses were performed using R v4.0.2 via RStudio. For differential expression analyses, the cut-off point for statistical significance was set at  $p < 0.05$ . qPCR results were analyzed using the Kruskal-Wallis test with the *kruskal.test* program of the *stats package* (R Core Team, 2025), followed by the Dunn multiple comparisons test with the *dunn\_test* program of the *rstatix* package (Kassambara, 2023), with FDR adjustment using the Benjamini-Hochberg method (1995).

## 5. RESULTS

### 5.1. Work scheme

The present research is divided into several stages, which include the development of bioinformatics tools, *in silico* transcriptomic analysis of patient samples and experimental models, and *in vivo* validation experiments using RT-qPCR. Transcriptomic data were obtained from public databases of patients with ulcerative colitis (UC), dysplasia, and colitis-associated colorectal cancer (CCRAC), and from experimental models of these pathologies, induced by DSS or AOM-DSS.

First, the bioinformatics tools that laid the foundations of the *pipeline* of this research were developed. They include an *online* application for visualizing differential expression results, and an *R* package to facilitate transcriptomic analysis, allowing a bioinformatic approach to a wider community.

Second, differential expression analyses of patient samples were performed. The first group compared were patients with ulcerative colitis compared to healthy subjects. This analysis was divided into two parts: the first, an exploration of all cohorts with some type of ulcerative colitis (active or in remission, also called quiescent) and healthy subjects to compare (GSE224758, GSE37283, GSE73661, GSE87473, GSE92415 and GSE206285). The second part focuses on explaining the characteristics of quiescent ulcerative colitis (GSE37283), to mark it as an exception and describe its unique characteristics, which distinguish this condition from active inflammation and tissue that never suffered from chronic inflammation.

Third, the differences between patients with dysplasia and active IBD were explored. This IBD was considered in two of its forms: left ulcerative colitis (E2) and pancolitis (E3). This section is analyzed based on a single dataset, the GSE47908.

Fourth, patients with colitis-associated colorectal cancer (CCRAC) were analyzed against patients with active left ulcerative colitis. Because the available data did not contain such a comparison explicitly, two different cohorts were used to arrive at the results indirectly: the GSE3629 cohort, which includes patient samples with CCRAC and pancolitis (E3), and the GSE47908 cohort, which contains samples of pancolitis (E3) and left UC (E2). This

combination allowed to indirectly investigate the molecular transition between left UC and CCRAC, comparing it with the pancolitis-neoplasia transition.

Fifth, the results of a murine model from the GSE31106 dataset were explored, specifically referring to the comparison of adenocarcinoma versus inflammatory tissue, both induced with AOM-DSS. It is at this point where, taking into account the results in patients, the murine glycobiological signature was determined for the transition from ulcerative colitis to colitis-associated colorectal cancer. This signature was determined using a statistical cut-off point and relative to the times of change between conditions. An "extended glycobiological signature" (**Annex 8.3**) was also described, to consider relevant genes that did not reach the level of astringency of this analysis for future research.

Finally, the results were validated *in silico* by RT-qPCR. Four target genes were selected taking into account relevance to the objectives of this Thesis and other lines of research within the laboratory.

## **5.2. Development of an interactive application and an *R* package for differential expression analysis**

A web application for the analysis of differential expression data was developed and published, which can be accessed through the following hyperlink: <https://lcastelli.shinyapps.io/degs/>. In it, the result of a differential expression analysis can be viewed interactively, on a desktop computer or a mobile device, by uploading a table in .csv format that has the following columns, respecting the names exactly: "gene" (with the name of each gene), "logFC" (with the logarithm of "fold change"), "p" (with the p-value adjusted), and "sense" (with values "up" or "down", depending on whether the times of change are positive or negative, respectively). There should be no special character (e.g. "á", "í", "!" ...). In the case of not uploading a custom file, the site has its own example, the same being the result of the "extended glycobiological signature" resulting from this research project (**Annex 8.3**).

The tabulated numerical values are shown with 3 significant figures in the first tab of the application, "Table" (**Fig. 10 A**) and can be filtered interactively by selecting a p-value and a logFC reactively, although the defaults are p-value = 0.05, and the absolute value of logFC =

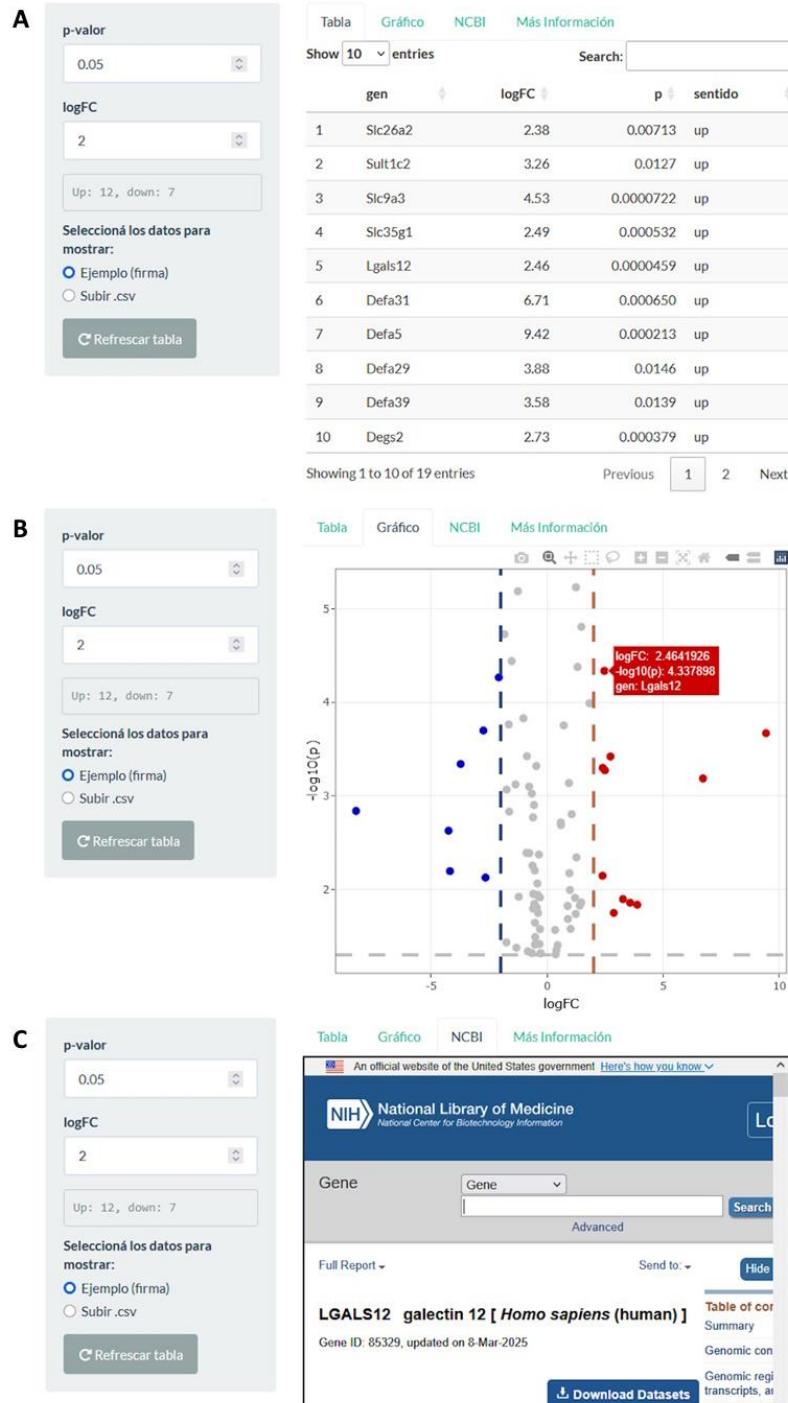
2. Depending on the values chosen by the user, not only do the eligible genes are listed, but the number of positively and negatively regulated genes is also reported in a text box in the sidebar of the website (in the format "*Up*: 12, *down*: 7" if the number of genes preserved after cutting is 12 positively regulated and 7 negatively regulated).

In the second tab, "Graph" (**Fig. 10 B**), a *volcano plot* is generated with the data preloaded or uploaded by the user that illustrates the differential expression and significance of each gene, with the *logFC* on the x-axis and the *-log10(p-value)* on the y-axis. The reference lines are an intersection on each axis for the same p-value (horizontal line, intersecting the y-axis) and *logFC* absolute value (two vertical lines, intersecting the x-axis) that filter the table from the first tab. They are modifiable and are updated every time the user chooses new values. In addition, by keeping the cursor suspended at a point, it is indicated with a label which gene it is and the corresponding value for each of the axes. When you *click* on a point, that gene is selected, which has two effects:

1. Filter the table on the first tab so that it shows the details of the corresponding gene. This can be undone by pressing the "refresh table" button on the side toolbar.
2. Search for the corresponding gene on the NCBI (<https://www.ncbi.nlm.nih.gov/>) website (National Center for Biotechnology Information, 2025), to display the detailed gene information on the third tab, "NCBI" (**Fig. 10 C**).

It is possible to undo the selection on the chart by double-clicking on any blank area of the chart. The graphic can be downloaded as a .png file by the user by pressing the camera button (📷) that becomes visible when the cursor is passed over it.

# UADE SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOCOBIOLICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía



**Figure 10.** App screenshots *Shiny* in operation. Images of the interactive application developed and published on the servers of *Shiny Apps*, showing preloaded data as an example. These data correspond to the "extended glycobiological signature" (Annex 8.3) obtained in this dissertation. The application interface is presented in different visualizations: **A)** data table, **B)** Volcano plot y **C)** information from NCBI. Source: <https://lcastelli.shinyapps.io/degs/>.

Second, an R package called *degfind* (**Annex 8.5**) was developed, for transcriptomic analysis of microarray and RNAseq expression data. It allowed to lay the foundations for the automation and optimization of multiple processes within the transcriptomic analysis of research, and is compatible to be built with the *devtools* build command from the command line, indicating the path to the files, published on *GitHub*. They contain a test .csv file in *Rd (R documentation)* format. This function manual (accessible via *degfind/man/* or using "?" followed by the name of the corresponding function as a command in *Rstudio*) has specific information on how to use each of the eight functions developed, which are as follows:

1. ***DEG\_table***: Generates a table of differentially expressed genes.
2. ***SDRs***: Processes expression data to extract differentially expressed gene names, and generates a Venn diagram with which they are positively and negatively regulated.
3. ***gene\_names***: Processes expression data to extract gene names.
4. ***make\_lm***: Processes expression data to generate a linear model and return results, as well as Bayesian statistics.
5. ***make\_toptable***: Processes Bayesian statistics from a linear model to generate a data framework such as *limma* main tables, plus a classification that considers low and high expression genes.
6. ***plot\_data\_quality***: Processes expression data to generate and save a violin graph that describes its quality.
7. ***sample\_names***: Processes expression data to extract sample names.
8. ***volcano***: Saves the *volcano plot* with differential expression data, where the *x-axis* is the *logFC* and the *y-axis* is the *p-set value*.

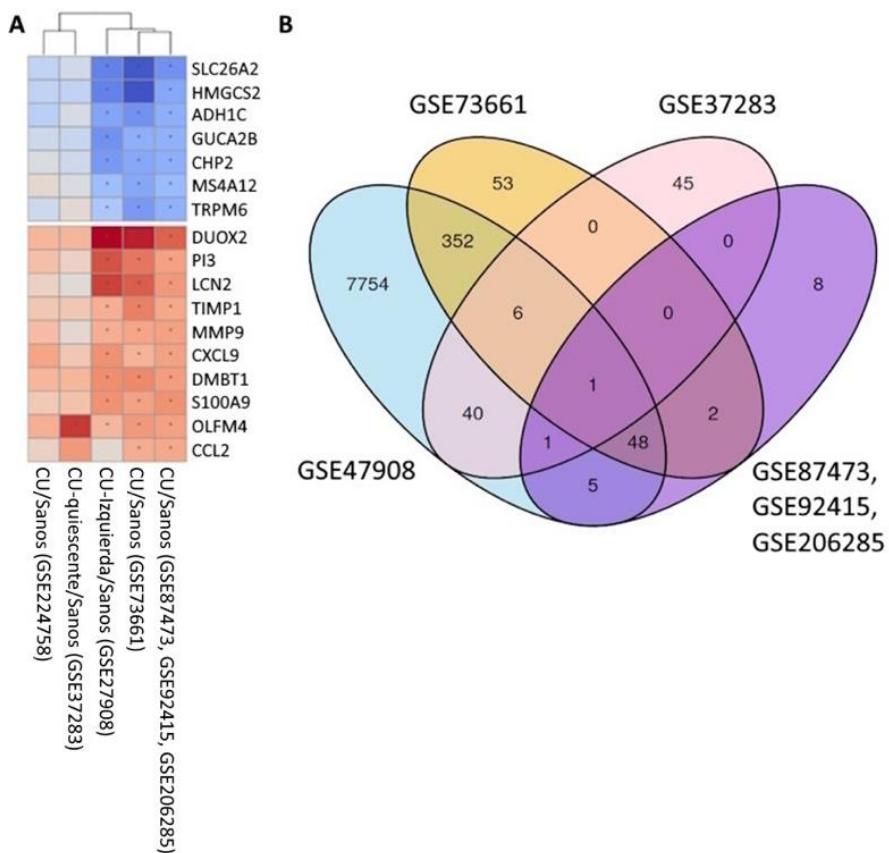
A sample of the use of each function can be found on YouTube through the following hyperlink: <https://youtu.be/7Vcm609bAac?si=cBIhvW-Qagzl2FEQ>.

### **5.3. Analysis of transcriptional profiles in public databases**

#### **5.3.1. Differential gene expression between patients with ulcerative colitis and healthy subjects**

##### ***5.3.1.1. Comprehensive cross-cohort analysis***

First, taking the results from the GSE73661 (Lanzillota, 2022; Arijs *et al.*, 2018) previously obtained, analyzing data from the following cohorts: GSE224758 (Digby-Bell, no associated paper), GSE37283 (Pekow *et al.*, 2013), GSE47908 (Bjerrum *et al.*, 2014), and from the published differential analysis for the combination of cohorts GSE87473, GSE92415, and GSE206285 (Zhang *et al.*, 2023) a profile of genes that are consistently dysregulated was established by comparing biopsies of patients with active and quiescent UC vs. healthy subjects (without IBD) throughout all cohorts. The results show high agreement between cohorts resulting in 7 decreased and 10 increased genes (**Fig. 11**).



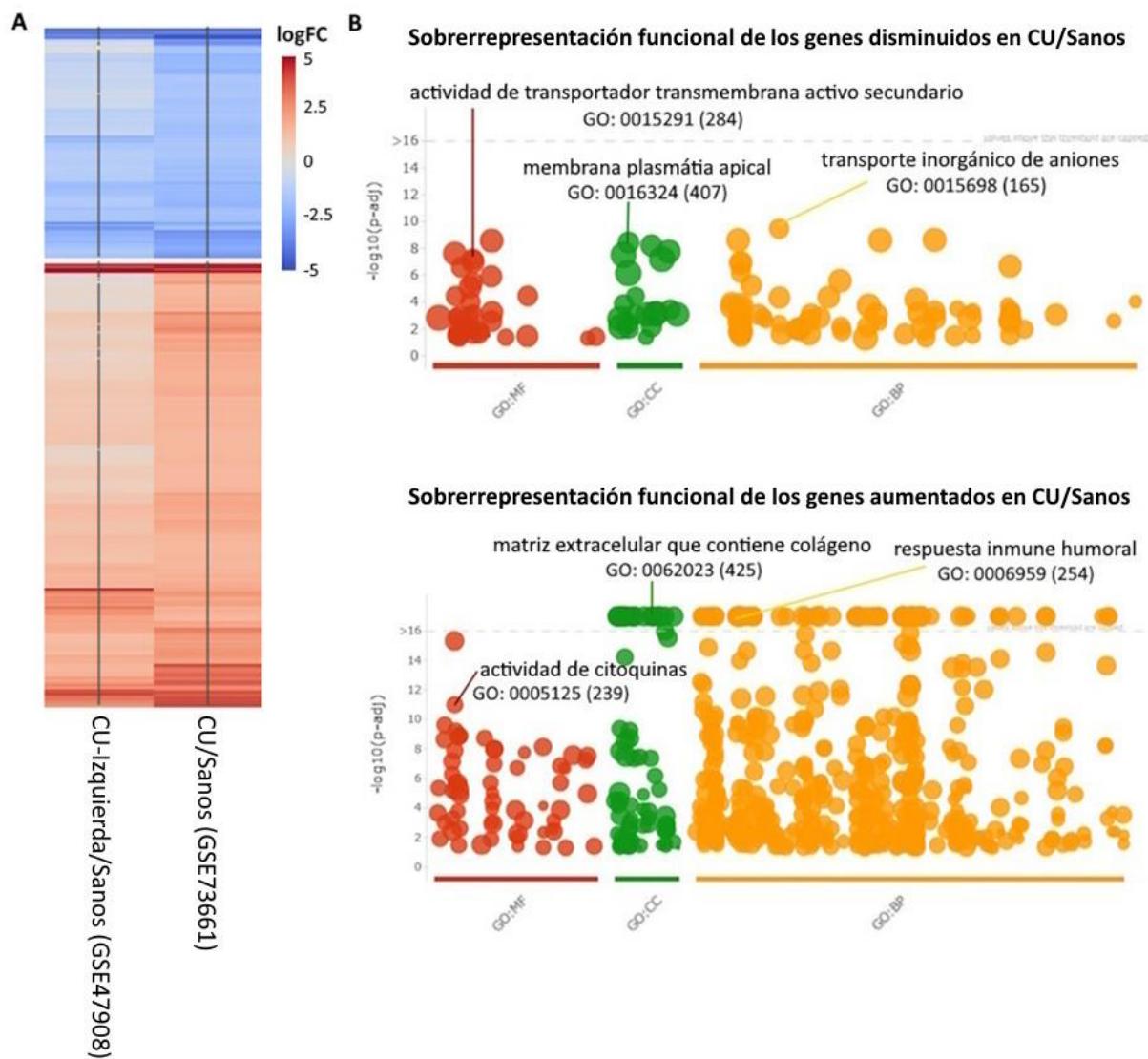
**Figure 11. Differentially expressed genes for the comparison of UC vs. healthy subjects throughout all human cohorts analyzed.** **A)** Heatmap with the genes shared by, from left to right, GSE224758 (Digby-Bell, no associated paper), GSE37283 (Pekow *et al.*, 2013), GSE47908 (Bjerrum *et al.*, 2014), and the results from GSE73661 (Lanzillotta, 2022; Arijs *et al.*, 2018), and the combination of GSE87473, GSE92415 and GSE206285 (Zhang *et al.*, 2023), which resulted in 7 negatively regulated and 10 positively regulated genes. For each cohort (columns) and each gene (rows), the value of the times of change (logFC) of the CU group with respect to the healthy controls is shown. Those genes that were significantly ( $p = 0.05$ ) differentially expressed in a cohort are denoted by an asterisk inside the cell. **B)** Venn diagram with significantly differentially expressed genes ( $p = 0.05$ ) in all cohorts, excluding GSE224758 because no significant value is obtained from it.

Decreased genes seem to indicate an alteration in intestinal mucosal homeostasis and metabolism in UC; among them, the following are observed: *SLC26A2* (Heneghan *et al.*, 2010; Shi *et al.*, 2020), which codes for a sulfate transporter; *HMGCS2*, a mitochondrial enzyme that plays a crucial role in ketogenesis (El-Sayed *et al.*, 2023) and *ADH1C*, a gene encoding the gamma subunit of class I alcohol dehydrogenase, which is involved in ethanol metabolism (Edenberg, 2007). On the other hand, the increased genes suggest a marked activation of epithelial regeneration and immune response. Among them, *OLFM4* stands out, which encodes the ollactomedina 4 protein and is overexpressed in the analyzed samples of active and

quiescent UC, known to be a marker of *intestinal stem cells* and participate in inflammatory processes (Xing *et al.*, 2024); *DUOX2*, which encodes the dual glycoprotein oxidase 2, of the NADPH oxidase family (capable of producing reactive oxygen species) (Poncelet *et al.*, 2019); and *LCN2*, which encodes lipocalin2, involved in innate immunity by limiting bacterial growth (Jaberi *et al.*, 2021). An increase in the expression of chemokine 9 (*CXCL9*), a chemokine that participates in the immune response particularly as an attractor of T lymphocytes to sites of inflammation, is also observed (Zhou *et al.*, 2023a); *MMP9*, a metalloproteinase; and *TIMP1*, a metalloproteinase inhibitor molecule (Gonzalez-Avila *et al.*, 2019). In general, quiescent UC (CUq) samples show dysregulated genes in the same sense as in active inflammation, but with lower levels of change and significance.

Taking this data into consideration, and looking specifically for dysregulation in actively inflamed mucosa vs. healthy tissue, the same comparison is made, but taking into account only samples from patients with active UC, so the GSE37283 cohort (which has UC samples in remission) and the GSE224758 cohort are not included, because they do not present significant results. Finally, the results of the combined cohorts reported by Zhang, *et al.* (2023), which because it is small and restrictive (21 genes decreased, 44 increased in their expression), could result in a large number of false negatives. This results in the analysis of the GSE47908 cohorts (pancolitis or E3 and left UC or E2 samples) and GSE73661 (active UC) and healthy subjects (without IBD) for comparison. The results obtained result in 144 genes generally decreased and another of 277 increased with high concordance (**Fig. 12 A**).

**UADE** SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía



**Figure 12. Analysis of differential expression and functional enrichment of genes shared between the GSE479008 and GSE73661 cohort for the active UC/healthy subjects comparison.** **A)** The heatmap shows the value of  $\log_2FC$  for each gene in the rows, corresponding to the comparisons in the columns, detailed at the bottom of the graph. The statistical significance, denoted by an asterisk in the corresponding cell, was determined as  $p < 0.05$ . **B)** The dotplot It shows the results of the overrepresentation analysis performed on the group of underexpressed (top) and overexpressed (bottom) genes. Each circle represents a pathway and its size, the number of genes it includes. Each pathway it is colored according to the database to which it belongs and the most relevant ones are labeled. On the axis y The p-value resulting from the analysis transformed into  $-\log_{10}$  is denoted, the graph is bounded at a maximum value of 16 demarcated by the dotted line. For both cases, statistical significance was determined as  $p < 0.05$ .

Among the most dysregulated genes for the comparison of active UC vs. tissue from healthy patients (without IBD) (**see Table VI**), the following are shared with the previous analysis: *SLC26A2*, *HMGCS2*, *ADH1C*, *CHP2* (decreased), *DUOX2* and *LCN2* (increased).

There are also dysregulations that have already been described in UC at the protein level, such as increased metalloprotease 3 (*MMP3*, Siloši *et al.*, 2014), and decreased aquaporin 8 (*AQP8*, Zahn, *et al.*, 2007). The decrease in the latter protein, which seals the tight paracellular junctions in the apical region of the plasma membrane, reflects an increase in epithelial permeability in CU (Zhu *et al.*, 2019). A decreased expression of galectin-4 (*LGALS4*), a gene of special interest for the laboratory (Mora Massaro's PhD thesis) (Massaro, 2025), is also observed in CU-Izquierda vs healthy patients (p-adjusted value = 1.68E-4, *logFC* = -0.59). In addition to this, the expression of *DUOXA2* is increased, which encodes the maturation factor of dual oxidase 2, which in turn enables the functioning of the *DUOX2* protein (Poncelet *et al.*, 2019), also increased. Another significantly increased gene for both cohorts is *CHI3L1*, a lectin relevant in inflammatory processes, which is involved in tissue repair and remodeling (Zhao *et al.*, 2020).

**TABLE VI. Significant genes with higher |logFC| for the comparison of active UC/healthy subjects (without IBD) in the GSE47908 and GSE73661 cohorts.** The orange cells mark the genes increased in CU and the blue cells, the decreased ones. The p-value corresponds to the adjusted p-value.

Gen		GSE47908		GSE73661	
Código	Nombre	logFC	p-valor	logFC	p-valor
ADH1C	Alcohol deshidrogenasa 1C (Clase I), polipéptido gamma	-2.80	4.70E-08	-3.24	1.95E-10
UGT2A3	UDP Glucuronosiltransferasa Familia 2 Miembro A3	-3.03	3.12E-07	-4.51	4.44E-21
CHP2	Proteína EF-Hand 2 similar a la calcineurina	-3.07	1.12E-08	-2.67	1.38E-12
HMGCS2	3-hidroxi-3-metilglutaril-CoA sintasa 2	-3.62	3.49E-10	-4.70	1.80E-16
SLC26A2	Solute Carrier Family 26 Miembro 2	-3.67	4.94E-09	-4.61	5.66E-16
ABCG2	Miembro 2 de la subfamilia G de casetes de unión a ATP (Grupo sanguíneo JR)	-3.78	9.91E-07	-3.68	1.50E-17
CLDN8	Claudina 8	-3.92	2.42E-05	-4.03	4.80E-14
PCK1	Fosfoenolpiruvato Carboxiquinasa 1	-3.99	7.99E-07	-3.67	7.13E-12
AQP8	Acuaporina 8	-5.64	9.83E-09	-5.43	2.37E-21
SLC6A14	Familia de portadores de solutos 6 miembro 14	6.09	1.08E-11	5.80	2.42E-22
REG1A	Miembro de la familia de regeneradores 1 Alfa	5.55	9.34E-07	4.75	9.40E-10
DUOX2	Doble oxidasa 2	5.51	7.98E-09	4.63	1.94E-13
REG1B	Miembro de la familia de regeneradores 1 Beta	5.28	7.56E-07	3.98	3.56E-07
LCN2	Lipocalina 2	4.27	2.41E-10	3.73	2.90E-15
DUOXA2	Factor 2 de maduración de la oxidasa dual	4.26	3.03E-10	4.54	2.79E-14
CXCL1	Ligando 1 de la quimiocina con motivo C-X-C	4.04	3.32E-09	3.62	9.67E-19
MMP3	Metalopeptidasa de matriz 3	4.04	2.28E-06	4.82	3.03E-10
TNIP3	Proteína 3 de interacción con TNFAIP3	3.95	1.11E-07	4.17	2.27E-12

An analysis of functional overrepresentation of the increased and decreased *clusters* between active UC and non-inflamed tissue was performed (see Table VII; Fig. 12). First, the *cluster* of diminished genes gave as the most significant terms related to transmembrane transport of inorganic anions, localization in the apical plasma membrane, and the metabolic

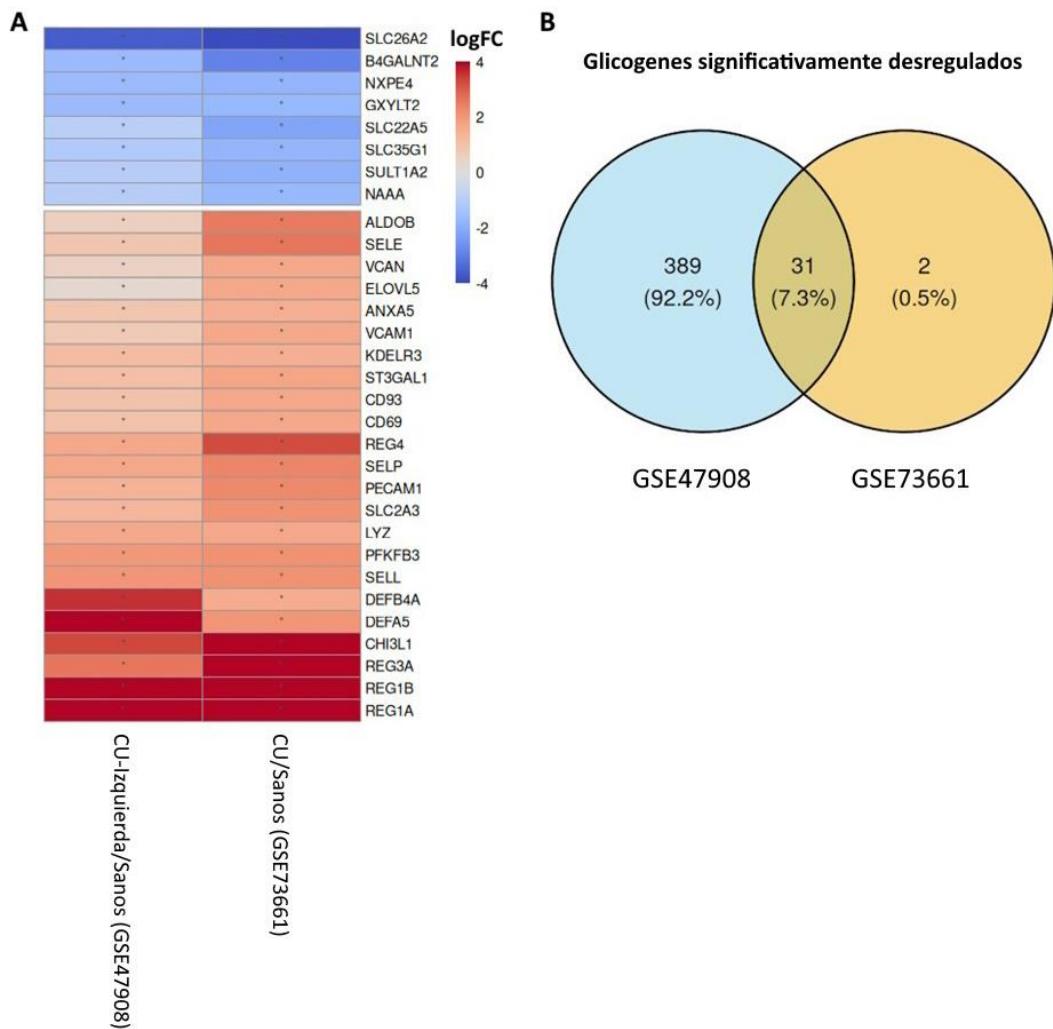
process of steroids. On the other hand, the functional analysis of augmented genes resulted in terms related to the humoral immune response, migration of leukocytes, granulocytes and neutrophils, chemotaxis, and response to molecules of bacterial origin, as well as the regulation of intercellular adhesion and binding of glycosaminoglycans, the activity of serine-type peptidase and endopeptidase, and localization in the extracellular matrix.

**TABLE VII. Main results of the overrepresentation analysis for the comparison of active UC/healthy subjects (without IBD) in cohorts GSE47908 and GSE73661.** The orange cells mark the increased terms in CU with the lowest adjusted p-value and the blue cells, the decreased ones with the lowest adjusted p-value. A filter was used to include pathways that had 500 genes or less.

ID	Fuente	Término	P-valor	Precisión	Recall
GO:0015698	GO:BP	Transporte de aniones inorgánicos	3.42E-10	0.110	0.091
GO:0016324	GO:CC	Membrana plasmática apical	3.64E-09	0.137	0.047
GO:0045177	GO:CC	Parte apical de la célula	5.98E-09	0.144	0.042
GO:0015291	GO:MF	Actividad del transportador transmembrana activo secundario	6.49E-08	0.116	0.056
GO:0006820	GO:BP	Transporte de aniones monoatómicos	1.05E-07	0.096	0.078
GO:0008514	GO:MF	Actividad del transportador transmembrana de aniones orgánicos	2.54E-07	0.109	0.056
GO:0006821	GO:BP	Transporte de cloruro	6.80E-07	0.081	0.090
GO:0022804	GO:MF	Actividad del transportador transmembrana activo	1.23E-06	0.130	0.040
GO:0015293	GO:MF	Actividad del simportador	3.20E-06	0.080	0.075
GO:0008202	GO:BP	Proceso metabólico de esteroides	5.40E-06	0.110	0.046
GO:0006959	GO:BP	Respuesta inmunitaria humoral	7.25E-26	0.139	0.150
GO:0050900	GO:BP	Migración de leucocitos	8.93E-24	0.158	0.109
GO:0097529	GO:BP	Migración de leucocitos mieloides	3.06E-22	0.125	0.143
GO:0097530	GO:BP	Migración de granulocitos	6.21E-22	0.106	0.186
GO:1990266	GO:BP	Migración de neutrófilos	1.02E-21	0.099	0.209
GO:0032496	GO:BP	Respuesta a lipopolisacáridos	2.83E-21	0.139	0.113
GO:0030595	GO:BP	Quimiotaxis leucocítica	4.38E-21	0.121	0.139
GO:0002237	GO:BP	Respuesta a moléculas de origen bacteriano	1.87E-20	0.139	0.107
GO:0071621	GO:BP	Quimiotaxis de granulocitos	2.23E-20	0.095	0.202
GO:0006935	GO:BP	Quimiotaxis	4.58E-20	0.154	0.091

Thirdly and finally, focusing on glycobiological genes (**Annex 8.1**), 31 significantly dysregulated glycogenes were found in both cohorts for the comparison of biopsies from patients with active UC vs. subjects without IBD (**Fig. 13**).

**UADE** SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía



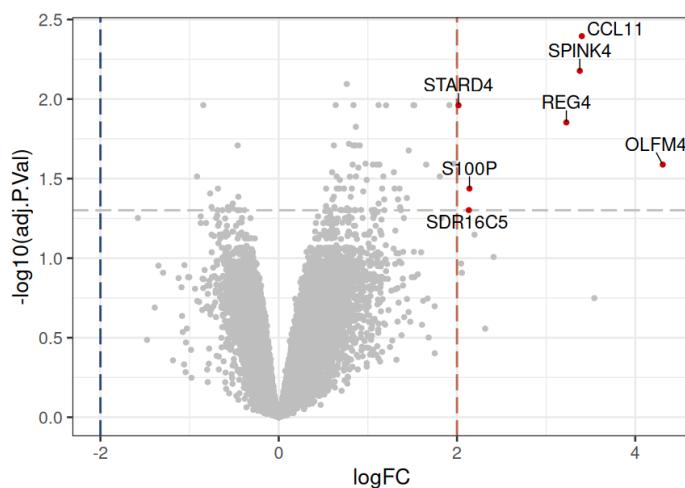
**Figure 13. Analysis of differential expression of glycogenes shared between the GSE479008 and GSE73661 cohorts for the CU/Control comparison.** **A)** The heatmap shows the value of  $\log_{2}FC$  for the 31 genes consistently dysregulated in the rows, corresponding to the comparisons in the columns, detailed at the bottom of the graph. The statistical significance, denoted by an asterisk in the corresponding cell, was determined as  $p < 0.05$ . **B)** The Venn diagram shows the significantly dysregulated shared glycogenes in both cohorts, indicating an intersection of 7.3% or 31 genes.

Among the genes with decreased expression, in addition to the transporter already described *SLC26A2*, is *B4GALNT2*, an N-acetylgalactosaminyltransferase that participates in the biosynthesis of the Sda antigen, of great importance in the intestine (Groux-Degroote *et al.*, 2021). We also found an increase in the expression of: selectins (*SELL*, *SELP* and *SELE*), key lectins in leukocyte migration; *VCAM1*, an adhesin that also participates in this process (Singh *et al.*, 2023); from *CHI3L1*, a lectin involved in inflammatory processes (Chen *et al.*, 2024), *ALDOB* and *PFKFB3*, enzymes that participate in glycolysis, and finally, members of the REG

family (*REG4*, *REG3A*, *REG1B*, *REG1A*), related to the immune response to bacteria (Sun *et al.*, 2021). Also related to this last function is increased defensin 5 (*DEFA5*), which acts as an antimicrobial peptide (Sabit *et al.*, 2024). Another increased gene is *ST3GAL1*, which codes for an α(2,3)-sialyltransferase, so its dysregulation poses a potential alteration in the sialome (Fan *et al.*, 2025).

### 5.3.1.2. Differences in expression between healthy gut, quiescent ulcerative colitis, and active inflammation

Given the results in the previous section, we were interested in delving into the differences and similarities between quiescent UC and healthy tissue. By analyzing quiescent UC samples ( $n = 4$ ) with respect to healthy subjects ( $n = 5$ ) from the GSE37283 dataset, 93 differentially expressed genes were found taking a p-value of less than 0.05 as a cut-off point, of which 17 had decreased expression and 76 increased (see Table VIII; Fig. 14).



**Figure 14. Differential expression analysis in biopsies of patients with quiescent ulcerative colitis and healthy subjects (without IBD).** The Volcano plot It shows the results of the differential expression analysis of patients with quiescent ulcerative colitis ( $n = 4$ ) and healthy subjects ( $n = 5$ ) from the GSE37283 cohort. The x-axis denotes the  $\log_{2}FC$  of the comparison and the blue and red dotted lines are located at the values -2 and 2. The y-axis corresponds to the logarithmic transformation in base 10 of the adjusted p-value and the gray dotted line denotes the cut-off point of statistical significance ( $p < 0.05$ ). Differentially overexpressed genes with a  $\log_{2}FC$  greater than 2 are marked as red dots with their names.

**UADE** SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía

---

**TABLE VIII.** Significantly ( $p < 0.05$ ) main genes were dysregulated between patients with quiescent ulcerative colitis and healthy subjects (without IBD). The p-value corresponds to the adjusted p-value. In orange, the augmented terms are denoted and in blue, the diminished ones. In addition, the rows corresponding to the glycogenes have the text in bold.

Gen	Nombre	logFC	p-valor
CCL11	Ligando de quimiocinas con motivo C-C 11	3.40	4.02E-03
SPINK4	Inhibidor de la serina peptidasa Kazal tipo 4	3.38	6.65E-03
KPNA1	Subunidad de Carioferina Alfa 1	0.76	8.04E-03
TPM4	Tropomiosina 4	0.64	1.09E-02
FAM241A	Familia con similitud de secuencia 241 Miembro A	1.51	1.09E-02
PRDX5	Peroxiredoxina 5	-0.85	1.09E-02
SERPINB5	Miembro 5 de la familia B de las Serpinas (O "Mapsin")	1.91	1.09E-02
BZW1	Cremallera de leucina básica y dominios W2 1	1.12	1.09E-02
<b>SEC24D</b>	<b>Proteína transportadora de proteínas Sec24D</b>	<b>0.84</b>	<b>1.09E-02</b>
TIGAR	Fosfatasa reguladora de la glucólisis inducida por TP53	1.20	1.09E-02
PNP	Purina Nucleósido Fosforilasa	1.52	1.09E-02
STARD4	Dominio de transferencia de lípidos 4 relacionado con StAR	2.02	1.09E-02
<b>REG4</b>	<b>Miembro de la familia de regeneradores 4</b>	<b>3.23</b>	<b>1.40E-02</b>

Among the genes with significantly increased expression between quiescent UC and healthy tissue, inflammation-related genes such as *CCL11*, an eosinophil-recruiting chemokine (Polosukhina *et al.*, 2021), *OLFM4*, an extracellular matrix glycoprotein that facilitates cell adhesion and is involved in intestinal inflammation (Xing *et al.*, 2024), *SPINK4*, a serine protease inhibitor that has been reported as a serological marker of IBD (Wang *et al.*, 2024), and the glycogen *REG4*, which acts in response to bacteria and tissue regeneration in the face of injury (Sun *et al.*, 2021). In addition, the expression of *STARD4*, which codes for a soluble protein involved in cholesterol transport, is increased (Talandashti *et al.*, 2024), *S100P*, a calcium-binding protein that is involved in multiple cellular processes, including cell cycle progress (Camara *et al.*, 2020), and *SDR16C5*, an alcohol dehydrogenase involved in the oxidation of retinol (vitamin A) to its biologically active form, retinoic acid, and is related to the functioning of the endoplasmic reticulum (Adams *et al.*, 2017).

Then, focusing on the differentially expressed glycogenes, we found 11 increased in CUq with respect to healthy tissue, including *REG4* (see Table IX).

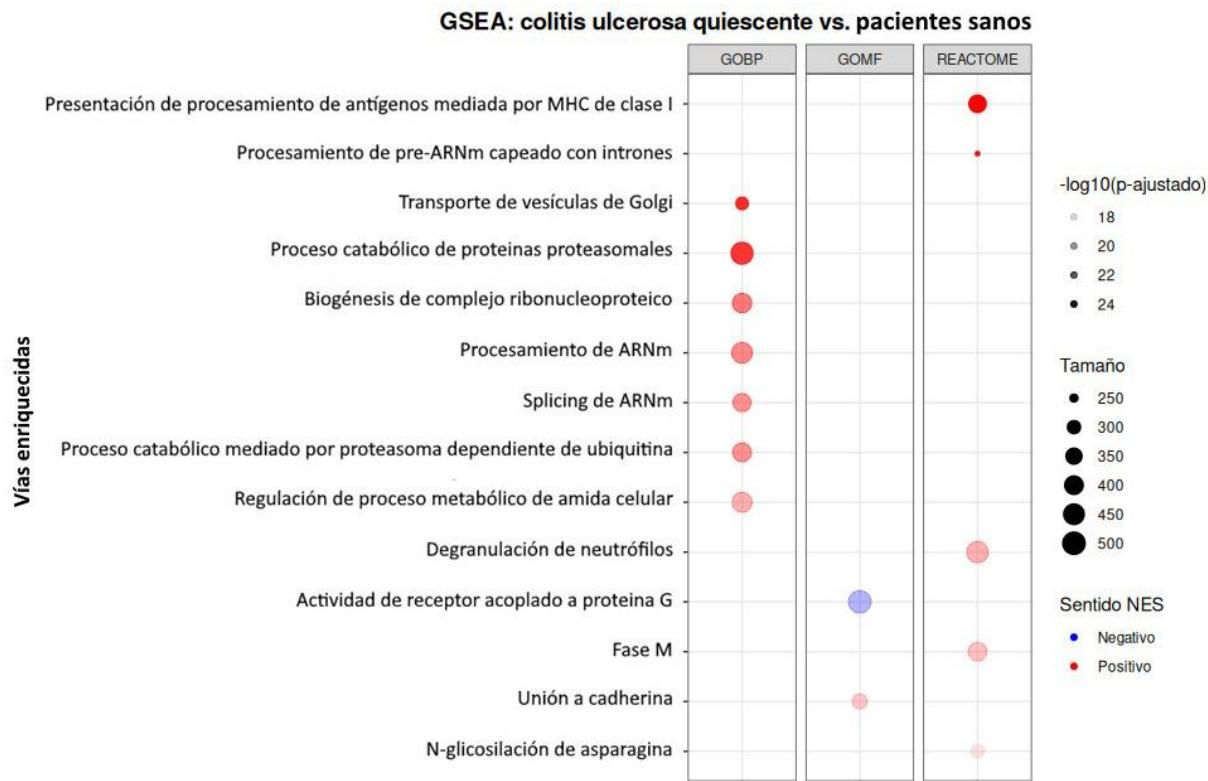
**TABLE IX.** Significantly ( $p < 0.05$ ) glycogenes were significantly dysregulated between patients with quiescent ulcerative colitis and healthy subjects (without IBD). The p-value corresponds to the adjusted p-value. Orange denotes augmented terms.

Gen	Nombre	logFC	p-valor
SEC24D	Proteína transportadora de proteínas Sec24D	0.84	1.09E-02
REG4	Miembro de la familia de regeneradores 4	3.23	1.40E-02
PDIA3	Miembro 3 de la familia A de las isomerasas disulfuro de proteínas	0.86	1.96E-02
KDELR3	Receptor de retención de proteínas del retículo endoplasmático KDEL 3	1.13	2.58E-02
SULT1C2	Miembro 2 de la familia 1C de las sulfotransferasas	1.65	2.58E-02
SAR1B	GTPasa 1B relacionada con Ras asociada a la secreción	0.87	3.07E-02
ATP6V0A1	Subunidad A1 de la ATPasa transportadora de H+ V0	0.65	3.65E-02
SGPP1	Esfingosina-1-Fosfato Fosfatasa 1	1.30	3.65E-02
GMPS	Guanina monofosfato sintasa	0.56	4.99E-02
TMED7	Proteína transmembrana de tráfico de P24 7	0.76	4.99E-02
SEC24A	SEC24 Homólogo A, Componente del Complejo de Cubierta COPII	0.85	4.99E-02

Three of these eleven increased glycogenes (*SAR1B*, *SEC24A*, and *SEC24D*) belong to the Coat II Protein Complex (COPII), a group of proteins that facilitates the formation of vesicles to transport proteins from the endoplasmic reticulum to the Golgi apparatus (Garbes *et al.*, 2015). The *SAR1B* protein is a GTPase that acts as a molecular switch, helping to assemble the COPII layer for the transport of proteins and lipids from the ER to the Golgi apparatus. In the cells of the small intestine (enterocytes), *SAR1B* is also crucial for the transport of chylomicrons, which are essential for the absorption and transport of fats and fat-soluble vitamins (Sané *et al.*, 2017). Likewise, proteins of the endoplasmic reticulum are found, such as *PDIA3*, which modulates glycoprotein folding, and *KDELR3* (Chichiarelli *et al.*, 2022). It also increases the expression of *ATP6V0A1*, a vacuolar ATPase (V-ATPase) that mediates the acidification of organelles such as lysosomes and participates in the absorption of cholesterol in CRC cells, resulting in mechanisms that mediate antitumor immune suppression (Huang *et al.*, 2024).

Functional enrichment analysis using the GSEA method in quiescent UC compared to healthy subjects indicates activation of pathways related to antigen presentation and neutrophil degranulation, protein degradation, RNA processing, vesicular transport and Golgi apparatus functionality, cadherin binding and N-glycosylation; on the other hand, there is a decrease in the activity of G-protein-coupled receptors —these results seem to indicate that, even in

quiescent UC, despite the absence of signs of macroscopic inflammation, the intestine is altered in highly relevant processes, including at the glycobiological level (**Fig. 15**).

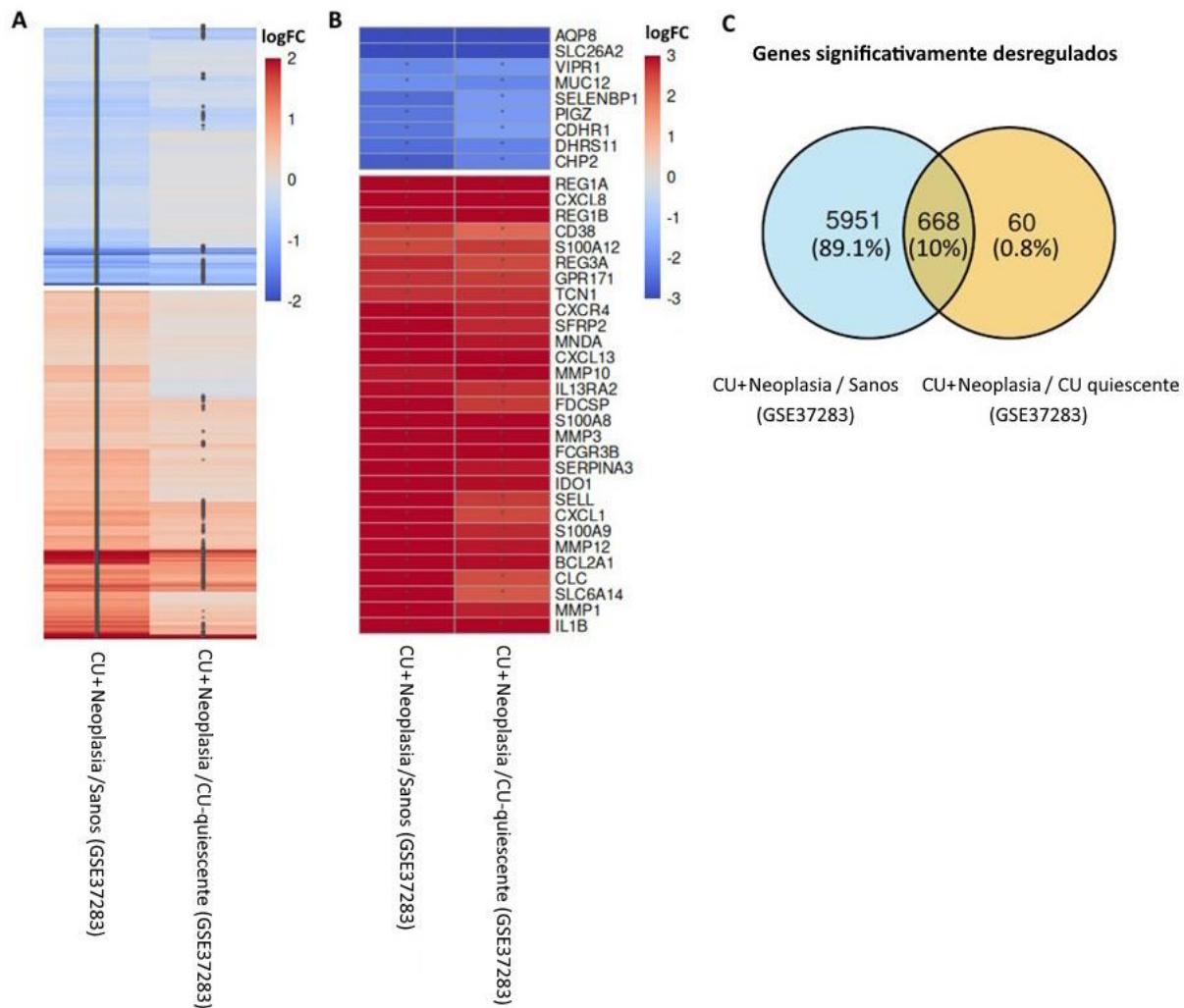


**Figure 15. Functional enrichment analysis comparing transcriptional profile of biopsies from patients with quiescent ulcerative colitis compared to healthy subjects (without IBD).** The bubble plot displays the results of the GSEA analysis. The rows show the pathways composed of more than 5 and less than 500 genes that showed the most statistically significant differences ( $p < 0.05$ ). The color of the circle denotes whether the pathway was found activated (red) or inhibited (blue) in the quiescent ulcerative colitis group, the intensity of the color shows the value of the adjusted p-value transformed into logarithm in base 10, and the size of the circle denotes the number of genes included in each group. The different columns divide the roads according to their source dataset: GO: BP, GO: MF, or Reactome.

As a second approach, the results of the differential expression of samples obtained from biopsies of UC with remote neoplasia ("UC+Neoplasia") were analyzed compared to quiescent UC samples and healthy subjects, using data from the GSE37283 cohort. We performed the comparison by grouping quiescent UC and healthy subjects, and then only against samples from healthy subjects. When comparing dysregulation against CU+ neoplasia, similar dysregulation patterns were observed in both comparisons, although the trends are clearer and more significant in the comparison with biopsies of subjects without IBD, where we found two

**UADE** SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía

delimited clusters, with 3837 genes of increased expression and 2842 of decreased expression (Fig. 16 A).



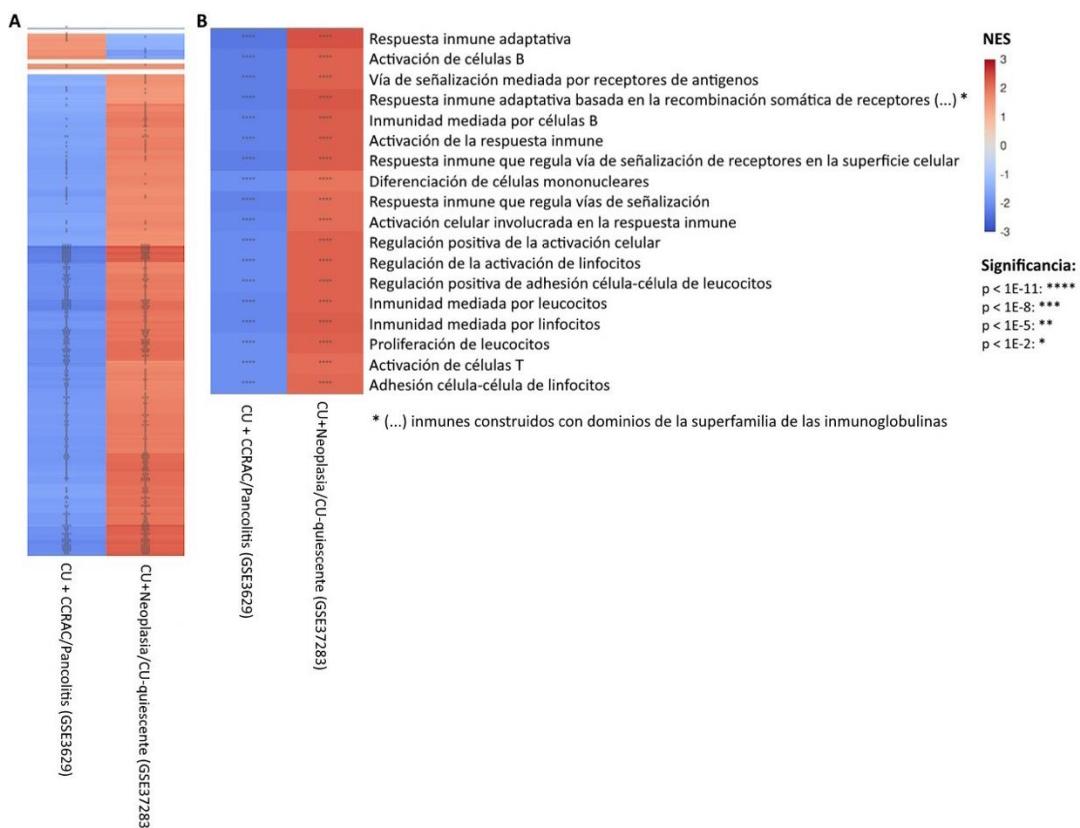
**Figure 16. Comparative analysis of dysregulation in patients with UC+neoplasia, compared to healthy patients (without IBD) and quiescent ulcerative colitis.** A) The heatmap with the significant genes for one or the other comparison, it shows the value of  $\log_{2}FC$  for each gene, in the rows, corresponding to the comparisons, in the columns, detailed at the bottom of the graph. The statistical significance, denoted by an asterisk in the corresponding cell, was determined as  $p < 0.05$ . B) Selection of 29 augmented genes and 9 inhibited genes, from among those with the highest  $|\log_{2}FC|$  for both comparisons. Cells marked with an asterisk are significant ( $p < 0.05$ ). C) Venn diagram demonstrating the number of significantly dysregulated genes between the two comparisons.

Of these genes, among the most significant (Fig. 16 B), several that we had seen dysregulated in active UC are dysregulated again in patients with UC and remote neoplasia: among them, *AQP8*, *SLC26A2* and *LGALS4* are again decreased, and members of the *REG1A*,

# UADE SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía

*REG3A*, and *REG1B* families of regulators, CXCL1, CXCL8, and CXCL13 chemokines are increased, the chemokine receptor CXCR4, metalloproteases MMP3, MMP10 and MMP12, and interleukins such as IL1B and the interleukin receptor IL13RA2. However, the behavior of healthy and remissionary mucosa again demonstrates a different transcriptional profile, since 60 genes were found to be significantly dysregulated in patients with UC and remote neoplasia vs. quiescent ulcerative colitis that were not dysregulated compared to the control group (Fig. 16 C), of which 9 were increased and 51 decreased (Annex 8.2). Among the latter, the decrease in GCNT2, a glycosyltransferase proposed as a biomarker in cancer (Perez *et al.*, 2021).

Finally, the results of the functional analysis of the samples of patients with UC+ neoplasia (a term that will be used to encompass the "CU+CCRAC" group from the GSE3629 and "CU+Neoplasia" group from GSE37283) were compared with the different forms of ulcerative colitis for each cohort (active pancolitis, or CUq). The result was 4 clusters with GO:BP terms with a generally opposite behavior between comparisons (Fig. 17 A).



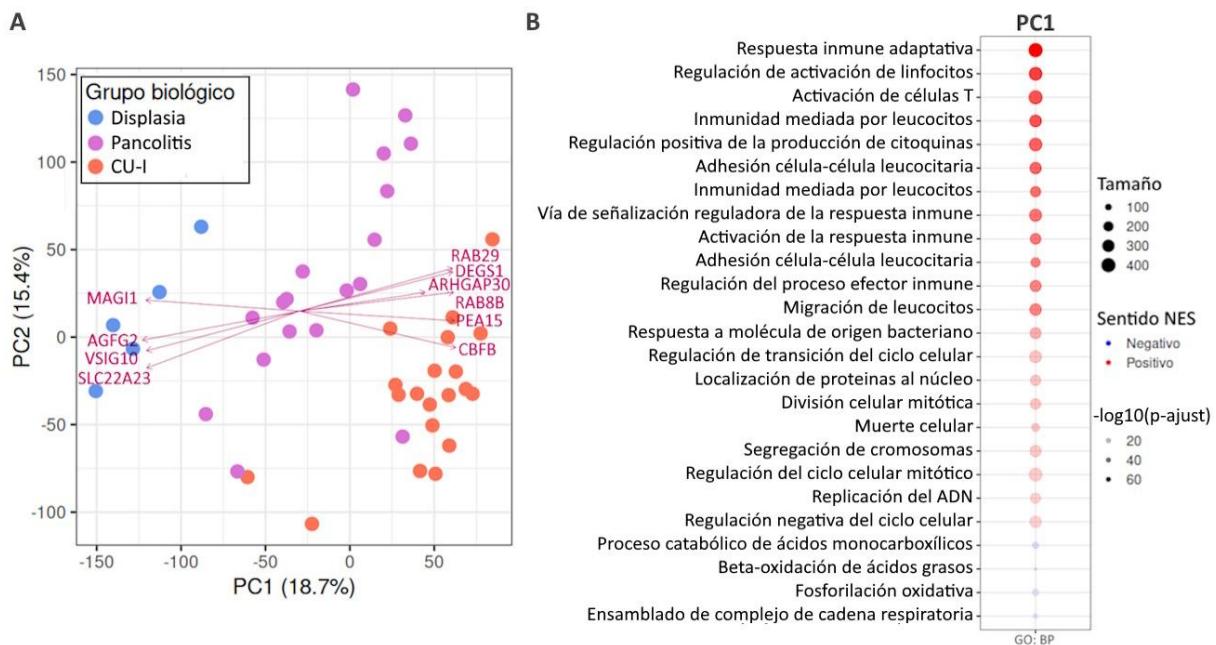
**Figure 17. Heatmaps Comparing Terms GO: BP of UC+tissue biopsies with neoplasia vs. extensive active UC (E3), and quiescent UC.** The graphs show the enrichment value (NES, *Normalized Enrichment Score*) for each pathway, in the rows, corresponding to each comparison, detailed in the columns, at the bottom of each graph. The statistical significance, denoted by an asterisk in the corresponding cell, was determined as  $p < 0.05$ . Terms containing between 5 and 500 genes are included. **A) Heatmap** with the terms GO: BP total significant. **B) Heatmap Reduced**, selecting the terms with  $p$ -value adjusted  $< 1E-11$  for both comparisons.

After selecting the most significantly enriched pathways (adjusted value  $< 1E-11$  for both comparisons), pathways related to the activation of the immune response were observed, positively enriched in CU+Neoplasia vs. CUq, and negatively enriched in CU+CCRAC vs. pancolitis (**Fig. 17 B**). In other words, inflammation in pancolitis is greater than in CU+ neoplasia; meanwhile, inflammation in CU+ neoplasm is greater than in quiescent UC (pancolitis  $>$  CU+tissue with neoplasm  $>$  CUq). These results can be explained by two phenomena: the induction of an immunosuppressive microenvironment in association with the tumor, and the decrease in inflammation in quiescence. The analysis suggests that tissue from biopsies of patients with neoplastic lesions, and despite being analyzed in areas remote from the tumor, immunosuppression is present.

### **5.3.2. Analysis of patients with dysplasia, left ulcerative colitis (E2), and pancolitis (E3)**

Next, we decided to perform a principal component analysis of the dysplasia, left ulcerative colitis, and pancolitis samples of the GSE47908 cohort, which based on gene dysregulation revealed a separate grouping for the three sample types, with very little overlap (**Fig. 18 A**).

# UADE SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía



**Figure 18. Principal component analysis for dysplasia, left ulcerative colitis, and pancolitis specimens. A)** The Principal Component Analysis (PCA, *Principal Components Analysis*) describes the two main components that contain 34.1% of the variance. Overlapping are the top 12 genes that contribute the most to PC1 variation, including *MAGI1* towards negative values, and *ARHGAP30* towards the positive ones. **B)** The *bubble plot* shows a selection of terms from *GO: BP* of GSEA analysis for PC1. The rows show the *pathways* composed of more than 5 and less than 500 genes that showed the most statistically significant differences ( $p < 0.05$ ). The color of the circle denotes whether the *pathway* was found to be activated/related to the most positive values of the x-axis of the PCA graph (red) or inhibited/related to the most negative values of the x-axis of the PCA graph (blue), the intensity of the color shows the value of the adjusted p-value transformed into logarithm in base 10, and the size of the circle denotes the number of genes included in each *pathway*.

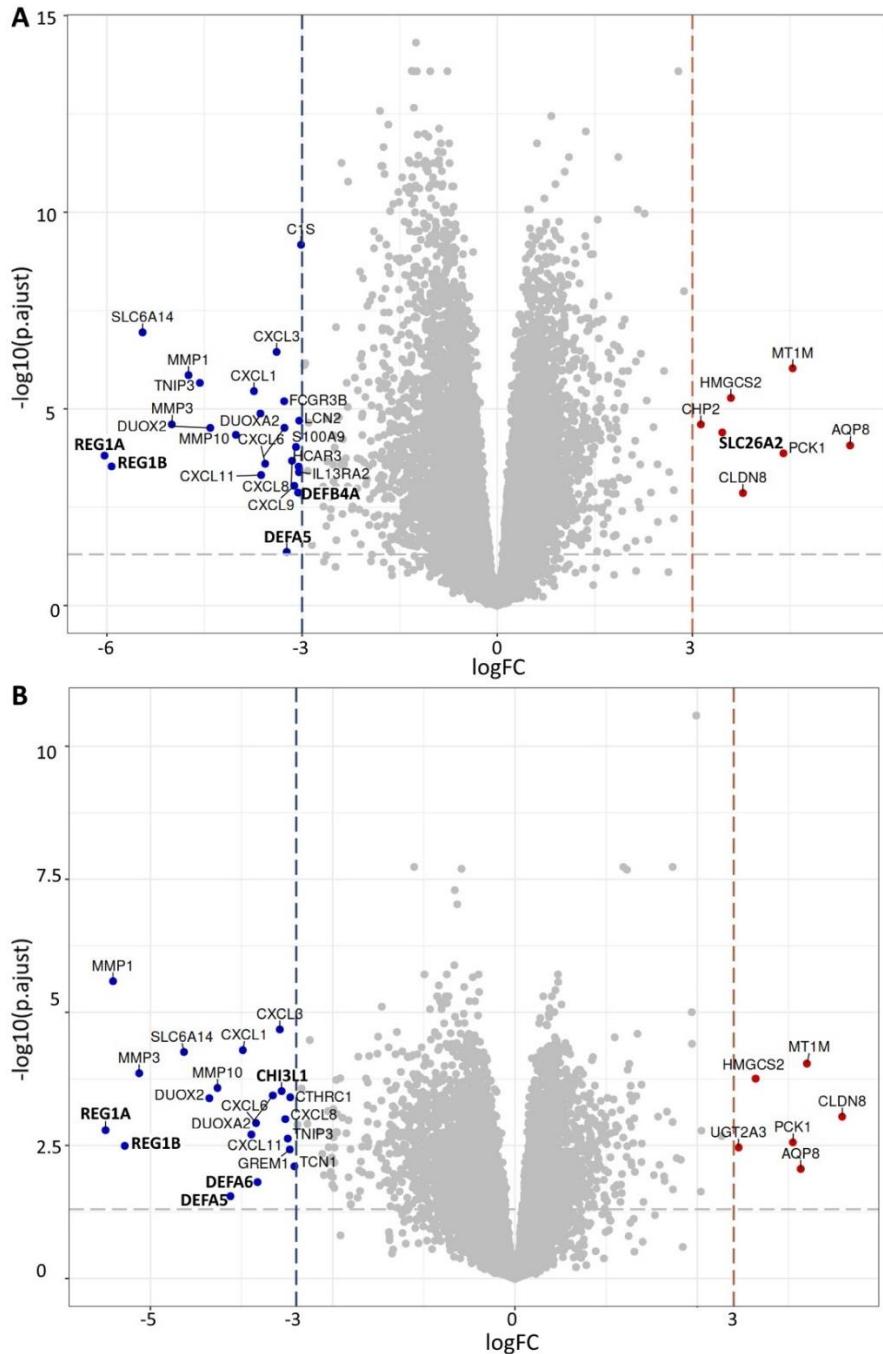
The first two main components (PCs) contain 34.1% of the variance and separation occurs mainly along PC1 (x-axis), which contains 18.7% of the variance: samples from patients with left UC are located towards positive values, with dysplasia towards negative values and with pancolitis in the center. The approach to the *cluster* of dysplastic samples by means of the *cluster* of pancolytic samples could indicate that dysregulation that occurs with increasing the extent and severity of inflammation (left UC to pancolitis) would "approach" the tissue to a stage more similar to dysplasia, which correlates with the higher frequency of dysplasia in these patients (Sambuelli et al., 2019).

Next, we performed the functional analysis of the first principal component using the *GSEA method* (Fig. 18 B). This analysis revealed an enrichment in the left UC samples of pathways associated with the immune response, cell cycle regulation and apoptosis (positive

values of the component). Among the genes that contribute most to this direction are *CBFB*, a subunit of a transcription factor involved in hematopoiesis and activation of immune cells, as well as *RAB8B* and *RAB29*, small GTPases, enzymes essential for immunological functions such as vesicular trafficking, signal transduction and pathogen elimination (Homma *et al.*, 2021). Contributing in the same direction as the first main component is *ARHGAP30*—GTPase Rho activating protein—, whose expression is necessary for the activation of p53 after stress due to DNA damage, and its level correlates with acetylation and functional activation of p53 in tissues with CRC. In addition, a low level of *ARHGAP30* expression is associated with low survival in CRC patients (Wang *et al.*, 2014). On the other hand, the negative values of the component (towards which the dysplasia samples tend) showed a clear metabolic change, with enrichment of pathways related to energy production, especially beta-oxidation of fatty acids and oxidative phosphorylation. In this direction, and therefore more expressed in dysplasia, genes involved in cell adhesion, such as *VSIG10* and *MAGII*, stand out.

When performing the differential gene expression analysis between dysplasia and left UC, 9885 significantly differentially expressed genes were identified, of which 5913 are overexpressed (n glycogenes = 152) and 3972 underexpressed (n glycogenes = 272) (**see Table X; Fig. 19 A**).

**UADE** SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía



**Figure 19. Main dysregulated genes between (A) dysplasia and ulcerative colitis E2 (left) and (B) dysplasia and ulcerative colitis E3 (extensive or pancolitis).** The volcano plots Results of differential expression analysis of patients with dysplasia vs A) CU-I, or B) pancolitis of cohort GSE47908. The x-axis denotes the  $\log_{2}FC$  of the comparison and the blue and red dotted lines are located at the values -3 and 3. The y-axis corresponds to the logarithmic transformation in base 10 of the adjusted p-value and the gray dotted line denotes the cut-off point of statistical significance ( $p < 0.05$ ). Differentially overexpressed genes with a  $\log_{2}FC$  greater than 3 are marked as red dots with their names, and inhibited ones with a  $\log_{2}FC$  less than -3, are indicated in blue.

**UADE** SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía

---

**TABLE X. Main general genes and significantly dysregulated glycogenes (p-value < 0.05, |logFC| > 2) among patients with dysplasia and left ulcerative colitis.** The p-value corresponds to the adjusted value and the glycogenes are shown in bold.

Genes totales (Displasia/CU-I)			
Gen	Nombre	LogFC	P-Valor
CCDC183	Proteína 183 que contiene un dominio de bobina enrollada	2.78	2.61E-14
TRAPP C1	Subunidad 1 del complejo de transporte de proteínas asociado al retículo trans-Golgi	-2.39	5.61E-12
<b>LMAN1</b>	<b>Lectina tipo manosa 1</b>	<b>-2.29</b>	<b>1.67E-11</b>
CTS Z	Catepsina Z	2.16	8.51E-11
ENTREP2	Proteína 2 relacionada con el receptor de endotelina	2.27	1.09E-10
C1S	Componente 1 subcomponente S del complemento	-3.02	6.71E-10
RRM2	Subunidad 2 de la ribonucleótido reductasa	-2.11	3.25E-09
EPHX1	Epóxido hidrolasa 1 (microsomal)	-2.06	4.81E-09
MT1F	Metalotioneína 1F	2.87	1.02E-08
WARS1	Sintetasa del triptófano ARNt 1	-2.48	8.39E-08
ANLN	Anilina (proteína de anillado)	-2.08	8.52E-08
SLC6A14	Transportador de aminoácidos dependiente de sodio y cloruro	-5.45	1.14E-07
Glicogenes (Displasia/CU-I)			
<b>LMAN1</b>	<b>Lectina tipo manosa 1</b>	<b>-2.29</b>	<b>1.67E-11</b>
<b>DEFB1</b>	<b>Beta-defensina 1</b>	<b>2.56</b>	<b>1.09E-06</b>
<b>CHI3L1</b>	<b>Proteína 1 similar a la quitinasa 3</b>	<b>-3.27</b>	<b>3.04E-05</b>
<b>SLC26A2</b>	<b>Transportador de sulfato de la familia SLC26 miembro 2</b>	<b>3.46</b>	<b>3.99E-05</b>
<b>NXPE4</b>	<b>Proteína similar a neurexofina 4</b>	<b>2.07</b>	<b>1.49E-04</b>
<b>REG1A</b>	<b>Gen relacionado con la regeneración 1 alfa</b>	<b>-6.04</b>	<b>1.55E-04</b>
<b>LAMP3</b>	<b>Proteína 3 asociada a lisosomas</b>	<b>-2.35</b>	<b>2.52E-04</b>
<b>REG1B</b>	<b>Gen relacionado con la regeneración 1 beta</b>	<b>-5.93</b>	<b>2.90E-04</b>
<b>SELL</b>	<b>Selectina L</b>	<b>-2.52</b>	<b>7.08E-04</b>
<b>REG4</b>	<b>Gen relacionado con la regeneración 4</b>	<b>-2.24</b>	<b>1.27E-03</b>
<b>DEFB4A</b>	<b>Beta-defensina 4A</b>	<b>-3.06</b>	<b>1.35E-03</b>
<b>REG3A</b>	<b>Gen relacionado con la regeneración 3 alfa</b>	<b>-2.56</b>	<b>1.72E-03</b>

On the other hand, the genes differentially significantly expressed between dysplasia and pancolitis reach a total of 5538 (56% less than against left UC), of which 2899 are overexpressed (n glycogenes = 82) and 2639 underexpressed (n glycogenes = 197) (see Table XI; Fig. 19 B).

**UADE** SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía

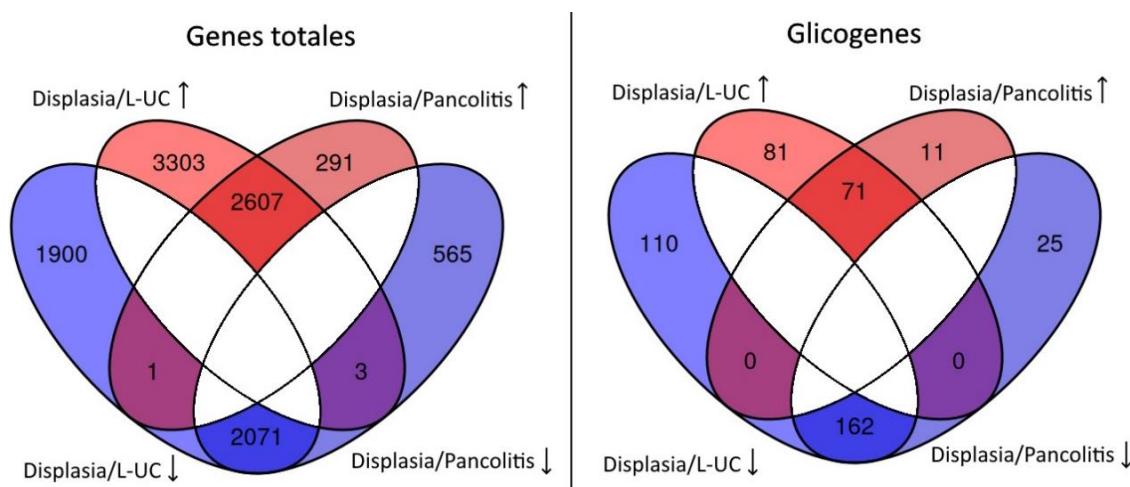
**TABLE XI. Major overall genes and significantly dysregulated glycogenes (p-value < 0.05, |logFC| > 2) among patients with dysplasia and pancolitis.** The p-value corresponds to the adjusted value and the glycogenes are shown in bold.

Genes totales (Displasia/Pancolitis)			
Gen	Nombre	LogFC	P-Valor
CCDC183	Proteína 183 que contiene un dominio de bobina enrollada	2.49	2.66E-11
ENTREP2	Proteína 2 relacionada con el receptor de endotelina	2.16	1.85E-08
MMP1	Metaloproteinasa de matriz 1	-5.51	2.60E-06
MT1F	Metalotioneína 1F	2.43	9.91E-06
CXCL3	Quimioquina ligando CXC 3	-3.23	2.09E-05
CD55	Molécula reguladora de la cascada del complemento (decay-accelerating factor)	-2.81	3.31E-05
CDHR1	Proteína 1 relacionada con cadherina	2.43	3.90E-05
CXCL1	Quimioquina ligando CXC 1	-3.73	5.13E-05
SLC6A14	Transportador de aminoácidos dependiente de sodio y cloruro	-4.54	5.54E-05
TMEM158	Proteína de membrana transmembrana 158	-2.15	6.80E-05
MT1M	Metalotioneína 1M	4.01	9.15E-05
MMP3	Metaloproteinasa de matriz 3	-5.15	1.39E-04
Glicogenes (Displasia/Pancolitis)			
CHI3L1	<b>Proteína 1 similar a la quitinasa 3</b>	-3.20	3.02E-04
NXPE4	<b>Proteína similar a neurexofina 4</b>	2.08	7.32E-04
LAMP3	<b>Proteína 3 asociada a lisosomas</b>	-2.44	8.15E-04
REG1A	<b>Gen relacionado con la regeneración 1 alfa</b>	-5.62	1.63E-03
SLC26A2	<b>Transportador de sulfato de la familia SLC26 miembro 2</b>	2.84	2.11E-03
REG1B	<b>Gen relacionado con la regeneración 1 beta</b>	-5.35	3.22E-03
REG4	<b>Gen relacionado con la regeneración 4</b>	-2.11	6.37E-03
REG3A	<b>Gen relacionado con la regeneración 3 alfa</b>	-2.48	6.68E-03
SELL	Selectina L	-2.13	9.28E-03
DEFB6	<b>Alfa-defensina 6</b>	-3.53	1.54E-02
DEFB4A	<b>Beta-defensina 4A</b>	-2.51	1.73E-02
DEFB5	<b>Alfa-defensina 5</b>	-3.90	2.84E-02

In particular, galectin-4 (*LGALS4*), of special interest to our laboratory, is significantly increased in dysplasia with respect to both CU-I (p-adjusted value = 7.98E-03, *logFC* = 0.58) and pancolitis (p-adjusted value = 4.50E-02, *logFC* = 0.48).

Beyond the analysis of the most significantly dysregulated genes, the profile of dysregulation between dysplasia and left colitis (E2) and extensive colitis or pancolitis (E3) shows differential patterns. With a significant percentage of dysregulated genes in common between dysplasia and both forms of UC, pancolitis has fewer differentially expressed genes with respect to dysplasia than left UC, which is consistent with those observed in principal component analysis. This pattern can also be observed in the focused analysis of glycogenes, and notably, no glycens going in one direction in the dysplasia/CU-I comparison go in the

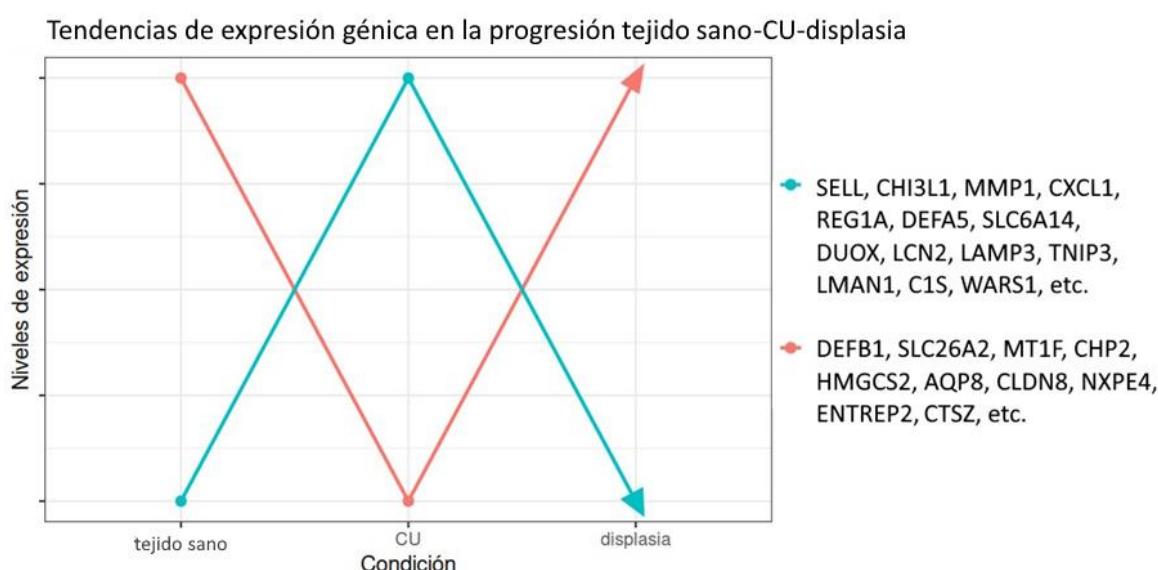
opposite direction in dysplasia/pancolitis, so it could be inferred that the glycobiological processes are gradual and consistent between E2, E3 and dysplasia (**Fig. 20**).



**Figure 20.** Venn diagrams indicating the number of significantly dysregulated genes ( $p < 0.05$ ) between dysplasia/CU-I and dysplasia/pancolitis. On the left are the total genes and on the right the glycogenes. In cases where the arrow points up, it means that the corresponding genes are increased in dysplasia, and the opposite is true for the down arrow.

It is notable that in a significant number of genes, when comparing dysplasia against UC (regardless of its classification in E2 or E3), the dysregulation patterns are opposite to those observed for UC/healthy tissue (**Section 5.3.1.1, of Results**). That is, they increase their expression in UC compared to biopsy of healthy intestine, and a decrease in UC towards dysplasia is observed, or vice versa. Primarily, this phenomenon is observed in genes related to the immune response that first increase and then decrease, such as *SELL* (a key lectin in leukocyte migration), *CHI3L1* (a lectin involved in inflammation), and gene families such as metalloproteases (*MMP1*, *MMP3*, *MMP10*), chemokines (*CXCL1*, *CXCL3*, *CXCL6*, *CXCL8*, *CXCL9*, *CXCL11*), members of the *REG family* (*REG1A*, *REG1B*) and defensins (*DEFA5*, *DEFB4A*). It should be noted that *DEFB1* is an exception in the defencin family, as it decreases in the transition from control to UC and increases in dysplasia. This is explained because, while defensins such as *DEFA5* can be regulated by inflammatory cytokines, *DEFB1* cannot, remaining constant, showed an opposite regulation pattern (low in inflammation and up in dysplasia). Other genes that increase in left UC and decrease in dysplasia are: *SLC6A14*, a solute transporter and genes involved in epithelial defense; *DUOX* and *DUOXA2*, which regulate hydrogen peroxide synthesis; *LCN2*, which inhibits bacterial growth; *LAMP3*, lysosomal

glycoprotein linked to apoptosis (Tanaka *et al.*, 2022); and *TNIP3*, a protein linked to inflammatory bowel disease that regulates the cellular response to lipopolysaccharides (Yang *et al.*, 2024). On the contrary, genes that decrease their expression in left UC, but increase in dysplasia are mostly related to ion transport (e.g. *SLC26A2*), metabolism (e.g. *HMGCS2*) and the structure of membranes and their junctions, such as *AQP8* and *CLDN8*, claudin component of tight junctions (Zhu *et al.*, 2019). The gene encoding an inflammation-associated lysosomal protein, cathepsin Z (*CTSZ*), is also increased (Campden *et al.*, 2022) (**Fig. 21**).



**Figure 21. Summary diagram in the variation of the expression of two groups of genes in the comparison of healthy tissue-left UC-dysplasia.** By way of illustration, the evolution of some genes is shown, which, according to the results of the analysis of this research, increase their expression of healthy tissue to ulcerative colitis and then decrease in dysplasia, or vice versa. The x-axis represents the phenotype or biological condition and the y-axis represents an approximation of the general expression levels of the genes involved. The different strict variations in expression between healthy tissue and dysplasia are not considered, but the general trends of the gene groups from the control group to UC and from the UC group to dysplasia, considering the progression observed in patients.

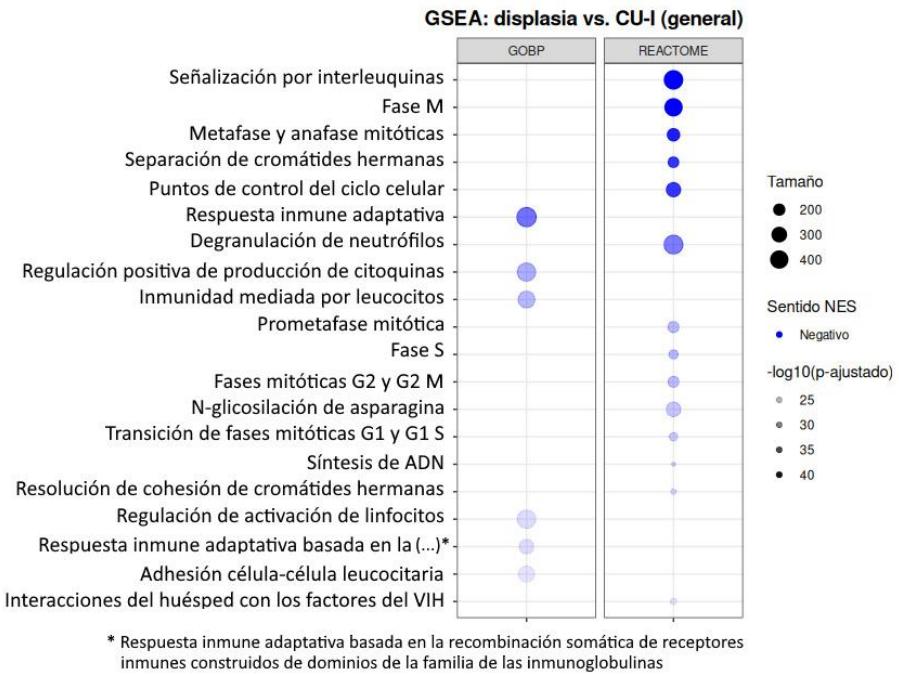
Regarding the functional enrichment analysis obtained using the GSEA method, the vast majority of pathways are negatively enriched in dysplasia and are usually similar when compared with both types of IBD, either left UC or pancolitis (**see Table XII**). The enriched pathways are related to the activation of the immune response, including to biological stimuli such as bacteria and viruses, cell division, regulation of apoptotic signaling and the localization of proteins in the nucleus.

**TABLE XII.** Most significant terms of the functional enrichment analysis using the GSEA method for the comparison of dysplasia with left UC or pancolitis. The p-value corresponds to the adjusted value.

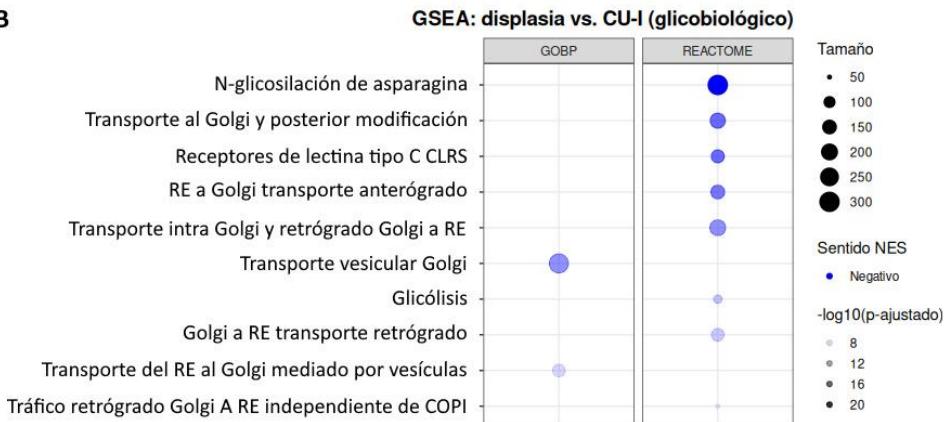
Displasia/CU-I			
Vía	Conjunto	NES	P-Valor
Fase M	Reactome	-2.82	6.25E-43
Señalización por interleucinas	Reactome	-2.75	6.25E-43
Metafase y anafase mitóticas	Reactome	-3.03	3.60E-40
Separación de cromátides hermanas	Reactome	-3.03	8.76E-38
Puntos de control del ciclo celular	Reactome	-2.90	8.76E-38
Respuesta inmune adaptativa	GO: BP	-2.54	4.69E-32
Degranulación de neutrófilos	Reactome	-2.50	4.97E-31
Regulación positiva de producción de citoquinas	GO: BP	-2.46	2.12E-26
Inmunidad mediada por leucocitos	GO: BP	-2.52	2.11E-25
Prometafase mitótica	Reactome	-2.76	3.12E-25
Fase S	Reactome	-2.82	3.31E-25
Fases mitóticas G2 M	Reactome	-2.76	4.91E-25
N-glicosilación de asparagina	Reactome	-2.55	4.67E-24
Fase mitótica G1 y transición G1 S	Reactome	-2.79	1.84E-23
Síntesis de ADN	Reactome	-2.85	1.88E-23
Resolución de cohesión de cromátides hermanas	Reactome	-2.83	1.76E-22
Displasia/Pancolitis			
Vía	Conjunto	NES	P-Valor
Señalización por interleucinas	Reactome	-2.82	1.67E-45
Fase M	Reactome	-2.67	8.50E-35
Puntos de control del ciclo celular	Reactome	-2.80	7.26E-34
Respuesta inmune adaptativa	GO: BP	-2.52	7.24E-32
Metafase y anafase mitóticas	Reactome	-2.85	1.21E-31
Degranulación de neutrófilos	Reactome	-2.46	8.58E-30
Separación de cromátides hermanas	Reactome	-2.89	5.58E-29
Regulación positiva de producción de citocinas	GO: BP	-2.50	1.14E-28
Inmunidad mediada por leucocitos	GO: BP	-2.54	1.89E-27
Adhesión célula-célula leucocitaria	GO: BP	-2.54	1.00E-25
Regulación de activación de linfocitos	GO: BP	-2.43	1.00E-25
Respuesta inmune adaptativa basada en la recombinación somática de receptores inmunes construidos de dominios de la superfamilia de las inmunoglobulinas	GO: BP	-2.59	3.93E-25
Migración de leucocitos	GO: BP	-2.49	1.21E-24
Vía de señalización mediada por citoquinas	GO: BP	-2.38	1.21E-24
Inmunidad mediada por linfocitos	GO: BP	-2.57	1.76E-24
Activación de células T	GO: BP	-2.29	7.23E-23

In specific relation to glycobiology, pathways related to vesicle-mediated transport, glycosylation, particularly N-glycosylation, and glycoprotein metabolism are negatively enriched in dysplasia. In particular, all these dysregulations have high significance in the comparison of dysplasia vs. left UC (**Fig. 22**), implying that glycobiological pathways, beyond the dysregulation of specific genes, are consistently affected in the transition to healthy tissue, → ulcerative colitis, E2→, E3→, dysplasia.

A



B



**Figure 22. Functional enrichment analysis in the dysplasia/UC-I comparison.** The bubble plot top panel (A) Shows Terms GO: BP The most significant general results obtained by GSEA for the comparison of dysplasia vs. UC-I, and the lower (B) The Terms GO: BP particularly related to glycobiology for the same comparison. The rows show the pathways composed of more than 5 and less than 500 genes that showed the most statistically significant differences ( $p < 0.05$ ). The color of the circle, always blue, denotes that the main pathways were found to be inhibited in the dysplasia group. The intensity of the color shows the value of the adjusted p-value transformed into logarithm in base 10, and the size of the circle denotes the number of genes included in each pathway.

### 5.3.3. Analysis of biopsies of patients with colitis-associated colorectal cancer and biopsies of patients with ulcerative colitis

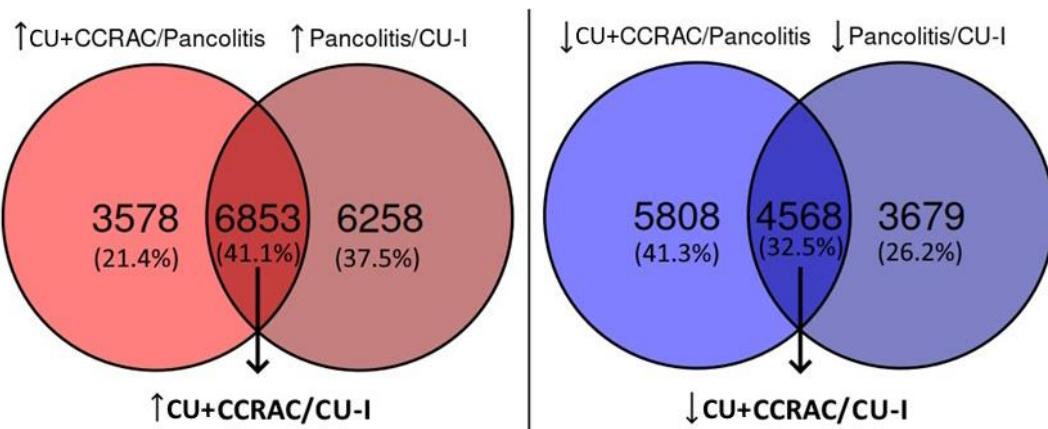
To explore transcriptomic changes associated with tumor development from left UC (UC-I), two complementary cohorts were used: the GSE3629 cohort, which includes samples

of UC + CCRAC and pancolitis, and the GSE47908 cohort, which contains samples of pancolitis and left UC. To this end, we leverage the concept of *Cumulative Inflammatory Burden* (CIB) as a driver of tumorigenesis: the cumulative effect of repeated inflammatory assaults significantly increases the risk of developing neoplasm in UC (Yalchin *et al.*, 2021). Because UC begins in the rectum and extends to the entire colon continuously (Lynch & Hsu, 2023), the colonic mucosa of patients with pancolitis (E3) is subjected on average to more inflammatory cycles than that of patients with left ulcerative colitis (E2). In this sense, the differential expression analysis between pancolitis and UC-I allowed us to analyze the molecular changes associated with the extension of the disease, and the comparison between UC+CCRAC and pancolitis, the possible alterations typical of the presence of tumor in a context of high inflammatory load.

In this context, four sets of dysregulated genes are defined:

1. Genes overexpressed in CU+CCRAC with respect to pancolitis
2. Genes underexpressed in CU+CCRAC with respect to pancolitis
3. Genes overexpressed in pancolitis compared to left UC
4. Genes underexpressed in pancolitis compared to left UC

The intersections between the groups with the same pattern of regulation (i.e., overexpression or underexpression) in both comparisons indicate genes that are dysregulated by the presence of the tumor regardless of the type of UC (**Fig. 23**). These genes are relevant because they could be defining a signature of gene dysregulation that appears early and is associated with the presence of a tumor. For these comparisons, 6853 were found to be overexpressed, 201 of them being glycogenes; and 4568 underexpressed, of which 270 were glycogenes.

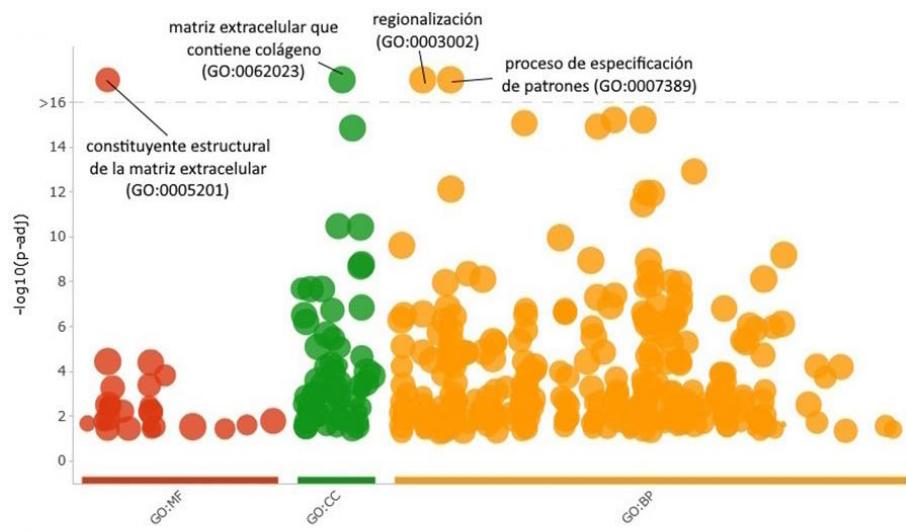


**Figure 23.** Venn diagrams indicating the number of dysregulated genes among patients with UC+CCRAC and pancolitis or pancolitis and UC-I, with their intersections. On the left, the genes increased in both comparisons are indicated in red, and on the right the decreased genes, in blue. The intersections between the sets with the same pattern of regulation (i.e., overexpression or underexpression) in both comparisons correspond to the differentially expressed genes between CU+CCRAC and CU-I.

The functional enrichment analysis of these genes in common was performed using the functional overrepresentation method for each group. The genes increased in CU+CCRAC compared to CU-I included *DEFA5*, defensin that presented an increase in UC compared to controls and a decrease in dysplasia vs pancolitis, and two galectins of particular relevance for our laboratory, *LGALS4* and *LGALS12* (PhD thesis by Mora Massaro and Sebastián Maller, respectively) (Maller, 2019; Maller *et al.*, 2020; Massaro, 2025). Although *LGALS4* had previously appeared among our results, increased in dysplasia vs colitis (regardless of the extent) and decreased between colitis and healthy tissue, *LGALS12* not. The functional analysis resulted in 367 activated terms. They are mainly associated with the organization and composition of the extracellular matrix, reactivation of embryonic pathways and development, binding between cells and membrane potential. Decreased genes in CU+CCRAC versus CU-I—which include *ST3GAL1*, increased between CU and healthy tissue—resulted in a total of 739 negatively enriched GO terms, mainly related to the immune response, but also including localization in the chromosomal region, and segregation of sister chromatids (see Table XIII; Fig. 24).

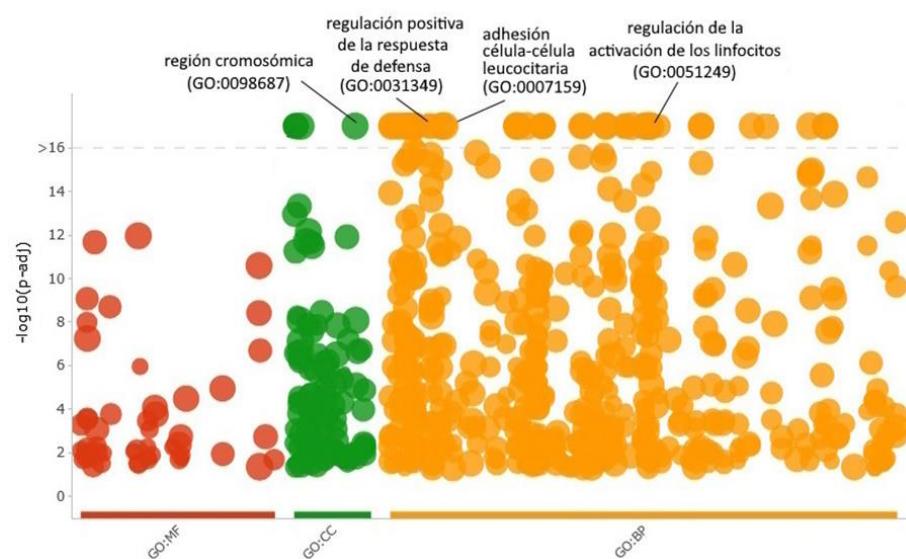
**A**

Sobrerrepresentación funcional de genes aumentados en CU+CCRAC vs CU-Izquierda



**B**

Sobrerrepresentación funcional de genes disminuidos en CU+CCRAC vs CU-Izquierda



**Figure 24. Result of overrepresentation analysis in CU+CCRAC/CU-Left.** The bubble charts visually represent the results of the functional enrichment analysis, showing the significance of the enriched terms on the axis  $y$  such as  $-\log_{10}(p\text{-adjusted value})$ , so that the most significant values are higher, proportionally, until they cross a threshold of very high significance (gray dotted line). Each point represents a *pathway*, and the size of the bubble represents the number of genes included in it (only those of less than or equal to 500 are considered). The color of the bubble represents whether the *pathway* belongs to the set GO: MF (in red), GO: CC (in green) or GO: BP (in yellow). **A)** Representation of overrepresented pathways in genes augmented in CU+CCRAC/CU-I. **B)** Representation of overrepresented pathways in genes decreased in CU+CCRAC/CU-I.

**UADE** SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía

---

**TABLE XIII. Augmented and decreased terms in the analysis of functional overrepresentation of dysregulated genes between CU+CCRAC/CU-I.** The p-value corresponds to the adjusted value.

Términos aumentados en CU+CCRAC/CU-I						
ID	Fuente	Término	P-Valor	Precisión	Recall	
GO:0062023	GO:CC	Matriz celular que contiene colágeno	3.50E-23	0.038	0.468	
GO:0007389	GO:BP	Proceso de especificación de patrones	1.45E-17	0.041	0.429	
GO:0005201	GO:MF	Constituyente estructural de la matriz extracelular	2.48E-17	0.019	0.593	
GO:0003002	GO:BP	Regionalización	2.55E-17	0.038	0.438	
GO:0048568	GO:BP	Desarrollo de órganos embrionarios	6.12E-16	0.039	0.423	
GO:0045229	GO:BP	Organización de la estructura encapsuladora externa	6.16E-16	0.030	0.462	
GO:0030198	GO:BP	Organización de la matriz extracelular	8.60E-16	0.030	0.462	
GO:0043062	GO:BP	Organización de la estructura extracelular	1.22E-15	0.030	0.460	
GO:0097060	GO:CC	Membrana sináptica	1.41E-15	0.033	0.433	
GO:0061448	GO:BP	Desarrollo de tejido conectivo	1.22E-13	0.027	0.459	
GO:0007409	GO:BP	Axonogénesis	7.42E-13	0.038	0.405	
GO:0048705	GO:BP	Morfogénesis del sistema esquelético	1.10E-12	0.023	0.479	
GO:0050808	GO:BP	Organización sináptica	1.18E-12	0.039	0.400	
GO:0048562	GO:BP	Morfogénesis de órganos embrionarios	3.51E-12	0.027	0.443	
GO:0045211	GO:CC	Membrana post-sináptica	3.35E-11	0.023	0.439	
GO:0098978	GO:CC	Sinapsis glutamatérgica	3.65E-11	0.032	0.398	
GO:0034329	GO:BP	Ensamblado de unión celular	1.11E-10	0.036	0.394	
GO:0001654	GO:BP	Desarrollo del ojo	2.44E-10	0.033	0.401	
GO:0150063	GO:BP	Desarrollo del sistema visual	6.51E-10	0.033	0.397	
GO:0042391	GO:BP	Regulación del potencial de membrana	1.14E-09	0.034	0.390	
Términos inhibidos en CU+CCRAC/CU-I						
ID	Fuente	Término	P-Valor	Precisión	Recall	
GO:0098687	GO:CC	Región cromosómica	8.09E-39	0.048	0.469	
GO:0051249	GO:BP	Regulación de la activación de los linfocitos	1.69E-37	0.057	0.439	
GO:0002757	GO:BP	Vía de señalización activadora de la respuesta inmunitaria	1.84E-37	0.055	0.444	
GO:0030098	GO:BP	Diferenciación de linfocitos	1.04E-36	0.052	0.454	
GO:0007059	GO:BP	Segregación de cromátides hermanas	1.50E-34	0.050	0.450	
GO:0042113	GO:BP	Activación de células B	1.43E-31	0.037	0.504	
GO:0051251	GO:BP	Regulación positiva de la activación de los linfocitos	4.82E-30	0.040	0.472	
GO:0002768	GO:BP	Vía de señalización del receptor de superficie celular que regula la respuesta inmunitaria	5.57E-30	0.041	0.467	
GO:0002696	GO:BP	Regulación positiva de la activación de leucocitos	3.97E-29	0.042	0.454	
GO:0031349	GO:BP	Regulación positiva de la respuesta de defensa	9.32E-29	0.052	0.409	
GO:0002429	GO:BP	Vía de señalización del receptor de superficie celular activador de la respuesta inmunitaria	1.46E-28	0.038	0.474	
GO:0050867	GO:BP	Regulación positiva de la activación celular	1.75E-28	0.043	0.444	
GO:0000819	GO:BP	Segregación de cromátides hermanas	1.82E-28	0.032	0.522	
GO:0045088	GO:BP	Regulación de la respuesta inmune innata	1.82E-28	0.048	0.423	
GO:0007159	GO:BP	Adhesión célula-célula leucocitaria	1.10E-27	0.045	0.428	
GO:0030217	GO:BP	Diferenciación de células T	2.03E-27	0.038	0.465	
GO:0033044	GO:BP	Regulación de la organización cromosómica	2.56E-27	0.032	0.511	
GO:0002366	GO:BP	Activación de leucocitos implicada en la respuesta inmune	6.56E-27	0.037	0.466	
GO:0002697	GO:BP	Regulación del proceso efector inmunitario	1.93E-26	0.043	0.430	
GO:0070661	GO:BP	Proliferación de leucocitos	2.32E-26	0.041	0.442	

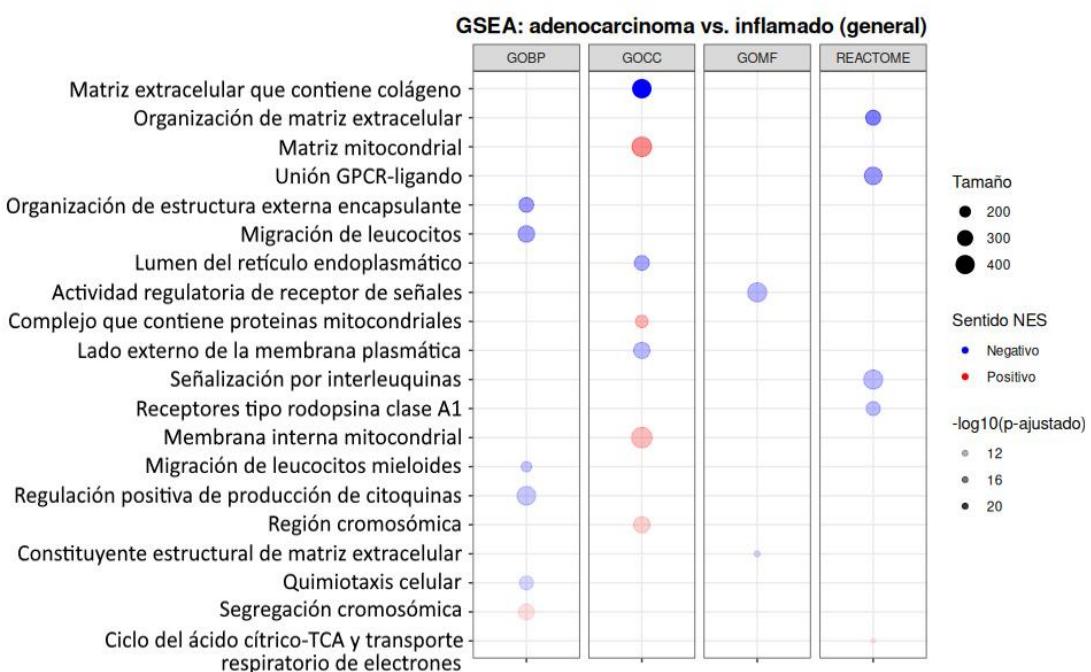
### **5.3.4. Study of Relevant Genetic Dysregulation in Human Pathologies with Transcriptomic Data from Murine Models**

To observe whether genes previously described as dysregulated early in UC (**Section 5.3.3, Results**) and correlating with tumor presence in patients are reproduced in murine models, publicly available data in the GSE31106 dataset corresponding to the experimental model of colitis-associated colorectal cancer (AOM-DSS, **Section 4.1.1, of Materials and Methods**) (Tang *et al.*, 2012).

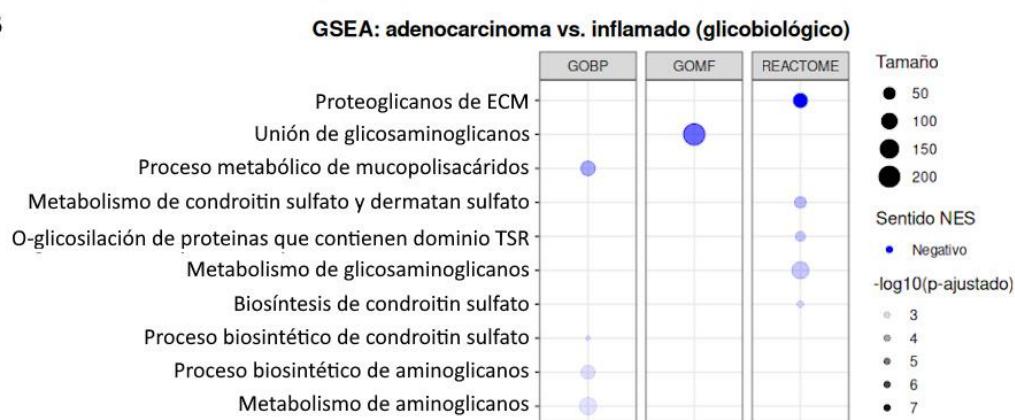
Functional enrichment analysis using the GSEA method in adenocarcinoma samples compared to inflamed tissue (week 20 vs week 2) revealed significant differences in general biological and glycobiological processes. Among the general enriched processes (**Fig. 25 A**), a negative enrichment of pathways related to leukocyte migration and activation, cytokine production, and cellular chemotaxis was observed in adenocarcinoma. In contrast, processes such as chromosomal segregation and extracellular matrix organization showed positive enrichment. Regarding glycobiological processes (**Fig. 25 B**), there was evidence of significant negative enrichment of pathways associated with the metabolism and biosynthesis of mucopolysaccharides and chondroitin sulfate, components of the extracellular matrix.

**UADE** SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía

**A**



**B**

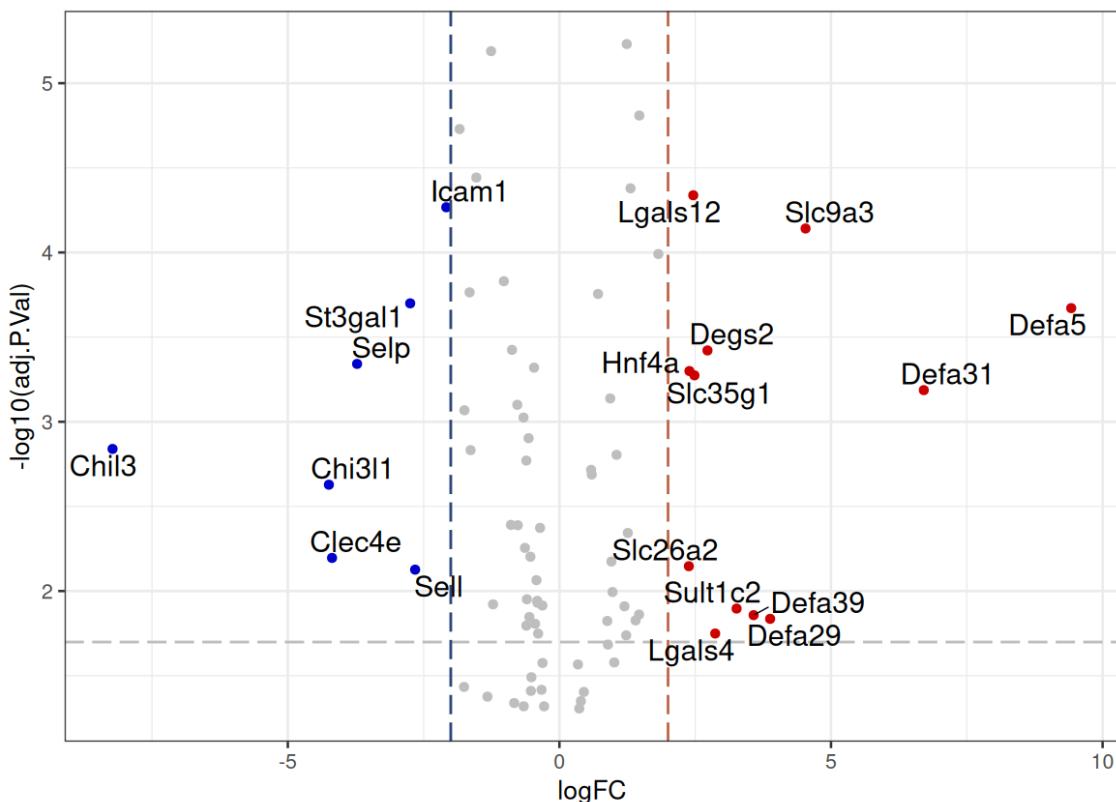


**Figure 25. Functional enrichment analysis in the murine model of AOM-DSS comparing adenocarcinoma (week 20)/inflamed tissue (week 2), with terms GO: BP resulting from the GSEA method.** The bubble plot displays the results of the GSEA analysis. The rows show the pathways composed of more than 5 and less than 500 genes that showed more statistically significant differences ( $p < 0.05$ ). The color of the circle denotes whether the pathway was found to be activated (red) or decreased (blue) in the adenocarcinoma group, the intensity of the color shows the value of the adjusted p-value transformed into logarithm in base 10, and the size of the circle denotes the number of genes included in each pathway. The different columns divide the roads according to their source dataset: GO: BP, GO: CC, GO: MF, or Reactome.

There are 1240 significant genes in the murine model that covary with the dysregulated genes in the CU+CCRAC/CU-I comparison obtained in **Section 5.2.3 of Results**, 639 of which are overexpressed in adenocarcinoma against inflamed tissue and 601 underexpressed. A total of 79 of these are glycogenes (**Annex 8.1**), of which 44 are underexpressed and 35

overexpressed, and all have a p-value adjusted  $< 0.05$ ; from now on we will call this group of genes "extended glycobiological signature" (**Annex 8.3**), are the data uploaded as an example for the user in the Shiny interactive application (**Section 5.2, of Results**).

These glycobiological genes are then filtered (p-value adjusted  $< 0.02$ ,  $|\log FC| > 2$ ) to obtain the most representative ones, which are shared between human pathologies and the experimental model of AOM-DSS (**Fig. 26**).



**Figure 26. Volcano plot of glycobiological genes common to human pathologies and the AOM-DSS model in the adenocarcinoma/inflamed tissue comparison.** The axis x denotes the  $\log_{2}FC$  of the comparison and the blue and red dotted lines are located at the values -2 and 2. The axis y corresponds to the logarithmic transformation in base 10 of the adjusted p-value and the gray dotted line denotes the cut-off point of statistical significance ( $p < 0.02$ ). Differentially overexpressed genes with a  $\log_{2}FC$  greater than 2 are marked as red dots with their names, and those underexpressed with a  $\log_{2}FC$  less than -2, are indicated in blue.

Finally, taking into account this last significance filter, those genes that will make up the final (or reduced) glycobiological signature are selected in the experimental model of adenocarcinoma with respect to inflamed tissue, based on their biological relevance and our analyses in patients. This signature includes 10 genes, the 4 augmented being: a defensine

(*Defa5*), 2 galectins (*Lgals4* and *Lgals12*), and a sulfate transporter (*Slc26a2*). On the other hand, the 6 decreased genes are: three adhesion molecules (the selectins *Sell* and *Selp*, and *Icam1*), a sialyltransferase (*St3gal1*), a chitinase (*Chil3*) and a pseudochitinase (*Chi3l1*). (see **Table XIV**).

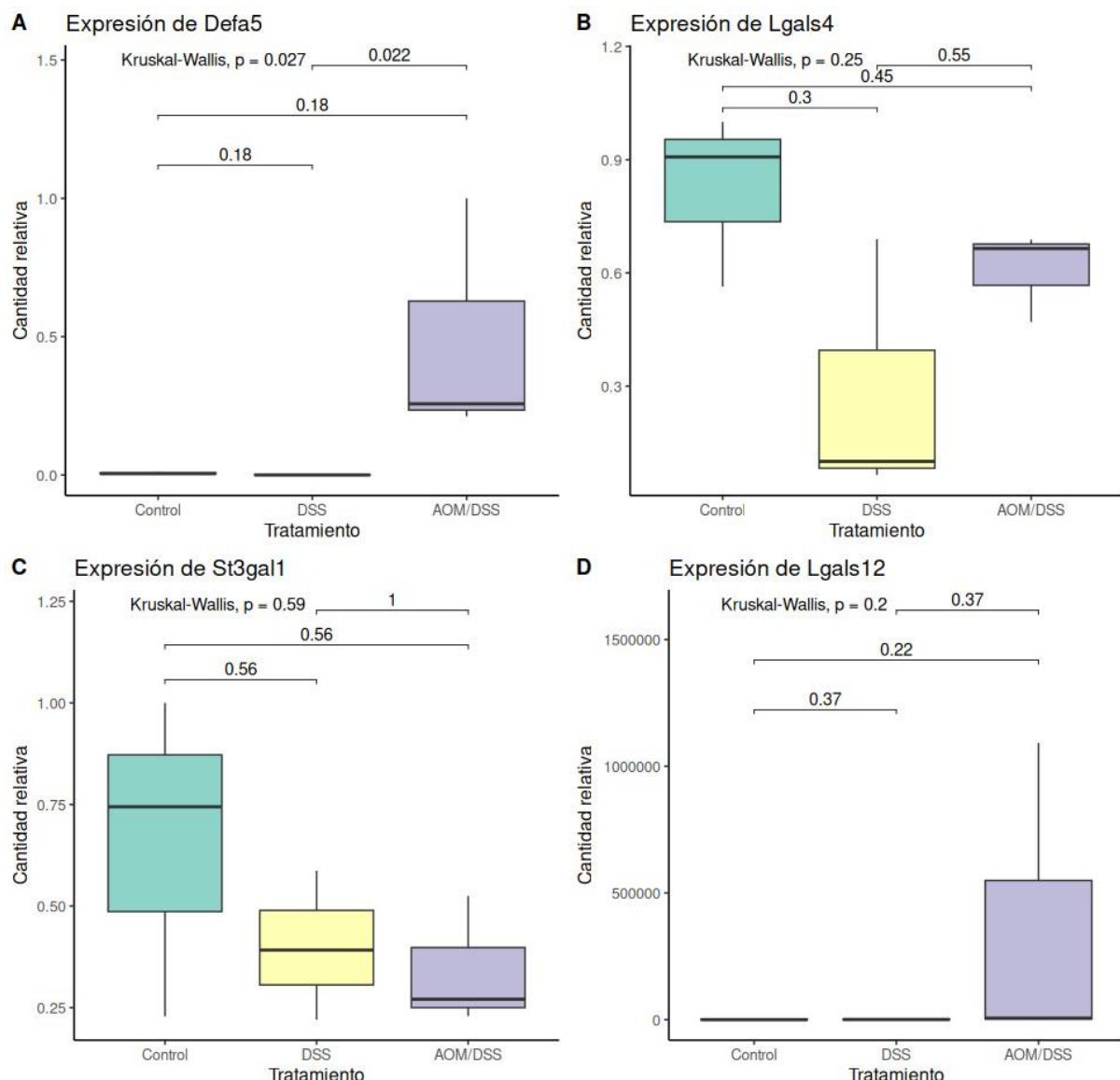
**TABLE XIV. Glycobiological signature for the transition from intestinal inflammation to neoplasia.** In pink, the cells of the genes that are increased in the adenocarcinoma model compared to the UC model are marked, and in blue, those that are decreased.

Ratón	Ortólogo humano	Nombre del gen	Sentido
<i>Defa5</i>	DEFA5	Defensina 5	↑
<i>Lgals4</i>	LGALS4	Galectina 4	↑
<i>Lgals12</i>	LGALS12	Galectina 12	↑
<i>Slc26a2</i>	SLC26A2	Miembro 2 de la familia de portadores de solutos 26	↑
<i>Icam1</i>	ICAM1	Molécula de adhesión intercelular 2	↓
<i>Sell</i>	SELL	Selectina L	↓
<i>Selp</i>	SELP	Selectina P	↓
<i>St3gal1</i>	ST3GAL1	ST3 Beta-galactósido alfa-2,3-sialiltransferasa 1	↓
<i>Chi3l1</i>	CHI3L1	Proteína 1 similar a la quitinasa 3	↓
<i>Chil3</i>	CHIA	Quitinasa ácida	↓

#### 5.4. Validation of results by RT-qPCR

To validate the results obtained *in silico*, RNA samples from the intestine of control female mice, undergoing acute colitis (inflammatory model, DSS) or undergoing colitis-associated colorectal cancer (AOM-DSS) were used. The details of the experimental groups can be found in **Materials and Methods, Section 4.2.1**. The genes chosen for *in vivo* validation were 4: *Defa5*, *St3gal1*, *Lgals4*, and *Lgals12*.

After reverse transcription and amplification according to the details in materials and methods (**Section 4.2.2, Materials and Methods**), normalized average expression values were obtained for each biological group according to each target gene (**Fig. 27**). Detailed amplification results are included in **Annex 8.4**.



**Figure 27. Expression results of *Defa5* (A), *Lgals4* (B), *St3gal1* (C) and *Lgals12* (D) in RT-qPCR.** Box plots representing the average normalized relative expression amount of each gene (axis y) for each biological group (axis x). The median is represented by a horizontal line across each box, the extreme values are represented by the vertical line across each box, the upper edge of the box corresponds to the third quartile of the ordered data set and the lower one to the first. For each target, the adjusted p-values resulting from the nonparametric Kruskal-Wallis test are reported, where the only significant result is in *Defa5* ( $p = 0.027$ ). For each pairwise comparison between treatments, the adjusted p-values resulting from the test are reported Post-hoc Dunn's method, where the only significant result is between DSS and AOM-DSS for *Defa5* ( $p = 0.02$ ).

The results indicate that the expression of *Defa5* increases significantly in AOM-DSS with respect to DSS, in accordance with what was anticipated by the analysis carried out in public databases of the AOM-DSS model (Section 5.3.4, of Results).

On the other hand, for the expression of *Lgals4*, *St3gall* and *Lgals12*, inconclusive results were obtained due to different technical difficulties: in the case of *St3gall*, there was a non-specific amplification. For *Lgals12*, out-of-scale values were obtained, with a negative value efficiency (-100) reported by the analytical software, indicating that it was not expressed in the samples, and a possible technical artifice as a result. Finally, in the case of *Lgals4* we believe that the result is not significant, but with an increase in n the difference between healthy and inflamed tissue could be significant, since other laboratory results obtained outside the framework of this Thesis show dysregulation in the same sense in DSS.

## 6. CONCLUSIONS AND DISCUSSION

### 6.1. Development of an online application and *R* package for transcriptomic data analysis

First, it was decided to establish a tool for the analysis and reporting of differential expression of transcriptomic data, which is reflected in an online application for visualization of differential expression results, and an *R* package to facilitate transcriptomic analysis. These developed tools combine open source software functionalities of scientific interest, such as the transcriptomic analysis packages available in Bioconductor (Huber *et al.*, 2015), as well as commercial platforms, such as *Thermo Fisher Scientific's Transcriptome Analysis Console (TAC)* (Thermo Fisher Scientific, 2020). CT is a graphical tool that allows analyzing gene expression data from microarrays, offering functions such as quality control, normalization, detection of differentially expressed genes, alternative *splicing* analysis, and analysis of biological annotations and interactive visualization of results. In addition, for statistical analyses he uses the *limma package*, the same one I used in my research. Other similar tools are *Omics Playground*, for the exploration of proteomics and transcriptomics data without the need to know how to program (Akhmedov *et al.*, 2020), and *GIANT*, for transcriptomics data (Vandel *et al.*, 2020).

They allow access to these analyses to users, without the need for high programming knowledge, providing tools of free use for the laboratory and the community in general. In addition, the establishment of these tools opens the door to future developments with broader applications that improve the communication and analysis of differential expression analysis results, for microarrays and RNA-seq, including for *single-cell sequencing* sample analysis. In addition, the tools are not restrictive for use in glycobiology, nor for specific species such as humans or mice, but are adaptable to any study of differential expression.

### 6.2. Inflammation and Healthy Tissue: Glycobiological Differences

The analysis of the bulk transcriptome of biopsies from patients with active UC compared to those obtained from subjects without IBD carried out in this Thesis continued the line of work initiated with the Thesis of Román Lanzilotta (2022), and served to validate the

dysregulation observed in a second cohort. Through this study, we were able to describe a list of 31 genes associated with glycosylation directly or indirectly that, consistently and independently of inter-cohort variability (associated with genetic, dietary, treatment factors, among others) are dysregulated in active ulcerative colitis. The increased inflammatory response is evidenced by the overexpression of adhesion molecules involved in leukocyte migration, such as cell adhesion molecule 1 (*VCAM1*), and selectins (*SELL*, *SELP*, *SELE*) (Smith & Bertozzi, 2021b; 2021a; Singh *et al.*, 2023). It is also seen in the overexpression of *CHI3L1*, a lectin involved in inflammatory processes (Chen *et al.*, 2024), overexpression of members of the regulatory protein family (*REG4*, *REG3A*, *REG1A*, *REG1B*), and of defensin 5 (*DEFA5*), which acts as an antimicrobial peptide (Sabit *et al.*, 2024). Although an increase in the expression of collagen-degrading metalloproteases is observed in IBD, their activity is counteracted by the overexpression of Tissue Inhibitors of Metalloproteinases (TIMPs) (De Souza & Fiocchi, 2016). This dysregulation is reflected in our results, where both metalloproteases and *MMP9* and *TIMP1* are significantly overexpressed in CU compared to healthy tissue.

In addition, dysregulation of glycosylation is observed in the face of chronic inflammation, as indicated by the decrease in *B4GALNT2*, which encodes β-(1,4)-N-acetylgalactosaminyltransferase 2, involved in the biosynthesis of the Sda antigen, which is mainly expressed in healthy colonic mucosa (Groux-Degroote *et al.*, 2021). Another robustly underexpressed gene in CU is *SLC26A2*, which could be associated with reduced sulfation of intestinal mucins, altering the protective capacity of the intestinal barrier (Ohana *et al.*, 2012). In a recent study, *SLC26A2* showed a negative correlation with the IL-17 signaling pathway and a positive association with tight junctions, so its decline is consistent with abnormal immune cell infiltration and inflammatory lesions during the pathogenesis of UC (Qian *et al.*, 2025). Another increased gene is *ST3GAL1*, which encodes an α(2,3)-sialyltransferase, and whose dysregulation suggests a possible alteration in the sialome (Fan *et al.*, 2025).

### **6.3. Remission status**

Secondly, we decided to study the bulk transcriptome of biopsies from patients with quiescent ulcerative colitis (UC) with respect to active UC and also with respect to healthy

tissue, demonstrating that molecularly and glycobiologically, tissue in remission is not equivalent to healthy tissue. In this inflammatory pathology, which develops in an orderly pattern from the rectum to the rest of the colon, the intestinal tissue in a quiescent or inactive state presents persistent functional alterations resulting from the cycles of inflammation that reflect: a) chronic damage and b) incomplete regeneration. In this regard, permanent changes in the epithelium have already been described, including dysregulation in the expression of genes that affect its secretory functions (e.g., in goblet cells, resulting in abnormal mucin secretion and defective protective mucus), antimicrobial, and regeneration and scarring (Planell *et al.*, 2013). In addition, even in the absence of signs of active inflammation at the endoscopic/histological level, in quiescent UC an increased inflammatory response similar to those observed in active UC is observed, but generally with fewer times of change and significance, reflecting that the intestinal tissue in quiescent UC exhibits chronic remodeling, epithelial dysfunction and an altered immune response. with persistence of molecular indicators of the inflammatory process and response to pathogenic organisms. In this sense, when analyzing the immune infiltrate in quiescent UC, it was shown that monocytes, T cells, and B cells are decreased with respect to active inflammation, but the myeloid fraction (macrophages, eosinophils) remained unchanged (Fenton *et al.*, 2021) so that although the signs of inflammation persist, they are not of the same nature as in active inflammation. In this sense, in our analyses in CUq samples, the overexpression of *CCL11*, an eosinophil-recruiting chemokine (Polosukhina *et al.*, 2021). Activation of this signaling has already been reported in UC in remission, along with overexpression of related genes, such as *CCL22* (Fenton *et al.*, 2021).

Our results are also supported by a study highlighting the overexpression of olfactomedina 4 (*OLFM4*) (Fenton *et al.*, 2021). *OLFM4* positively modulates IL-22, a cytokine crucial for epithelial repair and defense in the gut (Xing *et al.*, 2024), and in turn, it has been characterized as a marker of a subpopulation of stem cells that have the capacity for self-renewal and are essential for the repair of the intestinal epithelium, so their overexpression in quiescent UC could be associated with epithelial regeneration circuits.

All these processes are reflected in the functional enrichment analysis (**Section 5.3.1, Results**), where even in quiescence, an increase in proteolytic activity is observed, a process

that can occur in response to IBD and worsen its pathogenesis (Hou *et al.*, 2024). Activation of pathways related to the immune response, such as neutrophil degranulation, and antigen-presenting activity by MHC class I is also seen, coinciding with the previously described immune response. Focusing on glycobiology, it seems important to us that in quiescent CU vesicular transport pathways associated with the Golgi apparatus and *N-glycosylation are enriched, suggesting alterations in glycoprotein production, which will surely deserve a later and more detailed study outside this Thesis.*

Interestingly, *REG4*, dysregulated in active CU, is also one of the most significantly overexpressed genes in CUq with respect to healthy mucosal tissue. The increase in the family of REGs, with a lectin-C-like domain, has been widely reported in IBD where they have anti-inflammatory and antimicrobial effects, and participate in tissue recombination, helping to maintain mucosal integrity (Sun *et al.*, 2021). Regeneration of colonic tissue in remission is essential to improve the long-term prognosis of ulcerative colitis (Bjerrum *et al.*, 2022), which gives great significance to REG4's activity. However, due to their proliferative effects, proteins of the REGs family could have a pro-tumor role, as they are overexpressed in colorectal cancer and colorectal cancer cell lines (Takasawa *et al.*, 2018).

In summary, and considering that the samples analyzed in this Thesis correspond to patients in histological remission and without active inflammation (Pekow *et al.*, 2013), the mucosa in quiescent UC presents a particular condition, with separate characteristics of both patients with active UC and healthy patients.

#### **6.4. Extent of inflammation and dysplasia: molecular mechanisms**

The extent of ulcerative colitis is a key factor in the choice of treatment and the prognosis of long-term clinical outcome, including the risk of developing high-grade dysplasia and colorectal cancer (Roda *et al.*, 2017). Since UC is a pathology that starts from the rectum and evolves and progresses to the entire colon continuously, the extension and time of evolution are correlative, something that does not necessarily happen in Crohn's disease, where inflammation is characterized by inflamed areas that do not follow a pattern between inflammatory cycles (Kaz and Venu, 2025). Among the main risk factors for developing colitis-

associated colorectal cancer (CCRAC) are extensive colitis and a long duration of the disease; this could reflect the cumulative impact of chronic inflammation, which, associated with mutagenesis, remission-relapse cycles, and epithelial regeneration after injury, can exert selective pressure that accelerates carcinogenesis (Porter *et al.*, 2021). In accordance with all this information, our results indicate that the tissue of patients with pancolitis, being exposed to a greater number of inflammatory cycles, has a transcriptomic profile more similar to that of dysplasia, which indicates a greater pre-malignancy, compared to left ulcerative colitis.

Dysplasia is a non-invasive neoplastic cell proliferation, that is, it is limited by the basement membrane of the epithelium, and can appear before or accompany an invasive neoplasm. Despite often being called a " premalignant lesion", and indicating a high risk of developing cancer, there is no definitive evidence that dysplasia is always the precursor lesion to the tumor: the frequency with which they occur together could simply reflect an individual's predisposition to develop neoplasia, for genetic and environmental reasons (Salas Caudevilla, 2007). In addition, the progression from inflammation to low-grade dysplasia to high-grade dysplasia, and finally to colorectal cancer, is not always fulfilled, but a tumor can develop from low-grade dysplasia (Zhou *et al.*, 2023b).

However, there is evidence that carcinogenesis transforms the epithelium and its microenvironment early into the " precancerous stage," where mechanisms of microenvironment reorganization and immune escape can be observed prior to the development of carcinoma (Angelova *et al.*, 2021; Goto *et al.*, 2025). Our results are consistent with these findings, because the analysis of functional enrichment in dysplasia shows negatively enriched immune cell activation and response pathways to bacterial stimuli. In dysplasia, and at the glycobiological level, a decrease in the expression of glycobiological genes such as *CHI3L1* (lectin involved in inflammation), *SELL* (lymphocyte adhesion molecule), members of the REG family (*REG1A*, *REG1B*) and *DEFA5* (antimicrobial molecule, which acts in response to dysbiosis) is observed. Another commonality between dysplasia and cancer is dysregulated cell proliferation, which we could infer from functional enrichment analysis, which resulted in negative enrichment of pathways related to cell cycle checkpoints. In this sense, previous genetic studies indicate that most dysplastic lesions associated with colitis have mutated tumor suppressor gene, *P53*, so that it loses its function controlling the cell cycle. In addition,

chromosomal instability is present in low-grade dysplasia, which progresses more in high-grade dysplasia (Baars *et al.*, 2024).

Chronic intestinal inflammation favors oxidative stress, and the sustained release of pro-inflammatory cytokines (such as IL-6 and TNF- $\alpha$ ), growth factors (VEGF), which triggers cycles of tissue damage and repair, which increase the probability of errors during cell division, causing the proliferation of aberrant cells, increasing the risk of dysplasia and colon cancer (Otálora-Otálora *et al.*, 2024). In this sense, the regeneration of the intestinal epithelium (*mucosal healing*) after inflammatory intestinal lesions implies a proliferation of exacerbated epithelial cells mediated by the STAT3, STAT5 and IL-22 pathways, which could, together with the appearance of driver mutations and an inhibition of the apoptotic pathways, trigger a neoplastic process (Villablanca *et al.*, 2022). Following this line of thought, it has been shown that the tissue of patients with active UC who would later develop dysplasia, had increased expression of the active form of STAT3 (P-STAT3), suggesting that the activation of JAK-STAT3 (a key pathway in cell proliferation) indicates a pre-dysplastic inflammatory state. In IBD, P-STAT3 may play a role in cell proliferation in dysplasia, and help predict dysplasia development. However, and interestingly, once dysplasia is established, P-STAT3 levels drop relative to inflamed tissue. This could be because, as dysplasia progresses, the epithelium becomes less susceptible to inflammatory signals, such as IL-6, which activate the JAK-STAT3 pathway (Baars *et al.*, 2024).

Interestingly, in our analysis we saw many negatively enriched pathways related to cell division in dysplasia versus UC, which at first glance is paradoxical. However, we must not forget that we are analyzing bulk biopsies, so multiple cell types are included in the samples, among which we cannot discern. For this reason, although *a priori* we would expect to see greater cell proliferation in dysplasia samples, it is possible that at the bulk transcriptomic level this division is not as marked as the regenerative proliferation driven by tissue damage in UC (*wound healing*), or is masked by the effect of decreased immune infiltration. In addition, in the inflammatory context there is a high rate of cell death, and the cells that survive could "compensate" by trying to divide more, but the resulting final rate is negative. In this sense, it would be of interest to carry out *single-cell* RNA-seq research that analyzes gene expression in individual cells.

## **6.5. From inflammation to malignant transformation: analysis of patient databases and an experimental model of colitis-associated colorectal cancer**

Field cancerization, as explained above, implies that areas of the intestinal mucosa that are apparently unmalignant and far from the tumor present molecular and genetic changes due to repeated aggressions (such as inflammation). It has been shown that genetic instability is even observed in these areas, including mutations in genes such as p53, and epigenetic alterations (Patel, 2015).

In summary, field cancerization reflects a process of diffuse molecular damage that is not focused exclusively on the tumor, explaining the appearance of new independent tumors. In our functional enrichment analysis of the CU+CCRAC/left-left UC comparison in humans (**Section 5.3.3, Results**), notably and even with a remote neoplasm, an inhibition of immune pathways consistent with an evasion of the immune response promoted by the tumor microenvironment was observed. In turn, an activation of processes related to the remodeling of the extracellular matrix and connective tissue is observed, which would facilitate tumor invasion, metastasis and angiogenesis (Goto *et al.*, 2025). It is important to note that, in the absence of single-cohort data, an indirect analysis was chosen, comparing UC+CCRAC vs. pancolitis samples, on the one hand, and left UC samples against pancolitis on the other in another cohort, which *a priori* is a limitation in this study.

The second conclusion we draw from these studies is that when we compare CU+ remote neoplasia vs healthy tissue, as expected, the dysregulations are many and varied, but when we compare dysregulated genes with quiescent UC, not only is the number lower (reinforcing that the quiescent mucosa is different from normal, and that there are already activated programs, as described above) but remarkably, 60 genes are differentially and specifically dysregulated between UC with remote neoplasia and quiescent UC. These genes provide a more refined and clinically relevant molecular profile to prevent the appearance of secondary neoplasms in UC. In this sense, among these genes is the one that codes for a glycosyltransferase, *GCNT2*, involved in the branching into chains of oligosaccharides of glycoproteins and glycolipids. The levels of messenger RNA for this enzyme have already been

described as decreased in colorectal cancer, which was related to a process of epigenetic regulation of expression (Nakamura *et al.*, 2015).

## **6.6. Glycobiological signature potentially predictive of the transition from ulcerative colitis to colitis-associated colorectal cancer**

Based on the covariant gene dysregulation in CU+CCRAC vs UC-I in human pathologies, and in the absence of a longitudinal study in patients that allows studying tumor development in the context of UC, we sought to investigate the transcriptomics of samples obtained from public data from a model of colitis-associated colorectal cancer (AOM-DSS) (samples collected in week 20 of AOM-DSS administration) against inflamed tissue (samples collected in week 2, with acute inflammation). Thus, a relevant glycobiological signature was extracted whose mechanisms could be studied in this experimental model (**Section 5.3.4, of Results**). This firm is made up of: *Defa5*, *Slc26a2*, *Lgals4* and *Lgals12* (augmented, 4), and *Sell*, *Selp*, *Chi3l1*, *Chil3*, *Icam1*, and *St3gall* (diminished, 6). Of these genes, he decided to validate 4 *in vivo* by RT-qPCR using samples of inflammation models (DSS) and adenocarcinoma models (AOM-DSS): *Defa5*, *St3gall*, *Lgals4* and *Lgals12*. Although we have validated the dysregulation of DEFA5 significantly and the rest of the genes did not present significance in the experiments of this Thesis, we describe below the relevance of this signature in the context of intestinal inflammation and colorectal cancer associated with colitis.

### **6.6.1. *Defa5***

Defensins are antimicrobial peptides of innate and adaptive immunity that can defend the body from bacteria, fungi, and viruses, which are highly conserved in plants, insects, and mammals (Gao *et al.*, 2021). One way in which they exert their antimicrobial action is through their high proportion of positively charged residues, which allows them to interact with the negatively charged components of the membrane of gram-negative bacteria, penetrating the barrier and breaking into membrane phospholipids via amphipathic peptide domains (Fu *et al.*, 2023). This family of proteins plays an essential role in the pathogenesis of IBD, counteracting bacterial attack caused by dysbiosis (Ramasundara *et al.*, 2009) or through its

immunomodulatory role, regulating the function of neutrophils, monocytes, T lymphocytes, and dendritic and epithelial cells (Shi, 2007).

In this paper we focus on defensin-5 (*DEFA5*), an intestinal  $\alpha$ -defensin produced by Paneth cells, due to its relevance throughout all analyses and in the literature related to our topic of study. In particular, *DEFA5*, which is not expressed in normal colonic mucosa (Ramasundara *et al.*, 2009), is upregulated in response to proinflammatory cytokines (Fu *et al.*, 2023) and was proposed as a biomarker in UC to differentiate it from Crohn's disease (Williams *et al.*, 2017).

In addition, increased *DEFA5* transcript has been shown in sporadic colorectal cancer, and is associated with a favorable prognosis, which was related to its inhibitory effect on tumor cell growth without affecting normal cells, which is why it could act as a tumor suppressor (Zhao *et al.*, 2023b). It has been proposed that the tumor suppressive role of defensin-5 could be exerted by binding to phosphatidylinositol 3-kinase (PI3K) subunits, blocking its signaling pathway, and delaying cell growth and metastasis. In addition, overexpression of *DEFA5* was shown to suppress tumor growth in *in vivo* experiments of nude mice with xenografts and colon cancer cell lines (Qiao *et al.*, 2021; Sabit *et al.*, 2024). In patients with sporadic colorectal cancer, it was shown that the increase in mRNA for *DEFA5* correlates with longer survival in these patients (*Hazard Ratio* = 0.62, *P* = 0.02) (Zhao *et al.*, 2023b), but there are no studies in patients with CCRAC. In a mouse model of colitis-associated colorectal cancer, *DEFA5* overexpression was also observed, induced by treatment with AOM-DSS (Gobert *et al.*, 2022b; 2022a).

Our *in silico* results suggest that, in the transition between healthy tissue-inflammation-mouse adenocarcinoma, *Defa5* expression increases in inflammation and does so again in malignant tissue, in agreement with the literature (Zhao *et al.*, 2023b; Fu *et al.*, 2023). In RT-qPCR, we validated these results with significant results, suggesting that *Defa5* could be regulated in a similar way in both sporadic and colitis-associated colorectal cancer. Meanwhile, and considering the differences between CCRE and CCRAC, additional studies are required. It would be of interest to investigate whether its potential role as a tumor suppressor and the longer survival in patients with high expression of *DEFA5* is maintained in patients with CCRAC.

### 6.6.2. Galectins: *Lgals4* and *Lgals12*

Galectins are glycan-binding proteins that play a crucial role in various physiological and pathological processes; given their multiple functions in tumor development, they have been proposed as possible therapeutic targets in different neoplasms (Mariño *et al.*, 2023).

In this sense, galectin-4 (*LGALS4*) is mainly expressed in the digestive tract of healthy patients, but data on its role in intestinal inflammation are contradictory. On the one hand, it can promote inflammation in the gut, stimulating the production of IL-6 by CD4+ T cells, and on the other, it can appease inflammation by inducing T-cell apoptosis. In models of colitis, it also downregulates proinflammatory cytokines (Brinchmann *et al.*, 2018). In our analyses, galectin-4 dysregulation was observed in all transitions: from normal to inflamed mucosa (decreases, **Section 5.3.1, Results**), from inflamed to dysplastic mucosa (increases, both from left UC and from pancolitis, **Section 5.3.2, Results**), and finally in AOM-DSS, from inflamed mucosa to cancer (increases, **Section 5.3.4, of Results**). This is consistent with previous analyses showing that, in IBD, galectin-4 expression decreases in the intestinal mucosa (Mora Massaro's PhD thesis) (Massaro, 2025), but no data are available for CCRAC. Meanwhile, data from both ERCC patients and colorectal cancer cell lines indicate that Gal4 expression decreases in the tumor (Satelli *et al.*, 2011). In the validation carried out in this Thesis in the AOM-DSS model by RT-qPCR (**Section 5.4, of Results**), although not significant, the trend is verified: the expression of this gene is lower in inflammation (DSS) than in the control, and then Gal-4 increases between inflammation and adenocarcinoma, but not up to the levels of the control group. We hypothesized that the lack of significant results could be due to the low efficiency in the amplification of this target (**Annex 8.4**), or to the small sample size of this experiment. It remains for us to validate this result in the model and confirm whether this observed increase is a differential feature of CCRAC with respect to sporadic colorectal cancer.

On the other hand, galectin-12 (*LGALS12*), which is mainly expressed in adipose tissue (Sjöstedt *et al.*, 2020; The Human Protein Atlas, 2025b; Maller, 2019), could be involved in the development of inflammatory and autoimmune processes. In particular, it was shown that Gal12KO mice present lower clinical manifestations of TNBS hapten-induced colitis in models compared to WT mice (Maller, 2019). This galectin also participates in the control of the cell

cycle: its expression increases in cells arrested in the G1 phase or at the G1-S limit of the cell cycle, including tumor cells, which reduce their growth when ectopic expression of galectin-12 (Yang *et al.*, 2008). In particular, galectin-12 is present in normal and silenced colonic mucosa in multiple CRC cell lines, and its expression in CRC could have tumor-suppressing activity (Katzenmaier *et al.*, 2019). In our RT-qPCR analysis, there was an absence of galectin-12 in all the samples evaluated. Despite this, *Lgals12* is significantly overexpressed in the transcriptomics analysis in the AOM-DSS experimental model (**Section 5.3.4 of Results**), so we will confirm these results in subsequent works to this thesis.

### **6.6.3. *St3gal1***

The last gene we sought to validate by RT-qPCR was *ST3GAL1*, which encodes for a sialyltransferase that catalyzes the transfer of sialic acid in a  $\alpha(2,3)$  bond to the Serine or Threonine-bound Gal- $\beta$ 1,3-GalNAc residue (*core O-type 1*), which blocks chain elongation. Dysregulation of *ST3GAL1* could lead to alterations in the sialoma, which, in turn, can promote tumor invasion and metastasis (Fan *et al.*, 2025). In fact, the expression of *ST3GAL1* is very high in several types of cancer (including colorectal), and its high expression in breast cancer correlates with worse patient survival. ST3GAL1 could modulate tumor cell migration via the EGFR/NRP1 pathway (Fan *et al.*, 2025).

Our analyses indicate significant overexpression of *ST3GAL1* in patients with active UC vs. without IBD (**Section 5.3.1.1, Results**), which is consistent with previous transcriptional analyses, which link this gene and *O-GlcNAcylation with the regulation of immunity and the occurrence of inflammatory bowel diseases* (Lu *et al.*, 2024). However, this gene is decreased in our glycobiological signature (**Section 5.3.4, of Results**), which could reflect another difference of CCRAC with respect to CRC at the glycobiological level, or a difference specifically associated with the experimental model. Unfortunately, in quantitative PCR we could not reliably validate the results *in silico*, because as a target, *St3gal1* had non-specific amplification. In summary, and as for Gal-4 and Gal-12, additional studies are required to clarify the role of this gene in AOM-DSS, and in CCRAC.

### **6.6.4. Adhesion molecules: *Sell*, *Selp*, and *Icam1***

Selectins are a family of C-type lectins (calcium-dependent), known for their mediation in the adhesion of immune cells to the endothelium to facilitate entry into sites of inflammation. Lectin L (*SELL*) is constitutively expressed in circulating leukocytes, and selectin P (*SELP*) in activated endothelial cells and platelets. Both promote metastasis in different types of cancer (Smith & Bertozzi, 2021b; 2021a). As for the intercellular adhesion molecule 1 (*ICAM1*), this protein of the immunoglobulin family plays a key role in the accumulation of neutrophils in the mucosa in IBD: after leaving the circulatory system, neutrophils migrate from the basal to the apical part of the epithelium, where *ICAM1* allows them to bind, thus reaching the intestinal lumen (Qiu *et al.*, 2022; Kudelka *et al.*, 2020). Its role in colorectal cancer is controversial, because it seems to have a dual behavior, depending on the context: in its membrane-bound form, it has been associated with a favorable prognosis in CRC, with less tumor growth and longer survival; but in its secreted form, its overexpression is associated with a worse prognosis (Qiu *et al.*, 2022). One mechanism that may explain the association of *ICAM1* with favorable prognoses is that it favors the infiltration and cytotoxicity of antitumor CD8+ T cells, improving recognition in tumor cells where it is expressed (Qiu *et al.*, 2022).

In our *in silico* results in the AOM-DSS model (Section 5.3.4, Results), both selectins increase from healthy patients to patients with active UC (indicating inflammation), and then, together with *ICAM1*, decrease in the tissue of patients with CCRAC (suggesting immune evasion).

#### **6.6.5. *Chi3l1* and *Chil3***

Mammalian acid chitinase (AMCase, *Acidic Mammalian Chitinase*; encoded by the *CHIA* gene in humans, and *Chil3* in mice) and pseudochitinase *CHI3L1* (without hydrolase activity) are proteins that bind chitin, a polysaccharide present in fungi, insects, and parasitic nematodes, but completely absent in mammals, so these proteins are involved in the immune response to foreign organisms (Eurich *et al.*, 2009).

In this context, it was shown that, in the gut, commensal bacteria can stimulate epithelial cells to produce and secrete *CHI3L1*, which interacts with the microbiota through peptidoglycans (main components of the walls of gram-positive bacteria, whose chitin-like

structure allows its recognition by pseudochitinase). This interaction plays an anchoring role in the mucosa, favoring the colonization of beneficial gram-positive bacteria such as *Lactobacillus*, improving intestinal health. For this reason, deficiencies in CHI3L1 can lead to dysbiosis and exacerbate DSS-induced colitis (Chen *et al.*, 2024). In the context of IBD, CHI3L1 has been proposed as a biomarker for active disease, because it is almost not expressed in healthy individuals, but is elevated in the serum of patients with UC (Eurich *et al.*, 2009).

In the context of cancer, CHI3L1 is expressed in, and secreted by, several types of solid tumors; notably, in CRC its increase has been detected in patients' serum, but it has not been quantified in tissue, nor has it been studied in patients with CRCC. However, considering that CHI3L1 could play an important role in the development of CCRC, promoting chronic inflammation and activating the Wnt/β-catenin pathway (Eurich *et al.*, 2009).

The results of our transcriptomic analyses suggest that the expression of this pseudochitinase increases in active UC (vs. patients without IBD, **Section 5.3.1.1, of Results**) and in intestinal inflammation due to AOM-DSS (**Section 5.3.4, of Results**), and then decreases in dysplasia in patients (**Section 5.3.2, of Results**) and in adenocarcinoma in the AOM-DSS model(**Section 5.3.4, of Results**). This could indicate that, although the functionality of CHI3L1 has a characteristic increase in inflammation that would functionally favor neoplastic development, as its expression is associated with immune cells (e.g., macrophages) (Zhao *et al.*, 2023a), its expression decreases in the tumor due to immune evasion.

On the other hand, there is also evidence that AMCase is involved in IBD, as a 2022 study investigated the effect of administration of the selective inhibitor for AMCase, OAT-177, in a mouse model of DSS-induced colitis. The results were a significant decrease in inflammation and TNF-α mRNA levels, signaling the pro-inflammatory activity of *Chil3* (Mazur *et al.*, 2022). As this protein is also produced in macrophages, it can be deduced that the significant decrease in mRNA in the murine model of adenocarcinoma vs. inflammation (**Section 5.3.4, of the Results**), could reflect, as for *Chil3II*, the immune evasion exerted by the tumor.

In future work, we will analyze *single-cell* data to validate these hypotheses.

### **6.6.6. *SLC26A2***

The solute transporter family 26 gene, member 2 (*SLC26A2*), encoding a membrane sulfate/chloride antecarrier, is critical for inorganic sulfate uptake. The encoding protein (DTDST, *diastrophic dysplasia sulfate transporter*) is expressed in epithelial and calcithoform cells of the colon of healthy patients (The Human Protein Atlas, 2025a; Sjöstedt *et al.*, 2020; Qian *et al.*, 2025). *SLC26A2* plays a role in IBD, and the evidence (both obtained in this Thesis and in the literature) indicates that it decreases in the intestinal mucosa of patients with active UC, compared to healthy controls, a dysregulation that also occurs in models of intestinal inflammation induced by DSS compared to the control group (Qian *et al.*, 2025). *SLC26A2* has been shown to have a negative correlation with the IL-17 signaling pathway, and a positive association with tight junctions (Qian *et al.*, 2025). Interestingly, the overexpression of *SLC26A2* in Crohn's disease was reported, being proposed as an indicator that distinguishes patients with this disease from those with ulcerative colitis (Alper & Sharma, 2013).

In colon cancer, *SLC26A2* expression decreases, and patients with increased *SLC26A2* expression in tumor tissues have a better prognosis (Qian *et al.*, 2025). In addition, inhibition of *SLC26A2* in tumor cell cultures increases their proliferation (Alper & Sharma, 2013). *SLC26A2* also provides sulfates for the sulfation of the sialyl groups of the Lewis(x) antigen in normal colonic epithelial cells; these sialyl groups are not sulfated in colon cancer cells, which is associated with an epigenetic suppression of *SLC26A2* expression, reversible by histone deacetylase inhibitors (Alper and Sharma, 2013).

Another reason why *SLC26A2* is a gene of great interest is because of its ability to regulate intestinal mucins, because the decrease in sulfation in them alters their protective capacity (Corfield, 2018). In addition, sulfated glycosaminoglycans can regulate metalloproteinase activity significantly (Ra *et al.*, 2009). A particular metalloproteinase, metalloproteinase-2 (MMP2) increases with pathological progression from normal mucosa to adenoma without high-grade dysplasia, to adenoma with high-grade dysplasia and finally to cancer. Their expression levels make it possible to distinguish between adenomas with or without high-grade dysplasia (Murnane *et al.*, 2011). For this reason, elucidating the regulatory

action of SLC26A2 on MMP2 could be especially useful in predicting the future risk of colorectal cancer.

The results of our transcriptomic analyses indicate that *SLC26A2* is dysregulated at various stages of the transition from healthy tissue to the tumor associated with colitis. First, we observed their decrease in active and quiescent UC (although only for active inflammation with significant values), compared to healthy patients (**Section 5.3.1, Results**). This dysregulation is consistent with the current literature (Qian *et al.*, 2025). However, we then observed that *SLC26A2* significantly increased in the CU-dysplasia transition in patients (**Section 5.3.2, of Results**), and also in the adenocarcinoma-inflammation comparison in AOM-DSS (**Section 5.3.4, of Results**). In the Thesis of Christian Mazzeo (LU: 1049260) (Mazzeo, 2021), and through the analysis of public databases of biopsies of patients with sporadic colorectal cancer, we observed a decrease in the expression of this gene, a result that is reproduced in a recent work (Mamoor, 2023). Given the lack of data in patients with colorectal cancer associated with colitis, it remains to be validated whether the differences observed again correspond to the mechanics of the CCRAC or to a characteristic of the murine model.

## 6.7. Contributions and future perspectives

The analysis carried out in this Thesis in patient samples and in models of intestinal inflammation and colorectal cancer associated with colitis allowed to delineate a signature of 10 glycogenes potentially dysregulated in a differential manner with respect to intestinal inflammation: *DEFA5*, *LGALS4*, *LGALS12*, *SLC26A2* (increased), *ICAM1*, *SELL*, *SELP*, *ST3GAL1*, *CHI3L1*, and *CHIA* (decreased). We studied the dysregulation of four of these genes (*Defa5*, *Lgals4*, *Lgals12* and *St3gall*) by RT-qPCR in a model of DSS-induced colitis and CRRAC (AOM-DSS). This approach allowed validating the alteration of *Defa5*, and supports the investigation of the glycobiological signature obtained in samples from patients with CCRAC, with the aim of implementing early diagnosis strategies in patients with UC.

## 7. BIBLIOGRAPHY

- ADAMS, M.K., LEE, S.-A., BELYAEVA, O.V., WU, L., and KEDISHVILI, N.Y., 2017. Characterization of human short chain dehydrogenase/reductase SDR16C family members related to retinol dehydrogenase 10. *Chemico-Biological Interactions*, vol. 276, pp. 88-94. ISSN 00092797. DOI 10.1016/j.cbi.2016.10.019.
- ADEBAYO, A.S., AGBAJE, K., ADESINA, S.K. and OLAJUBUTU, O., 2023. Colorectal Cancer: Disease Process, Current Treatment Options, and Future Perspectives. *Pharmaceutics*, vol. 15, no. 11, pp. 2620. ISSN 1999-4923. DOI 10.3390/pharmaceutics15112620.
- AKHMEDOV, M., MARTINELLI, A., GEIGER, R., & KWEE, I., 2020. Omics Playground: a comprehensive self-service platform for visualization, analytics and exploration of Big Omics Data. *NAR Genomics and Bioinformatics* [online], vol. 2, no. 1. [Accessed: 8 July 2025]. ISSN 2631-9268. DOI 10.1093/nargab/lqz019. Available at: <https://academic.oup.com/nargab/article/doi/10.1093/nargab/lqz019/5661104>.
- ALPER, S.L. and SHARMA, A.K., 2013. The SLC26 gene family of anion transporters and channels. *Molecular Aspects of Medicine*, vol. 34, no. 2-3, pp. 494-515. ISSN 0098-2997. DOI 10.1016/j.mam.2012.07.009.
- ALZAHIRANI, S., AL DOGHAIATHER, H. and AL-GHAFARI, A., 2021. General insight into cancer: An overview of colorectal cancer (Review). *Molecular and Clinical Oncology*, vol. 15, no. 6, pp. 271. ISSN 2049-9450, 2049-9469. DOI 10.3892/mco.2021.2433.
- ANGELOVA, M., MASCAUX, C., & GALON, J., 2021. Evasion before invasion: Pre-cancer immunosurveillance. *OncoImmunology* [online], vol. 10, no. 1. [Accessed: 7 July 2025]. ISSN 2162-402X. DOI 10.1080/2162402x.2021.1912250. Available at: <https://www.tandfonline.com/doi/full/10.1080/2162402X.2021.1912250>.
- ARIJS, I., DE HERTOGH, G., LEMMENS, B., VAN LOMMEL, L., DE BRUYN, M., VANHOVE, W., CLEYNEN, I., MACHIELS, K., FERRANTE, M., SCHUIT, F., VAN ASSCHE, G., RUTGEERTS, P., and VERMEIRE, S., 2018. Effect of vedolizumab (anti- $\alpha$ 4 $\beta$ 7-integrin) therapy on histological healing and mucosal gene expression in patients with UC. *Gut*, vol. 67, no. 1, pp. 43-52. ISSN 0017-5749, 1468-3288. DOI 10.1136/gutjnl-2016-312293.
- ARNOLD, M., SIERRA, M.S., LAVERSANNE, M., SOERJOMATARAM, I., JEMAL, A., & BRAY, F., 2017. Global patterns and trends in colorectal cancer incidence and mortality. *Gut*, vol. 66, no. 4, pp. 683-691. ISSN 0017-5749, 1468-3288. DOI 10.1136/gutjnl-2015-310912.
- AXELRAD, J.E., LICHTIGER, S. and YAJNIK, V., 2016. Inflammatory bowel disease and cancer: The role of inflammation, immunosuppression, and cancer treatment. *World Journal of Gastroenterology*, vol. 22, no. 20, pp. 4794. ISSN 1007-9327. DOI 10.3748/wjg.v22.i20.4794.
- BAARS, M.J.D., FLOOR, E., SINHA, N., TER LINDE, J.J.M., VAN DAM, S., AMINI, M., NIJMAN, I.J., TEN HOVE, J.R., DRYLEWICZ, J., OFFERHAUS, G.J.A., LACLÉ, M.M., OLDENBURG, B., and VERCOULEN, Y., 2024. Multiplex spatial omics reveals changes in immune-epithelial crosstalk during inflammation and dysplasia development in chronic IBD patients. *iScience*, vol. 27, no. 8, pp. 110550. ISSN 2589-0042. DOI 10.1016/j.isci.2024.110550.

- BAEK, D.H. and KIM, G.H., 2020. Can transrectal ultrasound-guided core needle biopsy serve as an accurate diagnostic tool for rectal lesions? *Annals of Translational Medicine*, vol. 8, no. 4, pp. 66-66. ISSN 23055839, 23055847. DOI 10.21037/atm.2019.11.144.
- BEAUGERIE, L. & ITZKOWITZ, S.H., 2015. Cancers complicating inflammatory bowel disease. *The New England Journal of Medicine*, vol. 372, no. 15, pp. 1441-1452. ISSN 1533-4406. DOI 10.1056/NEJMra1403718.
- BENJAMINI, Y. and HOCHBERG, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 57, no. 1, pp. 289-300. ISSN 1369-7412, 1467-9868. DOI 10.1111/j.2517-6161.1995.tb02031.x.
- BIRAGYN, A., RUFFINI, P.A., LEIFER, C.A., KLYUSHNENKOVA, E., SHAKHOV, A., CHERTOV, O., SHIRAKAWA, A.K., FARBER, J.M., SEGAL, D.M., OPPENHEIM, J.J., and KWAK, L.W., 2002. Toll-Like Receptor 4-Dependent Activation of Dendritic Cells by  $\beta$ -Defensin 2. *Science*, vol. 298, no. 5595, pp. 1025-1029. ISSN 0036-8075, 1095-9203. DOI 10.1126/science.1075565.
- BJERRUM, J.T., NIELSEN, O.H., RIIS, L.B., PITTEL, V., MUELLER, C., ROGLER, G., & OLSEN, J., 2014. Transcriptional Analysis of Left-sided Colitis, Pancolitis, and Ulcerative Colitis-associated Dysplasia: *Inflammatory Bowel Diseases*, vol. 20, no. 12, pp. 2340-2352. ISSN 1078-0998. DOI 10.1097/MIB.0000000000000235.
- BJERRUM, J.T., WANG, Y., ZHANG, J., RIIS, L.B., NIELSEN, O.H., and SEIDELIN, J.B., 2022. Lipidomic Trajectories Characterize Delayed Mucosal Wound Healing in Quiescent Ulcerative Colitis and Identify Potential Novel Therapeutic Targets. *International Journal of Biological Sciences*, vol. 18, no. 5, pp. 1813-1828. ISSN 1449-2288. DOI 10.7150/ijbs.67112.
- BOAL CARVALHO, P. and COTTER, J., 2017. Mucosal Healing in Ulcerative Colitis: A Comprehensive Review. *Drugs*, vol. 77, no. 2, pp. 159-173. ISSN 0012-6667, 1179-1950. DOI 10.1007/s40265-016-0676-y.
- BOIVIN, G.P., WASHINGTON, K., YANG, K., WARD, J.M., PRETLOW, T.P., RUSSELL, R., BESSELSEN, D.G., GODFREY, V.L., DOETSCHMAN, T., DOVE, W.F., PITOT, H.C., HALBERG, R.B., ITZKOWITZ, S.H., GRODEN, J., and COFFEY, R.J., 2003. Pathology of mouse models of intestinal cancer: Consensus report and recommendations. *Gastroenterology*, vol. 124, no. 3, pp. 762-777. ISSN 00165085. DOI 10.1053/gast.2003.50094.
- BRINCHMANN, M.F., PATEL, D.M., and IVERSEN, M.H., 2018. The Role of Galectins as Modulators of Metabolism and Inflammation. *Mediators of Inflammation*, vol. 2018, pp. 1-11. ISSN 0962-9351, 1466-1861. DOI 10.1155/2018/9186940.
- CAGNONI, A.J., PÉREZ SÁEZ, J.M., RABINOVICH, G.A. and MARIÑO, K.V., 2016. Turning-Off Signaling by Siglecs, Selectins, and Galectins: Chemical Inhibition of Glycan-Dependent Interactions in Cancer. *Frontiers in Oncology* [online], vol. 6. [Accessed: 22 October 2024]. ISSN 2234-943X. DOI 10.3389/fonc.2016.00109. Available at: <https://www.frontiersin.org/journals/oncology/articles/10.3389/fonc.2016.00109/full>.

CAI, Z., WANG, S., & LI, J., 2021. Treatment of Inflammatory Bowel Disease: A Comprehensive Review. *Frontiers in Medicine*, vol. 8, pp. 765474. ISSN 2296-858X. DOI 10.3389/fmed.2021.765474.

CAMARA, R., OGBENI, D., GERSTMANN, L., OSTOVAR, M., HURER, E., SCOTT, M., MAHMOUD, N.G., RADON, T., CRNOGORAC-JURCEVIC, T., PATEL, P., MACKENZIE, L.S., CHAU, D.Y.S., KIRTON, S.B. and ROSSITER, S., 2020. Discovery of novel small molecule inhibitors of S100P with in vitro anti-metastatic effects on pancreatic cancer cells. *European Journal of Medicinal Chemistry*, vol. 203, pp. 112621. ISSN 02235234. DOI 10.1016/j.ejmech.2020.112621.

CAMPDEN, R.I., WARREN, A.L., GREENE, C.J., CHIRIBOGA, J.A., ARNOLD, C.R., AGGARWAL, D., MCKENNA, N., SANDALL, C.F., MACDONALD, J.A., and YATES, R.M., 2022. Extracellular cathepsin Z signals through the  $\alpha 5$  integrin and augments NLRP3 inflammasome activation. *Journal of Biological Chemistry*, vol. 298, no. 1, pp. 101459. ISSN 0021-9258. DOI 10.1016/j.jbc.2021.101459.

CHANDRASEKAR, D., GUERRIER, C., ALISSON-SILVA, F., DHAR, C., CAVAL, T., SCHWARZ, F., and HOMMES, D.W., 2023. Warning Signs From the Crypt: Aberrant Protein Glycosylation Marks Opportunities for Early Colorectal Cancer Detection. *Clinical and Translational Gastroenterology*, vol. 14, no. 7, pp. e00592. ISSN 2155-384X. DOI 10.14309/ctg.00000000000000592.

CHEN, Y., CHEN, L., LUN, A.T.L., BALDONI, P.L., and SMYTH, G.K., 2025. edgeR v4: powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets. *Nucleic Acids Research*, vol. 53, no. 2, pp. gkaf018. ISSN 0305-1048, 1362-4962. DOI 10.1093/nar/gkaf018.

CHEN, Y., YANG, R., QI, B., & SHAN, Z., 2024. Peptidoglycan-Chi3II interaction shapes gut microbiota in intestinal mucus layer [online]. 16 September 2024. S.l.: s.n. [Accessed: 30 June 2025]. Available at: <https://elifesciences.org/reviewed-preprints/92994v2>.

CHENG, F., ZHANG, R., SUN, C., RAN, Q., ZHANG, C., SHEN, C., YAO, Z., WANG, M., SONG, L., and PENG, C., 2023. Oxaliplatin-induced peripheral neurotoxicity in colorectal cancer patients: mechanisms, pharmacokinetics and strategies. *Frontiers in Pharmacology*, vol. 14, pp. 1231401. ISSN 1663-9812. DOI 10.3389/fphar.2023.1231401.

CHICHIARELLI, S., ALTIERI, F., PAGLIA, G., RUBINI, E., MINACORI, M., & EUFEMI, M., 2022. ERp57/PDIA3: new insight. *Cellular & Molecular Biology Letters*, vol. 27, no. 1, pp. 12. ISSN 1425-8153, 1689-1392. DOI 10.1186/s11658-022-00315-x.

COLLARD, M., GUEDJ, N., TOURNEUR-MARSILLE, J., ALBUQUERQUE, M., MAGGIORI, L., HAMMEL, P., TRETON, X., PANIS, Y., and OGIER-DENIS, E., 2020. Immune-checkpoint inhibitor anti-PD1 aggravates colitis-associated colorectal cancer without enhancing intestinal inflammation. *Integrative Cancer Science and Therapeutics* [online], vol. 7, no. 2. [Accessed: 28 June 2025]. ISSN 20564546. DOI 10.15761/ICST.1000334. Available at: <https://www.oatext.com/immune-checkpoint-inhibitor-anti-pd1-aggravates-colitis-associated-colorectal-cancer-without-enhancing-intestinal-inflammation.php#gsc.tab=0>.

- CONWAY, J.R., LEX, A., and GEHLENborg, N., 2017. UpSetR: an R package for the visualization of intersecting sets and their properties. In: J. HANCOCK (ed.), *Bioinformatics*, vol. 33, no. 18, pp. 2938-2940. ISSN 1367-4803, 1367-4811. DOI 10.1093/bioinformatics/btx364.
- CORFIELD, A.P., 2018. The Interaction of the Gut Microbiota with the Mucus Barrier in Health and Disease in Human. *Microorganisms*, vol. 6, no. 3, pp. 78. ISSN 2076-2607. DOI 10.3390/microorganisms6030078.
- CS50R Final Project: degfind (Bioinformatics tools for differential gene expression analysis) - YouTube. [online], [undated]. [Consulted: 27 May 2025]. Available at: <https://www.youtube.com/watch?si=0UAEnBc3huiGKCx&v=7Vcm609bAac&feature=youtu.be>.
- CUFFARI, C., PRESENT, D.H., BAYLESS, T.M. and LICHTENSTEIN, G.R., 2005. Optimizing therapy in patients with pancolitis. *Inflammatory Bowel Diseases*, vol. 11, no. 10, pp. 937-946. ISSN 1078-0998. DOI 10.1097/01.mib.0000179469.86500.ac.
- DE FREITAS JUNIOR, J.C.M. and MORGADO-DÍAZ, J.A., 2016. The role of N-glycans in colorectal cancer progression: potential biomarkers and therapeutic applications. *Oncotarget*, vol. 7, no. 15, pp. 19395-19413. ISSN 1949-2553. DOI 10.18632/oncotarget.6283.
- DE SOUZA, H.S.P. and FIOCCHI, C., 2016. Immunopathogenesis of IBD: current state of the art. *Nature Reviews Gastroenterology & Hepatology*, vol. 13, no. 1, pp. 13-27. ISSN 1759-5045, 1759-5053. DOI 10.1038/nrgastro.2015.186.
- DEUTSCHMANN, C., ROGGENBUCK, D., & SCHIERACK, P., 2019. The loss of tolerance to CHI3L1 – A putative role in inflammatory bowel disease? *Clinical Immunology*, vol. 199, pp. 12-17. ISSN 15216616. DOI 10.1016/j.clim.2018.12.005.
- DEWAL, M.B., DICHIARA, A.S., ANTONOPOULOS, A., TAYLOR, R.J., HARMON, C.J., HASLAM, S.M., DELL, A., and SHOULDERS, M.D., 2015. XBP1s Links the Unfolded Protein Response to the Molecular Architecture of Mature N-Glycans. *Chemistry & Biology*, vol. 22, no. 10, pp. 1301-1312. ISSN 10745521. DOI 10.1016/j.chembiol.2015.09.006.
- DIAS, A.M., CORREIA, A., PEREIRA, M.S., ALMEIDA, C.R., ALVES, I., PINTO, V., CATARINO, T.A., MENDES, N., LEANDER, M., OLIVA-TELES, M.T., MAIA, L., DELERUE-MATOS, C., TANIGUCHI, N., LIMA, M., PEDROTO, I., MARCOS-PINTO, R., LAGO, P., REIS, C.A., VILANOVA, M. and PINHO, S.S., 2018. Metabolic control of T cell immune response through glycans in inflammatory bowel disease. *Proceedings of the National Academy of Sciences* [online], vol. 115, no. 20. [Accessed: 26 May 2025]. ISSN 0027-8424, 1091-6490. DOI 10.1073/pnas.1720409115. Available at: <https://pnas.org/doi/full/10.1073/pnas.1720409115>.
- DURINCK, S., SPELLMAN, P.T., BIRNEY, E., and HUBER, W., 2009. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nature Protocols*, vol. 4, no. 8, pp. 1184-1191. ISSN 1754-2189, 1750-2799. DOI 10.1038/nprot.2009.97.

EDENBERG, H.J., 2007. The Genetics of Alcohol Metabolism: Role of Alcohol Dehydrogenase and Aldehyde Dehydrogenase Variants. *Alcohol Research & Health*, vol. 30, no. 1, pp. 5-13. ISSN 1535-7414.

EICHELE, D.D. and KHARBANDA, K.K., 2017. Dextran sodium sulfate colitis murine model: An indispensable tool for advancing our understanding of inflammatory bowel diseases pathogenesis. *World Journal of Gastroenterology*, vol. 23, no. 33, pp. 6016-6029. ISSN 1007-9327. DOI 10.3748/wjg.v23.i33.6016.

EL-SAYED, D., EL-KARAKSY, H., WALI, Y. and YOUSSEY, I., 2023. Mitochondrial 3-hydroxymethylglutaryl-CoA synthase-2 (HMGCS2) deficiency: a rare case with bicytopenia and coagulopathy. *BMJ Case Reports*, vol. 16, no. 11, pp. e257011. ISSN 1757-790X. DOI 10.1136/bcr-2023-257011.

EURICH, K., SEGAWA, M., TOEI-SHIMIZU, S., and MIZOGUCHI, E., 2009. Potential role of chitinase 3-like-1 in inflammation-associated carcinogenic changes of epithelial cells. *World Journal of Gastroenterology*, vol. 15, no. 42, pp. 5249. ISSN 1007-9327. DOI 10.3748/wjg.15.5249.

FAN, T., YEO, H.L., HUNG, T.-H., CHANG, N.-C., TANG, Y.-H., YU, J., CHEN, S.-H., and YU, A.L., 2025. ST3GAL1 regulates cancer cell migration through crosstalk between EGFR and neuropilin-1 signaling. *Journal of Biological Chemistry*, vol. 301, no. 4, pp. 108368. ISSN 00219258. DOI 10.1016/j.jbc.2025.108368.

FAN, Y., MENG, Y., HU, X., LIU, J., and QIN, X., 2024. Uncovering novel mechanisms of chitinase-3-like protein 1 in driving inflammation-associated cancers. *Cancer Cell International*, vol. 24, no. 1, pp. 268. ISSN 1475-2867. DOI 10.1186/s12935-024-03425-y.

FENTON, C.G., TAMAN, H., FLORHOLMEN, J., SØRBYE, S.W., and PAULSEN, R.H., 2021. Transcriptional Signatures That Define Ulcerative Colitis in Remission. *Inflammatory Bowel Diseases*, vol. 27, no. 1, pp. 94-105. ISSN 1078-0998, 1536-4844. DOI 10.1093/ibd/izaa075.

FERLAY, J., ERVIK, M., LAM, F., LAVERSANNE, M., COLOMBET, M., MERY, L., PIÑEROS, M., ZNAOR, A., SOERJOMATARAM, I. and BRAY, F., 2024. Cancer Today. [online]. [Accessed: 23 May 2025]. Available at: <https://gco.iarc.who.int/today/>.

FONSECA, L.M.D., SILVA, V.A.D., FREIRE-DE-LIMA, L., PREVIATO, J.O., MENDONÇA-PREVIATO, L. and CAPELLA, M.A.M., 2016. Glycosylation in Cancer: Interplay between Multidrug Resistance and Epithelial-to-Mesenchymal Transition? *Frontiers in Oncology* [online], vol. 6. [Consulted: 1 July 2025]. ISSN 2234-943X. DOI 10.3389/fonc.2016.00158. Available at: <http://journal.frontiersin.org/Article/10.3389/fonc.2016.00158/abstract>.

FRANKE, A.J., SKELTON, W.P., STARR, J.S., PAREKH, H., LEE, J.J., OVERMAN, M.J., ALLEGRA, C., & GEORGE, T.J., 2019. Immunotherapy for Colorectal Cancer: A Review of Current and Novel Therapeutic Approaches. *Journal of the National Cancer Institute*, vol. 111, no. 11, pp. 1131-1141. ISSN 1460-2105. DOI 10.1093/jnci/djz093.

FU, J., ZONG, X., JIN, M., MIN, J., WANG, F., and WANG, Y., 2023. Mechanisms and regulation of defensins in host defense. *Signal Transduction and Targeted Therapy*, vol. 8, no. 1, pp. 300. ISSN 2059-3635. DOI 10.1038/s41392-023-01553-x.

FUJII, H., SHINZAKI, S., IIJIMA, H., WAKAMATSU, K., IWAMOTO, C., SOBAJIMA, T., KUWAHARA, R., HIYAMA, S., HAYASHI, Y., TAKAMATSU, S., UOZUMI, N., KAMADA, Y., TSUJII, M., TANIGUCHI, N., TAKEHARA, T., and MIYOSHI, E., 2016. Core Fucosylation on T Cells, Required for Activation of T-Cell Receptor Signaling and Induction of Colitis in Mice, Is Increased in Patients With Inflammatory Bowel Disease. *Gastroenterology*, vol. 150, no. 7, pp. 1620-1632. ISSN 00165085. DOI 10.1053/j.gastro.2016.03.002.

GANDINI, A., PUGLISI, S., PIRRONE, C., MARTELLI, V., CATALANO, F., NARDIN, S., SEEBER, A., PUCCINI, A., & SCIALLERO, S., 2023. The role of immunotherapy in microsatellites stable metastatic colorectal cancer: state of the art and future perspectives. *Frontiers in Oncology*, vol. 13, pp. 1161048. ISSN 2234-943X. DOI 10.3389/fonc.2023.1161048.

GAO, X., DING, J., LIAO, C., XU, J., LIU, X., and LU, W., 2021. Defensins: The natural peptide antibiotic. *Advanced Drug Delivery Reviews*, vol. 179, pp. 114008. ISSN 0169-409X. DOI 10.1016/j.addr.2021.114008.

GARBES, L., KIM, K., RIESS, A., HOYER-KUHN, H., BELEGGIA, F., BEVOT, A., KIM, M.J., HUH, Y.H., KWEON, H.-S., SAVARIRAYAN, R., AMOR, D., KAKADIA, P.M., LINDIG, T., KAGAN, K.O., BECKER, J., BOYADJIEV, S.A., WOLNIK, B., SEMLER, O., BOHLANDER, S.K., KIM, J., and NETZER, C., 2015. Mutations in SEC24D, Encoding a Component of the COPII Machinery, Cause a Syndromic Form of Osteogenesis Imperfecta. *The American Journal of Human Genetics*, vol. 96, no. 3, pp. 432-439. ISSN 00029297. DOI 10.1016/j.ajhg.2015.01.002.

GAUTIER, L., COPE, L., BOLSTAD, B.M., and IRIZARRY, R.A., 2004. affy—analysis of *Affymetrix GeneChip* data at the probe level. *Bioinformatics*, vol. 20, no. 3, pp. 307-315. ISSN 1367-4811, 1367-4803. DOI 10.1093/bioinformatics/btg405.

GECSE, K.B. and VERMEIRE, S., 2018. Differential diagnosis of inflammatory bowel disease: imitations and complications. *The Lancet Gastroenterology & Hepatology*, vol. 3, no. 9, pp. 644-653. ISSN 24681253. DOI 10.1016/S2468-1253(18)30159-6.

GIESLER, S., RIEMER, R., LOWINUS, T., & ZEISER, R., 2025. Immune-mediated colitis after immune checkpoint inhibitor therapy. *Trends in Molecular Medicine*, vol. 31, no. 3, pp. 265-280. ISSN 14714914. DOI 10.1016/j.molmed.2024.09.009.

GIROTTI, M.R., SALATINO, M., DALOTTO-MORENO, T. and RABINOVICH, G.A., 2020. Sweetening the hallmarks of cancer: Galectins as multifunctional mediators of tumor progression. *Journal of Experimental Medicine*, vol. 217, no. 2, pp. e20182041. ISSN 0022-1007, 1540-9538. DOI 10.1084/jem.20182041.

GOBERT, A.P., LATOUR, Y.L., ASIM, M., BARRY, D.P., ALLAMAN, M.M., FINLEY, J.L., SMITH, T.M., MCNAMARA, K.M., SINGH, K., SIERRA, J.C., DELGADO, A., LUIS, P.B., SCHNEIDER, C., WASHINGTON, M.K., PIAZUELO, M.B., ZHAO, S., COBURN, L.A., and

WILSON, K.T., 2022a. Correction. *Gastroenterology*, vol. 162, no. 7, pp. 2138-2143. ISSN 0016-5085. DOI 10.1053/j.gastro.2022.04.006.

GOBERT, A.P., LATOUR, Y.L., ASIM, M., BARRY, D.P., ALLAMAN, M.M., FINLEY, J.L., SMITH, T.M., MCNAMARA, K.M., SINGH, K., SIERRA, J.C., DELGADO, A.G., LUIS, P.B., SCHNEIDER, C., WASHINGTON, M.K., PIAZUELO, M.B., ZHAO, S., COBURN, L.A., and WILSON, K.T., 2022b. Protective Role of Spermidine in Colitis and Colon Carcinogenesis. *Gastroenterology*, vol. 162, no. 3, pp. 813-827.e8. ISSN 0016-5085. DOI 10.1053/j.gastro.2021.11.005.

GONZÁLEZ-ÁVILA, G., SOMMER, B., MENDOZA-POSADA, D.A., RAMOS, C., GARCÍA-HERNÁNDEZ, A.A. and FALFÁN-VALENCIA, R., 2019. Matrix metalloproteinases participation in the metastatic process and their diagnostic and therapeutic applications in cancer. *Critical Reviews in Oncology/Hematology*, vol. 137, pp. 57-83. ISSN 10408428. DOI 10.1016/j.critrevonc.2019.02.010.

GOTO, N., AGUDO, J. and YILMAZ, Ö.H., 2025. Early immune evasion in colorectal cancer: interplay between stem cells and the tumor microenvironment. *Trends in Cancer*, pp. S2405803325001128. ISSN 24058033. DOI 10.1016/j.trecan.2025.04.016.

GRONDIN, J.A., KWON, Y.H., FAR, P.M., HAQ, S., & KHAN, W.I., 2020. Mucins in Intestinal Mucosal Defense and Inflammation: Learning From Clinical and Experimental Studies. *Frontiers in Immunology*, vol. 11, pp. 2054. ISSN 1664-3224. DOI 10.3389/fimmu.2020.02054.

GROUX-DEGROOTE, S., VICOGNE, D., COGEZ, V., SCHULZ, C. & HARDUIN-LEPERS, A., 2021. B4GALNT2 Controls Sda and SLex Antigen Biosynthesis in Healthy and Cancer Human Colon. *ChemBioChem*, vol. 22, no. 24, pp. 3381-3390. ISSN 1439-4227, 1439-7633. DOI 10.1002/cbic.202100363.

GUINNEY, J., DIENSTMANN, R., WANG, X., DE REYNIÈS, A., SCHLICKER, A., SONESON, C., MARISA, L., ROEPMAN, P., NYAMUNDANDA, G., ANGELINO, P., BOT, B.M., MORRIS, J.S., SIMON, I.M., GERSTER, S., FESSLER, E., DE SOUSA E MELO, F., MISSIAGLIA, E., RAMAY, H., BARRAS, D., HOMICKO, K., MARU, D., MANYAM, G.C., BROOM, B., BOIGE, V., PEREZ-VILLAMIL, B., LADERAS, T., SALAZAR, R., GRAY, J.W., HANAHAN, D., TABERNERO, J., BERNARDS, R., FRIEND, S.H., LAURENT-PUIG, P., MEDEMA, J.P., SADANANDAM, A., WESSELS, L., DELORENZI, M., KOPETZ, S., VERMEULEN, L., and TEJPAR, S., 2015. The consensus molecular subtypes of colorectal cancer. *Nature Medicine*, vol. 21, no. 11, pp. 1350-1356. ISSN 1546-170X. DOI 10.1038/nm.3967.

HÄNZELMANN, S., CASTELO, R. and GUINNEY, J., 2013. GSVA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics*, vol. 14, no. 1, pp. 7. ISSN 1471-2105. DOI 10.1186/1471-2105-14-7.

HENEGHAN, J.F., AKHAVEIN, A., SALAS, M.J., SHMUKLER, B.E., KARNISKI, L.P., VANDORPE, D.H. and ALPER, S.L., 2010. Regulated transport of sulfate and oxalate by SLC26A2/DTDST. *American Journal of Physiology-Cell Physiology*, vol. 298, no. 6, pp. C1363-C1375. ISSN 0363-6143, 1522-1563. DOI 10.1152/ajpcell.00004.2010.

- HERRERA-GÓMEZ, R.G., GRECEA, M., GALLOIS, C., BOIGE, V., PAUTIER, P., PISTILLI, B., PLANCHARD, D., MALKA, D., DUCREUX, M., & MIR, O., 2022. Safety and Efficacy of Bevacizumab in Cancer Patients with Inflammatory Bowel Disease. *Cancers*, vol. 14, no. 12, pp. 2914. ISSN 2072-6694. DOI 10.3390/cancers14122914.
- HIEMSTRA, P.S., 2006. DEFENSINS. *Encyclopedia of Respiratory Medicine* [online]. S.l.: Elsevier, pp. 7-10. [Accessed: 26 May 2025]. ISBN 978-0-12-370879-3. Available in: <https://linkinghub.elsevier.com/retrieve/pii/B0123708796001101>.
- HOKAMA, A., MIZOGUCHI, E. and MIZOGUCHI, A., 2008. Roles of galectins in inflammatory bowel disease. *World Journal of Gastroenterology*, vol. 14, no. 33, pp. 5133. ISSN 1007-9327. DOI 10.3748/wjg.14.5133.
- HOLUBAR, S.D., LIGHTNER, A.L., POYLIN, V., VOGEL, J.D., GAERTNER, W., DAVIS, B., DAVIS, K.G., MAHADEVAN, U., SHAH, S.A., KANE, S.V., STEELE, S.R., PAQUETTE, I.M., FEINGOLD, D.L., and PREPARED ON BEHALF OF THE CLINICAL PRACTICE GUIDELINES COMMITTEE OF THE AMERICAN SOCIETY OF COLON AND RECTAL SURGEONS, 2021. The American Society of Colon and Rectal Surgeons Clinical Practice Guidelines for the Surgical Management of Ulcerative Colitis. *Diseases of the Colon & Rectum*, vol. 64, no. 7, pp. 783-804. ISSN 0012-3706. DOI 10.1097/DCR.0000000000002037.
- HOMMA, Y., HIRAGI, S., & FUKUDA, M., 2021. Rab family of small GTPases: an updated view on their regulation and functions. *The FEBS Journal*, vol. 288, no. 1, pp. 36-55. ISSN 1742-464X, 1742-4658. DOI 10.1111/febs.15453.
- HOSSAIN, Md.S., KARUNIAWATI, H., JAIROUN, A.A., URBI, Z., OOI, D.J., JOHN, A., LIM, Y.C., KIBRIA, K.M.K., MOHIUDDIN, A.K.M., MING, L.C., GOH, K.W., and HADI, M.A., 2022. Colorectal Cancer: A Review of Carcinogenesis, Global Epidemiology, Current Challenges, Risk Factors, Preventive and Treatment Strategies. *Cancers*, vol. 14, no. 7, pp. 1732. ISSN 2072-6694. DOI 10.3390/cancers14071732.
- HOU, J.-J., DING, L., YANG, T., YANG, Y.-F., JIN, Y.-P., ZHANG, X.-P., MA, A.-H., and QIN, Y.-H., 2024. The proteolytic activity in inflammatory bowel disease: insight from gut microbiota. *Microbial Pathogenesis*, vol. 188, pp. 106560. ISSN 0882-4010. DOI 10.1016/j.micpath.2024.106560.
- HUANG, D.W., SHERMAN, B.T., and LEMPICKI, R.A., 2009. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, vol. 37, no. 1, pp. 1-13. ISSN 1362-4962, 0305-1048. DOI 10.1093/nar/gkn923.
- HUANG, T.-X., HUANG, H.-S., DONG, S.-W., CHEN, J.-Y., ZHANG, B., LI, H.-H., ZHANG, T.-T., XIE, Q., LONG, Q.-Y., YANG, Y., HUANG, L.-Y., ZHAO, P., BI, J., LU, X.-F., PAN, F., ZOU, C. and FU, L., 2024. ATP6V0A1-dependent cholesterol absorption in colorectal cancer cells triggers immunosuppressive signaling to inactivate memory CD8+ T cells. *Nature Communications*, vol. 15, no. 1, pp. 5680. ISSN 2041-1723. DOI 10.1038/s41467-024-50077-7.
- HUANG, Y.-F., AOKI, K., AKASE, S., ISHIHARA, M., LIU, Y.-S., YANG, G., KIZUKA, Y., MIZUMOTO, S., TIEMEYER, M., GAO, X.-D., AOKI-KINOSHITA, K.F., and FUJITA, M.,

2021. Global mapping of glycosylation pathways in human-derived cells. *Developmental Cell*, vol. 56, no. 8, pp. 1195- 1209.e7. ISSN 15345807. DOI 10.1016/j.devcel.2021.02.023.
- HUANG, Y.-L., CHASSARD, C., HAUSMANN, M., VON ITZSTEIN, M., and HENNET, T., 2015. Sialic acid catabolism drives intestinal inflammation and microbial dysbiosis in mice. *Nature Communications*, vol. 6, no. 1, pp. 8141. ISSN 2041-1723. DOI 10.1038/ncomms9141.
- HUBER, W., CAREY, V.J., GENTLEMAN, R., ANDERS, S., CARLSON, M., CARVALHO, B.S., BRAVO, H.C., DAVIS, S., GATTO, L., GIRKE, T., GOTTARDO, R., HAHNE, F., HANSEN, K.D., IRIZARRY, R.A., LAWRENCE, M., LOVE, M.I., MACDONALD, J., OBENCHAIN, V., OLEŚ, A.K., PAGÈS, H., REYES, A., SHANNON, P., SMYTH, G.K., TENENBAUM, D., WALDRON, L. & MORGAN, M., 2015. Orchestrating high-throughput genomic analysis with Bioconductor. *Nature Methods*, vol. 12, no. 2, pp. 115-121. ISSN 1548-7091, 1548-7105. DOI 10.1038/nmeth.3252.
- HUGO, W., ZARETSKY, J.M., SUN, L., SONG, C., MORENO, B.H., HU-LIESKOVAN, S., BERENT-MAOZ, B., PANG, J., CHMIELOWSKI, B., CHERRY, G., SEJA, E., LOMELI, S., KONG, X., KELLEY, M.C., SOSMAN, J.A., JOHNSON, D.B., RIBAS, A., and LO, R.S., 2016. Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. *Cell*, vol. 165, no. 1, pp. 35-44. ISSN 1097-4172. DOI 10.1016/j.cell.2016.02.065.
- HUNG, M.-S., CHEN, I.-C., LIN, P.-Y., LUNG, J.-H., LI, Y.-C., LIN, Y.-C., YANG, C.-T., and TSAI, Y.-H., 2016. Epidermal growth factor receptor mutation enhances expression of vascular endothelial growth factor in lung cancer. *Oncology Letters*, vol. 12, no. 6, pp. 4598-4604. ISSN 1792-1074, 1792-1082. DOI 10.3892/ol.2016.5287.
- JABERI, S.A., COHEN, A., D'SOUZA, C., ABDULRAZZAQ, Y.M., OJHA, S., BASTAKI, S., and ADEGHATE, E.A., 2021. Lipocalin-2: Structure, function, distribution and role in metabolic disorders. *Biomedicine & Pharmacotherapy*, vol. 142, pp. 112002. ISSN 07533322. DOI 10.1016/j.biopha.2021.112002.
- JOHANNES, L., JACOB, R., & LEFFLER, H., 2018. Galectins at a glance. *Journal of Cell Science*, vol. 131, no. 9, pp. jcs208884. ISSN 1477-9137, 0021-9533. DOI 10.1242/jcs.208884.
- JOJA, M., GRANT, E.T., and DESAI, M.S., 2025. Living on the edge: Mucus-associated microbes in the colon. *Mucosal Immunology*, pp. S1933021925000418. ISSN 19330219. DOI 10.1016/j.mucimm.2025.04.003.
- KAŁUŻNA, A., OLCZYK, P. and KOMOSIŃSKA-VASSEV, K., 2022. The Role of Innate and Adaptive Immune Cells in the Pathogenesis and Development of the Inflammatory Response in Ulcerative Colitis. *Journal of Clinical Medicine*, vol. 11, no. 2, pp. 400. ISSN 2077-0383. DOI 10.3390/jcm11020400.
- KANG, M. & MARTIN, A., 2017. Microbiome and colorectal cancer: Unraveling host-microbiota interactions in colitis-associated colorectal cancer development. *Seminars in Immunology*, vol. 32, pp. 3-13. ISSN 10445323. DOI 10.1016/j.smim.2017.04.003.
- KASSAMBARA, A., 2023. *rstatix: Pipe-Friendly Framework for Basic Statistical Tests* [online]. 1 February 2023. S.l.: s.n. [Accessed: 28 May 2025]. Available at: <https://cran.r-project.org/web/packages/rstatix/index.html>.

# UADE

## SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía

---

- KATZENMAIER, E.-M., FUCHS, V., WARNKEN, U., SCHNÖLZER, M., GEBERT, J., & KOPITZ, J., 2019. Deciphering the galectin-12 protein interactome reveals a major impact of galectin-12 on glutamine anaplerosis in colon cancer cells. *Experimental Cell Research*, vol. 379, no. 2, pp. 129-139. ISSN 0014-4827. DOI 10.1016/j.yexcr.2019.03.032.
- KAZ, A.M. and VENU, N., 2025. Diagnostic Methods and Biomarkers in Inflammatory Bowel Disease. *Diagnostics*, vol. 15, no. 11, pp. 1303. ISSN 2075-4418. DOI 10.3390/diagnostics15111303.
- KELM, M., QUIROS, M., AZCUTIA, V., BOERNER, K., CUMMINGS, R.D., NUSRAT, A., BRAZIL, J.C., and PARKOS, C.A., 2020. Targeting epithelium-expressed sialyl Lewis glycans improves colonic mucosal wound healing and protects against colitis. *JCI Insight*, vol. 5, no. 12, pp. e135843. ISSN 2379-3708. DOI 10.1172/jci.insight.135843.
- KLINKEN, B.J.-W.V., WAL, J.-W.G.V. der, EINERHAND, A.W.C., BÜLLER, H.A. and DEKKER, J., 1999. Sulphation and secretion of the predominant secretory human colonic mucin MUC2 in ulcerative colitis. *Gut*, vol. 44, no. 3, pp. 387-393. ISSN 0017-5749, 1468-3288. DOI 10.1136/gut.44.3.387.
- KOLDE, R., 2025. *pheatmap: Pretty Heatmaps* [online]. R. 2025. S.l.: s.n. Available at: <https://github.com/raivokolde/pheatmap>.
- KUDELKA, M.R., STOWELL, S.R., CUMMINGS, R.D., and NEISH, A.S., 2020. Intestinal epithelial glycosylation in homeostasis and gut microbiota interactions in IBD. *Nature Reviews Gastroenterology & Hepatology*, vol. 17, no. 10, pp. 597-617. ISSN 1759-5045, 1759-5053. DOI 10.1038/s41575-020-0331-7.
- KUFE, D.W., 2009. Mucins in cancer: function, prognosis and therapy. *Nature Reviews Cancer*, vol. 9, no. 12, pp. 874-885. ISSN 1474-175X, 1474-1768. DOI 10.1038/nrc2761.
- LANZILLOTA, R.G., 2022. *Search for glyco-biomarkers for patient stratification in inflammatory bowel diseases*. Bachelor's Thesis. S.l.: Universidad Argentina de la Empresa.
- LE BERRE, C., HONAP, S. and PEYRIN-BIROULET, L., 2023. Ulcerative colitis. *The Lancet*, vol. 402, no. 10401, pp. 571-584. ISSN 01406736. DOI 10.1016/S0140-6736(23)00966-2.
- LE, D.T., URAM, J.N., WANG, H., BARTLETT, B.R., KEMBERLING, H., EYRING, A.D., SKORA, A.D., LUBER, B.S., AZAD, N.S., LAHERU, D., BIEDRZYCKI, B., DONEOWER, R.C., ZAHEER, A., FISHER, G.A., CROCENZI, T.S., LEE, J.J., DUFFY, S.M., GOLDBERG, R.M., DE LA CHAPELLE, A., KOSHJI, M., BHAIJEE, F., HUEBNER, T., HRUBAN, R.H., WOOD, L.D., CUCA, N., PARDOLL, D.M., PAPADOPOULOS, N., KINZLER, K.W., ZHOU, S., CORNISH, T.C., TAUBE, J.M., ANDERS, R.A., ESHLEMAN, J.R., VOGELSTEIN, B., & DIAZ, L.A., 2015. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *The New England Journal of Medicine*, vol. 372, no. 26, pp. 2509-2520. ISSN 1533-4406. DOI 10.1056/NEJMoa1500596.
- LEITE-GOMES, E., DIAS, A.M., AZEVEDO, C.M., SANTOS-PEREIRA, B., MAGALHÃES, M., GARRIDO, M., AMORIM, R., LAGO, P., MARCOS-PINTO, R. and PINHO, S.S., 2022. Bringing to Light the Risk of Colorectal Cancer in Inflammatory Bowel Disease: Mucosal Glycosylation as a Key Player. *Inflammatory Bowel Diseases*, vol. 28, no. 6, pp. 947-962. ISSN 1078-0998. DOI 10.1093/ibd/izab291.

# UADE SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía

---

- LEPPKES, M. & NEURATH, M.F., 2020. Cytokines in inflammatory bowel diseases – Update 2020. *Pharmacological Research*, vol. 158, pp. 104835. ISSN 10436618. DOI 10.1016/j.phrs.2020.104835.
- LEWANDOWSKA, A., RUDZKI, G., LEWANDOWSKI, T., STRYJKOWSKA-GÓRA, A. and RUDZKI, S., 2022. Risk Factors for the Diagnosis of Colorectal Cancer. *Cancer Control*, vol. 29, pp. 10732748211056692. ISSN 1073-2748, 1526-2359. DOI 10.1177/10732748211056692.
- LEY, K., 2003. The role of selectins in inflammation and disease. *Trends in Molecular Medicine*, vol. 9, no. 6, pp. 263-268. ISSN 14714914. DOI 10.1016/S1471-4914(03)00071-6.
- LI, J., JI, Y., CHEN, N., DAI, L., and DENG, H., 2023. Colitis-associated carcinogenesis: crosstalk between tumors, immune cells and gut microbiota. *Cell & Bioscience*, vol. 13, no. 1, pp. 194. ISSN 2045-3701. DOI 10.1186/s13578-023-01139-8.
- LIBERZON, A., SUBRAMANIAN, A., PINCHBACK, R., THORVALDSDÓTTIR, H., TAMAYO, P., and MESIROV, J.P., 2011. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*, vol. 27, no. 12, pp. 1739-1740. ISSN 1367-4811, 1367-4803. DOI 10.1093/bioinformatics/btr260.
- LIN, A., ZHANG, J. and LUO, P., 2020. Crosstalk Between the MSI Status and Tumor Microenvironment in Colorectal Cancer. *Frontiers in Immunology*, vol. 11, pp. 2039. ISSN 1664-3224. DOI 10.3389/fimmu.2020.02039.
- LIU, F.-T. and RABINOVICH, G.A., 2005. Galectins as modulators of tumour progression. *Nature Reviews Cancer*, vol. 5, no. 1, pp. 29-41. ISSN 1474-1768. DOI 10.1038/nrc1527.
- LIU, H., PATEL, N.R., WALTER, L., INGERSOLL, S., SITARAMAN, S.V., and GARG, P., 2013. Constitutive expression of MMP9 in intestinal epithelium worsens murine acute colitis and is associated with increased levels of proinflammatory cytokine Kc. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, vol. 304, no. 9, pp. G793-G803. ISSN 0193-1857, 1522-1547. DOI 10.1152/ajpgi.00249.2012.
- LOW, D., SUBRAMANIAM, R., LIN, L., AOMATSU, T., MIZOGUCHI, A., NG, A., DEGRUTTOLA, A.K., LEE, C.G., ELIAS, J.A., ANDOH, A., MINO-KENUDSON, M., and MIZOGUCHI, E., 2015. Chitinase 3-like 1 induces survival and proliferation of intestinal epithelial cells during chronic inflammation and colitis-associated cancer by regulating S100A9. *Oncotarget*, vol. 6, no. 34, pp. 36535-36550. ISSN 1949-2553. DOI 10.18632/oncotarget.5440.
- LU, Y., SU, Y., WANG, N., LI, D., ZHANG, H., and XU, H., 2024. Identification of O-glycosylation related genes and subtypes in ulcerative colitis based on machine learning. In: A. PANDEY (ed.), *PLOS ONE*, vol. 19, no. 12, pp. e0311495. ISSN 1932-6203. DOI 10.1371/journal.pone.0311495.
- LYNCH, W.D. and HSU, R., 2023. Ulcerative Colitis. *StatPearls* [online]. Treasure Island (FL): StatPearls Publishing, [Accessed: 28 June 2025]. Available at: <http://www.ncbi.nlm.nih.gov/books/NBK459282/>. NBK459282

MALLER, S.M., 2019. *Galectin-12-mediated regulatory circuits in the interaction between adipose tissue and vascularization programs* [online]. Doctoral Thesis. S.I.: Universidad de Buenos Aires - Facultad de Ciencias Exactas y Naturales. Available at: [https://bibliotecadigital.exactas.uba.ar/download/tesis/tesis\\_n6766\\_Maller.pdf](https://bibliotecadigital.exactas.uba.ar/download/tesis/tesis_n6766_Maller.pdf).

MALLER, S.M., CAGNONI, A.J., BANNOUD, N., SIGAUT, L., PÉREZ SÁEZ, J.M., PIETRASANTA, L.I., YANG, R., LIU, F., CROCI, D.O., DI LELLA, S., SUNDBLAD, V., RABINOVICH, G.A., and MARIÑO, K.V., 2020. An adipose tissue galectin controls endothelial cell function via preferential recognition of 3-fucosylated glycans. *The FASEB Journal*, vol. 34, no. 1, pp. 735-753. ISSN 0892-6638, 1530-6860. DOI 10.1096/fj.201901817r.

MAMOOR, S., 2023. *Differential expression of SLC26A2 in colorectal cancer*. [online]. 8 March 2023. S.I.: Center for Open Science. [Accessed: 16 July 2025]. Available in: <https://osf.io/5hdnv/>.

MARIÑO, K., SALDOVA, R., ADAMCZYK, B. and RUDD, P.M., 2011. Changes in Serum N-Glycosylation Profiles: Functional Significance and Potential for Diagnostics. In: A. PILAR RAUTER and T. LINDHORST (eds.), *Carbohydrate Chemistry* [online]. 1. S.I.: The Royal Society of Chemistry, pp. 57-93. [Accessed: 26 May 2025]. ISBN 978-1-84973-154-6. Available in: <https://books.rsc.org/books/monograph/1911/chapter-abstract/2511990/>.

MARIÑO, K.V., CAGNONI, A.J., CROCI, D.O. and RABINOVICH, G.A., 2023. Targeting galectin-driven regulatory circuits in cancer and fibrosis. *Nature Reviews Drug Discovery*, vol. 22, no. 4, pp. 295-316. ISSN 1474-1776, 1474-1784. DOI 10.1038/s41573-023-00636-2.

MARKOWITZ, S.D. and BERTAGNOLLI, M.M., 2009. Molecular origins of cancer: Molecular basis of colorectal cancer. *The New England Journal of Medicine*, vol. 361, no. 25, pp. 2449-2460. ISSN 1533-4406. DOI 10.1056/NEJMra0804588.

MÁRMOL, I., SÁNCHEZ-DE-DIEGO, C., PRADILLA DIESTE, A., CERRADA, E. and RODRIGUEZ YOLDI, M., 2017. Colorectal Carcinoma: A General Overview and Future Perspectives in Colorectal Cancer. *International Journal of Molecular Sciences*, vol. 18, no. 1, pp. 197. ISSN 1422-0067. DOI 10.3390/ijms18010197.

MASSARO, M., 2025. *The Galectin-4-glycan axis and its relevance in intestinal inflammation*. S.I.: Universidad de Buenos Aires - Facultad de Ciencias Exactas y Naturales.

MAZUR, M., WŁODARCZYK, J., ŚWIERCZYŃSKI, M., KORDEK, R., GRZYBOWSKI, M.M., OLCZAK, J., and FICHNA, J., 2022. The Anti-Inflammatory Effect of Acidic Mammalian Chitinase Inhibitor OAT-177 in DSS-Induced Mouse Model of Colitis. *International Journal of Molecular Sciences*, vol. 23, no. 4, pp. 2159. ISSN 1422-0067. DOI 10.3390/ijms23042159.

MAZZEO, C.A., 2021. *Search for glycoimmunological markers for prognosis and treatment of colorectal cancer*. S.I.: Universidad Argentina de la Empresa.

MCGRANAHAN, N., FURNESS, A.J.S., ROSENTHAL, R., RAMSKOV, S., LYNGAA, R., SAINI, S.K., JAMAL-HANJANI, M., WILSON, G.A., BIRKBAK, N.J., HILEY, C.T., WATKINS, T.B.K., SHAFI, S., MURUGAESU, N., MITTER, R., AKARCA, A.U., LINARES, J., MARAFIOTI, T., HENRY, J.Y., VAN ALLEN, E.M., MIAO, D., SCHILLING, B., SCHADENDORF, D., GARRAWAY, L.A., MAKAROV, V., RIZVI, N.A., SNYDER, A.,

# UADE

## SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía

---

HELLMANN, M.D., MERGHOUB, T., WOLCHOK, J.D., SHUKLA, S.A., WU, C.J., PEGGS, K.S., CHAN, T.A., HADRUP, S.R., QUEZADA, S.A. and SWANTON, C., 2016. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science (New York, N.Y.)*, vol. 351, no. 6280, pp. 1463-1469. ISSN 1095-9203. DOI 10.1126/science.aaf1490.

MIZOGUCHI, E., 2006. Chitinase 3-Like-1 Exacerbates Intestinal Inflammation by Enhancing Bacterial Adhesion and Invasion in Colonic Epithelial Cells. *Gastroenterology*, vol. 130, no. 2, pp. 398-411. ISSN 00165085. DOI 10.1053/j.gastro.2005.12.007.

MOROSI, L.G., CUTINE, A.M., CAGNONI, A.J., MANSELLE-COCCO, M.N., CROCI, D.O., MERLO, J.P., MORALES, R.M., MAY, M., PÉREZ-SÁEZ, J.M., GIROTTI, M.R., MÉNDEZ-HUERGO, S.P., PUCCI, B., GIL, A.H., HUERNOS, S.P., DOCENA, G.H., SAMBUELLI, A.M., TOSCANO, M.A., RABINOVICH, G.A. and MARIÑO, K.V., 2021. Control of intestinal inflammation by glycosylation-dependent lectin-driven immunoregulatory circuits. *Science Advances*, vol. 7, no. 25, pp. eabf8630. DOI 10.1126/sciadv.abf8630.

MUNKHOLM, P., 2003. Review article: the incidence and prevalence of colorectal cancer in inflammatory bowel disease. *Alimentary Pharmacology & Therapeutics*, vol. 18, no. s2, pp. 1-5. ISSN 0269-2813, 1365-2036. DOI 10.1046/j.1365-2036.18.s2.2.x.

MUNKLEY, J. & ELLIOTT, D.J., 2016. Hallmarks of glycosylation in cancer. *Oncotarget*, vol. 7, no. 23, pp. 35478-35489. ISSN 1949-2553. DOI 10.18632/oncotarget.8155.

MURNANE, M.J., CAI, J., SHUJA, S., MCANENY, D., and WILLETT, J.B., 2011. Active matrix metalloproteinase-2 activity discriminates colonic mucosa, adenomas with and without high-grade dysplasia, and cancers. *Human Pathology*, vol. 42, no. 5, pp. 688-701. ISSN 0046-8177. DOI 10.1016/j.humpath.2010.08.021.

NAKAMURA, K., YAMASHITA, K., SAWAKI, H., WARAYA, M., KATOH, H., NAKAYAMA, N., KAWAMATA, H., NISHIMIYA, H., EMA, A., NARIMATSU, H., and WATANABE, M., 2015. Aberrant methylation of GCNT2 is tightly related to lymph node metastasis of primary CRC. *Anticancer Research*, vol. 35, no. 3, pp. 1411-1421. ISSN 1791-7530.

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION, 2025. National Center for Biotechnology Information. [online]. [Accessed: 6 July 2025]. Available in: <https://www.ncbi.nlm.nih.gov/>.

NEURATH, M.F., 2014. Cytokines in inflammatory bowel disease. *Nature Reviews Immunology*, vol. 14, no. 5, pp. 329-342. ISSN 1474-1733, 1474-1741. DOI 10.1038/nri3661.

NISHIHARA, Y., OGINO, H., TANAKA, M., IHARA, E., FUKAURA, K., NISHIOKA, K., CHINEN, T., TANAKA, Y., NAKAYAMA, J., KANG, D., and OGAWA, Y., 2021. Mucosa-associated gut microbiota reflects clinical course of ulcerative colitis. *Scientific Reports*, vol. 11, no. 1, pp. 13743. ISSN 2045-2322. DOI 10.1038/s41598-021-92870-0.

OHANA, E., SHCHEYNIKOV, N., PARK, M., and MUALLEM, S., 2012. Solute Carrier Family 26 Member a2 (Slc26a2) Protein Functions as an Electroneutral SO<sub>4</sub><sup>2-</sup>/OH<sup>-</sup>/Cl<sup>-</sup> Exchanger Regulated by Extracellular Cl<sup>-</sup>. *Journal of Biological Chemistry*, vol. 287, no. 7, pp. 5122-5132. ISSN 00219258. DOI 10.1074/jbc.M111.297192.

ORDÁS, I., ECKMANN, L., TALAMINI, M., BAUMGART, D.C. and SANDBORN, W.J., 2012. Ulcerative colitis. *The Lancet*, vol. 380, no. 9853, pp. 1606-1619. ISSN 01406736. DOI 10.1016/S0140-6736(12)60150-0.

OTÁLORA-OTÁLORA, B.A., PAYÁN-GÓMEZ, C., LÓPEZ-RIVERA, J.J., PEDROZA-ACONCHA, N.B., ARBOLEDA-MOJICA, S.L., ARISTIZÁBAL-GUZMÁN, C., ISAZA-RUGET, M.A. and ÁLVAREZ-MORENO, C.A., 2024. Interplay of Transcriptomic Regulation, Microbiota, and Signaling Pathways in Lung and Gut Inflammation-Induced Tumorigenesis. *Cells*, vol. 14, no. 1, pp. 1. ISSN 2073-4409. DOI 10.3390/cells14010001.

PAPA-GOBBI, R., MUGLIA, C.I., ROCCA, A., CURCIARELLO, R., SAMBUELLI, A.M., YANTORNO, M., CORREA, G., MOROSI, L.G., DI SABATINO, A., BIANCHERI, P., MACDONALD, T.T., TOSCANO, M.A., MARIÑO, K.V., RABINOVICH, G.A., and DOCENA, G.H., 2021. Spatiotemporal regulation of galectin-1-induced T-cell death in lamina propria from Crohn's disease and ulcerative colitis patients. *Apoptosis*, vol. 26, no. 5-6, pp. 323-337. ISSN 1360-8185, 1573-675X. DOI 10.1007/s10495-021-01675-z.

PAPAMICHAEL, D., AUDISIO, R.A., GLIMELIUS, B., DE GRAMONT, A., GLYNNE-JONES, R., HALLER, D., KÖHNE, C.-H., ROSTOFT, S., LEMMENS, V., MITRY, E., RUTTEN, H., SARGENT, D., SASTRE, J., SEYMOUR, M., STARLING, N., VAN CUTSEM, E., and AAPRO, M., 2015. Treatment of colorectal cancer in older patients: International Society of Geriatric Oncology (SIOG) consensus recommendations 2013. *Annals of Oncology*, vol. 26, no. 3, pp. 463-476. ISSN 09237534. DOI 10.1093/annonc/mdu253.

PARANG, B., BARRETT, C.W., and WILLIAMS, C.S., 2016. AOM/DSS Model of Colitis-Associated Cancer. In: A.I. IVANOV (ed.), *Gastrointestinal Physiology and Diseases* [online]. New York, NY: Springer New York, Methods in Molecular Biology, pp. 297-307. [Consulted: 30 June 2025]. ISBN 978-1-4939-3601-4. Available at: [http://link.springer.com/10.1007/978-1-4939-3603-8\\_26](http://link.springer.com/10.1007/978-1-4939-3603-8_26).

PATEL, A., 2015. Field cancerisation in colorectal cancer: A new frontier or pastures past? *World Journal of Gastroenterology*, vol. 21, no. 13, pp. 3763. ISSN 1007-9327. DOI 10.3748/wjg.v21.i13.3763.

PEKOW, J., DOUGHERTY, U., HUANG, Y., GOMETZ, E., NATHANSON, J., COHEN, G., LEVY, S., KOCHERGINSKY, M., VENU, N., WESTERHOFF, M., HART, J., NOFFSINGER, A.E., HANAUER, S.B., HURST, R.D., FICHERA, A., JOSEPH, L.J., LIU, Q., and BISSONNETTE, M., 2013. Gene signature distinguishes patients with chronic ulcerative colitis harboring remote neoplastic lesions. *Inflammatory Bowel Diseases*, vol. 19, no. 3, pp. 461-470. ISSN 1536-4844. DOI 10.1097/MIB.0b013e3182802bac.

PÉREZ, A., ANDRADE-DA-COSTA, J., DE SOUZA, W., DE SOUZA FERREIRA, M., BORONI, M., DE OLIVEIRA, I., FREIRE-NETO, C., FERNANDES, P., DE LANNA, C., SOUZA-SANTOS, P., MORGADO-D AZ, J. and DE-FREITAS-JUNIOR, J.C., 2020. N-glycosylation and receptor tyrosine kinase signaling affect claudin-3 levels in colorectal cancer cells. *Oncology Reports* [online], [Accessed: 2 July 2025]. ISSN 1021-335X, 1791-2431. DOI 10.3892/or.2020.7727. Available in: <http://www.spandidos-publications.com/10.3892/or.2020.7727>.

# UADE

## SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía

---

- PEREZ, M., CHAKRABORTY, A., LAU, L.S., MOHAMMED, N.B.B., & DIMITROFF, C.J., 2021. Melanoma-associated glycosyltransferase GCNT2 as an emerging biomarker and therapeutic target. *The British Journal of Dermatology*, vol. 185, no. 2, pp. 294-301. ISSN 1365-2133. DOI 10.1111/bjd.19891.
- PERŠE, M. & CERAR, A., 2012. Dextran Sodium Sulphate Colitis Mouse Model: Traps and Tricks. *Journal of Biomedicine and Biotechnology*, vol. 2012, pp. 1-13. ISSN 1110-7243, 1110-7251. DOI 10.1155/2012/718617.
- PINHO, S.S. and RABINOVICH, G.A., 2025. The glycoimmune landscape in health and disease. *Seminars in Immunology*, vol. 78, pp. 101965. ISSN 10445323. DOI 10.1016/j.smim.2025.101965.
- PLANELL, N., LOZANO, J.J., MORA-BUCH, R., MASAMUNT, M.C., JIMENO, M., ORDÁS, I., ESTELLER, M., RICART, E., PIQUÉ, J.M., PANÉS, J. and SALAS, A., 2013. Transcriptional analysis of the intestinal mucosa of patients with ulcerative colitis in remission reveals lasting epithelial cell alterations. *Gut*, vol. 62, no. 7, pp. 967-976. ISSN 0017-5749, 1468-3288. DOI 10.1136/gutjnl-2012-303333.
- POLOSUKHINA, D., SINGH, K., ASIM, M., BARRY, D.P., ALLAMAN, M.M., HARDBOWER, D.M., PIAZUELO, M.B., WASHINGTON, M.K., GOBERT, A.P., WILSON, K.T., and COBURN, L.A., 2021. CCL11 exacerbates colitis and inflammation-associated colon tumorigenesis. *Oncogene*, vol. 40, no. 47, pp. 6540-6546. ISSN 0950-9232, 1476-5594. DOI 10.1038/s41388-021-02046-3.
- PONCELET, L., DUMONT, J.-E., MIOT, F., & DE DEKEN, X., 2019. The Dual Oxidase Duox2 stabilized with DuoxA2 in an enzymatic complex at the surface of the cell produces extracellular H<sub>2</sub>O<sub>2</sub> able to induce DNA damage in an inducible cellular model. *Experimental Cell Research*, vol. 384, no. 1, pp. 111620. ISSN 00144827. DOI 10.1016/j.yexcr.2019.111620.
- POPOV, J., CAPUTI, V., NANDEESHA, N., RODRIGUEZ, D.A., & PAI, N., 2021. Microbiota-Immune Interactions in Ulcerative Colitis and Colitis Associated Cancer and Emerging Microbiota-Based Therapies. *International Journal of Molecular Sciences*, vol. 22, no. 21, pp. 11365. ISSN 1422-0067. DOI 10.3390/ijms222111365.
- PORTER, R.J., ARENDS, M.J., CHURCHHOUSE, A.M.D., and DIN, S., 2021. Inflammatory Bowel Disease-Associated Colorectal Cancer: Translational Risks from Mechanisms to Medicines. *Journal of Crohn's and Colitis*, vol. 15, no. 12, pp. 2131-2141. ISSN 1873-9946, 1876-4479. DOI 10.1093/ecco-jcc/jjab102.
- POSIT PBC, 2025. *rstudio/shiny* [online]. R. 27 May 2025. S.l.: RStudio. [Consulted: 27 May 2025]. Available at: <https://github.com/rstudio/shiny>.
- QIAN, L., HU, S., ZHAO, H., HAN, Y., DAI, C., ZAN, X., ZHI, Q., and XU, C., 2025. The Diagnostic Significance of SLC26A2 and Its Potential Role in Ulcerative Colitis. *Biomedicines*, vol. 13, no. 2, pp. 461. ISSN 2227-9059. DOI 10.3390/biomedicines13020461.
- QIAO, Q., BAI, R., SONG, W., GAO, H., ZHANG, M., LU, J., HONG, M., ZHANG, X., SUN, P., ZHANG, Q., and ZHAO, P., 2021. Human  $\alpha$ -defensin 5 suppressed colon cancer growth by targeting PI3K pathway. *Experimental Cell Research*, vol. 407, no. 2, pp. 112809. ISSN 00144827. DOI 10.1016/j.yexcr.2021.112809.

QIU, Z., WANG, Y., ZHANG, Z., QIN, R., PENG, Y., TANG, W., XI, Y., TIAN, G., and ZHANG, Y., 2022. Roles of intercellular cell adhesion molecule-1 (ICAM-1) in colorectal cancer: expression, functions, prognosis, tumorigenesis, polymorphisms and therapeutic implications. *Frontiers in Oncology* [online], vol. 12. [Accessed: 8 July 2025]. ISSN 2234-943X. DOI 10.3389/fonc.2022.1052672. Available at: <https://www.frontiersin.org/articles/10.3389/fonc.2022.1052672/full>.

R CORE TEAM, 2025. *stats: The R Stats Package* [online]. R. 2025. S.l.: s.n. Available at: <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/00Index.html>.

RA, H.-J., HARJU-BAKER, S., ZHANG, F., LINHARDT, R.J., WILSON, C.L., and PARKS, W.C., 2009. Control of Promatrilysin (MMP7) Activation and Substrate-specific Activity by Sulfated Glycosaminoglycans. *Journal of Biological Chemistry*, vol. 284, no. 41, pp. 27924-27932. ISSN 00219258. DOI 10.1074/jbc.M109.035147.

RADZIEJEWSKA, I., 2024. Tumor-associated carbohydrate antigens of MUC1 – Implication in cancer development. *Biomedicine & Pharmacotherapy*, vol. 174, pp. 116619. ISSN 07533322. DOI 10.1016/j.biopha.2024.116619.

RAMASUNDARA, M., LEACH, S.T., LEMBERG, D.A., and DAY, A.S., 2009. Defensins and inflammation: The role of defensins in inflammatory bowel disease. *Journal of Gastroenterology and Hepatology*, vol. 24, no. 2, pp. 202-208. ISSN 0815-9319, 1440-1746. DOI 10.1111/j.1440-1746.2008.05772.x.

RAUSCH, P., REHMAN, A., KÜNZEL, S., HÄSLER, R., OTT, S.J., SCHREIBER, S., ROSENSTIEL, P., FRANKE, A., and BAINES, J.F., 2011. Colonic mucosa-associated microbiota is influenced by an interaction of Crohn's disease and *FUT2* (*Secretor*) genotype. *Proceedings of the National Academy of Sciences*, vol. 108, no. 47, pp. 19030-19035. ISSN 0027-8424, 1091-6490. DOI 10.1073/pnas.1106408108.

REILY, C., STEWART, T.J., RENFROW, M.B., & NOVAK, J., 2019. Glycosylation in health and disease. *Nature Reviews Nephrology*, vol. 15, no. 6, pp. 346-366. ISSN 1759-5061, 1759-507X. DOI 10.1038/s41581-019-0129-4.

RIAZ, N., MORRIS, L., HAVEL, J.J., MAKAROV, V., DESRICHARD, A., & CHAN, T.A., 2016. The role of neoantigens in response to immune checkpoint blockade. *International Immunology*, vol. 28, no. 8, pp. 411-419. ISSN 1460-2377, 0953-8178. DOI 10.1093/intimm/dxw019.

RITCHIE, M.E., PHIPSON, B., WU, D., HU, Y., LAW, C.W., SHI, W., and SMYTH, G.K., 2015. Limma Powers Differential Expression Analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, vol. 43, no. 7, pp. e47-e47. ISSN 1362-4962, 0305-1048. DOI 10.1093/nar/gkv007.

RIZVI, N.A., HELLMANN, M.D., SNYDER, A., KVISTBORG, P., MAKAROV, V., HAVEL, J.J., LEE, W., YUAN, J., WONG, P., HO, T.S., MILLER, M.L., REKHTMAN, N., MOREIRA, A.L., IBRAHIM, F., BRUGGEMAN, C., GASMI, B., ZAPPASODI, R., MAEDA, Y., SANDER, C., GARON, E.B., MERGHOUB, T., WOLCHOK, J.D., SCHUMACHER, T.N. and CHAN, T.A., 2015. Cancer immunology. Mutational landscape determines sensitivity to PD-1

blockade in non-small cell lung cancer. *Science (New York, N.Y.)*, vol. 348, no. 6230, pp. 124-128. ISSN 1095-9203. DOI 10.1126/science.aaa1348.

RODA, G., NARULA, N., PINOTTI, R., SKAMNELOS, A., KATSANOS, K.H., UNGARO, R., BURISCH, J., TORRES, J., and COLOMBEL, J.-F., 2017. Systematic review with meta-analysis: proximal disease extension in limited ulcerative colitis. *Alimentary Pharmacology & Therapeutics*, vol. 45, no. 12, pp. 1481-1492. ISSN 0269-2813, 1365-2036. DOI 10.1111/apt.14063.

ROSENBERG, J.E., HOFFMAN-CENSITS, J., POWLES, T., VAN DER HEIJDEN, M.S., BALAR, A.V., NECCHI, A., DAWSON, N., O'DONNELL, P.H., BALMANOUKIAN, A., LORIOT, Y., SRINIVAS, S., RETZ, M.M., GRIVAS, P., JOSEPH, R.W., GALSKY, M.D., FLEMING, M.T., PETRYLAK, D.P., PEREZ-GRACIA, J.L., BURRIS, H.A., CASTELLANO, D., CANIL, C., BELLMUNT, J., BAJORIN, D., NICKLES, D., BOURGON, R., FRAMPTON, G.M., CUI, N., MARIATHASAN, S., ABIDOYE, O., FINE, G.D., and DREICER, R., 2016. Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: a single-arm, multicentre, phase 2 trial. *Lancet (London, England)*, vol. 387, no. 10031, pp. 1909-1920. ISSN 1474-547X. DOI 10.1016/S0140-6736(16)00561-4.

ROSHANDEL, G., GHASEMI-KEBRIA, F. and MALEKZADEH, R., 2024. Colorectal Cancer: Epidemiology, Risk Factors, and Prevention. *Cancers*, vol. 16, no. 8, pp. 1530. ISSN 2072-6694. DOI 10.3390/cancers16081530.

SABIT, H., PAWLIK, T.M., ABDEL-GHANY, S., and ARNETH, B., 2024. Defensins: Exploring Their Opposing Roles in Colorectal Cancer Progression. *Cancers*, vol. 16, no. 15, pp. 2622. ISSN 2072-6694. DOI 10.3390/cancers16152622.

SALAS CAUDEVILLA, A., 2007. Evaluation of dysplasia in digestive diseases. *Gastroenterology and Hepatology*, vol. 30, no. 10, pp. 602-611. ISSN 0210-5705. DOI 10.1157/13112598.

SAMBUELLI, A.M., NEGREIRA, S., GIL, A., GONCALVES, S., CHAVERO, P., TIRADO, P., BELLICOSO, M. and HUERNOS, S., 2019. Management of inflammatory bowel disease. Review and treatment algorithms. *ACTA Latin American Gastroenterological Journal* [online], vol. 49. ISSN 2469-1119. Available at: <https://www.actagastro.org/numeros-anteriores/2019/Vol-49-S2/Vol49S2.pdf>.

SANÉ, A.T., SEIDMAN, E., PERETTI, N., KLEME, M.L., DELVIN, E., DESLANDRES, C., GAROFALO, C., SPAHIS, S. and LEVY, E., 2017. Understanding Chylomicron Retention Disease Through Sar1b Gtpase Gene Disruption: Insight From Cell Culture. *Arteriosclerosis, Thrombosis, and Vascular Biology*, vol. 37, no. 12, pp. 2243-2251. ISSN 1079-5642, 1524-4636. DOI 10.1161/ATVBAHA.117.310121.

SANTANA, P.T., ROSAS, S.L.B., RIBEIRO, B.E., MARINHO, Y. and DE SOUZA, H.S.P., 2022. Dysbiosis in Inflammatory Bowel Disease: Pathogenic Role and Potential Therapeutic Targets. *International Journal of Molecular Sciences*, vol. 23, no. 7, pp. 3464. ISSN 1422-0067. DOI 10.3390/ijms23073464.

SATELLI, A., RAO, P.S., THIRUMALA, S., and RAO, U.S., 2011. Galectin-4 functions as a tumor suppressor of human colorectal cancer. *International Journal of Cancer*, vol. 129, no. 4, pp. 799-809. ISSN 0020-7136, 1097-0215. DOI 10.1002/ijc.25750.

SCHLESSINGER, A., ZATORSKI, N., HUTCHINSON, K., & COLAS, C., 2023. Targeting SLC transporters: small molecules as modulators and therapeutic opportunities. *Trends in Biochemical Sciences*, vol. 48, no. 9, pp. 801-814. ISSN 09680004. DOI 10.1016/j.tibs.2023.05.011.

SCHROEDER, K.W., TREMAINE, W.J. and ILSTRUP, D.M., 1987. Coated Oral 5-Aminosalicylic Acid Therapy for Mildly to Moderately Active Ulcerative Colitis. *New England Journal of Medicine*, vol. 317, no. 26, pp. 1625-1629. ISSN 0028-4793, 1533-4406. DOI 10.1056/NEJM198712243172603.

SCHULTHESS, J., PANDEY, S., CAPITANI, M., RUE-ALBRECHT, K.C., ARNOLD, I., FRANCHINI, F., CHOMKA, A., ILOTT, N.E., JOHNSTON, D.G.W., PIRES, E., MCCULLAGH, J., SANSOM, S.N., ARANCIBIA-CÁRCAMO, C.V., UHLIG, H.H., and POWRIE, F., 2019. The Short Chain Fatty Acid Butyrate Imprints an Antimicrobial Program in Macrophages. *Immunity*, vol. 50, no. 2, pp. 432-445.e7. ISSN 10747613. DOI 10.1016/j.jimmuni.2018.12.018.

SCHUMACHER, T.N. and SCHREIBER, R.D., 2015. Neoantigens in cancer immunotherapy. *Science*, vol. 348, no. 6230, pp. 69-74. DOI 10.1126/science.aaa4971.

SHAN, M., YANG, D., DOU, H., and ZHANG, L., 2019. Fucosylation in cancer biology and its clinical applications. *Progress in Molecular Biology and Translational Science* [online]. S.l.: Elsevier, pp. 93-119. [Consulted: 2 July 2025]. ISBN 978-0-12-817738-9. Available at: <https://linkinghub.elsevier.com/retrieve/pii/S187711731930002X>.

SHI, J., 2007. Defensins and Paneth cells in inflammatory bowel disease: *Inflammatory Bowel Diseases*, vol. 13, no. 10, pp. 1284-1292. ISSN 1078-0998. DOI 10.1002/ibd.20197.

SHI, L., HAN, X., LI, J.-X., LIAO, Y.-T., KOU, F.-S., WANG, Z.-B., SHI, R., ZHAO, X.-J., SUN, Z.-M., and HAO, Y., 2020. Identification of differentially expressed genes in ulcerative colitis and verification in a colitis mouse model by bioinformatics analyses. *World Journal of Gastroenterology*, vol. 26, no. 39, pp. 5983-5996. ISSN 1007-9327. DOI 10.3748/wjg.v26.i39.5983.

SIEGEL, R.L., MILLER, K.D., and JEMAL, A., 2018. Cancer statistics, 2018. *CA: a cancer journal for clinicians*, vol. 68, no. 1, pp. 7-30. ISSN 1542-4863. DOI 10.3322/caac.21442.

SIEVERT, C., 2025. shinyapps.io. *Shiny Apps* [online]. [Accessed: 6 July 2025]. Available at: <https://www.shinyapps.io/>.

SILOŞI, I., BOLDEANU, M.V., MOGOANTĂ, S.Ş., GHILUŞI, M., COJOCARU, M., BICIUŞCĂ, V., COJOCARU, I.M., AVRĂMESCU, C.S., GHEONEA, D.I., SILOŞI, C.A. and TURCULEANU, A., 2014. Matrix metalloproteinases (MMP-3 and MMP-9) implication in the pathogenesis of inflammatory bowel disease (IBD). *Romanian Journal of Morphology and Embryology = Revue Roumaine De Morphologie Et Embryologie*, vol. 55, no. 4, pp. 1317-1324. ISSN 2066-8279.

SILVA, M.C., FERNANDES, Â., OLIVEIRA, M., RESENDE, C., CORREIA, A., DE-FREITAS-JUNIOR, J.C., LAVELLE, A., ANDRADE-DA-COSTA, J., LEANDER, M., XAVIER-FERREIRA, H., BESSA, J., PEREIRA, C., HENRIQUE, R.M., CARNEIRO, F., DINIS-RIBEIRO, M., MARCOS-PINTO, R., LIMA, M., LEPEÑIES, B., SOKOL, H., MACHADO, J.C., VILANOVA, M. and PINHO, S.S., 2020. Glycans as Immune Checkpoints: Removal of Branched N-glycans Enhances Immune Recognition Preventing Cancer Progression. *Cancer Immunology Research*, vol. 8, no. 11, pp. 1407-1425. ISSN 2326-6066, 2326-6074. DOI 10.1158/2326-6066.CIR-20-0264.

SILVERBERG, M.S., SATSANGI, J., AHMAD, T., ARNOTT, I.D.R., BERNSTEIN, C.N., BRANT, S.R., CAPRILLI, R., COLOMBEL, J.-F., GASCHE, C., GEBOES, K., JEWELL, D.P., KARBAN, A., LOFTUS, E.V., PEÑA, A.S., RIDDELL, R.H., SACHAR, D.B., SCHREIBER, S., STEINHART, A.H., TARGAN, S.R., VERMEIRE, S. and WARREN, B.F., 2005. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. *Canadian Journal of Gastroenterology = Journal Canadien De Gastroenterologie*, vol. 19 Suppl A, pp. 5A-36A. ISSN 1916-7237. DOI 10.1155/2005/269076.

ŠIMURINA, M., DE HAAN, N., VUČKOVIĆ, F., KENNEDY, N.A., ŠTAMBUK, J., FALCK, D., TRBOJEVIĆ-AKMAČIĆ, I., CLERC, F., RAZDOROV, G., KHON, A., LATIANO, A., D'INCÀ, R., DANESE, S., TARGAN, S., LANDERS, C., DUBINSKY, M., MCGOVERN, D.P.B., ANNESE, V., WUHRER, M., LAUC, G., CAMPBELL, H., ZOLDOŠ, V., PERMBERTON, I.K., KOLARICH, D., FERNANDES, D.L., THEORODOROU, E., MERRICK, V., SPENCER, D.I., GARDNER, R.A., DORAN, R., SHUBHAKAR, A., BOYAPATI, R., RUDAN, I., LIONETTI, P., KRIŠTIĆ, J., NOVOKMET, M., PUČIĆ-BAKOVIĆ, M., GORNIK, O., ANDRIULLI, A., CANTORO, L., STURNILO, G., FIORINO, G., MANETTI, N., ARNOTT, I.D., NOBLE, C.L., LEES, C.W., SHAND, A.G., HO, G.-T., DUNLOP, M.G., MURPHY, L., GIBSON, J., EVENDEN, L., WROBEL, N., GILCHRIST, T., FAWKES, A., KAMMEIJER, G.S.M., VOJTA, A., SAMARŽIJA, I., MARKULIN, D., KLASIĆ, M., DOBRINIĆ, P., AULCHENKO, Y., VAN DEN HEUVE, T., JONKERS, D., and PIERIK, M., 2018. Glycosylation of Immunoglobulin G Associates With Clinical Features of Inflammatory Bowel Diseases. *Gastroenterology*, vol. 154, no. 5, pp. 1320-1333.e10. ISSN 00165085. DOI 10.1053/j.gastro.2018.01.002.

SINGH, V., KAUR, R., KUMARI, P., PASRICHA, C., and SINGH, R., 2023. ICAM-1 and VCAM-1: Gatekeepers in various inflammatory and cardiovascular disorders. *Clínica Chimica Acta*, vol. 548, pp. 117487. ISSN 00098981. DOI 10.1016/j.cca.2023.117487.

SJÖSTEDT, E., ZHONG, W., FAGERBERG, L., KARLSSON, M., MITSIOS, N., ADORI, C., OKSVOLD, P., EDFORS, F., LIMISZEWSKA, A., HIKMET, F., HUANG, J., DU, Y., LIN, L., DONG, Z., YANG, L., LIU, X., JIANG, H., XU, X., WANG, J., YANG, H., BOLUND, L., MARDINOGLU, A., ZHANG, C., VON FEILITZEN, K., LINDSKOG, C., PONTÉN, F., LUO, Y., HÖKFELT, T., UHLÉN, M., & MULDER, J., 2020. An atlas of the protein-coding genes in the human, pig, and mouse brain. *Science*, vol. 367, no. 6482, pp. eaay5947. ISSN 0036-8075, 1095-9203. DOI 10.1126/science.aay5947.

SMITH, B.A.H. and BERTOZZI, C.R., 2021a. Author Correction: The clinical impact of glycobiology: targeting selectins, Siglecs and mammalian glycans. *Nature Reviews Drug*

*Discovery*, vol. 20, no. 3, pp. 244-244. ISSN 1474-1776, 1474-1784. DOI 10.1038/s41573-021-00160-1.

SMITH, B.A.H. and BERTOZZI, C.R., 2021b. The clinical impact of glycobiology: targeting selectins, S Igcs and mammalian glycans. *Nature Reviews Drug Discovery*, vol. 20, no. 3, pp. 217-243. ISSN 1474-1776, 1474-1784. DOI 10.1038/s41573-020-00093-1.

SNYDER, A., MAKAROV, V., MERGHOUB, T., YUAN, J., ZARETSKY, J.M., DESRICHARD, A., WALSH, L.A., POSTOW, M.A., WONG, P., HO, T.S., HOLLMANN, T.J., BRUGGEMAN, C., KANNAN, K., LI, Y., ELIPENAHLI, C., LIU, C., HARBISON, C.T., WANG, L., RIBAS, A., WOLCHOK, J.D., and CHAN, T.A., 2014. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *The New England Journal of Medicine*, vol. 371, no. 23, pp. 2189-2199. ISSN 1533-4406. DOI 10.1056/NEJMoa1406498.

SONG, M., PAN, Q., YANG, J., HE, J., ZENG, J., CHENG, S., HUANG, Y., ZHOU, Z.-Q., ZHU, Q., YANG, C., HAN, Y., TANG, Y., CHEN, H., WENG, D.-S., and XIA, J.-C., 2020. Galectin-3 favours tumour metastasis via the activation of  $\beta$ -catenin signalling in hepatocellular carcinoma. *British Journal of Cancer*, vol. 123, no. 10, pp. 1521-1534. ISSN 0007-0920, 1532-1827. DOI 10.1038/s41416-020-1022-4.

SUBRAMANIAN, A., TAMAYO, P., MOOTHA, V.K., MUKHERJEE, S., EBERT, B.L., GILLETTE, M.A., PAULOVICH, A., POMEROY, S.L., GOLUB, T.R., LANDER, E.S. and MESIROV, J.P., 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, vol. 102, no. 43, pp. 15545-15550. ISSN 0027-8424, 1091-6490. DOI 10.1073/pnas.0506580102.

SUN, C., WANG, X., HUI, Y., FUKUI, H., WANG, B., and MIWA, H., 2021. The Potential Role of REG Family Proteins in Inflammatory and Inflammation-Associated Diseases of the Gastrointestinal Tract. *International Journal of Molecular Sciences*, vol. 22, no. 13, pp. 7196. ISSN 1422-0067. DOI 10.3390/ijms22137196.

TAKASAWA, S., TSUCHIDA, C., SAKURAMOTO-TSUCHIDA, S., TAKEDA, M., ITAYA-HIRONAKA, A., YAMAUCHI, A., MISU, M., SHOBATAKE, R., UCHIYAMA, T., MAKINO, M., and OHBAYASHI, C., 2018. Expression of human REG family genes in inflammatory bowel disease and their molecular mechanism. *Immunologic Research*, vol. 66, no. 6, pp. 800-805. ISSN 0257-277X, 1559-0755. DOI 10.1007/s12026-019-9067-2.

TALANDASHTI, R., VAN EK, L., GEHIN, C., XUE, D., MOQADAM, M., GAVIN, A.-C., and REUTER, N., 2024. Membrane specificity of the human cholesterol transfer protein STARD4. *Journal of Molecular Biology*, vol. 436, no. 11, pp. 168572. ISSN 00222836. DOI 10.1016/j.jmb.2024.168572.

TANAKA, T., WARNER, B.M., MICHAEL, D.G., NAKAMURA, H., ODANI, T., YIN, H., ATSUMI, T., NOGUCHI, M., & CHIORINI, J.A., 2022. LAMP3 inhibits autophagy and contributes to cell death by lysosomal membrane permeabilization. *Autophagy*, vol. 18, no. 7, pp. 1629-1647. ISSN 1554-8627, 1554-8635. DOI 10.1080/15548627.2021.1995150.

TANG, A., LI, N., LI, X., YANG, H., WANG, W., ZHANG, L., LI, G., XIONG, W., MA, J., and SHEN, S., 2012. Dynamic activation of the key pathways: linking colitis to colorectal

cancer in a mouse model. *Carcinogenesis*, vol. 33, no. 7, pp. 1375-1383. ISSN 1460-2180. DOI 10.1093/carcin/bgs183.

TEN HOORN, S., DE BACK, T.R., SOMMEIJER, D.W. and VERMEULEN, L., 2022. Clinical Value of Consensus Molecular Subtypes in Colorectal Cancer: A Systematic Review and Meta-Analysis. *JNCI: Journal of the National Cancer Institute*, vol. 114, no. 4, pp. 503-516. ISSN 0027-8874, 1460-2105. DOI 10.1093/jnci/djab106.

THE HUMAN PROTEIN ATLAS, 2025a. SLC26A2 protein expression summary - The Human Protein Atlas. [online]. [Consulted: 9 July 2025]. Available in: <https://www.proteinatlas.org/ENSG00000155850-SLC26A2>.

THE HUMAN PROTEIN ATLAS, 2025b. Tissue expression of LGALS12 - Summary - The Human Protein Atlas. [online]. [Accessed: 6 July 2025]. Available at: <https://www.proteinatlas.org/ENSG00000133317-LGALS12/tissue>.

THERMO FISHER SCIENTIFIC, 2020. *Transcriptome Analysis Console (TAC)* [online]. 2020. Carlsbad, CA: s.n. Available at: <https://www.thermofisher.com/ar/es/home/life-science/microarray-analysis/microarray-analysis-instruments-software-services/microarray-analysis-software/affymetrix-transcriptome-analysis-console-software.html>.

TOSCANO, M.A., MARTÍNEZ ALLO, V.C., CUTINE, A.M., RABINOVICH, G.A., & MARIÑO, K.V., 2018. Untangling Galectin-Driven Regulatory Circuits in Autoimmune Inflammation. *Trends in Molecular Medicine*, vol. 24, no. 4, pp. 348-363. ISSN 14714914. DOI 10.1016/j.molmed.2018.02.008.

ULLAH, F., PILLAI, A.B., OMAR, N., DIMA, D., and HARICHAND, S., 2023. Early-Onset Colorectal Cancer: Current Insights. *Cancers*, vol. 15, no. 12, pp. 3202. ISSN 2072-6694. DOI 10.3390/cancers15123202.

UZZAN, M., MARTIN, J.C., MESIN, L., LIVANOS, A.E., CASTRO-DOPICO, T., HUANG, R., PETRALIA, F., MAGRI, G., KUMAR, S., ZHAO, Q., ROSENSTEIN, A.K., TOKUYAMA, M., SHARMA, K., UNGARO, R., KOSOY, R., JHA, D., FISCHER, J., SINGH, H., KEIR, M.E., RAMAMOORTHI, N., O'GORMAN, W.E., COHEN, B.L., RAHMAN, A., COSSARINI, F., SEKI, A., LEYRE, L., VAQUERO, S.T., GURUNATHAN, S., GRASSET, E.K., LOSIC, B., DUBINSKY, M., GREENSTEIN, A.J., GOTTLIEB, Z., LEGNANI, P., GEORGE, J., IRIZAR, H., STOJMIROVIC, A., BRODMERKEL, C., KASARKIS, A., SANDS, B.E., FURTADO, G., LIRA, S.A., TUONG, Z.K., KO, H.M., CERUTTI, A., ELSON, C.O., CLATWORTHY, M.R., MERAD, M., SUÁREZ-FARIÑAS, M., ARGmann, C., HACKNEY, J.A., VICTORA, G.D., RANDOLPH, G.J., KENIGSBERG, E., COLOMBEL, J.F. and MEHANDRU, S., 2022. Ulcerative colitis is characterized by a plasmablast-skewed humoral response associated with disease activity. *Nature Medicine*, vol. 28, no. 4, pp. 766-779. ISSN 1078-8956, 1546-170X. DOI 10.1038/s41591-022-01680-y.

VAJARIA, B.N. and PATEL, P.S., 2017. Glycosylation: a hallmark of cancer? *Glycoconjugate Journal*, vol. 34, no. 2, pp. 147-156. ISSN 0282-0080, 1573-4986. DOI 10.1007/s10719-016-9755-2.

VAN ALLEN, E.M., MIAO, D., SCHILLING, B., SHUKLA, S.A., BLANK, C., ZIMMER, L., SUCKER, A., HILLEN, U., FOPPEN, M.H.G., GOLDINGER, S.M., UTIKAL, J., HASSEL,

# UADE

## SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía

---

- J.C., WEIDE, B., KAEHLER, K.C., LOQUAI, C., MOHR, P., GUTZMER, R., DUMMER, R., GABRIEL, S., WU, C.J., SCHADENDORF, D., and GARRAWAY, L.A., 2015. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science (New York, N.Y.)*, vol. 350, no. 6257, pp. 207-211. ISSN 1095-9203. DOI 10.1126/science.aad0095.
- VANDEL, J., GHEERAERT, C., STAELS, B., EECKHOUTE, J., LEFEBVRE, P., and DUBOIS-CHEVALIER, J., 2020. GIANT: galaxy-based tool for interactive analysis of transcriptomic data. *Scientific Reports* [online], vol. 10, no. 1. [Accessed: 8 July 2025]. ISSN 2045-2322. DOI 10.1038/s41598-020-76769-w. Available at: <https://www.nature.com/articles/s41598-020-76769-w>.
- VARKI, A., 2017. Biological roles of glycans. *Glycobiology*, vol. 27, no. 1, pp. 3-49. ISSN 0959-6658, 1460-2423. DOI 10.1093/glycob/cww086.
- VARKI, A., CUMMINGS, R.D., ESKO, J.D., HUDSON H., F., STANLEY, P., BERTOZZI, C.R., HART, G.W., & ETZLER, M.E., 2022. *Essentials of Glycobiology*. 4.a. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press. ISBN 978-1-62182-421-3.
- VILLABLANCA, E.J., SELIN, K. and HEDIN, C.R.H., 2022. Mechanisms of mucosal healing: treating inflammatory bowel disease without immunosuppression? *Nature Reviews Gastroenterology & Hepatology*, vol. 19, no. 8, pp. 493-507. ISSN 1759-5045, 1759-5053. DOI 10.1038/s41575-022-00604-y.
- VISCIDO, A., VALVANO, M., STEFANELLI, G., CAPANNOLO, A., CASTELLINI, C., ONORI, E., CICCONE, A., VERNIA, F., & LATTELLA, G., 2022. Systematic review and meta-analysis: the advantage of endoscopic Mayo score 0 over 1 in patients with ulcerative colitis. *BMC Gastroenterology*, vol. 22, no. 1, pp. 92. ISSN 1471-230X. DOI 10.1186/s12876-022-02157-5.
- WAN, Y., YANG, L., JIANG, S., QIAN, D., and DUAN, J., 2022. Excessive Apoptosis in Ulcerative Colitis: Crosstalk Between Apoptosis, ROS, ER Stress, and Intestinal Homeostasis. *Inflammatory Bowel Diseases*, vol. 28, no. 4, pp. 639-648. ISSN 1078-0998, 1536-4844. DOI 10.1093/ibd/izab277.
- WANG, J., QIAN, J., HU, Y., KONG, X., CHEN, H., SHI, Q., JIANG, L., WU, C., ZOU, W., CHEN, Y., XU, J., and FANG, J.-Y., 2014. ArhGAP30 promotes p53 acetylation and function in colorectal cancer. *Nature Communications*, vol. 5, no. 1, pp. 4735. ISSN 2041-1723. DOI 10.1038/ncomms5735.
- WANG, Y., HAN, J., YANG, G., ZHENG, S., ZHOU, G., LIU, X., CAO, X., LI, G., ZHANG, B., XIE, Z., LI, L., ZHANG, M., LI, X., CHEN, M., and ZHANG, S., 2024. Therapeutic potential of the secreted Kazal-type serine protease inhibitor SPINK4 in colitis. *Nature Communications*, vol. 15, no. 1, pp. 5874. ISSN 2041-1723. DOI 10.1038/s41467-024-50048-y.
- WATANABE, T., KOBUNAI, T., TODA, E., KANAZAWA, T., KAZAMA, Y., TANAKA, J., TANAKA, T., YAMAMOTO, Y., HATA, K., KOJIMA, T., YOKOYAMA, T., KONISHI, T., OKAYAMA, Y., SUGIMOTO, Y., OKA, T., SASAKI, S., AJIOKA, Y., MUTO, T., and NAGAWA, H., 2007. Gene expression signature and the prediction of ulcerative colitis-associated colorectal cancer by DNA microarray. *Clinical Cancer Research: An Official*

*Journal of the American Association for Cancer Research*, vol. 13, no. 2 Pt 1, pp. 415-420. ISSN 1078-0432. DOI 10.1158/1078-0432.CCR-06-0753.

WICKHAM, H., 2016. *ggplot2: Elegant Graphics for Data Analysis*. 2nd ed. 2016. Cham: Springer International Publishing: Imprint: Springer. Use R! ISBN 978-3-319-24277-4. 519.5

WILLIAMS, A.D., KOROLKOVA, O.Y., SAKWE, A.M., GEIGER, T.M., JAMES, S.D., MULDOON, R.L., HERLINE, A.J., GOODWIN, J.S., IZBAN, M.G., WASHINGTON, M.K., SMOOT, D.T., BALLARD, B.R., GAZOULI, M., and M'KOMA, A.E., 2017. Human alpha defensin 5 is a candidate biomarker to delineate inflammatory bowel disease. *PLOS ONE*, vol. 12, no. 8, pp. e0179710. ISSN 1932-6203. DOI 10.1371/journal.pone.0179710.

WONG, M.Y., CHEN, K., ANTONOPOULOS, A., KASPER, B.T., DEWAL, M.B., TAYLOR, R.J., WHITTAKER, C.A., HEIN, P.P., DELL, A., GENEREUX, J.C., HASLAM, S.M., MAHAL, L.K., and SHOULDER, M.D., 2018. XBP1s activation can globally remodel N-glycan structure distribution patterns. *Proceedings of the National Academy of Sciences* [online], vol. 115, no. 43. [Consulted: 30 June 2025]. ISSN 0027-8424, 1091-6490. DOI 10.1073/pnas.1805425115. Available in: <https://pnas.org/doi/full/10.1073/pnas.1805425115>.

WRIGHT, R. and TRUELOVE, S.R., 1966. Serial rectal biopsy in ulcerative colitis during the course of a controlled therapeutic trial of various diets. *The American Journal of Digestive Diseases*, vol. 11, no. 11, pp. 847-857. ISSN 0002-9211, 1573-2568. DOI 10.1007/BF02233941.

XING, Z., LI, X., HE, J., CHEN, Y., ZHU, L., ZHANG, X., HUANG, Z., TANG, J., GUO, Y. and HE, Y., 2024. OLFM4 modulates intestinal inflammation by promoting IL-22+ILC3 in the gut. *Communications Biology* [online], vol. 7, no. 1. [Accessed: 7 July 2025]. ISSN 2399-3642. DOI 10.1038/s42003-024-06601-y. Available in: <https://www.nature.com/articles/s42003-024-06601-y>.

YALCHIN, M., BAKER, A.-M., GRAHAM, T.A., & HART, A., 2021. Predicting Colorectal Cancer Occurrence in IBD. *Cancers*, vol. 13, no. 12, pp. 2908. ISSN 2072-6694. DOI 10.3390/cancers13122908.

YAN, L., 2021. *ggvenn: Draw Venn Diagram by "ggplot2"* [online]. 13 January 2021. S.l.: s.n. [Accessed: 6 July 2025]. Available at: <https://CRAN.R-project.org/package=ggvenn>.

YANG, C. & MERLIN, D., 2024. Unveiling Colitis: A Journey through the Dextran Sodium Sulfate-induced Model. *Inflammatory Bowel Diseases*, vol. 30, no. 5, pp. 844-853. ISSN 1078-0998, 1536-4844. DOI 10.1093/ibd/izad312.

YANG, H., SHEN, X., WANG, H., and SHUAI, W., 2024. Tumour necrosis factor alpha-induced protein 3-interacting protein 3 overexpression protects against arrhythmogenic remodelling in the heart failure mice. *Europace* [online], vol. 27, no. 1. [Accessed: 7 July 2025]. ISSN 1099-5129, 1532-2092. DOI 10.1093/europace/euaf002. Available at: <https://academic.oup.com/europace/article/doi/10.1093/europace/euaf002/7952882>.

YANG, R.-Y., RABINOVICH, G.A. and LIU, F.-T., 2008. Galectins: structure, function and therapeutic potential. *Expert Reviews in Molecular Medicine* [online], vol. 10. [Accessed: 8 July 2025]. ISSN 1462-3994. DOI 10.1017/s1462399408000719. Available at: [https://www.cambridge.org/core/product/identifier/S1462399408000719/type/journal\\_article](https://www.cambridge.org/core/product/identifier/S1462399408000719/type/journal_article).

ZAHN, A., 2007. Aquaporin-8 expression is reduced in ileum and induced in colon of patients with ulcerative colitis. *World Journal of Gastroenterology*, vol. 13, no. 11, pp. 1687. ISSN 1007-9327. DOI 10.3748/wjg.v13.i11.1687.

ZHANG, C., ZHANG, J., ZHANG, Y., SONG, Z., BIAN, J., YI, H., and MA, Z., 2023. Identifying neutrophil-associated subtypes in ulcerative colitis and confirming neutrophils promote colitis-associated colorectal cancer. *Frontiers in Immunology* [online], vol. 14. [Accessed: 21 October 2024]. ISSN 1664-3224. DOI 10.3389/fimmu.2023.1095098. Available at:

<https://www.frontiersin.org/journals/immunology/articles/10.3389/fimmu.2023.1095098/full>.

ZHANG, L., GALLUP, M., ZLOCK, L., CHEN, Y.T.F., FINKBEINER, W.E., and MCNAMARA, N.A., 2014. Pivotal role of MUC1 glycosylation by cigarette smoke in modulating disruption of airway adherens junctions in vitro. *The Journal of Pathology*, vol. 234, no. 1, pp. 60-73. ISSN 0022-3417, 1096-9896. DOI 10.1002/path.4375.

ZHAO, H., HUANG, M., and JIANG, L., 2023a. Potential Roles and Future Perspectives of Chitinase 3-like 1 in Macrophage Polarization and the Development of Diseases. *International Journal of Molecular Sciences*, vol. 24, no. 22, pp. 16149. ISSN 1422-0067. DOI 10.3390/ijms242216149.

ZHAO, T., SU, Z., LI, Y., ZHANG, X., and YOU, Q., 2020. Chitinase-3 like-protein-1 function and its role in diseases. *Signal Transduction and Targeted Therapy*, vol. 5, no. 1, pp. 201. ISSN 2059-3635. DOI 10.1038/s41392-020-00303-7.

ZHAO, X., LU, M., LIU, Z., ZHANG, M., YUAN, H., DAN, Z., WANG, D., MA, B., YANG, Y., YANG, F., SUN, R., LI, L., and DANG, C., 2023b. Comprehensive analysis of alpha defensin expression and prognosis in human colorectal cancer. *Frontiers in Oncology* [online], vol. 12. [Accessed: 8 July 2025]. ISSN 2234-943X. DOI 10.3389/fonc.2022.974654. Available at: <https://www.frontiersin.org/articles/10.3389/fonc.2022.974654/full>.

ZHOU, C., GAO, Y., DING, P., WU, T., and JI, G., 2023a. The role of CXCL family members in different diseases. *Cell Death Discovery*, vol. 9, no. 1, pp. 212. ISSN 2058-7716. DOI 10.1038/s41420-023-01524-9.

ZHOU, R.W., HARPAZ, N., ITZKOWITZ, S.H., and PARSONS, R.E., 2023b. Molecular mechanisms in colitis-associated colorectal cancer. *Oncogenesis*, vol. 12, no. 1, pp. 48. ISSN 2157-9024. DOI 10.1038/s41389-023-00492-0.

ZHU, L., HAN, J., LI, L., WANG, Y., LI, Y., and ZHANG, S., 2019. Claudin Family Participates in the Pathogenesis of Inflammatory Bowel Diseases and Colitis-Associated Colorectal Cancer. *Frontiers in Immunology*, vol. 10, pp. 1441. ISSN 1664-3224. DOI 10.3389/fimmu.2019.01441.

## 8. ANNEXES

### 8.1. List of glycobiological genes or "glycogenes"

- **General Glycogenes (877):** ARSA, ARSB, ARSD, ARSL, ARSF, STS, FUCA1, GLA, GLB1, GUSB, HPSE, HPSE2, HEXA, HEXB, GAA, HYAL1, HYAL2, HYAL3, HYAL4, OGA, IDUA, AGA, ASAHI, ASAHH, ACER2, NAAA, CTBS, CTNS, CTSA, GALC, GBA, GM2A, LAMP1, LAMP2, LAMP3, LCT, LIPA, NAGA, SMPD1, MAN1A1, MAN1A2, MAN1B1, MAN1C1, MAN2A1, MAN2A2, MAN2B1, MANBA, NAGLU, NPL, PPBP, SLC17A5, NEU1, NEU2, NEU3, NEU4, GALNS, GNS, IDS, SULF1, SULF2, SGSH, SLC35A1, SLC35A2, SLC35A3, SLC35A4, SLC35A5, SLC35B1, SLC35B2, SLC35B3, SLC35B4, SLC35C1, SLC35C2, SLC35D1, SLC35D2, SLC35D3, SLC35E1, SLC35E2B, SLC35E3, SLC35E4, SLC35F1, SLC35F2, SLC35F3, SLC35F4, SLC35F5, CMAS, CMAHP, FPGT, GALE, GALK1, GALK2, GALT, GFPT1, GFPT2, GMDS, GMPPA, GMPPB, GNB1, RACK1, GNE, NANP, GNPDAA1, GNPNAT1, GPI, HK1, HK3, KHK, MPI, NAGK, NANS, PAPSS1, PAPSS2, PGM1, PMM1, PMM2, RENBP, TSTA3, UAP1, UGDH, UGP2, UXS1, SULT1A2, SULT1A3, SULT1B1, SULT1C2, SULT1C3, SULT1C4, TPST1, TPST2, DOLK, ALG5, DPM1, DPM2, DPM3, MPDU1, DPAGT1, ALG13, ALG13, ALGg14, ALG1, ALG2, ALG11, RFT1, ALG3, ALG9, ALG12, ALG6, ALG8, ALG10, ALG10B, DOLPP1, FDPS, DHDDS, NUS1, SRD5A3, STT3A, RPN1, RPN2, DDOST, DAD1, OST4, TMEM258, KRTCAP2, OSTC, SELENOT, STT3B, MAGT1, TUSC3, MLEC, CANX, CALR, MOGS, GANAB, PRKCSH, UGGT1, UGGT2, MANEA, MANEAL, MAN2C1, MGAT1, MGAT3, FUT8, MGAT2, MGAT4A, MGAT4B, MGAT5, MGAT5B, GNPTAB, GNPTG, NAGPA, LMAN1, LMAN2L, LMAN2, M6PR, IGF2R, EDEM2, EDEM1, EDEM3, OS9, NGLY1, ENGASE, MGAT4D, B4GALNT3, B4GALNT4, ST6GAL2, ST3GAL3, ST3GAL4, ST3GAL6, FUT9, FUT3, FUT4, FUT5, FUT6, FUT10, FUT11, B4GALT1, B4GALT2, B4GALT3, B4GALT4, B4GALT5, ST6GAL1, A4GALT, ABO, GAL3ST2, GAL3ST3, FUT1, FUT2, CHST2, CHST4, B3GNT2, B3GNT4, B3GNT8, B3GNT3, GCNT2, GCNT3, B3GAT1, B3GAT2, CHST10, B4GALNT2, ST8SIA2, ST8SIA3, ST8SIA4, B3GALT1, B3GALT2, B3GALT5, B3GALNT2, CHST8, CHST9, FUT7, GALNT1, GALNT10, GALNT11, GALNT12, GALNT13, GALNT14, GALNT15, GALNT16, GALNT18, GALNTL5, GALNTL6, GALNT2, GALNT17, GALNT3, GALNT4, GALNT5, GALNT6, GALNT7, GALNT8, GALNT9, ST6GALNAC1, ST6GALNAC2, C1GALT1, C1GALT1C1, ST3GAL1, A4GNT, ST6GALNAC3, ST6GALNAC4, GCNT1, GCNT4, B3GNT6, ST8SIA6, GAL3ST4, POFUT1, POFUT2, B3GLCT, LFNG, MFNG, RFNG, POMT1, POMT2, POMGNT1, POMGNT2, POMK, FKTN, FKRP, RXYLT1, B4GAT1, LARGE1, LARGE2, ISPD, OGT, MGEA5, EOGT, POGLUT1, GXYLT1, GXYLT2, XXYLT1, KDELC1, KDELC1\_2, PL0D1, PL0D2, PL0D3, COLGALT1, COLGALT2, TMTC1, TMTC2, TMTC3, TMTC4, DAG1, SPTLC1, SPTLC2, ORMDL1, ORMDL2, ORMDL3, KDSR, CERS1, CERS2, CERS3, CERS4, CERS5, CERS6, DEGS1, DEGS2, ASAHH2B, ACER1, ACER3, SGPP1, SGPP2, SPHK1, SPHK2, SGPL1, CERK, PLPP1, PLPP3, PLPP2, SGMS1, SGMS2,

# UADE SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía

---

SMPD2, SMPD3, SMPD4, SMPDL3A, SMPDL3B, ENPP7, UGCG, GBA2, UGT8, PSAP, COL4A3BP, PLEKHA8, ELOVL1, ELOVL2, ELOVL3, ELOVL4, ELOVL5, ELOVL6, ELOVL7, HSD17B12, HACD1, HACD2, HACD3, HACD4, TECR, B4GALT6, B3GNT5, B4GALNT1, ST3GAL5, GAL3ST1, CPTP, B3GALNT1, GBGT1, ST3GAL2, ST8SIA1, ST6GALNAC5, ST6GALNAC6, ST8SIA5, B3GALT4, PGM2, HK2, G6PC, GCK, AMDHD2, NT5M, PGM3, FUK, CHGA, PFKL, PFKP, PFKM, GNPDA2, TGDS, ALDOB, H6PD, PGM5, ALDOC, FBP2, PFKFB2, PFKFB3, G6PC2, FBP1, CASD1, SIAE, SLC2A1, SLC2A2, SLC2A3, SLC2A4, SLC2A5, SLC2A8, SLC2A10, SLC2A12, SLC5A8 1, SLC26A2, SLC35E2A, SLC35F6, SLC35G1, SLC35G2, SLC35G3, SLC35G4, SLC35G5, SLC35G6, SLC37A4, SLC39A8, CANT1, IMPAD1, SLC33A1, GOLPH3, ATP6AP1, ATP6AP2, ATP6V0A2, ATP6V1E1, ATP6V1A, TMEM199, CCDC115, TMEM165, GLTP, SPPL3, ARSI, ARSG, ARSE, ARSH, ARSJ, ARSK, MAN2B2, KLB, KL, LCTL, AGL, AMY1A, AMY1B, AMY1C, AMY2A, AMY2B, SLC3A2, SLC3A1, GBE1, CHI3L2, LYZL6, HEXDC, LYZ, LYZL1, LYZL2, SPACA3, SPACA5, LALBA, LYG1, LYG2, FUCACA6 2, MGAM, GANC, MYORG, MGAM2, SI, GLB1L, GLB1L3, GLB1L2, TREH, SPAM1, PGGHG, CHID1, HNF4A, IL6ST, IL23R, IL2RA, SLC9A3, XBP1, SEC24D, SLC50A1, ARF4, SERP1, SDF2L1, SYVN1, SEC31A, SEL1L, UBE2J1, DERL2, ENTPD4, ERP44, PFKFB4, PGK1, PREB, RAB1A, SAR1B, SEC13, SEC23A, SEC24C, SLC2A6, SLC37A1, SLC37A3, SLC45A4, TMEM59, VCP, DCK, DNMT3B, ALDOA, ATP6V0A1, ATP7A, CAD, CALM2, CALM3, CLN6, CTPS2, DNMT11, ENO1, G6PC3, G6PD, GAPDH, GMPS, HSPA1A, HSPA8, LDHA, MDH1, MDH2, NME2, NME3, NME4, NME7, PGAM1, PGLS, RAB1B, RPIA, SLC29A1, SORT1, TALDO1, THBS1, TKT, TMEM115, TPI1, TRAK2, TXNDC9, UCK2, UMPS, UPRT, CALHM1, CALM1, COPA, COPB1, COPB2, COPE, COPG, ERDJ3, ERDJ4, ERDJ5, ERO1LP, ERP57, GCC1, GCC2, GLTPD1, GLTPD2, GOLGA1, GOLGA2, GOLGA3, GOLGA4, GOLGA5, GOLGA8A, GOLGA8B, GOLGA8F, GOLGA8G, GOLGA8H, GOLGA8J, GOLGA8K, GOLGA8M, GOLGA8N, GOLGA8O, GOLGA8Q, GOLGA8R, GOLGA8S, GOLGA8T, GOLGB1, GORASP1, HSPA1B, KDELR3, LMF1, PDIA5, PDIA6, SEC23B, SEC24A, SEC24B, SEC61A1, SEC61B, SEC61G, SEC62, SEC63, SERP2, SSR1, SSR2, SSR3, SSR4, TMF1, ATF6, TRIP11, GOLGA6A, GOLGA6B, GOLGA6C, GOLGA6D, GOLGA7, CUX1, USO1, BLZF1, GORASP2, VPS51, VPS52, VPS53, VPS54, STX5, GOSR1, GOSR2, YKT6, SNAP29, STX16, GORAB, SCYL1, HSPA5, CO g1, COG2, COG3, COG4, COG5, COG6, COG7, COG8, SAR1A, SEC31B, ARF1, ARFGAP1, COPG1, COPG2, COPZ2, COPZ1, ATRNL1, ATRN, PKD1L2, PKD1, PKD2, PRG2, PRG3, CD302, LY75, DGCR2, CD248, THBD, ACAN, BCAN, NCAN, VCAN, ASGR1, ASGR2, CD207, CLEC10A, CLEC4A, CLEC4C, CLEC4D, CLEC4E, CLEC4F, CLEC4G, CLEC4M, CLEC6A, CLECL1, COLEC12, CD209, CCL23, COLEC11, COLEC10, MBL2, SFTPA1, SFTPA2, SFTP, SELE, SELL, SELP, CD69, CD72, CLEC12A, CLEC12B MBL2, SFTPA1, CLEC1A, CLEC1B, CLEC2B, CLEC2D, CLEC2L, CLEC5A, CLEC7A, CLEC9A, KLRB1, KLRC1, KLRC3, KLRD1, KLRF1, KLRG1, KLRK1, OLR1, MRC1, PLA2R1, CLEC11A, CLEC3A, REG1A, REG1B, REG3A, REG4, CHODL, CLEC3B, LAYN, CLEC14A, CD83, FREM1, CLEC18A, ICAM1, ICAM2, ICAM3, NCAM1, VCAM1, CD22, CD33, MAG, SIGLEC1,

SIGLEC10, SIGLEC11, SIGLEC14, SIGLEC15, SIGLEC5, SIGLEC6, SIGLEC7, SIGLEC8, SIGLEC9, SIGLECL1, SIGLEC17P, CLC, LGALSL, LGALS1, LGALS12, , LGALS13, LGALS2, LGALS3, LGALS7, LGALS8, LGALS9, LGALS14, LGALS3BP, CLGN, CALR3, FCER2, CLEC17A, SFTPB, SFTPC, KLRG2, CLEC2A, KLRC2, KLRC4, MRC2, REG3G, CLEC19A, PKD1L3, CD93, CLEC16A, CLEC18B, CLEC18C, FBXO2, FBXO6, FBXO17, FBXO27, FBXO44, ICAM5, L1CAM, PECAM1, SIGLEC12, SIGLEC16, MASP1, MASP2, MCFD2, PDIA3, LMAN1L, TMED1, TMED10, TMED2, TMED3, TMED4, TMED5, TMED6, TMED7, TMED8, TMED9, ERLEC1, GRID3, TMED1 ifin, LGALS16, LGALS4, LGALS9B, LGALS9C, AGRN, LYZL4, ANXA4, ANXA5, ANXA7, ANXA9, CHI3L1, CHIA, CHIT1, OVGP1, GPNMB, HABP4, MDK, SDCBP, SDCBP2, HMMR, FCN1, FCN2, FCN3, ITLN1, ITLN2, CD1D, IKZF1, TMEM121, SMARCB1, RUNX3, SLC9A9, ASXL2, RNF168, OVOL1, TNFRSF13B, KDELR2, MEF2B, BORCS8, MGME1, BANF2, SNX5, MIR4425, ELL2, IRF1, SLC22A5, MICB-DT, MICB, HCG26, TXLNB, HIVEP2, ABCF2, NXPE4, NXPE2, MIR4708, IKZF3, GSDMB, LRRC3C, ZPBP2, MAPT, TBX21, CEP131, RUNX1, TAB1, HNF1A, IGH, DERL3, CHCHD10, PRRC2A, KREMEN1

- **Alpha Defensesins (6):** DEFA1, DEFA1B, DEFA3, DEFA4, DEFA5, DEFA6
- **Beta Defensesins (82):** DEFB1, DEFB4A, DEFB4B, DEFB103A, DEFB103B, DEFB104A, DEFB104B, DEFB105A, DEFB105B, DEFB106A, DEFB106B, DEFB107A, DEFB107B, DEFB108A, DEFB108B, DEFB109P1, DEFB110, DEFB111, DEFB112, DEFB113, DEFB114, DEFB115, DEFB116, DEFB117, DEFB118, DEFB119, DEFB120, DEFB121, DEFB122, DEFB123, DEFB124, DEFB125, DEFB126, DEFB127, DEFB128, DEFB129, DEFB130A, DEFB130B, DEFB131A, DEFB131B, DEFB132, DEFB133, DEFB134, DEFB135, DEFB136, DEFB137, DEFB138, DEFB139, DEFB140, DEFB141, DEFB142, DEFB143, DEFB144, DEFB145, DEFB146, DEFB147, DEFB148, DEFB149, DEFB150, DEFB151, DEFB152, DEFB153, DEFB154, DEFB155, DEFB156, DEFB157, DEFB158, DEFB159, DEFB160, DEFB161, DEFB162, DEFB163, DEFB164, DEFB165, DEFB166, DEFB167, DEFB168, DEFB169, DEFB170, DEFB171, DEFB172, DEFB173

## 8.2. Dysregulated genes in quiescent neoplasm/UC but not in neoplasm/healthy patients

TABLE XV. The 60 genes significantly dysregulated in neoplasia with respect to quiescent ulcerative colitis and not with respect to the group of healthy patients. Result obtained in the Section 5.3.1.2, from Results.

Aumentados	Inhibidos
RAB13, SPI1, CEP19, RHEX, LY6E, PIK3CD, TOX2, CAVIN3, TRGC2	<b>GCNT2</b> , SMIM5, MIR570HG, NMRAL2P, FAM118A, CCDC9B, MAVS, HOOK2, FBLIM1, MOGAT3, SRSF6, EPPK1, TSPAN3, ALS2CL, SIMC1, ACAA1, SIRT7, ALDH18A1, MICAL2, NOXA1, CLDN15, MKNK1, POLR2J4, ENOSF1, TMEM120A, PTPN18, TOE1, ATG16L1, CAMK2G, ZYG11B, AFG3L1P, KIAA1217, PRR13, PITPNM3, CAPN2, BAIAP2L2, DCAF7, SREBF1, HOXB3, STIM1, DUSP16, PER3, CLN8, ZBTB43, CDC14A, APH1A, ADGRL1, APP, FLCN, CDK20, CEP72

### 8.3. Extended glycobiological signature

**TABLE XVI. Extended glycobiological signature of logFC-positive genes.** In bold, the genes included in the final signature. The p-value corresponds to the adjusted value and is presented with two significant figures, while the *logFC* it is rounded to two decimal places.

Gen murino	Gen humano ortólogo	LogFC	P-Valor
<i>Ormdl2</i>	ORMDL2	1.24	5.87E-06
<i>Gcc2</i>	GCC2	1.47	1.55E-05
<i>Gnpat1</i>	GNPNAT1	1.31	4.18E-05
<b><i>Lgals12</i></b>	<b>LGALS12</b>	<b>2.46</b>	<b>4.59E-05</b>
<i>Slc9a3</i>	SLC9A3	4.53	7.22E-05
<i>Gal3st1</i>	GAL3ST1	1.82	1.02E-04
<i>Gorasp1</i>	GORASP1	0.71	1.76E-04
<b><i>Defa5</i></b>	<b>DEFA4, DEFA5, DEFA6</b>	<b>9.42</b>	<b>2.13E-04</b>
<i>Degs2</i>	DEGS2	2.73	3.79E-04
<i>Hnf4a</i>	HNF4A	2.39	5.02E-04
<i>Slc35g1</i>	SLC35G1	2.49	5.32E-04
<i>Defa31</i>	DEFA4, DEFA5, DEFA6	6.71	6.50E-04
<i>C1galt1c1</i>	C1GALT1C1	0.94	7.28E-04
<i>Chchd10</i>	CHCHD10	1.05	1.57E-03
<i>Vps53</i>	VPS53	0.58	1.92E-03
<i>Copg2</i>	COPG2	0.59	2.05E-03
<i>Slc35a1</i>	SLC35A1	1.26	4.54E-03
<i>Papss2</i>	PAPSS2	0.96	6.69E-03
<b><i>Slc26a2</i></b>	<b>SLC26A2</b>	<b>2.38</b>	<b>7.13E-03</b>
<i>Slc22a21</i>	SLC22A5	0.98	1.01E-02
<i>Plekha8</i>	PLEKHA8	1.20	1.23E-02
<i>Sult1c2</i>	SULT1C2	3.26	1.27E-02
<i>Slc3a1</i>	SLC3A1	1.47	1.38E-02
<i>Defa39</i>	DEFA4, DEFA5, DEFA6	3.58	1.39E-02
<i>Defa29</i>	DEFA4, DEFA5, DEFA6	3.88	1.46E-02
<i>A4gnt</i>	A4GNT	1.40	1.49E-02
<i>Mgat4b</i>	MGAT4B	0.88	1.50E-02
<b><i>Lgals4</i></b>	<b>LGALS4</b>	<b>2.87</b>	<b>1.78E-02</b>
<i>Acer2</i>	ACER2	1.23	1.83E-02
<i>Large2</i>	LARGE2	0.89	2.07E-02
<i>St6galnac2</i>	ST6GALNAC2	1.01	2.64E-02
<i>Golga2</i>	GOLGA6B, GOLGA2, GOLGA8N	0.34	2.71E-02
<i>Cog6</i>	COG6	0.45	3.94E-02
<i>Smpd3</i>	SMPD3	0.40	4.47E-02
<i>Dolpp1</i>	DOLPP1	0.37	4.94E-02

**UADE** SEARCH FOR A POTENTIALLY PREDICTIVE GLYCOBIOLOGICAL SIGNATURE OF THE TRANSITION FROM ULCERATIVE COLITIS TO COLITIS-ASSOCIATED COLORECTAL CANCER – Castelli, Lucía

---

**Table XVII. Extended glycobiological signature of logFC-negative genes.** In bold, the genes included in the final signature. The p-value corresponds to the adjusted value and is presented with two significant figures, while the *logFC* it is rounded to two decimal places.

Gen murino	Gen humano ortólogo	LogFC	P-Valor
<i>Psap</i>	PSAP	-1.26	6.47E-06
<i>A4galt</i>	A4GALT	-1.84	1.87E-05
<i>Degs1</i>	DEGS1	-1.53	3.61E-05
<b><i>Icam1</i></b>	<b>ICAM1</b>	<b>-2.08</b>	<b>5.41E-05</b>
<i>Mfng</i>	MFNG	-1.02	1.48E-04
<i>Galnt2</i>	GALNT2	-1.65	1.72E-04
<b><i>St3gal1</i></b>	<b>ST3GAL1</b>	<b>-2.75</b>	<b>2.00E-04</b>
<i>Il2ra</i>	IL2RA	-0.87	3.76E-04
<b><i>Selp</i></b>	<b>SELP</b>	<b>-3.72</b>	<b>4.55E-04</b>
<i>Sgpl1</i>	SGPL1	-0.47	4.79E-04
<i>Atp6v1a</i>	ATP6V1A	-0.78	7.94E-04
<i>Pfkfb3</i>	PFKFB3	-1.75	8.55E-04
<i>Ctbs</i>	CTBS	-0.66	9.44E-04
<i>Atp6ap1</i>	ATP6AP1	-0.57	1.25E-03
<b><i>Chil3</i></b>	<b>CHIA</b>	<b>-8.23</b>	<b>1.45E-03</b>
<i>Icam2</i>	ICAM2	-1.64	1.47E-03
<i>Slc35b4</i>	SLC35B4	-0.61	1.70E-03
<b><i>Chi3l1</i></b>	<b>CHI3L1</b>	<b>-4.24</b>	<b>2.35E-03</b>
<i>Ugcg</i>	UGCG	-0.89	4.07E-03
<i>Fbxo6</i>	FBXO6	-0.77	4.09E-03
<i>Golga7</i>	GOLGA7	-0.36	4.24E-03
<i>Lipa</i>	LIPA	-0.63	5.56E-03
<i>Eno1</i>	ENO1	-0.53	6.27E-03
<i>Clec4e</i>	CLEC4E	-4.19	6.37E-03
<b><i>Sell</i></b>	<b>SELL</b>	<b>-2.66</b>	<b>7.47E-03</b>
<i>Il6st</i>	IL6ST	-0.42	8.62E-03
<i>Tmem121</i>	TMEM121	-0.60	1.12E-02
<i>Sar1a</i>	SAR1A	-0.41	1.14E-02
<i>Hexa</i>	HEXA	-0.40	1.17E-02
<i>Clec7a</i>	CLEC7A	-1.22	1.20E-02
<i>Gosr2</i>	GOSR2	-0.31	1.22E-02
<i>Fut11</i>	FUT11	-0.55	1.42E-02
<i>Sts</i>	STS	-0.45	1.56E-02
<i>St8sia3</i>	ST8SIA3	-0.61	1.60E-02
<i>Eno1b</i>	ENO1	-0.39	1.78E-02
<i>Abcf2</i>	ABCF2	-0.31	2.66E-02
<i>Gns</i>	GNS	-0.52	3.23E-02
<i>Fbp1</i>	FBP1	-1.76	3.68E-02
<i>Clec16a</i>	CLEC16A	-0.33	3.83E-02
<i>Sel1l</i>	SEL1L	-0.52	3.89E-02
<i>Vcam1</i>	VCAM1	-1.32	4.20E-02
<i>Lyz2</i>	LYZ	-0.83	4.59E-02
<i>Gla</i>	GLA	-0.66	4.79E-02
<i>M6pr</i>	M6PR	-0.28	4.80E-02

#### 8.4. Detailed real-time PCR results

**TABLE XVIII. Results of quantitative PCR amplification.** The data correspond to the *output* of the program *CFX Master*, and are the basis of statistical analysis in the **Section 5.4**, from **Results**.

Target	Muestra	Cantidad Relativa	Tratamiento	Cantidad relativa promedio	Eficiencia (ideal: 90-110)
Defa5	A1	0.21	AOM/DSS	0.49	121.1
Defa5	A2	0.26	AOM/DSS		
Defa5	A3	1.00	AOM/DSS		
Defa5	C1	0.01	Control		
Defa5	C2	0.00	Control		
Defa5	C3	0.01	Control		
Defa5	D1	0.00	DSS		
Defa5	D2	0.00	DSS		
Defa5	D3	0.00	DSS		
Lgals12	A1	6085.37	AOM/DSS	366125.78	-100
Lgals12	A2	1092031.00	AOM/DSS		
Lgals12	A3	260.56	AOM/DSS		
Lgals12	C1	1.00	Control		
Lgals12	C2	4.43	Control		
Lgals12	C3	1674.65	Control		
Lgals12	D1	688.40	DSS		
Lgals12	D2	165.66	DSS		
Lgals12	D3	2907.82	DSS		
Lgals4	A1	0.69	AOM/DSS	0.61	76.2
Lgals4	A2	0.66	AOM/DSS		
Lgals4	A3	0.47	AOM/DSS		
Lgals4	C1	0.91	Control		
Lgals4	C2	0.56	Control		
Lgals4	C3	1.00	Control		
Lgals4	D1	0.10	DSS		
Lgals4	D2	0.06	DSS		
Lgals4	D3	0.69	DSS		
St3Gal1	A1	0.52	AOM/DSS	0.34	90.3
St3Gal1	A2	0.27	AOM/DSS		
St3Gal1	A3	0.23	AOM/DSS		
St3Gal1	C1	0.74	Control		
St3Gal1	C2	1.00	Control		
St3Gal1	C3	0.23	Control		
St3Gal1	D1	0.39	DSS		
St3Gal1	D2	0.22	DSS		
St3Gal1	D3	0.59	DSS		
B2M	A1	0.26	AOM/DSS	0.25	101.9
B2M	A2	0.23	AOM/DSS		
B2M	A3	0.27	AOM/DSS		
B2M	C1	1.00	Control		
B2M	C2	0.46	Control		
B2M	C3	0.45	Control		
B2M	D1	0.25	DSS		
B2M	D2	0.18	DSS		
B2M	D3	0.46	DSS		

## **8.5. Appendix scripts, full tables, and *degfind***

The *scripts* used for data analysis, processing and graph generation, as well as for developing the application in *Shiny* and the tables of results relevant to this research, are available in the GitHub repository corresponding to the project ([https://github.com/luciacastelli/firma\\_ccrac\\_tesina2025](https://github.com/luciacastelli/firma_ccrac_tesina2025)); the developed *R* package, *degfind*, It can be accessed via <https://github.com/luciacastelli/degfind>.